IBM SPSS Modeler 18.2.1 - Guide d'utilisation

Important
Avant d'utiliser le présent document et le produit associé, prenez connaissance des informations générales figurant à la section «Remarques», à la page 257.

Informations produit

LE PRESENT DOCUMENT EST LIVRE EN L'ETAT SANS AUCUNE GARANTIE EXPLICITE OU IMPLICITE. IBM DECLINE NOTAMMENT TOUTE RESPONSABILITE RELATIVE A CES INFORMATIONS EN CAS DE CONTREFACON AINSI QU'EN CAS DE DEFAUT D'APTITUDE A L'EXECUTION D'UN TRAVAIL DONNE.

Ce document est mis à jour périodiquement. Chaque nouvelle édition inclut les mises à jour. Les informations qui y sont fournies sont susceptibles d'être modifiées avant que les produits décrits ne deviennent eux-mêmes disponibles. En outre, il peut contenir des informations ou des références concernant certains produits, logiciels ou services non annoncés dans ce pays. Cela ne signifie cependant pas qu'ils y seront annoncés.

Pour plus de détails, pour toute demande d'ordre technique, ou pour obtenir des exemplaires de documents IBM, référez-vous aux documents d'annonce disponibles dans votre pays, ou adressez-vous à votre partenaire commercial.

Vous pouvez également consulter les serveurs Internet suivants :

- http://www.fr.ibm.com (serveur IBM en France)
- http://www.ibm.com/ca/fr (serveur IBM au Canada)
- http://www.ibm.com (serveur IBM aux Etats-Unis)

Compagnie IBM France
Direction Qualité
17, avenue de l'Europe
92275 Bois-Colombes Cedex

© Copyright IBM France 2019. Tous droits réservés.

Cette édition s'applique à la version 18, édition 2, modification 0 d'IBM SPSS Modeler ainsi qu'à toutes les éditions et modifications suivantes, sauf indication contraire dans les nouvelles éditions.

Table des matières

Avis aux lecteurs canadiens vii

Chapitre 1. A propos d'IBM SPSS Modeler 1

Produits IBM SPSS Modeler 1
IBM SPSS Modeler 1
IBM SPSS Modeler Server 1
IBM SPSS Modeler Administration Console . . . 2
IBM SPSS Modeler Batch 2
IBM SPSS Modeler Solution Publisher 2
Adaptateurs IBM SPSS Modeler Server pour IBM SPSS Collaboration and Deployment Services . . 2
Editions d'IBM SPSS Modeler 2
Documentation 3
Documentation de SPSS Modeler Professional . . 3
Documentation de SPSS Modeler Premium . . . 4
Exemples d'application 4
Dossier Demos 4
Suivi des licences 5

Chapitre 2. Nouvelles fonctionnalités dans IBM SPSS Modeler 18.2.1 7

Chapitre 3. Présentation du produit . . . 9

Démarrage 9
Démarrage d'IBM SPSS Modeler 9
Lancement de l'application à partir de la ligne de commande 9
Connexion à IBM SPSS Modeler Server 10
Connexion à Analytic Server. 12
Changement de répertoire temporaire 13
Démarrage de plusieurs sessions IBM SPSS Modeler 13
Interface IBM SPSS Modeler en un clin d'oeil . . . 14
Espace de travail de flux IBM SPSS Modeler . . 14
Palette de noeuds 15
Gestionnaires IBM SPSS Modeler 16
Projets IBM SPSS Modeler 18
Barre d'outils IBM SPSS Modeler 18
Personnalisation de la barre d'outils 20
Personnalisation de la fenêtre IBM SPSS Modeler 20
Modification de la taille des icônes d'un flux . . 21
Utilisation de la souris dans IBM SPSS Modeler 22
Utilisation des touches de raccourci 22
Impression 23
Automatisation d'IBM SPSS Modeler 23

Chapitre 4. Compréhension de l'exploration de données 25

Présentation de l'exploration de données 25
Evaluation des données 26
Stratégie pour l'exploration de données 27
Méthodologie CRISP-DM 28
Types de modèle 29
Exemples d'exploration de données 35

Chapitre 5. Création de flux 37

Présentation de la création de flux 37
Création de flux de données 37
Utilisation des noeuds 38
Utilisation des flux 43
Descriptions de flux 57
Exécution des flux 58
Utilisation des modèles 59
Ajout de commentaires et d'annotations à des noeuds et à des flux 59
Enregistrement des flux de données 65
Chargement des fichiers 67
Mappage de flux de données 67
Astuces et raccourcis 69

Chapitre 6. Travail avec les données . . 71

Construction de graphiques 71
Présentation et terminologie 71
Construction d'un graphique à partir de la galerie 72
Types de graphique. 72
Tableau de bord 95
Préférences de visualisation globales 96

Chapitre 7. Utilisation de la sortie . . . 99

Visualiseur 99
Affichage et masquage des résultats 99
Déplacement, suppression et copie de sortie . . 100
Modification de l'alignement initial 100
Changement de l'alignement des éléments de sortie 100
Panneau de légende du Visualiseur 100
Ajout d'éléments dans le Visualiseur 101
Recherche et remplacement d'informations dans le Visualiseur 102
Copie de sorties dans d'autres applications . . . 103
Sortie interactive 104
Exportation de sortie 105
Options HTML 106
Options du rapport Web. 107
Options Word/RTF 107
Options Excel 108
Options PowerPoint 109
Options PDF 110
Options texte 110
Options graphiques uniquement 111
Options de format des graphiques 111
Impression du Visualiseur 112
Impression des sorties et des graphiques . . . 113
Aperçu avant impression 113
Attributs de page : En-têtes et pieds de page 113
Attributs de page : Options. 114
Enregistrement des sorties 114

Enregistrement d'un document Visualiseur . . 114
Tableaux croisés 116
Tableaux croisés 116
Manipulation d'un tableau croisé dynamique 116
Utilisation des couches 119
Affichage et masquage d'éléments 119
TableLooks 120
Propriétés du tableau. 120
Propriétés de cellule 123
Notes de bas de page et légendes 124
Largeur des cellules de données 126
Modification de la largeur de colonne 126
Affichage des bordures masquées dans un tableau croisé dynamique 126
Sélection de lignes, de colonnes et de cellules dans un tableau croisé dynamique 126
Impression des tableaux croisés dynamiques 126
Création d'un graphique à partir d'un tableau croisé dynamique 128
Tableaux de version antérieure 128
Options 128
Options 128
Options générales 128
Options du Visualiseur 129
Options de tableau croisé dynamique 129
Options relatives aux sorties 130

Chapitre 8. Traitement des valeurs manquantes 131

Description des valeurs manquantes. 131
Traitement des valeurs manquantes 132
Traitement des enregistrements contenant des valeurs manquantes 132
Traitement des champs contenant des valeurs manquantes 132
Traitement des enregistrements contenant des valeurs manquantes système 133
Attribution ou remplacement de valeurs manquantes 135
Fonctions CLEM pour les valeurs manquantes . . 136

Chapitre 9. Création d'expressions CLEM. 139

A propos de CLEM 139
Exemples CLEM 139
Valeurs et types de données 141
Expressions et conditions 142
Paramètres de flux, de session et de super noeud 142
Utilisation de chaînes. 143
Traitement des valeurs non renseignées et des valeurs manquantes 144
Utilisation de nombres 144
Utilisation de l'heure et de la date 144
Récapitulation de plusieurs champs 145
Utilisation des données à réponses multiples . . . 146
Générateur de formules 147
Accès au Générateur de formules. 147
Création d'expressions 147
Sélection de fonctions 147
Sélection de champs, de paramètres et de variables globales 150
Affichage ou sélection de valeurs. 151
Vérification des expressions CLEM 151
Rechercher et remplacer 152

Chapitre 10. Référence du langage CLEM. 155

Présentation de la référence CLEM 155
Types de données CLEM 155
Entiers 155
Réels 156
Caractères 156
Chaînes 156
Listes 156
Champs 157
Dates 157
Heure 158
Opérateurs CLEM 158
Informations de référence sur les fonctions . . . 160
Conventions pour la description des fonctions 161
Fonctions d'informations 162
Fonctions de conversion. 162
Fonctions de comparaison 164
Fonctions logiques. 166
Fonctions numériques 166
Fonctions trigonométriques. 167
Fonctions de probabilité 168
Fonctions spatiales 168
Opérations d'entiers sur les bits 169
Fonctions aléatoires 171
Fonctions sur chaînes. 171
Fonctions SoundEx 176
Fonctions date et heure 176
Fonctions séquentielles 180
Fonctions globales 186
Traitement des valeurs nulles et non renseignées 187
Champs spéciaux 188

Chapitre 11. Utilisation d'IBM SPSS Modeler avec un référentiel 191

A propos du IBM SPSS Collaboration and Deployment Services Repository 191
Stockage et déploiement des objets du référentiel 192
Connexion au référentiel 192
Saisie des informations d'identification pour le référentiel 192
Parcours des données d'identification de référentiel 193
Accès au contenu du référentiel 193
Stockage des objets dans le référentiel 193
Définition des propriétés d'un objet 193
Stockage des flux 196
Stockage de projets 196
Stockage des noeuds 197
Stockage des objets de sortie 197
Stockage des modèles et des palettes de modèles 197
Extraction d'objets du référentiel 198
Choix d'un objet à extraire 198

Sélection de la version d'objet 199
Recherche d'objets dans le référentiel 199
Modification des objets du référentiel 200
Création, changement de nom et suppression de dossiers 200
Verrouillage et déverrouillage d'objets de référentiel 201
Suppression d'objets du référentiel 201
Gestion des propriétés des objets du référentiel . . 202
Affichage des propriétés d'un dossier 202
Affichage et modification des propriétés des objets 202
Gestion des libellés de version d'objet 204
Déploiement de flux 204
Options de déploiement des flux 205
La branche de scoring 206

Chapitre 12. Exportation vers des applications externes 211
A propos de l'exportation vers des applications externes 211
Ouverture d'un flux dans IBM SPSS Modeler Advantage 211
Importation et exportation de modèles au format PMML 212
Types de modèle prenant en charge PMML . . 213

Chapitre 13. Projets et rapports . . . 215
Introduction aux projets 215
Vue CRISP-DM 215
Vue Classes 216
Création d'un projet 216
Création d'un projet 216
Ajout à un projet 216
Transfert de projets vers le IBM SPSS Collaboration and Deployment Services Repository 217
Définition des propriétés d'un projet 218
Annotation d'un projet 218
Propriétés de l'objet 219
Fermeture d'un projet 219
Génération d'un rapport 220
Enregistrement et exportation de rapports générés 221

Chapitre 14. Personnalisation d'IBM SPSS Modeler 223
Personnalisation des options d'IBM SPSS Modeler 223
Définition des options d'IBM SPSS Modeler . . . 223
Options du système 223
Définition des répertoires par défaut 224
Définition des options utilisateur 224
Personnalisation de la palette Noeuds 232
Personnalisation du gestionnaire de palette. . . 233
Modification d'une vue de l'onglet Palette . . . 235

Chapitre 15. Remarques relatives aux performances des flux et des noeuds . 237
Ordre des noeuds 237
Caches de noeud 238
Performances: noeuds d'exécution 239
Performances : noeuds de modélisation. 240
Performances : expressions CLEM 241

Chapitre 16. Accessibilité dans IBM SPSS Modeler 243
Présentation de l'accessibilité dans IBM SPSS Modeler 243
Types de prise en charge pour l'accessibilité . . . 243
Accessibilité pour les utilisateurs ayant une déficience visuelle 243
Accessibilité pour les utilisateurs aveugles. . . 244
Accessibilité via le clavier 244
Utilisation d'un lecteur d'écran 252
Conseils d'utilisation 253
Interférence avec d'autres logiciels 254
JAWS et Java 254
Utilisation des graphiques dans IBM SPSS Modeler 254

Chapitre 17. prise en charge d'Unicode 255
Prise en charge d'Unicode dans IBM SPSS Modeler 255

Remarques 257
Marques 258
Dispositions applicables à la documentation du produit 259

Index 261

Avis aux lecteurs canadiens

Le présent document a été traduit en France. Voici les principales différences et particularités dont vous devez tenir compte.

Illustrations

Les illustrations sont fournies à titre d'exemple. Certaines peuvent contenir des données propres à la France.

Terminologie

La terminologie des titres IBM peut différer d'un pays à l'autre. Reportez-vous au tableau ci-dessous, au besoin.

IBM France	**IBM Canada**
ingénieur commercial	représentant
agence commerciale	succursale
ingénieur technico-commercial	informaticien
inspecteur	technicien du matériel

Claviers

Les lettres sont disposées différemment : le clavier français est de type AZERTY, et le clavier français-canadien de type QWERTY.

OS/2 et Windows - Paramètres canadiens

Au Canada, on utilise :

- les pages de codes 850 (multilingue) et 863 (français-canadien),
- le code pays 002,
- le code clavier CF.

Nomenclature

Les touches présentées dans le tableau d'équivalence suivant sont libellées différemment selon qu'il s'agit du clavier de la France, du clavier du Canada ou du clavier des États-Unis. Reportez-vous à ce tableau pour faire correspondre les touches françaises figurant dans le présent document aux touches de votre clavier.

France	Canada	Etats-Unis
(Pos1)		Home
Fin	Fin	End
(PgAr)		PgUp
(PgAv)		PgDn
Inser	Inser	Ins
Suppr	Suppr	Del
Echap	Echap	Esc
Attn	Intrp	Break
Impr écran	ImpEc	PrtSc
Verr num	Num	Num Lock
Arrêt défil	Défil	Scroll Lock
(Verr maj)	FixMaj	Caps Lock
AltGr	AltCar	Alt (à droite)

Brevets

Il est possible qu'IBM détienne des brevets ou qu'elle ait déposé des demandes de brevets portant sur certains sujets abordés dans ce document. Le fait qu'IBM vous fournisse le présent document ne signifie pas qu'elle vous accorde un permis d'utilisation de ces brevets. Vous pouvez envoyer, par écrit, vos demandes de renseignements relatives aux permis d'utilisation au directeur général des relations commerciales d'IBM, 3600 Steeles Avenue East, Markham, Ontario, L3R 9Z7.

Assistance téléphonique

Si vous avez besoin d'assistance ou si vous voulez commander du matériel, des logiciels et des publications IBM, contactez IBM direct au 1 800 465-1234.

Chapitre 1. A propos d'IBM SPSS Modeler

IBM® SPSS Modeler est un ensemble d'outils d'exploration de données qui vous permet de développer rapidement, grâce à vos compétences professionnelles, des modèles prédictifs et de les déployer dans des applications professionnelles afin de faciliter la prise de décision. Conçu autour d'un modèle confirmé, le modèle CRISP-DM, IBM SPSS Modeler prend en charge l'intégralité du processus d'exploration de données, des données à l'obtention de meilleurs résultats commerciaux.

IBM SPSS Modeler propose différentes méthodes de modélisation issues des domaines de l'apprentissage automatique, de l'intelligence artificielle et des statistiques. Les méthodes disponibles dans la palette Modélisation vous permettent d'extraire de nouvelles informations de vos données et de développer des modèles prédictifs. Chaque méthode possède ses propres avantages et est donc plus adaptée à certains types de problème spécifiques.

Il est possible d'acquérir SPSS Modeler comme produit autonome ou de l'utiliser en tant que client en combinaison avec SPSS Modeler Server. Plusieurs autres options sont également disponibles, telles que décrites dans les sections suivantes. Pour plus d'informations, voir https://www.ibm.com/analytics/us/en/technology/spss/.

Produits IBM SPSS Modeler

La famille des produits IBM SPSS Modeler et les logiciels associés sont composés des éléments suivants.

- IBM SPSS Modeler
- IBM SPSS Modeler Server
- IBM SPSS Modeler Administration Console (inclus avec IBM SPSS Deployment Manager)
- IBM SPSS Modeler Batch
- IBM SPSS Modeler Solution Publisher
- Adaptateurs IBM SPSS Modeler Server pour IBM SPSS Collaboration and Deployment Services

IBM SPSS Modeler

SPSS Modeler est une version complète du produit que vous installez et exécutez sur votre ordinateur personnel. Pour obtenir de meilleures performances lors du traitement de jeux de données volumineux, vous pouvez exécuter SPSS Modeler en mode local, comme produit autonome, ou l'utiliser en mode réparti, en association avec IBM SPSS Modeler Server.

Avec SPSS Modeler, vous pouvez créer des modèles prédictifs précis rapidement et de manière intuitive, sans aucune programmation. L'interface visuelle unique vous permet de visualiser facilement le processus d'exploration de données. Grâce aux analyses avancées intégrées au produit, vous pouvez découvrir des motifs et tendances masqués dans vos données. Vous pouvez modéliser les résultats et comprendre les facteurs qui les influencent, afin d'exploiter les opportunités commerciales et de réduire les risques.

SPSS Modeler est disponible en deux éditions : SPSS Modeler Professional et SPSS Modeler Premium. Pour plus d'informations, voir la rubrique «Editions d'IBM SPSS Modeler», à la page 2.

IBM SPSS Modeler Server

Grâce à une architecture client/serveur, SPSS Modeler adresse les demandes d'opérations très consommatrices de ressources à un logiciel serveur puissant. Il offre ainsi des performances accrues sur des jeux de données plus volumineux.

© Copyright IBM Corp. 1994, 2019

SPSS Modeler Server est un produit avec licence distincte qui s'exécute en permanence en mode d'analyse réparti sur un hôte de serveur en combinaison avec une ou plusieurs installations d'IBM SPSS Modeler. Ainsi, SPSS Modeler Server fournit des performances supérieures sur de grands jeux de données car les opérations nécessitant beaucoup de mémoire peuvent être effectuées sur le serveur sans télécharger de données sur l'ordinateur client. IBM SPSS Modeler Server prend également en charge l'optimisation SQL et propose des capacités de modélisation dans la base de données pour des performances et une automatisation améliorées.

IBM SPSS Modeler Administration Console

Modeler Administration Console est une interface graphique permettant de gérer de nombreuses options de SPSS Modeler Server qui peuvent également être configurées au moyen d'un fichier d'options. La console est incluse dans IBM SPSS Deployment Manager et peut être utilisée pour surveiller et configurer vos installations SPSS Modeler Server ; elle est disponible gratuitement pour les clients actuels de SPSS Modeler Server. L'application ne peut être installée que sur des ordinateurs Windows ; en revanche, elle peut administrer un serveur installé sur n'importe quelle plate-forme prise en charge.

IBM SPSS Modeler Batch

Alors que l'exploration de données est généralement un processus interactif, il est également possible d'exécuter SPSS Modeler à partir d'une ligne de commande sans recourir à l'interface utilisateur graphique. Par exemple, vous pouvez avoir des tâches longue durée ou répétitives à exécuter sans intervention de l'utilisateur. SPSS Modeler Batch est une version spécifique du produit qui prend en charge toutes les capacités d'analyse de SPSS Modeler sans avoir besoin d'accéder à l'interface utilisateur standard. SPSS Modeler Server est requis pour utiliser SPSS Modeler Batch.

IBM SPSS Modeler Solution Publisher

SPSS Modeler Solution Publisher est un outil qui permet de créer une version « packagée » d'un flux SPSS Modeler qui peut être exécutée par un moteur Runtime externe ou intégrée dans une application externe. Ainsi, vous pouvez publier et déployer des flux SPSS Modeler complets dans des environnements où SPSS Modeler n'est pas installé. SPSS Modeler Solution Publisher est fourni avec le service IBM SPSS Collaboration and Deployment Services - Scoring et nécessite une licence distincte. Avec cette licence, vous recevez SPSS Modeler Solution Publisher Runtime qui vous permet d'exécuter les flux publiés.

Pour plus d'informations sur SPSS Modeler Solution Publisher, reportez-vous à la documentation IBM SPSS Collaboration and Deployment Services. Le Knowledge Center d'IBM SPSS Collaboration and Deployment Services contient des sections intitulées "IBM SPSS Modeler Solution Publisher" et "IBM SPSS Analytics Toolkit."

Adaptateurs IBM SPSS Modeler Server pour IBM SPSS Collaboration and Deployment Services

Différents adaptateurs pour IBM SPSS Collaboration and Deployment Services sont disponibles et permettent à SPSS Modeler et SPSS Modeler Server d'interagir avec un référentiel IBM SPSS Collaboration and Deployment Services. Ainsi, un flux SPSS Modeler déployé sur le référentiel peut être partagé par différents utilisateurs ou peut être accessible depuis l'application client léger IBM SPSS Modeler Advantage. Installez l'adaptateur sur le système qui héberge le référentiel.

Editions d'IBM SPSS Modeler

SPSS Modeler est disponible dans les éditions suivantes.

SPSS Modeler Professional

SPSS Modeler Professional offre tous les outils nécessaires à l'utilisation de la plupart des types de données structurées, tels que les comportements et interactions suivis dans les systèmes CRM, les

caractéristiques sociodémographiques, les comportements d'achat et les données de vente.

SPSS Modeler Premium

SPSS Modeler Premium est un produit avec licence distincte qui étend le champ d'applications de SPSS Modeler Professional afin de pouvoir traiter des données spécialisées et des données de texte non structurées. SPSS Modeler Premium inclut IBM SPSS Modeler Text Analytics :

IBM SPSS Modeler Text Analytics utilise des technologies linguistiques avancées et le traitement du langage naturel pour traiter rapidement une large variété de données textuelles non structurées, en extraire les concepts clés et les organiser pour les regrouper dans des catégories. Les concepts extraits et les catégories peuvent ensuite être combinés aux données structurées existantes, telles que les données démographiques, et appliqués à la modélisation grâce à la gamme complète d'outils d'exploration de données d'IBM SPSS Modeler, afin de favoriser une prise de décision précise et efficace.

IBM SPSS Modeler Subscription

IBM SPSS Modeler Subscription propose les mêmes fonctionnalités d'analyse prédictive que le client IBM SPSS Modeler traditionnel. L'édition Subscription vous permet de télécharger régulièrement des mises à jour de produit.

Documentation

Une documentation est disponible dans le menu Aide de SPSS Modeler. Elle ouvre le composant Knowledge Center en ligne, qui est toujours disponible en dehors du produit.

La documentation complète de chaque produit (y compris les instructions d'installation) au format PDF est également disponible dans un dossier compressé distinct, avec le téléchargement du produit. Vous pouvez également télécharger les derniers documents PDF sur Internet à l'adresse http://www.ibm.com/support/docview.wss?uid=ibm10874788.

Documentation de SPSS Modeler Professional

La suite de documentation SPSS Modeler Professional (à l'exception des instructions d'installation) est la suivante.

- **IBM SPSS Modeler - Guide d'utilisation.** Introduction générale à SPSS Modeler : création de flux de données, traitement des valeurs manquantes, création d'expressions CLEM, utilisation de projets et de rapports et regroupement des flux pour le déploiement dans IBM SPSS Collaboration and Deployment Services ou IBM SPSS Modeler Advantage.
- **Noeuds source, de processus et de sortie d'IBM SPSS Modeler.** Descriptions de tous les noeuds utilisés pour lire, traiter et renvoyer les données de sortie dans différents formats. En pratique, cela signifie tous les noeuds autres que les noeuds de modélisation.
- **Noeuds modélisation d'IBM SPSS Modeler.** Descriptions de tous les noeuds utilisés pour créer des modèles d'exploration de données. IBM SPSS Modeler propose différentes méthodes de modélisation issues des domaines de l'apprentissage automatique, de l'intelligence artificielle et des statistiques.
- **Guide des applications IBM SPSS Modeler.** Les exemples de ce guide fournissent des introductions brèves et ciblées aux méthodes et techniques de modélisation. Une version en ligne de ce guide est également disponible dans le menu Aide. Pour plus d'informations, voir la rubrique «Exemples d'application», à la page 4.
- **IBM SPSS Modeler - Génération de scripts Python et d'automatisation** Ce manuel fournit des informations sur l'automatisation du système via des scripts Python, notamment sur les propriétés pouvant être utilisées pour manipuler les noeuds et les flux.
- **Guide de déploiement d'IBM SPSS Modeler.** Informations sur l'exécution des flux IBM SPSS Modeler comme étapes des travaux d'exécution sous IBM SPSS Deployment Manager.

- **Guide du développeur IBM SPSS Modeler CLEF.** CLEF permet d'intégrer des programmes tiers tels que des programmes de traitement de données ou des algorithmes de modélisation en tant que noeuds dans IBM SPSS Modeler.
- **Guide d'exploration des bases de données IBM SPSS Modeler.** Informations sur la manière de tirer parti de la puissance de votre base de données pour améliorer les performances et étendre la gamme des capacités d'analyse via des algorithmes tiers.
- **Guide des performances et d'administration d'IBM SPSS Modeler Server.** Informations sur le mode de configuration et d'administration d'IBM SPSS Modeler Server.
- **IBM SPSS Deployment Manager - Guide d'utilisation.** Informations sur l'utilisation de l'interface utilisateur de la console d'administration incluses dans l'application Deployment Manager pour la surveillance et la configuration d'IBM SPSS Modeler Server.
- **Guide CRISP-DM d'IBM SPSS Modeler.** Guide détaillé sur l'utilisation de la méthodologie CRISP-DM pour l'exploration de données avec SPSS Modeler
- **IBM SPSS Modeler Batch - Guide d'utilisation.** Guide complet sur l'utilisation d'IBM SPSS Modeler en mode de traitement par lots, avec des détails sur l'exécution en mode de traitement par lots et les arguments de ligne de commande. Ce guide est disponible au format PDF uniquement.

Documentation de SPSS Modeler Premium

La suite de documentation SPSS Modeler Premium (à l'exception des instructions d'installation) est la suivante.

- **SPSS Modeler Text Analytics - Guide d'utilisation.** Informations sur l'utilisation des analyses de texte avec SPSS Modeler, notamment sur les noeuds Text Mining, l'espace de travail interactif, les modèles et d'autres ressources.

Exemples d'application

Tandis que les outils d'exploration de données de SPSS Modeler peuvent vous aider à résoudre une grande variété de problèmes métier et organisationnels, les exemples d'application fournissent des introductions brèves et ciblées aux méthodes et aux techniques de modélisation. Les jeux de données utilisés ici sont beaucoup plus petits que les énormes entrepôts de données gérés par certains Data Miners, mais les concepts et les méthodes impliqués peuvent être adaptés à des applications réelles.

Pour accéder aux exemples, cliquez sur **Exemples d'application** dans le menu Aide de SPSS Modeler.

Les fichiers de données et les flux d'échantillons sont installés dans le dossier `Demos`, sous le répertoire d'installation du produit. Pour plus d'informations, voir «Dossier Demos».

Exemples de modélisation de bases de données. Consultez les exemples dans le document *IBM SPSS Modeler Guide d'exploration de base de données*.

Exemples de génération de scripts. Consultez les exemples dans le document *IBM SPSS Modeler Guide de génération de scripts et d'automatisation*.

Dossier Demos

Les fichiers de données et les flux d'échantillons utilisés avec les exemples d'application sont installés dans le dossier `Demos`, sous le répertoire d'installation du produit (par exemple : `C:\Program Files\IBM\SPSS\Modeler\<version>\Demos`). Ce dossier est également accessible à partir du groupe de programmes IBM SPSS Modeler, dans le menu Démarrer de Windows ou en cliquant sur `Demos` dans la liste des répertoires récents de la boîte de dialogue **Fichier** > **Ouvrir un flux**.

Suivi des licences

Lorsque vous utilisez SPSS Modeler, l'utilisation des licences est suivie et consignée à intervalles réguliers. Les métriques de licence consignées sont *AUTHORIZED_USER* et *CONCURRENT_USER* et le type de métrique consigné dépend du type de licence dont vous disposez pour SPSS Modeler.

Les fichiers journaux générés peuvent être traités par IBM License Metric Tool, à partir duquel vous pouvez générer des rapports d'utilisation de licence.

Les fichiers journaux des licences sont créés dans le répertoire dans lequel les fichiers journaux de SPSS Modeler Client sont enregistrés (par défaut, `%ALLUSERSPROFILE%/IBM/SPSS/Modeler/<version>/log`).

Chapitre 2. Nouvelles fonctionnalités dans IBM SPSS Modeler 18.2.1

L'édition IBM SPSS Modeler ajoute les fonctionnalités ci-après.

Après avoir cliqué avec le bouton droit de la souris sur un noeud de données et sélectionné l'option Afficher les données, les nouveaux graphiques suivants ont été ajoutés à l'onglet Graphiques. Le tableau de bord a également été repensé.

- Cercles imbriqués
- Bulle
- Graphique de type rivière
- Tracé de temps

© Copyright IBM Corp. 1994, 2019

Chapitre 3. Présentation du produit

Démarrage

En tant qu'application d'exploration de données, IBM SPSS Modeler constitue une méthode stratégique de recherche de relations utiles dans les grands jeux de données. Contrairement aux méthodes statistiques plus traditionnelles, il n'est pas indispensable de savoir ce que vous recherchez avant de commencer. Vous pouvez explorer vos données, créer divers modèles et explorer diverses relations, jusqu'à ce que vous trouviez des informations utiles.

Démarrage d'IBM SPSS Modeler

Pour démarrer l'application, cliquez :

Démarrer > [Tous] Programmes > IBM SPSS Modeler <version> > IBM SPSS Modeler <version>

La fenêtre principale apparaît après quelques secondes.

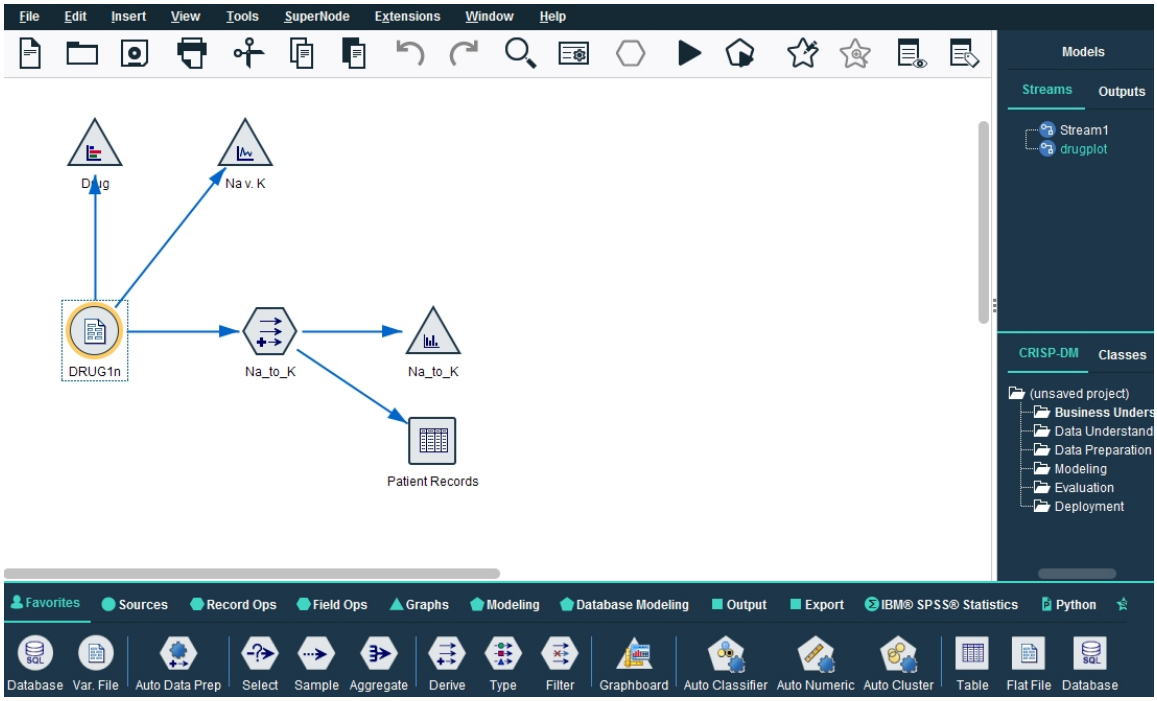


Figure 1. fenêtre d'application principale d'IBM SPSS Modeler

Lancement de l'application à partir de la ligne de commande

Vous pouvez utiliser la ligne de commande de votre système d'exploitation pour lancer IBM SPSS Modeler comme suit :

1. Dans le cas d'un ordinateur sur lequel est installé IBM SPSS Modeler, ouvrez une fenêtre DOS ou une invite de commande.
2. Pour lancer l'interface IBM SPSS Modeler en mode interactif, tapez la commande `modelerclient` suivie des arguments souhaités, par exemple :

```
modelerclient -stream report.str -execute
```

Les arguments disponibles (indicateurs) vous permettent de vous connecter à un serveur, de charger des flux, d'exécuter des scripts, ou d'indiquer les autres paramètres nécessaires.

Connexion à IBM SPSS Modeler Server

Il est possible d'exécuter IBM SPSS Modeler comme une application autonome ou un comme un client connecté directement à IBM SPSS Modeler Server ou à IBM SPSS Modeler Server or à un cluster de serveurs par le biais du Coordinateur des processus connecté à partir d'IBM SPSS Collaboration and Deployment Services. Le statut de la connexion apparaît en bas à gauche de la fenêtre IBM SPSS Modeler.

Lorsque vous souhaitez vous connecter à un serveur, vous pouvez saisir manuellement son nom ou sélectionner un nom que vous aurez préalablement défini. En revanche, si vous avez IBM SPSS Collaboration and Deployment Services, vous avez la possibilité de chercher dans une liste de serveurs ou de clusters de serveurs à partir de la boîte de dialogue Connexion au serveur. Vous pouvez naviguer via les services Statistiques s'exécutant sur un réseau grâce au Coordinateur des processus.

Pour vous connecter à un serveur

1. Dans le menu Outils, cliquez sur **Connexion au serveur**. La boîte de dialogue Connexion au serveur s'affiche. Vous pouvez également double-cliquez sur la zone de statut de la connexion dans la fenêtre IBM SPSS Modeler.
2. Dans la boîte de dialogue, indiquez les options de connexion à l'ordinateur du serveur local ou sélectionnez une connexion dans le tableau.
   - Cliquez sur **Ajouter** ou **Modifier** pour ajouter ou modifier une connexion. Pour plus d'informations, voir la rubrique «Ajout et modification d'une connexion à IBM SPSS Modeler Server», à la page 11.
   - Cliquez sur **Rechercher** pour accéder au serveur ou à un cluster de serveurs dans le Coordinateur de processus. Pour plus d'informations, voir la rubrique «Recherche de serveurs dans IBM SPSS Collaboration and Deployment Services», à la page 11.

   Tableau de serveur. Ce tableau comprend un ensemble de connexions au serveur définies. Il affiche la connexion par défaut, le nom du serveur, sa description et le numéro du port. Vous pouvez ajouter manuellement une nouvelle connexion ainsi que sélectionner ou rechercher une connexion existante. Pour définir un serveur particulier comme connexion par défaut, cochez la case dans la colonne Par défaut du tableau de la connexion.

   Chemin de données par défaut. Indiquez le chemin d'accès aux données situées sur l'ordinateur serveur. Cliquez sur le bouton **...** pour accéder à l'emplacement requis.

   Définir les données d'identification. Laissez cette case décochée pour permettre à la fonction de **connexion unique** de se connecter au serveur à l'aide de vos informations de nom d'utilisateur et de mot de passe locaux. Si la connexion unique n'est pas disponible, ou si vous cochez la case pour désactiver la connexion unique (par exemple pour vous connecter à un compte administrateur), les champs suivants sont activés et vous permettent d'entrer vos informations d'identification.

   ID utilisateur. Entrez le nom d'utilisateur avec lequel effectuer la connexion au serveur.

   Mot de passe. Entrez le mot de passe associé au nom d'utilisateur défini.

   Domaine. Indiquez le domaine utilisé pour la connexion au serveur. Le nom de domaine n'est requis que si l'ordinateur serveur se trouve dans un autre domaine Windows que l'ordinateur client.
3. Cliquez sur **OK** pour terminer la connexion.

Pour se déconnecter d'un serveur

1. Dans le menu Outils, cliquez sur **Connexion au serveur**. La boîte de dialogue Connexion au serveur s'affiche. Vous pouvez également double-cliquez sur la zone de statut de la connexion dans la fenêtre IBM SPSS Modeler.
2. Dans la boîte de dialogue, sélectionnez le serveur local puis cliquez sur **OK**.

Ajout et modification d'une connexion à IBM SPSS Modeler Server

Dans la boîte de dialogue Connexion au serveur, vous pouvez modifier ou ajouter une connexion au serveur. Cliquez sur Ajouter pour accéder à une boîte de dialogue Ajouter/Modifier un serveur non renseignée, dans laquelle vous pourrez entrer les données de la connexion au serveur. Si vous sélectionnez une connexion existante et cliquez sur Modifier, dans la boîte de dialogue Connexion au serveur, la boîte de dialogue Ajouter/Modifier un serveur s'ouvre, affichant les données de cette connexion, vous permettant ainsi d'apporter toutes les modifications que vous souhaitez.

Remarque : Vous ne pouvez pas modifier une connexion au serveur qui a été ajoutée à partir d'IBM SPSS Collaboration and Deployment Services, car le nom, le port et d'autres détails sont définis dans IBM SPSS Collaboration and Deployment Services. Les pratiques recommandées indiquent d'utiliser les mêmes ports pour communiquer avec IBM SPSS Collaboration and Deployment Services et avec SPSS Modeler Client. Ces ports peuvent être définis via les paramètres `max_server_port` et `min_server_port` dans le fichier options.cfg.

Pour ajouter des connexions au serveur

1. Dans le menu Outils, cliquez sur **Connexion au serveur**. La boîte de dialogue Connexion au serveur s'affiche.
2. Dans la boîte de dialogue, cliquez sur **Ajouter**. La boîte de dialogue Ajouter/Modifier un serveur pour la connexion au serveur s'ouvre.
3. Saisissez les données de connexion au serveur puis cliquez sur **OK** pour enregistrer la connexion et retourner à la boîte de dialogue Connexion au serveur.

- **Serveur.** Indiquez un serveur disponible ou sélectionnez-en un dans la liste. L'ordinateur serveur peut être identifié par un nom alphanumérique (par exemple, *monserveur*) ou une adresse IP qui lui est affectée (par exemple, 202.123.456.78).
- **Port.** Indiquez le numéro de port d'écoute du serveur. Si ce port par défaut ne fonctionne pas, demandez à l'administrateur système le numéro de port correct.
- **Description.** Saisissez une description optionnelle pour la connexion à ce serveur.
- **Coder la connexion (utiliser SSL).** Indique si une connexion SSL (**Secure Sockets Layer**) doit être utilisée. Le protocole SSL est fréquemment utilisé pour la sécurisation des données sur un réseau. Pour pouvoir utiliser cette fonction, vous devez activer le protocole SSL sur le serveur hébergeant IBM SPSS Modeler Server. Si nécessaire, contactez votre administrateur local pour plus d'informations.

Pour modifier des connexions au serveur

1. Dans le menu Outils, cliquez sur **Connexion au serveur**. La boîte de dialogue Connexion au serveur s'affiche.
2. Dans la boîte de dialogue, sélectionnez la connexion que vous souhaitez modifier puis cliquez sur **Modifier**. La boîte de dialogue Ajouter/Modifier un serveur pour la connexion au serveur s'ouvre.
3. Modifiez les données de connexion au serveur puis cliquez sur **OK** pour enregistrer les changements et retourner à la boîte de dialogue Connexion au serveur.

Recherche de serveurs dans IBM SPSS Collaboration and Deployment Services

Au lieu d'entrer manuellement une connexion au serveur, vous pouvez sélectionner un serveur ou un cluster de serveurs disponible sur le réseau par le biais du Coordinateur de processus, disponible dans IBM SPSS Collaboration and Deployment Services. Un cluster de serveurs contient plusieurs serveurs, et permet au Coordinateur de processus de déterminer le serveur qui répond le mieux à la demande de traitement.

Bien que vous ne vous puissiez pas ajouter manuellement de serveurs dans la boîte de dialogue Connexion au serveur, la recherche de serveurs disponibles vous permet de vous connecter aux serveurs sans que vous ayez besoin de connaître le nom du serveur et le numéro du port. Ces informations sont fournies automatiquement. Il vous faut néanmoins corriger les données de connexion telles que le nom de l'utilisateur, le domaine et le mot de passe.

Remarque : Si vous n'avez pas accès au Coordinateur de processus, vous pouvez tout de même saisir manuellement le nom du serveur auquel vous souhaitez vous connecter ou sélectionner un nom que vous aurez défini au préalable. Pour plus d'informations, voir la rubrique «Ajout et modification d'une connexion à IBM SPSS Modeler Server», à la page 11.

Pour rechercher des serveurs et des clusters de serveurs

1. Dans le menu Outils, cliquez sur **Connexion au serveur**. La boîte de dialogue Connexion au serveur s'affiche.
2. Cliquez sur **Rechercher** pour ouvrir la boîte de dialogue Recherche de serveurs. Si vous n'êtes plus connecté à IBM SPSS Collaboration and Deployment Services, lors de votre tentative d'accès au Coordinateur de processus, il vous sera demandé de vous reconnecter.
3. Sélectionnez le serveur ou le cluster de serveurs dans la liste.
4. Cliquez sur **OK** pour fermer la boîte de dialogue et ajouter cette connexion au tableau de la boîte de dialogue Connexion au serveur.

Connexion à Analytic Server

Si plusieurs serveurs Analytic Server sont disponibles, vous pouvez utiliser la boîte de dialogue Connexion à Analytic Server pour définir plusieurs serveurs à utiliser dans IBM SPSS Modeler. Votre administrateur a peut-être déjà configuré un serveur Analytic Server par défaut dans le fichier `<chemin_installation_Modeler>/config/options.cfg`. Toutefois, vous pouvez utiliser d'autres serveurs disponibles après les avoir définis. Par exemple, si vous utilisez les noeuds Source et Exportation d'Analytic Server, vous pouvez utiliser des connexions Analytic Server dans différentes branches d'un flux afin que lorsque chaque branche est exécutée, elle utilise son propre serveur Analytic Server, sans extraction de données dans IBM SPSS Modeler Server. Notez que si une branche contient plusieurs connexions Analytic Server, les données seront extraites des serveurs Analytic Server vers IBM SPSS Modeler Server. Pour plus d'informations, notamment sur les restrictions, voir «Propriétés du flux Analytic Server», à la page 50.

Pour créer une connexion Analytic Server, accédez à **Outils** > **Connexions Analytic Server** et spécifiez les informations requises dans les sections ci-après de la boîte de dialogue.

Connexion

URL. Entrez l'URL d'Analytic Server au format `https://nomhôte:port/racinecontexte`, `nomhôte` correspondant à l'adresse IP ou au nom d'hôte d'Analytic Server, `port`, au numéro de port, et `racinecontexte`, à la racine de contexte d'Analytic Server.

Locataire. Entrez le nom du locataire dont IBM SPSS Modeler Server est membre. Si vous ne connaissez pas le locataire, contactez votre administrateur.

Authentification

Mode. Sélectionnez l'un des modes d'authentification ci-après.

- **Nom d'utilisateur et mot de passe**. Vous devez entrer le nom d'utilisateur et le mot de passe.
- **Données d'identification stockées**. Vous devez sélectionner des informations d'identification dans IBM SPSS Collaboration and Deployment Services Repository.
- **Kerberos**. Vous devez entrer le nom principal du service et le chemin d'accès au fichier de configuration. Si vous ne connaissez pas ces informations, contactez votre administrateur.

Nom d'utilisateur. Entrez le nom d'utilisateur d'Analytic Server.

Domaines. Sélectionnez le domaine à utiliser pour la connexion à Analytic Server.

Mot de passe. Entrez le mot de passe d'Analytic Server.

Connecter. Cliquez sur **Connecter** pour tester la nouvelle connexion.

Connexions

Une fois que vous avez spécifié les informations ci-dessus et cliqué sur **Connecter**, la connexion est ajoutée à cette table Connexions. Si vous devez supprimer une connexion, sélectionnez-la et cliquez sur **Supprimer**.

Si votre administrateur a défini une connexion Analytic Server par défaut dans le fichier `options.cfg`, vous pouvez cliquer sur **Ajouter une connexion par défaut** pour l'ajouter également à vos connexions disponibles. Vous serez invité à entrer le nom d'utilisateur et le mot de passe.

Changement de répertoire temporaire

Certaines opérations effectuées par IBM SPSS Modeler Server peuvent nécessiter la création de fichiers temporaires. Par défaut, IBM SPSS Modeler crée les fichiers temporaires dans le répertoire temporaire du système. Vous pouvez modifier l'emplacement du répertoire temporaire en effectuant les opérations suivantes.

1. Créez un répertoire intitulé `spss` et un sous-répertoire intitulé *servertemp*.
2. Editez le fichier *options.cfg*, situé dans le répertoire */config* du dossier d'installation d'IBM SPSS Modeler. Editez le paramètre `temp_directory` de ce fichier en saisissant : `temp_directory, "C:/spss/servertemp"`.
3. Redémarrez ensuite le service IBM SPSS Modeler Server. Pour ce faire, cliquez sur **Services** dans les outils d'administration du Panneau de configuration Windows. Il vous suffit d'arrêter le service et de le redémarrer pour appliquer les modifications apportées. Vous pouvez redémarrer l'ordinateur pour redémarrer le service.

Tous les fichiers temporaires sont désormais écrits dans ce nouveau répertoire.

Remarque :

- Des barres obliques doivent être utilisées.
- Le paramètre `temp_directory` n'est pas applicable lors de l'exécution de flux d'évaluation via des travaux IBM SPSS Collaboration and Deployment Services. Lorsque vous exécutez ce type de travail, un fichier temporaire est créé. Par défaut, le fichier est enregistré dans le répertoire d'installation d'IBM SPSS Modeler Server. Vous pouvez changer le dossier de données par défaut dans lequel les fichiers temporaires sont enregistrés lorsque vous créez la connexion IBM SPSS Modeler Server dans IBM SPSS Modeler.

Démarrage de plusieurs sessions IBM SPSS Modeler

Si vous devez lancer plus d'une session IBM SPSS Modeler à la fois, vous devez effectuer certaines modifications de paramètres d'IBM SPSS Modeler et de Windows. Par exemple, il vous faudra effectuer ces modifications si vous avez deux licences de serveur distinctes et que vous souhaitez exécuter deux flux pour deux serveurs distincts du même ordinateur client.

Pour activer plusieurs sessions IBM SPSS Modeler :

1. Cliquez sur :

   Démarrer > **[Tous] Programmes** > **IBM SPSS Modeler**
2. Dans le raccourci d'IBM SPSS Modeler (celui avec l'icône), cliquez avec le bouton droit de la souris et sélectionnez **Propriétés**.
3. Dans la zone de texte **Cible**, ajoutez `-noshare` à la fin de la chaîne.
4. Dans Windows Explorer, sélectionnez :

   Outils > **Options des dossiers...**
5. Dans l'onglet Types de fichier, sélectionnez l'option Flux d'IBM SPSS Modeler et cliquez sur **Avancé**.

6. Dans la boîte de dialogue Modifier le type de fichier, sélectionnez Ouvrir avec IBM SPSS Modeler et cliquez sur **Modifier**.
7. Dans la zone de texte **Application utilisée pour effectuer l'action**, ajoutez `-noshare` avant l'argument **-flux**.

Interface IBM SPSS Modeler en un clin d'oeil

A chaque étape du processus d'exploration de données, l'interface conviviale d'IBM SPSS Modeler met à contribution vos compétences professionnelles. Les algorithmes de modélisation, comme la prévision, la classification, la segmentation et la détection d'association, permettent l'obtention de modèles performants et précis. Les résultats du modèle peuvent facilement être déployés et lus dans des bases de données, dans IBM SPSS Statistics et dans de nombreuses autres applications.

Travailler avec IBM SPSS Modeler est un processus en trois étapes de travail avec les données.

- Pour commencer, lisez les données d'IBM SPSS Modeler.
- Ensuite, exécutez les données par une série de manipulations.
- Et pour finir, envoyez les données vers une destination choisie.

Cette séquence d'opérations est appelée **flux de données** car les données circulent, enregistrement par enregistrement, de la source à la destination (modèle ou type de sortie de données), en passant par chaque manipulation.

Figure 2. Un flux simple

Espace de travail de flux IBM SPSS Modeler

L'espace de travail de flux est la plus grande zone de la fenêtre IBM SPSS Modeler. C'est dans cette zone que vous créez et manipulez les flux de données.

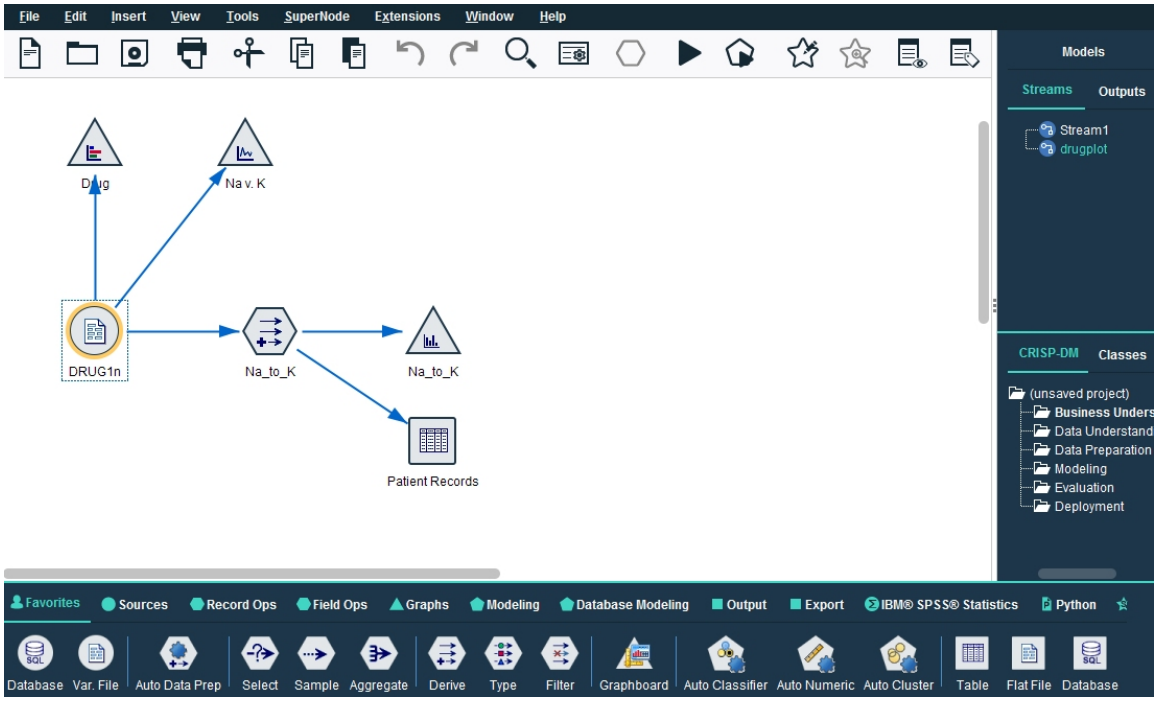


Figure 3. Espace de travail d'IBM SPSS Modeler (vue par défaut)

Les flux sont créés en dessinant des diagrammes des opérations de données nécessaires à votre entreprise sur l'espace de travail principal de l'interface. Chaque opération est représentée par une icône ou un **noeud**, et les noeuds sont reliés entre eux dans un **flux** représentant le flux de données passant par chaque opération.

Vous pouvez utiliser plusieurs flux à la fois dans IBM SPSS Modeler, que ce soit dans le même espace de travail de flux ou par l'ouverture d'un nouveau flux. Au cours d'une session, les flux sont stockés dans le gestionnaire de flux, en haut à droite de la fenêtre IBM SPSS Modeler.

Remarque : si vous utilisez un MacBook et que le paramètre **Force Click and haptic feedback** du trackpad intégré est activé, le déplacement à l'aide de la souris de noeuds de la palette de noeuds vers le canevas de flux peut entraîner l'ajout de noeuds en double au canevas. Pour éviter ce problème, il est recommandé de désactiver la préférence système **Force Click and haptic feedback** du trackpad.

Palette de noeuds

La plupart des données et des outils de modélisation de SPSS Modeler sont disponibles dans la *Palette de noeuds*, au bas de la fenêtre sous l'espace de travail de flux.

Par exemple, l'onglet de la palette **Ops. sur lignes** contient des noeuds permettant d'effectuer des opérations sur les *lignes*, telles que la sélection, la fusion et l'ajout.

Pour ajouter des noeuds à l'espace de travail, double-cliquez sur les icônes de la palette de noeuds ou faites glisser les icônes vers l'espace de travail. Vous pouvez ensuite les relier afin de créer un *flux* représentant le flux des données.

Figure 4. Onglet Ops sur lignes de la palette de noeuds

Chaque onglet de palette contient un ensemble de noeuds connexes employés pour différentes étapes des opérations de flux, comme :

- Les noeuds **Sources** amènent les données dans SPSS Modeler.
- Les noeuds **Ops sur lignes** effectuent des opérations sur les *lignes* de données, comme la sélection, la fusion et l'ajout.
- Les noeuds **Ops sur champs** effectuent des opérations sur les *champs* de données, comme le filtrage, le calcul de nouveaux champs et la détermination du niveau de mesure de champs particuliers.
- Les noeuds **Graphiques** affichent graphiquement les données avant et après la modélisation. Les graphiques peuvent être des tracés, des histogrammes, des noeuds Relations, ainsi que des graphiques d'évaluation.
- Les noeuds **Modélisation** utilisent les algorithmes de modélisation disponibles dans SPSS Modeler, comme les réseaux de neurones, les arbres décision, les algorithmes de classification et la création de séquences de données.
- Les noeuds de **Modélisation de la base de données** utilisant les algorithmes de modélisation disponibles dans les bases de données Microsoft SQL Server, IBM Db2, Oracle et Netezza.
- Les noeuds **Sortie** produisent diverses sorties pour les données, les graphiques et les résultats de modèles qui peuvent être affichés dans SPSS Modeler.
- Les noeuds **Exporter** produisent diverses sorties qui peuvent être affichées dans des applications externes telles que IBM SPSS Data Collection ou Excel.
- Les noeuds **IBM SPSS Statistics** importent des données à partir de, ou exportent des données vers IBM SPSS Statistics, et exécutent aussi des procédures IBM SPSS Statistics.
- Les noeuds **Python** permettent d'exécuter des algorithmes Python.
- **Spark.** Les noeuds peuvent être utilisés pour exécuter des algorithmes Spark.

Au fur et à mesure que vous maîtrisez mieux l'application SPSS Modeler, vous pouvez personnaliser le contenu de la palette en fonction de vos besoins.

A gauche de la palette des noeuds, vous pouvez filtrer les noeuds qui s'affichent en sélectionnant Supervisé, Association ou Segmentation.

Situé sous la palette de noeuds, un panneau de rapports fournit des informations sur la progression des diverses opérations, telles que la lecture des données dans le flux de données. Egalement situé sous la palette de noeuds, un panneau de statut fournit des informations sur l'activité actuelle de l'application, ainsi que des indications lorsqu'une saisie par l'utilisateur est requise.

Remarque : si vous utilisez un MacBook et que le paramètre **Force Click and haptic feedback** du trackpad intégré est activé, le déplacement à l'aide de la souris de noeuds de la palette de noeuds vers le canevas de flux peut entraîner l'ajout de noeuds en double au canevas. Pour éviter ce problème, il est recommandé de désactiver la préférence système **Force Click and haptic feedback** du trackpad.

Gestionnaires IBM SPSS Modeler

En haut à droite de la fenêtre se trouve le panneau des gestionnaires. Il contient trois onglets qui permettent de gérer les flux, les sorties et les modèles.

Vous pouvez utiliser l'onglet Flux pour ouvrir, renommer, enregistrer et supprimer les flux créés dans une session.

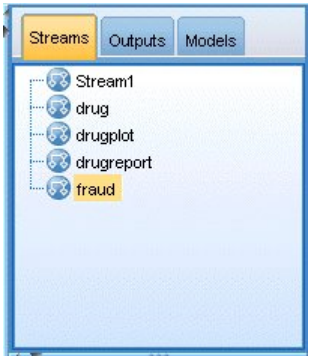


Figure 5. Onglet Flux

Figure 6. Onglet Sorties

L'onglet Sortie contient différents fichiers, tels que des graphiques et des tableaux, produits par des opérations de flux dans IBM SPSS Modeler. Vous pouvez afficher, sauvegarder, renommer et fermer les tableaux, les graphiques et les rapports qui figurent dans cet onglet.

Figure 7. Onglet Modèles qui contient des nugget de modèles

L'onglet Modèle est le plus puissant des onglets du gestionnaire. Cet onglet contient tous les **nuggets** de modèle qui contiennent les modèles générés dans IBM SPSS Modeler, pour la session en cours. Vous pouvez accéder à ces modèles directement à partir de l'onglet Modèles ou les ajouter au flux dans l'espace de travail.

Projets IBM SPSS Modeler

Dans la partie inférieure droite de la fenêtre se trouve le panneau de projet qui permet de créer et de gérer les **projets** d'exploration de données (groupes de fichiers en rapport avec une tâche d'exploration de données). Vous pouvez afficher les projets créés de deux façons dans IBM SPSS Modeler dans la vue Classes et dans la vue CRISP-DM.

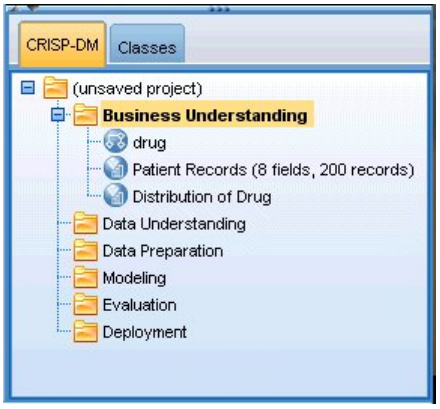


Figure 8. Vue CRISP-DM

L'onglet CRISP-DM permet d'organiser les projets en fonction de la méthodologie Cross-Industry Standard Process for Data mining commune utilisée dans le domaine. Que vous soyez un utilisateur chevronné ou novice, l'outil CRISP vous aidera à mieux organiser et communiquer vos efforts.

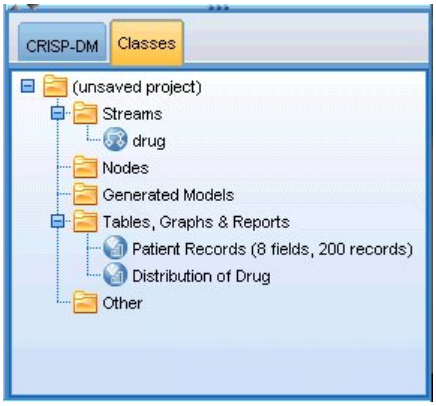


Figure 9. Vue Classes

L'onglet Classes permet d'organiser votre travail dans IBM SPSS Modeler en catégories, selon les types d'objet que vous créez. Cette vue est utile lorsque vous effectuez l'inventaire des données, des flux et des modèles.

Barre d'outils IBM SPSS Modeler

Une barre d'outils, composée d'icônes fournissant des options très utiles, se trouve en haut de la fenêtre IBM SPSS Modeler. Voici les boutons de la barre d'outils et leurs fonctions.

Créer un flux

Permet d'ouvrir un flux

Enregistrer le flux

Imprimer le flux en cours

Déplacer la sélection vers le Presse-papiers

Copier dans le Presse-papiers

Coller le contenu du Presse-papiers dans la sélection

Annuler la dernière action

Rétablir la dernière action

Recherche de noeuds

Editer les propriétés du flux

Aperçu de génération SQL

Exécuter le flux en cours

Exécuter la sélection de flux

Arrêter le flux (actif uniquement pendant l'exécution du flux)

Ajouter un super noeud

Zoom avant (super noeuds uniquement)

Zoom arrière (super noeuds uniquement)

Aucun balisage dans le flux

Insérer un commentaire

Masquer le balisage de flux (le cas échéant)

Afficher le balisage de flux masqué

Ouvrir un flux dans IBM SPSS Modeler Advantage

Le balisage de flux se compose des commentaires de flux, des liens de modèle et des indications de branche de scoring.

Les liens de modèle sont décrits dans le guide *Noeuds de modélisation IBM SPSS*.

Personnalisation de la barre d'outils

Vous pouvez modifier plusieurs aspects de la barre d'outils, tels que :

- Choisir si elle sera affichée ou non
- Choisir si les icônes comporteront ou non des info-bulles
- Choisir si elle utilisera des petites ou des grandes icônes

activer ou désactiver l'affichage de la barre d'outils :

1. Dans le menu principal, cliquez sur :
   Vue > **Barre d'outils** > **Afficher**

Pour modifier les paramètres des info-bulles ou de la taille des icônes :

1. Dans le menu principal, cliquez sur :
   Vue > **Barre d'outils** > **Personnaliser**

Cliquez sur **Afficher les info-bulles** ou **Gros boutons** le cas échéant.

Personnalisation de la fenêtre IBM SPSS Modeler

Vous pouvez utiliser les séparateurs situés entre les différentes zones de l'interface SPSS Modeler pour redimensionner ou fermer des outils en fonction de vos besoins. Par exemple, si vous travaillez avec un flux volumineux, vous pouvez utiliser les petites flèches situées sur chaque séparateur pour fermer la palette de noeuds, le panneau des gestionnaires et le panneau des projets. Ainsi, vous agrandissez l'espace de travail de flux et libérez suffisamment d'espace pour les flux volumineux ou multiples.

A partir du menu Vue, vous pouvez aussi cliquer sur **Palette de noeuds**, **Gestionnaires** ou **Projet** pour activer ou désactiver l'affichage de ces éléments.

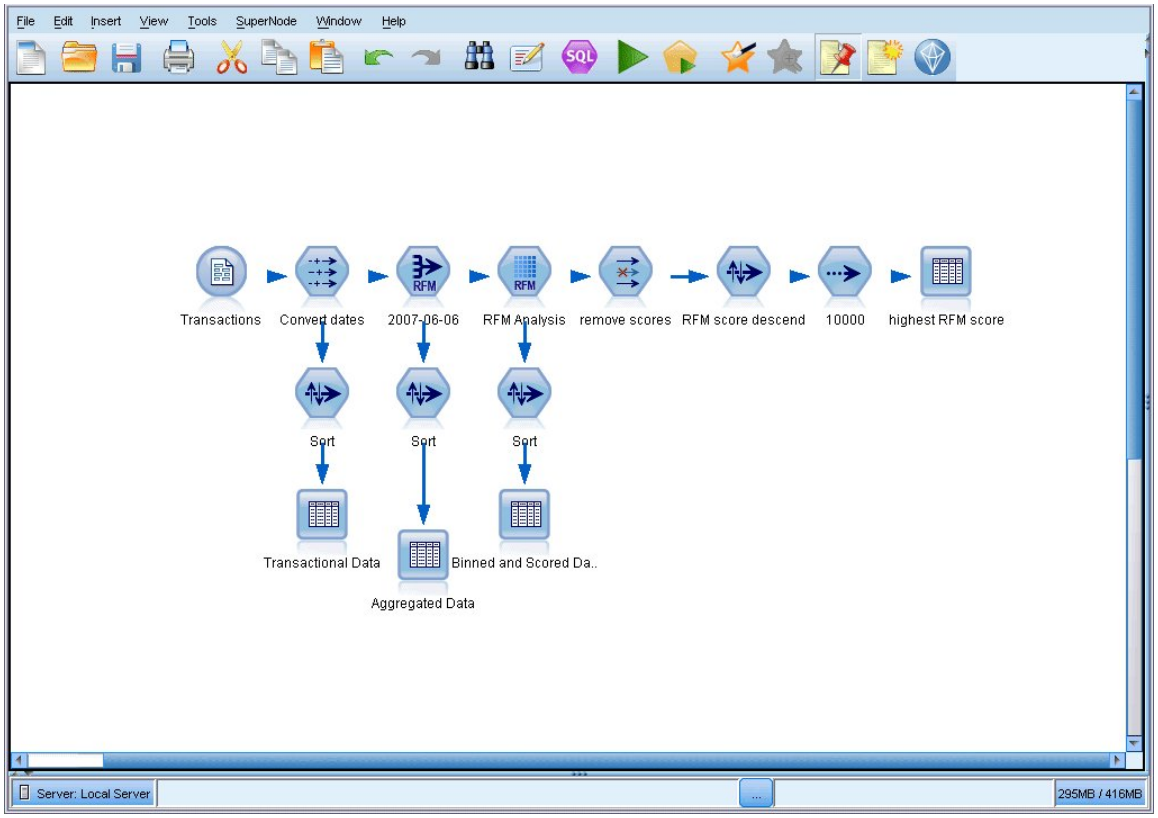


Figure 10. Espace de travail de flux agrandi

Vous pouvez également garder ouvertes la palette des noeuds, et les panneaux des gestionnaires et des projets, et utiliser les barres de défilement de l'espace de travail de flux pour vous déplacer dans cet espace ; ces barres sont situées sur le côté et en bas de la fenêtre SPSS Modeler.

Vous pouvez aussi commander l'affichage du balisage de l'écran, lequel se compose des commentaires de flux, des liens de modèles et des indications de branche de scoring. Pour activer ou désactiver cet affichage, cliquez sur :

Vue > **Balisage de flux**

Modification de la taille des icônes d'un flux

Vous pouvez changer la taille des icônes de flux par l'une des méthodes suivantes.

- A l'aide d'un paramètre de propriété de flux
- A l'aide d'un menu contextuel dans le flux
- A l'aide du clavier

Vous pouvez redimensionner la vue entière du flux à une taille comprise entre 8 % et 200 % de la taille d'icône standard.

Pour redimensionner le flux entier (méthode des propriétés du flux)

1. Dans le menu principal, sélectionnez :
   Outils > **Propriétés de flux** > **Options** > **Présentation**.
2. Sélectionnez la taille souhaitée dans le menu Taille d'icône.
3. Cliquez sur **Appliquer** pour afficher les résultats.
4. Cliquez sur **OK** pour enregistrer cette modification.

Pour redimensionner le flux entier (méthode du menu)

1. Cliquez avec le bouton droit de la souris sur l'arrière-plan du flux dans l'espace de travail.
2. Sélectionnez l'option **Taille d'icône** puis la taille souhaitée.

Pour redimensionner le flux entier (méthode du clavier)

1. Appuyez sur Ctrl + [-] sur le clavier pour effectuer un zoom arrière et réduire la vue d'une taille.
2. Appuyez sur Ctrl + Shift + [+] sur le clavier pour effectuer un zoom avant et agrandir la vue d'une taille.

Cette fonction est particulièrement utile pour obtenir une vue globale d'un flux complexe. Vous pouvez aussi l'utiliser pour réduire le nombre de pages nécessaires à l'impression d'un flux.

Utilisation de la souris dans IBM SPSS Modeler

Dans IBM SPSS Modeler, les utilisations les plus courantes de la souris sont les suivantes :

- **Clic simple.** Utilisez le bouton droit ou le bouton gauche de la souris pour sélectionner des options dans les menus, ouvrir des menus contextuels, ou accéder à diverses autres commandes et options standard. Cliquez avec la souris et maintenez le bouton de la souris enfoncé pour faire glisser des noeuds.
- **Double-clic.** Double-cliquez avec le bouton gauche de la souris pour placer des noeuds dans l'espace de travail de flux et éditer des noeuds existants.
- **Clic à l'aide du bouton central.** Cliquez avec le bouton central de la souris et faites glisser le curseur pour connecter des noeuds dans l'espace de travail de flux. Double-cliquez avec le bouton central de la souris pour déconnecter un noeud. Si vous ne possédez pas de souris à trois boutons, vous pouvez simuler cette fonction en appuyant sur la touche Alt tout en cliquant avec la souris et en la faisant glisser.

Utilisation des touches de raccourci

Dans IBM SPSS Modeler, de nombreuses opérations de programmation visuelle sont associées à des touches de raccourci. Par exemple, vous pouvez supprimer un noeud en cliquant dessus et en appuyant sur la touche Suppr de votre clavier. De la même façon, vous pouvez enregistrer rapidement un flux en appuyant sur la touche S tout en maintenant la touche Ctrl enfoncée. Les commandes de ce type sont indiquées par Ctrl et une autre touche (par exemple, Ctrl+S).

De nombreuses touches de raccourci sont utilisées dans les opérations Windows standard, telles que Ctrl+X pour couper un élément. Ces raccourcis sont pris en charge dans IBM SPSS Modeler, parallèlement à ceux présentés ci-après, propres à l'application.

Remarque : Dans certains cas, les anciennes touches de raccourci utilisées dans IBM SPSS Modeler sont en conflit avec les touches de raccourci Windows standard. Pour que ces anciennes touches de raccourci fonctionnent, il faut utiliser la touche Alt. Par exemple, Ctrl+Alt+C peut activer ou désactiver la mise en cache.

Tableau 1. Touches de raccourci prises en charge

Touche de raccourci	Fonction
Ctrl+A	Tout sélectionner
Ctrl+X	Couper
Ctrl+N	Permet de créer un flux
Ctrl+O	Permet d'ouvrir un flux
Ctrl+P	Imprimer
Ctrl+C	Copier
Ctrl+V	Coller

Tableau 1. Touches de raccourci prises en charge (suite)

Touche de raccourci	Fonction
Ctrl+Z	Annuler
Ctrl+Q	Permet de sélectionner tous les noeuds situés en aval du noeud sélectionné.
Ctrl+W	Permet de désélectionner tous les noeuds en aval (bascule du raccourci Ctrl+Q)
Ctrl+E	Exécuter à partir d'un noeud sélectionné
Ctrl+S	Permet d'enregistrer le flux en cours
Alt+flèches	Permettent de déplacer les noeuds sélectionnés dans l'espace de travail de flux dans le sens indiqué par la flèche utilisée
Maj+F10	Permet d'ouvrir le menu contextuel du noeud sélectionné

Tableau 2. Touches de raccourci prises en charge pour les anciennes touches d'accès rapide

Touche de raccourci	Fonction
Ctrl+Alt+D	Permet de dupliquer un noeud
Ctrl+Alt+L	Permet de charger un noeud
Ctrl+Alt+R	Permet de renommer un noeud
Ctrl+Alt+U	Permet de créer un noeud Utilisateur
Ctrl+Alt+C	Permet d'activer et de désactiver le cache
Ctrl+Alt+F	Permet de vider le cache
Ctrl+Alt+X	Développer le super noeud
Ctrl+Alt+Z	Permet d'effectuer un zoom avant/arrière
Suppr	Permet de supprimer un noeud ou une connexion

Impression

Les objets suivants peuvent être imprimés dans IBM SPSS Modeler :

- Diagrammes de flux
- Graphiques
- Tables
- Rapports (à partir du noeud Rapport et des rapports de projet)
- Scripts (à partir des boîtes de dialogue Propriétés du flux, Script autonome ou Script Super noeud)
- Modèles (navigateurs de modèle, onglets de boîte de dialogue avec élément en cours, Visualiseurs d'arbres)
- Annotations (à partir de l'onglet Annotations de la sortie)

Pour imprimer un objet :

- Pour imprimer sans afficher d'aperçu, cliquez sur le bouton Imprimer de la barre d'outils.
- Pour définir la mise en page avant d'imprimer, sélectionnez **Mise en page** dans le menu Fichier.
- Pour afficher un aperçu avant d'imprimer, sélectionnez **Aperçu avant impression** dans le menu Fichier.
- Pour afficher la boîte de dialogue d'impression standard vous permettant de sélectionner les imprimantes et de définir des options d'aspect, sélectionnez **Imprimer** dans le menu Fichier.

Automatisation d'IBM SPSS Modeler

Etant donné que l'exploration de données avancé peut être complexe et parfois long, IBM SPSS Modeler comprend plusieurs types d'assistance au codage et à l'automatisation.

- **Control Language for Expression Manipulation** (CLEM) est un langage permettant d'analyser et de manipuler les données circulant au sein des flux d'IBM SPSS Modeler. Les data miners utilisent beaucoup le langage CLEM dans les opérations de flux pour exécuter des tâches aussi simples que le calcul du profit à partir des données de coûts et de revenus, ou aussi complexes que la transformation de données du log Web en un ensemble de champs et d'enregistrements contenant des informations utilisables.
- **La génération de scripts** est un outil performant pour automatiser les processus dans l'interface utilisateur. Les scripts effectuent des opérations semblables à celles qui peuvent être exécutées à la souris ou au clavier. Vous pouvez également définir une sortie et manipuler des modèles générés.

Chapitre 4. Compréhension de l'exploration de données

Présentation de l'exploration de données

A l'aide de plusieurs techniques, l'**exploration de données** identifie des portions d'informations dans des bases de données. L'exploration de données permet d'extraire des informations de manière à pouvoir les utiliser dans des domaines tels que l'aide à la décision, les prévisions et les estimations. Les données sont souvent volumineuses, mais peu significatives et ne sont pas directement utilisables si elles restent à l'état brut. Seules les informations masquées des données ont de la valeur.

En matière d'exploration de données, vous atteindrez votre objectif en associant vos connaissances des données ou celles de votre expert à des techniques d'analyse élaborées et actives où l'ordinateur identifie les relations et options sous-jacentes des données. Le processus d'exploration de données génère des modèles à partir des données historiques utilisées ensuite pour les prévisions, la détection de caractéristiques, etc. La technique de création de ces modèles est appelée **apprentissage automatique** ou **modélisation**.

Méthodes de modélisation

IBM SPSS Modeler offre plusieurs technologies de modélisation et d'apprentissage automatique, qui peuvent être approximativement regroupées selon les types de problème qu'elles sont censées résoudre.

- Les méthodes de modélisation prédictive incluent les arbres décision, les réseaux de neurones et les modèles statistiques.
- Les modèles de classification se chargent essentiellement d'identifier des groupes d'enregistrements similaires et de répertorier les enregistrements en fonction du groupe auquel ils appartiennent. Les méthodes de classification non supervisée sont les suivantes : Kohonen, *k*-means et TwoStep.
- Les règles d'association associent une conclusion particulière (l'achat d'un produit particulier) à un ensemble de conditions (l'achat de plusieurs autres produits).
- Vous pouvez utiliser les modèles de filtrage pour filtrer les données afin de repérer les champs et enregistrements présentant le plus d'intérêt pour une modélisation et d'identifier les valeurs éloignées qui ne correspondent pas aux motifs connus. Vous disposez pour cela de deux méthodes : la sélection de fonction et la détection d'anomalies.

Découverte et manipulation des données

IBM SPSS Modeler inclut aussi des fonctions vous permettant d'appliquer vos connaissances aux données :

- **Manipulation des données.** Construit des éléments de données dérivés d'éléments existants et scinde les données en sous-ensembles plus compréhensibles. Vous pouvez fusionner et filtrer des données provenant de différentes sources.
- **Navigation et visualisation.** Affiche certains aspects des données à l'aide du noeud Audit données pour exécuter un audit initial, y compris les graphiques et les statistiques. La visualisation avancée comprend des graphiques interactifs, qui peuvent être exportés pour être inclus dans les rapports de projet.
- **Statistiques.** Confirme les liens soupçonnés entre des variables dans les données. Vous pouvez également utiliser les statistiques de IBM SPSS Statistics dans IBM SPSS Modeler.
- **Test des hypothèses.** Crée des modèles sur le comportement des données et vérifie ces modèles.

En règle générale, vous utiliserez ces fonctions pour identifier une série d'attributs de données pouvant se révéler très utile. Vous pouvez ensuite intégrer ces attributs aux techniques de modélisation en vue de l'identification des règles et des relations sous-jacentes.

Applications standard

Les applications standard utilisées pour les méthodes d'exploration de données sont les suivantes :

Publipostage. Déterminer les groupes démographiques qui ont la plus grande fréquence de réponses. Utilisez ces informations pour optimiser le nombre de réponses aux publipostages ultérieurs.

Cote de solvabilité. Utiliser l'historique des crédits d'un individu pour les prises de décisions concernant des acceptations ou des refus de crédit.

Ressources humaines. Comprendre les anciennes pratiques et créer des règles de décision visant à simplifier le processus d'embauche.

Recherche médicale. Créer des règles de décision qui suggèrent des procédures sur la base de preuves médicales.

Analyse de marché. Déterminer les variables (caractéristiques géographiques, tarifaires et liées aux clients) qui sont associées aux ventes.

Contrôle qualité. Analyser les données de la fabrication d'un produit et identifier les variables qui déterminent les défauts sur les produits.

Etudes stratégiques. Utiliser des données d'enquête pour établir une stratégie en appliquant des règles de décision qui permettent de sélectionner les variables les plus importantes.

Soins de santé. Associer des enquêtes et des données cliniques afin de déterminer les variables à prendre en compte dans le cadre de soins de santé.

Terminologie

Les termes **attribut**, **champ** et **variable** se rapportent à un élément de données unique commun à toutes les observations en cours. Un ensemble de valeurs d'attribut se rapportant à une observation spécifique est appelé **enregistrement**, **exemple** ou **observation**.

Evaluation des données

Pour qu'un travail d'exploration de données soit fructueux, les données à utiliser doivent répondre à certains critères. Les sections suivantes présentent certains aspects des données et de leur application à prendre en considération.

Assurez-vous que les données sont disponibles

Ce point peut sembler évident, mais il faut savoir que même si les données sont disponibles, elles peuvent ne pas être exploitables immédiatement. IBM SPSS Modeler peut importer des données à partir de bases de données (via ODBC) ou de fichiers. Les données peuvent se présenter sous une forme différente sur une machine à laquelle on ne peut pas accéder directement. Elles devront pour cela être téléchargées ou transférées dans un format convenable avant de pouvoir être utilisées. Au besoin, elles devront être dispersées parmi plusieurs bases de données et sources avant d'être de nouveau regroupées. Ce processus n'a pas forcément lieu en ligne. Si les données sont inscrites sur papier, une phase de saisie des données est nécessaire avant de procéder à l'exploration de données.

Vérifiez que les données couvrent tous les attributs nécessaires

L'objectif de l'exploration de données est d'identifier les attributs qui conviennent, par conséquent, ce point peut sembler étrange. Cependant, il s'avère très utile de consulter les données disponibles et d'essayer d'identifier les facteurs pertinents non enregistrés. Par exemple, si vous souhaitez prévoir la

vente de glaces, vous disposerez d'une quantité importante d'informations via les magasins de vente au détail et l'historique des ventes, ce qui n'est pas le cas des conditions météorologiques dont l'impact peut être significatif. Le fait que des attributs manquent ne signifie pas nécessairement que l'exploration de données ne produira pas de résultats fiables. Par contre, l'exactitude des prévisions peut être limitée.

L'exécution d'un audit complet de vos données permet d'évaluer rapidement la situation. Avant de continuer, associez un noeud Audit données à votre source de données, puis exécutez-le pour générer un rapport complet.

Faites attention aux données parasites

Les données contiennent souvent des erreurs ou peuvent contenir des points de vue subjectifs qui peuvent changer dans le temps. Ces phénomènes sont généralement désignés sous le nom de **bruits parasites**. Parfois, ces parasites sont normaux. Même s'il existait une règle générale, elle ne pourrait pas s'appliquer dans cent pour cent des cas.

En résumé, plus les données contiennent de parasites, plus il est difficile d'obtenir des résultats précis. Cependant, les méthodes d'apprentissage automatique d'IBM SPSS Modeler sont capables de gérer les données parasites, et ont été utilisées avec succès sur les jeux de données contenant environ 50 % de données parasites.

Assurez-vous d'avoir assez de données

Dans l'exploration de données, ce n'est pas nécessairement la taille d'un jeu de données qui importe. La représentativité du jeu de données est bien plus significative, ainsi que la couverture des résultats possibles et des combinaisons de variables.

En général, plus le nombre d'attributs considérés est important, plus le nombre d'enregistrements nécessaires pour obtenir une couverture représentative est élevé.

Si les données sont représentatives et s'il existe des règles sous-jacentes générales, un échantillon de données contenant plusieurs milliers d'enregistrements (ou plusieurs centaines), peut donner des résultats aussi performants qu'un échantillon contenant un million d'enregistrements. De plus, vous obtiendrez les résultats plus rapidement.

Adressez-vous aux spécialistes de ces données

Dans la plupart des cas, vous travaillez sur vos propres données et vous maîtrisez donc leur sens et contenu. Cependant, si vous travaillez sur des données pour un autre service de votre entreprise ou pour un client, il est fortement conseillé d'entrer en contact avec les experts qui maîtrisent ces données. Ils peuvent aider à identifier les attributs pertinents et à interpréter les résultats d'exploration de données, en faisant la différence entre des informations pertinentes et des informations erronées (informations sans aucune valeur) dues à des anomalies dans les jeux de données.

Stratégie pour l'exploration de données

Comme dans le cas de la plupart des efforts commerciaux, l'exploration de données est beaucoup plus profitable si vous procédez de manière méthodique. Même si vous utilisez des outils d'exploration de données de pointe comme IBM SPSS Modeler, vous avez besoin de la présence d'un analyste en informatique de gestion expérimenté pour assurer le suivi du processus. Pour vous guider dans votre planification, répondez aux questions suivantes :

- Quel réel problème souhaitez-vous résoudre ?
- De quelles sources de données disposez-vous et quelles sont les données nécessaires au problème en cours ?

- Quel genre de prétraitement et de nettoyage de données devez-vous effectuer avant d'utiliser l'exploration de données ?
- Quelles techniques d'exploration de données pensez-vous utiliser ?
- Comment pensez-vous évaluer les résultats de l'analyse d'exploration de données ?
- Comment tirer profit au maximum des informations obtenues suite à l'exploration de données ?

L'exploration de données peut devenir très rapidement compliqué. Vous devez surveiller beaucoup d'éléments à la fois, tels que les difficultés métier complexes, les sources de données multiples, les différents niveaux de qualité des données selon la source, la multiplicité des techniques d'exploration de données et des méthodes d'évaluation des résultats, etc.

Il est également important d'avoir à disposition un modèle de processus clairement défini pendant l'exploration de données. Le modèle de processus vous aide à répondre aux questions répertoriées précédemment dans cette section et permet de s'assurer que les points importants sont traités. Il vous indique la marche à suivre pour que vous restiez concentré sur le fond du problème tout en faisant face à la complexité des données à explorer.

Le modèle de processus d'exploration de données recommandé avec SPSS Modeler est le processus CRISP-DM (Cross-Industry Standard Process for Data Mining). En d'autres termes, ce modèle est conçu pour résoudre une grande diversité de problèmes propres aux entreprises ou au métier.

Méthodologie CRISP-DM

La méthodologie CRISP-DM comprend six étapes au cours desquelles vous aborderez les principaux problèmes d'exploration de données. Les six étapes s'ajustent au sein d'un processus cyclique conçu pour incorporer l'exploration de données dans vos pratiques métier générales.

Les six étapes comprennent :

- **Compréhension de l'entreprise.** Il s'agit peut-être de l'étape d'exploration de données la plus importante. La compréhension de l'entreprise vise à déterminer des objectifs commerciaux, à évaluer la situation, à déterminer les objectifs d'exploration de données et à établir une planification de projet.
- **Compréhension des données.** Les données sont la "matière première" de l'exploration de données. Cette étape vise à vous sensibiliser à l'importance de bien maîtriser les ressources de données et les caractéristiques de ces ressources. Elle aborde la collecte de données initiale, la description et l'exploration des données, et la vérification de la qualité de ces données. Le noeud Audit données disponible dans la palette de noeuds Sortie est un outil indispensable pour comprendre les données.
- **Préparation des données.** Après avoir catalogué vos ressources de données, vous devez préparer les données à explorer. La préparation comprend la sélection, le nettoyage, la construction, l'intégration et le formatage des données.
- **Modélisation.** Cette étape est la plus intéressante. Des méthodes d'analyse élaborées sont utilisées pour extraire des informations à partir des données. Lors de cette étape, il s'agit de sélectionner des techniques de modélisation, de générer des conceptions de test, et de construire et d'évaluer des modèles.
- **Evaluation.** Une fois que vous avez choisi vos modèles, vous pouvez évaluer l'aide que peuvent apporter les résultats d'exploration de données à la concrétisation de vos objectifs commerciaux. Cette étape aborde l'évaluation des résultats, la vérification du processus d'exploration de données et la prise de décision des étapes à suivre.
- **Déploiement.** Cette étape permet de rentabiliser pleinement tous vos efforts. L'objectif de cette phase est d'intégrer vos nouvelles connaissances aux processus métier quotidiens afin de résoudre votre problème métier initial. Elle comprend le déploiement du plan, la surveillance et la maintenance, la production d'un rapport final et la révision du projet.

Il y a des éléments clés dans ce modèle de processus. Tout d'abord, il faut savoir que, même si le processus respecte l'ordre des étapes présentées dans les paragraphes précédents, il existe des situations

où les étapes ont une influence les unes sur les autres dans un ordre aléatoire. Par exemple, la préparation des données précède généralement la modélisation. Cependant, les décisions prises et les informations collectées lors de la modélisation remettent en question certaines parties de la préparation des données, ce qui entraîne de nouveaux problèmes de modélisation. Les deux étapes sont alors interdépendantes jusqu'à ce qu'elles soient satisfaisantes. De même, l'étape d'évaluation peut vous amener à réévaluer votre compréhension initiale de l'entreprise et vous aider à prendre conscience du fait que vous êtes en train de répondre à la mauvaise question. A ce stade, vous pouvez revoir votre compréhension de l'entreprise, puis passer au reste du processus dans le but d'atteindre un objectif clairement défini.

Le second élément-clé est la nature itérative de l'exploration de données. Ce serait un tort de croire que l'exploration de données se limite à planifier un projet, à le réaliser et à rassembler vos données. L'utilisation de l'exploration de données pour satisfaire aux exigences de votre clientèle est un effort continu. Les connaissances acquises, suite à un cycle d'exploration de données, aboutissent presque systématiquement à de nouvelles questions, de nouveaux problèmes et de nouvelles opportunités, permettant d'identifier les besoins de votre clientèle et d'y répondre. Ces nouvelles questions, opportunités et nouveaux problèmes peuvent être abordés en renouvelant l'extraction de données. L'exploration de données et l'identification de nouvelles opportunités doivent désormais faire partie de votre métier et être un élément-clé de votre stratégie commerciale globale.

Cette introduction n'offre qu'une présentation concise du modèle de processus CRISP-DM. Pour plus d'informations sur le modèle, consultez les ressources suivantes :

- Le *Guide CRISP-DM*, accessible de même que d'autres documents à partir du dossier *\Documentation* du disque d'installation.
- Le système d'aide de CRISP-DM, disponible à partir du menu Démarrer ou en cliquant sur **Aide CRISP-DM** à partir du menu Aide d'IBM SPSS Modeler.

Types de modèle

IBM SPSS Modeler propose différentes méthodes de modélisation issues des domaines de l'apprentissage automatique, de l'intelligence artificielle et des statistiques. Les méthodes disponibles dans la palette Modélisation vous permettent d'extraire de nouvelles informations de vos données et de développer des modèles prédictifs. Chaque méthode possède ses propres avantages et est donc plus adaptée à certains types de problème spécifiques.

Le manuel *IBM SPSS Modeler - Guide des applications* fournit des exemples pour nombreuses de ces méthodes, ainsi qu'une présentation générale du processus de modélisation. Ce guide est disponible en tant que tutoriel en ligne, ainsi qu'au format PDF. Pour plus d'informations, voir «Exemples d'application», à la page 4.

Les méthodes de modélisation sont divisées en plusieurs catégories :

- Supervisé
- Association
- Segmentation

Modèles supervisés

Les *modèles supervisés* utilisent les valeurs d'un ou plusieurs champs d'**entrée** pour prévoir la valeur d'un ou plusieurs champs de sortie ou **cible**. En voici quelques exemples : les arbres de décision (algorithmes C&R Tree, QUEST, CHAID et C5.0), la régression (algorithmes linéaires, logistiques, linéaires généralisés et régression de Cox), les réseaux de neurones, les machines à vecteurs de prise en charge et les réseaux Bayésiens.

Les modèles supervisés aident les organisations à prévoir un résultat connu (par exemple, si un client doit ou non effectuer un achat, ou si une transaction correspond à un modèle connu ou relève de la

fraude). Les techniques de modélisation comprennent le processus d'apprentissage automatique, l'induction de règles, l'identification de sous-groupes, les méthodes statistiques et la génération de modèles multiples.

Noeuds supervisés

Le noeud Discriminant automatique crée et compare les résultats binaires de plusieurs modèles différents (oui ou non, avec ou sans attrition, etc.), ce qui vous permet de choisir la meilleure approche pour une analyse donnée. Plusieurs algorithmes de modélisation sont pris en charge. Vous pouvez alors sélectionner les méthodes que vous souhaitez utiliser, les options spécifiques pour chacune d'elles et le critère de comparaison des résultats. Le noeud génère un ensemble de modèles basé sur les options spécifiées et classe les meilleurs candidats en fonction des critères indiqués.

Le noeud Numérisation automatique évalue et compare des modèles pour des résultats d'intervalle numérique continus par le biais de différentes méthodes. Le noeud fonctionne de la même manière que le noeud Discriminant automatique, vous permettant ainsi de choisir les algorithmes à utiliser et à tester avec différentes combinaisons d'options en un seul passage de modélisation. Les algorithmes pris en charge comprennent les réseaux neuraux, l'arborescence C&R Tree, CHAID, la régression linéaire, la régression linéaire généralisée et support vector machines (SVM). Les modèles peuvent être comparés selon la corrélation, l'erreur relative ou le nombre de variables utilisées.

Le noeud Arbre Classification et Regression, (C&R) Tree, génère un arbre décision qui vous permet de prévoir ou de classifier les observations futures. La méthode utilise la technique de partition récursive afin de diviser les données d'apprentissage en segments en réduisant l'index d'impureté à chaque étape, un noeud de l'arbre étant considéré comme "pur" si 100 % de ses observations appartiennent à une catégorie spécifique du champ cible. Les champs cible et les champs d'entrée peuvent être des champs d'intervalle numériques ou des champs catégoriels numériques (nominal,ordinal ou indicateur). Toutes les divisions sont binaires (deux sous-groupes uniquement).

Le noeud QUEST est une méthode de classification supervisée binaire permettant de créer des arbres décision, développée pour réduire le temps de traitement nécessaire aux analyses C&R Tree importantes, tout en limitant la tendance, observée parmi les méthodes d'arbre de classification, à favoriser les entrées autorisant un nombre supérieur de divisions. Les champs d'entrée peuvent être des intervalles numériques (continues) mais les champs cible doivent être catégoriels. Toutes les divisions sont binaires.

Le noeud CHAID génère des arbres décision à l'aide des statistiques du khi-deux pour identifier les séparations optimales. Contrairement aux noeuds C&R Tree et QUEST, CHAID peut générer des arbres non binaires, ce qui implique que certaines divisions possèdent plusieurs branches. Les champs cibles et les champs d'entrée peuvent être d'intervalle numérique (continu) ou catégoriques. La méthode Exhaustive CHAID correspond à une modification du CHAID qui examine plus en détail toutes les divisions possibles, mais dont les calculs sont plus longs.

Le noeud C5.0 crée un arbre décision ou un ensemble de règles. Le fonctionnement de ce modèle repose sur un découpage de l'échantillon basé sur le champ qui fournit le gain d'informations le plus important à chaque niveau. Le champ cible doit être catégoriel. Les divisions multiples en plus de deux sous-groupes sont autorisées.

Le noeud Liste de décision identifie les sous-groupes, ou les segments, qui présentent une probabilité plus élevée ou plus faible d'un résultat binaire donné par rapport à la population globale. Vous pouvez, par exemple, rechercher les clients qui ont une faible probabilité d'attrition ou ceux qui ont une plus forte probabilité de répondre favorablement à une campagne. Vous pouvez incorporer vos connaissances métier dans le modèle en ajoutant vos propres segments personnalisés et en prévisualisant des modèles alternatifs côte à côte de façon à comparer les résultats. Les modèles Liste de décision se composent d'une liste de règles dans laquelle chaque règle présente une condition et un résultat. Les règles sont appliquées dans l'ordre et la première règle correspondante détermine le résultat.

Les modèles de régression linéaire prédisent une cible continue en fonction de relations linéaires entre la cible et un ou plusieurs prédicteurs.

Le noeud ACP/Analyse factorielle propose des techniques de factorisation puissantes qui vous permettent de réduire la complexité de vos données. L'analyse en composantes principales (ACP) recherche les combinaisons linéaires des champs d'entrée qui permettent de capturer au mieux la variance dans l'ensemble de champs, où les composantes sont orthogonales (perpendiculaires) les unes par rapport aux autres. L'analyse factorielle a pour but d'identifier les facteurs sous-jacents qui expliquent la tendance des corrélations dans un ensemble de champs observés. Quelle que soit l'approche choisie, le but consiste à trouver un nombre limité de champs dérivés récapitulant les informations contenues dans l'ensemble de champs d'origine.

Le noeud Sélection de fonction filtre les champs d'entrée en vue de leur suppression, en fonction d'un ensemble de critères donné (tel que le pourcentage de valeurs manquantes) ; il classe ensuite les entrées restantes selon leur importance par rapport à la cible indiquée. Si l'on prend, par exemple, un jeu de données comportant des centaines d'entrées potentielles, quelles sont celles susceptibles d'être les plus utiles dans la modélisation des résultats de patients ?

L'analyse discriminante crée des hypothèses plus strictes que la régression logistique mais peut constituer une alternative ou un complément précieux à une analyse de régression logistique lorsque ces hypothèses sont réunies.

La régression logistique est une technique statistique de classification des enregistrements sur la base des valeurs des champs d'entrée. Excepté le fait qu'elle utilise un champ cible catégoriel et non pas numérique, cette régression est similaire à la régression linéaire.

La procédure Modèles linéaires généralisés développe le modèle linéaire général de sorte que la variable dépendante soit linéairement reliée aux facteurs et covariables via une fonction de lien précise. En outre, le modèle permet à la variable dépendante de suivre une distribution non normale. Il couvre les fonctionnalités d'un grand nombre de modèles statistiques, notamment le modèle de régression linéaire, le modèle de régression logistique, le modèle log-linéaire pour les données d'effectif et le modèle de survie avec censure par intervalle.

Un modèle mixte linéaire généralisé (MMLG) élargit le modèle linéaire de sorte que la cible puisse avoir une distribution non normale, qu'elle soit liée linéairement aux facteurs et covariables via une fonction de lien spécifiée, et que les observations puissent être corrélées. Les modèles mixtes linéaires généralisés couvrent une large variété de modèles, depuis les modèles de régression linéaire simple aux modèles multi-niveaux complexes destinés aux données longitudinales non normales.

Le noeud de régression de Cox vous permet de créer un modèle de survie pour les données de durée jusqu'à l'événement en présence d'enregistrements censurés. Ce modèle produit une fonction de survie qui prédit la probabilité que l'événement en question se soit produit à un moment (*t*) pour des valeurs données des variables d'entrée.

Le noeud Support Vector Machine (SVM) vous permet de classer les données dans l'un de deux groupes sans surajustement. SVM fonctionne bien avec les grands jeux de données, comme ceux qui disposent d'un très grand nombre de champs d'entrée.

Le noeud Réseau Bayésien permet de créer un modèle de probabilité en combinant les preuves observées et enregistrées avec les connaissances réelles pour établir la probabilité des occurrences. Le noeud est axé sur le Tree Augmented Naïve Bayes (TAN) et sur les réseaux Couverture de Markov qui servent principalement à la classification.

Le noeud Modèle de réponse en auto-apprentissage (SLRM) vous permet de créer un modèle dans lequel une nouvelle observation unique, ou un petit nombre de nouvelles observations, peuvent être utilisés pour réestimer un modèle sans qu'un recyclage de toutes les données soit nécessaire.

Le noeud Séries temporelles estime les modèles de lissage exponentiel, d'ARIMA (Autoregressive Integrated Moving Average) univariable et d'ARIMA multivariable (ou fonction de transfert) pour les données de séries temporelles et génère des prévisions d`une performance future. Ce noeud Séries temporelles est similaire au noeud Séries temporelles précédent rendu obsolète dans SPSS Modeler version 18. Toutefois, ce noeud Série temporelle plus récent est conçu pour exploiter la puissance d'IBM SPSS Analytic Server afin de traiter les données volumineuses et d'afficher le modèle résultant dans l'afficheur de résultat ajouté dans SPSS Modeler version 17.

Le noeud *k*-Voisin le plus proche (KNN) associe une nouvelle observation à la catégorie ou à la valeur des objets *k* les plus proches dans l'espace du prédicteur, où *k* est un entier. Les observations semblables sont proches l'une de l'autre et les observations dissemblables sont éloignées l'une de l'autre.

Le noeud de prévision spatio-temporelle (STP) utilise des données contenant des données d'emplacement, des champs d'entrée pour les prévisions (prédicteurs), un champ Temps et un champ Cible. Chaque emplacement comporte plusieurs lignes dans les données, lesquelles représentent les valeurs de chaque prédicteur à chaque mesure. Une fois les données analysées, elles peuvent être utilisées pour anticiper des valeurs cibles à chaque emplacement dans les données de forme utilisées dans l'analyse.

Modèles d'association

Les *modèles d'association* recherchent des modèles dans vos données où une ou plusieurs entités (telles que des événements, des achats ou des attributs) sont associées à une ou plusieurs autres entités. Les modèles établissent des ensembles de règles qui définissent ces relations. Ici, les champs au sein des données peuvent se comporter à la fois comme des entrées et comme des cibles. Vous ne pouvez pas découvrir ces associations manuellement, mais des algorithmes de règles font bien plus et plus rapidement, de sorte que vous êtes en mesure d'explorer des modèles plus complexes. Les modèles Apriori et Carma sont des exemples de l'utilisation de tels algorithmes. Un autre type de modèle d'association est un modèle de détection de séquences qui recherche des motifs séquentiels dans des données à structure temporelle.

Les modèles d'association sont particulièrement utiles pour prévoir des résultats multiples, par exemple, les clients qui ont acheté le produit X ont également acheté les produits Y et Z. Les modèles d'association associent une conclusion particulière (telle que la décision d'acheter un produit) à un ensemble de conditions. L'avantage des algorithmes de règles d'association par rapport aux algorithmes d'arbre décision standard (C5.0 et Arbre C&RT) est le fait qu'il puisse exister des associations entre tous les attributs. Un algorithme d'arbre décision peut construire une règle uniquement avec une seule conclusion. En revanche, les algorithmes d'association tentent d'en trouver plusieurs, chaque règle pouvant avoir une conclusion différente.

Noeuds Association

Le noeud Apriori extrait des données un ensemble de règles et retient les règles contenant la plus grande quantité d'informations. Le noeud Apriori fournit cinq méthodes de sélection de règles et utilise un modèle d'indexation sophistiqué pour traiter efficacement les volumes de données importants. Pour les problèmes importants, l'apprentissage du noeud Apriori est généralement plus rapide ; il n'existe aucune limite quant au nombre de règles pouvant être conservées et il peut prendre en charge des règles faisant l'objet de 32 pré-conditions. Le noeud Apriori exige que les champs d'entrée et de sortie soient tous catégoriels, mais fournit de meilleures performances car il est optimisé de ce type de données.

Le modèle CARMA extrait un ensemble de règles des données sans que vous ayez à définir les champs d'entrée ou les champs cible. Au contraire du noeud Apriori le noeud CARMA offre des paramètres de création pour la prise en charge de la règle (à la fois pour les antécédents et les conséquences), plutôt qu'une simple prise en charge d'antécédents. Cela signifie que les règles générées peuvent être utilisées dans un grand nombre d'applications, par exemple pour rechercher une liste des produits ou des services (antécédents) dont la conséquence correspond à l'élément que vous souhaitez promouvoir à l'occasion de cette période de congés.

Le noeud Séquence recherche des règles d'association dans des données dotées d'une dimension temporelle. Une séquence est une liste de jeux d'éléments ayant tendance à survenir dans un ordre prévisible. Par exemple, un client qui achète un rasoir et une lotion après-rasage achètera vraisemblablement de la crème à raser. Le noeud Séquence est basé sur l'algorithme de règles d'association CARMA, qui utilise une méthode efficace de double lecture pour rechercher des séquences.

Le noeud Règles d'association est similaire au noeud Apriori ; toutefois, à la différence de celui-ci, il peut traiter les données de liste. Par ailleurs, le noeud Règles d'association peut être utilisé avec IBM SPSS Analytic Server pour traiter des données volumineuses et exploiter un traitement parallèle, lequel est plus rapide.

Modèles de segmentation

Les *modèles de segmentation* divisent les données en segments, ou clusters, d'enregistrement ayant des profils similaires de champs d'entrée. Comme ils ne s'occupent que des champs d'entrée, les modèles de

segmentation n'ont aucun concept de champs de sortie ou cible. Parmi les exemples de modèles de segmentation, on trouve les réseaux Kohonen, la classification k moyennes, la classification en deux étapes et la détection d'anomalies.

Les modèles de Segmentation (appelés aussi « modèles de classification ») sont utiles dans les cas où le résultat précis est inconnu (par exemple, lorsque vous identifiez de nouveaux types de fraude ou lorsque vous identifiez des groupes d'intérêt dans votre clientèle). Les modèles de classification se chargent essentiellement d'identifier des groupes d'enregistrements similaires et de répertorier les enregistrements en fonction du groupe auquel ils appartiennent. Cette opération peut s'effectuer sans la connaissance préalable des groupes et de leurs caractéristiques, et distingue les modèles de classification non supervisée des autres techniques de modélisation par le fait qu'aucun champ de sortie ni champ cible n'est prédéfini pour le modèle à prévoir. Il n'y a pas de réponse vraie ou fausse pour ces modèles. Leur valeur est déterminée par leur capacité à capturer des groupements intéressants dans les données et ils fournissent des descriptions utiles de ces mêmes groupements. Les modèles de classification non supervisée sont souvent utilisés pour créer des clusters ou des segments qui sont ensuite utilisés en tant qu'entrées dans les analyses suivantes (par exemple, par la segmentation de clients potentiels dans des sous-groupes homogènes).

Noeuds Segmentation

Le noeud Cluster automatique évalue et compare les modèles de classification identifiant des groupes d'enregistrements ayant des caractéristiques similaires. Le noeud fonctionne de la même manière que les autres noeuds de modélisation automatiques, vous permettant de tester plusieurs combinaisons d'options en une seule modélisation. Les modèles peuvent être comparés à l'aide de mesures de bases permettant d'essayer de filtrer et de classer l'utilité des modèles de classification et de fournir une mesure en fonction de l'importance de champs particuliers.

Le noeud k moyenne classe le jeu de données dans différents groupes (ou clusters). La méthode définit un nombre de clusters fixe, affecte à plusieurs reprises des enregistrements à des clusters et ajuste les centres de cluster, jusqu'à ce que le modèle ne puisse plus être amélioré. Au lieu de tenter de prédire un résultat, le modèle *k*-means utilise un processus connu sous le nom d'apprentissage non supervisé pour découvrir des tendances dans l'ensemble de champs d'entrée.

Le noeud Kohonen génère un type de réseau de neurones qui peut être utilisé pour classer les données en groupes distincts. Lorsque l'apprentissage du réseau est terminé, les enregistrements similaires doivent être regroupés dans la connexion de sortie, tandis que les enregistrements différents sont à l'opposé. Vous pouvez étudier le nombre d'observations capturées par chaque unité du nugget de modèle afin d'identifier les unités fortes. Vous pouvez ainsi vous faire une idée du nombre de clusters approprié.

Le noeud TwoStep utilise une méthode de classification non supervisée en deux étapes. La première étape consiste en une exploration des données visant à compresser les données d'entrée brutes en sous-clusters plus faciles à manipuler. Au cours de la seconde étape, l'utilisation d'une méthode de classification hiérarchique permet de fusionner progressivement les sous-clusters en clusters de plus en plus importants. La technique TwoStep a l'avantage d'évaluer automatiquement le nombre de clusters optimal pour les données d'apprentissage. Il peut prendre en charge de manière efficace des types de champ mixtes et des jeux de données volumineux.

Le noeud Détection des anomalies identifie les observations inhabituelles, ou valeurs éloignées, qui ne se conforment pas aux motifs de données "normales". Il vous permet d'identifier les valeurs éloignées même si celles-ci ne correspondent pas aux motifs connus précédemment et même si vous ne savez pas exactement ce que vous recherchez.

Exploration des bases de données des modèles

IBM SPSS Modeler prend en charge l'intégration aux outils d'exploration de données et de modélisation disponibles auprès de fournisseurs de bases de données, dont Oracle Data Miner et Microsoft Analysis Services. Il est possible de construire, d'évaluer et de stocker des modèles dans la base de données, le tout depuis l'application IBM SPSS Modeler. Pour plus d'informations, voir le document *IBM SPSS Modeler - Guide d'exploration de base de données*.

Modèles IBM SPSS Statistics

Si vous avez installé une copie sous licence de IBM SPSS Statistics sur votre ordinateur, vous pouvez accéder et exécuter certains programmes IBM SPSS Statistics depuis IBM SPSS Modeler pour créer et évaluer des modèles.

Exemples d'exploration de données

Pour en savoir plus sur l'exploration de données, le meilleur moyen est de commencer par un exemple. Un certain nombre d'exemples d'application sont disponibles dans *IBM SPSS Modeler Guide des applications*, qui fournit des présentations brèves et ciblées de méthodes et de techniques de modélisation spécifiques. Pour plus d'informations, voir la rubrique «Exemples d'application», à la page 4.

Chapitre 5. Création de flux

Présentation de la création de flux

L'exploration de données avec IBM SPSS Modeler est basée sur l'exécution de données via une série de noeuds, appelée **flux.** Cette série de noeuds représente les opérations à réaliser sur les données et les liens entre les noeuds indiquent la direction du flux de données. Généralement, un flux de données est utilisé pour lire des données dans IBM SPSS Modeler, effectuer leur exécution par des manipulations diverses et envoyer ces dernières vers une destination, telle qu'un tableau ou un Visualiseur.

Par exemple, supposons que vous souhaitiez ouvrir une source de données, ajouter un nouveau champ, sélectionner des enregistrements en fonction des valeurs contenues dans ce nouveau champ, puis afficher les résultats dans un tableau. Dans ce cas, votre flux de données serait composé de quatre noeuds :

Un noeud Délimité défini de façon à pouvoir lire les données à partir de la source de données.

Un noeud Calculer qui permet d'ajouter le nouveau champ calculé au jeu de données.

Un noeud Sélectionner que vous utilisez pour configurer des critères de sélection destinés à exclure certains enregistrements du flux de données.

Un noeud Table qui permet d'afficher les résultats de vos manipulations à l'écran.

Création de flux de données

L'interface unique de SPSS Modeler vous permet d'explorer les données visuellement en travaillant avec des graphiques de flux de données. Au niveau le plus élémentaire, vous pouvez créer un flux de données en utilisant la procédure suivante :

- Ajoutez des noeuds dans l'espace de travail de flux.
- Connectez les noeuds de façon à former un flux.
- Définissez toutes les options de noeud ou de flux.
- Exécutez le flux.

Cette section contient des informations plus détaillées sur l'utilisation des noeuds pour créer des flux de données plus complexes. Elle aborde également les options et paramètres des noeuds et des flux. Pour obtenir des exemples pas à pas de création de flux en utilisant les données fournies avec SPSS Modeler (dossier `Demos` du répertoire d'installation du programme), reportez-vous à «Exemples d'application», à la page 4.

© Copyright IBM Corp. 1994, 2019

Utilisation des noeuds

Dans IBM SPSS Modeler, les noeuds permettent d'explorer des données. Les différents noeuds dans l'espace de travail représentent différents objets et actions. La palette située au bas de la fenêtre IBM SPSS Modeler contient tous les noeuds pouvant être utilisés pour la création des flux.

Il existe plusieurs types de noeuds. Les **Noeuds source** introduisent les données dans le flux et se trouvent dans l'onglet Sources de la palette des noeuds. Les **Noeuds d'exécution** exécutent des opérations sur des enregistrements et des champs de données individuels et se trouvent dans les onglets de la palette Ops sur lignes et Ops sur champs. Les **noeuds de sortie** produisent plusieurs résultats pour les données, les graphiques et les résultats des modèles et sont inclus dans les onglets Graphiques, Sortie et Exporter de la palette de noeuds. Les **noeuds de modélisation** utilisent des algorithmes statistiques pour créer des nuggets de modèle et se trouvent dans l'onglet Modélisation, et (s'il est activé) dans l'onglet Modélisation de base de données; de la palette des noeuds. Pour plus d'informations, voir la rubrique «Palette de noeuds», à la page 15.

Ces noeuds doivent être connectés pour former des flux qui, une fois exécutés, vous permettent de visualiser des relations et de tirer des conclusions. Les flux sont similaires à des scripts (vous pouvez les enregistrer et les réutiliser avec d'autres fichiers de données).

Un noeud exécutable qui traite des données du flux est appelé **noeud terminal**. Un noeud de modélisation ou de sortie est un noeud terminal s'il se trouve à la fin d'un flux ou d'une branche de flux. Vous ne pouvez pas connectez davantage de noeuds à un noeud terminal.

Remarque : Vous pouvez personnaliser la palette de noeuds. Pour plus d'informations, voir la rubrique «Personnalisation de la palette Noeuds», à la page 232.

Ajout de noeuds à un flux

Pour ajouter des noeuds à un flux à partir de la palette de noeuds, vous disposez de plusieurs méthodes :

- Double-cliquez sur un noeud de la palette. *Remarque* : un double-clic sur un noeud permet de le relier automatiquement au flux en cours. Pour plus d'informations, voir la rubrique «Connexion de noeuds dans un flux».
- Faites glisser un noeud à partir de la palette vers l'espace de travail de flux.
- Cliquez sur un noeud de la palette, puis cliquez sur l'espace de travail de flux.
- Choisissez une option appropriée à partir du menu Insérer d'IBM SPSS Modeler.

Une fois que vous avez ajouté un noeud à l'espace de travail de flux, double-cliquez sur le noeud pour afficher sa boîte de dialogue. Les options disponibles dépendent du type de noeud que vous ajoutez. Pour obtenir des informations sur les commandes de cette boîte de dialogue, cliquez sur le bouton **Aide**.

Suppression de noeuds

Pour supprimer un noeud du flux de données, cliquez dessus, puis appuyez sur la touche Suppr, ou cliquez dessus avec le bouton droit de la souris et sélectionnez **Supprimer** dans le menu.

Connexion de noeuds dans un flux

Pour que les noeuds ajoutés à l'espace de travail de flux forment un flux de données, ils doivent être reliés. Les connexions entre les noeuds indiquent la direction des données transmises d'une opération à la suivante. Il existe plusieurs façons de relier des noeuds pour former un flux : en double-cliquant, en utilisant le bouton central de la souris ou manuellement.

Pour ajouter et relier des noeuds en double-cliquant

La façon la plus simple de former un flux consiste à double-cliquer sur les noeuds de la palette. Cette méthode permet de relier automatiquement le nouveau noeud au noeud sélectionné dans l'espace de

travail de flux. Par exemple, si l'espace de travail contient un noeud SGBD, vous pouvez le sélectionner et double-cliquer sur le noeud à ajouter depuis la palette (par exemple, un noeud Calculer). Cette action relie automatiquement le noeud Calculer au noeud SGBD existant. Vous pouvez répéter ce processus jusqu'à ce que vous ayez connecté un noeud terminal, comme un noeud Histogramme ou Table. A ce stade, tout nouveau noeud sera connecté au dernier noeud non terminal en amont.

Pour relier des noeuds en utilisant le bouton central de la souris

Dans l'espace de travail de flux, cliquez sur un noeud et faites-le glisser vers un autre noeud en utilisant le bouton central de la souris (si votre souris ne dispose pas d'un bouton central, vous pouvez simuler l'action de ce bouton en appuyant sur la touche Alt tout en faisant glisser la souris d'un noeud vers un autre).

Pour relier des noeuds manuellement

Si votre souris ne dispose pas d'un bouton central et que vous préférez relier les noeuds manuellement, vous pouvez utiliser le menu contextuel d'un noeud pour le relier à un autre noeud qui figure déjà dans l'espace de travail.

1. Faites un clic droit sur le noeud à partir duquel vous souhaitez démarrer la connexion. Cette action ouvre le menu du noeud.
2. Dans le menu, cliquez sur **Connecter**.
3. L'icône de connexion apparaît sur le noeud de départ et sur le curseur. Cliquez sur le second noeud dans l'espace de travail pour relier les deux noeuds.

Certaines règles doivent être respectées pour la connexion des noeuds. Vous recevrez un message d'erreur si vous tentez d'effectuer les types de connexion suivants :

- Connexion vers un noeud source
- Connexion à partir d'un noeud terminal
- Noeud ayant plus de connexions d'entrée que le nombre maximal autorisé
- Connexion de deux noeuds déjà connectés
- Circularité (renvoi des données au noeud dont elles sont issues)

Contournement de noeuds dans un flux

Lorsque vous contournez un noeud du flux de données, toutes ses connexions d'entrée et de sortie sont remplacées par des connexions directes entre ses noeuds d'entrée et ses noeuds de sortie. Si le noeud ne dispose pas à la fois de connexions d'entrée et de sortie, toutes ses connexions sont alors supprimées sans être redirigées.

Par exemple, un flux qui calcule un nouveau champ, filtre des champs, puis explore les résultats dans un histogramme et un tableau. Si vous souhaitez également voir le même graphique et le même tableau de données *avant* le filtrage des champs, vous pouvez ajouter de nouveaux noeuds Histogramme et Table au flux ou contourner le noeud Filtrer. Lorsque vous contournez le noeud Filtrer, les connexions au graphique et au tableau passent directement par le noeud Calculer. Le noeud Filtrer est déconnecté du flux.

Pour contourner un noeud

1. Dans l'espace de travail de flux, double-cliquez sur le noeud à contourner à l'aide du bouton central de la souris. Vous pouvez également utiliser Alt+double-clic.

Remarque : vous pouvez annuler cette action en cliquant sur **Annuler** dans le menu Edition ou en appuyant sur Ctrl+Z.

Désactivation de noeuds dans un flux

Les noeuds d'exécution ayant une entrée unique au sein de flux peuvent être désactivés, ce qui permet d'ignorer le noeud au cours de l'exécution du flux. Cela vous évite d'avoir à supprimer ou à contourner le noeud et signifie que vous pouvez le laisser connecté aux noeuds restants. Vous pouvez toujours ouvrir et modifier les paramètres du noeud ; cependant les modifications ne prennent pas effet tant que vous n'avez pas de nouveau activé le noeud.

Par exemple, vous pouvez avoir un flux qui filtre plusieurs champs puis qui crée des modèles avec le jeu de données réduit. Si vous souhaitez créer aussi les mêmes modèles *sans* que les champs ne soient filtrés, afin de vérifier s'ils améliorent les résultats du modèle, vous pouvez désactiver le noeud Filtrer. Lorsque vous désactivez le noeud Filtrer, les connexions aux noeuds de modélisation sont directement transmises du noeud Calculer au noeud Typer

Pour désactiver un noeud

1. Dans l'espace de travail de flux, faites un clic droit sur le noeud à désactiver.
2. Cliquez sur **Désactiver le noeud** dans le menu contextuel.

Vous pouvez aussi cliquer sur **Noeud** > **Désactiver le noeud** à partir du menu Edition. Si vous souhaitez de nouveau inclure le noeud dans le flux, cliquez sur **Activer le noeud** de la même manière.

Remarque : vous pouvez annuler cette action en cliquant sur **Annuler** dans le menu Edition ou en appuyant sur Ctrl+Z.

Ajout de noeuds à des connexions existantes

Vous pouvez ajouter un nouveau noeud entre deux noeuds connectés en faisant glisser la flèche qui les relie.

1. Avec le bouton central de la souris, cliquez sur la flèche de connexion dans laquelle vous souhaitez insérer le noeud et faites-la glisser. Vous pouvez également simuler l'action du bouton du milieu de la souris en maintenant la touche Alt enfoncée pendant que vous cliquez et effectuez le déplacement.
2. Faites glisser la connexion jusqu'au noeud à inclure et relâchez le bouton de la souris.

Remarque : vous pouvez supprimer les nouvelles connexions du noeud et restaurer les connexions initiales en **contournant** le noeud.

Suppression de connexions entre des noeuds

Pour supprimer la connexion entre deux noeuds :

1. Faites un clic droit sur la flèche de connexion.
2. Dans le menu, cliquez sur **Supprimer la connexion**.

Pour supprimer toutes les connexions en direction et en provenance d'un noeud, procédez comme suit :

- Sélectionnez le noeud et appuyez sur F3.
- Sélectionnez le noeud puis, dans le menu principal, cliquez sur :

Edition > **Noeud** > **Déconnecter**

Définition des options des noeuds

Une fois que vous avez créé et connecté les noeuds, vous pouvez les personnaliser de plusieurs façons. Cliquez sur le noeud avec le bouton droit de la souris, puis sélectionnez l'une des options de menu.

- Cliquez sur **Editer** pour ouvrir la boîte de dialogue correspondant au noeud sélectionné.
- Cliquez sur **Connecter** pour connecter manuellement deux noeuds.
- Cliquez sur l'option **Déconnecter** pour supprimer tous les liens provenant et partant du noeud.
- Cliquez sur l'option **Renommer et annoter** pour faire apparaître l'onglet Annotations de la boîte de dialogue de modification.

- Cliquez sur **Nouveau commentaire** pour ajouter un commentaire relatif au noeud. Pour plus d'informations, voir la rubrique «Ajout de commentaires et d'annotations à des noeuds et à des flux», à la page 59.
- Cliquez sur **Désactiver le noeud** pour « masquer » le noeud au cours du traitement. Pour rendre le noeud visible de nouveau pour le traitement, cliquez sur **Activer le noeud**. Pour plus d'informations, voir la rubrique «Désactivation de noeuds dans un flux», à la page 40.
- Cliquez sur **Couper** ou **Supprimer** pour supprimer les noeuds sélectionnés de l'espace de travail de flux. *Remarque* : cliquer sur **Couper** vous permet de coller des noeuds, mais pas l'option **Supprimer**.
- Cliquez sur l'option **Copier le noeud** pour effectuer une copie du noeud sans ses connexions. Il peut être ajouté à un nouveau flux ou à un flux existant.
- Cliquez sur l'option **Charger le noeud** pour ouvrir un noeud enregistré et charger ses options dans le noeud sélectionné. Les noeuds doivent être du même type.
- Cliquez sur **Extraire le noeud** pour extraire le noeud à partir d'un référentiel IBM SPSS Collaboration and Deployment Services Repository connecté.
- Cliquez sur **Enregistrer le noeud** pour enregistrer les détails du noeud dans un fichier. Vous ne pouvez charger les détails du noeud que dans un autre noeud du même type.
- Cliquez sur **Stocker le noeud** pour stocker le noeud sélectionné dans un référentiel IBM SPSS Collaboration and Deployment Services Repository connecté.
- Cliquez sur **Cache** pour faire apparaître les options de mise en cache du noeud sélectionné.
- Cliquez sur **Mappage des données** pour faire apparaître les options de mappage des données sur une nouvelle source ou les options de spécifications des champs obligatoires.
- Cliquez sur **Créer un super noeud** pour faire apparaître les options de création d'un super noeud dans le flux en cours.
- Cliquez sur **Générer le noeud Utilisateur** pour remplacer le noeud sélectionné. Les exemples générés par ce noeud auront les mêmes champs que le noeud en cours.
- Cliquez sur **Exécuter à partir d'ici** pour exécuter tous les noeuds terminaux situés en aval du noeud sélectionné.

Options de mise en cache des noeuds

Pour optimiser l'exécution du flux, vous pouvez définir un *cache* sur n'importe quel nœud non terminal. Lorsque vous définissez un cache sur un noeud, le cache est alimenté par les données passant par le noeud lors de l'exécution suivante du flux de données. Les données sont alors lues à partir du cache (stocké sur le disque dans un répertoire temporaire) et non à partir de la source de données.

La mise en cache est plus utile après une opération de longue durée telle que le tri, la fusion ou l'agrégation. Par exemple, supposez que vous ayez un noeud source configuré pour la lecture des données relatives aux ventes à partir de la base de données et un noeud Agréger qui fournit des résultats de ventes par zone géographique. Vous pouvez définir un cache sur le noeud Agréger plutôt que sur le noeud source pour que le cache stocke uniquement les données agrégées et non le jeu de données intégral.

Remarque : La mise en cache au niveau des noeuds source, qui stocke simplement une copie des données originales telle qu'elle est lue dans IBM SPSS Modeler, n'améliorera pas les performances dans la plupart des cas.

Lorsque la mise en cache d'un noeud est activée, une petite icône en forme de document apparaît dans l'angle supérieur droit. Une fois les données mises en cache au niveau du noeud, l'icône de document devient verte.

Pour activer un cache

1. Dans l'espace de travail de flux, cliquez avec le bouton droit de la souris sur le noeud, puis cliquez sur **Cache** dans le menu.
2. Dans le sous-menu de mise en cache, cliquez sur **Activer**.

3. Vous pouvez désactiver le cache en cliquant sur le noeud avec le bouton droit de la souris et en cliquant sur **Désactiver** dans le sous-menu de mise en cache.

Mise en mémoire cache des noeuds dans une base de données

Dans le cas des flux exécutés dans une base de données, les données peuvent être mises en mémoire cache en milieu de flux dans une table temporaire de la base de données, plutôt que dans le système de fichiers. Vous pouvez combiner cela à la fonction d'optimisation SQL et accroître ainsi les performances de manière significative. Par exemple, la sortie d'un flux qui fusionne plusieurs tables afin de créer une vue d'exploration de données peut être mise en mémoire cache et réutilisée selon les besoins. La génération automatique du code SQL pour tous les noeuds en aval peut améliorer encore les performances.

Pour tirer profit de la mise en mémoire cache de base de données, vous devez activer les fonctions d'optimisation SQL et de mise en mémoire cache de base de données. Les paramètres d'optimisation du serveur annulent ceux définis pour le client. Pour plus d'informations, voir la rubrique «Définition des options d'optimisation pour les flux», à la page 47.

Une fois la mise en mémoire cache de base de données activée, il vous suffit de cliquer à l'aide du bouton droit de la souris sur un noeud non terminal afin de mettre les données en mémoire cache à ce stade du flux ; le cache est automatiquement créé dans la base de données lors de l'exécution suivante du flux. Si la mise en mémoire cache de base de données ou l'optimisation SQL n'est pas activée, le cache est écrit dans le système de fichiers.

Remarque : Les bases de données suivantes prennent en charge les tables temporaires à des fins de mise en mémoire cache : Db2, Oracle, SQL Server et Teradata. D'autres bases de données, comme Netezza, utilisent une table standard pour la mise en mémoire cache de base de données. Si nécessaire, le code SQL peut être personnalisé pour des bases de données spécifiques : contactez l'assistance technique.

Pour vider un cache

Lorsque le cache d'un noeud est vide, une icône en forme de document blanc apparaît sur le noeud. Lorsque le cache est saturé, l'icône en forme de document devient entièrement verte. Pour remplacer le contenu du cache, vous devez au préalable vider le cache, puis réexécuter le flux de données de façon à le remplir de nouveau.

1. Dans l'espace de travail de flux, cliquez avec le bouton droit de la souris sur le noeud, puis cliquez sur **Cache** dans le menu.
2. Dans le sous-menu de mise en cache, cliquez sur **Vider**.

Pour enregistrer un cache

Vous pouvez enregistrer le contenu d'un cache dans un fichier de données IBM SPSS Statistics (`*.sav`). Vous pouvez ensuite recharger le fichier en tant que cache ou définir un noeud utilisant le fichier de cache comme source de données. Vous pouvez également charger un cache que vous avez enregistré pour un autre projet.

1. Dans l'espace de travail de flux, cliquez avec le bouton droit de la souris sur le nœud, puis cliquez sur **Cache** dans le menu.
2. Dans le sous-menu de mise en cache, cliquez sur **Enregistrer le cache**.
3. Dans la boîte de dialogue Enregistrer le cache, choisissez l'emplacement d'enregistrement du fichier de cache.
4. Entrez un nom dans la zone de texte Fichier.
5. Assurez-vous que l'extension ***.sav** est sélectionnée dans la liste Type de fichiers, puis cliquez sur **Enregistrer**.

Pour charger un cache

Si vous avez enregistré un fichier de cache avant de le supprimer du noeud, vous pouvez le recharger.

1. Dans l'espace de travail de flux, cliquez avec le bouton droit de la souris sur le nœud, puis cliquez sur **Cache** dans le menu.
2. Dans le sous-menu de mise en cache, cliquez sur **Charger le cache**.
3. Dans la boîte de dialogue Charger le cache, sélectionnez l'emplacement du fichier de cache et cliquez sur le bouton **Charger**.

Prévisualisation des données dans les noeuds

Pour vous assurer que les données seront modifiées comme vous le désirez lors de la création du flux, vous pouvez envoyer vos données dans un noeud Table à chaque étape importante. Pour vous éviter cela, vous pouvez générer un aperçu de chaque noeud qui affiche un échantillon des données qui seront créées, ce qui permet de gagner du temps passé à la création de chaque noeud.

Pour des noeuds en amont d'un nugget de modèle, l'aperçu affiche les champs d'entrée ; pour un nugget de modèles ou des noeuds en aval du nugget (sauf pour les noeuds terminaux), l'aperçu affiche les champs d'entrée et les champs générés.

Le nombre de lignes affichées par défaut est de 10, mais vous pouvez le modifier dans les propriétés du flux. Pour plus d'informations, voir la rubrique «Définition d'options générales pour les flux», à la page 44.

Dans le menu **Générer**, vous pouvez créer plusieurs types de noeuds.

Remarque : Lorsque vous prévisualisez les données générées par ce noeud, toutes les modifications de propriété sont appliquées à ce noeud et ne peuvent pas être annulées (même comportement que si vous cliquez sur **Appliquer**).

Verrouillage des noeuds

Pour empêcher les autres utilisateurs de modifier les paramètres d'un ou de plusieurs noeuds dans un flux, vous pouvez encapsuler le ou les noeuds dans un type de noeud spécial appelé super noeud, puis verrouiller le super noeud en appliquant une protection par mot de passe.

Utilisation des flux

Une fois les noeuds source, exécution et terminal connectés dans l'espace de travail de flux, le flux est créé. En tant qu'ensembles de noeuds, les flux peuvent être enregistrés, annotés et ajoutés à des projets. Vous pouvez également définir de nombreuses options de flux, telles que l'optimisation, les options de date et d'heure, les paramètres et les scripts. Ces propriétés seront abordées dans les rubriques suivantes.

Dans IBM SPSS Modeler, vous pouvez utiliser et modifier plusieurs flux de données dans la même session IBM SPSS Modeler. La partie droite de la fenêtre principale contient le panneau des Gestionnaires qui vous aide à parcourir les flux, les résultats et les modèles actuellement ouverts. Si vous ne pouvez pas consulter le panneau des Gestionnaires, cliquez sur **Gestionnaires** dans le menu Vue, puis cliquez sur l'onglet **Flux**.

Dans cet onglet, vous pouvez effectuer les opérations suivantes :

- Accéder aux flux.
- Enregistrer les flux.
- Enregistrer les flux dans le projet en cours.
- Fermer les flux.
- Ouvrir de nouveaux flux.

- Stocker et extraire les flux d'un référentiel IBM SPSS Collaboration and Deployment Services (si vous en disposez). Pour plus d'informations, voir la rubrique «A propos du IBM SPSS Collaboration and Deployment Services Repository», à la page 191.

Dans l'onglet Flux, cliquez avec le bouton droit de la souris sur un flux pour accéder à ces options.

Définition des options pour des flux

Vous pouvez spécifier plusieurs options à appliquer au flux en cours. Vous pouvez aussi enregistrer ces options comme options par défaut à appliquer à tous vos flux. Les options sont les suivantes.

- **Général.** Diverses options comme des symboles et du codage de texte à utiliser dans le flux. Pour plus d'informations, voir «Définition d'options générales pour les flux».
- **Date/Heure.** Options relatives au format des expressions de date et d'heure. Pour plus d'informations, voir «Définition des options de date et heure pour les flux», à la page 46.
- **Formats de nombres.** Options contrôlant le format des expressions numériques. Pour plus d'informations, voir «Définition des options de format des nombres pour les flux», à la page 47.
- **Optimisation.** Options d'optimisation de la performance de flux. Pour plus d'informations, voir «Définition des options d'optimisation pour les flux», à la page 47.
- **Consignation et statut.** Options contrôlant la consignation SQL et le statut d'enregistrement. Pour plus d'informations, voir «Définition des options de statut de consignation et d'enregistrement SQL pour les flux», à la page 49.
- **Présentation.** Options relatives à la présentation du flux sur l'espace de travail. Pour plus d'informations, voir «Définition d'options de présentation pour les flux», à la page 49.
- **Analytic Server.** Options concernant l'utilisation de Analytic Server avec SPSS Modeler. Pour plus d'informations, voir «Propriétés du flux Analytic Server», à la page 50.
- **Données géospatiales.** Options concernant la mise en forme de données géospatiales pour leur utilisation dans le flux. Pour plus d'informations, voir «Définition d'options géospatiales pour les flux», à la page 51.

Définition des options de flux

1. Dans le menu Fichier, cliquez sur **Propriétés du flux** (ou sélectionnez le flux dans l'onglet Flux du panneau des Gestionnaires, faites un clic droit puis cliquez sur **Propriétés du flux** dans le menu contextuel).
2. Cliquez sur l'onglet **Options**.

Sinon, dans le menu Outils, cliquez sur :

Propriétés de flux > **Options**

Définition d'options générales pour les flux : Les options générales sont un ensemble d'options diverses s'appliquant à divers aspects du flux en cours.

La section **De base** inclut les options de base suivantes :

- **Symbole décimal.** Sélectionnez une virgule (,) ou un point (.) en tant que séparateur décimal.
- **Symbole de regroupement.** Pour les formats d'affichage des nombres, sélectionnez le symbole permettant de regrouper des valeurs (par exemple, l'espace dans 3 000,00). Vous avez le choix entre les options suivantes : aucun, point, virgule, espace et paramètres régionaux définis (auquel cas la valeur par défaut des paramètres régionaux actuels est utilisée).
- **Codage.** Spécifiez la méthode par défaut du flux pour le codage du texte. (*Remarque* : s'applique à variable. Noeud source de fichier et noeud Export Fichier plat uniquement. Aucun autre noeud n'utilise ce paramètre ; la majorité des fichiers de données contiennent des informations de codage intégrées.) Vous pouvez choisir la valeur système par défaut ou UTF-8. Si le système est exécuté en mode réparti,

sa valeur par défaut est spécifiée dans le Panneau de configuration de Windows de l'ordinateur serveur. Pour plus d'informations, voir la rubrique «Prise en charge d'Unicode dans IBM SPSS Modeler», à la page 255.

- **Mode d'évaluation des ensembles de règles.** Détermine le mode d'évaluation des modèles des ensembles de règles. Par défaut, les ensembles de règles utilisent la valeur **Vote** pour combiner les prévisions à partir de règles individuelles et déterminer la prévision définitive. Pour que les ensembles de règles utilisent par défaut la première règle, sélectionnez **Premier résultat**. Notez que cette option ne s'applique pas aux modèles Liste de décision, qui utilisent toujours la première correspondance telle que définie par l'algorithme.

Nombre maximum de lignes à afficher dans l'aperçu des données. Spécifiez le nombre de lignes à afficher lorsqu'un aperçu des données est demandé pour un noeud. Pour plus d'informations, voir la rubrique «Prévisualisation des données dans les noeuds», à la page 43.

Nombre maximum de membres pour les champs nominaux. Sélectionnez cette option pour spécifier le nombre maximal de membres pour les champs nominaux (ensemble) au-delà duquel le type du champ devient **Sans type**. Cette option peut être utile lorsque vous utilisez des champs nominaux volumineux. *Remarque* : lorsque le niveau de mesure d'un champ est configuré sur **Sans type**, son rôle est automatiquement configuré sur **Aucun**. Cela signifie que les champs ne sont pas disponibles pour la modélisation.

Taille limitée pour la modélisation Kohonen et K-MeansSélectionnez cette option pour spécifier le nombre maximal de membres des champs nominaux utilisés dans les réseaux Kohonen et les modélisations *K*-Means. La taille d'ensemble par défaut est 20. Si un nombre est supérieur à cette valeur, le champ est ignoré et un avertissement apparaît fournissant des informations sur le champ en question.

A noter que pour des raisons de compatibilité, cette option s'applique également à l'ancien noeud de réseaux de neurones, remplacés dans la version 14 d'IBM SPSS Modeler ; certains flux peuvent encore contenir ce noeud.

Rafraîchir les noeuds source lors de l'exécution. Sélectionnez cette option pour rafraîchir automatiquement tous les noeuds source lors de l'exécution du flux en cours. Cette action s'apparente à celle obtenue en cliquant sur le bouton **Rafraîchir** d'un noeud source, à la différence que cette option rafraîchit automatiquement tous les noeuds source du flux en cours (à l'exception des noeuds Utilisateur).

Remarque : cette option permet de vider les caches des noeuds en aval même si les données n'ont pas été modifiées. Si vous sélectionnez l'option **Exécuter le flux en cours** dans la barre d'outils, le vidage ne se produit qu'une fois par exécution du flux, ce qui signifie que vous pouvez quand même utiliser les caches en aval comme stockage temporaire lors d'une exécution unique. Par exemple, supposons que vous ayez défini un cache de milieu de flux après une opération de calcul complexe, et que vous disposez de plusieurs graphiques et rapports liés en aval de ce noeud Calculer. Lors de l'exécution du flux, le cache situé au niveau du noeud Calculer est vidé, puis de nouveau rempli, mais uniquement pour le premier graphique ou rapport. Les noeuds terminaux suivants lisent les données du cache du noeud Calculer. Si vous choisissez d'exécuter chaque noeud terminal individuellement (lorsque vous disposez de plusieurs noeuds terminaux) au lieu d'utiliser l'option **Exécuter le flux en cours**, le vidage du cache se produit à chaque exécution d'un noeud terminal.

Afficher les libellés de champ et de valeur dans le résultat. Affiche les libellés de champ et de valeur dans les tableaux, graphiques et autres sorties. S'il n'existe pas de libellé, ce sont les noms de champ et les valeurs de données qui apparaissent. Les libellés sont désactivés par défaut. Toutefois, vous pouvez activer ou désactiver chacun d'eux dans les autres composants d'IBM SPSS Modeler. Vous pouvez également afficher les libellés dans la fenêtre de sortie via un bouton bascule disponible dans la barre d'outils.

Figure 11. Icône de la barre d'outils permettant d'activer ou de désactiver les libellés de champ et de valeur

Afficher les durées d'exécution. Affiche les durées d'exécution individuelles pour les noeuds de flux dans l'onglet Durées d'exécution après l'exécution du flux. Pour plus d'informations, voir la rubrique «Affichage des durées d'exécution des noeuds», à la page 52.

La section **Création automatique de noeud** inclut les options suivantes pour création automatique de noeuds dans les flux individuels. Ces options déterminent si les nuggets de modélisation doivent être insérés dans le canevas du flux lors de la génération de nouveaux nuggets. Par défaut, ces options ne s'appliquent qu'aux flux créés dans la version 16 ou ultérieure. Dans IBM SPSS Modeler 16 et les versions ultérieures, si vous ouvrez un flux créé dans la version 15 ou une version antérieure, et exécutez un noeud de modélisation, le nugget n'est pas placé sur le canevas du flux comme c'était le cas dans les éditions précédentes. Si vous créez un nouveau flux à l'aide d'IBM SPSS Modeler version 16 (ou ultérieure) et exécutez un noeud modélisation, le nugget généré est placé sur le canevas du flux. Ceci est délibéré car, par exemple, l'option **Créer un noeud d'application de modèle pour la nouvelle sortie de modèle** disloquerait probablement les flux antérieurs à la version 16 s'exécutant par lots, dans des environnements IBM SPSS Collaboration and Deployment Services d'exécution par lots et dans d'autres environnements où l'interface utilisateur du client IBM SPSS Modeler Server n'est pas présente.

- **Créer un noeud d'application de modèle pour la nouvelle sortie de modèle.** Crée automatiquement des noeuds d'application de modèle pour la nouvelle sortie de modèle. Si vous sélectionnez cette option, vous pouvez toujours choisir dans l'option **Créer des liens de mise à jour de modèle** si les liens doivent être activés, désactivés, ou ne pas être créés.

  Lorsqu'un nouveau applicateur de modèle ou un noeud source est créé, les options de liens dans les menus déroulants déterminent si les liens de mise à jour entre le noeud de création et le nouveau noeud doivent être créés et, dans ce cas, leur mode. Si des liens sont créés, il est probable que vous voudrez les activer, mais ces options offrent néanmoins à l'utilisateur un contrôle total.
- **Créer des noeuds source à partir des générateurs de source.** Crée automatiquement des noeuds source à partir des générateurs de source. Similaire à l'option précédente. Si vous sélectionnez cette option, vous pouvez également déterminer dans le menu déroulant **Créer des liens d'actualisation source** si les liens d'actualisation doivent être activés, désactivés, ou ne pas être créés.

Enregistrer par défaut. Les options spécifiées s'appliquent uniquement au flux en cours. Cliquez sur ce bouton pour définir ces options comme paramètre par défaut de tous les flux.

Définition des options de date et heure pour les flux : Ces options spécifient le format à utiliser pour diverses expressions de date et d'heure dans le flux en cours.

Importer date/heure en tant que Indiquez si vous utilisez le stockage date/heure pour les champs date/heure ou si vous importez ces champs en tant que variables de chaîne.

Format de date Sélectionnez le format de date à utiliser pour les champs de stockage de date ou lorsque les chaînes sont interprétées comme des dates par les fonctions de date CLEM.

Format d'heure Sélectionnez le format d'heure à utiliser pour les champs de stockage d'heure ou lorsque les chaînes sont interprétées comme des heures par les fonctions d'heure CLEM.

Remise à zéro jours/minutes En ce qui concerne les formats temporels, indiquez si les différences horaires négatives sont interprétées comme se référant au jour précédent ou à l'heure précédente.

Date de référence (1er janv) Sélectionnez la date de référence (toujours le 1er janvier) utilisée par les fonctions de date CLEM opérant avec une date unique.

Année de référence des dates à deux chiffres Spécifiez l'année pivot pour ajouter les chiffres correspondant au siècle lorsque l'année est indiquée uniquement par deux chiffres. Par exemple, spécifiez 1930 comme année pivot pour supposer que la date 05/11/02 est en 2002. Le même paramètre utilise le 20ème siècle pour les dates après le 30 ; ainsi il sera supposé que la date 05/11/73 est en 1973.

Fuseau horaire Indiquez comment le fuseau horaire est choisi pour l'expression CLEM datetime_now.

- Si vous sélectionnez **Serveur**, le fuseau horaire dépend des éléments suivants :
  - Si le flux en cours utilise une source de données Analytic Server, l'expression datetime_now utilise l'heure d'Analytic Server ; par défaut, le serveur utilise l'heure du temps universel coordonné.
  - Si le flux en cours utilise un noeud source de base de données, les bases de données prises en charge utilisent la conversion SQL et l'expression datetime_now utilise l'heure de la base de données.
  - Pour tous les autres flux, le fuseau horaire utilise l'heure de SPSS Modeler Server.
- Si vous sélectionnez **Client Modeler**, le fuseau horaire est celui de la machine sur laquelle SPSS Modeler est installé.
- Vous pouvez également sélectionner l'une des valeurs de temps universel coordonné pour le fuseau horaire.

Enregistrer par défaut. Les options spécifiées s'appliquent uniquement au flux en cours. Cliquez sur ce bouton pour définir ces options comme paramètre par défaut de tous les flux.

Définition des options de format des nombres pour les flux : Ces options spécifient le format à utiliser pour diverses expressions numériques dans le flux en cours.

Format d'affichage des nombres. Vous pouvez sélectionner les formats d'affichage standard (**####.###**), scientifique (**#.###E+##**) ou monétaire (**### ## €**).

Nombre de décimales (au format standard, scientifique ou monétaire). Pour les formats d'affichage des nombres, indique le nombre de décimales à utiliser lorsque des nombres réels sont affichés ou imprimés. Cette option apparaît séparément pour chaque format d'affichage.

Calculs en. Sélectionnez **Radians** ou **Degrés** comme unité de mesure à utiliser dans les expressions CLEM trigonométriques. Pour plus d'informations, voir la rubrique «Fonctions trigonométriques», à la page 167.

Enregistrer par défaut. Les options spécifiées s'appliquent uniquement au flux en cours. Cliquez sur ce bouton pour définir ces options comme paramètre par défaut de tous les flux.

Définition des options d'optimisation pour les flux : Vous pouvez utiliser les paramètres d'optimisation pour optimiser la performance du flux. Les paramètres de performances et d'optimisation d'IBM SPSS Modeler Server (si ce dernier est utilisé) annulent les paramètres équivalents définis pour le client. Si ces paramètres sont désactivés sur le serveur, le client ne peut pas les activer. Par contre, s'ils sont activés sur le serveur, le client peut choisir de les désactiver.

Remarque : La modélisation de base de données et l'optimisation SQL requièrent l'activation de la connectivité à IBM SPSS Modeler Server sur l'ordinateur IBM SPSS Modeler. Avec ce paramètre activé, vous pouvez accéder aux algorithmes de la base de données, effectuer le push back de SQL directement depuis IBM SPSS Modeler et accéder à IBM SPSS Modeler Server. Pour vérifier le statut actuel de la licence, choisissez ce qui suit dans le menu IBM SPSS Modeler.

Aide > **A propos de** > **Détails supplémentaires**

Si la connectivité est activée, vous voyez l'option **Activation du serveur** dans l'onglet Etat de la licence.

Pour plus d'informations, voir «Connexion à IBM SPSS Modeler Server», à la page 10.

Remarque : La prise en charge des conversions et de l'optimisation SQL dépend du type de base de données utilisé. Pour obtenir les informations les plus récentes sur les bases de données et pilotes ODBC pris en charge et testés pour une utilisation avec IBM SPSS Modeler, consultez le site Web de support technique de l'entreprise à l'adresse http://www.ibm.com/support.

Autoriser la réécriture des flux. Sélectionnez cette option pour activer la réécriture de flux dans IBM SPSS Modeler. Quatre types de réécriture sont disponibles et vous pouvez en sélectionner un ou plusieurs. La réécriture de flux réorganise les noeuds d'un flux en arrière-plan, sans pour autant modifier sa sémantique, afin qu'il soit exécuté de manière plus efficace.

- **Optimiser la génération SQL.** Cette option permet de réorganiser les noeuds du flux afin d'augmenter le nombre d'opérations pouvant être répercutées, à l'aide d'une génération SQL, pour y être exécutées dans la base de données. Lorsqu'il détecte un noeud qui ne peut pas être affiché sous forme de code SQL, le programme d'optimisation recherche plus en avant tout noeud qui peut l'être et qu'il est possible de ramener avant le noeud problématique sans risquer de modifier la sémantique du flux. Non seulement la base de données peut effectuer les opérations plus efficacement qu'IBM SPSS Modeler, mais de telles conversions contribuent également à réduire le volume des données renvoyées à IBM SPSS Modeler pour être traitées. Vous réduisez ainsi le trafic sur le réseau et accélérez les opérations de flux. Veuillez noter que la case **Générer SQL** doit être sélectionnée pour que l'optimisation SQL ait le moindre effet.
- **Optimiser l'expression CLEM** Cette option permet à l'optimiseur de rechercher des expressions CLEM qui peuvent être prétraités avant l'exécution du flux, afin d'accroître la vitesse de traitement. Exemple simple : si vous avez une expression telle que *log(salary)*, l'optimiseur calcule la valeur de salary et la transmet pour traitement. Cela peut être utilisé pour améliorer les conversions SQL et les performances IBM SPSS Modeler Server.
- **Optimiser l'exécution de syntaxe.** Cette méthode de réécriture de flux augmente l'efficacité des opérations qui incorporent plus d'un noeud contenant une syntaxe IBM SPSS Statistics. L'optimisation est réalisée en combinant des commandes de syntaxe dans une opération unique, au lieu d'exécuter chacune d'elle en tant qu'opération séparée.
- **Optimiser les autres exécutions.** Cette méthode de réécriture de flux augmente l'efficacité des opérations exécutées (c'est-à-dire des opérations dont l'exécution ne peut pas être déléguée à la base de données). L'optimisation est obtenue en réduisant dès que possible la quantité de données dans le flux. Tout en conservant l'intégrité des données, le flux est réécrit de façon à rapprocher les opérations de la source de données, ceci afin de réduire le volume des données en aval pour les opérations coûteuses, telles que les jointures.

Activer le traitement parallèle. Lorsque vous exécutez ce programme sur un ordinateur doté de plusieurs processeurs, cette option permet au système d'équilibrer la charge entre les différents processeurs et d'augmenter ainsi les performances. Le traitement parallèle peut s'avérer bénéfique en cas d'utilisation de plusieurs noeuds ou des noeuds individuels suivants : C5.0, Fusionner (par clé), Trier, Discrétiser (méthodes des rangs et des quantiles) et Agréger (à l'aide d'un ou de plusieurs champs-clés).

Générer SQL. Sélectionnez cette option pour activer la génération SQL et répercuter ainsi les opérations de flux dans la base de données à l'aide du code SQL pour générer les processus d'exécution ; cette opération peut améliorer les performances. Pour augmenter encore les performances, vous pouvez également sélectionner l'option **Optimiser la génération SQL**, afin d'accroître au maximum le nombre d'opérations répercutées dans la base de données. Si les opérations d'un noeud ont été retransmises à la base de données, le noeud apparaît en violet au cours de l'exécution du flux.

- **Mise en cache de la base de données.** Dans le cas des flux exécutés dans la base de données qui génèrent du code SQL, les données peuvent être mises en mémoire cache en milieu de flux dans une table temporaire de la base de données, plutôt que dans le système de fichiers. Vous pouvez combiner cela à la fonction d'optimisation SQL et accroître ainsi les performances de manière significative. Par exemple, la sortie d'un flux qui fusionne plusieurs tables afin de créer une vue d'exploration de données peut être mise en mémoire cache et réutilisée selon les besoins. Une fois la mise en mémoire cache de base de données activée, il vous suffit de cliquer à l'aide du bouton droit de la souris sur un noeud non terminal afin de mettre les données en mémoire cache à ce stade du flux ; le cache est

automatiquement créé dans la base de données lors de l'exécution suivante du flux. Le code SQL peut ainsi être généré pour les noeuds en aval, ce qui améliore encore les performances. Vous pouvez également désactiver cette option si besoin est, par exemple lorsque des règles ou des permissions empêchent l'écriture des données dans la base de données. Si la mise en mémoire cache de base de données ou l'optimisation SQL n'est pas activée, le cache est écrit dans le système de fichiers. Pour plus d'informations, voir «Options de mise en cache des noeuds», à la page 41.

- **Utiliser une conversion élargie.** Cette option permet la conversion des données de chaînes à nombres ou de nombres à chaînes, si elles sont stockées dans un format approprié. Par exemple, si les données sont conservées dans la base de données en tant que chaîne, mais contiennent en fait un nombre important, les données peuvent être converties pour être utilisées au moment des conversions.

Remarque : En raison de différences d'implémentation SQL mineures, les flux exécutés dans une base de données peuvent renvoyer des résultats légèrement différents de ceux renvoyés lors de leur exécution dans IBM SPSS Modeler. Pour ces mêmes raisons, ces différences peuvent également varier en fonction du fournisseur de base de données.

Enregistrer par défaut. Les options spécifiées s'appliquent uniquement au flux en cours. Cliquez sur ce bouton pour définir ces options comme paramètre par défaut de tous les flux.

Définition des options de statut de consignation et d'enregistrement SQL pour les flux : Ces paramètres incluent diverses options contrôlant l'affichage des instructions SQL générées par le flux et l'affichage du nombre d'enregistrements traités par le flux.

Afficher SQL dans le fichier journal des messages lors de l'exécution du flux. Détermine l'insertion dans le journal des messages du code SQL généré lors de l'exécution du flux.

Afficher les détails de la génération SQL dans le fichier journal des messages lors de la préparation du flux. Au cours de la prévisualisation du flux, détermine l'insertion dans le journal des messages d'un aperçu du code SQL susceptible d'être généré.

Afficher SQL. Indique si le code SQL affiché dans le journal doit contenir des fonctions SQL natives ou des fonctions ODBC standard sous la forme `{fn FUNC(...)}`, telles que générées par SPSS Modeler. Les premières reposent sur certaines fonctionnalités du pilote ODBC, qui ne sont peut-être pas implémentées.

Reformater SQL pour une meilleure lisibilité. Indique si le SQL affiché dans le journal doit être formaté pour être lu.

Mettre à jour le compteur d'enregistrements. Indique le moment où les enregistrements doivent être enregistrés à mesure qu'ils arrivent dans les noeuds terminaux. Indiquez le chiffre déterminant la mise à jour du statut tous les *N* enregistrements.

Enregistrer par défaut. Les options spécifiées s'appliquent uniquement au flux en cours. Cliquez sur ce bouton pour définir ces options comme paramètre par défaut de tous les flux.

Définition d'options de présentation pour les flux : Ces paramètres donnent plusieurs options liées à l'affichage et à l'utilisation des espaces de travail de flux.

Largeur minimale de l'espace de travail de flux. Précisez la largeur minimale de l'espace de travail de flux en pixels.

Hauteur minimale de l'espace de travail de flux. Précisez la hauteur minimale de l'espace de travail de flux en pixels.

Vitesse de défilement du flux. Spécifiez la vitesse de défilement de l'espace de travail du flux pour contrôler la vitesse à laquelle défile le panneau de l'espace de travail de flux lorsque vous faites glisser un noeud d'un endroit à l'autre dans l'espace de travail. Plus les chiffres sont élevés, plus la vitesse de défilement est rapide.

Nb car. nom des noeuds. Indiquez la limite en nombre de caractères pour le nom des noeuds de l'espace de travail de flux.

Taille des noeuds. Sélectionnez une option pour redimensionner la vue entière du flux à une taille comprise entre 8 % et 200 % de la taille d'icône standard.

Taille des cellules de la grille. Sélectionnez la taille de cellule de grille dans la liste. Ce nombre est utilisé pour l'alignement des noeuds dans l'espace de travail de flux via une grille invisible. La taille de cellule de grille par défaut est de 0,25.

Ajuster à la grille. Sélectionnez cette option pour aligner les icônes sur une grille invisible (option sélectionnée par défaut).

Emplacement d'icône généré. Dans l'espace de travail, choisissez où placer les icônes des noeuds générés à partir des nuggets de modèle. L'emplacement par défaut se situe en haut à gauche.

Enregistrer par défaut. Les options spécifiées s'appliquent uniquement au flux en cours. Cliquez sur ce bouton pour définir ces options comme paramètre par défaut de tous les flux.

Propriétés du flux Analytic Server : Ces propriétés offrent diverses options pour travailler avec Analytic Server.

Nombre maximal d'enregistrements à traiter en dehors de Analytic Server
Indiquez le nombre maximum d'enregistrements à importer dans le serveur SPSS Modeler depuis une source de données Analytic Server.

Notification lorsqu'un noeud ne peut pas être traité dans Analytic Server
Ce paramètre détermine ce qui se passe lorsqu'un flux allant être soumis à Analytic Server contient un noeud ne pouvant pas être traité dans Analytic Server. Indiquez si le système doit émettre un avertissement et continuer de traiter le flux ou s'il doit générer une erreur et interrompre le traitement.

Paramètres de stockage du modèle Scission

Stockage de modèles de scission par référence sur Analytic Server en cas de dépassement de la taille de modèle (Mo)
Les nuggets de modèles sont généralement stockés comme une partie du flux. Les modèles de scission comportant de nombreuses scissions peuvent générer des nuggets de grande taille et le fait de déplacer le nugget continuellement entre le flux et l'Analytic Server peut avoir un impact sur la performance. Pour remédier à ce problème, les modèles de scission sont stockés sur l'Analytic Server lorsqu'ils dépassent la taille spécifiée et le nugget du serveur SPSS Modeler contient alors une référence au modèle.

Dossier par défaut pour le stockage des modèles par référence sur Analytic Server une fois l'exécution terminée
Indiquez le chemin par défaut où vous souhaitez stocker les modèles de scission sur Analytic Server. Le chemin doit commencer par un nom de projet Analytic Server valide.

Dossier pour le stockage des modèles promus
Indiquez le chemin par défaut où vous souhaitez stocker les modèles "promus". Un modèle promu n'est pas nettoyé lorsque la session SPSS Modeler se termine.

Définition d'options géospatiales pour les flux : N'importe quel champ géospatial, qu'il s'agisse d'une forme, d'une coordonnée, ou d'une valeur d'axe unique (telle que x ou y, ou une latitude et longitude), a un système de coordonnées associé. Ce système de coordonnées définit des attributs tels que le point d'origine (0,0) et les unités associées aux valeurs.

Il existe un certain nombre de système de coordonnées et ils sont de deux types : Géographiques et Projetées. Toutes les fonctions spatiales dans SPSS Modeler ne peuvent être utilisées qu'avec un système de coordonnées projetées.

En raison de la nature des systèmes de coordonnées, la fusion ou l'ajout de données de deux sources de données géospatiales distinctes requiert que les sources utilisent le même système de coordonnées. C'est pourquoi vous devez spécifier une valeur de coordonnée pour toute donnée géospatiale utilisée dans le flux.

Les données sont reprojetées automatiquement afin d'utiliser le système de coordonnées du flux choisi dans les situations suivantes :

- Pour les fonctions spatiales (telles que area, closeto, within), le paramètre transmis à la fonction est reprojeté automatiquement ; cependant, les données de la ligne d'origine sont conservées inchangées.
- Lors d'une utilisation des noeuds de génération ou de scoring (nugget) dans STP (prévision spatio-temporelle), la zone d'emplacement est reprojetée automatiquement. Lors du scoring, l'emplacement issu du nugget correspond à l'emplacement d'origine.
- Lors de l'utilisation du noeud Visualisation de carte.

Système de coordonnées de flux. Disponible uniquement si vous cochez la case. Cliquez sur **Modifier** pour afficher la liste des Systèmes de coordonnées projetées et sélectionner celui à utiliser pour le flux en cours.

Enregistrer par défaut. Le système de coordonnées que vous sélectionnez ne s'applique qu'au flux en cours. Pour sélectionner ce système comme système par défaut pour tous les flux, cliquez sur ce bouton.

Sélection de système de coordonnées géospatiales : Toutes les fonctions spatiales dans SPSS Modeler ne peuvent être utilisées qu'avec un système de coordonnées projetées.

La boîte de dialogue Select Stream Coordinate System contient une liste de tous les systèmes de coordonnées projetées que vous pouvez sélectionner pour des données géospatiales utilisées dans un flux.

Les informations suivantes sont fournies pour chaque système de coordonnées.

- **WKID** ID notoire unique à chaque système de coordonnées.
- **Nom** Nom du système de coordonnées.
- **Unités** Unité de mesure associée au système de coordonnées.

Outre la liste de tous les systèmes de coordonnées, la boîte de dialogue comporte également un contrôle **Filtrage**. Si vous connaissez le nom complet ou partiel du système de coordonnées dont vous avez besoin, entrez-le dans la zone **Nom**, en bas de la boîte de dialogue. La liste des systèmes de coordonnées disponibles est filtrée afin d'afficher uniquement ceux dont le nom contient le texte que vous avez saisi.

Affichage des messages d'opération de flux

Les messages concernant les opérations de flux comme l'exécution, l'optimisation, la durée de création et d'évaluation d'un modèle peuvent être facilement consultés grâce à l'onglet Messages de la boîte de dialogue des propriétés du flux. Les messages d'erreur sont également indiqués dans ce tableau.

Pour afficher les messages de flux

1. Dans le menu Fichier, cliquez sur **Propriétés du flux** (ou sélectionnez le flux dans l'onglet Flux du panneau des Gestionnaires, faites un clic droit puis cliquez sur **Propriétés du flux** dans le menu contextuel).
2. Cliquez sur l'onglet **Messages**.

Sinon, dans le menu Outils, cliquez sur :

Propriétés du flux > **Messages**

Les messages d'erreur sont affichés ici, avec les messages concernant les opérations sur le flux. Lorsque l'exécution du flux est interrompue à cause d'une erreur, cette boîte de dialogue apparaît dans l'onglet Messages et affiche le message d'erreur. En outre, le noeud comportant les erreurs est mis en évidence en rouge dans l'espace de travail de flux.

Si les options de consignation et d'optimisation SQL sont activées dans la boîte de dialogue Options utilisateur, les informations relatives au code SQL généré apparaissent également. Pour plus d'informations, voir la rubrique «Définition des options d'optimisation pour les flux», à la page 47.

Vous pouvez enregistrer les messages indiqués ici se rapportant à un flux en cliquant sur **Enregistrer les messages** dans la liste déroulante du bouton d'enregistrement (à gauche, juste en-dessous de l'onglet Messages).

Vous pouvez effacer tous les messages d'un flux donné en sélectionnant **Supprimer tous les messages** dans le menu déroulant.

Notez que l'utilisation de l'unité centrale indique le temps d'utilisation de l'unité centrale par le processus serveur. Le temps écoulé est le temps total entre le début et la fin de l'exécution, et inclut donc également des opérations comme le transfert de fichiers et le rendu des résultats. Le temps d'utilisation de l'unité centrale peut être supérieur au temps écoulé lorsque un flux exploite plusieurs unités centrales (exécution parallèle). Lorsqu'un flux renvoie complètement à la base de données utilisée comme source de données, l'utilisation de l'unité centrale est égale à zéro.

Affichage des durées d'exécution des noeuds

Dans l'onglet Messages, vous pouvez choisir d'afficher les durées d'exécution, ce qui vous permet de consulter les durées d'exécution individuelles pour tous les noeuds du flux qui s'exécutent dans IBM SPSS Modeler Server. Notez que les durées peuvent ne pas être précises pour les exécutions de flux dans d'autres zones, telles que R ou Analytic Server.

Remarque : Pour que cette fonction fonctionne, la case **Afficher les durées d'exécution** doit être sélectionnée dans le paramètre **Général** de l'onglet **Options**.

Dans la table des durées d'exécution des noeuds, les colonnes sont les suivantes. Cliquez sur un en-tête de colonne pour trier les entrées par ordre croissant ou décroissant (par exemple, pour voir les noeuds avec les plus longues durées d'exécution).

Noeud terminal. L'identifiant de la branche à laquelle le noeud appartient. L'identifiant est le nom du noeud terminal à la fin de la branche.

Libellé de noeud. Le nom du noeud auquel la durée d'exécution se rapporte.

ID de noeud. L'identifiant unique du noeud auquel la durée d'exécution se rapporte. Cet identifiant est généré par le système lorsque le noeud est créé.

Durée(s) d'exécution. La durée en secondes nécessaire à l'exécution de ce noeud.

Définition des paramètres de flux et de session

Vous pouvez définir des paramètres afin de les utiliser dans les expressions CLEM et pour la génération de scripts. Il s'agit, en effet, de variables définies par l'utilisateur qui sont enregistrées et conservées dans le flux, la session ou le super noeud en cours ; elles sont accessibles à partir de l'interface utilisateur, ainsi que via la fonction de génération de scripts. Si vous enregistrez un flux, par exemple, les paramètres définis pour ce dernier sont également enregistrés. (Cela permet de les différencier des variables de script locales qui ne peuvent être utilisées que dans le script dans lequel elles sont déclarées.) Les paramètres sont souvent utilisés lors de la génération de scripts pour contrôler le comportement du script, en fournissant des informations sur les champs et les valeurs qui n'ont pas besoin d'être codées en dur dans le script.

La portée d'un paramètre dépend de l'endroit où ce dernier est défini :

- Les paramètres de flux peuvent être définis dans un script de flux ou dans la boîte de dialogue des propriétés du flux ; ils sont disponibles pour tous les noeuds du flux. Ils sont affichés sur la liste des paramètres dans le Générateur de formules.
- Les paramètres de session peuvent être définis dans un script autonome ou dans la boîte de dialogue des paramètres de session. Ils sont disponibles pour tous les flux utilisés dans la session en cours (tous les flux répertoriés dans l'onglet Flux du panneau des gestionnaires).

Il est également possible de définir des paramètres pour les super noeuds ; dans ce cas, les paramètres ne sont visibles que par les noeuds encapsulés au sein de ce super noeud.

Pour définir des paramètres de flux et de session via l'interface utilisateur

1. Pour définir des paramètres de flux, dans le menu principal, cliquez sur :

   Outils > **Propriétés de flux** > **Paramètres**
2. Pour définir des paramètres de session, cliquez sur **Définir les paramètres de session** dans le menu Outils.

Invite ?. Cochez cette case si vous souhaitez que l'utilisateur soit invité à saisir une valeur pour ce paramètre au moment de l'exécution.

Nom. Les noms des paramètres sont répertoriés ici. Vous pouvez créer un paramètre en entrant un nom dans ce champ. Par exemple, pour créer un paramètre relatif à la température minimale, vous pouvez saisir `Valeur min.`. N'insérez pas le préfixe `$P-`, qui indique un paramètre dans les expressions CLEM. Ce nom est également utilisé pour l'affichage dans le Générateur de formules de CLEM.

Nom complet. Répertorie le nom descriptif de chaque paramètre créé.

Stockage. Sélectionnez le type de stockage dans la liste. Indique le mode de stockage des valeurs de données dans le paramètre. Par exemple, si vous utilisez des valeurs commençant par des zéros à conserver (comme 008), vous devez sélectionner **Chaîne** comme type de stockage. Sinon, les zéros seront supprimés de la valeur. Les types de stockage disponibles sont les suivants : chaîne, entier, réel, temps, date et horodatage. Pour les paramètres de date, les valeurs doivent être définies à l'aide de la notation standard ISO telle qu'elle est présentée dans le paragraphe suivant.

Valeur. Indique la valeur en cours du paramètre sélectionné. Modifiez ce paramètre selon les besoins. Pour les paramètres de date, les valeurs doivent être définies à l'aide de la notation standard ISO (soit, `YYYY-MM-DD`). Toute date définie dans un autre format est refusée.

Type (facultatif). Si vous prévoyez de déployer le flux vers une application externe, sélectionnez un niveau de mesure dans la liste. Sinon, il est conseillé de laisser la colonne *Type* en l'état. Si vous souhaitez spécifier des contraintes de valeur pour le paramètre, telles que des limites supérieures et inférieures d'un intervalle numérique, sélectionnez **Spécifier** dans la liste.

Vous ne pouvez définir les options de nom complet, de stockage et de type pour les paramètres que dans l'interface utilisateur. Il est impossible de définir ces options à l'aide de scripts.

Cliquez sur les flèches à droite pour déplacer le paramètre sélectionné vers le haut ou le bas de la liste des paramètres disponibles. Utilisez le bouton de suppression (indiqué par un *X*) pour supprimer le paramètre sélectionné.

Spécification des invites d'exécution pour les valeurs de paramètre

Si vous possédez des flux où il est possible que vous soyez amené à entrer différentes valeurs pour le même paramètre en différentes occasions, vous pouvez spécifier des invites d'exécution pour un ou plusieurs flux ou des valeurs de paramètre de session.

Paramètres. (Facultatif) Saisissez une valeur pour le paramètre, ou conservez la valeur par défaut s'il y en a une.

Désactivez ces invites. Cochez cette case si vous ne souhaitez pas que ces invites apparaissent lorsque vous exécutez le flux. Vous pouvez les réafficher en cochant la case **Invite ?** dans la boîte de dialogue Propriétés des flux ou Propriétés de la session où les paramètres ont été définis. Pour plus d'informations, voir la rubrique «Définition des paramètres de flux et de session», à la page 53.

Spécification des contraintes des valeurs d'un type de paramètre

Vous pouvez faire en sorte que des contraintes de valeur pour un paramètre soient disponibles au cours du déploiement du flux dans une application externe lisant les flux de modélisation des données. Cette boîte de dialogue vous permet de spécifier les valeurs mises à la disposition des utilisateurs externes exécutant le flux. Dans la boîte de dialogue, les contraintes de valeur varient de manière dynamique en fonction du type de données. Les options présentées ici sont identiques à celles proposées pour les valeurs du noeud Typer.

Type. Affiche le type de niveau de mesure actuellement sélectionné. Vous pouvez modifier cette valeur pour indiquer la façon dont vous souhaitez utiliser les paramètres dans IBM SPSS Modeler.

Stockage. Affiche le type de stockage, s'il est connu. Les types de stockage ne sont pas affectés par le niveau de mesure (Continu, Nominal ou Indicateur) choisi pour travailler dans IBM SPSS Modeler. Vous pouvez changer de type de stockage dans l'onglet Paramètres principal.

La partie inférieure de la boîte de dialogue change de manière dynamique en fonction du niveau de mesure sélectionné dans le champ **Type**.

Niveaux de mesure continus

Inférieur. Indiquez la limite inférieure des valeurs de paramètre.

Supérieur. Indiquez la limite supérieure des valeurs de paramètre.

Libellés. Vous pouvez spécifier des libellés pour toutes les valeurs d'un champ d'intervalle. Cliquez sur le bouton **Libellés** pour ouvrir une boîte de dialogue distincte permettant de spécifier les libellés de valeur.

Niveaux de mesure nominaux

Valeurs. Cette option vous permet de spécifier les valeurs d'un paramètre utilisé en tant que champ nominal. Ces valeurs ne sont pas introduites automatiquement dans le flux IBM SPSS Modeler, mais utilisées dans une liste déroulante destinée aux applications de déploiement externes. A l'aide des flèches et du bouton Supprimer, vous pouvez modifier, réorganiser ou supprimer les valeurs existantes.

Niveaux de mesure indicateurs

Vrai. Spécifiez la valeur indicateur du paramètre lorsque la condition est respectée.

Faux. Spécifiez la valeur indicateur du paramètre lorsque la condition n'est pas respectée.

Libellés. Vous pouvez spécifier des libellés pour les valeurs d'un champ indicateur.

Options de déploiement des flux

L'onglet Déploiement de la boîte de dialogue Propriétés du flux vous permet de spécifier des options pour le déploiement du flux dans IBM SPSS Collaboration and Deployment Services afin de rafraîchir un modèle, de programmer un travail automatisé ou une autre utilisation par IBM Analytical Decision Management. Tous les flux nécessitent une branche de scoring pour pouvoir être déployés. Pour plus d'informations, voir la rubrique «Stockage et déploiement des objets du référentiel», à la page 192.

Exécution de bouclage pour les flux

A l'aide de l'onglet Exécution de la boîte de dialogue des propriétés du flux, vous pouvez définir des conditions d'exécution de bouclage pour automatiser les tâches répétitives dans le flux en cours.

Une fois que vous avez défini ces conditions, vous pouvez les utiliser comme introduction à la génération de scripts car elles alimentent la fenêtre de script à l'aide de données de script de base pour votre flux que vous pouvez ensuite modifier (pour l'utiliser peut-être comme base à partir de laquelle générer de meilleurs scripts). Pour plus d'informations, voir la rubrique «Fonctions globales», à la page 186.

Pour définir un bouclage pour un flux

1. Dans le menu Fichier, cliquez sur **Propriétés du flux** (ou sélectionnez le flux dans l'onglet Flux du panneau des Gestionnaires, faites un clic droit puis cliquez sur **Propriétés du flux** dans le menu contextuel).
2. Cliquez sur l'onglet **Exécution**.
3. Sélectionnez le mode d'exécution **Bouclage / Exécution conditionnelle**.
4. Cliquez sur l'onglet **Bouclage**.

Sinon, dans le menu Outils, cliquez sur :

Propriétés de flux > **Exécution**

Vous pouvez également cliquer avec le bouton droit sur le noeud puis, dans le menu contextuel, sélectionner :

Bouclage / Exécution conditionnelle > **Edition des paramètres de bouclage**

Itération. Vous ne pouvez pas éditer la valeur de ce numéro de ligne, mais vous pouvez ajouter, supprimer ou déplacer une itération vers le haut ou vers le bas à l'aide des boutons situés à droite du tableau.

En-têtes du tableau. Ils indiquent la clé d'itération et les variables d'itération que vous avez créées lors de la définition de la boucle.

Affichage des valeurs globales des flux

A l'aide de l'onglet Valeurs globales de la boîte de dialogue des propriétés du flux, vous pouvez afficher les valeurs globales définies pour le flux en cours. Les valeurs globales sont créées à l'aide d'un noeud V. globales (Valeurs globales). Elles permettent de déterminer des statistiques, telles que la moyenne, la somme ou l'écart-type des champs sélectionnés.

Une fois le noeud V. globales (Valeurs globales) exécuté, ces valeurs sont alors disponibles pour diverses utilisations dans les opérations de flux. Pour plus d'informations, voir la rubrique «Fonctions globales», à la page 186.

Pour afficher les valeurs globales d'un flux

1. Dans le menu Fichier, cliquez sur **Propriétés du flux** (ou sélectionnez le flux dans l'onglet Flux du panneau des Gestionnaires, faites un clic droit puis cliquez sur **Propriétés du flux** dans le menu contextuel).
2. Cliquez sur l'onglet **Globales**.

Sinon, dans le menu Outils, cliquez sur :

Propriétés de flux > **Valeurs globales**

Valeurs globales disponibles. Les valeurs globales disponibles sont répertoriées dans ce tableau. Vous ne pouvez pas éditer les valeurs globales à ce niveau. Toutefois, vous pouvez supprimer toutes les valeurs globales pour un flux en vous servant du bouton Effacer toutes les valeurs situé à droite du tableau.

Recherche de noeuds dans un flux

Vous pouvez rechercher des noeuds dans un flux en spécifiant un certain nombre de critères, comme le nom, la catégorie ou l'identificateur de ces noeuds. Cette option est particulièrement utile pour les flux complexes qui contiennent un grand nombre de noeuds.

Pour rechercher des noeuds dans un flux

1. Dans le menu Fichier, cliquez sur **Propriétés du flux** (ou sélectionnez le flux dans l'onglet Flux du panneau des Gestionnaires, faites un clic droit puis cliquez sur **Propriétés du flux** dans le menu contextuel).
2. Cliquez sur l'onglet **Rechercher**.

Sinon, dans le menu Outils, cliquez sur :

Propriétés du flux > **Rechercher**

Vous pouvez spécifier plusieurs options afin de limiter la recherche, mais la recherche par ID de noeud (à l'aide du champ **ID égal à**) exclut les autres options.

Le libellé de noeud contient. Cochez cette case et saisissez la totalité ou une partie du libellé du noeud pour rechercher un noeud spécifique. Les recherches ne sont pas sensibles à la casse, et plusieurs mots peuvent être traités comme un seul extrait de texte.

Catégorie de noeud. Cochez cette case et cliquez sur une catégorie de la liste pour rechercher un type de noeud spécifique. **Noeud d'exécution** signifie un noeud de l'onglet Ops sur lignes ou Ops sur champs de la palette des noeuds ; **Appliquer le noeud du modèle** se rapporte à un nugget de modèle.

Les mots clés comprennent. Cochez cette case et saisissez un ou plusieurs mots-clés complets pour rechercher des noeuds qui contiennent ce texte dans le champ Mots-clés dans l'onglet Annotations de la boîte de dialogue Noeud. Le texte des mots-clés que vous saisissez doit correspondre parfaitement. Séparez plusieurs mots-clés par des points-virgules pour rechercher des alternatives (par exemple, saisissez `proton;neutron` pour rechercher tous les noeuds contenant l'un ou l'autre de ces mots-clés. Pour plus d'informations, voir la rubrique «Annotations», à la page 64.

L'annotation contient. Cochez cette case et saisissez un ou plusieurs mots pour rechercher des noeuds qui contiennent ce texte dans la zone de texte principale de l'onglet Annotations de la boîte de dialogue Noeud. Les recherches ne sont pas sensibles à la casse, et plusieurs mots peuvent être traités comme un seul extrait de texte. Pour plus d'informations, voir la rubrique «Annotations», à la page 64.

Génère un champ nommé. Cochez cette case et saisissez le nom d'un champ généré (par exemple, drug-$C). Vous pouvez utiliser cette option pour rechercher des noeuds de modélisation qui génèrent un champ particulier. Saisissez un seul nom de champ, qui doit correspondre parfaitement.

L'ID est égale à. Cochez cette case et saisissez l'ID du noeud pour rechercher un noeud particulier avec cet identificateur (sélectionner cette option désactive toutes les options précédentes). Les ID de noeud sont attribués par le système lorsqu'un noeud est créé et peuvent être utilisés pour référencer un noeud à des fins de générations de scripts ou d'automatisation. Saisissez un seul ID de noeud, qui doit correspondre parfaitement. Pour plus d'informations, voir la rubrique «Annotations», à la page 64.

Rechercher dans les super noeuds. Cette case est cochée par défaut, ce qui signifie que la recherche est effectuée sur des noeuds situés à la fois à l'intérieur et à l'extérieur des super noeuds. Décochez cette case si vous souhaitez effectuer une recherche sur les noeuds à l'extérieur des super noeuds uniquement, au niveau supérieur du flux.

Rechercher. Après avoir spécifié toutes les options désirées, cliquez sur ce bouton pour commencer la recherche.

Les noeuds correspondants aux options spécifiées sont répertoriés dans la partie inférieure de la boîte de dialogue. Sélectionnez un noeud dans la liste pour le mettre en surbrillance dans l'espace de travail du flux.

Changement de nom des flux

A l'aide de l'onglet Annotations de la boîte de dialogue des propriétés du flux, vous pouvez ajouter des annotations descriptives à un flux et donner un nom personnalisé à ce dernier. Ces options sont utiles plus particulièrement lorsque vous générez des rapports sur les flux ajoutés au panneau de projet. Pour plus d'informations, voir la rubrique «Annotations», à la page 64.

Descriptions de flux

Pour chaque flux que vous créez, IBM SPSS Modeler crée une description de flux contenant des informations concernant le contenu du flux. Cela peut s'avérer utile si vous essayez de déterminer les actions d'un flux mais que vous n'avez pas installé IBM SPSS Modeler, par exemple lorsque vous accédez à un flux via IBM SPSS Collaboration and Deployment Services.

La description du flux s'affiche sous la forme d'un document HTML se composant de plusieurs sections.

Informations générales sur le flux

Cette section contient le nom du flux ainsi que des détails sur le moment de la création du flux et de son dernier enregistrement.

Description et commentaires

Cette section comprend :

- toutes les annotations du flux (reportez-vous à «Annotations», à la page 64)
- tous les commentaires non connectés à des noeuds spécifiques
- tous les commentaires connectés à des noeuds dans les branches de modélisation et de scoring du flux

Informations de scoring

Cette section contient des informations sous divers en-têtes en rapport avec la branche de scoring du flux.

- **Commentaires.** Comprend des commentaires liés uniquement à des noeuds dans la branche de scoring.
- **Entrées.** Répertorie les champs d'entrée, ainsi que leur type de stockage (par exemple, chaîne, entier, réel, etc.).
- **Sorties.** Répertorie les champs de sortie, y compris les champs supplémentaires générés par le noeud de modélisation, ainsi que leur type de stockage.

- **Paramètres.** Répertorie tous les paramètres en rapport avec la branche de scoring du flux et ceux qui peuvent être affichés ou modifiés à chaque fois que le modèle est évalué. Ces paramètres sont identifiés lorsque vous cliquez sur le bouton **Paramètres de scoring** dans l'onglet **Déploiement** de la boîte de dialogue Propriétés du flux.
- **Noeud du modèle.** Affiche le nom et le type du modèle (par exemple, Réseau de neurones, Arbre C&R, etc.). Il s'agit du nugget de modèle sélectionné pour le champ **noeud Modèle** dans l'onglet **Déploiement** de la boîte de dialogue Propriétés du flux.
- **Détails de modèle.** Affiche des détails relatifs au nugget de modèle identifié sous l'en-tête précédent. Lorsque c'est possible, l'importance des variables prédicteur et des graphiques d'évaluation du modèle sont inclus.

Informations de modélisation

Contient des informations concernant la branche modélisation du flux.

- **Commentaires.** Répertorie tous les commentaires ou annotations connectés aux noeuds de la branche de modélisation.
- **Entrées.** Répertorie les champs d'entrée ainsi que leur rôle dans la branche de modélisation (sous la forme d'une valeur de rôle de champ, par exemple, Entrée, Cible, Division, etc.).
- **Paramètres.** Répertorie tous les paramètres en rapport avec la branche de modélisation du flux et ceux qui peuvent être affichés ou modifiés à chaque fois que le modèle est mis à jour. Ces paramètres sont identifiés lorsque vous cliquez sur le bouton **Paramètres du concepteur de modèle** sur l'onglet **Déploiement** de la boîte de dialogue Propriétés du flux.
- **Noeud de modélisation.** Affiche le nom et le type du noeud de modélisation utilisé pour générer ou mettre à jour le modèle.

Prévisualisation des descriptions de flux

Vous pouvez afficher le contenu de la description d'un flux dans un navigateur Web en cliquant sur une option dans la boîte de dialogue Propriétés du flux. Le contenu de la description dépend des options que vous avez spécifiées dans l'onglet Déploiement de la boîte de dialogue. Pour plus d'informations, voir la rubrique «Options de déploiement des flux», à la page 205.

Pour afficher une description de flux :

1. Dans le menu IBM SPSS Modeler principal, cliquez sur :
   Outils > **Propriétés de flux** > **Déploiement**
2. Définissez le type de déploiement, le noeud de scoring désigné et tous les paramètres de scoring.
3. Si le type de déploiement est Rafraîchissement de modèle, vous pouvez aussi sélectionner :
   - un noeud de modélisation et tous les paramètres de création de modèle
   - un nugget de modèle sur la branche de scoring du flux
4. Cliquez sur le bouton **Aperçu de la description du flux**.

Exportation des descriptions de flux

Vous pouvez exporter le contenu de la description de flux dans un fichier HTML.

Pour exporter une description de flux :

1. Dans le menu principal, cliquez sur :
   Fichier > **Exporter la description de flux**
2. Saisissez un nom pour le fichier HTML et cliquez sur **Enregistrer**.

Exécution des flux

Après avoir spécifié les options nécessaires pour les flux et connecté les noeuds requis, vous pouvez exécuter le flux en exécutant les données via les noeuds du flux. Il existe plusieurs manières d'exécuter un flux dans IBM SPSS Modeler. Vous pouvez :

- Cliquez sur **Exécuter** dans le menu Outils.
- Cliquez sur un des boutons **Exécuter...** de la barre d'outils. Ces boutons vous permettent d'exécuter la totalité du flux ou simplement le noeud terminal sélectionné. Pour plus d'informations, voir la rubrique «Barre d'outils IBM SPSS Modeler», à la page 18.
- Exécutez un flux de données unique en cliquant sur un noeud terminal à l'aide du bouton droit de la souris et en cliquant sur **Exécuter** dans le menu contextuel.
- Exécutez une partie d'un flux de données en cliquant sur un noeud non terminal à l'aide du bouton droit de la souris et en cliquant sur **Exécuter à partir d'ici** dans le menu contextuel. Cette action permet de n'exécuter que les opérations après le noeud sélectionné.

Pour arrêter l'exécution d'un flux en cours, vous pouvez cliquer sur le bouton d'arrêt rouge dans la barre d'outils ou cliquer sur **Arrêter l'exécution** du menu Outils.

Si un flux prend plus de trois secondes à s'exécuter, la boîte de dialogue Commentaires d'exécution s'affiche pour indiquer la progression.

Certains noeuds disposent d'affichages supplémentaires fournissant davantage d'informations sur l'exécution du flux. Vous pouvez les afficher en sélectionnant la ligne correspondante dans la boîte de dialogue. La première ligne est automatiquement sélectionnée.

Utilisation des modèles

Si un flux comprend un noeud de modélisation (c'est à dire, d'un onglet Modélisation ou Base de données d'une palette de noeuds), un **nugget de modèle** est créé lors de l'exécution du flux. Un nugget de modèle est un conteneur pour un **modèle**, c'est-à-dire l'ensemble des règles, des formules ou des équations qui vous permettent de générer des prédictions concernant des données source et qui se trouve au coeur de l'analyse prédictive.

Figure 12. Nugget de modèle

Lorsque vous êtes parvenu à exécuter un noeud de modélisation, un nugget de modèle correspondant est placé dans l'espace de travail du flux, où il est représenté par une icône en forme de diamant doré (d'où le nom « nugget » (pépite en français)). Vous pouvez ouvrir le nugget et parcourir son contenu pour afficher les détails du modèle. Pour afficher les prédictions, vous attachez et exécutez un ou plusieurs noeuds terminaux, le résultat à partir duquel les prédictions sont présentées sous une forme lisible.

Un flux de modélisation standard se compose de deux branches. La **branche de modélisation** contient le noeud de modélisation ainsi que les noeuds source et d'exécution qui le précèdent. La **branche de scoring** est créée lorsque vous exécutez le noeud de modélisation et qu'il contient le nugget de modèle et le noeud terminal ou les noeuds que vous utilisez pour afficher les prédictions.

Pour plus d'informations, voir le guide des *noeuds de modélisation d'IBM SPSS Modeler*.

Ajout de commentaires et d'annotations à des noeuds et à des flux

Il se peut que vous ayez à décrire un flux à d'autres personnes de votre entreprise. Pour vous aider, vous pouvez joindre des explications aux flux, noeuds et nuggets de modèle.

Les autres peuvent ensuite afficher ces commentaires à l'écran et vous pouvez imprimer une image du flux qui inclut ces commentaires.

Vous pouvez lister tous les commentaires d'un flux ou d'un super noeud, modifier l'ordre des commentaires dans la liste, modifier le texte du commentaire et modifier la couleur du premier plan ou de l'arrière-plan d'un commentaire. Pour plus d'informations, voir la rubrique «Liste des commentaires sur un flux», à la page 63.

Vous pouvez également ajouter des remarques sous la forme d'annotations de texte aux flux, noeuds et nuggets au moyen de l'onglet Annotations d'une boîte de dialogue Propriétés du flux, d'une boîte de dialogue Noeud ou d'une fenêtre de nugget de modèle. Ces remarques sont visibles uniquement lorsque l'onglet Annotations est ouvert, mais les annotations de flux peuvent être affichées sous forme de commentaires à l'écran. Pour plus d'informations, voir la rubrique «Annotations», à la page 64.

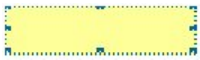
Commentaires

Les commentaires prennent la forme de zones de texte dans lesquelles vous pouvez saisir n'importe quelle taille de texte et ajouter tous les commentaires souhaités. Un commentaire peut être autonome (non attaché à un objet de flux) ou il peut être connecté à un ou plusieurs noeuds ou nuggets de modèle dans le flux. Les commentaires autonomes sont généralement utilisés pour décrire l'objectif général du flux ; les commentaires connectés décrivent le noeud ou le nugget auquel ils sont attachés. Les noeuds et les nuggets peuvent avoir plusieurs commentaires attachés, et le flux peut contenir n'importe quel nombre de commentaires autonomes.

Remarque : vous pouvez également afficher les annotations de flux sous forme de commentaires à l'écran, bien que ceux-ci ne puissent pas être joints aux noeuds ou aux nuggets. Pour plus d'informations, voir la rubrique «Conversion d'annotations en commentaires», à la page 64.

L'apparence de la zone de texte change pour indiquer le mode de commentaire en cours (ou l'annotation sous forme de commentaire), comme l'indique le tableau suivant :

Tableau 3. Modes des zones de texte de commentaires et d'annotations

Zone de texte des commentaires	Zone de texte des annotations	Mode	Indique	Obtenu par...
		Editer	Le commentaire est ouvert pour l'édition.	Création d'un nouveau commentaire ou annotation ou double-clic sur un commentaire ou une annotation existante.
		Dernier sélectionné	Il est possible de déplacer, de redimensionner ou de supprimer un commentaire.	Cliquez sur l'arrière-plan du flux après l'édition ou cliquez une fois sur un commentaire ou une annotation existante.
		Afficher	L'édition est terminée.	Cliquez sur un autre noeud, commentaire ou annotation après l'édition.

Lorsque vous créez un nouveau commentaire autonome, il apparaît initialement en haut à gauche de l'espace de travail du flux.

Si vous attachez un commentaire à un noeud ou un nugget, le commentaire est initialement affiché au-dessus de l'objet de flux auquel il est attaché.

La zone de texte est en blanc pour indiquer que du texte peut être saisi. Une fois que vous avez saisi le texte, cliquez hors de la zone de texte. L'arrière-plan du commentaire devient jaune pour indiquer que la saisie de texte est terminée. Le commentaire reste sélectionné, vous permettant de le déplacer, de le redimensionner ou de le supprimer.

Lorsque vous cliquez de nouveau, la bordure se transforme en lignes pleines pour indiquer que l'édition est terminée.

En double-cliquant sur un commentaire, vous faites passer la zone de texte en mode d'édition : l'arrière-plan devient blanc et le texte du commentaire peut être édité.

Vous pouvez également attacher des commentaires aux super noeuds.

Opérations impliquant des commentaires : Vous pouvez exécuter plusieurs opérations sur les commentaires. Vous pouvez :

- ajouter un commentaire autonome
- attacher un commentaire à un noeud ou un nugget
- éditer un commentaire
- redimensionner un commentaire
- déplacer un commentaire
- déconnecter un commentaire
- supprimer un commentaire
- afficher ou masquer tous les commentaires sur un flux

Pour ajouter un commentaire autonome

1. Assurez-vous que rien n'est sélectionné dans le flux.
2. Effectuez l'une des actions suivantes :
   - Dans le menu principal, cliquez sur :

     Insérer > **Nouveau commentaire**
   - Cliquez avec le bouton droit de la souris sur l'arrière-plan du flux et cliquez sur **Nouveau commentaire** dans le menu contextuel.
   - Cliquez sur le bouton de **Nouveau commentaire** de la barre d'outils.
3. Saisissez le texte du commentaire (ou coller le texte dans le presse-papiers).
4. Cliquez sur un noeud dans le flux pour enregistrer le commentaire.

Pour attacher un commentaire à un noeud ou un nugget

1. Sélectionnez un ou plusieurs noeuds ou nuggets dans l'espace de travail de flux.
2. Effectuez l'une des actions suivantes :
   - Dans le menu principal, cliquez sur :

     Insérer > **Nouveau commentaire**
   - Cliquez avec le bouton droit de la souris sur l'arrière-plan du flux et cliquez sur **Nouveau commentaire** dans le menu contextuel.
   - Cliquez sur le bouton de **Nouveau commentaire** de la barre d'outils.
3. Entrez le texte du commentaire.
4. Cliquez sur un autre noeud dans le flux pour enregistrer le commentaire.

   Vous pouvez également :
5. Insérer un commentaire autonome (voir section précédente).
6. Effectuez l'une des actions suivantes :

- Sélectionnez le commentaire, appuyez sur F2, puis sélectionnez le noeud ou le nugget.
- Sélectionnez le noeud ou le nugget, appuyez sur F2, puis sélectionnez le commentaire.
- (Souris à trois boutons uniquement) Déplacez le pointeur de la souris sur le commentaire, maintenez enfoncé le bouton central, faites glisser le pointeur de la souris sur le noeud ou le nugget et relâchez le bouton de la souris.

Pour attacher un commentaire à un noeud ou un nugget supplémentaire

Si un commentaire est déjà attaché à un noeud ou un nugget, ou s'il se trouve actuellement au niveau du flux et que vous voulez l'attachez à un noeud ou un nugget supplémentaire, effectuez l'une des actions suivantes :

- Sélectionnez le commentaire, appuyez sur F2, puis sélectionnez le noeud ou le nugget.
- Sélectionnez le noeud ou le nugget, appuyez sur F2, puis sélectionnez le commentaire.
- (Souris à trois boutons uniquement) Déplacez le pointeur de la souris sur le commentaire, maintenez enfoncé le bouton central, faites glisser le pointeur de la souris sur le noeud ou le nugget et relâchez le bouton de la souris.

Pour modifier un commentaire existant

1. Effectuez l'une des actions suivantes :
   - Double-cliquez sur la zone de texte du commentaire.
   - Sélectionnez la zone de texte et appuyez sur Entrée.
   - Faites un clic droit sur la zone de texte pour afficher son menu et choisissez Modifier.
2. Modifiez le texte du commentaire. Vous pouvez utiliser les touches de raccourci standard de Windows lors de l'édition, par exemple, Ctrl+C pour copier du texte. Les autres options existantes pendant la modification sont répertoriées dans le menu contextuel du commentaire.
3. Cliquez une fois hors de la zone de texte pour afficher les commandes de redimensionnement, puis à nouveau pour terminer le commentaire.

Pour redimensionner une zone de texte de commentaire

1. Sélectionnez le commentaire pour afficher les commandes de redimensionnement.
2. Cliquez sur une commande et faites-la glisser pour redimensionner la zone.
3. Cliquez hors de la zone de texte pour enregistrer la modification.

Pour déplacer un commentaire existant

Si vous souhaitez déplacer un commentaire mais pas ses objets attachés (le cas échéant), effectuez l'une des actions suivantes :

- Déplacez le pointeur de la souris sur le commentaire, maintenez enfoncé le bouton gauche de la souris et faites glisser le commentaire vers sa nouvelle position.
- Sélectionnez le commentaire, maintenez enfoncée la touche Alt et déplacez le commentaire à l'aide des flèches.

Si vous souhaitez déplacer un commentaire avec des noeuds ou des nuggets auxquels le commentaire est attaché :

1. Sélectionnez tous les objets à déplacer.
2. Effectuez l'une des actions suivantes :

- Déplacez le pointeur de la souris sur l'un des objets, maintenez enfoncé le bouton gauche de la souris, et faites glisser les objets vers leur nouvelle position.
- Sélectionnez un des objets, maintenez enfoncée la touche Alt, et déplacez les objets à l'aide des flèches.

Pour déconnecter un commentaire d'un noeud ou d'un nugget

1. Sélectionnez au moins un commentaire à déconnecter.
2. Effectuez l'une des actions suivantes :

- Appuyez sur la touche F3.
- Faites un clic droit sur un commentaire sélectionné et cliquez sur Déconnecter sur son menu.

Pour supprimer un commentaire

1. Sélectionnez au moins un commentaire à supprimer.
2. Effectuez l'une des actions suivantes :

- Appuyez sur la touche Supprimer.
- Faites un clic droit sur un commentaire sélectionné et cliquez sur Supprimer sur son menu.

Si le commentaire était attaché à un noeud ou un nugget, la ligne de liaison est supprimée également.

Si le commentaire était à l'origine un flux ou une annotation de super noeud ayant été converti en commentaire indépendant, ce commentaire est supprimé de l'espace de travail mais son texte est conservé dans l'onglet Annotations du flux ou du super noeud.

Pour afficher ou masquer des commentaires sur un flux

1. Effectuez l'une des actions suivantes :

- Dans le menu principal, cliquez sur :
  Vue > **Commentaires**
- Cliquez sur le bouton **Afficher/Masquer les commentaires** de la barre d'outils.

Liste des commentaires sur un flux : Vous pouvez afficher une liste de tous les commentaires qui ont été faits pour un flux ou un super noeud particulier.

Dans la liste, vous pouvez

- Modifier l'ordre des commentaires
- Modifier le texte du commentaire
- Modifier la couleur du premier plan ou de l'arrière-plan d'un commentaire

Liste de commentaires

Pour lister des commentaires effectués sur un flux, procédez de l'une des manières suivantes :

- Dans le menu principal, cliquez sur :
  Outils > **Propriétés de flux** > **Commentaires**
- Cliquez avec le bouton droit de la souris dans le panneau des gestionnaires et cliquez sur **Propriétés du flux**, puis sur**Commentaires**.
- Cliquez avec le bouton droit de la souris sur l'arrière-plan d'un flux de l'espace de travail et cliquez sur **Propriétés du flux**, puis sur **Commentaires**.

Texte. Le texte du commentaire. Double-cliquez sur le texte pour modifier le champ en une zone de texte modifiable.

Liens. Le nom du noeud auquel le commentaire est attaché. Si ce champ est vide, le commentaire s'applique au flux.

Boutons de positionnement. Ils permettent de déplacer un commentaire sélectionné vers le haut ou vers le bas dans la liste.

Couleurs de commentaires. Pour modifier la couleur de premier plan ou d'arrière-plan d'un commentaire, sélectionnez le commentaire, cochez la case **Couleurs personnalisées**, puis sélectionnez une couleur dans la liste **Arrière-plan** ou **Premier plan** (ou les deux). Cliquez sur **Appliquer**, puis cliquez sur l'arrière-plan du flux pour visualiser l'effet de la modification. Cliquez sur **OK** pour enregistrer cette modification.

Conversion d'annotations en commentaires : Il est possible de convertir des annotations faites sur des flux ou des super noeuds en commentaires.

Dans le cas de flux, l'annotation est convertie en un commentaire autonome (c'est-à-dire qu'il n'est joint à aucun noeud) dans l'espace de travail du flux.

Lorsqu'une annotation de super noeud est convertie en un commentaire, celui-ci n'est pas joint au super noeud dans l'espace de travail du flux, mais il est visible lorsque vous effectuez un zoom avant dans le super noeud.

Pour convertir une annotation de flux en un commentaire

1. Cliquez sur **Propriétés du flux** dans le menu Outils. (Vous pouvez également cliquer avec le bouton droit de la souris sur un flux dans le panneau des gestionnaires et cliquer sur **Propriétés du flux**.)
2. Cliquez sur l'onglet **Annotations**.
3. Cochez la case **Afficher une annotation comme un commentaire**.
4. Cliquez sur **OK**.

Pour convertir une annotation de super noeud en un commentaire

1. Double-cliquez sur l'icône du super noeud dans l'espace de travail.
2. Cliquez sur l'onglet **Annotations**.
3. Cochez la case **Afficher une annotation comme un commentaire**.
4. Cliquez sur **OK**.

Annotations

Il est possible d'annoter des noeuds, des flux et des modèles de différentes manières. Vous pouvez ajouter des annotations descriptives et spécifier un nom personnalisé. Ces options sont utiles plus particulièrement lorsque vous générez des rapports sur les flux ajoutés au panneau de projet. Pour les noeuds et les nuggets de modèle, vous pouvez également ajouter une info-bulle afin de pouvoir faire la différence entre les noeuds similaires dans l'espace de travail de flux.

Ajout d'annotations

Lorsque vous éditez un noeud ou un nugget de modèle, une boîte de dialogue à onglets apparaît ; elle contient l'onglet Annotations qui permet de définir différentes options d'annotation. Vous pouvez également ouvrir directement l'onglet Annotations.

1. Pour annoter un noeud ou un nugget, cliquez dessus avec le bouton droit de la souris dans l'espace de travail de flux, puis cliquez sur **Renommer et annoter**. La boîte de dialogue d'édition s'ouvre et l'onglet Annotations s'affiche.
2. Pour annoter un flux, cliquez sur **Propriétés du flux** dans le menu Outils. (Vous pouvez également cliquer avec le bouton droit de la souris sur un flux dans le panneau des gestionnaires et cliquer sur **Propriétés du flux**.) Cliquez sur l'onglet Annotations.

Nom. Sélectionnez **Personnalisé** pour modifier le nom généré automatiquement ou pour créer un nom unique pour le noeud selon son affichage dans l'espace de travail.

Texte de l'info-bulle. (Uniquement pour les noeuds et les nuggets de modèle) Saisissez le texte pour l'info-bulle dans l'espace de travail du flux. Cette option s'avère particulièrement utile lorsque vous travaillez avec plusieurs noeuds similaires.

Mots-clés. Permet d'indiquer les mots-clés à utiliser dans les rapports de projet, ou lors de la recherche de noeuds dans un flux ou du suivi des objets stockés dans le référentiel (reportez-vous à «A propos du IBM SPSS Collaboration and Deployment Services Repository», à la page 191). Vous pouvez séparer les mots-clés par des points-virgules, comme l'illustre l'exemple suivant : `revenu; type de culture; valeur de la demande`. Les espaces utilisés au début et à la fin de chaque mot-clé sont supprimés (par exemple,

`revenu ; type de culture` génère les mêmes résultats que `revenu;type de culture`. (Toutefois, les espaces utilisés à l'intérieur des mots-clés ne sont pas supprimés. Par exemple, l'occurrence `type de culture` dotée d'un espace et l'occurrence `type de culture` incluant deux espaces sont différentes.)

La zone de texte principale peut être utilisée pour entrer des annotations longues concernant les opérations ou les décisions relatives au noeud. Par exemple, lorsque vous partagez et réutilisez des flux, il est utile de garder une trace des décisions, telles que l'utilisation d'un noeud Filtrer pour supprimer un champ comportant de nombreux blancs. Les informations saisies lors de l'annotation sont stockées avec le noeud. Vous pouvez également inclure ces annotations dans un rapport de projet créé à partir du panneau de projet. Pour plus d'informations, voir la rubrique «Introduction aux projets», à la page 215.

Afficher l'annotation comme commentaire. (Uniquement pour des annotations sur un flux ou un super noeud) Cochez cette case pour convertir l'annotation en un commentaire autonome qui sera visible sur l'espace de travail du flux. Pour plus d'informations, voir la rubrique «Ajout de commentaires et d'annotations à des noeuds et à des flux», à la page 59.

ID. Affiche un ID unique qui peut être utilisé pour référencer le noeud à des fins de génération de scripts ou d'automatisation. Cette valeur est générée automatiquement lorsque le noeud est créé et elle ne sera pas modifiée. Notez également que, pour éviter toute confusion avec la lettre "O", les zéros ne sont pas utilisés dans les ID de noeud. Utilisez le bouton Copier, situé à droite, pour copier et coller l'ID dans des scripts ou ailleurs, selon les besoins.

Enregistrement des flux de données

Après avoir créé un flux, vous pouvez l'enregistrer pour un usage ultérieur.

Pour enregistrer un flux

1. Dans le menu Fichier, cliquez sur **Enregistrer le flux** ou sur **Enregistrer le flux sous**.
2. Dans la boîte de dialogue Enregistrer, sélectionnez le répertoire dans lequel enregistrer le fichier de flux.
3. Entrez le nom du flux dans la zone de texte Fichier.
4. Sélectionnez **Ajouter au projet** pour ajouter le flux enregistré au projet en cours.

Cliquez sur **Enregistrer** pour stocker le flux avec l'extension `*.str` dans le répertoire spécifié.

Fichiers de sauvegarde automatique. Lorsqu'un flux est enregistré, la version précédemment enregistrée du fichier est conservée automatiquement en tant que copie de sauvegarde. Un tiret est ajouté au nom du fichier (par exemple, `monflux.str-`). Pour restaurer cette version sauvegardée, supprimez simplement le tiret et rouvrez le fichier.

Enregistrement des états

Outre les flux, vous pouvez enregistrer les **états** qui incluent le diagramme de flux affiché et tout nugget de modèle que vous avez créé (répertorié dans l'onglet Modèles du panneau des gestionnaires).

Pour enregistrer un état

1. Dans le menu Fichier, cliquez sur :

   Etat > **Enregistrer l'état ou Enregistrer l'état sous**
2. Dans la boîte de dialogue Enregistrer, sélectionnez le répertoire dans lequel enregistrer le fichier d'état.

Cliquez sur **Enregistrer** pour stocker l'état avec l'extension **.cst* dans le répertoire spécifié.

Enregistrement de noeuds

Vous pouvez également enregistrer un noeud précis en cliquant dessus avec le bouton droit de la souris dans l'espace de travail de flux et en cliquant sur **Enregistrer le noeud** dans le menu contextuel. Utilisez l'extension de fichier **.nod*.

Enregistrement de plusieurs objets de flux

Si vous fermez IBM SPSS Modeler alors que plusieurs objets n'ont pas été enregistrés (comme des flux, des projets ou des nuggets de modèles), le système vous invite à les enregistrer avant que vous quittiez définitivement le logiciel. Si vous choisissez d'enregistrer les éléments, une boîte de dialogue affiche les options permettant d'enregistrer chaque objet.

1. Cochez les cases correspondant aux objets à enregistrer.
2. Cliquez sur **OK** pour enregistrer chaque objet à l'emplacement requis.

Une boîte de dialogue Enregistrer standard apparaît alors pour chaque objet. Une fois que vous avez terminé l'enregistrement, l'application se ferme.

Enregistrement des résultats

Les tableaux, graphiques et rapports générés depuis les noeuds de sortie IBM SPSS Modeler peuvent être enregistrés au format d'objet de sortie (**.cou*).

1. Lorsque vous affichez la sortie à enregistrer, dans les menus de la fenêtre de sortie, cliquez sur : **Fichier** > **Enregistrer**
2. Indiquez les nom et emplacement du fichier de sortie.
3. Si vous le souhaitez, sélectionnez **Ajouter le fichier au projet** dans la boîte de dialogue d'enregistrement pour inclure le fichier dans le projet en cours. Pour plus d'informations, voir la rubrique «Introduction aux projets», à la page 215.

Vous pouvez également cliquer avec le bouton droit de la souris sur un objet de sortie répertorié dans le panneau des gestionnaires et sélectionner **Enregistrer** dans le menu contextuel.

Codage et décodage des informations

Lorsque vous enregistrez un flux, un noeud, un projet, un fichier de sortie ou un nugget de modèle, vous pouvez le coder afin d'empêcher toute personne non autorisée de l'utiliser. Pour ce faire, sélectionnez une option supplémentaire à l'enregistrement et ajoutez un mot de passe à l'élément enregistré. Ce codage peut être paramétré pour n'importe quel élément que vous enregistrez et lui ajoute une sécurité supplémentaire ; ce type de codage est différent du codage SSL utilisé si vous transmettez des fichiers entre IBM SPSS Modeler et IBM SPSS Modeler Server.

Lorsque vous essayez d'ouvrir un élément codé, une invite vous demande d'entrer le mot de passe. Une fois que vous avez saisi le mot de passe approprié, l'élément est automatiquement décodé et s'ouvre comme d'habitude.

Pour coder un élément

1. Dans la boîte de dialogue Enregistrer de l'élément à coder, cliquez sur **Options**. La boîte de dialogue Options de codage apparaît.
2. Sélectionnez **Coder ce fichier**.
3. Pour plus de sécurité, vous pouvez également sélectionner **Masquer le mot de passe**. Tout ce que vous saisissez apparaît alors sous la forme d'une série de points.
4. Entrez le mot de passe. *Avertissement* : si vous oubliez le mot de passe, le fichier ou le modèle ne peut pas être ouvert.
5. Si vous avez sélectionné **Masquer le mot de passe**, saisissez de nouveau le mot de passe pour confirmer que vous l'avez correctement entré.
6. Cliquez sur **OK** pour revenir à la boîte de dialogue Enregistrer.

Remarque : si vous enregistrez une copie d'un élément protégé par un code, le nouvel élément est automatiquement enregistré dans un format codé en utilisant le mot de passe d'origine, sauf si vous modifiez les paramètres de la boîte de dialogue Options de codage.

Chargement des fichiers

Vous pouvez recharger un certain nombre d'objets enregistrés dans IBM SPSS Modeler :

- Flux (.str)
- Etats (.cst)
- Modèles (.gm)
- Palette de modèles (.gen)
- Noeuds (.nod)
- Sortie (.cou)
- Projets (.cpj)

Ouverture de nouveaux fichiers

Il est possible de charger les flux directement à partir du menu Fichier.

- Dans le menu Fichier, cliquez sur **Ouvrir un flux**.

Tous les autres types de fichier peuvent être ouverts via les sous-menus du menu Fichier. Par exemple, pour charger un modèle, dans le menu Fichier, cliquez sur :

Modèles > **Ouvrir un modèle ou Charger la palette Modèles**

Ouverture de fichiers récemment utilisés

Pour charger rapidement les derniers fichiers utilisés, utilisez les options figurant en bas du menu Fichier.

Sélectionnez **Flux récents**, **Projets récents** ou **Etats récents** pour développer la liste des derniers fichiers utilisés.

Mappage de flux de données

L'outil de mappage permet de connecter une nouvelle source de données à un flux existant. Cet outil configure la connexion et vous aide à définir la manière dont les noms des champs de la nouvelle source remplaceront ceux du flux existant. Au lieu de recréer intégralement un flux de données pour une nouvelle source de données, vous pouvez simplement établir une connexion avec un flux existant.

L'outil de mappage des données vous permet de relier deux fragments de flux et de vous assurer que le nom de tous les champs (essentiels) correspondent parfaitement. En résumé, le mappage des données équivaut simplement à créer un noeud Filtrer qui fait correspondre les champs appropriés en leur donnant un nouveau nom.

Il existe deux méthodes équivalentes pour mapper les données :

Sélectionner le noeud de remplacement. Cette méthode débute par le noeud à remplacer. D'abord, faites un clic droit sur le noeud à remplacer ; puis, avec l'option **Mappage de données** > **Sélectionner le noeud de remplacement** dans le menu contextuel, sélectionnez le noeud avec lequel le remplacer.

Mapper sur. Cette méthode débute par le noeud à introduire dans le flux. D'abord, faites un clic droit sur le noeud à introduire ; puis, avec l'option **Mappage de données** > **Mapper sur** dans le menu contextuel, sélectionnez le noeud auquel le joindre. Cette méthode est particulièrement utile pour le mappage sur un noeud terminal. *Remarque* : il est impossible d'effectuer des mappages sur des noeuds Fusionner ou Ajouter. En revanche, vous pouvez utiliser la méthode de connexion standard pour relier le flux au noeud Fusionner.

Le mappage des données est étroitement lié à la création de flux. Si vous essayez de vous connecter à un noeud qui comporte déjà une connexion, vous aurez la possibilité de remplacer la connexion existante ou d'effectuer le mappage sur ce noeud.

Mappage des données sur un modèle

Pour remplacer la source de données d'un modèle de flux par un nouveau noeud source en apportant vos propres données dans IBM SPSS Modeler, vous devez utiliser l'option **Sélectionner le noeud de remplacement** du menu contextuel Mappage des données. Cette option est disponible pour tous les noeuds à l'exception de Fusionner, Agréger et de tous les noeuds terminaux. L'utilisation de l'outil de mappage de données pour effectuer cette action permet de vous assurer que les champs correspondent parfaitement entre les opérations du flux existantes et la nouvelle source de données. La procédure suivante vous fournira une présentation du processus de mappage de données.

Etape 1 : Spécification des champs essentiels dans le noeud source d'origine. Pour que les opérations de flux soient correctement exécutées, vous devez spécifier des champs essentiels. Pour plus d'informations, voir la rubrique «Spécification de champs essentiels», à la page 69.

Etape 2 : Ajout d'une nouvelle source de données à l'espace de travail de flux. En vous servant de l'un des noeuds source, introduisez les nouvelles données de remplacement.

Etape 3 : Remplacement du noeud source du modèle. A l'aide de l'option Mappage des données du menu contextuel du noeud source du modèle, cliquez sur **Sélectionner le noeud de remplacement**, puis sélectionnez le noeud source pour les données de remplacement.

Etape 4 : Vérification des champs mappés. Dans la boîte de dialogue qui apparaît, vérifiez que le logiciel mappe correctement les champs de la source de données de remplacement sur le flux. Tous les champs essentiels non mappés sont affichés en rouge. Ces champs sont utilisés dans les opérations de flux et doivent être remplacés par un champ similaire de la nouvelle source de données afin que les opérations en aval fonctionnent correctement. Pour plus d'informations, voir la rubrique «Vérification des champs mappés», à la page 69.

Une fois que vous avez utilisé la boîte de dialogue pour vous assurer que tous les champs essentiels étaient correctement mappés, l'ancienne source de données est déconnectée. La nouvelle source de données est alors connectée au flux grâce à un noeud Filtrer appelé *Mapper*. Ce noeud Filtrer dirige le mappage réel des champs du flux. Un noeud Filtrer appelé *Démapper* est également inclus dans l'espace de travail de flux. Le noeud Filtrer appelé *Démapper* peut être utilisé pour annuler le mappage des noms de champs si vous l'ajoutez au flux. Le mappage sera annulé, mais vous devrez éditer les noeuds terminaux en aval pour resélectionner les champs et les superpositions.

Mappage entre flux

Comme pour la connexion de noeuds, cette méthode de mappage des données n'exige pas que vous définissiez des champs essentiels au préalable. Avec cette méthode, vous pouvez simplement relier un flux à l'autre via l'option **Mapper sur** du menu contextuel Mappage des données. Ce type de mappage de données est utile pour le mappage à des noeuds terminaux, ainsi que pour la copie et le collage entre flux. *Remarque* : lorsque vous utilisez l'option **Mapper sur**, vous ne pouvez pas effectuer de mappage vers les noeuds Fusionner et Ajouter, et tous les types de noeud source.

Pour mapper des données entre des flux

1. Cliquez avec le bouton droit de la souris sur le noeud que vous souhaitez utiliser pour la connexion au nouveau flux.
2. Dans le menu, cliquez sur :

   Mappage des données > **Mapper sur**
3. Utilisez le curseur pour sélectionner un noeud de destination dans le flux cible.
4. Dans la boîte de dialogue qui apparaît, assurez-vous que les champs correspondent parfaitement, puis cliquez sur **OK**.

Spécification de champs essentiels

Lorsque vous effectuez un mappage sur un flux existant, les champs essentiels sont spécifiés par l'auteur du flux. Ces champs essentiels indiquent si un champ particulier est utilisé dans les opérations en aval. Par exemple, le flux existant peut créer un modèle qui utilise un champ appelé *Attrition*. Dans ce flux, *Attrition* est un champ essentiel requis pour la création du modèle. De même, les champs utilisés dans les noeuds de manipulation, tels que les noeuds Calculer, sont requis pour le calcul du nouveau champ. En définissant explicitement ces champs comme obligatoires, vous avez la certitude qu'ils seront mappés sur les champs appropriés contenus dans la nouvelle source de données. Si des champs obligatoires ne sont pas mappés, un message d'erreur est envoyé. Si certains noeuds de manipulation ou de sortie vous semblent inutiles, vous pouvez les supprimer du flux et retirer les champs correspondants de la liste des champs obligatoires.

Pour définir des champs essentiels

1. Cliquez avec le bouton droit de la souris sur le noeud source du modèle de flux à remplacer.
2. Dans le menu, cliquez sur :

   Mappage des données > **Spécifier les champs essentiels**
3. A l'aide du sélecteur de champs, vous pouvez ajouter des champs à la liste ou en supprimer. Pour ouvrir le sélecteur de champs, cliquez sur l'icône située à droite de la liste des champs.

Vérification des champs mappés

Une fois que vous avez choisi le point à partir duquel un flux (ou une source) de données sera mappé sur un autre flux (une autre source), une boîte de dialogue apparaît. Elle vous permet de sélectionner les champs à mapper ou de vous assurer que le mappage par défaut du système est correct. Si des champs essentiels ont été définis pour le flux ou la source de données et qu'ils ne présentent aucune correspondance, ils sont affichés en rouge. Tous les champs non mappés de la source de données seront inclus tels quels dans le noeud Filtrer, mais vous pouvez également mapper des champs non essentiels.

Original. Répertorie tous les champs du modèle ou du flux existant (tous les champs présents plus en aval). Les champs provenant de la nouvelle source de données seront mappés sur les champs existants.

Mappé. Affiche la liste des champs sélectionnés à mapper sur les champs du modèle. Il s'agit des champs dont les noms doivent être modifiés pour correspondre aux champs d'origine utilisés dans les opérations de flux. Cliquez sur la cellule du tableau correspondant à un champ pour activer la liste des champs disponibles.

Avant toute opération de mappage, il est conseillé d'étudier la source de données avec attention, afin d'identifier précisément les champs à mapper. Par exemple, vous pouvez utiliser l'onglet Types du noeud source pour consulter le récapitulatif des données source.

Astuces et raccourcis

Pour une utilisation rapide et simple, familiarisez-vous avec les astuces et raccourcis suivants :

- **Créez des flux rapidement à l'aide d'un double-clic.** Double-cliquez simplement sur un noeud de la palette pour l'ajouter et le connecter au flux en cours.
- **Utilisez des combinaisons de touches pour sélectionner des noeuds en aval.** Appuyez sur Ctrl+Q et sur Ctrl+W pour activer ou désactiver la sélection de tous les noeuds en aval.
- **Utilisez des touches de raccourci pour connecter et déconnecter les noeuds.** Lorsqu'un noeud est sélectionné dans l'espace de travail, appuyez sur la touche F2 pour activer une connexion, sur la touche Tab pour accéder au noeud souhaité et sur la touche Maj+Espace pour mettre un terme à la connexion. Appuyez sur la touche F3 pour déconnecter toutes les entrées et sorties du noeud sélectionné.
- **Personnalisez l'onglet Palette de noeuds avec vos noeuds favoris.** Dans le menu Outils, cliquez sur **Gérer les palettes** pour ouvrir une boîte de dialogue vous permettant d'ajouter, de supprimer ou de déplacer des noeuds affichés dans la Palette de noeuds.

- **Renommez les noeuds et ajoutez des info-bulles.** Chaque boîte de dialogue de noeud contient un onglet Annotations. Il vous permet de spécifier un nom personnalisé pour chaque noeud de l'espace de travail et d'ajouter des info-bulles pour l'organisation de votre flux. Vous pouvez également inclure des annotations longues pour suivre la progression, enregistrer les détails du processus, et indiquer toute décision commerciale prise ou à prendre.
- **Insérez des valeurs automatiquement dans une expression CLEM.** Le Générateur de formules, accessible à partir d'un grand nombre de boîtes de dialogue (telles que celles des noeuds Calculer et Remplacer), vous permet d'insérer automatiquement des valeurs de champ dans une expression CLEM. Dans le Générateur de formules, cliquez sur le bouton des valeurs pour sélectionner les valeurs de champ.

Figure 13. Bouton Valeurs

- **Accès rapide aux fichiers.** Lorsque vous recherchez des fichiers dans une boîte de dialogue Ouvrir, utilisez la liste Fichier (bouton représentant un diamant jaune dans la partie supérieur de la boîte de dialogue, près du champ Chercher dans) pour accéder aux répertoires précédemment utilisés et aux répertoires IBM SPSS Modeler par défaut. Utilisez les boutons Suivant et Précédent pour parcourir les répertoires utilisés.
- **Réduisez l'encombrement de la fenêtre de sortie.** Vous pouvez fermer et supprimer rapidement une sortie à l'aide du bouton *X* rouge situé dans le coin supérieur droit de toutes les fenêtres de sortie. Vous pouvez ainsi conserver uniquement les résultats prometteurs ou intéressants dans l'onglet Sorties du panneau des gestionnaires.

Un ensemble complet de raccourcis clavier est disponible pour le logiciel. Pour plus d'informations, voir la rubrique «Accessibilité via le clavier», à la page 244.

Savez-vous que vous pouvez...

- faire glisser et sélectionner un groupe de noeuds sur l'espace de travail à l'aide de la souris.
- copier et coller des noeuds d'un flux à l'autre.
- accéder à l'aide à partir de chaque boîte de dialogue et fenêtre de sortie.
- obtenir de l'aide sur le processus CRISP-DM (Cross-Industry Standard Process for Data Mining). (Dans le menu Aide, cliquez sur **Aide CRISP-DM**.)

Chapitre 6. Travail avec les données

Vous pouvez examiner les données d'un flux en cliquant avec le bouton droit de la souris sur un noeud de données et en sélectionnant **Afficher les données**. Dans la fenêtre qui s'affiche :

- L'onglet **Graphique** permet de créer des visualisations de données avancées pour explorer vos données à partir de différentes perspectives et identifier des modèles, des connexions et des relations dans vos données.
- L'onglet **Feuille de calcul** affiche une vue en lecture seule de vos données sous forme de table.
- L'onglet **Audit des données** affiche la fréquence et les statistiques de chaque colonne.
- Dans l'onglet **Tableau de bord**, vous pouvez cliquer sur l'option permettant de **démarrer la conception de la présentation** et créer une présentation pour afficher plusieurs graphiques sur une seule page. Vous pouvez enregistrer la présentation comme modèle, puis faire glisser les graphiques enregistrés et les déposer aux positions voulues sur la présentation.
- L'onglet **Préférences** permet de définir les préférences de l'interface utilisateur, comme la langue et l'aspect général.

Notez que cette fonction utilise le port 28900 par défaut. Si vous devez utiliser un autre port, modifiez la valeur du paramètre de configuration `data_view_port_number` dans le fichier `options.cfg`.

Construction de graphiques

Dans l'onglet **Graphiques**, vous pouvez créer des graphiques à partir de graphiques prédéfinis de la galerie ou à partir de composants spécifiques (par exemple, des axes et des barres). Pour construire un graphique, vous devez sélectionner un type de graphique dans la galerie ou des éléments de base dans les options de type de graphique fournies.

Lors de la création du graphique, le canevas en affiche l'aperçu. L'aperçu utilise les libellés de variable et les niveaux de mesure réels représentatifs de vos données effectives.

Les nouveaux utilisateurs préfèrent utiliser la galerie. Pour obtenir des informations sur l'utilisation de la galerie, reportez-vous à «Construction d'un graphique à partir de la galerie», à la page 72.

Lancement du Générateur de graphiques

- Cliquez avec le bouton droit de la souris sur le noeud de données que vous souhaitez utiliser, puis sélectionnez **Afficher les données**.

Présentation et terminologie

Ecran de bienvenue
: Au démarrage du générateur de graphiques, vous avez la possibilité de sélectionner un type de graphique ou de sélectionner des colonnes d'un jeu de données actif. A mesure que vous ajoutez des colonnes à visualiser, les options de types de graphique sont actualisées de manière à proposer les types de graphique les mieux adaptés aux colonnes sélectionnées.

Canevas
: Zone dans laquelle vous construisez le graphique dans la boîte de dialogue Générateur de graphiques.

Types de diagrammes
: Répertorie les types de graphiques disponibles. Les éléments graphiques sont les éléments qui représentent des données dans le graphique. Il s'agit de barres, points, lignes, etc.

Panneau Détails

Le panneau Détails décrit les blocs de construction élémentaires du graphique.

Paramètres du graphique

Fournit des options permettant de sélectionner les variables utilisées pour générer le graphique, la méthode de distribution, des champs de titre et de sous-titre, etc. Selon le type de graphique sélectionné, les options proposées dans le panneau Détails varient. Pour plus d'informations sur les options disponibles pour chaque type de graphique, voir «Types de graphique».

Actions

Fournit des options de téléchargement de fichiers de configuration de graphique, de téléchargement de graphiques sous forme de fichiers image, de réinitialisation de graphiques et de définition de préférences globales pour les graphiques.

Construction d'un graphique à partir de la galerie

La galerie est la méthode la plus simple de création de graphiques. Les étapes de base permettant de générer un graphique à partir de la galerie sont présentées ci-après.

1. Dans la section **Types de graphique**, sélectionnez une catégorie de graphiques. Un aperçu du type de graphique sélectionné s'affiche sur la grille du diagramme.
2. Si le canevas affiche déjà un graphique, le nouveau graphique remplace les axes et les éléments du graphique.
   a. Selon le type de graphique sélectionné, les variables disponibles sont présentées dans le panneau Détails sous plusieurs en-têtes différents (par exemple, **Catégorie** pour les graphiques à barres, **Axe des X** et **Axe des Y** pour les graphiques à courbes). Sélectionnez les variables appropriées pour le type de graphique sélectionné.

Types de graphique

La galerie contient un ensemble de graphiques les plus couramment utilisés. A savoir les types suivants :

Nuages de points et tracés de points

Nuages de points 1-D, simples, groupés, superposés et 3-D ; tracés de points simples, graphiques 1-D et en lignes de chute. Pour plus d'informations, voir la rubrique «Nuages de points et tracés de points», à la page 74.

Graphiques à courbes

Simples. Pour plus d'informations, voir la rubrique «Graphiques à courbes», à la page 75.

Graphiques de séries multiples

Simples. Pour plus d'informations, voir la rubrique «Graphiques de séries multiples», à la page 76.

Histogrammes

Simples, empilés, polygones de fréquences. Pour plus d'informations, voir la rubrique «Histogrammes», à la page 77.

Graphiques de type pyramide de population

Simples. Pour plus d'informations, voir la rubrique Graphiques de type pyramide de population.

Graphiques Q-Q

Simples. Pour plus d'informations, voir la rubrique «Graphiques Q-Q», à la page 73.

Graphiques circulaires

Simples. Pour plus d'informations, voir la rubrique «Graphiques circulaires», à la page 78.

Graphiques à barres

Simples, empilés et en cluster. Pour plus d'informations, voir la rubrique «Graphiques à barres», à la page 79.

Graphiques de parallèles
Simples. Pour plus d'informations, voir la rubrique «Graphiques de parallèles», à la page 80.

Tableaux de relation
Simples. Pour plus d'informations, voir la rubrique «Tableaux de relation», à la page 81.

Boîtes à moustaches
Simple et en cluster. Pour plus d'informations, voir la rubrique «Boîtes à moustaches», à la page 82.

Graphiques de type carte d'arbre
Simples et en rayons de soleil. Pour plus d'informations, voir la rubrique Graphiques de type carte d'arbre.

Graphiques de type Carte
Simples. Pour plus d'informations, voir la rubrique «Graphiques de type Carte», à la page 84.

Cartes thermiques
Simples. Pour plus d'informations, voir la rubrique «Cartes thermiques», à la page 86.

Graphiques t-SNE
Simples. Pour plus d'informations, voir la rubrique «Graphiques t-SNE», à la page 87.

Nuages de mots
Simples. Pour plus d'informations, voir la rubrique «Graphiques de nuages de mots», à la page 88.

Graphiques de barre d'erreurs
Simples. Pour plus d'informations, voir la rubrique «Graphiques de barre d'erreurs», à la page 89.

Graphiques 3D
Simples. Pour plus d'informations, voir la rubrique «Graphiques 3D», à la page 90.

Graphiques de matrice de nuage de points
Simples. Pour plus d'informations, voir la rubrique «Graphiques de matrice de nuage de points», à la page 91.

Graphiques en chandeliers
Simples. Pour plus d'informations, voir la rubrique «Graphiques en chandeliers», à la page 92.

Graphiques à double axe des Y
Simples. Pour plus d'informations, voir la rubrique Graphiques à double axe des Y.

Graphiques personnalisés
Pour plus d'informations, voir la rubrique «Graphiques personnalisés», à la page 95.

Graphiques Q-Q

Les graphiques Q-Q (quantile-quantile) comparent deux distributions de probabilité en traçant leurs quantiles les uns par rapport aux autres. Un graphique Q–Q permet de comparer les formes des distributions en affichant une vue graphique qui montre les similitudes et différences de propriétés telles qu'emplacement, échelle et asymétrie de deux distributions.

Création d'un graphique Q-Q simple

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Graphique Q-Q**.

   Le canevas est mis à jour en affichant un modèle de graphique Q-Q.
2. Sélectionnez une variable telle que **Axe des X**.

Fonctions supplémentaires

Axe des X
Répertorie les variables du jeu de données qui sont disponibles pour l'axe *x* du graphique.

Distribution
La liste déroulante affiche toutes les méthodes de distribution disponibles.

Type de tracé
Sélectionnez un tracé Q-Q (quantile-quantile) ou un tracé P-P (pourcentage-pourcentage).

Estimer la distribution depuis les données
Lorsque cette fonction est activée, les paramètres de données de la **Distribution** sélectionnée sont automatiquement estimés. Lorsqu'elle est désactivée, les valeurs de distribution **Forme** et **Echelle** s'affichent.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Nuages de points et tracés de points

L'élément graphique de point permet de créer plusieurs catégories étendues de graphiques :

- **Nuage de points.** Ces graphiques sont utiles pour tracer des données à multiples variables. Ils peuvent aider à déterminer les relations potentielles entre des variables d'échelle. Un nuage de points simple utilise un système de coordonnées 2D pour tracer deux variables. Un nuage de points 3D utilise un système de coordonnées 3D pour tracer trois variables. Lorsque vous devez tracer plus de variables, vous pouvez essayer d'utiliser des nuages de points superposés et des matrices de nuage de points (SPLOM). Un nuage de points superposé affiche des paires superposées de variables *x-y*, chaque paire étant identifiée par une couleur ou une forme. Une matrice SPLOM crée une matrice de nuages de points 2D, chaque variable y étant tracée par rapport à l'autre.
- **Tracés de points :** Tout comme les histogrammes, ces graphiques sont utiles pour représenter la distribution d'une variable d'échelle unique. Les données sont regroupées par casiers, mais, au lieu d'une valeur par casier (comme dans le cas d'un effectif), tous les points de chaque casier sont affichés et empilés. Ces graphiques sont parfois appelés "tracés de densité".
- **Tracés de points récapitulatifs :** Ces graphiques sont semblables aux graphiques à barres, à la différence que les points sont affichés à l'endroit où apparaîtrait le sommet des barres. Etant donné que les tracés de points récapitulatifs sont semblables aux graphiques à barres, voir «Graphiques à barres», à la page 79 pour plus d'informations sur leur création.
- **Graphiques à lignes de chute :** Il s'agit d'un type spécial de tracé de points récapitulatifs. Les points sont regroupés et une ligne est tracée pour relier les points de chaque catégorie. Le graphique à lignes de chute est utile pour comparer une statistique sur plusieurs variables catégorielles.

Création d'un nuage de points simple

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Nuage de points**.

   Le canevas est mis à jour en affichant un modèle de graphique en nuage de points.
2. Sélectionnez une variable d'échelle telle que **Axe des X**.
3. Sélectionnez une variable d'échelle telle que **Axe des Y**. Il n'est pas nécessaire de spécifier de statistiques, car le nuage de points affiche généralement des valeurs brutes.

Fonctions supplémentaires

Axe des X
Répertorie les variables du jeu de données qui sont disponibles pour l'axe *x* du graphique.

Axe des Y
Répertorie les variables du jeu de données qui sont disponibles pour l'axe *y* du graphique.

Courbe d'ajustement
Sur une courbe d'ajustement, les points de données sont adaptés à une courbe ne couvrant généralement pas tous les points de données. La courbe d'ajustement fait ressortir la tendance des données. Certaines courbes d'ajustement sont basées sur le principe de régression. D'autres reposent sur la méthode itérative des moindres carrés pondérés. Sélectionnez une option de courbe d'ajustement dans la liste déroulante.

Carte de couleurs
Répertorie les variables de carte de couleurs disponibles. Ces variables utilisent un dégradé de couleurs, basé sur la plage de valeurs de la colonne indiquée, pour leur représentation dans les points de tracé. Les cartes de couleurs sont également dénommées cartes choroplèthes.

Carte de tailles
Répertorie les variables de carte de tailles disponibles. Ces variables utilisent différentes tailles pour leur représentation dans les points de tracé.

Carte de formes
Répertorie les variables de carte de formes disponibles. Ces variables utilisent différentes formes pour leur représentation dans les points de tracé.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Graphiques à courbes

Un graphique à courbes est une série de points de données reliés par des lignes. Un graphique à courbes est très utile pour présenter des courbes de tendance dont les différences sont subtiles ou dont les lignes de données en croisent une autre. Vous pouvez utiliser un graphique à courbes pour récapituler des variables catégorielles, le graphique étant, dans ce cas, similaire à un graphique à barres (voir «Graphiques à barres», à la page 79). Les graphiques à courbes sont également utiles pour les données de séries temporelles.

Création d'un graphique à courbes de série temporelle simple

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Ligne**.

   Le canevas est mis à jour en affichant un modèle de graphique à courbes.
2. Sélectionnez une variable de date telle que **Axe des X**.
3. Sélectionnez une variable d'échelle telle que **Axe des Y**. Il s'agit de la variable dont les valeurs ont été enregistrées dans le temps.

Fonctions supplémentaires

Axe des X
Répertorie les variables du jeu de données qui sont disponibles pour l'axe *x* du graphique.

Axe des Y
Répertorie les variables du jeu de données qui sont disponibles pour l'axe *y* du graphique.

Aire Lorsque cette option est activée, la zone située en dessous de la ligne est remplie avec une autre couleur.

Scinder par
Sélectionnez une variable catégorielle pour créer un tableau de graphiques, avec une cellule pour chaque catégorie dans la variable Scinder par. Comme le regroupement, la fonction Scinder par variables ajoute essentiellement plus de dimensions à votre graphique en affichant des informations pour chaque catégorie de variable. Reportez-vous à la rubrique Ajout de Scinder par variables pour plus d'informations.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Graphiques de séries multiples

Les graphiques de séries multiples sont similaires aux graphique à courbes, sauf que vous pouvez représenter graphiquement plusieurs variables sur l'axe des *y*.

Création d'un graphique de séries multiples simple

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Séries multiples**.

   Le canevas est mis à jour en affichant un modèle de graphique de séries multiples.
2. Sélectionnez une variable telle que **Axe des X**.
3. Sélectionnez au moins deux variables d'échelle telles les variables **Axe des Y**.

Fonctions supplémentaires

Axe des X
Répertorie les variables du jeu de données qui sont disponibles pour l'axe *x* du graphique.

Axe des Y
Répertorie les variables du jeu de données disponibles pour les axes *y* du graphique.

Sélectionnez le type de graphique (en courbe ou à barres) dans la liste déroulante.

Cliquez sur **Ajouter une autre colonne** pour inclure d'autres colonnes dans le graphique.

Normalisation des données
Lorsque cette fonction est activée, elle transforme les données en distribution normale ce qui permet de comparer les données de plusieurs jeux de données ou de plusieurs colonnes. Cette fonction crée des empilements à 100% des nombres et convertit les statistiques en pourcentages.

Lisse Lorsque cette fonction est activée, les lignes droites du graphique ont un aspect plus incurvé.

Afficher le point d'inflexion
Lorsque cette fonction est activée, les points d'inflexion du graphique sont visibles.

Axe des Y secondaire
Répertorie les variables du jeu de données disponibles pour les axes *y* secondaires du graphique.

Sélectionnez le type de graphique (en courbe ou à barres) dans la liste déroulante.

Cliquez sur **Ajouter une autre colonne** pour inclure d'autres colonnes dans le graphique.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Histogrammes

Un histogramme se présente comme un graphique à barres, mais au lieu de comparer des catégories ou de cerner des tendance sur une durée, chaque barre indique comment les données sont réparties dans une seule catégorie. Chaque barre représente une gamme continue de données ou le nombre de fréquences pour un point de données spécifique.

Les histogrammes sont utiles pour représenter la distribution d'une variable d'échelle unique. Les données sont regroupées par casiers et récapitulées à l'aide d'une statistique d'effectif ou de pourcentage. Le polygone d'effectifs est une variante de l'histogramme. Il ressemble à un histogramme standard où la surface se substitue à la barre.

La pyramide de population est une autre variante de l'histogramme. Son nom reflète son utilisation la plus fréquente, à savoir récapituler des données sur une population. Lorsque ce type de graphique est utilisé avec des données de population, il est scindé en fonction de la variable Sexe pour fournir deux histogrammes horizontaux de données d'âge côte à côte. Dans les pays dont la population est jeune, la forme du graphique obtenu ressemble à une pyramide.

Création d'un histogramme

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Histogramme**.

   Le canevas est mis à jour en affichant un modèle de graphique d'histogramme.
2. Sélectionnez une variable d'échelle telle que **Axe des X**.

   Remarque : La statistique d'un histogramme est Histogramme ou Histogramme - Pourcentage. Ces statistiques regroupent les données par casiers et calculent l'effectif de chaque casier.

Fonctions supplémentaires

Axe des X
Répertorie les variables du jeu de données qui sont disponibles pour l'axe *x* du graphique.

Scinder par
Sélectionnez une variable catégorielle pour créer un tableau de graphiques, avec une cellule pour chaque catégorie dans la variable Scinder par. Comme le regroupement, la fonction Scinder par variables ajoute essentiellement plus de dimensions à votre graphique en affichant des informations pour chaque catégorie de variable. Reportez-vous à la rubrique Ajout de Scinder par variables pour plus d'informations.

Afficher la courbe de distribution
Lorsque cette fonction est activée, la courbe d'ajustement de distribution s'affiche sur le graphique.

Largeur de casier
Le curseur contrôle la taille de l'intervalle utilisé pour fractionner les données en groupes.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Graphiques circulaires

Le graphique circulaire sert à comparer des proportions. Par exemple, vous pouvez utiliser un graphique circulaire pour démontrer qu'une plus grande proportion de femmes est intégrée dans une classe spécifique.

Création d'un graphique circulaire simple

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Circulaire**.

   Le canevas est mis à jour en affichant un modèle de graphique circulaire.
2. Sélectionnez une variable catégorielle (nominale ou ordinale) dans la liste **Catégorie**. Les catégories de cette variable déterminent le nombre de tranches du graphique circulaire.
3. Sélectionnez une fonction statistique récapitulative pour l'élément graphique. Pour les graphiques circulaires, vous utilisez généralement une statistique ou une somme basée sur l'effectif. Le résultat de la statistique détermine la taille de chaque tranche.

Fonctions supplémentaires

Catégorie
Sélectionnez une variable catégorielle (nominale ou ordinale) déterminant le nombre de tranches dans le graphique circulaire.

Type de graphique circulaire
Les styles disponibles sont les suivants :

Normal
Les segments du graphique circulaire sont affichés sous forme de tranches ordinaires.

Anneau
Les segments du graphique circulaire sont présentés sous forme d'anneaux. Egalement dénommé graphique en anneau.

Rose Contrairement au graphique circulaire ordinaire, lequel utilise un rayon commun, les tailles des segments du graphique circulaire varient selon leur valeur.

Récapitulatif
Sélectionnez une fonction statistique récapitulative pour l'élément graphique. Pour les graphiques circulaires, vous utilisez généralement une statistique ou une somme basée sur l'effectif. Le résultat de la statistique détermine la taille de chaque tranche.

Il existe deux types de fonction de récapitulatif statistique. Cette distinction est importante vu qu'elle détermine si vous devez spécifier une variable **Valeur**.

- **Fonctions ne nécessitant pas une variable de valeur.** Il s'agit des fonctions qui ne nécessitent pas de variable **Valeur**. Toutes les statistiques d'effectif et de pourcentage se trouvent dans cette catégorie. Ces statistiques sont disponibles lorsqu'aucune variable **Valeur** n'a été définie.
- **Fonctions nécessitant une variable de valeur.** Il s'agit des fonctions qui nécessitent une variable **Valeur**. Par exemple, la fonction *Moyenne* nécessite une variable avec laquelle la moyenne est calculée. Ces statistiques sont disponibles lorsqu'une variable **Valeur** a été définie.

Valeur Ce champ s'affiche lorsqu'une fonction **Récapitulatif** (laquelle requiert une variable d'échelle) a été sélectionnée. Sélectionnez une variable faisant office de variable d'échelle.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Graphiques à barres

Les graphiques à barres sont utiles pour récapituler des variables catégorielles. Par exemple, vous pouvez utiliser un graphique à barres pour représenter le nombre d'hommes et de femmes qui ont participé à une enquête, ou encore le salaire moyen des hommes et des femmes.

Création d'un graphique à barres simples

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Barres**.

   Le canevas est mis à jour en affichant un modèle de graphique à barres.
2. Sélectionnez une variable (nominale ou ordinale) catégorielle telle la variable **Catégorie**. Vous pouvez utiliser une variable d'échelle, mais les résultats ne seront utiles que dans quelques cas spécifiques. Le graphique à barres offre de meilleurs résultats s'il contient un nombre limité de valeurs distinctes. Si vous créez un graphique à barres avec un axe d'échelle **Catégorie**, les barres seront très fines vu que chacune sera tracée avec sa valeur exacte et qu'elles ne pourront pas chevaucher d'autres valeurs continues.
3. Sélectionnez une statistique dans la liste **Récapitulatif**. Le résultat de la statistique détermine la hauteur des barres. Si la statistique recherchée ne figure pas dans la liste **Récapitulatif**, elle peut nécessiter une variable. Sélectionnez une variable dans la liste **Valeur** et vérifiez si à présent la statistique est disponible. Il peut exister d'autres limitations liées au type du graphique. Par exemple, les graphiques de barre d'erreur ne peuvent être calculés que pour des statistiques particulières.

Fonctions supplémentaires

Catégorie
Répertorie les variables du jeu de données qui sont disponibles pour l'axe *x* du graphique.

Ordre basé sur
Sélectionnez une option de tri pour les catégories de la variable.

Nom de catégorie
Utilisez les libellés de catégorie pour trier les catégories de la variable. Voici les libellés qui apparaissent dans le graphique, généralement sous la forme de libellés de graduation ou de légende.

Valeur de la catégorie
Utilisez la valeur stockée dans le jeu de données pour le tri des catégories de la variable. La valeur de la catégorie permet d'identifier cette catégorie dans le jeu de données. Elle est souvent différente de son libellé et n'est pas nécessairement descriptive. Par exemple, la valeur peut être un nombre (par exemple, *1*), alors que le libellé est une description textuelle de la catégorie (par exemple, *Féminin*).

Ordre des catégories
Sélectionnez l'ordre de tri des catégories de variable.

Telles qu'elles
Les catégories de variable sont présentées telles qu'elles figurent dans le jeu de données.

Ordre croissant
Les catégories de variable sont triées en ordre ascendant.

Ordre décroissant
Les catégories de variable sont triées en ordre descendant.

Récapitulatif
Sélectionnez une fonction statistique récapitulative pour l'élément graphique. Le résultat de la statistique détermine la position des éléments graphiques sur l'axe des *Y*. Dans un graphique 2D, la statistique est calculée pour chaque valeur figurant sur l'axe des *X*. Dans un graphique 3D, elle est calculée pour l'intersection des valeurs figurant sur les axes des *X* et des *Z*.

Il existe deux types de fonction de récapitulatif statistique. La distinction est importante car elle détermine si vous devez spécifier une variable **Valeur** comme axe *y*.

- **Fonctions ne nécessitant pas une variable de valeur.** Il s'agit des fonctions qui ne nécessitent pas de variable d'axe des *Y*. Toutes les statistiques d'effectif et de pourcentage se trouvent dans cette catégorie. Ces statistiques sont disponibles lorsqu'aucune variable **Valeur** n'a été définie.
- **Fonctions nécessitant une variable de valeur.** Il s'agit des fonctions qui nécessitent une variable **Valeur** pour l'axe des *y*. Par exemple, la fonction *Moyenne* nécessite une variable avec laquelle la moyenne est calculée. Ces statistiques sont disponibles lorsqu'une variable **Valeur** a été définie.

Valeur Ce champ s'affiche lorsqu'une fonction **Récapitulatif**, laquelle requiert une variable pour l'axe des *y*, est sélectionnée. Sélectionnez une variable comme valeur pour l'axe des *y*.

Scinder par
Sélectionnez une variable catégorielle pour créer un tableau de graphiques, avec une cellule pour chaque catégorie dans la variable Scinder par. Comme le regroupement, la fonction Scinder par variables ajoute essentiellement plus de dimensions à votre graphique en affichant des informations pour chaque catégorie de variable. Reportez-vous à la rubrique Ajout de Scinder par variables pour plus d'informations.

Type de scission
Lorsqu'une variable **Scinder par** est sélectionnée, vous pouvez choisir d'afficher les barres de la catégorie résultante en les empilant ou bien en cluster. La mise en cluster et l'empilement ajoutent la dimensionnalité au graphique. La mise en cluster scinde une barre en plusieurs barres, tandis que l'empilement crée des segments dans chaque barre. Veillez à choisir la statistique adaptée à l'empilement. Lorsque les valeurs sont ajoutées les unes aux autres (empilées), le résultat obtenu doit être cohérent. Par exemple, ajouter et empiler des valeurs moyennes ne produit généralement pas de résultats significatifs.

Transposition
Lorsque cette fonction est activée, les axes *x* et *y* du graphique sont transposés.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Graphiques de parallèles

Les graphiques de parallèles sont utiles pour visualisation de géométrie avec dimensions nombreuses et lors de l'analyse de données multivariées. Les graphiques de parallèles ressemblent aux graphiques à courbes des données de séries temporelles, mais les axes ne correspondent pas à des points dans le temps (il n'y a aucun ordre naturel).

Création d'un graphique de parallèles simple

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Parallèles**.

   Le canevas est mis à jour en affichant un modèle de graphique de parallèles.
2. Sélectionnez au moins deux variables telles les variables **Colonnes**. Chaque colonne représente une verticale, axe parallèle sur le graphique.

Remarque : L'ordre des colonnes est important pour la recherche des fonctions. Dans une analyse de données standard, vous devrez peut-être réorganiser les colonnes à de nombreuses reprisee.

Fonctions supplémentaires

Colonnes
Répertorie les variables du jeu de données disponibles pour les axes *y* du graphique.

Cliquez sur **Ajouter une autre colonne** pour ajouter des colonnes.

Carte de couleurs
Répertorie les variables de carte de couleurs disponibles. Ces variables utilisent un dégradé de couleurs, basé sur la plage de valeurs de la colonne indiquée, pour leur représentation dans les points de tracé. Les cartes de couleurs sont également dénommées cartes choroplèthes.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Tableaux de relation

Un tableau de relation sert à déterminer la relation entre des variables.

Création d'un graphique de relation simple

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Tableau de relation**.

   Le canevas est mis à jour en affichant un modèle de graphique relationnel.
2. Sélectionnez au moins deux variables telles les variables **Colonnes**.

Fonctions supplémentaires

Colonnes
Répertorie les variables de jeux de données disponibles.

Cliquez sur **Ajouter une autre colonne** pour ajouter des colonnes.

Style de ligne
Contrôle le style de ligne entre les points de données associées.

Courbe
Lorsque cette option est sélectionnée, des courbes sont tracées entre les points de données associés.

Droite Lorsque cette option est sélectionnée, des lignes droites sont tracées entre points de données associées.

Seuil de libellés
Affiche des libellés pour les points de données dont les valeurs excèdent la valeur définie.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Boîtes à moustaches

Une boîte à moustaches affiche les cinq statistiques (minimum, premier quartile, médiane, troisième quartile et maximum). Cette fonction est utile pour afficher la distribution d'une variable d'échelle et mettre en évidence les valeurs extrêmes.

Création d'une boîte à moustaches

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Boîte à moustaches**.

   Le canevas est mis à jour en affichant un modèle de boîte à moustaches.
2. Sélectionnez une ou plusieurs variables d'échelle telle la variable **Colonne**.

Remarque : Les statistiques d'un tracé de points sont représentées par une boîte à moustaches. Ceci n'est pas modifiable.

Fonctions supplémentaires

Colonnes
Répertorie les variables du jeu de données qui sont disponibles pour l'axe *x* du graphique.

Cliquez sur **Ajouter une autre colonne** pour ajouter des colonnes.

Scinder par
Sélectionnez une variable catégorielle pour créer un tableau de graphiques, avec une cellule pour chaque catégorie dans la variable Scinder par. Comme le regroupement, la fonction Scinder par variables ajoute essentiellement plus de dimensions à votre graphique en affichant des informations pour chaque catégorie de variable. Reportez-vous à la rubrique Ajout de Scinder par variables pour plus d'informations.

Ordre des catégories
Sélectionnez l'ordre de tri des catégories de variable.

Telles qu'elles
Les catégories de variable sont présentées telles qu'elles figurent dans le jeu de données.

Ordre croissant
Les catégories de variable sont triées en ordre ascendant.

Ordre décroissant
Les catégories de variable sont triées en ordre descendant.

Normalisation des données
Lorsque cette fonction est activée, elle transforme les données en distribution normale ce qui

permet de comparer les données de plusieurs jeux de données ou de plusieurs colonnes. Cette fonction crée des empilements à 100% des nombres et convertit les statistiques en pourcentages.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

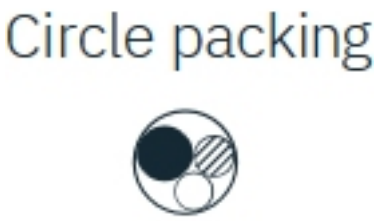
Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Graphiques à cercles imbriqués

Les graphiques à cercles imbriqués affichent les données hiérarchiques sous la forme d'un ensemble de zones imbriquées, ce qui vous permet de visualiser une grande quantité de données structurées de manière hiérarchique. Ils sont similaires à une carte d'arbre, mais utilise des cercles à la place de rectangles. Les graphiques à cercles imbriqués utilisent le confinement (imbrication) pour afficher les données de hiérarchie.

Création d'un graphique à cercles imbriqués

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Série temporelle**.



   Le canevas est mis à jour pour afficher un modèle de graphique à cercles imbriqués.
2. Sélectionnez une variable **Colonnes** dans la liste déroulante.

   Remarque : Cliquez sur **Ajouter une autre colonne** pour inclure d'autres variables de colonne.

Fonctions supplémentaires

Couleur de groupe
Activer ou désactiver les regroupements de couleurs.

Récapitulatif
Sélectionnez une fonction de récapitulatif statistique. Il s'agit de la méthode utilisée pour résumer chaque catégorie.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Graphique à bulles

Les graphiques à bulles affichent des catégories dans vos groupes comme cercles pleins non hiérarchiques. La taille de chaque cercle (bulle) est proportionnelle à sa valeur. Les graphiques à bulles sont utiles pour comparer les relations dans vos données.

Création d'un graphique à bulles simple

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Bulle**.

   Le canevas est mis à jour pour afficher un modèle de graphique à bulles.
2. Sélectionnez une variable **Colonnes** dans la liste déroulante.

   Remarque : Cliquez sur **Ajouter une autre colonne** pour inclure d'autres variables de colonne.

Fonctions supplémentaires

Couleur de groupe
Activer ou désactiver les regroupements de couleurs.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Graphiques de type Carte

Les graphiques de type Carte sont généralement utilisés pour comparer des valeurs et afficher des catégories de zones géographiques. Ils sont particulièrement adaptés lorsque les données contiennent des informations géographiques (pays, régions, états, départements, codes postaux, etc.).

Création d'un graphique de type Carte simple

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Carte**.

   Le canevas est mis à jour en affichant un modèle de graphique de type Carte.
2. Sélectionnez dans la liste déroulante **Longitude** une variable qui servira de valeur de longitude.
3. Sélectionnez dans la liste déroulante **Latitude** une variable qui servira de valeur de latitude.

Fonctions supplémentaires

Service de carte
Répertorie les services disponibles pour fournir des images de carte. Voir «Options du service de carte», à la page 85.

Type Répertorie les types de graphique disponibles pour représenter les données.

Longitude
Répertorie les variables disponibles pour servir de valeur de longitude.

Latitude
Répertorie les variables disponibles pour servir de valeur de latitude.

Infobulle
Répertorie les variables utilisables pour générer des informations d'infobulle lorsque vous survolez un point de données.

Carte de tailles
Répertorie les variables de carte de tailles disponibles. Ces variables utilisent différentes tailles pour leur représentation dans les points de tracé.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Options du service de carte : Dans le panneau Détails des graphiques de type carte, vous pouvez sélectionner le service de carte à utiliser pour fournir les images de carte.

Configuration des options du service de carte

1. Arrêtez IBM SPSS Modeler Server.
2. Pour appliquer des fichiers de carte locaux au format geojson, placez-les dans le répertoire `<répertoire_installation_Modeler>/dataview/conf/public/mapfiles`. Si ce répertoire n'existe pas, créez-le.
3. Ouvrez le fichier `<répertoire_installation_Modeler>/dataview/conf/application.conf` dans un éditeur de texte, puis ajoutez les options ci-après à la section `map.resources`. Dans cet exemple, deux nouvelles cartes intitulées `World - Local Map` et `World - Online Map` sont ajoutées.

```
map.resources {
  "mapservices" :
    [

      {
        "type"      : "geojson",
        "name"      : "World (Bulit-in)",
        "location"  : "built_in_world",
        "id"        : "built_in_world"
      },
      {
        "type"      : "geojson",
        "name"      : "Blank (Built-in)",
        "location"  : "built_in_blank",
        "id"        : "built_in_blank"
      },
      {
        "type"      : "geojson",
        "name"      : "World - Local Map",
        "location"  : "/mapfiles/world.json",
        "useProxy"  : "false",
        "id"        : "world_local"
      },
      {
        "type"      : "geojson",
        "name"      : "World - Online Map",
        "location"  : "http://map.example.com/map/world.json",
        "useProxy"  : "true",
        "id"        : "world_online"
      }    ]
}
```

Les paramètres suivants sont disponibles pour les options de carte :

Tableau 4. Paramètres du service de carte

Option	Valeur	Type	Obligatoire ?	Description
`type`	`geojson`	Chaîne	Oui	Actuellement, seule la valeur `geojson` est prise en charge pour cette option.
`name`	Nom de la carte	Chaîne	Oui	Entrez un nom pour la carte. Ce nom apparaît dans le menu déroulant **Service de carte** du panneau **Détails** des graphiques de type carte.
`location`	Emplacement de la carte	Chaîne	Oui	Entrez l'URL ou l'emplacement du fichier local de la carte. L'URL doit commencer par `http://` ou `https://`. Le fichier local doit commencer par le chemin `/mapfiles/`.
`useProxy`	`true` ou `false`	Chaîne	Oui	Utilisez `true` si l'emplacement de la carte correspond à une URL. Utilisez `false` si l'emplacement de la carte correspond à un fichier local.
`id`	ID carte	Chaîne	Oui	Utilisez un ID unique pour chaque carte.

Remarque :

- Ne supprimez pas les options **World (Built-in)** et **Blank (Built-in)** par défaut.
- N'oubliez pas d'utiliser une virgule entre les blocs de contenu, comme illustré dans l'exemple, et assurez-vous que votre nouvelle section `map.resources` correspond à un contenu au format JSON valide.
- IBM SPSS Modeler Server (ou le client SPSS Modeler autonome sans serveur) doit pouvoir accéder à l'emplacement de la carte.

4. Démarrez IBM SPSS Modeler Server.

Désormais, lorsque vous utilisez des graphiques de type carte, vous pouvez voir le nouveau service de carte que vous avez ajouté dans le panneau **Détails**.

Cartes thermiques

Les cartes thermiques présentent les données de sorte que les valeurs individuelles d'une matrice soient représentées par des couleurs.

Création d'une carte thermique simple

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Carte thermique**.

Le canevas est mis à jour en affichant un modèle de carte thermique.

2. Sélectionnez une variable telle que **Colonne**. Chaque catégorie de variable est représentée sous forme de colonne individuelle dans le graphique.
3. Sélectionnez une variable telle que **LIgne**. Chaque catégorie de variable est représentée sous forme de ligne individuelle dans le graphique.

Fonctions supplémentaires

Colonne
Répertorie les variables du jeu de données disponibles comme colonnes du graphique. Chaque catégorie de variable est représentée sous forme de colonne individuelle dans le graphique.

Ligne Répertorie les variables du jeu de données disponibles pour les lignes du graphique. Chaque catégorie de variable est représentée sous forme de ligne individuelle dans le graphique.

Ordre des catégories
Sélectionnez l'ordre de tri des catégories de variable.

Telles qu'elles
Les catégories de variable sont présentées telles qu'elles figurent dans le jeu de données.

Ordre croissant
Les catégories de variable sont triées en ordre ascendant.

Ordre décroissant
Les catégories de variable sont triées en ordre descendant.

Récapitulatif
Sélectionnez une fonction statistique récapitulative pour l'élément graphique. Le résultat de la statistique détermine la position des éléments graphiques sur l'axe des *Y*. Dans un graphique 2D, la statistique est calculée pour chaque valeur figurant sur l'axe des *X*. Dans un graphique 3D, elle est calculée pour l'intersection des valeurs figurant sur les axes des *X* et des *Z*.

Il existe deux types de fonction de récapitulatif statistique. La distinction est importante car elle détermine si vous devez spécifier une variable **Valeur** comme axe *y*.

- **Fonctions ne nécessitant pas une variable de valeur.** Il s'agit des fonctions qui ne nécessitent pas de variable d'axe des *Y*. Toutes les statistiques d'effectif et de pourcentage se trouvent dans cette catégorie. Ces statistiques sont disponibles lorsqu'aucune variable **Valeur** n'a été définie.
- **Fonctions nécessitant une variable de valeur.** Il s'agit des fonctions qui nécessitent une variable **Valeur** pour l'axe des *y*. Par exemple, la fonction *Moyenne* nécessite une variable avec laquelle la moyenne est calculée. Ces statistiques sont disponibles lorsqu'une variable **Valeur** a été définie.

Valeur Ce champ s'affiche lorsqu'une fonction **Récapitulatif**, laquelle requiert une variable pour l'axe des *y*, est sélectionnée. Sélectionnez une variable comme valeur pour l'axe des *y*.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Graphiques t-SNE

t-SNE (T-distributed Stochastic Neighbor Embedding) est un algorithme d'apprentissage automatique pour la visualisation de données. Les graphiques t-SNE modélisent chaque objet de grande dimension en

un point à deux ou trois dimensions de sorte que les objets similaires soient modélisés par des points voisins et les objets dissemblables par des points distants avec haute probabilité.

Création d'un graphique t-SNE simple

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **t-SNE**.

   Le canevas est mis à jour en affichant un modèle de graphique t-SNE.
2. Définissez les valeurs **Perplexité**, **Taux d'apprentissage** et **Nombre maximal d'itérations**.
3. Sélectionnez éventuellement une variable **Carte de couleur**.

Fonctions supplémentaires

Perplexité
: Définit un nombre qui établit une estimation éclairée comme étant le nombre de voisins proches pour chaque point de données. Cette fonction vise à équilibrer les aspects locaux et globaux de vos données.

Taux d'apprentissage
: Cette valeur affecte la vitesse d'apprentissage en spécifiant les modifications de taille et de poids à chaque itération.

Nombre maximal d'itérations
: Nombre maximal d'itérations à effectuer.

Carte de couleurs
: Répertorie les variables de carte de couleurs disponibles. Ces variables utilisent un dégradé de couleurs, basé sur la plage de valeurs de la colonne indiquée, pour leur représentation dans les points de tracé. Les cartes de couleurs sont également dénommées cartes choroplèthes.

Titre principal
: Titre du graphique.

Note de bas de page
: Note de bas de page du graphique qui figure en-dessous de celui-ci.

Graphiques de nuages de mots

Les graphiques de nuages de mots présentent les données sous forme de mots où la taille et l'emplacement de chaque mot est déterminé par sa pondération.

Création d'un graphique de nuage de mots simples

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Nuage de mots**.

   Le canevas est mis à jour en affichant un modèle de graphique de nuage de mots.
2. Sélectionnez une variable **Source**. Chaque catégorie de variable est représentée dans le graphique en fonction de sa valeur pondérée.
3. Sélectionnez une valeur **Forme** pour le graphique. Les données du graphique résultant sont présentées sous la forme sélectionnée.

Fonctions supplémentaires

Source
: Recense les variables du jeu de données disponibles comme source du graphique. Chaque catégorie de variable est représentée dans le graphique en fonction de sa valeur pondérée.

Forme Répertorie les formes de graphique disponibles. Les données du graphique résultant sont présentées sous la forme sélectionnée.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Graphiques de barre d'erreurs

Les graphiques de barre d'erreur représentent la variabilité des données et indiquent l'erreur (ou l'incertitude) dans une mesure signalée. Les barres d'erreur permettant de déterminer si les différences sont significatives du point de vue statistique. Les barres d'erreur peut également suggérer la qualité de l'adéquation pour une fonction donnée.

Création d'un graphique de barre d'erreur simple

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Barre d'erreur**.

   Le canevas est actualisé en affichant un canevas de graphique de barre d'erreur.
2. Sélectionnez une variable telle que **Axe des Y**. Il s'agit de la variable dont les données sont représentées sur l'axe des *y*.
3. Sélectionnez une variable d'échelle telle la variable **Catégorie**. Il s'agit de la variable dont les données sont représentées sur l'axe des *x*.

Fonctions supplémentaires

Axe des Y
Répertorie les variables du jeu de données qui sont disponibles pour l'axe *y* du graphique.

Catégorie
Répertorie les variables du jeu de données qui sont disponibles pour l'axe *x* du graphique.

Ordre des catégories
Sélectionnez l'ordre de tri des catégories de variable.

Telles qu'elles
Les catégories de variable sont présentées telles qu'elles figurent dans le jeu de données.

Ordre croissant
Les catégories de variable sont triées en ordre ascendant.

Ordre décroissant
Les catégories de variable sont triées en ordre descendant.

Scinder par
Sélectionnez une variable catégorielle pour créer un tableau de graphiques, avec une cellule pour chaque catégorie dans la variable Scinder par. Comme le regroupement, la fonction Scinder par variables ajoute essentiellement plus de dimensions à votre graphique en affichant des informations pour chaque catégorie de variable. Reportez-vous à la rubrique Ajout de Scinder par variables pour plus d'informations.

Ligne de référence
Lorsque cette option est activée, affiche une ligne de référence sur le graphique. La ligne de référence est corrélée à la **méthode statistique** sélectionnée.

Barres d'erreur
Lorsque cette fonction est activée, les lignes qui représentent la marge d'erreur s'affichent dans le graphique.

Mesure
Sélectionnez le type de mesure représenté par les barres d'erreur :

Intervalles de confiance
Définit l'intervalle de confiance pour les variables sélectionnées. Valeur par défaut : 0,95 (95 %), comme reflété dans le champ **Valeur de représentation**.

Erreur standard
Mesure l'erreur standard des variables sélectionnées.

Ecart type
Mesure les écarts type des variables sélectionnées.

Niveau de confiance
Cette valeur représente les intervalles de confiance de la **Mesure** sélectionnée. Valeur par défaut 0,95 (95 %).

Méthode statistique
Sélectionnez la méthode pour décrire la tendance centrale :

Moyenne
Résultat de la somme des ratios divisée par le nombre total de ratios.

Médiane
Valeur telle que le nombre de ratios qui lui sont inférieurs est égal au nombre de rations qui lui sont supérieurs.

Mode d'affichage
Sélectionnez comment s'affiche la sélection de la **Méthode statistique** (barre, ligne ou cercle).

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Graphiques 3D

Les graphiques 3D sont généralement utilisés pour représenter des fonctions à plusieurs variables et incluent une variable axe des Z qui est une fonction des variables axe des X et axe des Y.

Création d'un graphique 3D simple

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Graphique 3D**.

   Le canevas est mis à jour en affichant un modèle de graphique 3D.
2. Sélectionnez le **Type** de graphique dans la liste déroulante.
3. Sélectionnez une variable **axe des X** dans la liste déroulante.
4. Sélectionnez une variable **axe des Y** dans la liste déroulante.
5. Sélectionnez une variable **axe des Z** dans la liste déroulante.

Fonctions supplémentaires

Type Répertorie les types de graphique disponibles pour représenter les données.

Axe des X
Répertorie les variables du jeu de données qui sont disponibles pour l'axe *x* du graphique.

Axe des Y
Répertorie les variables du jeu de données qui sont disponibles pour l'axe *y* du graphique.

Axe des Z
Répertorie les variables du jeu de données qui sont disponibles pour l'axe *z* du graphique.

Infobulle
Répertorie les variables utilisables pour générer des informations d'infobulle lorsque vous survolez un point de données.

Carte de couleurs
Répertorie les variables de carte de couleurs disponibles. Ces variables utilisent un dégradé de couleurs, basé sur la plage de valeurs de la colonne indiquée, pour leur représentation dans les points de tracé. Les cartes de couleurs sont également dénommées cartes choroplèthes.

Carte de tailles
Répertorie les variables de carte de tailles disponibles. Ces variables utilisent différentes tailles pour leur représentation dans les points de tracé.

Rapport Z
Définit l'échelle des valeurs de données de l'axe *z*, par rapport aux axes *x* et y.

Rotation
Active et désactive la rotation du graphique.

Infobulles de point de données
Définit comment s'affichent les infobulles des points de données (à droite des points de données, en haut à droite du graphique ou masquées).

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Graphiques de matrice de nuage de points

Les matrices de nuage de points sont un bon outil pour déterminer s'il existe des corrélations linéaires entre plusieurs variables.

Création d'un graphique de matrice de nuage de points

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Matrice de nuage de points**.

   Le canevas est mis à jour en affichant un modèle de graphique de matrice de nuage de points.
2. Sélectionnez plusieurs variables d'échelle **Colonnes**.

Chaque variable sélectionnée est tracée par rapport à chaque autre variable afin de créer une matrice de nuage de points individuels.

Fonctions supplémentaires

Corrélation
Lorsque cette option est activée, des informations de corrélation linéaire (forte, moyenne, faible) sont affichées pour les variables sélectionnées.

Colonnes
Sélectionnez au moins deux variables de matrice. Les variables doivent être numériques (mais non au format date).

Cliquez sur **Ajouter une autre colonne** pour ajouter des colonnes.

Carte de couleurs
Répertorie les variables de carte de couleurs disponibles. Ces variables utilisent un dégradé de couleurs, basé sur la plage de valeurs de la colonne indiquée, pour leur représentation dans les points de tracé. Les cartes de couleurs sont également dénommées cartes choroplèthes.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Graphiques en chandeliers

Les graphiques en chandeliers sont un type de graphiques financiers utilisés pour décrire les variations de prix d'une garantie, d'une dérivée ou d'une devise. Chaque chandelier correspond généralement à un jour. Un graphique sur un mois affiche les 20 jours de bourse sous la forme de 20 chandeliers. Les graphiques en chandeliers sont le plus souvent utilisés dans l'analyse des tendances de prix des capitaux propres et des devises et sont similaires aux boîtes à moustaches.

Le jeu de données utilisé pour créer un graphique en chandeliers doit contenir des valeurs d'ouverture, maximale, minimale et de clôture pour chaque période à afficher.

Création d'un graphique en chandeliers simple

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Chandelier**.

   Le canevas est mis à jour en affichant un modèle de graphique en chandeliers.
2. Sélectionnez une variable telle que **Axe des X**.
3. Sélectionnez une variable telle que **Max**.
4. Sélectionnez une variable telle que **Min**.

Fonctions supplémentaires

Axe des X
Répertorie les variables du jeu de données qui sont disponibles pour l'axe *x* du graphique.

Max Répertorie les variables du jeu de données disponibles comme valeur de prix élevée du graphique.

Récapitulatif de la zone Elevé
Sélectionnez une fonction statistique récapitulative pour la variable Elevé sélectionnée.

Min Répertorie les variables du jeu de données disponibles comme valeur de prix faible du graphique.

Récapitulatif de la zone Faible
Sélectionnez une fonction statistique récapitulative pour la variable Faible sélectionnée.

Ouverture
Répertorie les variables du jeu de données disponibles comme valeur de prix d'ouverture du graphique.

Clôture
Répertorie les variables du jeu de données disponibles comme valeur de prix de clôture du graphique.

Volume
Répertorie les variables du jeu de données disponibles comme barres de volume du graphique.

Chandelier
Fait basculer les données du graphique d'un affichage en chandeliers à un affichage en courbes ou inversement.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Graphique de type rivière

Une graphique de type rivière est un graphique de flux spécialisé qui affiche des changements dans le temps.

Création d'un graphique de type rivière simple

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Graphique de type rivière**.

   Le canevas est mis à jour pour afficher un modèle de graphique de type rivière.
2. Sélectionnez une variable **Axe des X**.
3. Sélectionnez une valeur pour **Catégorie**.

   Restriction : Si un champ de catégorie comporte plus de 50 catégories distinctes, seules les 50 premières catégories maximum seront utilisées comme événements dans le graphique.

Fonctions supplémentaires

Ordre basé sur
En fonction de la valeur de l'axe des X que vous sélectionnez, vous pouvez être en mesure de spécifier si l'ordre des catégories est basé sur le nom de catégorie ou la valeur de catégorie.

Ordre des catégories de l'axe des X
En fonction de la valeur de l'axe des X que vous sélectionnez, vous pouvez être en mesure de spécifier si l'ordre des catégories est croissant, décroissant ou tel qu'il est lu à partir du jeu de données.

Récapitulatif
Sélectionnez une fonction de récapitulatif statistique. Il s'agit de la méthode utilisée pour résumer chaque catégorie.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

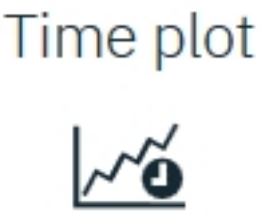
Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Tracés horaires

Les tracés horaires illustrent des points de données à des intervalles de temps successifs. Les séries temporelles représentées doivent contenir des valeurs numériques et sont supposées avoir lieu sur une durée au sein de laquelle les périodes sont uniformes. Les tracés horaires fournissent une analyse préliminaire des caractéristiques hte des séries temporelles sur les statistiques de base et les tests, et génèrent ainsi des informations utiles sur vos données avant la modélisation. Les séries temporelles comprennent des méthodes d'analyse telles que la décomposition, le test de Dickey-Fuller augmenté (ADF), les corrélations (ACF/PACF) et l'analyse spectrale.

Création d'une série temporelle simple

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Série temporelle**.



   Le canevas est mis à jour et affiche un modèle de graphique en série temporelle.
2. Dans la liste déroulante **Valeurs**, sélectionnez une valeur pour l'axe des Y.

Fonctions supplémentaires

Date Sélectionnez une date dans la liste déroulante, le cas échéant. Chaque observation est séparée par le même intervalle de temps. Si vous sélectionnez une variable de date, une option de rééchantillonnage s'affiche. Vous pouvez utiliser cette option pour agréger les champs de valeur afin qu'ils correspondent à l'intervalle spécifié.

Algorithme de tracé de temps
Algorithme de tracé de temps à utiliser pour l'analyse des données de séries temporelles :

- **Décomposition.** Décompose une série temporelle en trois composants (cycle de tendance, saisonnier, et irrégulier). La décomposition s'effectue de façon additive.
- **Test ADF.** Les tests augmentés de Dickey–Fuller (ADF) testent l'hypothèse nulle selon laquelle il y a une racine unitaire dans la série et que la série n'est pas stationnaire. Si le résultat du test rejette l'hypothèse nulle, cela signifie que la série est stationnaire ou peut être représentée comme étant stationnaire avec un modèle de différence.
- **ACF/PACF.** Corrélations de séries.
- S**Anlyse spectrale.** Un outil analytique dans le domaine des fréquences. La fréquence la plus élevée est marquée par un diamant.

Permuter la position du graphique
Inverse les positions des graphiques sur le système de coordonnées rectangulaires planimétriques et le système de coordonnées polaires.

Afficher le point décisif
Affiche ou masque les points de retournements dans les graphiques. En se basant sur les composantes du cycle de tendance de la décomposition des séries temporelles, nous explorons si la série a une tendance générale ou a des points qui modifient la direction du motif de tendance.

Afficher la valeur aberrante
Affiche ou masque les valeurs aberrantes. Les valeurs aberrantes de la série temporelle sont analysées à partir de la composante irrégulière de la décomposition des séries temporelles.

Titre principal
Titre du graphique.

Sous-titre
Sous-titre du graphique, figure directement en-dessous de son titre.

Note de bas de page
Note de bas de page du graphique qui figure en-dessous de celui-ci.

Graphiques personnalisés

L'option Graphiques personnalisés vous permet de coller/éditer le code JSON afin de créer le graphique souhaité.

Création d'un graphique personnalisé

1. Dans la section **Types de graphique** du Générateur de graphiques, cliquez sur l'icône **Personnalisé**.

2. Collez le code JSON contenant les spécifications du graphique dans le champ prévu à cet effet dans le panneau Détails.
3. Cliquez sur **Générer un graphique**.

Tableau de bord

Vous pouvez créer une présentation de tableau de bord pour visualiser plusieurs graphiques à la fois. Enregistrez la présentation comme modèle, puis faites glisser les graphiques enregistrés et déposez-les aux positions voulues sur la présentation.

1. Après avoir cliqué avec le bouton droit de la souris sur un noeud de données et sélectionné l'option **Afficher les données**, cliquez sur la commande Tableau de bord dans la section **Actions**.

Figure 14. Commande Tableau de bord

Le tableau de bord s'affiche et propose les paramètres ci-après.

Modèle
Dans cette section, vous pouvez sélectionner l'un des modèles de présentation prédéfinis ou enregistrés ou en créer un nouveau.

Sélectionner un modèle de présentation
Choisissez parmi les modèles de présentation disponibles ou utilisez un nouveau modèle.

Actions

Cliquez sur l'icône **Editer la présentation** pour modifier le modèle de présentation que vous avez sélectionné. Vous pouvez également importer un fichier de tableau de bord. Lorsque vous avez terminé, enregistrez votre modèle.

Cliquez sur l'icône **Quitter l'édition de la présentation** une fois l'édition terminée.

Contenu

Utilisez cette section pour faire glisser des éléments vers la présentation du tableau de bord.

Choisir un graphique enregistré

Cette section contient la liste des graphiques que vous avez enregistrés. Faites glisser des graphiques et déposez-les aux positions voulues sur la présentation de tableau de bord.

Choisir un objet

Vous pouvez également faire glisser le texte ou les images HTML vers la présentation du tableau de bord.

Préférences de visualisation globales

Vous pouvez remplacer les paramètres par défaut des titres, du curseur de plage, des quadrillages et du suivi de la souris. Vous pouvez également indiquer un modèle de schéma de couleurs différent.

1. Après avoir cliqué avec le bouton droit de la souris sur un noeud de données et sélectionné l'option **Afficher les données**,cliquez sur la commande des préférences de visualisation globale dans la section **Actions**.

Figure 15. Contrôle des préférences de visualisation globales

La boîte de dialogue Préférences de visualisation globales, qui contient les paramètres suivants, s'affiche.

Titres Cette section fournit des paramètres globaux de titre de graphique.

Titres globaux

Active ou désactive les titres globaux pour tous les graphiques.

Titre principal global

Active ou désactive l'affichage des titres de graphique principaux, globaux. Lorsque ce paramètre est activé, le titre de graphique de niveau supérieur que vous entrez ici est appliqué à tous les graphiques, se substituant au paramètre **Titre principal** de chaque graphique.

Sous-titre global

Active ou désactive l'affichage des sous-titres de graphique globaux. Lorsque ce paramètre est activé, le sous-titre de graphique que vous entrez ici est appliqué à tous les graphiques, se substituant au paramètre **Sous-titre** de chaque graphique.

Titres par défaut

Active ou désactive les titres par défaut pour tous les graphiques.

Outils Cette section fournit des options de contrôle du comportement des graphiques.

Curseur de plage

Active ou désactive le curseur de plage pour chaque graphique. Lorsque ce paramètre est activé, vous pouvez contrôler la quantité de données du graphique qui s'affiche à l'aide du curseur de plage affiché sous chaque graphique.

Quadrillages
Contrôle l'affichage des quadrillages de l'axe des X (vertical) et des Y (horizontal).

Suivi de la souris
Lorsque cette option est activée, l'emplacement du curseur, par rapport aux données du graphique, est suivi et affiché lorsqu'il est situé sur le graphique.

Boîte à outils
Active ou désactive la boîte à outils pour chaque graphique. En fonction du type de graphique, la boîte à outils à droite de l'écran fournit des outils, tels que le zoom, l'enregistrement en tant qu'image, la restauration, la sélection de données et l'annulation d'une sélection.

Thème
Sélectionnez un modèle pour modifier les couleurs utilisées dans les graphiques ayant une variable de regroupement ou d'empilement. Tout attribut d'élément défini dans le fichier modèle sélectionné remplace les paramètres du modèle par défaut pour ces attributs d'éléments.

Notification
Indiquez si vous souhaitez recevoir un avertissement lors d'un changement de type de graphique ou de la navigation vers la page de démarrage.

2. Cliquez sur **Appliquer** pour enregistrer vos paramètres ou sur **Annuler** pour ne pas prendre en compte les modifications.

Chapitre 7. Utilisation de la sortie

Lorsque vous exécutez certains flux, les résultats sont disponibles dans le Visualiseur via l'onglet **Modèle** ou l'onglet **Avancé** des noeuds nugget du modèle. Vous pouvez accéder facilement dans le Visualiseur à la sortie que vous désirez examiner. Vous pouvez également manipuler la sortie et créer un document contenant exactement la sortie voulue. Certaine sorties graphiques utilisent le Visualiseur.

Le Visualiseur est utilisé pour les sorties suivantes dans IBM SPSS Modeler:

- Nuggets de modèle TCM
- Nuggets de modèle STP
- Nuggets de modèle de cluster TwoStep-AS
- Nuggets de modèle GSAR
- Noeud de graphique Visualisation de carte

Visualiseur

Les résultats sont affichés dans le Visualiseur . Vous pouvez utiliser l' du Visualiseur pour :

- Parcourir les résultats
- Afficher ou masquer les tableaux et graphiques sélectionnés.
- Modifier l'ordre d'affichage des résultats en déplaçant les éléments sélectionnés.
- Déplacer des éléments entre le Visualiseur et d'autres applications

Le Visualiseur est divisé en deux panneaux :

- Le panneau gauche contient la légende du contenu du résultat.
- Le panneau droit contient les tableaux statistiques, les graphiques et les sorties textes.

Vous pouvez cliquer sur un élément affiché dans la légende pour accéder directement au tableau ou au graphique correspondant. Vous pouvez cliquer et faire glisser le bord droit du panneau de légende pour modifier la largeur de la fenêtre.

Affichage et masquage des résultats

Dans le Visualiseur , vous pouvez sélectionner individuellement des tableaux ou les résultats d'une procédure pour les afficher ou les masquer. Ce processus est utile si vous voulez limiter la proportion de sorties visibles dans le panneau de contenu.

Pour masquer les tableaux et les graphiques

1. Double-cliquez sur l'icône Livre de l'élément dans le panneau de légende du Visualiseur.
   ou
2. Cliquez sur l'élément pour le sélectionner.
3. A partir des menus, sélectionnez :
   Affichage > **Masquer**
   ou
4. Cliquez sur l'icône Livre fermé (Masquer) dans la barre d'outils Légende.

L'icône Livre ouvert (Montrer) devient l'icône active, indiquant que l'élément est désormais masqué.

Masquage des résultats de la procédure

1. Cliquez sur la boîte à gauche du nom de la procédure dans le panneau de légende.

Cela masque tous les résultats de la procédure et réduit l'affichage de la légende.

Déplacement, suppression et copie de sortie

Vous pouvez réarranger les résultats en copiant, en déplaçant ou en effaçant un élément ou un groupe d'éléments.

Déplacement des sorties dans le Visualiseur

1. Sélectionnez les éléments affichés dans le panneau de légende ou de contenu.
2. Faites glisser les éléments sélectionnés dans un emplacement différent.

Suppression de la sortie dans le Visualiseur

1. Sélectionnez les éléments affichés dans le panneau de légende ou de contenu.
2. Appuyez sur la touche **Supprimer**.

   ou
3. A partir des menus, sélectionnez :

   Edition > **Supprimer**

Modification de l'alignement initial

Par défaut, tous les résultats sont initialement alignés à gauche. Pour changer l'alignement initial des nouveaux éléments de sortie :

1. A partir des menus, sélectionnez :

   Editer > **Options**
2. Cliquez sur l'onglet **Visualiseur**.
3. Dans le groupe Etat initial de sortie, sélectionnez le type d'élément (par exemple : tableau croisé dynamique, graphique, sortie texte).
4. Sélectionnez l'option d'alignement souhaitée.

Changement de l'alignement des éléments de sortie

1. Dans le panneau de légende ou de contenu, sélectionnez les éléments à aligner.
2. A partir des menus, sélectionnez :

   Format > **Aligner à gauche**

   ou

   Format > **Centre**

   ou

   Format > **Aligner à droite**

Panneau de légende du Visualiseur

Le panneau de légende fournit une table des matières du document du Visualiseur . Vous pouvez utiliser le panneau de légende pour naviguer dans votre résultat et contrôler l'affichage du résultat. La plupart des actions dans le panneau de légende ont un effet correspondant sur le panneau de contenu.

- Le fait de sélectionner un élément dans le panneau de légende affiche l'élément correspondant dans le panneau de contenu.
- Déplacer un élément dans le panneau de légende déplace l'élément correspondant dans le panneau de contenu.
- Réduire la ligne de légende masque le résultat de tous les éléments des niveaux réduits.

Contrôler la légende du Résultat : Pour contrôler la légende du résultat, vous pouvez :

- Développer et réduire l'affichage de la légende.
- Modifier le niveau de ligne de légende pour les éléments sélectionnés.
- Modifier la taille des éléments dans l'affichage de la ligne de légende.
- Modifier la police utilisée dans l'affichage de la ligne de légende.

Développement et réduction de l'affichage des légendes

1. Cliquez sur la boîte à gauche de l'élément de ligne de légende à réduire ou à étendre.
   ou
2. Cliquez sur l'élément dans la ligne de légende.
3. A partir des menus, sélectionnez :
   Affichage > **Réduire**
   ou

Affichage > **Développer**

Modification du niveau de légende

1. Cliquez sur l'élément dans le panneau de légende.
2. A partir des menus, sélectionnez :
   Edition > **Légende** > **Promouvoir**
   ou
   Edition > **Légende** > **Rétrograder**

Modification de la taille des éléments dans la légende

1. A partir des menus, sélectionnez :
   Affichage > **Taille de la ligne de légende**
2. Sélectionnez la taille de la ligne de légende (**Petite**, **Moyenne**, ou **Grande**).

Modification de la police dans la légende

1. A partir des menus, sélectionnez :
   Affichage > **Police de la ligne de légende...**
2. Sélectionnez une police.

Ajout d'éléments dans le Visualiseur

Dans le Visualiseur , vous pouvez ajouter des éléments tels que des titres, du texte, des graphiques ou des éléments provenant d'autres applications.

Ajout d'un titre ou d'un texte

Les éléments de texte qui ne sont pas connectés à un tableau ou un graphique peuvent être ajoutés dans le Visualiseur .

1. Cliquez sur le tableau, graphique ou autre objet qui précédera le titre ou texte.
2. A partir des menus, sélectionnez :
   Insérer > **Nouveau titre**
   ou
   Insérer > **Nouveau texte**
3. Double-cliquez sur le nouvel objet.
4. Entrez le texte.

Ajout d'un fichier texte

1. Dans le panneau de légende ou le panneau de contenu du Visualiseur, cliquez sur le tableau, le graphique ou tout autre objet qui précédera le texte.
2. A partir des menus, sélectionnez :

   Insérer > **Fichier texte...**
3. Sélectionnez un fichier texte.

Pour modifier le texte, double-cliquez dessus.

Collage d'objets dans le Visualiseur

Vous pouvez coller dans le Visualiseur des objets provenant d'autres applications. Pour se faire, vous utilisez l'option **Coller après** ou **Collage spécial**. Les deux types de collage placent le nouvel objet à la suite de l'objet actuellement sélectionné dans le Visualiseur . Utilisez **Collage spécial** lorsque vous voulez choisir le format de l'objet à coller.

Recherche et remplacement d'informations dans le Visualiseur

1. Pour rechercher et remplacer des informations dans le Visualiseur. à partir des menus, sélectionnez :

   Editer > **Rechercher**

   ou

   Editer > **Remplacer**

Vous pouvez utiliser l'option Rechercher et remplacer pour :

- Effectuer une recherche dans l'ensemble du document ou uniquement les éléments sélectionnés.
- Effectuer une recherche vers le haut ou vers le bas à partir de l'emplacement en cours.
- Effectuer une recherche dans les deux panneaux ou restreindre la recherche au panneau de contenu ou de légende.
- Rechercher des éléments masqués. Il s'agit de tout élément masqué dans le panneau de contenu (par exemple des tableaux de remarques masqués par défaut) ainsi que des lignes et colonnes masquées dans les tableaux croisés dynamiques.
- Restreindre les critères de recherche aux résultats respectant la casse.
- Restreindre les critères de recherche dans les tableaux croisés dynamiques aux correspondances du contenu global des cellules.
- Restreindre les critères de recherche dans les tableaux croisés dynamiques aux marqueurs de note de bas de page uniquement. Cette option n'est pas disponible si la sélection dans le Visualiseur comprend autre chose que des tableaux croisés dynamiques.

Eléments masqués et couches de tableau croisé dynamique

- Les couches se trouvant sous la couche actuellement visible d'un tableau croisé dynamique multidimensionnel ne sont pas considérées comme masquées et sont incluses dans la zone de recherche même quand les éléments masqués ne sont pas inclus dans la recherche.
- Les éléments masqués incluent les éléments masqués du panneau de contenu (éléments accompagnés d'icônes de livre fermé dans le panneau de légende ou inclus dans des blocs réduits du panneau de légende) ainsi que les lignes et colonnes de tableaux croisés dynamiques masquées par défaut (par exemple, les lignes et colonnes vides sont masquées par défaut) ou masquées manuellement en modifiant le tableau et en masquant de façon sélective des lignes ou colonnes spécifiques. Les éléments masqués ne sont inclus dans la recherche que si vous avez explicitement sélectionné **Inclure les éléments masqués**.
- Dans les deux cas, l'élément masqué ou non visible contenant le texte ou la valeur de recherche est affiché une fois trouvé, mais cet élément revient ensuite à son état d'origine.

Recherche d'une plage de valeurs dans des tableaux croisés dynamiques

Pour rechercher des valeurs qui tombent dans une plage spécifique dans des tableaux croisés dynamiques :

1. Activez un tableau croisé dynamique ou sélectionnez un ou plusieurs tableaux croisés dynamiques dans le Visualiseur. Vérifiez que seuls des tableaux croisés dynamiques sont sélectionnés. Si tout autre objet est sélectionné, l'option Plage n'est pas disponible.
2. A partir des menus, sélectionnez :

   Editer > **Rechercher**
3. Cliquez sur l'onglet **Plage**.
4. Sélectionnez le type de plage : Entre, Supérieur ou égal à ou Inférieur ou égal à.
5. Sélectionnez la ou les valeurs qui définissent la plage.

- Si une des valeurs contient des caractères non numériques, les deux valeurs sont traitées en tant que chaînes.
- Si les deux valeurs sont de nombres, seules les valeurs numériques sont recherchées.
- Vous ne pouvez pas utiliser l'onglet Plage pour remplacer des valeurs.

Cette fonction n'est pas disponible pour les tableaux classiques. Pour plus d'informations, voir la rubrique «Tableaux de version antérieure», à la page 128.

Copie de sorties dans d'autres applications

Les objets de sortie peuvent être copiés et collés dans d'autres applications telles que les traitements de texte ou les tableurs. Vous pouvez coller les sorties de différentes façons. En fonction de l'application cible et des objets de sortie sélectionnés, certains formats suivants (ou tous les formats) peuvent être disponibles :

Métafichier : Format de métafichier WMF et EMF. Ces formats ne sont disponibles que sous les systèmes d'exploitation Windows.

RTF (texte enrichi) : Les objets sélectionnés, les sorties texte et les tableaux croisés dynamiques peuvent être copiés et collés au format RTF. Dans la plupart des applications, cette procédure colle le tableau croisé dynamique en tant que tableau pouvant être modifié dans une autre application. Les tableaux croisés dynamiques qui sont trop larges par rapport au document, seront tronqués, redimensionnés pour s'ajuster à la largeur du document, ou laissés tels quels selon les paramètres du tableau croisé dynamique. Pour plus d'informations, voir la rubrique «Options de tableau croisé dynamique», à la page 129.

Remarque : Microsoft Word peut ne pas afficher correctement les tableaux de très grande largeur.

Image : Format d'image JPG et PNG.

BIFF : Les sorties sous forme de texte et les tableaux croisés dynamiques peuvent être collés dans un tableur au format BIFF. Les nombres des tableaux croisés dynamiques gardent leur précision numérique. Ce format n'est disponible que sous les systèmes d'exploitation Windows.

Texte : Les sorties texte et les tableaux croisés dynamiques peuvent être copiés et collés en tant que texte. Cette option peut être utile pour les applications telles que le courrier électronique, où l'application ne peut accepter ou transmettre que du texte.

Objet graphique Microsoft Office. Les graphiques qui prennent en charge ce format peuvent être copiés dans des applications Microsoft Office et édités dans ces applications en tant que graphiques Microsoft Office natifs. En raison des différences entre les graphiques SPSS Statistics/SPSS Modeler et les graphiques Microsoft Office, certaines fonctionnalités des graphiques SPSS Statistics/SPSS Modeler ne

sont pas conservées dans la version copiée. La copie de plusieurs graphiques sélectionnés dans le format Objet graphique Microsoft Office n'est pas prise en charge.

Si l'application cible prend en charge plusieurs formats disponibles, elle possède certainement un élément de menu Collage spécial vous permettant de sélectionner le format, ou elle affiche probablement automatiquement une liste des formats disponibles.

Remarque : Microsoft Office version 16 (ou version supérieure) est requis pour la copie et le collage de boîtes à moustaches et d'histogrammes.

Copie et collage de plusieurs objets de sortie

Les restrictions suivantes s'appliquent au collage de plusieurs objets de sortie dans d'autres applications :

- **Format RTF :** Dans la plupart des applications, cette procédure colle les tableaux croisés dynamiques en tant que tableaux pouvant être modifié dans une autre application. Les graphiques, graphiques d'arbre et les vues de modèle sont collés en tant qu'images.
- **Formats de métafichier et d'image :** Tous les objets de sortie sélectionnés sont collés comme s'il s'agissait d'un seul objet dans l'application cible.
- **Format BIFF :** Les graphiques, graphiques d'arbre et les vues de modèle sont exclus.

Copier spécial

Lors de la copie et du collage de grands volumes de sorties, par exemple des tableaux croisés dynamiques particulièrement grands, vous pouvez augmenter la vitesse de l'opération en utilisant la fonction **Modifier** > **Copier spécial** pour limiter le nombre de formats copiés vers le Presse-papier.

Il est aussi possible d'enregistrer les formats sélectionnés comme ensemble de formats par défaut à utiliser pour la copie vers le Presse-papiers. Ce paramètre sera conservé de session en session.

Copier en tant que

Vous pouvez cliquer à l'aide du bouton droit de la souris sur un objet sélectionné dans le Visualiseur de sortie et sélectionner **Editer** > **Copier en tant que** pour copier effectuer la copie dans l'un des formats les plus populaires (par exemple, **Tous**, **Image** ou **Objet graphique Microsoft Office**). La sélection de **Editer** > **Copier** copie vers **tous** les formats. Si **Copier en tant que** est grisé ou n'est pas disponible pour un objet sélectionné, cela signifie que ce format de copie n'est pas disponible pour cet objet.

Sortie interactive

Les objets de sortie interactive contiennent plusieurs objets associés. La sélection dans un objet peut modifier ce qui est affiché ou mis en évidence dans l'autre. Par exemple, la sélection d'une ligne dans un tableau peut mettre en évidence une zone dans une carte ou afficher un graphique pour une catégorie différente.

Les objets de sortie interactive ne gèrent pas les fonctions d'édition (comme la modification du texte, des couleurs, des polices ou des bordures du tableau). Les objets individuels peuvent être copiés depuis l'objet interactif vers le Visualiseur. Les tableaux copiés depuis la sortie interactive peuvent être modifiés dans l'éditeur de tableaux croisés.

Copie d'objets depuis la sortie interactive

Fichier>**Copier dans le Visualiseur** copie les objets de sortie sélectionnés vers la fenêtre Visualiseur.

- Les options disponibles dépendent du contenu de la sortie interactive.
- **Graphique** et **Mappe** créent des objets graphiques.

- **Tableau** crée un tableau croisé qui peut être modifié dans l'éditeur de tableaux croisés.
- **Instantané** crée une image de la vue en cours.
- **Modèle** crée une copie de l'objet de sortie interactive en cours.

Editer>Copier l'objet copie dans le presse-papiers l'objet de sortie interactive sélectionné.
- Le collage de l'objet copié dans le Visualiseur est équivalent à **Fichier>Copier dans le Visualiseur**.
- Le collage de l'objet dans une autre application l'insère en tant qu'image.

Zoom et panoramique

Pour les cartes, vous pouvez utiliser **Vue>Zoom** pour effectuer un zoom sur la vue de la carte. Depuis une vue en zoom d'une carte, vous pouvez utiliser **Vue>Panoramique** pour faire bouger la vue.

Paramètres d'impression

Fichier>Paramètres d'impression détermine comment imprimer les objets interactifs.
- **Imprimer la vue visible uniquement**. N'imprime que la vue affichée actuellement. Il s'agit de l'option par défaut.
- **Imprimer toutes les vues**. Imprime toutes les vues contenues dans la sortie interactive.
- L'option sélectionnée détermine également l'action par défaut pour exportation de l'objet de sortie.

Exportation de sortie

L'option Exporter sortie enregistre la sortie Visualiseur aux formats HTML, texte, Word/RTF, Excel, PowerPoint (requiert PowerPoint 97 ou version ultérieure), et PDF. Les graphiques peuvent également être exportés dans différents formats graphiques.

Remarque : L'option Exporter vers PowerPoint n'est disponible que sous les systèmes d'exploitation Windows.

Exportation de la sortie

1. Cliquez n'importe où dans la boîte de dialogue pour définir le Visualiseur comme actif.
2. Cliquez sur le bouton **Exporter** sur la barre d'outils ou cliquez avec le bouton droit de la souris dans la fenêtre de sortie et sélectionnez **Exporter**.
3. Entrez un nom de fichier (ou un préfixe pour les graphiques) et sélectionnez un format d'exportation.

Objets à exporter : Vous pouvez exporter tous les objets du Visualiseur. tous les objets visibles ou seulement les objets sélectionnés.

Type de document : Les options suivantes sont disponibles :
- **Word/RTF (*.doc).** Les tableaux croisés sont exportés sous forme de tableaux Word en préservant tous les attributs de mise en forme (par exemple, les bordures de cellule, les styles de police et les couleurs d'arrière-plan). La sortie texte est exportée au format RTF. Les graphiques, diagrammes d'arbre et les vues de modèle sont inclus au format PNG. Notez que Microsoft Word peut ne pas afficher correctement les tableaux de très grande largeur.
- **Excel 97-2004 (*.xls)/Excel 2007 et version ultérieure (*.xlsx).** Les lignes, colonnes et cellules de tableaux croisés sont exportées sous forme de lignes, colonnes et cellules Excel, en préservant tous leurs attributs de mise en forme (par exemple, bordures de cellule, styles de police et couleur d'arrière-plan). La sortie texte est exportée avec tous les attributs de police. Chaque ligne dans la sortie texte constitue une ligne dans le fichier Excel, le contenu complet de la ligne résidant dans une seule cellule. Les graphiques, diagrammes d'arbre et les vues de modèle sont inclus au format PNG. La sortie peut être exporté au format *Excel 97-2004* ou *Excel 2007 et version ultérieure*.

- **HTML (*.htm).** Les tableaux croisés dynamiques sont exportés comme des tableaux HTML. Les sorties texte sont publiées au format HTML pré-formaté. Les graphiques, les diagrammes d'arbre et les vues de modèle sont imbriqués dans le document au format graphique sélectionné. Un navigateur compatible avec HTML 5 est requis pour l'affichage d'une sortie exportée sous le format HTML.
- **PDF (*.pdf).** Toutes les sorties sont exportées tels qu'elles apparaissent dans l'aperçu avant impression, avec tous les attributs de formatage intacts.
- **Text - Plain/UTF8/UTF16 (*.txt).** Les formats de sortie texte comprennent les formats texte normal, UTF-8 et UTF-16. Les tableaux croisés dynamiques peuvent être exportés en format tabulé ou avec espaces. Toutes les sorties texte sont exportées en format avec espaces. Pour les graphiques, les vues d'arborescence et les vues de modèle, une ligne indiquant le nom du fichier image est insérée dans le fichier texte pour chaque graphique.
- **Aucun (graphiques uniquement) :** Les formats d'exportation disponibles sont les suivants : EPS, JPEG, TIFF, PNG et BMP. Sous les systèmes d'exploitation Windows, le format EMF (métafichier amélioré) est également disponible.

Ouvrir le dossier conteneur : Ouvre le dossier qui contient les fichiers créés par l'exportation.

Options HTML

L'exportation au format HTML nécessite un navigateur compatible HTML 5.

Les options suivantes sont disponibles pour exporter les sorties au format HTML :

Couches des tableaux croisés dynamiques : Par défaut, l'inclusion ou l'exclusion de couches de tableaux croisés dynamiques est contrôlée par les propriétés du tableau pour chaque tableau croisé dynamique. Vous pouvez remplacer ce paramètre et inclure toutes les couches, ou les exclure toutes sauf la couche actuellement visible. Pour plus d'informations, voir la rubrique «Propriétés du tableau : Impression», à la page 123.

Exporter les tableaux en couches comme étant interactifs : Les tableaux en couches sont affichés tels qu'ils apparaissent dans le Visualiseur et vous pouvez modifier de manière interactive la couche affichée dans le navigateur. Si cette option n'est pas sélectionnée, chaque couche de tableau est affichée en tant que tableau séparé.

Tableaux au format HTML : Cette option contrôle les informations de style incluses pour les tableaux croisés dynamiques exportés.

- **Exporter avec les styles et la largeur de colonne fixe :** Toutes les informations de style du tableau croisé (polices, couleurs d'arrière-plan, etc.) et de largeur de colonne sont préservées.
- **Exporter sans les styles :** Les tableaux croisés dynamiques sont convertis en tableaux HTML par défaut. Aucun attribut de style n'est conservé. La largeur des colonnes est automatiquement déterminée.

Inclure des notes de bas de page et des légendes : Contrôle l'inclusion ou l'exclusion des notes de bas de page et légendes de tous les tableaux croisés dynamiques.

Vues de modèles : Par défaut, l'inclusion ou l'exclusion de vues de modèles est contrôlée par les propriétés du modèle, pour chaque modèle. Vous pouvez remplacer ce paramètre et inclure toutes les vues, ou les exclure toutes sauf la vue actuellement visible. Pour plus d'informations, voir la rubrique Propriétés du modèle. (Remarque : toutes les vues de modèles, y compris les tableaux, sont exportées comme graphiques.)

Remarque : Pour le format HTML, vous pouvez également contrôler le format de fichier image pour les graphiques exportés. Pour plus d'informations, voir la rubrique «Options de format des graphiques», à la page 111.

Pour définir les options d'exportation HTML

1. Sélectionnez **HTML** comme format d'exportation.
2. Cliquez sur **Change Options**.

Options du rapport Web

Un rapport Web est un document interactif compatible avec la plupart des navigateurs. La plupart des fonctions interactives associées aux tableaux croisés dynamiques qui sont disponibles dans le Visualiseur sont également disponibles dans les rapports Web.

Titre du rapport : Titre affiché dans l'en-tête du rapport. Par défaut, il s'agit du nom de fichier. Vous pouvez indiquer un titre personnalisé à la place d'un nom de fichier.

Format : Il existe deux options pour le format des rapports :

- **Rapport Web SPSS (HTML 5) :** Ce format requiert un navigateur compatible HTML 5.
- **Rapport actif Cognos (mht) :** Ce format requiert un navigateur qui prend en charge les fichiers au format MHT ou l'application des rapports actifs Cognos.

Exclure des objets : Vous pouvez exclure les types d'objet sélectionnés du rapport :

- **Texte :** Objets de texte qui ne correspondent pas à des journaux. Cette option inclut les objets de texte qui contiennent des informations sur le jeu de données actif.
- **Journaux :** Objets texte répertoriant la syntaxe de commande exécutée. Les éléments du journal incluent également des messages d'avertissement et d'erreur rencontrés par des commandes qui ne génèrent pas de sortie dans le Visualiseur.
- **Tableaux de remarques :** La sortie des procédures statistiques et de génération de graphiques inclut un tableau Remarques. Ce tableau contient des informations sur le jeu de données utilisé, les valeurs manquantes et la syntaxe de commande utilisée pour exécuter la procédure.
- **Messages d'erreur et d'avertissement :** Messages d'erreur et d'avertissement associés aux procédures statistiques ou de représentation graphique.

Restyle the tables and charts to match the Web Report. Cette option applique le style de rapport Web standard à tous les tableaux et graphiques. Elle prévaut sur les polices, couleurs ou autres styles dans la sortie telle que l'affiche le Visualiseur. Vous ne pouvez pas modifier le style de rapport Web standard.

Web Server Connection. Vous pouvez inclure l'URL d'emplacement d'un ou de plusieurs serveurs d'application exécutant IBM SPSS Statistics Web Report Application Server. Le serveur d'application Web dispose de fonctionnalités permettant de faire pivoter les tableaux, de modifier les graphiques et d'enregistrer les rapports Web modifiés.

- Sélectionnez **Utiliser** pour chaque serveur d'application que vous désirez inclure dans le rapport Web.
- Si un rapport Web contient une spécification d'URL, il se connecte à ce serveur d'application pour offrir ses fonctions d'édition supplémentaires.
- Si vous spécifiez plusieurs URL, le rapport Web tente de se connecter à chaque serveur dans l'ordre où ils ont été spécifiés.

Le serveur IBM SPSS Statistics Web Report Application Server peut être téléchargé depuis le site http://www.ibm.com/developerworks/spssdevcentral.

Options Word/RTF

Les options suivantes sont disponibles pour exporter les sorties au format Word :

Couches des tableaux croisés dynamiques : Par défaut, l'inclusion ou l'exclusion de couches de tableaux croisés dynamiques est contrôlée par les propriétés du tableau pour chaque tableau croisé dynamique.

Vous pouvez remplacer ce paramètre et inclure toutes les couches, ou les exclure toutes sauf la couche actuellement visible. Pour plus d'informations, voir la rubrique «Propriétés du tableau : Impression», à la page 123.

Larges tableaux croisés dynamiques : Contrôle le traitement des tableaux trop larges pour la largeur définie du document. Par défaut, le tableau est réduit aux bonnes dimensions. Le tableau est divisé en sections et les libellés de ligne sont répétés pour chaque section du tableau. Vous pouvez également réduire ou modifier les tableaux larges et les autoriser à s'étendre au-delà de la largeur définie du document.

Conserver les points de rupture : Si vous avez défini des points de rupture, ces paramètres seront conservés dans les tableaux Word.

Inclure des notes de bas de page et des légendes : Contrôle l'inclusion ou l'exclusion des notes de bas de page et légendes de tous les tableaux croisés dynamiques.

Vues de modèles : Par défaut, l'inclusion ou l'exclusion de vues de modèles est contrôlée par les propriétés du modèle, pour chaque modèle. Vous pouvez remplacer ce paramètre et inclure toutes les vues, ou les exclure toutes sauf la vue actuellement visible. Pour plus d'informations, voir la rubrique Propriétés du modèle. (Remarque : toutes les vues de modèles, y compris les tableaux, sont exportées comme graphiques.)

Mise en page pour l'exportation : Ceci ouvre une boîte de dialogue qui vous permet de définir la taille et les marges de la page du document exporté. La largeur du document utilisée pour déterminer le comportement d'ajustement et de réduction est la largeur de la page moins les marges de gauche et de droite.

Pour définir les options d'exportation Word

1. Sélectionnez **Word/RTF** comme format d'exportation.
2. Cliquez sur **Change Options**.

Options Excel

Les options suivantes sont disponibles pour exporter les sorties au format Excel :

Créez une feuille de calcul ou un classeur, ou modifiez une feuille de calcul existante : Par défaut, un nouveau classeur est créé. Si un fichier portant le même nom existe, il sera écrasé. Si vous choisissez de créer une nouvelle feuille de calcul, si une feuille de calcul portant le même nom existe dans le fichier spécifié, elle sera écrasée. Si vous choisissez de modifier une feuille de calcul existante, vous devez en spécifier le nom. (Cela est facultatif si vous créez une feuille de calcul). Les noms de feuille de calcul ne peuvent pas dépasser 31 caractères et ne doivent pas comporter de barre oblique, de barre oblique inversée, de crochet, de point d'interrogation ou d'astérisque.

Lors de l'exportation vers Excel 97-2004, si vous modifiez une feuille de calcul existante, les graphiques, les modèles de vue et les diagrammes d'arbre ne seront pas inclus dans les sorties exportées.

Emplacement dans la feuille de calcul : Contrôle l'emplacement dans la feuille de calcul destiné aux sorties exportées. Par défaut, les sorties exportés seront ajoutés après la dernière colonne à contenu, en partant de la première ligne, sans modifier le contenu existant. C'est une bonne solution pour ajouter de nouvelles colonnes à une feuille de calcul existante. L'ajout des sorties exportées après la dernière ligne est une bonne solution pour ajouter de nouvelles lignes à une feuille de calcul existante. L'ajout des sorties exportées en partant d'un emplacement de cellule spécifique écrasera le contenu dans la zone d'ajout des sorties exportées.

Couches des tableaux croisés dynamiques : Par défaut, l'inclusion ou l'exclusion de couches de tableaux croisés dynamiques est contrôlée par les propriétés du tableau pour chaque tableau croisé dynamique.

Vous pouvez remplacer ce paramètre et inclure toutes les couches, ou les exclure toutes sauf la couche actuellement visible. Pour plus d'informations, voir la rubrique «Propriétés du tableau : Impression», à la page 123.

Inclure des notes de bas de page et des légendes : Contrôle l'inclusion ou l'exclusion des notes de bas de page et légendes de tous les tableaux croisés dynamiques.

Vues de modèles : Par défaut, l'inclusion ou l'exclusion de vues de modèles est contrôlée par les propriétés du modèle, pour chaque modèle. Vous pouvez remplacer ce paramètre et inclure toutes les vues, ou les exclure toutes sauf la vue actuellement visible. Pour plus d'informations, voir la rubrique Propriétés du modèle. (Remarque : toutes les vues de modèles, y compris les tableaux, sont exportées comme graphiques.)

Pour définir les options d'exportation Excel

1. Sélectionnez **Excel** comme format d'exportation.
2. Cliquez sur **Change Options**.

Options PowerPoint

Les options suivantes sont disponibles pour PowerPoint :

Couches des tableaux croisés dynamiques : Par défaut, l'inclusion ou l'exclusion de couches de tableaux croisés dynamiques est contrôlée par les propriétés du tableau pour chaque tableau croisé dynamique. Vous pouvez remplacer ce paramètre et inclure toutes les couches, ou les exclure toutes sauf la couche actuellement visible. Pour plus d'informations, voir la rubrique «Propriétés du tableau : Impression», à la page 123.

Larges tableaux croisés dynamiques : Contrôle le traitement des tableaux trop larges pour la largeur définie du document. Par défaut, le tableau est réduit aux bonnes dimensions. Le tableau est divisé en sections et les libellés de ligne sont répétés pour chaque section du tableau. Vous pouvez également réduire ou modifier les tableaux larges et les autoriser à s'étendre au-delà de la largeur définie du document.

Inclure des notes de bas de page et des légendes : Contrôle l'inclusion ou l'exclusion des notes de bas de page et légendes de tous les tableaux croisés dynamiques.

Utiliser les entrées de légendes du Visualiseur comme titres de diapositive : Inclut un titre sur chaque diapositive créée par l'exportation. Chaque diapositive contient un élément unique exporté depuis le Visualiseur. Le titre est formé à partir de l'entrée correspondant à l'élément dans le panneau de légende du Visualiseur.

Vues de modèles : Par défaut, l'inclusion ou l'exclusion de vues de modèles est contrôlée par les propriétés du modèle, pour chaque modèle. Vous pouvez remplacer ce paramètre et inclure toutes les vues, ou les exclure toutes sauf la vue actuellement visible. Pour plus d'informations, voir la rubrique Propriétés du modèle. (Remarque : toutes les vues de modèles, y compris les tableaux, sont exportées comme graphiques.)

Mise en page pour l'exportation : Ceci ouvre une boîte de dialogue qui vous permet de définir la taille et les marges de la page du document exporté. La largeur du document utilisée pour déterminer le comportement d'ajustement et de réduction est la largeur de la page moins les marges de gauche et de droite.

Pour définir les options d'exportation PowerPoint

1. Sélectionnez **PowerPoint** comme format d'exportation.
2. Cliquez sur **Change Options**.

Remarque : L'option Exporter vers PowerPoint n'est disponible que sous les systèmes d'exploitation Windows.

Options PDF

Les options suivantes sont disponibles pour le format PDF :

Intégrer les signets : Cette option inclut dans le document PDF des signets qui correspondent aux entrées de légendes. Tout comme le panneau de légende du Visualiseur, les signets peuvent faciliter considérablement la navigation parmi les documents comportant un nombre élevé d'objets de sortie.

Intégrer les polices : L'option d'intégration des polices rend le document PDF identique sur tous les ordinateurs. Dans le cas contraire, si certaines polices utilisées dans le document ne sont pas disponibles sur l'ordinateur où le document PDF est affiché (ou imprimé), la substitution de polices risque de générer des résultats inférieurs.

Couches des tableaux croisés dynamiques : Par défaut, l'inclusion ou l'exclusion de couches de tableaux croisés dynamiques est contrôlée par les propriétés du tableau pour chaque tableau croisé dynamique. Vous pouvez remplacer ce paramètre et inclure toutes les couches, ou les exclure toutes sauf la couche actuellement visible. Pour plus d'informations, voir la rubrique «Propriétés du tableau : Impression», à la page 123.

Vues de modèles : Par défaut, l'inclusion ou l'exclusion de vues de modèles est contrôlée par les propriétés du modèle, pour chaque modèle. Vous pouvez remplacer ce paramètre et inclure toutes les vues, ou les exclure toutes sauf la vue actuellement visible. Pour plus d'informations, voir la rubrique Propriétés du modèle. (Remarque : toutes les vues de modèles, y compris les tableaux, sont exportées comme graphiques.)

Pour définir les options d'exportation PDF

1. Sélectionnez **PDF** comme format d'export.
2. Cliquez sur **Change Options**.

Autres paramètres affectant la sortie PDF

Mise en page/Attributs de page : La taille de page, l'orientation, les marges, le contenu et l'affichage des en-têtes et bas de page, et la taille des graphiques imprimés dans les documents PDF sont contrôlés par les options de mise en page et d'attribut de page.

Propriétés du tableau/Modèles de tableaux : Le redimensionnement de tableaux larges et/ou longs et l'impression des couches de tableau sont contrôlés par les propriétés de chaque tableau. Ces propriétés peuvent également être enregistrées dans Modèles de tableaux.

Imprimante par défaut/active : La résolution (PPP) du document PDF correspond au paramètre de résolution en cours de l'imprimante par défaut ou sélectionnée (qui peut être changée à l'aide de l'option Mise en page). La résolution maximale est de 1 200 ppp. Si le paramètre de résolution de l'imprimante est supérieur, la résolution du document PDF sera de 1 200 ppp.

Remarque : Une haute résolution peut générer des résultats médiocres si les documents sont imprimés sur des imprimantes de résolution inférieure.

Options texte

Les options suivantes sont disponibles pour l'exportation de texte :

Format des tableaux croisés dynamiques : Les tableaux croisés dynamiques peuvent être exportés en format tabulé ou avec espaces. Pour le format avec espaces, vous pouvez également contrôler :

- **Largeur des colonnes : Ajustement automatique** ne permet pas un renvoi à la ligne du contenu des colonnes, et chacune d'entre elles est aussi large que le plus grand libellé ou la plus grande valeur de cette colonne. **Personnalisé** permet de définir une largeur de colonne maximale qui est appliquée à toutes les colonnes de la table, et les valeurs qui dépassent cette largeur sont renvoyées à la ligne suivante de cette colonne.
- **Caractère de bordure de ligne/colonne :** Permet de contrôler les caractères utilisés pour créer des bordures de ligne et de colonne. Pour supprimer l'affichage de ces bordures de ligne et de colonne , indiquez des espaces vides pour les valeurs.

Couches des tableaux croisés dynamiques : Par défaut, l'inclusion ou l'exclusion de couches de tableaux croisés dynamiques est contrôlée par les propriétés du tableau pour chaque tableau croisé dynamique. Vous pouvez remplacer ce paramètre et inclure toutes les couches, ou les exclure toutes sauf la couche actuellement visible. Pour plus d'informations, voir la rubrique «Propriétés du tableau : Impression», à la page 123.

Inclure des notes de bas de page et des légendes : Contrôle l'inclusion ou l'exclusion des notes de bas de page et légendes de tous les tableaux croisés dynamiques.

Vues de modèles : Par défaut, l'inclusion ou l'exclusion de vues de modèles est contrôlée par les propriétés du modèle, pour chaque modèle. Vous pouvez remplacer ce paramètre et inclure toutes les vues, ou les exclure toutes sauf la vue actuellement visible. Pour plus d'informations, voir la rubrique Propriétés du modèle. (Remarque : toutes les vues de modèles, y compris les tableaux, sont exportées comme graphiques.)

Pour définir les options d'exportation de texte

1. Sélectionnez **Texte** comme format d'exportation.
2. Cliquez sur **Change Options**.

Options graphiques uniquement

Les options suivantes sont disponibles pour l'exportation de graphiques uniquement :

Vues de modèles : Par défaut, l'inclusion ou l'exclusion de vues de modèles est contrôlée par les propriétés du modèle, pour chaque modèle. Vous pouvez remplacer ce paramètre et inclure toutes les vues, ou les exclure toutes sauf la vue actuellement visible. Pour plus d'informations, voir la rubrique Propriétés du modèle. (Remarque : toutes les vues de modèles, y compris les tableaux, sont exportées comme graphiques.)

Options de format des graphiques

Pour les documents HTML et texte ainsi que pour l'exportation de graphiques uniquement, vous pouvez sélectionner le format graphique, et, pour chaque format graphique, vous pouvez contrôler divers paramètres facultatifs.

Pour sélectionner le format graphique et des options pour les graphiques exportés :

1. Sélectionnez **HTML**, **Text** ou **None (Graphics only)** comme type de document.
2. Sélectionnez un type de fichier graphique dans la liste déroulante proposée.
3. Cliquez sur **Change Options** pour modifier les options du format de fichier graphique sélectionné.

Options d'exportation de graphiques au format JPEG

Taille d'image (%)
Pourcentage de la taille d'origine du graphique, jusqu'à 200 pour cent.

Convertir en niveaux de gris
Convertit les couleurs en nuances de gris.

Options d'exportation de graphiques au format BMP

Taille d'image (%)
Pourcentage de la taille d'origine du graphique, jusqu'à 200 pour cent.

Compresser l'image pour réduire la taille du fichier
Technique de compression sans perte qui crée des fichiers plus petits sans affecter la qualité de l'image.

Options d'exportation de graphiques au format PNG

Taille d'image (%)
Pourcentage de la taille d'origine du graphique, jusqu'à 200 pour cent.

Profondeur des couleurs
Détermine le nombre de couleurs du graphique exporté. Quelle que soit la profondeur de couleur utilisée lors de l'enregistrement d'un graphique, le nombre de couleurs du graphique est compris entre le nombre de couleurs réellement utilisées et le nombre maximal de couleurs autorisé par la profondeur. Par exemple, si vous enregistrez un graphique contenant trois couleurs (rouge, blanc et noir) en mode 16 couleurs, ce graphique conserve ses trois couleurs.

Remarque : Lorsque le nombre de couleurs du graphique est supérieur au nombre prévu pour cette profondeur, les couleurs sont converties de manière à reproduire celles du graphique.

Options d'exportation de graphiques aux formats EMF et TIFF

Taille d'image : Pourcentage de la taille d'origine du graphique, jusqu'à 200 pour cent.

Remarque : Le format EMF (métafichier amélioré) n'est disponible que sous les systèmes d'exploitation Windows.

Options d'exportation de graphiques au format EPS

Taille d'image : Vous pouvez spécifier la taille sous forme de pourcentage de la taille d'origine de l'image (jusqu'à 200 pour cent) ou vous pouvez spécifier la largeur d'une image en pixels (la hauteur étant déterminée par la valeur de largeur et le rapport d'aspect). L'image exportée est toujours proportionnelle à celle d'origine.

Inclure une image d'aperçu TIFF : Enregistre un aperçu avec l'image EPS au format TIFF pour un affichage dans les applications qui ne peuvent pas afficher les images EPS à l'écran.

Polices : Contrôle le traitement des polices dans les images EPS.

- **Utiliser des polices de référence :** Si les polices utilisées dans le graphique sont disponibles sur le périphérique de sortie, elles sont utilisées. Sinon, le périphérique de sortie utilise des polices alternatives.
- **Remplacer les polices avec courbes :** Transforme les polices en données de courbe PostScript. Vous ne pouvez plus modifier le texte en tant que tel dans les applications de retouche des graphiques EPS. Cette option est utile si les polices utilisées dans le graphique ne sont pas disponibles sur le périphérique de sortie.

Impression du Visualiseur

Il y a deux options pour l'impression du contenu de la fenêtre du Visualiseur.:

Toutes les sortie affichées : Seuls s'impriment les éléments affichés dans le panneau de contenu. Les éléments masqués (éléments accompagnés d'une icône représentant un livre fermé dans le panneau de ligne de légende ou masqués dans les couches réduites des lignes de légende) ne s'impriment pas.

Sélection : Seuls s'impriment les éléments sélectionnés dans les panneaux de ligne de légende et/ou de contenu.

Impression des sorties et des graphiques

1. Rendez la fenêtre du Visualiseur active (cliquez n'importe où dans la fenêtre).
2. A partir des menus, sélectionnez :

   Fichier > **Imprimer...**
3. Choisissez les paramètres d'impression voulus.
4. Cliquez sur **OK** pour imprimer.

Aperçu avant impression

L'Aperçu avant impression vous montre ce qui va s'imprimer sur chaque page des documents du Visualiseur. Il est conseillé d'utiliser l'Aperçu avant impression avant d'imprimer réellement un document du Visualiseur car cela permet de détecter les éléments qui ne seront pas visibles, tout simplement en regardant le panneau de contenu du Visualiseur. notamment :

- Les ruptures de page
- Les couches masquées des tableaux croisés dynamiques
- Les ruptures dans les tableaux de grande largeur
- Les en-têtes et pieds de page imprimés sur chaque page

Si des sorties sont sélectionnées dans le Visualiseur. l'aperçu n'affiche que les sorties sélectionnées. Pour avoir un aperçu de tous les sorties, vérifiez que rien n'est sélectionné dans le Visualiseur.

Attributs de page : En-têtes et pieds de page

Les en-têtes et pieds de page sont les informations qui s'impriment en haut et en bas de chaque page. Vous pouvez saisir le texte à afficher dans les en-têtes et les pieds de page. Vous pouvez également utiliser la barre d'outils qui se trouve au milieu de la boîte de dialogue pour insérer :

- Date et heure
- Des numéros de page
- Le nom de fichier du Visualiseur
- Les libellés des en-têtes de légende
- Les titres et les sous-titres
- **Définir par défaut** utilise les paramètres indiqués ici comme paramètres par défaut pour les nouveaux documents du Visualiseur. (Remarque : Cela définit les paramètres en cours à la fois sur l'onglet En-tête/Pied de page et l'onglet Options comme paramètres par défaut.)
- Les libellés d'en-tête des lignes de légende indiquent l'en-tête de ligne de légende de premier, deuxième, troisième et/ou quatrième niveau pour le premier élément de chaque page.
- Les titres et sous-titres de page impriment les titres et sous-titres de la page en cours. Ils peuvent être créés avec les options Nouveau titre de page du menu Insérer du Visualiseur ou avec les commandes `TITLE` et `SUBTITLE`. Si vous n'avez indiqué aucun titre ou sous-titre de page, ce paramètre est ignoré.

  Remarque : Les caractéristiques de police des nouveaux titres de page et des sous-titres sont vérifiées dans l'onglet Visualiseur de la boîte de dialogue Options (accessible en choisissant Options dans le menu Edition). Les caractéristiques de police des titres de page et des sous-titres existants peuvent être modifiées en changeant les titres à l'aide du Visualiseur.

Pour voir comment vos en-têtes et pieds de page vont apparaître sur la page imprimée, choisissez Aperçu avant impression dans le menu Fichier.

Pour insérer des en-têtes et des pieds de page

1. Rendez la fenêtre du Visualiseur active (cliquez n'importe où dans la fenêtre).
2. Dans les menus, sélectionnez :

   Fichier > **En-tête et pied de page...**
3. Saisissez l'en-tête et/ou le pied de page devant figurer sur chaque page.

Attributs de page : Options

Cette boîte de dialogue contrôle la taille du graphique imprimé, l'espace entre les éléments de sortie imprimés, ainsi que la numérotation des pages.

- **Taille imprimée graphique :** Contrôle la taille du graphique imprimé par rapport à la taille définie pour la page. Le rapport d'aspect du graphique n'est pas affecté par la taille du graphique imprimé. La taille globale d'un graphique imprimé est limitée par sa hauteur et par sa largeur. Lorsque les bordures externes d'un graphique atteignent la bordure gauche et droite de la page, la taille du graphique ne peut pas croître en hauteur.
- **Espace entre éléments :** Contrôle l'espace entre des éléments imprimés. Chaque tableau croisé dynamique, graphique et objet de texte constitue un élément distinct. Ce paramètre n'affecte pas l'affichage des éléments dans le Visualiseur.
- **Commencer page :** Numéros de page consécutifs commençant à partir d'un numéro spécifié.
- **Définir par défaut :** Cette option utilise les paramètres définis ici comme paramètres par défaut pour les nouveaux documents Visualiseur. (Notez que cette option définit les paramètres d'en-tête et de pied de page en cours comme paramètres par défaut).

Modification de la taille du graphique imprimé, de la numérotation des pages et de l'espace entre les éléments imprimés

1. Rendez la fenêtre du Visualiseur active (cliquez n'importe où dans la fenêtre).
2. A partir des menus, sélectionnez :

   Fichier > **Attributs de page...**
3. Cliquez sur l'onglet **Options**.
4. Modifiez les paramètres et cliquez sur **OK**.

Enregistrement des sorties

Le contenu du Visualiseur peut être enregistré.

- **Objet en sortie (*.cou)**. Ce format enregistre le conteneur de sortie complet, y-compris le graphique, les tabulations, les annotations, etc. Ce format peut être ouvert et lu dans IBM SPSS Modeler, ajouté à des projets, et publié et suivi à l'aide d'IBM SPSS Collaboration and Deployment Services Repository. Ce format n'est pas compatible avec IBM SPSS Statistics.
- **Fichiers Visualiseur (*.spv)**. Format utilisé pour afficher les fichiers dans la fenêtre du Visualiseur . Lorsque vous enregistrez sous ce format depuis un nugget de modèle dans IBM SPSS Modeler, seul le contenu du Visualiseur dans l'onglet **Modèle** est enregistré.

Pour contrôler les options d'enregistrement des rapports Web ou enregistrer les résultats dans un autre format (par exemple, texte, Word, Excel), utilisez **Exporter** dans le menu **Fichier**.

Enregistrement d'un document Visualiseur

1. A partir du menu de la fenêtre Visualiseur, sélectionnez :

   Fichier > **Enregistrer**
2. Entrez le nom du document et cliquez sur **Enregistrer**.

   Vous pouvez au besoin procéder aux actions suivantes :

Verrouiller des fichiers pour empêcher leur modification dans IBM SPSS Smartreader

Si un document du Visualiseur est verrouillé, vous pouvez manipuler les tableaux croisés (permuter les lignes et les colonnes, modifier la couche affichée, etc.), mais vous ne pouvez pas modifier la sortie ou enregistrer vos modifications du document du Visualiseur dans IBM SPSS Smartreader (produit distinct pour exploitation des documents du Visualiseur). Ce paramètre est sans effet sur les documents du Visualiseur ouverts dans IBM SPSS Statistics ou dans IBM SPSS Modeler.

Chiffrer les fichiers avec un mot de passe

Vous pouvez protéger les informations confidentielles stockées dans un document du Visualiseur en chiffrant le document à l'aide d'un mot de passe. Une fois chiffré, le document ne peut être ouvert qu'en soumettant le mot de passe. Les utilisateurs d'IBM SPSS Smartreader devront également soumettre le mot de passe pour ouvrir le fichier.

Pour chiffrer un document du Visualiseur.:

a. Sélectionnez **Chiffrer le fichier avec un mot de passe** dans la boîte de dialogue Enregistrer les sorties sous.
b. Cliquez sur **Enregistrer**.
c. Dans la boîte de dialogue Chiffrer le fichier, indiquez un mot de passe de votre choix puis entrez-le à nouveau dans le champ Confirmer le mot de passe. Les mots de passe ne peuvent contenir plus de 10 caractères et sont sensibles à la casse.

Avertissement : Les mots de passe ne peuvent pas être récupérés s'ils sont perdus. Le fichier ne peut pas être ouvert si le mot de passe associé est perdu.

Création de mots de passe fiables

- Utilisez huit caractères ou plus.
- Utilisez des chiffres, des symboles et même des signes de ponctuation dans votre mot de passe.
- Evitez les nombres ou caractères qui se suivent, comme `"123"` et `"abc"`, et évitez les répétitions de telles séquences, comme `"111aaa"`.
- Ne créez pas de mots de passe qui utilisent des informations personnelles telles que des dates d'anniversaire ou des surnoms.
- Modifiez régulièrement votre mot de passe.

Remarque : Le stockage de fichiers chiffrés dans un IBM SPSS Collaboration and Deployment Services Repository n'est pas pris en charge.

Modification des fichiers chiffrés

- Si vous ouvrez un fichier chiffré, effectuez des modifications dans celui-ci, puis sélectionnez Fichier > Enregistrer, le fichier modifié sera enregistré avec le même mot de passe.
- Vous pouvez changer le mot de passe d'un fichier chiffré en l'ouvrant, puis en répétant les étapes de chiffrement ci-dessous, et en indiquant un mot de passe différent dans la boîte de dialogue Chiffrer le fichier.
- Vous pouvez enregistrer une version non chiffrée d'un fichier chiffré en ouvrant le fichier, puis en sélectionnant Fichier > Enregistrer sous, puis en décochant l'option **Chiffrer le fichier avec un mot de passe** dans la boîte de dialogue Enregistrer les sorties sous.

Remarque : Les fichiers de données et les documents de sortie chiffrés ne peuvent pas être ouverts dans des versions IBM SPSS Statistics antérieures à la version 21. Les fichiers de syntaxe chiffrés ne peuvent pas être ouverts dans les versions antérieures à la version 22.

Store required model information with the output document

Cette option ne s'applique que si des éléments du Visualiseur de modèle dans le document de sortie requièrent des informations supplémentaires pour activer certaines fonctionnalités interactives. Cliquez sur **Plus d'informations** pour afficher la liste de ces éléments et les fonctionnalités interactives nécessitant des informations supplémentaires. Si vous stockez ces

informations avec le document de sortie, la taille du fichier peut croître significativement. Si vous décidez de ne pas stocker ces informations, vous pourrez toujours ouvrir les éléments en sortie, mais les fonctionnalités interactives spécifiées ne seront pas alors disponibles.

Tableaux croisés

Tableaux croisés

De nombreux résultats sont présentés sous la forme de tableaux qui peuvent être croisés de manière interactive. Ce qui signifie que vous pouvez réorganiser les lignes, les colonnes et les couches.

Manipulation d'un tableau croisé dynamique

Les options de manipulation d'un tableau croisé dynamique sont les suivantes :

- Transposition des lignes et des colonnes
- Déplacement des lignes et des colonnes
- Création de couches multidimensionnelles
- Association et dissociation des lignes et des colonnes
- Affichage et masquage des lignes, des colonnes et d'autres informations
- Rotation des libellés de ligne et de colonne
- Recherche de définitions de termes

Activation d'un tableau croisé

Avant de pouvoir manipuler ou modifier un tableau croisé, vous devez l'**activer**. Pour activer un tableau :

1. Double-cliquez sur le tableau.

   ou

2. Cliquez avec le bouton droit de la souris sur le tableau et sélectionnez l'option de menu **Modifier le contenu**.
3. Dans le sous-menu, sélectionnez **Dans le Visualiseur** ou **Dans une fenêtre séparée.**

Pivotement de tableau

Un tableau comporte trois dimensions : lignes, colonnes et couches. Une dimension peut contenir plusieurs éléments (ou aucun). Vous pouvez modifier l'organisation du tableau en déplaçant les éléments entre les dimensions ou à l'intérieur. Pour déplacer un élément, faites-le glisser puis déposez-le où vous le souhaitez.

Modification de l'ordre d'affichage des éléments dans une dimension

Pour modifier l'ordre d'affichage des éléments dans une dimension de table (ligne, colonne, ou couche) :

1. Si les plateaux de croisement ne sont pas déjà activés, dans le menu Tableau croisé dynamique sélectionnez :

   Tableau croisé dynamique > **Structure pivotante**
2. Faites glisser les éléments vers la dimension dans le plateau de croisement.

Déplacement de lignes et de colonnes dans un élément de dimension

1. Dans le tableau (et non dans les plateaux de croisement), cliquez sur le libellé de la ligne ou de la colonne à déplacer.
2. Faites glisser le libellé vers la nouvelle position.

Transposition des lignes et des colonnes

Si vous souhaitez simplement retourner les lignes et les colonnes, il existe une alternative simple à l'utilisation des plateaux de croisement :

1. A partir des menus, sélectionnez :

Tableau croisé dynamique > **Transposer lignes et colonnes**

Cette action a le même effet que de glisser tous les éléments de ligne dans la dimension colonne et glisser tous les éléments de colonnes dans la dimension ligne.

Groupement de lignes ou de colonnes

1. Sélectionnez les libellés de ligne ou de colonne que vous souhaitez grouper (cliquez sur les libellés et faites-les glisser ou utilisez la combinaison Maj+clic pour sélectionner plusieurs libellés).
2. A partir des menus, sélectionnez :

   Edition > **Groupe**

Un libellé de groupe est inséré automatiquement. Double-cliquez sur le libellé de groupe pour en modifier le texte.

Remarque : Pour ajouter des lignes ou des colonnes à un groupe existant, vous devez d'abord dissocier les éléments actuellement dans le groupe. Vous pouvez ensuite créer un nouveau groupe contenant les éléments supplémentaires.

Dissociation de lignes ou de colonnes

La dissociation supprime automatiquement le libellé de groupe.

Rotation des libellés de ligne ou de colonne

Vous pouvez faire pivoter des libellés entre l'affichage horizontal et vertical des libellés de colonne les plus internes et les libellés de ligne les plus externes dans un tableau.

1. A partir des menus, sélectionnez :

   Format > **Faire pivoter les libellés des colonnes internes**

   ou

   Format > **Faire pivoter les libellés des lignes externes**

Seules les libellés de colonne extérieure et les libellés de ligne intérieure peuvent être pivotées.

Tri des lignes

Pour trier les lignes d'un tableau croisé dynamique :

1. Activez le tableau.
2. Sélectionnez n'importe quelle cellule de la colonne à trier. Pour trier seulement un groupe de lignes sélectionné, choisissez au moins deux cellules contiguës dans la colonne à trier.
3. Dans les menus, sélectionnez :

   Editer > **Trier les lignes**
4. Sélectionnez l'ordre de tri **Croissant** ou **Décroissant** dans le sous-menu.

- Si la dimension de ligne contient des groupes, le tri n'affecte que le groupe contenant la sélection.
- Il est impossible de trier au sein du groupe.
- Il est impossible de trier des tableaux contenant plus d'un élément dans la dimension de ligne.

Insertion de lignes et de colonnes

Pour insérer une ligne ou une colonne dans un tableau croisé dynamique :

1. Activez le tableau.
2. Sélectionnez n'importe quelle cellule dans le tableau.
3. Dans les menus, sélectionnez :

   Insérer avant

   ou

   Insérer après

Dans le sous-menu, sélectionnez :

Ligne

ou

colonne

- Un signe plus (+) est inséré dans chaque cellule de la nouvelle ligne ou de la nouvelle colonne pour l'empêcher d'être automatiquement masquée du fait qu'elle est vide.
- Dans un tableau contenant des dimensions en couches ou imbriquées, une colonne ou une ligne est insérée à chaque niveau de dimension correspondant.

Contrôle de l'affichage des libellés de variable et de valeur

Si des variables contiennent des libellés de valeur ou de variable, vous pouvez contrôler l'affichage des noms et des libellés des variables, ainsi que les valeurs de données et les libellés de valeurs dans les tableaux croisés dynamiques.

1. Activer le tableau croisé dynamique.
2. Dans les menus, sélectionnez :

   Vue > **Libellés de variable**

   ou

   Vue > **Libellés de valeurs**
3. Sélectionnez une des options suivantes dans le sous-menu :

- **Nom** ou **Valeur**. Seuls les noms de variable (ou les valeurs) sont affichés. Les libellés descriptifs ne sont pas affichés.
- **Libellé**. Seuls les libellés descriptifs sont affichés. Les noms de variable (ou les valeurs) ne sont pas affichés.
- **Les deux**. Les noms (ou les valeurs), ainsi que les libellés descriptifs, sont affichés.

Modification de la langue de sortie

Pour modifier la langue de sortie d'un tableau croisé dynamique :

1. Activez le tableau
2. Dans les menus, sélectionnez :

   Vue > **Langue**
3. Sélectionnez l'une des langues disponibles.

La modification de la langue n'affecte que le texte généré par l'application, comme le titre des tableaux, les libellés de ligne et de colonnes, ou le texte des notes de bas de page. Les noms de variables et les libellés de variables descriptives et de valeurs ne sont pas affectés .

Navigation dans les grands tableaux

Pour utiliser la fenêtre de navigation pour les grands tableaux :

1. Activez le tableau.
2. A partir des menus, sélectionnez :

   Vue > **Navigation**

Annulation des modifications

Vous pouvez annuler les dernières modifications ou toutes les modifications d'un tableau croisé dynamique activé. Ces deux actions s'appliquent uniquement aux modifications apportées depuis l'activation la plus récente du tableau.

Pour annuler la dernière modification :

1. Dans les menus, sélectionnez :

   Editer > **Annuler**

Pour annuler toutes les modifications :

2. Dans les menus, sélectionnez :

   Editer > **Restaurer**

Utilisation des couches

Vous pouvez afficher un tableau bidimensionnel distinct pour chaque catégorie ou combinaison de catégories. Le tableau peut être considéré comme une superposition de couches, dont seule celle du dessus est visible.

Création et affichage de couches

Pour créer des couches :

1. Si des structures pivotantes ne sont pas encore activées, sélectionnez depuis le menu **Tableau croisé** :

   Tableau croisé dynamique > **Structure pivotante**

2. Faites glisser un élément de la dimension ligne ou colonne dans la dimension couche.

Le déplacement d'éléments vers la dimension couche crée un tableau multidimensionnel, mais seule une "coupe" bidimensionnelle est affichée. Le tableau visible est le tableau de la couche supérieure. Par exemple, si une variable catégorielle yes/no existe dans la dimension couche, le tableau multidimensionnel comporte deux couches : une pour la catégorie "yes" et l'autre pour la catégorie "no".

Modification de la couche affichée

1. Sélectionnez une catégorie dans la liste déroulante de couches (dans le tableau croisé dynamique et non dans le plateau de croisement).

Atteindre la catégorie de la couche

Aller à la catégorie de couche vous permet de modifier les couches d'un tableau croisé dynamique. Cette boîte de dialogue est particulièrement utile lorsqu'il existe plusieurs couches ou que la couche sélectionnée possède plusieurs catégories.

Affichage et masquage d'éléments

De nombreux types de cellules peuvent être masqués, notamment :

- Libellés de dimension
- Catégories, notamment la cellule de libellé et les cellules de données dans une ligne ou une colonne
- Libellés de catégorie (sans masquer les cellules de données)
- Notes de bas de page, titres et légendes

Masquage de lignes et de colonnes dans un tableau

Affichage des lignes et des colonnes masquées dans un tableau

1. A partir des menus, sélectionnez :

   Affichage > **Afficher toutes les catégories**

Cette option affiche toutes les lignes et les colonnes masquées du tableau. (Si l'option **Masquer les lignes et les colonnes vides** est sélectionnée pour ce tableau dans Propriétés du tableau, une ligne ou une colonne complètement vide reste masquée.)

Masquage et affichage des libellés de dimension

1. Sélectionnez le libellé de dimension ou tout étiquette de catégorie dans la dimension.
2. Dans le menu Affichage ou dans le menu contextuel, sélectionnez **Masquer le libellé de dimension** ou **Afficher le libellé de dimension**.

Masquage ou affichage des titres de tableau

Pour masquer un titre :

1. Activer le tableau croisé dynamique.
2. Sélectionnez le titre.
3. Dans le menu Affichage, choisissez **Masquer**.

   Pour afficher les titres masqués :
4. Dans le menu Affichage, choisissez **Montrer Tout**.

TableLooks

Un TableLook est un ensemble de propriétés qui définissent l'apparence d'une table. Vous pouvez sélectionner un TableLook prédéfini ou créer votre propre TableLook.

- Avant ou après l'application d'un TableLook, vous pouvez modifier les formats de cellules pour des cellules individuelles ou des groupes de cellules en utilisant les propriétés des cellules. Les formats de cellule modifiés restent intacts, même si vous appliquez un nouveau modèle de tableau.
- Vous pouvez également restaurer toutes les cellules conformément aux formats de cellule définis par le TableLook en cours. Cette action réinitialise toutes les cellules qui ont été modifiées. Si l'option **Tel qu'affiché** est sélectionnée dans la liste de fichiers TableLook, toutes les cellules modifiées sont réinitialisées suivant les propriétés en cours du tableau.
- Seules les propriétés de tableau définies dans la boîte de dialogue Propriétés de tableau sont enregistrées dans des modèles de tableau. Les modèles de tableaux ne comprennent pas les modifications des cellules individuelles.

Application d'un modèle de tableau

1. Activer le tableau croisé dynamique.
2. Dans les menus, sélectionnez :

   Format > **Modèles de tableaux...**
3. Sélectionnez un TableLook dans la liste de fichiers. Pour sélectionner un fichier dans un autre répertoire, cliquez sur **Parcourir**.
4. Cliquez sur **OK** pour appliquer le modèle de tableau au tableau croisé sélectionné.

Modification ou création d'un modèle de tableau

1. Dans la boîte de dialogue TableLook, sélectionnez un TableLook dans la liste de fichiers.
2. Cliquez sur **Modifier modèle**.
3. Modifiez les propriétés de tableau pour obtenir les attributs souhaités et cliquez sur **OK**.
4. Cliquez sur **Enregistrer modèle** pour enregistrer le modèle de tableau modifié ou sur **Enregistrer sous** pour l'enregistrer comme nouveau modèle.

- La modification d'un TableLook affecte uniquement le tableau croisé dynamique sélectionné. Un modèle de tableau n'est pas appliqué à d'autres tableaux qui utilisent ce modèle de tableau sauf si vous les sélectionnez et leur réappliquez le modèle de tableau.
- Seules les propriétés de tableau définissent dans la boîte de dialogue Propriétés de tableau sont enregistrées dans des modèles de tableau. Les modèles de tableaux ne comprennent pas les modifications des cellules individuelles.

Propriétés du tableau

Les Propriétés du tableau vous permettent de définir les propriétés générales d'un tableau, et de définir les styles de cellule pour plusieurs parties d'une table. Vous pouvez :

- Contrôler les propriétés générales, telles que le masquage des lignes ou des colonnes vides, et le réglage des propriétés d'impression.
- Contrôler le format et la position des marqueurs de note de bas de page.

- Déterminer des formats spécifiques pour les cellules de la zone de données, pour les libellés de ligne et de colonne, ou pour d'autres zones du tableau.
- Contrôler la largeur et la couleur des lignes de bordure de chaque zone du tableau.

Modification des propriétés d'un tableau croisé dynamique

1. A partir des menus, sélectionnez :

   Format > **Propriétés du tableau...**
2. Sélectionnez un onglet (**Général**, **Notes de bas de page**, **Formats de cellule**, **Bordures**, ou **Impression**).
3. Sélectionnez les options de votre choix.
4. Cliquez sur **OK** ou sur **Appliquer**.

Les nouvelles propriétés sont appliquées au tableau croisé dynamique sélectionné.

Propriétés du tableau : Général

De nombreuses propriétés s'appliquent à la table dans son ensemble. Vous pouvez :

- Afficher ou masquer les lignes et colonnes vides. (Une ligne ou une colonne vide ne contient aucun élément dans les cellules de données.)
- Contrôler le positionnement de libellés de ligne, qui peuvent se trouver dans l'angle supérieur gauche ou être imbriquées.
- Contrôler la largeur de colonne maximale et minimale (exprimée en points).

Pour modifier les propriétés générales de la table :

1. Cliquez sur l'onglet **Général**.
2. Sélectionnez les options de votre choix.
3. Cliquez sur **OK** ou sur **Appliquer**.

Définition des lignes à afficher :

Remarque : Cette fonction ne s'applique qu'aux tableaux existants.

Par défaut, les tableaux comportant un grand nombre de lignes sont affichées par section de 100 lignes. Pour contrôler le nombre de lignes affichées dans un tableau :

1. Sélectionnez **Afficher les tableaux par lignes**.
2. Cliquez sur **Définir les lignes à afficher**.

   ou
3. Dans le menu Affichage d'un tableau croisé dynamique activé, choisissez **Afficher les tableaux par lignes** et **Définir les lignes à afficher**.

Lignes à afficher : Contrôle le nombre maximal de lignes à afficher en une fois. Les contrôles de navigation vous permettent de vous déplacer vers les différentes sections du tableau. Valeur minimum : 10. La valeur par défaut est 100.

Tolérance Veuve/Orpheline : Contrôle le nombre maximal de lignes de la dimension de ligne la plus interne du tableau à diviser entre les vues affichées du tableau. Par exemple, s'il existe six catégories dans chaque groupe de la dimension de ligne la plus interne, en spécifiant une valeur de six, vous évitez qu'un groupe ne soit divisé entre les vues affichées. Ce paramètre peut engendrer un nombre total de lignes supérieur au nombre de lignes maximal à afficher dans une vue.

Propriétés du tableau : Notes

L'onglet associé aux notes de la boîte de dialogue Propriété du tableau contrôle le formatage des notes de bas de page ainsi que les textes de commentaires des tableaux.

Notes de bas de page : Les propriétés des marqueurs des notes de bas de page comprennent le style et la position par rapport au texte.

- Le style des marqueurs des notes de bas de page peut être des numéros (1, 2, 3, ...) ou des lettres (a, b, c, ...).
- Les marqueurs des notes de bas de page sont attachés au texte sous la forme d'exposants ou d'indices.

Texte de commentaire : Vous pouvez ajouter un texte de commentaire à chaque tableau.

- Le texte du commentaire apparaît sous la forme d'une infobulle lorsque vous passez le curseur de la souris sur un tableau dans le Visualiseur.
- Les lecteurs d'écran lisent le texte du commentaire lorsque le tableau est mis en évidence.
- L'infobulle du Visualiseur affiche uniquement les 200 premiers caractères du caractères, tandis que les lecteurs d'écran lisent le texte dans son intégralité.
- Lorsque vous exportez une sortie au format HTML ou de rapport Web, le texte du commentaire est utilisé en tant que texte secondaire.

Propriétés du tableau : Formats de cellule

Pour le formatage, un tableau est divisé en zones : titre, couches, libellés de coins, libellés de lignes, libellés de colonnes, données, légende et notes de bas de page. Pour chaque zone d'un tableau, vous pouvez modifier les formats de cellule associés. Les formats de cellule comprennent les caractéristiques du texte (telles que la police, la taille, la couleur, le style), l'alignement horizontal et vertical, les couleurs d'arrière-plan, et les marges de cellule internes.

Les formats de cellule sont appliqués aux zones (catégories d'informations). Il n'existe pas de caractéristiques de cellules individuelles. Cette distinction est un élément important au moment du croisement d'un tableau.

Par exemple,

- Si vous spécifiez une police en gras comme format de cellule pour les libellés de colonne, les libellés de colonne apparaîtront en gras quelles que soient les informations en cours d'affichage dans la dimension colonne. Si vous déplacez un élément de la dimension colonne vers une autre dimension, il ne retient pas la caractéristique gras des libellés de colonne.
- Si vous mettez les libellés de colonne en gras simplement en surlignant les cellules dans un tableau croisé dynamique activé et en cliquant sur le bouton Gras de la barre d'outils, le contenu des cellules restera en gras quelle que soit la dimension vers laquelle vous déplacez ces cellules, et les libellés de colonne ne retiendront pas la caractéristique grasse pour les autres éléments déplacés dans la dimension colonne.

Pour modifier les formats de cellule :

1. Sélectionnez l'onglet **Formats de cellule**.
2. Sélectionnez une zone dans la liste déroulante ou cliquez sur une zone de l'échantillon.
3. Sélectionnez des caractéristiques pour la zone. Vos sélections sont reflétées dans l'exemple.
4. Cliquez sur **OK** ou sur **Appliquer**.

Couleurs de ligne en alternance

Pour appliquer une couleur d'arrière-plan et/ou de texte différente et faire alterner les couleurs des lignes dans la zone Données de la table :

1. Sélectionnez **Données** dans la liste déroulante Surfaces.
2. Cliquez sur (cochez) **Autre couleur de ligne** dans le groupe Couleur d'arrière-plan.
3. Sélectionnez les couleurs à utiliser en alternance pour le texte et l'arrière-plan des lignes.

Les couleurs de lignes en alternance affectent seulement la zone Données du tableau. Elles n'ont pas d'effet sur les zones de libellés de ligne ou de colonne.

Propriétés du tableau : Bordures

Pour chaque emplacement de bordure dans un tableau, vous pouvez sélectionner un style de ligne et une couleur. Si vous sélectionnez **Aucune** comme style, il n'y aura aucune ligne à l'emplacement sélectionné.

Pour modifier les bordures de la table :

1. Cliquez sur l'onglet **Bordures**.
2. Sélectionnez un emplacement de bordure, soit en cliquant sur son nom dans la liste soit en cliquant sur une ligne dans la zone exemple.
3. Sélectionnez un style de ligne ou **Aucun**.
4. Sélectionnez une couleur.
5. Cliquez sur **OK** ou sur **Appliquer**.

Propriétés du tableau : Impression

Vous pouvez contrôler les propriétés suivantes pour les tableaux croisés dynamiques imprimés :

- Imprimer toutes les couches ou uniquement la couche supérieure du tableau, et imprimer chaque couche sur une page différente.
- Réduire un tableau horizontalement ou verticalement pour l'adapter à la page avant d'imprimer.
- Contrôler les lignes veuves/orphelines en contrôlant le nombre minimum de lignes et de colonnes qui seront contenues dans une section imprimée d'un tableau, si le tableau est trop large et/ou trop long pour la taille de page définie.

  Remarque : Si un tableau est trop long pour tenir sur le restant de la page en cours car il y a un d'autres sorties au-dessus dans la page, mais qu'il tient dans la longueur de page définie, il est automatiquement imprimé sur une nouvelle page, quel que soit le paramétrage des lignes veuves/orphelines.
- Inclure le texte de continuation pour les tableaux qui ne tiennent pas sur une seule page. Vous pouvez afficher le texte de lignes incomplètes au bas de chaque page et en haut de chaque page. Si aucune option n'est sélectionnée, le texte de lignes incomplètes ne sera pas affiché.

Pour contrôler les propriétés d'impression des tableaux croisés dynamiques :

1. Cliquez sur l'onglet **Impression**.
2. Sélectionnez les options d'impression de votre choix.
3. Cliquez sur **OK** ou sur **Appliquer**.

Propriétés de cellule

Les propriétés de cellule s'appliquent à une cellule sélectionnée. Vous pouvez modifier la police, le format des valeurs, l'alignement, les marges et les couleurs. Les propriétés de cellule remplacent les propriétés de tableau ; par conséquent si vous modifiez les propriétés de tableau, vous ne modifiez pas les propriétés de cellule appliquées individuellement.

Pour modifier les propriétés de cellule :

1. Sélectionnez les cellules dans le tableau.
2. Dans le menu Format ou le menu contextuel, choisissez **Propriétés de la cellule**.

Police et arrière-plan

L'onglet Couleur et arrière-plan contrôle le style et la couleur de la police et la couleur d'arrière-plan pour les cellules sélectionnées dans le tableau.

Valeur de format

L'onglet Format valeur contrôle les formats de valeur des cellules sélectionnées. Vous pouvez sélectionner des formats pour les nombres, les dates, l'heure ou les devises, et vous pouvez régler le nombre de décimales affichées.

Alignement et marges

L'onglet Alignement et marges contrôle l'alignement horizontal et vertical des valeurs et les marges supérieure, inférieure, gauche et droite pour les cellules sélectionnées. L'alignement horizontal **Mixte** aligne le contenu de chaque cellule en fonction de son type. Par exemple, les dates sont alignées à droite et les valeurs de texte sont alignées à gauche.

Notes de bas de page et légendes

Vous pouvez ajouter des notes de bas de page et des légendes à un tableau. Vous pouvez également masquer les notes de bas de page ou les légendes, modifier les marqueurs des notes de bas de page et renuméroter les notes de bas de page.

Ajout de notes de bas de page et de légendes

Pour ajouter une légende à un tableau :

1. Dans le menu Insérer, sélectionnez **Légende**.

Une note de bas de page peut être attachée à n'importe quel élément d'un tableau. Pour ajouter une note de bas de page :

1. Cliquez sur un titre, une cellule ou une légende dans un tableau croisé dynamique activé.
2. Dans le menu Insérer, sélectionnez **Note de bas de page**.
3. Insérez le texte de la note de bas de page dans la zone appropriée.

Affichage ou masquage d'une légende

Pour masquer une légende :

1. Sélectionnez la légende.
2. Dans le menu Affichage, choisissez **Masquer**.

Pour afficher les légendes masquées :

1. Dans le menu Affichage, choisissez **Montrer Tout**.

Masquage ou affichage d'une note de bas de page dans un tableau

Pour masquer une note de bas de page :

1. Cliquez avec le bouton droit sur la cellule contenant la référence de la note de bas de page et sélectionnez **Masquer les notes de bas de page** dans le menu contextuel.

   ou

2. Sélectionnez la note de bas de page dans leur zone du tableau, puis **Masquer** dans le menu contextuel.

   Remarque : pour les tableaux classiques, sélectionnez la zone des notes de bas de page, puis **Modifier les notes de bas de page** dans le menu contextuel et désélectionnez (décochez) la propriété Visible des notes de bas de page que vous désirez masquer.

   Si une cellule contient plusieurs notes de bas de page, utilisez cette dernière méthode pour masquer les notes de bas de page de votre choix.

Pour masquer toutes les notes de bas de page du tableau :

1. Sélectionnez tous les éléments dans la zone des notes de bas de page du tableau (en cliquant dessus et en les faisant glisser ou à l'aide de Maj+Clic pour les sélectionner) et sélectionnez **Masquer** dans le menu Vue.

Pour afficher les notes de bas de page masquées

1. Sélectionnez **Afficher toutes les notes de bas de page** dans le menu Affichage.

Marque de bas de page

Le marqueur de bas de page modifie les caractères pouvant être utilisés pour dénoter une marque de bas de page. Par défaut, les marques de notes de bas de page standard sont des séquences de lettres ou de chiffres, en fonction des paramètres de propriétés du tableau. Vous pouvez également attribuer une marque spéciale. Les marques spéciales ne sont pas affectées lorsque vous numérotez de nouveau les notes de bas de page ou que vous passez du mode chiffres au mode lettres pour les marques standard. L'affichage des chiffres ou des lettres des marques standard et la position de l'indice sont contrôlés par l'onglet Notes de bas de page de la boîte de dialogue Propriétés du tableau.

Pour modifier les marqueurs des notes de bas de page :

1. Sélectionnez une note de bas de page.
2. Dans le menu **Format**, sélectionnez **Marque de bas de page**.

Les marqueurs spéciaux sont limités à 2 caractères. Les marques de base de page avec des marqueurs spéciaux précèdent celles avec des lettres ou des nombres séquentiels dans la zone des notes de bas de page du tableau ; leur modification pour un marqueur spécial peut donc réorganiser la liste des notes de bas de page.

Renumérotation des notes de bas de page

Lorsque vous avez croisé un tableau en permutant les lignes, les colonnes et les couches, il se peut que les notes de bas de page soient dans le désordre. Pour renuméroter les notes de bas de page :

1. Dans le menu Format, sélectionnez **Renuméroter les notes de bas de page**.

Modification des notes de bas de page dans les tableaux classiques

Pour les tableaux classiques, vous pouvez utiliser la boîte de dialogue Modifier les notes de bas de page pour modifier le texte et la police, modifier les marqueurs de note de bas de page et masquer ou supprimer des notes de bas de page de manière sélective.

Lorsque vous insérez une nouvelle note de bas de page dans un tableau classique, la boîte de dialogue Modifier les notes de bas de page s'ouvre automatiquement. Afin d'utiliser la boîte de dialogue Modifier les notes de bas de page pour modifier des notes de bas de page existantes (sans en créer de nouvelles) :

Marque : Par défaut, les marques de notes de bas de page standard sont des séquences de lettres ou de chiffres, en fonction des paramètres de propriétés du tableau. Pour attribuer une marque spéciale, saisissez simplement la nouvelle valeur de marque dans la colonne correspondante. Les marques spéciales ne sont pas affectées lorsque vous numérotez de nouveau les notes de bas de page ou que vous passez du mode chiffres au mode lettres pour les marques standard. L'affichage des chiffres ou des lettres des marques standard et la position de l'indice sont contrôlés par l'onglet Notes de bas de page de la boîte de dialogue Propriétés du tableau. Pour plus d'informations, voir la rubrique «Propriétés du tableau : Notes», à la page 121.

Pour modifier une marque spéciale afin qu'elle redevienne standard, cliquez avec le bouton droit sur la marque dans la boîte de dialogue Modifier les notes de bas de page, et sélectionnez **Marque de bas de page** dans le menu contextuel, puis Marque standard dans la boîte de dialogue Marque de bas de page.

Note de bas de page. Le contenu de la note de bas de page. L'affichage reflète les paramètres de police et d'arrière-plan en cours. Les paramètres de police peuvent être modifiés pour chaque note de bas de page à l'aide de la sous-boîte de dialogue Format. Pour plus d'informations, voir la rubrique «Paramètres de couleur et de police des notes de bas de page», à la page 126. Une unique couleur d'arrière-plan est appliquée à toutes les notes de bas de page et peut être modifiée dans l'onglet Police et arrière-plan de la boîte de dialogue Propriétés de la cellule. Pour plus d'informations, voir la rubrique «Police et arrière-plan», à la page 123.

Visible : Toutes les notes de bas de page sont visibles par défaut. Désélectionnez (décochez) la case Visible pour masquer une note de bas de page.

Paramètres de couleur et de police des notes de bas de page : Pour les tableaux classiques, vous pouvez utiliser la boîte de dialogue Format pour modifier la famille, le style, la taille et la couleur de police d'une ou de plusieurs notes de bas de page sélectionnées :

1. Dans la boîte de dialogue Modifier les notes de bas de page, sélectionnez (cliquez sur) une ou plusieurs notes de bas de page dans la grille Notes de bas de page.
2. Cliquez sur le bouton **Format**.

La famille, le style, la taille et les couleurs sont appliqués à toutes les notes de bas de page sélectionnées.

La couleur d'arrière-plan, l'alignement et les marges peuvent être définis dans la boîte de dialogue Propriétés de la cellule et appliqués à toutes les notes de bas de page. Vous ne pouvez pas modifier ces paramètres pour chaque note de bas de page individuelle. Pour plus d'informations, voir la rubrique «Police et arrière-plan», à la page 123.

Largeur des cellules de données

La boîte de dialogue Définir la largeur des cellules de données permet de définir la même largeur pour toutes les cellules de données.

Pour définir la largeur de toutes les cellules de données :

1. A partir des menus, sélectionnez :

   Format > **Définition de la largeur des cellules de données...**
2. Saisissez une valeur pour la largeur des cellules.

Modification de la largeur de colonne

1. Cliquez et faites glisser la bordure de la colonne.

Affichage des bordures masquées dans un tableau croisé dynamique

Pour les tables contenant peu de bordures visibles, vous pouvez afficher les bordures masquées. Cette option peut simplifier certaines tâches comme la modification de la largeur des colonnes.

1. Dans le menu Affichage, choisissez **Quadrillage**.

Sélection de lignes, de colonnes et de cellules dans un tableau croisé dynamique

Vous pouvez sélectionner une ligne ou une colonne en entier ou un ensemble de cellules de données et de libellés spécifié.

Pour sélectionner plusieurs cellules :

Sélectionner > **Cellules de données et de libellés**

Impression des tableaux croisés dynamiques

Plusieurs facteurs peuvent affecter l'apparence des tableaux croisés dynamiques imprimés, et ces facteurs peuvent être contrôlés en modifiant les attributs des tableaux croisés dynamiques.

- Pour les tableaux croisés dynamiques multidimensionnels (tableaux avec couches), vous pouvez imprimer toutes les couches ou imprimer seulement la couche supérieure (visible). Pour plus d'informations, voir la rubrique «Propriétés du tableau : Impression», à la page 123.

- Pour les tableaux croisés dynamiques longs ou larges, vous pouvez redimensionner automatiquement le tableau pour l'ajuster à la page ou contrôler l'emplacement des ruptures de tableau ou des sauts de page. Pour plus d'informations, voir la rubrique «Propriétés du tableau : Impression», à la page 123.

Utilisez l'Aperçu avant impression dans le menu Fichier pour voir l'apparence des tableaux croisés dynamiques imprimés.

Contrôle des ruptures de tableau pour les tableaux longs et larges

Les tableaux croisés dynamiques qui sont trop larges ou trop longs pour être imprimés dans la taille de page définie sont automatiquement fractionnés et imprimés en plusieurs parties. Vous pouvez :

- Contrôler l'emplacement des lignes et des colonnes où les grands tableaux sont fractionnés.
- Indiquer des lignes et des colonnes qui doivent être gardées ensemble lorsque les tableaux sont fractionnés.
- Remettre à l'échelle les grands tableaux pour les ajuster à la taille de page définie.

Pour spécifier des sauts de ligne et de colonne pour les tableaux croisés imprimés :

1. Activer le tableau croisé dynamique.
2. Cliquez sur une cellule quelconque dans la colonne à gauche de laquelle vous désirez insérer un saut ou sur une cellule quelconque dans la ligne avant celle où vous désirez insérer le saut.
3. Dans les menus, sélectionnez :

   Format > **Points de rupture** > **Point de rupture vertical**

   ou

   Format > **Points de rupture** > **Point de rupture horizontal**

1. Activer le tableau croisé dynamique.
2. Cliquez sur une cellule quelconque dans la colonne à gauche de laquelle vous désirez insérer un saut ou sur une cellule quelconque dans la ligne avant celle où vous désirez insérer le saut.
3. Dans les menus, sélectionnez :

   Format > **Points de rupture** > **Point de rupture vertical**

   ou

   Format > **Points de rupture** > **Point de rupture horizontal**

Pour spécifier des lignes ou des colonnes à conserver ensemble :

1. Sélectionnez les libellés des lignes ou des colonnes à garder ensemble. Cliquez et faites glisser ou utilisez la combinaison Maj+Clic pour sélectionner plusieurs libellés de lignes ou de colonnes.
2. Dans les menus, sélectionnez :

   Format > **Points de rupture** > **Conserver ensemble**

Pour afficher les points d'arrêt et conserver les groupes ensemble :

1. Dans les menus, sélectionnez :

Format > **Points de rupture** > **Afficher les points de rupture**

Les points de rupture apparaissent sous la forme de lignes verticales ou horizontales. Les groupes conservés ensemble apparaissent sous forme de régions rectangulaires grisées avec une bordure plus sombre.

Remarque : l'affichage des points d'arrêt et des groupes conservés ensemble n'est pas pris en charge pour les tableaux des anciennes versions.

Pour effacer les points d'arrêt et conserver les groupes ensemble :

Pour effacer un point d'arrêt :

1. Cliquez sur une cellule quelconque dans la colonne à gauche d'un point d'arrêt vertical ou sur une cellule dans la ligne au-dessus d'un point d'arrêt horizontal.
2. Dans les menus, sélectionnez :

   Format > **Points de rupture** > **Effacer le point de rupture ou le groupe**

   Pour supprimer l'option Conserver un groupe ensemble :
3. Sélectionnez les libellés de colonne ou de ligne qui spécifient le groupe.
4. Dans les menus, sélectionnez :

   Format > **Points de rupture** > **Effacer le point de rupture ou le groupe**

Tous les points d'arrêt et les groupes conservés ensemble sont automatiquement effacés lorsque vous faites pivoter ou réorganisez une ligne ou une colonne. Ce comportement ne s'applique qu'aux tableaux classiques.

Création d'un graphique à partir d'un tableau croisé dynamique

1. Double-cliquez sur le tableau croisé dynamique pour l'activer.
2. Sélectionnez les lignes, les colonnes ou les cellules à afficher dans le graphique.
3. Cliquez avec le bouton droit de la souris à n'importe quel endroit de la zone sélectionnée.
4. Sélectionnez **Créer un graphique** dans le menu contextuel, puis choisissez un type de graphique.

Tableaux de version antérieure

Vous pouvez sélectionner un rendu de tableaux sous forme de tableaux classiques (dénommés tableaux complets dans la version 19) qui sont alors totalement compatibles avec les éditions de IBM SPSS Statistics antérieures à la version 20. Ce rendu peut prendre du temps et n'est recommandé que si une compatibilité avec les éditions antérieures à la version 20 est requise. Pour plus d'informations sur la procédure de création de tableaux de version antérieure, voir «Options de tableau croisé dynamique», à la page 129.

Options

Options

Les options contrôlent divers paramètres.

Pour modifier les paramètres d'option

1. Dans les menus, sélectionnez :

   Edition > **Options...**
2. Cliquez sur les onglets correspondant aux paramètres que vous souhaitez modifier.
3. Modifiez les paramètres.
4. Cliquez sur **OK** ou sur **Appliquer**.

Options générales

Nombre maximal d'unités d'exécution

Nombre d'unités d'exécution que les procédures à unités d'exécution multiples utilisent lors du calcul des résultats. Le paramètre **Automatique** repose sur le nombre de coeurs de traitement disponibles. Spécifiez une valeur inférieure si vous voulez mettre à la disposition d'autres applications davantage de ressources de traitement alors que des procédures à unités d'exécution multiples sont en cours. Cette option est désactivée en mode analyse distribuée.

Sortie

Afficher un zéro non significatif pour les valeurs décimales. Affiche les zéros non significatifs des valeurs numériques qui ne comportent que des décimales. Par exemple, si les zéros non significatifs sont affichés, la valeur .123 est affichée sous la forme 0.123. Ce paramètre ne s'applique pas aux valeurs numériques correspondant à une devise ou un pourcentage. A l'exception des fichiers ASCII fixes (*.dat), les zéros non significatifs ne sont pas inclus lorsque les données sont sauvegardées dans un fichier externe.

Unités de mesure : Système de mesure utilisé (Points, Pouces, ou Centimètres) pour spécifier les attributs tels que les marges de cellules du tableau croisé dynamique, les largeurs de cellules et l'espace d'impression entre les tableaux

Options du Visualiseur

Les options d'affichage de la sortie avec le Visualiseur affectent seulement la nouvelle sortie produite après modification de la configuration. Une sortie qui est déjà affichée dans le Visualiseur n'est pas affectée par les changements de ces paramètres.

Etat initial sortie : Cette procédure contrôle quels éléments seront automatiquement affichés ou masqués chaque fois que vous exécuterez une procédure et combien d'éléments sont initialement alignés. Vous pouvez contrôler l'affichage des éléments suivants : journal, avertissements, remarques, titres, tableaux croisés dynamiques, graphiques, diagrammes d'arbre et sortie texte. Vous pouvez également démarrer ou arrêter l'affichage des commandes dans le journal. Vous pouvez copier la syntaxe de commande à partir du journal et l'enregistrer dans un fichier de syntaxe.

Remarque : Tous les éléments de sortie sont affichés à gauche dans le Visualiseur. Seul l'alignement des sorties imprimées est affecté par les paramètres d'alignement. Les éléments centrés et alignés à droite sont identifiés par un petit symbole.

Titre : Cette option permet de contrôler le style de police, la taille, et la couleur des titres de sortie.

Titre de la page : Cette option permet de contrôler tous les styles de police, la taille et la couleur des nouveaux titres de page et de ceux générés par la syntaxe de commande `TITLE` et `SUBTITLE` ou créés dans le menu **Insertion** à l'aide de l'option **Nouveau titre de page**.

Sortie texte : Police utilisée pour les sorties texte. La sortie texte est prévue pour être utilisée dans une police monoespacée (à espacement fixe). Si vous sélectionnez une police proportionnelle, le tableau de sortie ne sera pas correctement aligné.

Mise en page par défaut : Cette option permet de contrôler les options d'orientation et de marges d'impression par défaut.

Options de tableau croisé dynamique

Les options des tableaux croisés dynamiques définissent différentes options d'affichage des tableaux croisés dynamiques.

Modèle de tableaux

Sélectionnez un modèle de tableau dans la liste des fichiers et cliquez sur **OK** ou sur **Appliquer**. Vous pouvez utiliser l'un des modèles de tableaux offerts dans IBM SPSS Statistics, ou vous pouvez en créer un dans l'Editeur de tableaux croisés dynamiques (sélectionnez **Modèles de tableaux** dans le menu Format).

- **Parcourir :** Cette procédure permet de sélectionner un modèle de tableau à partir d'un autre répertoire.

- **Définir le répertoire des modèles de tableaux :** Cette procédure permet de modifier le répertoire des modèles de tableau par défaut. Utiliser l'option **Parcourir** pour accéder au répertoire souhaité, sélectionnez dans ce répertoire un modèle de tableau, puis sélectionnez **Définir le répertoire des modèles de tableau**.

Remarque : Les modèles de tableaux créés dans des versions précédentes d'IBM SPSS Statistics ne peuvent pas être utilisés dans la version 16.0 ou ultérieure.

Largeur des colonnes

Les options suivantes permettent de contrôler l'ajustement automatique des largeurs de colonnes dans les tableaux croisés dynamiques.

- **Ajuster pour les libellés uniquement :** Cette procédure permet d'ajuster la largeur de colonne par rapport à la largeur du libellé colonne. Ceci permet de générer des tableaux plus compacts, mais les valeurs de données plus grandes que le libellé risquent d'être tronquées.
- **Ajuster pour les libellés et les données pour tous les tableaux :** Ajuste la largeur des colonnes selon la valeur la plus grande, à savoir celle du libellé de la colonne ou celle des données. Cela produit des tableaux plus larges, mais garantit que toutes les valeurs seront affichées.

Mode de modification par défaut

Cette option contrôle l'activation des tableaux croisés dynamiques dans la fenêtre du Visualiseur ou dans une fenêtre séparée. Par défaut, le fait de double-cliquer sur un tableau croisé dynamique active tout sauf les tableaux de grande taille dans la fenêtre du Visualiseur. Vous pouvez choisir d'activer les tableaux croisés dynamiques dans une fenêtre séparée ou de sélectionner une définition de taille qui ouvrira de plus petits tableaux croisés dynamiques dans la fenêtre du Visualiseur et de plus grands tableaux croisés dynamiques dans une fenêtre séparée.

Copie de tableaux larges dans le Presse-papiers au format texte enrichi

Lorsque des tableaux croisés dynamiques sont collés au format Word/RTF, les tableaux trop larges pour la largeur du document seront ajustés ou réduits pour correspondre à la largeur du document, soit ils seront inchangés.

Options relatives aux sorties

Les options relatives aux sorties contrôlent le paramètre par défaut d'un certain nombre d'options concernant les sorties.

Accessibilité avec lecteur d'écran. Contrôle la méthode de lecture des libellés de lignes et de colonnes par le lecteur d'écran. Vous pouvez lire la totalité des libellés de lignes et de colonnes pour chaque cellule de données ou lire uniquement les libellés de lignes et de colonnes qui sont différents lorsque vous changez de cellule.

Chapitre 8. Traitement des valeurs manquantes

Description des valeurs manquantes

Au cours de l'étape de préparation des données d'exploration de données, vous aurez souvent à remplacer des valeurs manquantes. Les **valeurs manquantes** sont des valeurs du jeu de données qui sont inconnues, non recueillies ou entrées de façon incorrecte. En général, ces valeurs ne sont pas valides pour leur champ. Par exemple, le champ *Sexe* doit contenir les valeurs *M* et *F*. Si les valeurs *Y* ou *Z* apparaissent dans ce champ, il y a de fortes chances qu'elles ne soient pas valides et, par conséquent, elles devraient être considérées comme des blancs. De même, une valeur négative pour le champ *Age* ne signifie rien et devrait également être considérée comme blanc. Il arrive fréquemment que de telles valeurs soient entrées volontairement ou ne soient pas renseignées au cours d'un questionnaire pour indiquer une absence de réponse. Il est recommandé d'apporter une attention particulière à ces blancs pour déterminer si une absence de réponse, telle que le refus d'indiquer son âge, est un facteur dans la prévision d'un résultat spécifique.

Certaines méthodes de modélisation traitent mieux les données manquantes que d'autres. Par exemple, C5.0 et Apriori parviennent à analyser les valeurs qui sont explicitement déclarées "manquantes" dans un noeud Typer. D'autres méthodes de modélisation rencontrent des problèmes de traitement des données manquantes, entraînant ainsi des durées d'apprentissage prolongées et des modèles moins précis.

Plusieurs types de valeur manquante sont reconnus par IBM SPSS Modeler :

- **Valeurs système nulles ou manquantes.** Ces valeurs sont des valeurs non-chaîne qui ne sont pas renseignées dans la base de données ou dans le fichier source, et qui n'ont pas été spécifiquement définies comme "manquantes" dans un noeud source ou un noeud Typer. Les valeurs système manquantes sont affichées sous la forme **$null$**. Les chaînes vides ne sont pas considérées comme des valeurs nulles dans IBM SPSS Modeler, même si elles peuvent être traitées comme telles par certaines bases de données.
- **Chaînes vides et espaces blancs.** Les chaînes vides et les espaces blancs (chaînes sans caractère visible) sont traités différemment des valeurs nulles. Dans la plupart des cas, les chaînes vides sont considérées comme des espaces blancs. Par exemple, si vous choisissez de traiter les espaces blancs comme blancs dans un noeud source ou un noeud Typer, ce paramètre s'applique également aux chaînes vides.
- **Valeurs manquantes définies par l'utilisateur ou vides.** Ces valeurs sont des valeurs, telles que `inconnu`, `99` ou `-1`, qui sont explicitement définies comme manquantes dans un noeud source ou type. Vous pouvez également, si vous le souhaitez, préciser si les valeurs nulles et les espaces blancs doivent être traités comme des blancs ; un traitement spécial leur est alors appliqué et ils sont exclus de la plupart des calculs. Par exemple, vous pouvez utiliser la fonction `@BLANK` pour traiter comme des blancs ces valeurs, ainsi que d'autres types de valeur manquante.

Lecture de données mixtes. Au cours de la lecture des champs de stockage numérique (entier, nombre réel, heure, horodatage ou date), toutes les valeurs non numériques sont définies comme étant *nulles* ou *manquantes dans le système*. En effet, contrairement à certaines applications, les types de stockage mixtes au sein d'un champ ne sont pas autorisés. Pour éviter ce type de problème, faites en sorte que les champs comportant des données mixtes soient lus en tant que chaînes ; pour cela, modifiez le type de stockage dans le noeud source ou dans l'application externe.

Lecture de chaînes vides issues d'Oracle. Lorsque vous lisez une base de données Oracle ou que vous écrivez dedans, souvenez-vous que, contrairement à IBM SPSS Modeler et à la plupart des autres bases de données, Oracle traite et stocke les valeurs de chaîne vides comme des valeurs nulles. Autrement dit, les mêmes données extraites d'une base de données Oracle, ou d'un fichier ou d'une autre base de données peuvent se comporter différemment, et donc renvoyer des résultats différents.

© Copyright IBM Corp. 1994, 2019

Traitement des valeurs manquantes

Vous devez décider du traitement à appliquer aux valeurs manquantes en fonction de vos connaissances métier ou de votre domaine. Afin de réduire la durée d'apprentissage et d'améliorer l'exactitude, il est recommandé de supprimer les blancs du jeu de données. Par ailleurs, la présence de blancs peut mener à de nouvelles opportunités commerciales ou de nouvelles idées. Pour déterminer la technique la plus appropriée, prenez en considération les aspects suivants de vos données :

- Taille du jeu de données
- Nombre de champs contenant des blancs
- Quantité d'informations manquantes

En règle générale, vous pouvez appliquer deux méthodes :

- Vous pouvez exclure des champs ou des enregistrements avec des valeurs manquantes.
- Vous pouvez attribuer, remplacer ou forcer les valeurs manquantes avec un grand nombre de méthodes.

Ces deux méthodes peuvent être en grande partie automatisées grâce au noeud Audit données. Par exemple, vous pouvez générer un noeud Filtrer qui exclut les champs contenant trop de valeurs manquantes pour être utiles à la modélisation et générer un super noeud qui attribue les valeurs manquantes à l'un des champs ou à tous les champs restants. C'est ici que la puissance réelle de l'audit intervient, vous permettant non seulement d'évaluer l'état actuel de vos données, mais également d'agir sur la base de cette évaluation.

Traitement des enregistrements contenant des valeurs manquantes

Si la majorité des valeurs manquantes sont concentrées dans un nombre limité d'enregistrements, vous pouvez simplement les exclure. Par exemple, une banque conserve en général des enregistrements complets et détaillés sur ses clients prêteurs. Si, toutefois, la banque restreint moins le nombre de prêts accordés à son propre personnel, les données recueillies pour les prêts au personnel comporteraient plusieurs champs vides. Dans ce cas, deux options de traitement des valeurs manquantes sont à votre disposition :

- Vous pouvez utiliser un noeud Sélectionner pour supprimer les enregistrements relatifs au personnel.
- Si le jeu de données est volumineux, vous pouvez supprimer tous les enregistrements contenant des blancs.

Traitement des champs contenant des valeurs manquantes

Si la majorité des valeurs manquantes est concentrée dans un petit nombre de champs, vous pouvez les traiter au niveau du champ plutôt qu'au niveau de l'enregistrement. Cette méthode vous permet de tester l'importance relative de certains champs avant de décider d'une méthode de traitement des valeurs manquantes. Si un champ n'est pas significatif en termes de modélisation, il n'est probablement pas utile de le conserver, quel que soit le nombre de valeurs manquantes qu'il comporte.

Par exemple, une société effectuant des études de marché peut recueillir des données à l'aide d'un questionnaire général de 50 questions. Deux des questions abordent l'âge et les idéologies politiques, informations que bon nombre de personnes hésitent à fournir. Dans ce cas, *Age* et *Convictions_politiques* comportent de nombreuses valeurs manquantes.

Niveau de mesure de champ

Pour choisir la méthode à utiliser, vous devez également tenir compte du niveau de mesure de champs contenant des valeurs manquantes.

Champs numériques. Pour les champs de type numérique, tels que les champs *Continus*, éliminez toutes les valeurs non numériques avant de construire un modèle. En effet, de nombreux modèles ne fonctionneront pas si des blancs sont inclus dans les champs numériques.

Champs catégoriels. Pour les champs catégoriels, tels que les champs *Nominaux* et *Indicateurs*, il n'est pas nécessaire de modifier les valeurs manquantes, bien que les modifier augmenterait l'exactitude du modèle. Par exemple, un modèle qui utilise le champ *Sexe* fonctionnera toujours avec des valeurs insignifiantes, telles que *Y* et *Z*, mais le fait de supprimer toutes les valeurs autres que *M* et *F* amplifiera l'exactitude du modèle.

Filtrage ou suppression de champs

Pour filtrer les champs présentant un nombre trop élevé de valeurs manquantes, vous disposez de plusieurs options :

- Vous pouvez utiliser le noeud Audit données pour filtrer les champs en fonction de la qualité.
- Vous pouvez utiliser un noeud Sélection de fonction pour filtrer les champs dont le nombre de valeurs manquantes est supérieur à un pourcentage défini et les classer en fonction de l'importance par rapport à une cible donnée.
- Au lieu de supprimer ces champs, vous pouvez utiliser un noeud Typer pour définir leur rôle sur la valeur **Aucun**. Ainsi, les champs resteront dans le jeu de données mais seront supprimés du processus de modélisation.

Traitement des enregistrements contenant des valeurs manquantes système

Qu'est-ce que les valeurs manquantes système ?

Les valeurs manquantes système représentent des valeurs de données inconnues ou non applicables. Dans les bases de données, ces valeurs sont souvent appelées valeurs *NULL*.

Les valeurs manquantes système sont différentes des valeurs non renseignées. Les valeurs non renseignées sont en général définies dans le noeud Type comme des valeurs particulières, ou des plages de valeurs, qui peuvent être considérées comme des valeurs manquantes définies par l'utilisateur. Les valeurs non renseignées sont traitées différemment dans le contexte de la modélisation.

Construction des valeurs manquantes système

Les valeurs manquantes système peuvent être présentes dans les données lues à partir d'une source de données (par exemple les tables de base de données peuvent contenir des valeurs *NULL*).

Les valeurs manquantes système peuvent être construites à l'aide de la valeur **undef** dans les expressions. Par exemple, l'expression CLEM suivante renvoie l'âge, s'il est inférieur ou égal à 30, ou une valeur manquante s'il est supérieur à 30 :

```
if Age > 30 then undef else Age endif
```

Les valeurs manquantes peuvent aussi être créées lorsqu'une jointure externe est effectuée, lorsqu'un nombre est divisé par zéro, lorsque la racine carrée d'un nombre négatif est calculée, et dans d'autres cas.

Affichage des valeurs manquantes système

Les valeurs manquantes système s'affichent dans des tableaux et d'autres entrées sous la forme `$null$`.

Test des valeurs manquantes système

Utilisez la fonction spéciale **@NULL** pour renvoyer `true` si la valeur de l'argument est une valeur manquante système, par exemple :

```
if @NULL(MyFieldName) then 'It is null' else 'It is not null' endif
```

Transmission de valeurs manquantes système à des fonctions

En général, les valeurs manquantes système transmises aux fonctions propagent les valeurs système à la sortie. Par exemple, si la valeur du champ `f1` est une valeur manquante système sur une ligne particulière, l'expression `log(f1)` aura également pour valeur une valeur manquante système pour cette ligne. Il existe une exception : la fonction **@NULL**.

Valeurs manquantes système dans les expressions impliquant des opérateurs arithmétiques

L'application d'opérateurs arithmétiques aux valeurs qui incluent une valeur manquante système a pour résultat une valeur manquante système. Par exemple, si la valeur du champ `f1` est une valeur manquante système sur une ligne particulière, l'expression `f1 + 10` aura également pour valeur une valeur manquante système pour cette ligne.

Valeurs manquantes système dans les expressions impliquant des opérateurs logiques

Lorsque vous utilisez des valeurs manquantes système dans des expressions qui impliquent des opérateurs logiques, les règles de logique des trois valeurs (*true* (vrai), *false* (faux) et *missing* (manquant)) s'appliquent et peuvent être décrites dans les tables de vérité. Vous trouverez les tables de vérité des opérateurs logiques courant de *not*, *and* et *or* dans les tableaux ci-après.

Tableau 5. Table de vérité pour NOT

Opérande	**Opérande NOT**
true	false
false	true
missing	missing

Tableau 6. Table de vérité pour AND

Opérande1	**Opérande2**	**Opérande1 AND Opérande2**
true	true	true
true	false	false
true	missing	missing
false	true	false
false	false	false
false	missing	false
missing	true	missing
missing	false	false
missing	missing	missing

Tableau 7. Table de vérité pour OR

Opérande1	**Opérande2**	**Opérande1 OR Opérande2**
true	true	true

Tableau 7. Table de vérité pour OR (suite)

Opérande1	Opérande2	Opérande1 OR Opérande2
true	false	true
true	missing	true
false	true	true
false	false	false
false	missing	missing
missing	true	true
missing	false	missing
missing	missing	missing

Valeurs manquantes système dans les expressions impliquant des opérateurs de comparaison

Lorsque vous comparez une valeur manquante système et une valeur manquante non système, le résultat obtenu est une valeur manquante système et non une valeur true ou false. Les valeurs manquantes système peuvent être comparées les unes avec les autres ; deux valeurs manquantes système sont considérées comme égales.

Valeurs manquantes système dans les expressions if/then/else/endif

Lorsque vous utilisez des expressions conditionnelles et qu'elles renvoient une valeur manquante système, la valeur renvoyée est celle de la clause `else`.

Valeurs manquantes système dans le noeud Sélectionner (Select)

Pour un enregistrement, lorsque l'expression de sélection a pour résultat une valeur manquante, l'enregistrement n'est pas inclus dans la sortie du noeud Sélectionner (cette action s'applique aux modes Inclure et Supprimer).

Valeurs manquantes système dans le noeud Fusionner (Merge)

Lorsque vous effectuez une fusion à l'aide d'une clé, les enregistrements comprenant des valeurs manquantes système dans un champ clé ne sont pas fusionnés.

Valeurs manquantes système dans les agrégations

Lorsque vous agrégez des données dans des colonnes, les valeurs manquantes ne sont pas incluses dans le calcul. Par exemple, dans une colonne comportant trois valeurs { 1, 2 et undef }, la somme des valeurs de la colonne calculée est 3 ; la valeur moyenne est 1,5.

Attribution ou remplacement de valeurs manquantes

Lorsque le nombre des valeurs manquantes est faible, il peut être utile d'insérer des valeurs à la place des blancs. Vous pouvez le faire à partir du rapport Audit données, qui permet, selon les besoins, de définir des options pour des champs particuliers, puis de générer un super noeud qui attribue des valeurs selon un certain nombre de méthodes. Il s'agit de la méthode la plus souple, qui permet également de définir le traitement d'un grand nombre de champs dans un seul noeud.

Les méthodes suivantes sont disponibles pour attribuer des valeurs manquantes :

Fixe. Remplacement par une valeur fixe (soit la moyenne du champ, soit la moitié de l'intervalle, soit une constante que vous indiquez).

Aléatoire. Remplacement par une valeur aléatoire fondée sur une loi normale ou uniforme.

Expression. Permet d'indiquer une expression personnalisée. Par exemple, vous pourriez remplacer les valeurs par une variable globale créée par le noeud V. globales.

Algorithme. Remplacement par une valeur prévue par un modèle fondé sur l'algorithme C&RT. Chaque champ auquel une valeur est attribuée à l'aide de cette méthode est associé à un modèle C&RT distinct et à un noeud Remplacer qui remplace les valeurs non renseignées et les valeurs nulles par la valeur prédite par le modèle. Ensuite, un noeud Filtrer est utilisé pour supprimer les champs de prévision générés par le modèle.

Vous pouvez également, pour forcer la valeur de certains champs, utiliser un noeud Typer pour que les types de champ ne couvrent que des valeurs correctes, puis définir la colonne *Vérifier* sur **Forcer** pour les champs dont les blancs doivent être remplacés.

Fonctions CLEM pour les valeurs manquantes

Plusieurs fonctions peuvent être utilisées pour traiter les valeurs manquantes. Les fonctions suivantes sont souvent utilisées dans les noeuds Sélectionner et Remplacer pour supprimer ou renseigner des valeurs manquantes :

- `count_nulls(LIST)`
- `@BLANK(FIELD)`
- `@NULL(FIELD)`
- `undef`

Les fonctions `@` peuvent être utilisées en association avec la fonction `@FIELD` pour identifier la présence de blancs ou de valeurs nulles dans un ou plusieurs champs. Les champs peuvent être simplement signalés lorsqu'ils contiennent des blancs ou des valeurs nulles, ou ils peuvent être renseignés par des valeurs de remplacement ou utilisés dans un certain nombre d'autres opérations.

Vous pouvez compter les valeurs nulles dans une liste de champs de la manière suivante :

```
count_nulls(['cardtenure' 'card2tenure' 'card3tenure'])
```

Lors de l'utilisation des fonctions acceptant une liste de champs comme entrée, vous pouvez utiliser les fonctions spéciales `@FIELDS_BETWEEN` et `@FIELDS_MATCHING`, comme l'illustre l'exemple suivant :

```
count_nulls(@FIELDS_MATCHING('card*'))
```

Vous pouvez utiliser la fonction `undef` pour renseigner les champs avec la valeur système manquante, affichée sous la forme **$null$**. Par exemple, pour remplacer une valeur numérique, vous pouvez utiliser une instruction conditionnelle telle que :

```
if not(Age > 17) or not(Age < 66) then undef else Age endif
```

Elle remplace n'importe quel élément non compris dans l'intervalle par une valeur système manquante, affichée sous la forme **$null$**. En utilisant la fonction `not()`, vous pouvez récupérer toutes les autres valeurs numériques, y compris les valeurs négatives. Pour plus d'informations, voir la rubrique «Traitement des valeurs nulles et non renseignées», à la page 187.

Remarque sur la suppression d'enregistrements

Lorsque vous utilisez un noeud Sélectionner pour supprimer des enregistrements, la syntaxe utilise une logique à trois valeurs et inclut automatiquement des valeurs nulles dans les instructions SELECT. Pour exclure les valeurs nulles (manquantes système) dans une expression SELECT, vous devez le spécifier

explicitement en utilisant `and not` dans l'expression. Par exemple, pour sélectionner et inclure tous les enregistrements dans lesquels le type de médicament de prescription est Drug C (Médicament C), utilisez l'instruction SELECT suivante :

```
Drug = 'drugC' and not(@NULL(Drug))
```

Dans des cas semblables, les précédentes versions excluaient les valeurs nulles.

Chapitre 9. Création d'expressions CLEM

A propos de CLEM

Le langage CLEM (Clementine Language for Expression Manipulation) CLEM est efficace pour la compréhension et la manipulation des données circulant au sein des flux IBM SPSS Modeler. Les data miners utilisent beaucoup le langage CLEM dans les opérations de flux pour exécuter des tâches aussi simples que le calcul du profit à partir des données de coûts et de revenus, ou aussi complexes que la transformation de données du log Web en un ensemble de champs et d'enregistrements contenant des informations utilisables.

CLEM est utilisé dans IBM SPSS Modeler pour :

- Comparer et évaluer les conditions sur les champs d'enregistrement.
- Calculer des valeurs pour les nouveaux champs.
- Calculer de nouvelles valeurs pour les champs existants.
- Comprendre la séquence des enregistrements.
- Insérer les données des enregistrements dans des rapports.

Les expressions CLEM sont indispensables à la préparation des données dans IBM SPSS Modeler et peuvent être utilisées dans un grand nombre de noeuds, des opérations sur les enregistrements et les champs (Sélectionner, Equilibrer, Remplacer) aux graphiques et aux sorties (Analyse, Rapport, Table). Par exemple, vous pouvez utiliser CLEM dans un noeud Calculer pour créer un champ basé sur une formule telle qu'un rapport.

Les expressions CLEM peuvent également être utilisées pour les opérations de recherche et de remplacement globales. Par exemple, l'expression `@NULL(@FIELD)` peut être utilisée dans un noeud Remplacer pour remplacer les **valeurs système manquantes** par la valeur entière 0. (Pour remplacer les **valeurs utilisateur manquantes**, également appelées blancs, utilisez la fonction `@BLANK`.)

Des expressions CLEM plus complexes peuvent également être créées. Vous pouvez ainsi calculer de nouveaux champs en fonction d'un ensemble conditionnel de règles, par exemple une nouvelle catégorie de valeur créée à l'aide des expressions suivants : `If: CardID = @OFFSET(CardID,1), Then: @OFFSET(ValueCategory,1), Else: 'exclude'`.

Cet exemple utilise la fonction `@OFFSET` pour indiquer au système de renvoyer la valeur du champ *CatégorieValeur* de l'enregistrement précédent si la valeur du champ *IDCarte* d'un enregistrement donné est la même que celle de l'enregistrement précédent. Dans le cas contraire, la chaîne « exclude » doit être attribuée. En d'autres termes, si les champs *IDCarte* des enregistrements adjacents sont identiques, la même catégorie de valeur doit leur être attribuée. (Les enregistrements associés à la chaîne « exclude » peuvent être sélectionnés plus tard à l'aide d'un noeud Sélectionner.)

Exemples CLEM

Les exemples d'expressions suivants illustrent la syntaxe et les types d'expression à employer dans CLEM.

Expressions simples

Les formules peuvent être très simples, comme celle ci-dessous qui calcule un nouveau champ en fonction des valeurs des champs *After (Après)* et *Before (Avant)* :

```
(After - Before) / Before * 100.0
```

© Copyright IBM Corp. 1994, 2019

Les noms de champ ne sont pas entourés de guillemets si vous faites référence à leur valeur.

De même, l'expression suivante renvoie simplement le journal de chaque valeur du champ *salary* (salaire).

```
log(salary)
```

Expressions complexes

Les expressions peuvent également être longues et complexes. L'expression suivante renvoie la valeur *true (vrai)* si la valeur des deux champs (*$KX-Kohonen* et *$KY-Kohonen*) est comprise dans les intervalles spécifiés. Dans cet exemple, les noms de champ sont entourés de guillemets simples car ils contiennent des caractères spéciaux.

```
('$KX-Kohonen' >= -0.2635771036148072 and '$KX-Kohonen' <= 0.3146203637123107
and '$KY-Kohonen' >= -0.18975617885589602 and
'$KY-Kohonen' <= 0.17674794197082522) -> T
```

Diverses fonctions, comme les fonctions de chaîne, nécessitent la saisie de plusieurs paramètres dans une syntaxe correcte. Dans l'exemple suivant, la fonction `subscrs` est utilisée pour renvoyer le premier caractère du champ *produce_ID*, et indiquer si un élément est organique, génétiquement modifié ou traditionnel. Les résultats de l'expression sont décrits par `-> ` result`.

```
subscrs(1,produce_ID) -> `c`
```

De même, l'expression suivante sera :

```
stripchar(`3`,`123`) -> `12`
```

Il est important de noter que les caractères sont toujours entourés de guillemets simples inversés.

Combinaison de fonctions dans une expression

Les expressions CLEM sont fréquemment constituées d'une combinaison de fonctions. La fonction ci-dessous combine `subscr` et `lowertoupper` pour renvoyer le premier caractère de *produce_ID* et le convertir en lettre majuscule.

```
lowertoupper(subscr(1,produce_ID)) -> `C`
```

Cette expression peut également être écrite en sténographie :

```
lowertoupper(produce_ID(1)) -> `C`
```

Voici une autre combinaison de fonctions fréquemment utilisée :

```
locchar_back(`n`, (length(web_page)), web_page)
```

Cette expression repère le caractère `n` dans la valeur du champ *web_page* en commençant par le dernier caractère de la valeur, puis en remontant vers le début. En incluant également la fonction `length`, l'expression calcule de manière dynamique la longueur de la valeur en cours au lieu d'utiliser un nombre statique, comme 7, qui ne fonctionnerait pas pour les valeurs de moins de sept caractères.

Fonctions spéciales

De nombreuses fonctions spéciales (précédées du symbole @) sont disponibles. Voici quelques fonctions couramment utilisées :

```
@BLANK('referrer ID') -> T
```

Les fonctions spéciales sont fréquemment utilisées conjointement, méthode souvent utilisée pour signaler les blancs dans plusieurs champs simultanément.

```
@BLANK(@FIELD)-> T
```

Pour consulter d'autres exemples, reportez-vous à la documentation CLEM. Pour plus d'informations, voir la rubrique «Présentation de la référence CLEM», à la page 155.

Valeurs et types de données

Comme les formules, les expressions CLEM sont construites à partir de valeurs, de noms de champs, d'opérateurs et de fonctions. L'expression CLEM valide la plus simple est une valeur ou un nom de champ. Exemples de valeurs valides :

```
3
1.79
'banana'
```

Exemples de noms de champs :

```
Product_ID
'$P-NextField'
```

Product étant le nom d'un champ issu d'un le jeu de données du panier d'achats et *$P-NextField*, le nom d'un paramètre. La valeur de l'expression est la valeur du champ nommé. En général, les noms de champs commencent par une lettre, et peuvent comporter des chiffres et des traits de soulignement (_). Toutefois, vous n'êtes pas obligé de suivre ces règles si vous placez le nom entre guillemets. Les valeurs CLEM sont les suivantes :

- Chaînes : par exemple, `« c1 »`, `« Type 2 »`, `« extrait d'un texte libre »`
- Nombres entiers : par exemple, `12`, `0`, `189`
- Nombres réels : par exemple, `12,34`, `0,0`, `0,0045`
- Champs Date/heure : par exemple, `05/12/2002`, `12/05/2002`, `12/05/02`

Vous pouvez également utiliser les éléments suivants :

- Caractères : par exemple, `` `a` `` ou `3`
- Listes d'éléments : par exemple, `[1 2 3]`, `['Type 1' 'Type 2']`

Les listes et les codes de caractère ne représentent généralement pas des valeurs de champ. En général, ils jouent le rôle d'arguments de fonctions CLEM.

Règles des guillemets

Même si le logiciel est flexible quant à la détermination des champs, des valeurs, des paramètres et des chaînes utilisés dans une expression CLEM, les règles générales suivantes dressent la liste des conseils à suivre lorsque vous créez des expressions :

- **Chaînes** utilisez toujours des guillemets doubles lorsque vous écrivez des chaînes (`"Type 2"` ou `"value"`). Vous pouvez également utiliser des guillemets simples mais cela risque de créer une confusion avec les champs placés entre guillemets.
- **Caractères** utilisez toujours des guillemets simples inversés (`` ` ``). Prenons l'exemple du caractère `d` dans la fonction `stripchar(`d`,"drugA")`. La seule exception concerne l'utilisation d'un entier pour désigner un caractère spécifique d'une chaîne. Prenons l'exemple du caractère `5` dans la fonction `lowertoupper("druga"(5)) –> "A"`. *Remarque* : sur les claviers britanniques et américains standard, la touche du guillemet simple inversé (accent grave, unicode 0060) se trouve juste à côté de la touche Echap.
- **Champs** les champs utilisés dans les expressions CLEM ne sont généralement pas entourés de guillemets (`subscr(2,arrayID)) –> CHAR`). Vous pouvez utiliser des guillemets simples si le champ comprend un espace ou un caractère spécial (`'Order Number'`). Les champs placés entre guillemets mais non définis dans le jeu de données seront lus comme des chaînes.
- **Paramètres** utilisez toujours des guillemets simples (`'$P-threshold'`).

Expressions et conditions

Les expressions CLEM utilisées pour le calcul de nouvelles valeurs peuvent renvoyer un résultat. Par exemple :

```
Weight * 2.2
Age + 1
sqrt(Signal-Echo)
```

Elles peuvent également indiquer un résultat *vrai* ou *faux* quand elles sont utilisées pour sélectionner une condition. Par exemple :

```
Drug = "drugA"
Age < 16
not(PowerFlux) and Power > 2000
```

Vous pouvez combiner des opérateurs et des fonctions de façon arbitraire dans les expressions CLEM. Par exemple :

```
sqrt(abs(Signal)) * max(T1, T2) + Baseline
```

Les parenthèses et l'ordre de priorité des opérateurs déterminent l'ordre dans lequel l'expression est évaluée. Dans cet exemple, l'ordre d'évaluation est le suivant :

- `abs(Signal)` est évalué, puis `sqrt` est appliqué au résultat.
- `max(T1, T2)` est évalué.
- Les deux résultats sont multipliés : **x** est prioritaire par rapport à +.
- Enfin, `Baseline` est ajoutée au résultat.

L'ordre décroissant de priorité (à savoir, des opérations exécutées en premier aux opérations exécutées en dernier) est le suivant :

- arguments de fonctions
- appels de fonctions
- **xx**
- **x / mod div rem**
- + –
- > < >= <= /== == = /=

Si vous voulez ignorer l'ordre de priorité, ou si vous doutez de l'ordre de l'évaluation, vous pouvez utiliser des parenthèses pour rendre ce dernier explicite. Par exemple :

```
sqrt(abs(Signal)) * (max(T1, T2) + Baseline)
```

Paramètres de flux, de session et de super noeud

Vous pouvez définir des paramètres afin de les utiliser dans les expressions CLEM et pour la génération de scripts. Il s'agit, en effet, de variables définies par l'utilisateur qui sont enregistrées et conservées dans le flux, la session ou le super noeud en cours ; elles sont accessibles à partir de l'interface utilisateur, ainsi que via la fonction de génération de scripts. Si vous enregistrez un flux, par exemple, les paramètres définis pour ce dernier sont également enregistrés. (Cela permet de les différencier des variables de script locales qui ne peuvent être utilisées que dans le script dans lequel elles sont déclarées.) Les paramètres sont souvent utilisés lors de la génération de scripts pour contrôler le comportement du script, en fournissant des informations sur les champs et les valeurs qui n'ont pas besoin d'être codées en dur dans le script.

La portée d'un paramètre dépend de l'endroit où ce dernier est défini :

- Les paramètres de flux peuvent être définis dans un script de flux ou dans la boîte de dialogue des propriétés du flux ; ils sont disponibles pour tous les noeuds du flux. Ils sont affichés sur la liste des paramètres dans le Générateur de formules.
- Les paramètres de session peuvent être définis dans un script autonome ou dans la boîte de dialogue des paramètres de session. Ils sont disponibles pour tous les flux utilisés dans la session en cours (tous les flux répertoriés dans l'onglet Flux du panneau des gestionnaires).

Il est également possible de définir des paramètres pour les super noeuds ; dans ce cas, les paramètres ne sont visibles que par les noeuds encapsulés au sein de ce super noeud.

Utilisation de paramètres dans les expressions CLEM

Les paramètres sont représentés dans CLEM les expressions par `$P-pname`, où `pname` est le nom du paramètre. Lorsqu'ils sont utilisés dans des expressions CLEM, les paramètres doivent être placés entre guillemets simples, par exemple `'$P-scale'`.

Vous pouvez afficher les paramètres disponibles en utilisant le Générateur de formules. Pour afficher les paramètres en cours :

1. Dans toute boîte de dialogue acceptant des expressions CLEM, cliquez sur le bouton Générateur de formules.
2. Dans la liste Champs, sélectionnez **Paramètres**.

Vous pouvez sélectionner dans la liste les paramètres à insérer dans l'expression CLEM. Pour plus d'informations, voir la rubrique «Sélection de champs, de paramètres et de variables globales», à la page 150.

Utilisation de chaînes

Il existe un grand nombre d'opérations disponibles à utiliser pour les chaînes, notamment :

- Conversion d'une chaîne en majuscules ou en minuscules : `uppertolower(CHAR)`.
- Suppression de caractères indiqués, tels que `` `ID_` `` ou `` `$` ``, d'une variable de chaîne `stripchar(CHAR,STRING)`.
- Détermination de la longueur (nombre de caractères) d'une variable de chaîne `length(STRING)`.
- Vérification de l'ordre alphabétique des valeurs de chaîne `alphabefore(STRING1, STRING2)`.
- Suppression des espaces blancs situés au début et à la fin des valeur `trim(STRING)`, `trim_start(STRING)` ou `trimend(STRING)`.
- Extraction des premiers ou derniers *n* caractères d'une chaîne `startstring(LENGTH, STRING)` ou `endstring(LENGTH, STRING)`. Par exemple, imaginons un champ *élément* qui associe un nom de produit à un code d'identification de 4 caractères (`ACME CAMERA-D109`). Pour créer un nouveau champ qui contient uniquement le code à 4 caractères, spécifiez la formule suivante dans un noeud Calculer.

  ```
  endstring(4, item)
  ```
- Mise en correspondance avec un motif spécifique `STRING matches PATTERN`. Par exemple, pour sélectionner des personnes avec "marché" dans leur intitulé de poste, vous pouvez spécifier l'élément suivant dans un noeud Sélectionner :

  ```
  job_title matches "*market*"
  ```
- Remplacement de toutes les instances d'une sous-chaîne faisant partie d'une chaîne `replace(SUBSTRING, NEWSUBSTRING, STRING)`. Par exemple, pour remplacer toutes les instances d'un caractère non pris en charge, comme une barre verticale (|) par un point virgule, avant l'exploration du texte, utilisez la fonction `replace` dans un noeud Remplacer. Dans**Renseigner les champs :**, sélectionnez tous les champs où le caractère peut être présent. Pour la condition **Remplacer :**, sélectionnez **Toujours**, et spécifiez la condition suivante dans **Remplacer par :**

  ```
  replace('|',';',@FIELD)
  ```

- Calcul d'un champ indicateur basé sur la présence d'une sous-chaîne spécifique. Par exemple, vous pouvez utiliser un noeud Calculer afin de générer un champ indicateur distinct pour chaque réponse ; entrez pour cela une expression semblable à ce qui suit :

```
hassubstring(museums,"museum_of_design")
```

Pour plus d'informations, voir la rubrique «Fonctions sur chaînes», à la page 171.

Traitement des valeurs non renseignées et des valeurs manquantes

Le remplacement de blancs ou de valeurs manquantes est une tâche de préparation des données courante pour les data miners. CLEM offre un grand nombre d'outils permettant d'automatiser le traitement des blancs. Les valeurs non renseignées sont généralement traitées dans le noeud Remplacer ; cependant, les fonctions suivantes peuvent être utilisées dans n'importe quel noeud acceptant les expressions CLEM :

- `@BLANK(FIELD)` peut être utilisée pour déterminer les enregistrements dont les valeurs sont vides pour un champ particulier tel que *Age*.
- `@NULL(FIELD)` peut être utilisée pour déterminer les enregistrements dont les valeurs sont absentes du système pour les champs concernés. Dans IBM SPSS Modeler, les valeurs système manquantes sont affichées en tant que valeurs **$null$**.

Pour plus d'informations, voir la rubrique «Traitement des valeurs nulles et non renseignées», à la page 187.

Utilisation de nombres

De nombreuses opérations standard sur des valeurs numériques sont disponibles dans IBM SPSS Modeler, notamment :

- Calcul du sinus de l'angle spécifié — `sin(NUM)`
- Calcul du logarithme naturel de champs numérique — `log(NUM)`
- Calcul de la somme de deux nombres — `NUM1 + NUM2`

Pour plus d'informations, voir la rubrique «Fonctions numériques», à la page 166.

Utilisation de l'heure et de la date

Les formats d'heure et de date peuvent varier en fonction de votre source de données et de vos paramètres régionaux. Les formats de date et d'heure utilisés sont propres à chaque flux et sont définis dans la boîte de dialogue Propriétés du flux. Les exemples suivants sont des fonctions couramment utilisées avec les champs de date et d'heure.

Calcul de la durée écoulée

Vous pouvez facilement calculer la durée écoulée à partir d'une date de référence en utilisant une famille de fonctions identique à la suivante. Cette fonction indique le temps en mois en utilisant d'abord la date de référence, puis celle représentée par la chaîne de date `DATE`, sous forme de nombre réel. Ce chiffre est approximatif et utilise une base de 30 jours.

```
date_in_months(Date)
```

Comparaison de valeurs Date/heure

Les valeurs des champs Date/heure peuvent être comparées dans les enregistrements en utilisant des fonctions identiques à la suivante. Cette fonction renvoie une valeur vraie *(true)* si la chaîne de date `DATE1` représente une date antérieure à celle représentée par la chaîne de date `DATE2`. Sinon, cette fonction renvoie la valeur 0.

```
date_before(Date1, Date2)
```

Calcul des différences

Vous pouvez également calculer la différence entre deux heures et deux dates en utilisant des fonctions du type :

```
date_weeks_difference(Date1, Date2)
```

Cette fonction indique le temps en semaines en partant de la date représentée par la chaîne de date DATE1 jusqu'à la date représentée par la chaîne de date DATE2, sous forme de nombre réel. La semaine de base utilisée comprend 7,0 jours. Si DATE2 est antérieure à DATE1, cette fonction renvoie un nombre négatif.

Date du jour

La date du jour peut être ajoutée au jeu de données en utilisant la fonction @TODAY. La date du jour est ajoutée en tant que chaîne au champ spécifié ou au nouveau champ en utilisant le format de date sélectionné dans la boîte de dialogue des propriétés du flux. Pour plus d'informations, voir la rubrique «Fonctions date et heure», à la page 176.

Récapitulation de plusieurs champs

Le langage CLEM comporte un certain nombre de fonctions renvoyant des statistiques récapitulatives sur plusieurs champs. Ces fonctions s'avèrent particulièrement utiles lors de l'analyse des données d'enquête, où les diverses réponses à une question peuvent être stockées dans plusieurs champs. Pour plus d'informations, voir la rubrique «Utilisation des données à réponses multiples», à la page 146.

Fonctions de comparaison

Vous pouvez comparer les valeurs de plusieurs champs en utilisant les fonctions min_n et max_n, par exemple :

```
max_n(['card1fee' 'card2fee''card3fee''card4fee'])
```

Vous pouvez également utiliser diverses fonctions de comptage pour dénombrer les valeurs qui répondent à des critères spécifiques, même quand ces valeurs sont stockées dans plusieurs champs. Par exemple, pour compter le nombre de cartes conservées depuis plus de cinq ans :

```
count_greater_than(5, ['cardtenure' 'card2tenure' 'card3tenure'])
```

Pour compter les valeurs nulles sur le même ensemble de champs :

```
count_nulls(['cardtenure' 'card2tenure' 'card3tenure'])
```

Remarque : Cet exemple compte le nombre de cartes conservées, pas le nombre de personnes qui les possèdent. Pour plus d'informations, voir la rubrique «Fonctions de comparaison», à la page 164.

Pour compter la fréquence d'une valeur spécifique dans plusieurs champs, vous pouvez utiliser la fonction count_equal. L'exemple suivant compte le nombre de champs, dans la liste, qui contiennent la valeur Y.

```
count_equal("Y",[Answer1, Answer2, Answer3])
```

En fonction des valeurs suivantes des champs de la liste, la fonction renvoie les résultats pour la valeur Y comme indiqué.

Tableau 8. Valeurs de fonction

Answer1	Answer2	Answer3	Comptage
Y	N	Y	2
Y	N	N	1

Fonctions numériques

Vous pouvez obtenir des statistiques relatives à plusieurs champs en utilisant les fonctions sum_n, mean_n et sdev_n ; par exemple :

```
sum_n(['card1bal' 'card2bal''card3bal'])
mean_n(['card1bal' 'card2bal''card3bal'])
```

Pour plus d'informations, voir la rubrique «Fonctions numériques», à la page 166.

Génération de listes de champs

Lors de l'utilisation des fonctions acceptant une liste de champs comme entrée, vous pouvez utiliser les fonctions spéciales @FIELDS_BETWEEN(start, end) et @FIELDS_MATCHING(pattern) comme entrées. Par exemple, en supposant que l'ordre des champs correspond à celui illustré dans l'exemple sum_n ci-dessus, la formule suivante reviendrait au même :

```
sum_n(@FIELDS_BETWEEN(card1bal, card3bal))
```

Sinon, pour compter le nombre de valeurs nulles dans tous les champs commençant par « *card* » :

```
count_nulls(@FIELDS_MATCHING('card*'))
```

Pour plus d'informations, voir la rubrique «Champs spéciaux», à la page 188.

Utilisation des données à réponses multiples

Plusieurs fonctions de comparaison permettent d'analyser des données à réponses multiples, y compris :

- value_at
- first_index / last_index
- first_non_null / last_non_null
- first_non_null_index / last_non_null_index
- min_index / max_index

Par exemple, imaginons une question à réponses multiples posée comme première, deuxième et troisième plus importantes raisons pour décider d'un achat précis (par exemple, le prix, la recommandation personnelle, revue, fournisseur local, etc.) Dans ce cas, vous pouvez déterminer l'importance du prix en calculant l'index du champ dans lequel il a été inclus pour la première fois :

```
first_index("price", [Reason1 Reason2 Reason3])
```

De même, supposez que vous avez demandé à vos clients de classer trois voitures par ordre de préférence et que vous avez codé les réponses en trois champs séparés, comme suit :

Tableau 9. Exemple de classement de voitures

id client	voiture1	voiture2	voiture3
101	1	3	2
102	3	2	1
103	2	3	1

Dans ce cas, vous pouvez déterminer l'index du champ de la voiture préférée (classée n°1 ou le rang le plus bas) à l'aide de la fonction min_index :

```
min_index(['car1' 'car2' 'car3'])
```

Pour plus d'informations, voir la rubrique «Fonctions de comparaison», à la page 164.

Référencement des ensembles à réponses multiples

La fonction spéciale `@MULTI_RESPONSE_SET` peut être utilisée pour référencer tous les champs dans un ensemble à réponses multiples. Par exemple, si les trois champs *voiture* de l'exemple précédent sont compris dans un ensemble à réponses multiples appelé *car_rankings*, le même résultat serait renvoyé :

```
max_index(@MULTI_RESPONSE_SET("car_rankings"))
```

Générateur de formules

Vous pouvez saisir des expressions CLEM manuellement ou utiliser le générateur de formules qui affiche une liste complète de fonctions et d'opérateurs CLEM, ainsi que des champs de données du flux en cours, vous permettant de construire rapidement des expressions sans mémoriser les noms exacts des champs ou des fonctions. Par ailleurs, les commandes du Générateur ajoutent automatiquement les guillemets appropriés autour des champs et des valeurs, ce qui facilite la création d'expressions à la syntaxe correcte.

Remarque : il n'est pas possible d'utiliser la boîte de dialogue Générateur de formules pour générer des scripts ou définir des paramètres.

Remarque : Si vous désirez changer votre source de données, vérifiez d'abord que le Générateur de formules gère toujours les fonctions que vous avez sélectionnées. Comme certaines bases de données ne gèrent pas toutes les fonctions, vous risquez de rencontrer une erreur lors de l'exécution face à une nouvelle source de données.

Accès au Générateur de formules

Le Générateur de formules est disponible dans tous les noeuds où des expressions CLEM sont utilisées, y compris les noeuds Sélectionner, Equilibrer, Calculer, Remplacer, Analyse, Rapport et Table. Vous pouvez l'ouvrir en cliquant sur le bouton représentant une calculatrice situé juste à droite du champ de formule.

Création d'expressions

Non seulement la boîte de dialogue Générateur de formules contient des listes complètes de champs, de fonctions et d'opérateurs, mais elle vous permet également d'accéder à des valeurs de données si vos données sont instanciées.

Pour créer une formule à l'aide du générateur de formules

1. Saisissez votre expression dans le champ en vous servant des listes de fonctions et de champs comme références.

   ou

2. Sélectionnez les champs et fonctions requis dans les listes déroulantes.
3. Pour ajouter le champ ou la fonction au champ des formules, cliquez sur le bouton en forme de flèche jaune ou double-cliquez directement sur la fonction ou le champ.
4. Utilisez les boutons des opérandes situés au milieu de la boîte de dialogue pour insérer les opérandes dans l'expression.

Sélection de fonctions

La liste des fonctions affiche toutes les fonctions et tous les opérateurs CLEM disponibles. Faites défiler la liste pour sélectionner une fonction ou utilisez la liste déroulante pour afficher un sous-ensemble des fonctions ou des opérateurs. Les fonctions disponibles sont regroupées en catégories pour faciliter vos recherches.

La plupart de ces catégories sont décrites dans la section Référence de la description de langage CLEM. Pour obtenir plus d'informations, reportez-vous au «Informations de référence sur les fonctions», à la page 160.

Les autres catégories sont les suivantes.

- La catégorie **Fonctions générales** contient une sélection des fonctions les plus couramment utilisées.
- **Récemment utilisé** contient la liste des fonctions CLEM utilisées dans la session en cours.
- **@ Functions** contient une liste de toutes les fonctions spéciales, dont le nom est précédé par le signe @.

  Remarque : Les fonctions `@DIFF1(FIELD1,FIELD2)` et `@DIFF2(FIELD1,FIELD2)` requièrent que les deux types de champ soient identiques (par exemple, tous deux Entier ou Long ou Réel).
- **Fonctions de base de données**. Si le flux comprend une connexion à une base de données (via un noeud source SGBD), cette sélection répertorie les fonctions disponibles depuis la base de données, y compris les fonctions définies par l'utilisateur (FDU). Pour obtenir plus d'informations, reportez-vous au «Fonctions de base de données».
- **Aggrégats de base de données**. Si le flux comprend une connexion à une base de données (via un noeud source SGBD), cette sélection répertorie les options d'agrégation disponibles à partir de cette base de données. Ces options sont disponibles dans le Générateur de formules du noeud Agréger.
- **Aggrégats de fenêtre de base de données**. Si le flux comprend une connexion à une base de données (via un noeud source SGBD), cette sélection répertorie les options d'agrégation de fenêtre que vous pouvez utiliser avec cette base de données. Ces options ne sont disponibles dans le Générateur de formules que dans les noeuds de la palette **Opérations sur les champs**.

  Remarque : Comme SPSS Modeler obtient les **fonctions Agrégat de fenêtres** de la vue Système de base de données, les options disponibles dépendent du comportement de la base de données.

  Bien qu'elles soient appelées "aggrégats",; ces options ne sont pas destinées à une utilisation dans le noeud Agréger ; elles s'appliquent plutôt aux noeuds Dériver ou Sélectionner. Cedi est dû au fait que leurs sorties sont scalaires et non pas de véritables agrégats, ce qui signifie qu'elles ne réduisent pas la quantité de données affichées dans la sortie de la même façon que le noeud Agréger. Par exemple, vous pouvez utiliser ce type d'agrégation pour fournir une moyenne mobile via des lignes de données, par exemple la "moyenne de la ligne en cours et de toutes les lignes précédentes".
- **Agrégats intégrés**. Contient une liste des modes d'agrégation possibles qui peuvent être utilisés.
- La catégorie **Opérateurs** répertorie tous les opérateurs disponibles lors de la génération d'expressions. Les opérateurs sont également disponibles via les boutons situés au centre de la boîte de dialogue.
- **Toutes les fonctions** contient la liste complète des fonctions CLEM disponibles.

Une fois que vous avez sélectionné un groupe de fonctions, double-cliquez dessus pour les insérer dans le champ des formules à l'endroit indiqué par la position du curseur.

Fonctions de base de données

Les fonctions de base de données peuvent être répertoriées dans de nombreux emplacements ; le tableau ci-après affiche les emplacements dans lesquels SPSS Modeler effectue les recherches lorsqu'il doit trouver les détails d'une fonction. Les administrateurs de base de données peuvent se servir de ce tableau pour vérifier que les utilisateurs disposent des privilèges d'accès aux zones requises leur permettant d'utiliser les différentes fonctions.

De plus, le tableau répertorie les conditions qui sont utilisées pour le filtrage lorsqu'une fonction est disponible en vue de son utilisation, selon le type de base de données et de fonction.

Remarque : Si vous utilisez des fonctions de base de données d'Amazon Redshift, il se peut que votre administrateur de base de données doive vous octroyer des droits d'accès aux six objets de base de données ci-après. Les quatre premiers sont des tables de catalogue système et les deux derniers, des schémas.

- pg_type
- pg_proc
- pg_namespace
- pg_aggregate

- information_schema
- pg_catalog

Tableau 10. Fonctions de base de données dans le générateur de formules

Base de données	Type de fonction	Où trouver les fonctions	Conditions de filtrage des fonctions
Db2 LUW	UDF	SYSCAT.ROUTINES SYSCAT.ROUTINEPARMS	ROUTINETYPE est F et FUNCTIONTYPE est S
Db2 LUW	UDA	SYSCAT.ROUTINES SYSCAT.ROUTINEPARMS	ROUTINETYPE est F et FUNCTIONTYPE est C
Db2 iSeries	UDF	QSYS2.SYSROUTINES QSYS2.SYSPARMS	ROUTINE_TYPE est F et FUNCTION_TYPE est S
Db2 iSeries	UDA	QSYS2.SYSROUTINES QSYS2.SYSPARMS	ROUTINE_TYPE est F et FUNCTION_TYPE est C
Db2 z/OS	UDF	SYSIBM.SYSROUTINES SYSIBM.SYSPARMS	ROUTINETYPE est F et FUNCTIONTYPE est S
Db2 z/OS	UDA	SYSIBM.SYSROUTINES SYSIBM.SYSPARMS	ROUTINETYPE est F et FUNCTIONTYPE est C
SQL Server	UDF	SYS.ALL_OBJECTS SYS.ALL_PARAMETERS SYS.TYPES	TYPE est FN ou FS
SQL Server	UDA	SYS.ALL_OBJECTS SYS.ALL_PARAMETERS SYS.TYPES	TYPE est AF
Oracle	UDF	ALL_ARGUMENTS ALL_PROCEDURES	Toutes les conditions suivantes sont remplies : • OBJECT_TYPE est FUNCTION • AGGREGATE est NO • PLS_TYPE n'est pas NULL
Oracle	UDA	ALL_ARGUMENTS ALL_PROCEDURES	Toutes les conditions suivantes sont remplies : • ARGUMENT_NAME est NULL • AGGREGATE est YES • PLS_TYPE n'est pas NULL
Teradata	UDF	DBC.FUNCTIONS DBC.ALLRIGHTS	Toutes les conditions suivantes sont remplies : • FUNCTIONTYPE est F • COLUMNNAME est RETURN0 • SPPARAMETERTYPE est O • ACCESSRIGHT est EF
Teradata	UDA	DBC.FUNCTIONS DBC.ALLRIGHTS	Toutes les conditions suivantes sont remplies : • FUNCTIONTYPE est A • COLUMNNAME est RETURN0 • SPPARAMETERTYPE est O • ACCESSRIGHT est EF

Tableau 10. Fonctions de base de données dans le générateur de formules (suite)

Base de données	Type de fonction	Où trouver les fonctions	Conditions de filtrage des fonctions
Netezza	UDF	####.._V_FUNCTION NZA.._V_FUNCTION INZA.._V_FUNCTION	Pour ####.._V_FUNCTION, les conditions suivantes s'appliquent : • RESULT ne contient pas de chaîne avec des valeurs telles que : TABLE% • FUNCTION ne contient pas de chaîne avec des valeurs telles que : '/_%' escape '/' • VARARGS est FALSE Pour NZA.._V_FUNCTION et INZA.._V_FUNCTION, les conditions suivantes s'appliquent : • RESULT ne contient pas de chaîne avec des valeurs telles que : TABLE% • FUNCTION ne contient pas de chaîne avec des valeurs telles que : '/_%' escape '/' • BUILTIN est f • VARARGS est FALSE
Netezza	UDA	####.._V_AGGREGATE NZA.._V_FUNCTION INZA.._V_FUNCTION	Les deux conditions suivantes sont remplies : • AGGTYPE est ANY ou GROUPED • VARARGS est FALSE
Netezza	WUDA	####.._V_AGGREGATE NZA.._V_FUNCTION INZA.._V_FUNCTION	Pour ####.._V_AGGREGATE, les conditions suivantes s'appliquent : • AGGTYPE est ANY ou ANALYTIC • AGGREGATE n'est pas MAX_LABEL • VARARGS est FALSE Pour NZA.._V_FUNCTION et INZA.._V_FUNCTION, les conditions suivantes s'appliquent : • AGGTYPE est ANY ou ANALYTIC • BUILTIN est f • VARARGS est FALSE

Signification des termes utilisés dans le tableau

- UDF Fonction définie par l'utilisateur
- UDA Agrégat défini par l'utilisateur
- WUDA Fonction d'agrégation de fenêtre définie par l'utilisateur
- #### Base de données à laquelle vous êtes connecté.

Sélection de champs, de paramètres et de variables globales

La liste des champs affiche tous les champs actuellement disponibles dans le flux de données. Sélectionnez un champ dans la liste. Pour ajouter un champ à l'expression, cliquez sur le bouton en forme de flèche jaune ou double-cliquez directement sur le champ.

Pour plus d'informations, voir la rubrique «Paramètres de flux, de session et de super noeud», à la page 142.

En plus des champs, vous pouvez également choisir parmi les éléments suivants :

Ensembles à réponses multiples. Pour plus d'informations, reportez-vous au guide *IBM SPSS Modeler - Noeuds source, de processus et de sortie*.

Récemment utilisés contient une liste des champs, des ensembles à réponses multiples, des paramètres et des valeurs globales utilisés dans la session en cours.

Paramètres. Pour plus d'informations, voir la rubrique «Paramètres de flux, de session et de super noeud», à la page 142.

Valeurs globales. Pour plus d'informations, reportez-vous au guide *IBM SPSS Modeler - Noeuds source, de processus et de sortie*.

Affichage ou sélection de valeurs

Les valeurs de champs peuvent être affichées à partir de plusieurs endroits dans le système, y compris le générateur de formules, les rapports d'audits de données et lors de la modification de valeurs futures dans un noeud Intervalles de temps. Veuillez noter que pour utiliser cette fonction, les données doivent être entièrement instanciées dans une source ou un noeud Typer afin que le stockage, les types et les valeurs soient connus.

Pour afficher les valeurs d'un champ à partir du générateur de formules ou d'un noeud Intervalles de temps, sélectionnez le champ requis et cliquez sur le bouton de sélection des valeurs pour ouvrir une boîte de dialogue des valeurs du champ sélectionné. Vous pouvez ensuite sélectionner une valeur et cliquer sur **Insérer** pour coller la valeur dans la formule ou la liste en cours.

Figure 16. Bouton de sélection des valeurs

Pour les champs indicateurs et les champs nominaux, toutes les valeurs définies sont présentées. Pour les champs d'intervalles continus (intervalle numérique), les valeurs minimales et maximales sont affichées.

Vérification des expressions CLEM

Cliquez sur **Vérifier** dans le Générateur de formules (angle inférieur droit) pour valider l'expression. Expressions qui n'ont pas encore été vérifiées et affichées en rouge. En cas d'erreurs, un message indiquant la cause s'affiche.

La vérification porte sur les éléments suivants :

- Utilisation correcte des guillemets dans les noms de valeurs et de champs
- Utilisation correcte des paramètres et des variables globales
- Utilisation correcte des opérateurs
- Existence des champs référencés
- Existence et définition des valeurs globales référencées

Si votre syntaxe comporte des erreurs, essayez de recréer l'expression en utilisant les listes et les boutons des opérateurs au lieu de taper l'expression manuellement. Cette méthode ajoute automatiquement les guillemets corrects pour les champs et les valeurs.

Prenez en compte les limitations suivantes lors de la construction d'expressions dans IBM Analytical Decision Management. Les expressions ne peuvent contenir aucun des éléments suivants :

- Une référence à un paramètre de flux IBM SPSS Modeler
- Une référence à une variable globale IBM SPSS Modeler
- Une référence à une fonction de base de données

- Une référence à l'une des fonctions de zone spéciales ou à la valeur de zone `@` :
  - `@TARGET`
  - `@PREDICTED`
  - `@FIELD`
  - `@PARTITION_FIELD`
  - `@TRAINING_PARTITION`
  - `@TESTING_PARTITION`
  - `@VALIDATION_PARTITION`

Rechercher et remplacer

La boîte de dialogue Rechercher/Remplacer est disponible aux emplacements où vous modifier un script ou le texte d'une expression, par exemple l'éditeur de script ou le générateur d'expression CLEM, ou encore lors de la définition d'un modèle dans le noeud Rapport. Lorsque vous éditez du texte dans l'un de ces champs, appuyez sur `Ctrl+F` pour accéder à la boîte de dialogue, en vous assurant que le curseur est centré sur une zone de texte. Si vous travaillez dans un noeud Remplacer, par exemple, vous pouvez accéder à la boîte de dialogue depuis toute zone de texte de l'onglet Paramètres, ou depuis le champ de texte du générateur d'expression.

1. Lorsque le curseur se trouve sur une zone de texte, appuyez sur Ctrl+F pour accéder à la boîte de dialogue Rechercher/Remplacer.
2. Entrez le texte que vous souhaitez rechercher ou faites un choix dans la liste déroulante des éléments récemment consultés.
3. Saisissez éventuellement un texte de remplacement.
4. Cliquez sur **Suivant** pour lancer la recherche.
5. Cliquez sur **Remplacer** pour remplacer la sélection en cours, ou **Remplacer tout** pour mettre à jour certaines instances ou les instances sélectionnées.
6. La boîte de dialogue se ferme après chaque opération. Appuyez sur F3 depuis toute zone de texte pour répéter la dernière opération de recherche ou sur Ctrl+F pour accéder de nouveau à la boîte de dialogue.

Options de recherche

Respecter la casse. Spécifie si l'opération de recherche est sensible à la casse ; par exemple, si *myvar* correspond à *myVar*. Le texte de remplacement est toujours inséré exactement tel qu'il a été saisi, quel que soit le réglage de ce paramètre.

Mot entier. Spécifie si l'opération de recherche doit porter sur le texte inséré dans des mots. Si cette option est sélectionnée, par exemple, une recherche portant sur *spider* ne produira pas la réponse *spiderman* ou *spider-man*.

Caractères génériques. Spécifie si la syntaxe des caractères génériques est utilisée (voir la section suivante). Lorsqu'elle est sélectionnée, l'option **Mot entier** est désactivée et sa valeur est ignorée.

Texte sélectionné. Contrôle la portée de la recherche lorsque vous utilisez l'option **Remplacer tout**.

Syntaxe des caractères génériques

Les caractères génériques vous permettent de rechercher des caractères spéciaux tels que les tabulations ou les sauts de ligne, des classes ou des intervalles de caractères telles que *a* à *d*, toute valeur numérique ou non et les limites telles que le début ou la fin d'une ligne. Les types d'expression suivants sont pris en charge :

Tableau 11. Correspondances de caractères

Caractères	Correspondances
x	Le caractère x
\\	Le caractère barre oblique inversée
\0n	Le caractère présentant la valeur octale 0n (0 <= n <= 7)
\0nn	Le caractère présentant la valeur octale 0nn (0 <= n <= 7)
\0mnn	Le caractère présentant la valeur octale 0mnn (0 <= m <= 3, 0 <= n <= 7)
\xhh	Le caractère présentant la valeur hexadécimale 0xhh
\uhhhh	Le caractère présentant la valeur hexadécimale 0xhhhh
\t	Le caractère tabulation ('\u0009')
\n	Le caractère saut de ligne (retour à la ligne) ('\u000A')
\r	Le caractère retour chariot ('\000D')
\f	Le caractère alimentation de formulaire ('\u000C')
\a	Le caractère alerte (sonnerie) ('\u0007')
\e	Le caractère d'échappement('\u001B')
\cx	Le caractère de contrôle correspondant à x

Tableau 12. Classes de caractères

Classes de caractères	Correspondances
[abc]	a, b ou c (classe simple)
[^abc]	Tout caractère excepté a, b ou c (soustraction)
[a-zA-Z]	Caractères de a à z ou de A à Z compris (intervalle)
[a-d[m-p]]	Caractères de a à d ou de m à p (union). Cette option peut aussi être spécifiée comme [a-dm-p]
[a-z&&[def]]	Caractères de a à z ainsi que d, e ou f (intersection)
[a-z&&[^bc]]	Caractères de a à z à l'exception de b et c (soustraction). Cette option peut aussi être spécifiée comme [ad-z]
[a-z&&[^m-p]]	Caractères de a à z, mais pas de m à p (soustraction). Cette option peut aussi être spécifiée comme [a-lq-z]

Tableau 13. Classes de caractères prédéfinies

Classes de caractères prédéfinies	Correspondances
.	Tout caractère (peut correspondre ou non aux terminaisons de ligne)
\d	Tout chiffre : [0-9]
\D	Un caractère non numérique : [^0-9]
\s	Un espace blanc : [\t\n\x0B\f\r]
\S	Un espace non blanc : [^\s]
\w	Un caractère alphanumérique : [a-zA-Z_0-9]
\W	Un caractère non alphanumérique : [^\w]

Tableau 14. Correspondances de limites

Correspondances de limites	Correspondances
^	Le début d'une ligne

Tableau 14. Correspondances de limites (suite)

Correspondances de limites	**Correspondances**
$	La fin d'une ligne
\b	Une limite de mot
\B	Une limite autre que celle d'un mot
\A	Le début de la saisie
\Z	La fin de la saisie à l'exception de la terminaison finale éventuelle
\z	La fin de la saisie

Chapitre 10. Référence du langage CLEM

Présentation de la référence CLEM

Cette section décrit le décrit le Langage de contrôle (CLEM) pour la Manipulation d'Expressions , puissant outil permettant d'analyser et de manipuler les données utilisées dans les flux IBM SPSS Modeler. Vous pouvez utiliser le langage CLEM au sein des noeuds pour exécuter entre autres les tâches suivantes : évaluation des conditions, calcul de valeurs, insertion de données dans les rapports.

Les expressions CLEM se composent de valeurs, de noms de champs, d'opérateurs et de fonctions. En utilisant une syntaxe correcte, vous pouvez créer des opérations de données nombreuses et puissantes.

Types de données CLEM

Les types de données CLEM peuvent se composer des éléments suivants :

- Entiers
- Réels
- Caractères
- Chaînes
- Listes
- Champs
- Date/Heure

Règles des guillemets

Même si IBM SPSS Modeler est flexible lorsque vous déterminez des champs, des valeurs, des chaînes et des paramètres utilisés dans une expression CLEM, les règles générales suivantes fournissent la liste des conseils à suivre lorsque vous créez des expressions :

- Chaînes : utilisez toujours des guillemets doubles lorsque vous écrivez des chaînes, comme `"Type 2"`. Vous pouvez également utiliser des guillemets simples mais cela risque de créer une confusion avec les champs placés entre guillemets.
- Champs : utilisez des guillemets simples, uniquement lorsque cela est rendu nécessaire par la présence d'espaces ou d'autres caractères spéciaux, comme `'Numéro de commande'`. Les champs placés entre guillemets mais non définis dans le jeu de données seront lus comme des chaînes.
- Paramètres : utilisez toujours des guillemets simples lorsque vous employez des paramètres, comme `'$P-threshold'`.
- Caractères : utilisez toujours des guillemets simples inversés (`), comme dans `stripchar(`d`, "drugA")`.

Ces règles sont détaillées dans les rubriques suivantes.

Entiers

Les entiers sont représentés par une séquence de chiffres décimaux. (Facultatif) Vous pouvez placer un signe moins (-) devant un entier afin d'obtenir un nombre négatif, par exemple : `1234`, `999`, `–77`.

Le langage CLEM permet de manipuler les entiers de précision arbitraire. La taille maximale de l'entier varie selon votre plate-forme. Si les valeurs sont trop élevées pour être affichées dans un champ de nombre entier, changez le type de champ en `Réel` pour restaurer la valeur.

Réels

Un nombre *Réel* est un nombre à virgule flottante. Les réels sont représentés par un ou plusieurs chiffres, suivis d'un séparateur décimal, suivi d'un ou de plusieurs autres chiffres. Les réels CLEM sont maintenus en double précision.

(Facultatif) Vous pouvez placer un signe moins (-) devant le réel afin d'obtenir un nombre négatif, par exemple : `1.234`, `0.999`, `–77.001`. Respectez la syntaxe <*nombre*> e <*exposant*> pour exprimer un nombre réel en notation exponentielle, par exemple : `1234.0e5`, `1.7e–2`. Lorsque l'application IBM SPSS Modeler lit des chaînes numériques en provenance de fichiers et les convertit automatiquement en nombres, les nombres sans chiffre avant ou après le séparateur décimal sont acceptés ; par exemple, `999.` ou `.11`. Cependant, ces formes sont incorrectes dans les expressions CLEM.

Remarque : lors du référencement de nombres réels dans les expressions CLEM, un point doit être utilisé comme séparateur décimal, quels que soient les paramètres pour le flux en cours ou les paramètres régionaux. Par exemple précisez

```
Na > 0.6
```

plutôt que :

```
Na > 0,6
```

Cela s'applique même lorsque le symbole décimal choisi dans la boîte de dialogue Propriétés du flux est la virgule. En outre, cela est conforme à l'instruction générale selon laquelle la syntaxe de code doit être indépendante des conventions et paramètres régionaux.

Caractères

Des caractères (habituellement désignés par `CHAR`) sont généralement employés avec une expression CLEM pour réaliser des tests sur les chaînes. Ainsi, vous pouvez utiliser la fonction `isuppercode` pour déterminer si le premier caractère d'une chaîne est une majuscule. L'expression CLEM suivante se sert d'un caractère pour indiquer que le test doit être effectué sur le premier caractère de la chaîne :

```
isuppercode(subscrs(1, "MyString"))
```

Pour exprimer le code (et non l'emplacement) d'un caractère particulier dans une expression CLEM, utilisez des guillemets simples inversés : `` `<*caractère*>` `` par exemple, `` `A` ``, `` `Z` ``.

Remarque : Il n'existe pas de type de stockage `CHAR` pour les champs. Par conséquent, si un champ est calculé ou rempli avec une expression générant un caractère `CHAR`, ce résultat est converti en chaîne.

Chaînes

En général, vous devez placer les chaînes entre guillemets doubles, Par exemple, `"c35product2"` et `"referrerID"`. Pour signaler des caractères spéciaux dans une chaîne, utilisez une barre oblique inverse, par exemple, `"\$65443"`. (Pour indiquer une barre oblique inversée, utilisez deux barres obliques inversées, \\.) Vous pouvez placer une chaîne entre guillemets simples mais elle risque alors d'être confondue avec un champ placé entre guillemets (`'referrerID'`). Pour plus d'informations, voir la rubrique «Fonctions sur chaînes», à la page 171.

Listes

Une liste est une séquence d'éléments triés, qui peuvent être de nature différente. Les listes sont placées entre crochets ([]), Par exemple, `[1 2 4 16]` et `["abc" "def"]`. Les listes ne sont pas utilisées pour les valeurs des champs IBM SPSS Modeler. Elles sont utilisées pour fournir des arguments aux fonctions, comme `member` et `oneof`.

Remarque : Les listes ne peuvent être composées que d'objets statiques (par exemple, une chaîne, un nombre ou un nom de zone) et ne doivent pas contenir d'appels à des fonctions.

Champs

Dans les expressions CLEM, les noms autres que les noms de fonction sont considérés comme des noms de champ. Vous pouvez les écrire comme suit : `Power`, `val27`, `state_flag`, etc. En revanche, si le nom commence par un chiffre ou contient un caractère non alphabétique, comme un espace (à l'exception du trait de soulignement), placez le nom entre guillemets simples, par exemple `'Power Increase'`, `'2nd answer'`, `'#101'`, `'$P-NextField'`.

Remarque : les champs placés entre guillemets mais non définis dans le jeu de données seront lus comme des chaînes.

Dates

Les calculs de date reposent sur la date de référence spécifiée dans la boîte de dialogue des propriétés du flux. La date de référence par défaut est le 1er janvier 1900.

Le langage CLEM prend en charge les formats de date suivants.

Tableau 15. Formats de l'heure en langage CLEM

Format	Exemples
JJMMAA	150163
MMJJAA	011563
AAMMJJ	630115
AAAAMMJJ	19630115
AAAAJJJ	Année à quatre chiffres, suivie d'une valeur à trois chiffres représentant le jour de l'année (par exemple, 2000032 correspond au 32e jour de l'année 2000, soit le 1er février 2000).
JOUR	Jour de la semaine utilisé dans les paramètres régionaux actuels (par exemple Monday, Tuesday, etc. (en anglais).
MOIS	Mois utilisé dans les paramètres régionaux actuels (par exemple, janvier, février, etc.
JJ/MM/AA	15/01/63
JJ/MM/AAAA	15/01/1963
MM/JJ/AA	01/15/63
MM/JJ/AAAA	01/15/1963
JJ-MM-AA	15-01-63
JJ-MM-AAAA	15-01-1963
MM-JJ-AA	01-15-63
MM-JJ-AAAA	01-15-1963
JJ.MM.AA	15.01.63
JJ.MM.AAAA	15.01.1963
MM.JJ.AA	01.15.63
MM.JJ.AAAA	01.15.1963
JJ-MOI-AA	15-JAN-63, 15-jan-63, 15-Jan-63
JJ/MOI/AA	15/JAN/63, 15/jan/63, 15/Jan/63
JJ.MOI.AA	15.JAN.63, 15.jan.63, 15.Jan.63
JJ-MOI-AAAA	15-JAN-1963, 15-jan-1963, 15-Jan-1963
JJ/MOI/AAAA	15/JAN/1963, 15/jan/1963, 15/Jan/1963
JJ.MOI.AAAA	15.JAN.1963, 15.jan.1963, 15.Jan.1963

Tableau 15. Formats de l'heure en langage CLEM (suite)

Format	Exemples
MOI AAAA	Jan 2004
t T AAAA	Date représentée par un chiffre (1–4) correspondant au trimestre, suivi de la lettre *T* et d'une année à quatre chiffres (par exemple, le 25 décembre 2004 est représenté par 4 Q 2004.
ss SM AAAA	Valeur à deux chiffres représentant la semaine de l'année, suivie des lettres *SM*, puis d'une année à quatre chiffres. La semaine de l'année est calculée en supposant que le premier jour de la semaine est un lundi et que la première semaine comprend au moins un jour.

Heure

Le langage CLEM prend en charge les formats d'heure suivants.

Tableau 16. Formats de l'heure en langage CLEM

Format	Exemples
HHMMSS	120112, 010101, 221212
HHMM	1223, 0745, 2207
MMSS	5558, 0100
HH:MM:SS	12:01:12, 01:01:01, 22:12:12
HH:MM	12:23, 07:45, 22:07
MM:SS	55:58, 01:00
(H)H:(M)M:(S)S	12:1:12, 1:1:1, 22:12:12
(H)H:(M)M	12:23, 7:45, 22:7
(M)M:(S)S	55:58, 1:0
HH.MM.SS	12.01.12, 01.01.01, 22.12.12
HH.MM	12.23, 07.45, 22.07
MM.SS	55.58, 01.00
(H)H.(M)M.(S)S	12.1.12, 1.1.1, 22.12.12
(H)H.(M)M	12.23, 7.45, 22.7
(M)M.(S)S	55.58, 1.0

Opérateurs CLEM

Les opérateurs suivants sont disponibles.

Tableau 17. Opérateurs de langage CLEM

Opération	Commentaires	Priorité (voir section suivante)
or	Utilisé entre deux expressions CLEM. Renvoie la valeur true (vrai) si l'une des deux, ou les deux expressions présentent la valeur true (vrai).	10
and	Utilisé entre deux expressions CLEM. Renvoie la valeur true (vrai) si les deux expressions présentent la valeur true (vrai).	9

Tableau 17. Opérateurs de langage CLEM (suite)

Opération	Commentaires	Priorité (voir section suivante)
=	Utilisé entre deux éléments comparables. Renvoie la valeur true (vrai) si ITEM1 est égal à ITEM2.	7
==	Identique à =.	7
/=	Utilisé entre deux éléments comparables. Renvoie la valeur true (vrai) si ITEM1 *n'est pas* égal à ITEM2.	7
/==	Identique à /=.	7
>	Utilisé entre deux éléments comparables. Renvoie la valeur true (vrai) si ITEM1 est supérieur à ITEM2.	6
>=	Utilisé entre deux éléments comparables. Renvoie la valeur true (vrai) si ITEM1 est supérieur ou égal à ITEM2.	6
<	Utilisé entre deux éléments comparables. Renvoie la valeur true (vrai) si ITEM1 est inférieur à ITEM2.	6
<=	Utilisé entre deux éléments comparables. Renvoie la valeur true (vrai) si ITEM1 est inférieur ou égal à ITEM2.	6
&&=_0	Utilisé entre deux entiers. Equivaut à l'expression booléenne INT1 && INT2 = 0.	6
&&/=_0	Utilisé entre deux entiers. Equivaut à l'expression booléenne INT1 && INT2 /= 0.	6
+	Additionne deux nombres : NUM1 + NUM2.	5
><	Concatène deux chaînes, par exemple, CHAINE1 >< CHAINE2.	5
-	Soustrait un nombre à un autre : NUM1 - NUM2. Peut être également utilisé devant un nombre : - NUM.	5
*	Multiplie deux nombres : NUM1 * NUM2.	4
&&	Utilisé entre deux entiers. Le résultat est l'opération 'et' (and) au niveau des bits des entiers INT1 et INT2.	4
&&~~	Utilisé entre deux entiers. Le résultat est l'opération 'et' (and) au niveau des bits de INT1 et du complément au niveau des bits de INT2.	4
||	Utilisé entre deux entiers. Le résultat de l'opération est l'opérateur sur les bits 'OU inclusif' des entiers INT1 et INT2.	4
~~	Utilisé devant un entier. Produit le complément sur les bits de INT.	4
||/&	Utilisé entre deux entiers. Le résultat de l'opération est l'opérateur sur les bits 'OU exclusif' des entiers INT1 et INT2.	4
INT1 << N	Utilisé entre deux entiers. Produit le motif binaire de l'entier INT décalé vers la gauche de N positions.	4

Tableau 17. Opérateurs de langage CLEM (suite)

Opération	Commentaires	Priorité (voir section suivante)
`INT1 >> N`	Utilisé entre deux entiers. Produit le motif binaire de l'entier INT décalé vers la droite de N positions.	4
`/`	Utilisé pour diviser un nombre par un autre : NUM1 / NUM2.	4
`**`	Utilisé entre deux nombres : BASE ** POWER. Renvoie la valeur de base BASE élevée à la puissance POWER.	3
`rem`	Utilisé entre deux entiers : INT1 rem INT2. Renvoie le reste, INT1 - (INT1 div INT2) * INT2.	2
`div`	Utilisé entre deux entiers : INT1 div INT2. Effectue la division des entiers.	2

Priorité des opérateurs

La priorité des opérateurs détermine l'ordre dans lequel les opérations des expressions complexes sont traitées, notamment dans les expressions sans parenthèses avec plusieurs opérateurs infixes. Par exemple, `3 + 4 * 5`

est calculé de la manière suivante : `3 + (4 * 5)` et non `(3 + 4) * 5`. En effet, selon la règle de priorité, la multiplication est prioritaire sur l'addition. A chaque opérateur du langage CLEM est affectée une valeur de priorité. Plus cette valeur est faible, plus l'opérateur est prioritaire. Il sera donc traité avant ceux ayant un niveau de priorité plus élevé.

Informations de référence sur les fonctions

Les fonctions CLEM suivantes sont disponibles lorsque vous travaillez avec des données dans IBM SPSS Modeler. Vous pouvez saisir ces fonctions sous forme de code dans diverses boîtes de dialogue, par exemple dans les noeuds Dériver et Binariser, ou utiliser le Générateur de formules pour créer des expressions CLEM valides sans avoir à mémoriser des listes de fonctions ou des noms de champs.

Tableau 18. Fonctions CLEM à utiliser avec les données IBM SPSS Modeler

Type de fonction	Description
Information	Servent à obtenir un aperçu des valeurs de champ. Par exemple, la fonction `is_string` renvoie une valeur true (vrai) pour tous les enregistrements de type chaîne.
Conversion	Servent à créer des champs ou à convertir un type de stockage. Par exemple, la fonction `to_timestamp` convertit le champ sélectionné en un horodatage.
Comparaison	Servent à comparer des valeurs de champ, entre elles ou par rapport à la chaîne spécifiée. Par exemple, la fonction `<=` permet de savoir si la valeur d'un champ est inférieure ou égale à celle d'un autre.
Logique	Servent à exécuter des opérations logiques, comme `if`, `then` et `else`.
Numérique	Servent à effectuer des calculs numériques, comme le calcul du logarithme naturel de valeurs de champ.
Trigonométrique	Servent à effectuer des calculs trigonométriques, comme le calcul du cosinus inverse d'un angle.

Tableau 18. Fonctions CLEM à utiliser avec les données IBM SPSS Modeler (suite)

Type de fonction	Description
Probabilité	Renvoie des probabilités basées sur diverses distributions, comme la probabilité qu'une valeur de la fonction de distribution Student *t* soit inférieure à une valeur spécifique.
Spatiale	Utilisée pour effectuer des calculs spatials sur des données géospatiales.
Sur les bits	Servent à manipuler des nombres entiers comme éléments binaires.
Aléatoire	Servent à sélectionner des éléments ou à générer des nombres de façon aléatoire.
Chaîne	Utilisée pour effectuer diverses opérations sur des chaînes, telles que `stripchar`, laquelle vous permet de supprimer un caractère spécifié.
SoundEx	Sert à trouver des chaînes lorsque l'orthographe exacte est inconnue, sur la base d'hypothèses phonétiques (modalités de prononciation de certaines lettres).
Date et heure	Permet d'effectuer diverses opérations sur ces champs de date, heure et horodatage.
Séquence	Utilisée pour défricher la séquence des enregistrements d'un jeu de données ou pour effectuer des opérations basées sur cette séquence.
Globale	Utilisée pour accéder à des valeurs globales créées par un noeud Valeurs globales. Par exemple, `@MEAN` sert à faire référence à la moyenne de toutes les valeurs d'un champ dans tout le jeu de données.
Nul et non renseigné	Servent à atteindre, à marquer et souvent à remplir des blancs définis par l'utilisateur ou des valeurs système manquantes. Par exemple, `@BLANK(FIELD)` permet d'afficher un indicateur ayant la valeur true (vrai) pour les enregistrements qui comportent des blancs.
Champs spéciaux	Servent à signaler les champs spéciaux en cours d'examen. Par exemple, `@FIELD` sert lors du calcul de plusieurs champs.

Conventions pour la description des fonctions

Les conventions suivantes sont utilisées dans ce guide pour faire référence aux éléments d'une fonction.

Tableau 19. Conventions pour la description des fonctions

Convention	Description
BOOL	Un indicateur, comme true ou false (vrai, faux).
NUM, NUM1, NUM2	Tout nombre.
REAL, REAL1, REAL2	Nombre réel, comme `1,234` ou `–77,01`.
INT, INT1, INT2	Entier, comme 1 ou `–77`.
CHAR	Code de caractère, comme `` `A` ``.
STRING	Chaîne, comme `"referrerID"`.
LIST	Liste d'éléments, comme `["abc" "def"]`.
ITEM	Champ, comme `Customer` ou `extract_concept`.
DATE	Champ de date, comme `start_date`, dont les valeurs sont au format `DD-MON-YYYY`.
TIME	Champ d'heure, comme `power_flux`, dont les valeurs sont au format `HHMMSS`.

Les fonctions couvertes dans ce guide sont répertoriées dans un tableau à trois colonnes : la première colonne indique le nom de la fonction, la deuxième le type de résultat (entier, chaîne, etc.) et la troisième, une description (si elle est disponible). Par exemple, la description suivante est la description d'une fonction `rem`.

Tableau 20. rem fonction description

Fonction	Résultat	Description
`INT1 rem INT2`	*Nombre*	Renvoie le reste de la division de l'entier *INT1* par l'entier *INT2*. Par exemple, `INT1 – (INT1 div INT2) * INT2`.

Les informations relatives aux conventions d'utilisation, comme la manière de répertorier des éléments ou de spécifier les caractères d'une fonction, sont décrites dans une autre rubrique. Pour plus d'informations, voir la rubrique «Types de données CLEM», à la page 155.

Fonctions d'informations

Les fonctions d'informations sont utilisées pour obtenir un aperçu des valeurs d'un champ particulier. Elles sont généralement utilisées pour calculer des champs indicateurs. Par exemple, vous pouvez utiliser la fonction `@BLANK` pour créer un champ indicateur indiquant les enregistrements dont les valeurs sont vides pour le champ sélectionné. De même, vous pouvez vérifier le type de stockage d'un champ à l'aide de l'une des fonctions de type de stockage, comme `is_string`.

Tableau 21. fonctions d'informations CLEM.

Fonction	Résultat	Description
`@BLANK(FIELD)`	*Booléen*	Renvoie la valeur true (vrai) pour tous les enregistrements dont les valeurs sont vides, d'après les règles de traitement des blancs définies dans un noeud Typer ou un noeud source en amont (onglet Types).
`@NULL(ITEM)`	*Booléen*	Renvoie la valeur true (vrai) pour tous les enregistrements dont les valeurs ne sont pas définies. Les valeurs non définies sont des valeurs système nulles, affichées dans IBM SPSS Modeler sous la forme `$null$`.
`is_date(ITEM)`	*Booléen*	Renvoie la valeur true (vrai) pour tous les enregistrements de type date.
`is_datetime(ITEM)`	*Booléen*	Renvoie la valeur vraie (True) pour tous les enregistrements de type date, heure ou horodatage.
`is_integer(ITEM)`	*Booléen*	Renvoie la valeur vraie (True) pour tous les enregistrements de type entier.
`is_number(ITEM)`	*Booléen*	Renvoie la valeur true (vrai) pour tous les enregistrements de type nombre.
`is_real(ITEM)`	*Booléen*	Renvoie la valeur true (vrai) pour tous les enregistrements de type réel.
`is_string(ITEM)`	*Booléen*	Renvoie la valeur true (vrai) pour tous les enregistrements de type chaîne.
`is_time(ITEM)`	*Booléen*	Renvoie la valeur true (vrai) pour tous les enregistrements de type heure.
`is_timestamp(ITEM)`	*Booléen*	Renvoie la valeur true (vrai) pour tous les enregistrements de type horodatage.

Fonctions de conversion

Les fonctions de conversion permettent de créer des champs et de convertir le type de stockage des champs existants. Par exemple, vous pouvez former de nouvelles chaînes en joignant ou en séparant des

chaînes existantes. Pour joindre des chaînes, utilisez l'opérateur ><. Par exemple, si le champ Site a la valeur "BRAMLEY", alors "xx" >< Site renvoie "xxBRAMLEY". Le résultat de >< est toujours une chaîne, même si les arguments n'en sont pas. Ainsi, si le champ V1 a la valeur 3 et si le champ V2 contient 5, V1 >< V2 renvoie "35" (qui est une chaîne, pas un nombre).

Les fonctions de conversion (et toutes les autres fonctions qui nécessitent un type spécifique d'entrée, par exemple une valeur de date ou d'heure) dépendent des formats actuels indiqués dans la boîte de dialogue des options de flux. Par exemple, si vous souhaitez convertir un champ de type chaîne avec des valeurs *Jan 2003*, *Fév 2003*, etc., sélectionnez **MOIS AAAA** comme format de date par défaut pour le flux.

Tableau 22. Fonctions de conversion CLEM

Fonction	Résultat	Description
ITEM1 >< ITEM2	*Chaîne*	Concatène les valeurs pour deux champs et renvoie la chaîne résultante sous la forme *ITEM1ITEM2*.
to_integer(ITEM)	*Entier*	Convertit le stockage du champ spécifié en un entier.
to_real(ITEM)	*Réel*	Convertit le stockage du champ spécifié en un réel.
to_number(ITEM)	*Nombre*	Convertit le stockage du champ spécifié en un nombre.
to_string(ITEM)	*Chaîne*	Convertit le stockage du champ spécifié en une chaîne. Lorsqu'un réel est converti en chaîne via cette fonction, il renvoie une valeur avec 6 chiffres après le point radix.
to_time(ITEM)	*Temps*	Convertit le stockage du champ spécifié en une heure.
to_date(ITEM)	*Date*	Convertit le stockage du champ spécifié en une date.
to_timestamp(ITEM)	*Horodatage*	Convertit le stockage du champ spécifié en un horodatage.
to_datetime(ITEM)	*Dateheure*	Convertit le stockage du champ spécifié en une valeur d'heure, de date ou d'horodatage.
datetime_date(ITEM)	*Date*	Renvoie la valeur de date pour un *nombre*, une *chaîne*, ou un *horodatage*. Veuillez noter que cette fonction est la seule fonction vous permettant de reconvertir un nombre (en secondes) en date. Si ITEM est une chaîne, crée une date en analysant une chaîne au format de date sélectionné. Pour que cette fonction réussisse, vous devez spécifier un format de date correct dans la boîte de dialogue des propriétés du flux. Si ITEM est un nombre, il est interprété comme un nombre de secondes écoulées depuis la date de départ (ou période). Les fractions de jour sont supprimées. Si ITEM est un horodatage, la partie date de cet horodatage est renvoyée. Si ITEM est une date, il est renvoyé sans modification.
stb_centroid_latitude(ITEM)	*Entier*	Renvoie une valeur d'entier pour la latitude correspondant au centroïde de l'argument geohash.
stb_centroid_longitude(ELEMENT)	*Entier*	Renvoie une valeur d'entier pour la longitude correspondant au centroïde de l'argument geohash.

Tableau 22. Fonctions de conversion CLEM (suite)

Fonction	Résultat	Description
`to_geohash(ELEMENT)`	*Chaîne*	Renvoie la chaîne géo-hachée correspondant à la latitude et à la longitude avec le nombre de bits spécifié pour la densité. Un géo-hachage est un code utilisé pour identifier un jeu de coordonnées géographiques basé sur les informations de latitude et de longitude. Les trois paramètres pour `to_geohash` sont les suivants : • *latitude* : plage (-180, 180) ; les unités sont les degrés dans le système de coordonnées WGS84 • *longitude* : plage (-90, 90) ; les unités sont les degrés dans le système de coordonnées WGS84 • *bits* : nombre de bits à utiliser pour stocker le hachage. Plage [1,75]. Ce paramètre a un impact sur la longueur de la chaîne renvoyée (1 caractère utilisé pour 5 bits) et l'exactitude du hachage. Par exemple, 5 bits (1 caractère) représentent environ 2500 kilomètres ou 45 bits (9 caractères) représentent environ 2,3 mètres.

Fonctions de comparaison

Les fonctions de comparaison sont utilisées pour comparer des valeurs de champ, entre elles ou par rapport à la chaîne spécifiée. Par exemple, vous pouvez vérifier l'égalité de deux chaînes en utilisant `=`. Voici un exemple de vérification de l'égalité des chaînes : `Class = "class 1"`.

Lors d'une comparaison de nombres, *supérieur à* signifie que le nombre est plus proche de l'infini positif et *inférieur à* signifie qu'il se rapproche de l'infini négatif. Autrement dit, tous les nombres négatifs sont inférieurs à tous les nombres positifs.

Tableau 23. Fonctions de comparaison CLEM

Fonction	Résultat	Description
`count_equal(ITEM1, LIST)`	*Entier*	Renvoie le nombre de valeurs, issues d'une liste de champs, égales à *ITEM1* ou une valeur nulle si *ITEM1* est nul.
`count_greater_than(ITEM1, LIST)`	*Entier*	Renvoie le nombre de valeurs, issues d'une liste de champs, supérieures à *ITEM1* ou une valeur nulle si *ITEM1* est nul.
`count_less_than(ITEM1, LIST)`	*Entier*	Renvoie le nombre de valeurs, issues d'une liste de champs, inférieures à *ITEM1* ou une valeur nulle si *ITEM1* est nul.
`count_not_equal(ITEM1, LIST)`	*Entier*	Renvoie le nombre de valeurs, issues d'une liste de champs, différentes de *ITEM1* ou une valeur nulle si *ITEM1* est nul.
`count_nulls(LIST)`	*Entier*	Renvoie le nombre de valeurs nulles issues d'une liste de champs.
`count_non_nulls(LIST)`	*Entier*	Renvoie le nombre de valeurs non nulles issues d'une liste de champs.
`date_before(DATE1, DATE2)`	*Booléen*	Utilisée pour vérifier l'ordre des valeurs de date. Renvoie une valeur true (vrai) si *DATE1* est antérieure à *DATE2*.
`first_index(ITEM, LIST)`	*Entier*	Renvoie l'index du premier champ contenant ITEM depuis une LISTE de champs ou 0 si aucune valeur n'est trouvée. Pris en charge uniquement pour les types de chaînes, d'entiers et de réels.
`first_non_null(LIST)`	*Tout*	Renvoie la première valeur non nulle de la liste de champs fournie. Tous les types de stockage sont pris en charge.

Tableau 23. Fonctions de comparaison CLEM (suite)

Fonction	Résultat	Description
`first_non_null_index(LIST)`	*Entier*	Renvoie l'index du premier champ dans la LISTE spécifiée contenant une valeur non nulle ou 0 si toutes les valeurs sont nulles. Tous les types de stockage sont pris en charge.
`ITEM1 = ITEM2`	*Booléen*	Renvoie la valeur true (vrai) pour les enregistrements où *ITEM1* est égal à *ITEM2*.
`ITEM1 /= ITEM2`	*Booléen*	Renvoie la valeur true (vrai) si les deux chaînes sont différentes ou renvoie 0 si elles sont identiques.
`ITEM1 < ITEM2`	*Booléen*	Renvoie la valeur true (vrai) pour les enregistrements où *ITEM1* est inférieur à *ITEM2*.
`ITEM1 <= ITEM2`	*Booléen*	Renvoie la valeur true (vrai) pour les enregistrements où *ITEM1* est inférieur ou égal à *ITEM2*.
`ITEM1 > ITEM2`	*Booléen*	Renvoie la valeur true (vrai) pour les enregistrements où *ITEM1* est supérieur à *ITEM2*.
`ITEM1 >= ITEM2`	*Booléen*	Renvoie la valeur true (vrai) pour les enregistrements où *ITEM1* est supérieur ou égal à *ITEM2*.
`last_index(ITEM, LIST)`	*Entier*	Renvoie l'index du dernier champ contenant ITEM depuis une LISTE de champs ou 0 si aucune valeur n'est trouvée. Pris en charge uniquement pour les types de chaînes, d'entiers et de réels.
`last_non_null(LIST)`	*Tout*	Renvoie la dernière valeur non nulle de la liste de champs fournie. Tous les types de stockage sont pris en charge.
`last_non_null_index(LIST)`	*Entier*	Renvoie l'index du dernier champ dans la LISTE spécifiée contenant une valeur non nulle ou 0 si toutes les valeurs sont nulles. Tous les types de stockage sont pris en charge.
`max(ITEM1, ITEM2)`	*Tout*	Renvoie le plus grand des deux éléments *ITEM1* et *ITEM2*.
`max_index(LIST)`	*Entier*	Renvoie l'index du champ contenant la valeur maximale d'une liste de champs numériques ou 0 si toutes les valeurs de champ sont nulles. Par exemple, si le troisième champ de la liste contient la valeur maximale, la valeur 3 de l'index est renvoyée. Si plusieurs champs contiennent la valeur maximale, le premier de la liste (le plus à gauche) est renvoyé.
`max_n(LIST)`	*Nombre*	Renvoie la valeur maximale d'une liste de champs numériques ou une valeur nulle si toutes les valeurs de champ sont nulles.
`member(ITEM, LIST)`	*Booléen*	Renvoie la valeur true (vrai) si *ITEM* est un membre de la liste (*LIST*) spécifiée. Sinon, renvoie une valeur fausse (False). Il est également possible d'indiquer une liste de noms de champ.
`min(ITEM1, ITEM2)`	*Tout*	Renvoie le plus petit des deux éléments *ITEM1* et *ITEM2*.
`min_index(LIST)`	*Entier*	Renvoie l'index du champ contenant la valeur minimale d'une liste de champs numériques ou 0 si toutes les valeurs de champ sont nulles. Par exemple, si le troisième champ de la liste contient la valeur minimale, la valeur 3 de l'index est renvoyée. Si plusieurs champs contiennent la valeur minimale, le premier de la liste (le plus à gauche) est renvoyé.
`min_n(LIST)`	*Nombre*	Renvoie la valeur minimale d'une liste de champs numériques ou une valeur nulle si toutes les valeurs de champ sont nulles.
`time_before(TIME1, TIME2)`	*Booléen*	Utilisée pour vérifier l'ordre des valeurs d'heure. Renvoie une valeur true (vrai) si *TIME1* est antérieure à *TIME2*.

Tableau 23. Fonctions de comparaison CLEM (suite)

Fonction	Résultat	Description
`value_at(INT, LIST)`		Renvoie la valeur de décalage INT ou NULL de chaque champ de la liste si le décalage est en-dehors de l'intervalle de valeurs valides (c'est-à-dire, inférieur à 1 ou supérieur au nombre de champs de la liste). Tous les types de stockage sont pris en charge.

Fonctions logiques

Vous pouvez utiliser des expressions CLEM pour effectuer des opérations logiques.

Tableau 24. Fonctions logiques CLEM

Fonction	Résultat	Description
`COND1 et COND2`	*Booléen*	Cette opération est une conjonction logique et renvoie une valeur true (vrai) si *COND1* et *COND2* sont toutes les deux vraies. Si *COND1* est fausse, *COND2* n'est pas évaluée ; il est alors possible d'avoir des conjonctions où *COND1* teste d'abord si une opération dans *COND2* est autorisée. Par exemple, `length(Label) >=6 et Label(6) = 'x'`.
`COND1 or COND2`	*Booléen*	Cette opération est une disjonction logique (inclusive) et renvoie une valeur true (vrai) si l'une des deux conditions *COND1* ou *COND2* est vraie ou si les deux le sont. Si *COND1* est vraie, *COND2* n'est pas évaluée.
`not(COND)`	*Booléen*	Cette opération est une négation logique et renvoie une valeur true (vrai) si *COND* est fausse. Sinon, cette opération renvoie la valeur 0.
`if COND then EXPR1 else EXPR2 endif`	*Tout*	Cette opération est une évaluation conditionnelle. Si *COND* est vraie, cette opération renvoie le résultat de l'évaluation de l'expression *EXPR1*. Sinon, elle renvoie le résultat de l'évaluation de l'expression *EXPR2*.
`if COND1 then EXPR1 elseif COND2 then EXPR2 else EXPR_N endif`	*Tout*	Cette opération est une évaluation conditionnelle à plusieurs branches. Si *COND1* est vraie, cette opération renvoie le résultat de l'évaluation de l'expression *EXPR1*. En revanche, si *COND2* est vraie, l'opération renvoie le résultat de l'évaluation de l'expression *EXPR2*. Sinon, le résultat de l'évaluation de l'expression *EXPR_N* est renvoyé.

Fonctions numériques

CLEM contient un certain nombre de fonctions numériques courantes.

Tableau 25. Fonctions numériques CLEM

Fonction	Résultat	Description
`–NUM`	*Nombre*	Utilisée pour inverser *NUM*. Renvoie le même nombre, avec le signe opposé.
`NUM1 + NUM2`	*Nombre*	Renvoie la somme de *NUM1* et *NUM2*.
`NUM1 –NUM2`	*Nombre*	Renvoie la valeur *NUM2*, soustraite de *NUM1*.
`NUM1 * NUM2`	*Nombre*	Renvoie la valeur *NUM1* multipliée par *NUM2*.
`NUM1 / NUM2`	*Nombre*	Renvoie la valeur *NUM1* divisée par *NUM2*.
`INT1 div INT2`	*Nombre*	Utilisée pour la division en nombres entiers. Renvoie la valeur de l'entier *INT1* divisé par l'entier *INT2*.

Tableau 25. Fonctions numériques CLEM (suite)

Fonction	Résultat	Description
`INT1 rem INT2`	*Nombre*	Renvoie le reste de la division de l'entier *INT1* par l'entier *INT2*. Par exemple, `INT1 – (INT1 div INT2) * INT2`.
`INT1 mod INT2`	*Nombre*	Cette fonction a été remplacée. Utilisez la fonction `rem` à la place.
`BASE ** POWER`	*Nombre*	Renvoie la valeur *BASE* élevée à la puissance *POWER*. Les deux valeurs peuvent être tout nombre (en revanche, *BASE* ne doit pas avoir la valeur zéro si *POWER* est elle-même égale à une valeur nulle autre que l'entier 0). Si *POWER* est un entier, le calcul est effectué en multipliant successivement les puissances de *BASE*. Ainsi, si *BASE* est un entier, le résultat l'est également. Si *POWER* est l'entier 0, le résultat est toujours un 1 du même type que *BASE*. Sinon, si *POWER* n'est pas un entier, le résultat est calculé selon la formule `exp(POWER * log(BASE))`.
`abs(NUM)`	*Nombre*	Renvoie la valeur absolue de *NUM*, qui est toujours un nombre du même type.
`exp(NUM)`	*Réel*	Renvoie *e* élevé à la puissance *NUM*, *e* étant la base des logarithmes naturels.
`fracof(NUM)`	*Réel*	Renvoie la partie fractionnelle de *NUM*, définie selon la formule `NUM–intof(NUM)`.
`intof(NUM)`	*Entier*	Tronque son argument pour obtenir un entier. Il renvoie l'entier du même signe que *NUM* ayant la valeur absolue la plus élevée, comme `abs(INT) <= abs(NUM)`.
`log(NUM)`	*Réel*	Renvoie le logarithme naturel (base *e*) de *NUM*, qui ne doit en aucun cas avoir la valeur zéro.
`log10(NUM)`	*Réel*	Renvoie le logarithme décimal de *NUM* qui ne doit en aucun cas avoir la valeur zéro. Cette fonction est définie selon la formule `log(NUM) / log(10)`.
`negate(NUM)`	*Nombre*	Utilisée pour inverser *NUM*. Renvoie le même nombre, avec le signe opposé.
`round(NUM)`	*Entier*	Utilisée pour arrondir *NUM* à un entier en utilisant `intof(NUM+0.5)` si *NUM* est positif ou `intof(NUM–0.5)` si*NUM* est négatif.
`sign(NUM)`	*Nombre*	Permet de déterminer le signe de *NUM*. Cette opération renvoie 1, 0, ou 1 si *NUM* est un entier. Si *NUM* est un nombre réel, elle renvoie 1,0, 0,0 ou 1,0 selon que *NUM* est négatif, positif ou égal à zéro.
`sqrt(NUM)`	*Réel*	Renvoie la racine carrée de *NUM*. *NUM* doit être positif.
`sum_n(LIST)`	*Nombre*	Renvoie la somme des valeurs d'une liste de champs numériques ou une valeur nulle si toutes les valeurs de champ sont nulles.
`mean_n(LIST)`	*Nombre*	Renvoie la valeur moyenne d'une liste de champs numériques ou une valeur nulle si toutes les valeurs de champ sont nulles.
`sdev_n(LIST)`	*Nombre*	Renvoie l'écart-type d'une liste de champs numériques ou une valeur nulle si toutes les valeurs de champ sont nulles.

Fonctions trigonométriques

Toutes les fonctions de cette section utilisent un angle comme argument ou en renvoient un comme résultat. Dans les deux cas, les unités d'angle (radians ou degrés) sont contrôlées par les paramètres de l'option de flux appropriée.

Tableau 26. Fonctions trigonométriques CLEM

Fonction	Résultat	Description
`arccos(NUM)`	*Réel*	Calcule le cosinus inverse de l'angle spécifié.
`arccosh(NUM)`	*Réel*	Calcule le sinus inverse hyperbolique de l'angle spécifié.
`arcsin(NUM)`	*Réel*	Calcule le sinus inverse de l'angle spécifié.
`arcsinh(NUM)`	*Réel*	Calcule le sinus inverse hyperbolique de l'angle spécifié.
`arctan(NUM)`	*Réel*	Calcule la tangente inverse de l'angle spécifié.
`arctan2(NUM_Y, NUM_X)`	*Réel*	Calcule la tangente inverse de `NUM_Y / NUM_X` et utilise les signes des deux nombres pour calculer des informations du cadran. Le résultat est un réel dans l'intervalle `- pi < ANGLE <= pi (radians)` – `180 < ANGLE <= 180 (degrés)`
`arctanh(NUM)`	*Réel*	Calcule la tangente inverse hyperbolique de l'angle spécifié.
`cos(NUM)`	*Réel*	Calcule le cosinus de l'angle spécifié.
`cosh(NUM)`	*Réel*	Calcule le cosinus hyperbolique de l'angle spécifié.
`pi`	*Réel*	Cette constante est la meilleure approximation réelle de pi.
`sin(NUM)`	*Réel*	Calcule le sinus de l'angle spécifié.
`sinh(NUM)`	*Réel*	Calcule le sinus hyperbolique de l'angle spécifié.
`tan(NUM)`	*Réel*	Calcule la tangente de l'angle spécifié.
`tanh(NUM)`	*Réel*	Calcule la tangente hyperbolique de l'angle spécifié.

Fonctions de probabilité

Les fonctions de probabilité renvoient des probabilités basées sur différentes distributions, comme la probabilité qu'une valeur de la distribution *t* de Student soit inférieure à une valeur spécifique.

Tableau 27. Fonctions de probabilité CLEM

Fonction	Résultat	Description
`cdf_chisq(NUM, DF)`	*Réel*	Renvoie la probabilité qu'une valeur de la distribution du khi-deux, avec les degrés de liberté indiqués, soit inférieure au nombre spécifié.
`cdf_f(NUM, DF1, DF2)`	*Réel*	Renvoie la probabilité qu'une valeur de la distribution *F*, avec les degrés de liberté *DF1* et *DF2* soit inférieure au nombre spécifié.
`cdf_normal(NUM, MEAN, STDDEV)`	*Réel*	Renvoie la probabilité qu'une valeur de la distribution normale, avec la moyenne et l'écart-type indiqués, soit inférieure au nombre spécifié.
`cdf_t(NUM, DF)`	*Réel*	Renvoie la probabilité qu'une valeur de la distribution *t* de Student, avec les degrés de liberté indiqués, soit inférieure au nombre spécifié.

Fonctions spatiales

Les fonctions spatiales peuvent être utilisées avec des données géospatiales. Elles vous permettent, par exemple, de calculer la distance entre deux points, l'aire d'un polygone, etc. Certaines situations nécessitent également une fusion de plusieurs jeux de données géospatiales, basée sur un prédicat spatial (dans, à proximité de, etc.), qui peut être réalisée via une condition de fusion.

Ces fonctions spatiales opèrent en conjonction avec le système de coordonnées spécifié dans **Outils** > **Propriétés du flux** > **Options** > **Géospatial**.

Remarque : Ces fonctions spatiales ne s'appliquent pas aux données tridimensionnelles. Si des données tridimensionnelles sont importées dans le flux, seules les deux premières dimensions sont utilisées par ces fonctions. Les valeurs de l'axe des Z sont ignorées.

Tableau 28. Fonctions spatiales CLEM

Fonction	Résultat	Description
`close_to(FORME,FORME,NOMBRE)`	*Booléen*	Teste si deux formes sont à une certaine DISTANCE l'une de l'autre. Si un système de coordonnées projetées est utilisé, DISTANCE est exprimée en mètres. Si aucun système de coordonnées n'est utilisé, l'unité est arbitraire.
`crosses(FORME,FORME)`	*Booléen*	Teste si deux formes se croisent. Cette fonction convient pour deux formes de type chaîne, ou une chaîne et un polygone.
`overlap(FORME,FORME)`	*Booléen*	Teste l'existence d'une intersection entre 2 polygones et si l'intersection est intérieure aux deux formes.
`within(FORME,FORME)`	*Booléen*	Teste si l'intégralité de FORME1 est contenue dans un POLYGONE.
`area(FORME)`	*Réel*	Renvoie l'aire du POLYGONE spécifié. Si un système projeté est utilisé, la fonction renvoie des mètres carrés. Si aucun système de coordonnées n'est utilisé, l'unité est arbitraire. La forme doit être un POLYGONE ou un MULTIPOLYGONE.
`num_points(FORME,LISTE)`	*Entier*	Renvoie le nombre de points d'un champ de points (MULTIPOINT) qui sont contenus dans les limites d'un POLYGONE. FORME1 doit être un POLYGONE ou un MULTIPOLYGONE.
`distance(FORME,FORME)`	*Réel*	Renvoie la distance entre FORME1 et FORME2. Si un système projeté est utilisé, la fonction renvoie des mètres. Si aucun système de coordonnées n'est utilisé, l'unité est arbitraire. FORME1 et FORME2 peuvent être n'importe quel type de géo-mesure.

Opérations d'entiers sur les bits

Ces fonctions permettent aux entiers d'être manipulés en tant que profils binaires représentant des valeurs de complément à 2, où la position binaire `N` a la pondération `2**N`. Les bits sont numérotés en commençant par 0. Ces opérations se comportent comme si le bit de signe d'un entier était décalé vers la gauche à l'infini. Ainsi, partout au-delà du bit le plus significatif, un entier positif a le bit 0 et un entier négatif a le bit 1.

Tableau 29. Opérations d'entiers sur les bits CLEM

Fonction	Résultat	Description
`~~ INT1`	*Entier*	Produit le complément sur les bits de l'entier *INT1*. C'est-à-dire qu'il existe un 1 dans le résultat pour chaque position binaire pour laquelle *INT1* contient un 0. `~~ INT = -(INT + 1) est toujours vraie.`
`INT1 || INT2`	*Entier*	Le résultat de cette opération est l'opérateur sur les bits "inclusif ou" des entiers *INT1* et *INT2*. Autrement dit, un 1 est inséré dans le résultat pour chaque position binaire où existe un 1 dans *INT1*, dans *INT2* ou dans les deux entiers.
`INT1 ||/& INT2`	*Entier*	Le résultat de cette opération est l'opérateur sur les bits "exclusif ou" des entiers *INT1* et *INT2*. Autrement dit, un 1 est inséré dans le résultat pour chaque position binaire où existe un 1 soit dans *INT1*, soit dans *INT2*, mais pas dans les deux.

Tableau 29. Opérations d'entiers sur les bits CLEM (suite)

Fonction	Résultat	Description
`INT1 && INT2`	*Entier*	Produit l'opération booléenne ET sur les bits des entiers *INT1* et *INT2*. Cela signifie qu'un 1 est présent dans le résultat pour chaque position binaire où il existe un 1 dans les deux entiers *INT1* et *INT2*.
`INT1 &&~~ INT2`	*Entier*	Produit le booléen ET de l'entier *INT1* et le complément sur les bits de l'entier *INT2*. Cela signifie qu'un 1 est présent dans le résultat pour chaque position binaire où il existe un 1 dans *INT1* et un 0 dans *INT2*. Cette opération est identique à `INT1 && (~~INT2)`. Elle s'avère utile pour l'élimination des bits d'*INT1* définis dans *INT2*.
`INT << N`	*Entier*	Produit le motif binaire de l'entier *INT1*, décalé vers la gauche de *N* positions. Utilisez une valeur *N* négative pour produire un décalage vers la droite.
`INT >> N`	*Entier*	Produit le motif binaire de l'entier *INT1*, décalé vers la droite de *N* positions. Utilisez une valeur *N* négative pour produire un décalage vers la gauche.
`INT1 &&=_0 INT2`	*Booléen*	Equivaut à l'expression booléenne `INT1 && INT2 /==0`, mais est plus efficace.
`INT1 &&/=_0 INT2`	*Booléen*	Equivaut à l'expression booléenne `INT1 && INT2 == 0` mais est plus efficace.
`integer_bitcount(INT)`	*Entier*	Compte le nombre de bits 1 ou 0 dans la représentation de complément à 2 de l'entier *INT*. Si *INT* n'est pas négatif, *N* représente le nombre de bits 1. Si *INT* est négatif, N correspond au nombre de bits 0. En raison de l'extension de signe, il existe un nombre infini de bits 0 dans un entier positif ou de bits 1 dans un entier négatif. L'expression `integer_bitcount(INT) = integer_bitcount(-(INT+1))` est toujours vraie.
`integer_leastbit(INT)`	*Entier*	Renvoie la position binaire *N* du bit le moins significatif défini dans l'entier *INT*. *N* représente la puissance de 2 la plus élevée par laquelle il est possible de diviser *INT* sans reste.
`integer_length(INT)`	*Entier*	Renvoie la longueur en bits de l'entier *INT*, sous forme d'entier de complément à 2. Cela signifie que *N* est le plus petit entier pour lequel l'expression `INT < (1 << N) if INT >= 0 INT >= (–1 << N) if INT < 0` est vérifiée. Si *INT* n'est pas négatif, la représentation d'*INT* sous forme d'entier non signé nécessite un champ d'au moins *N* bits. Sinon, un minimum de *N+1* bits est nécessaire pour représenter *INT* en tant qu'entier signé, quel que soit son signe.
`testbit(INT, N)`	*Booléen*	Teste le bit de position *N* dans l'entier *INT* et renvoie l'état du bit *N* sous forme de valeur booléenne, vraie (True) s'il s'agit de 1 et fausse (False) pour 0.

Fonctions aléatoires

Les fonctions suivantes sont utilisées pour sélectionner des éléments ou générer des nombres.

Tableau 30. fonctions aléatoires CLEM

Fonction	Résultat	Description
`oneof(LIST)`	*Tout*	Renvoie un élément choisi au hasard dans *LIST*. Vous devez entrer les éléments de liste au format `[ITEM1,ITEM2,...,ITEM_N]`. Il est également possible d'indiquer une liste de noms de champ.
`random(NUM)`	*Nombre*	Renvoie un nombre aléatoire distribué uniformément et du même type (*INT* ou *REAL*), de 1 à *NUM*. Si vous utilisez un entier, seuls des entiers sont renvoyés. Si vous utilisez un nombre réel (décimal), des nombres réels sont renvoyés (précision décimale déterminée par les options de flux). Le plus grand nombre aléatoire qui peut être renvoyé par la fonction est égal à *NUM*.
`random0(NUM)`	*Nombre*	Possède les mêmes propriétés que `random(NUM)`, mais commence à 0. Le plus grand nombre aléatoire renvoyé par la fonction sera toujours différent de *NUM*.

Fonctions sur chaînes

Dans le langage CLEM, vous pouvez effectuer les opérations suivantes avec des chaînes :

- Comparer des chaînes
- Créer des chaînes
- Accéder à des caractères

Dans une expression CLEM, une chaîne est une suite de caractères insérée entre guillemets doubles (`"chaînes"`). Tout caractère alphanumérique unique peut être un caractère (`CHAR`). Ces caractères sont déclarés dans les expressions CLEM à l'aide de guillemets simples inversés : `` `<caractère>` ``, (par exemple, `` `z` ``, `` `A` ``, ou `` `2` ``). Les caractères hors limites ou les indices négatifs d'une chaîne génèrent un comportement indéfini.

Remarque : La comparaison entre deux chaînes dont l'une utilise la conversion SQL et l'autre non peut générer des résultats différents lorsqu'il existe des espace de fin de chaîne.

Tableau 31. Fonctions sur chaînes CLEM

Fonction	Résultat	Description
`allbutfirst(N, STRING)`	*Chaîne*	Renvoie la chaîne *STRING* où les *N* premiers caractères ont été supprimés.
`allbutlast(N, STRING)`	*Chaîne*	Renvoie la chaîne *STRING* où les derniers caractères ont été supprimés.
`alphabefore(STRING1, STRING2)`	*Booléen*	Utilisée pour vérifier l'ordre alphabétique des chaînes. Renvoie une valeur vraie (True) si *STRING1* précède *STRING2*.
`endstring(LENGTH, STRING)`	*Chaîne*	Extrait les *N* derniers caractères de la chaîne indiquée. Si la longueur de la chaîne est inférieure ou égale à la valeur spécifiée, aucune modification n'a lieu.
`hasendstring(STRING, SUBSTRING)`	*Entier*	Cette fonction est identique à `isendstring(SUBSTRING, STRING)`.
`hasmidstring(STRING, SUBSTRING)`	*Entier*	Cette fonction est identique à `ismidstring(SUBSTRING, STRING)` (sous-chaîne incorporée).

Tableau 31. Fonctions sur chaînes CLEM (suite)

Fonction	Résultat	Description
`hasstartstring(STRING, SUBSTRING)`	*Entier*	Cette fonction est identique à `isstartstring(SUBSTRING, STRING)`.
`hassubstring(STRING, N, SUBSTRING)`	*Entier*	Cette fonction est identique à `issubstring(SUB_STRING, N, STRING)`, où *N* prend par défaut la valeur 1.
`count_substring(STRING, SUBSTRING)`	*Entier*	Renvoie le nombre de fois où la sous-chaîne indiquée apparaît dans la chaîne. Par exemple, `count_substring("foooo.txt", "oo")` renvoie 3.
`hassubstring(STRING, SUBSTRING)`	*Entier*	Cette fonction est identique à `issubstring(SUBSTRING, 1, STRING)` où *N* prend par défaut la valeur 1.
`isalphacode(CHAR)`	*Booléen*	Renvoie la valeur true (vrai) si *CHAR* est un caractère dont le code de caractère est une lettre dans la chaîne indiquée (souvent un nom de champ). Ou alors, cette fonction renvoie une valeur 0. Par exemple, `isalphacode(produce_num(1))`.
`isendstring(SUBSTRING, STRING)`	*Entier*	Si la chaîne *STRING* se termine par la sous-chaîne *SUBSTRING*, cette fonction renvoie l'indice d'entier de *SUBSTRING* dans *STRING*. Sinon, cette fonction renvoie la valeur 0.
`islowercode(CHAR)`	*Booléen*	Renvoie la valeur true (vrai) si *CHAR* est un caractère en minuscule dans la chaîne indiquée (souvent un nom de champ). Sinon, la fonction renvoie la valeur 0. Par exemple, les expressions `islowercode(``)` et `islowercode(country_name(2))` sont valides.
`ismidstring(SUBSTRING, STRING)`	*Entier*	Si *SUBSTRING* est une sous-chaîne de la chaîne *STRING*, mais qu'elle ne commence pas par le premier caractère de *STRING* et ne se termine pas par le dernier, cette fonction renvoie l'indice auquel la sous-chaîne commence. Sinon, cette fonction renvoie la valeur 0.
`isnumbercode(CHAR)`	*Booléen*	Renvoie la valeur true (vrai) si *CHAR* est un caractère dont le code de caractère est un chiffre dans la chaîne indiquée (souvent un nom de champ). Sinon, cette fonction renvoie la valeur 0 (par exemple, `isnumbercode(product_id(2))`).
`isstartstring(SUBSTRING, STRING)`	*Entier*	Si la chaîne *STRING* commence par la sous-chaîne *SUBSTRING*, cette fonction renvoie l'indice 1. Sinon, elle renvoie une valeur 0.
`issubstring(SUBSTRING, N, STRING)`	*Entier*	Recherche dans la chaîne *STRING*, à partir du caractère numéro *N*, la sous-chaîne *SUBSTRING*. Si cette sous-chaîne est trouvée, la fonction renvoie l'indice d'entier auquel la sous-chaîne correspondante commence. Sinon, cette fonction renvoie la valeur 0. Si la valeur de *N* n'est pas fournie, cette fonction prend par défaut la valeur 1.

Tableau 31. Fonctions sur chaînes CLEM (suite)

Fonction	Résultat	Description
`issubstring(SUBSTRING, STRING)`	*Entier*	Recherche dans la chaîne *STRING*, à partir du caractère numéro *N*, la sous-chaîne *SUBSTRING*. Si cette sous-chaîne est trouvée, la fonction renvoie l'indice d'entier auquel la sous-chaîne correspondante commence. Sinon, cette fonction renvoie la valeur 0. Si la valeur de *N* n'est pas fournie, cette fonction prend par défaut la valeur 1.
`issubstring_count(SUBSTRING, N, STRING):`	*Entier*	Renvoie l'index de la *N*ième occurrence de la sous-chaîne *SUBSTRING* au sein de la chaîne *STRING* indiquée. Si le nombre d'occurrences de la sous-chaîne *SUBSTRING* est inférieur à *N* occurrences, 0 est renvoyé.
`issubstring_lim(SUBSTRING, N, STARTLIM, ENDLIM, STRING)`	*Entier*	Cette fonction est identique à `issubstring`, mais la correspondance doit commencer à l'indice *STARTLIM* ou avant, et doit se terminer à l'indice *ENDLIM* ou avant. Les contraintes *STARTLIM* ou *ENDLIM* peuvent être désactivées en donnant une valeur fausse à l'un des arguments ; par exemple, `issubstring_lim(SUBSTRING, N, false, false, STRING)` est égale à `issubstring`.
`isuppercode(CHAR)`	*Booléen*	Renvoie une valeur true (vrai) si *CHAR* est un caractère majuscule. Sinon, la fonction renvoie la valeur 0. Par exemple, les expressions `isuppercode(``)` et `isuppercode(country_name(2))` sont valides.
`last(CHAR)`	*Chaîne*	Renvoie le dernier caractère *CHAR* de la chaîne *STRING* (qui doit contenir au moins un caractère).
`length(STRING)`	*Entier*	Renvoie la longueur de la chaîne *STRING*, en nombre de caractères.
`locchar(CHAR, N, STRING)`	*Entier*	Sert à identifier l'emplacement de caractères dans des champs symboliques. Cette fonction recherche le caractère *CHAR* dans la chaîne *STRING*, en commençant par le caractère numéro *N* de *STRING*. Cette fonction renvoie une valeur indiquant l'emplacement (à partir de *N*) où se trouve le caractère. Si le caractère est introuvable, la fonction renvoie la valeur 0. Si la fonction a un décalage *(N)* incorrect (par exemple, décalage dépassant la longueur de la chaîne), elle renvoie `$null$`. Par exemple, `locchar(`n`, 2, web_page)` recherche dans le champ*web_page* le caractère `` `n` `` en commençant par le deuxième caractère de la valeur du champ. *Remarque* : veillez à encadrer le caractère spécifié de guillemets simples inversés.

Tableau 31. Fonctions sur chaînes CLEM (suite)

Fonction	Résultat	Description
`locchar_back(CHAR, N, STRING)`	*Entier*	Semblable à `locchar`, cette fonction effectue la recherche vers l'arrière, en commençant par le caractère numéro *N*. Par exemple, `locchar_back(`n`, 9, web_page)` effectue une recherche dans le champ *web_page* en commençant par le neuvième caractère, puis en repartant vers le début de la chaîne. Si cette fonction a un décalage incorrect (par exemple, décalage dépassant la longueur de la chaîne), elle renvoie `$null$`. Pour plus d'efficacité, employez `locchar_back` en association avec la fonction `length(<field>)` afin d'utiliser la longueur de la valeur en cours du champ de manière dynamique (par exemple, `locchar_back(`n`, (length(web_page)), web_page)`).
`lowertoupper(CHAR)` `lowertoupper (STRING)`	*CHAR* ou *Chaîne*	L'entrée peut être soit une chaîne soit un caractère, qui est utilisé dans cette fonction pour renvoyer un nouvel élément de même type, avec les caractères minuscules convertis dans leurs équivalents majuscules. Par exemple, les expressions `lowertoupper(`a`)`, `lowertoupper("Ma chaîne")` et `lowertoupper(field_name(2))` sont valides.
`matches`	*Booléen*	Renvoie la valeur True (vrai) si une chaîne correspond à un motif spécifié. Le motif doit être un littéral de chaîne ; il ne peut pas s'agir d'un nom de champ contenant un motif. Il est possible d'inclure un point d'interrogation (?) dans le motif pour obtenir une correspondance exacte avec un caractère ; un astérisque (*) implique une correspondance avec zéro ou plusieurs caractères. Pour obtenir une correspondance avec un point d'interrogation ou un astérisque de littéral (plutôt que de les utiliser comme caractères génériques), il est possible d'utiliser une barre oblique inverse (\) en tant que caractère d'échappement.
`replace(SUBSTRING, NEWSUBSTRING, STRING)`	*Chaîne*	Dans la chaîne *STRING* indiquée, remplace toutes les instances de la sous-chaîne *SUBSTRING* par la nouvelle sous-chaîne *NEWSUBSTRING*.
`replicate(COUNT, STRING)`	*Chaîne*	Renvoie une chaîne qui se compose de la chaîne d'origine copiée le nombre de fois indiqué.
`stripchar(CHAR,STRING)`	*Chaîne*	Cette fonction permet de supprimer des caractères d'une chaîne ou d'un champ. Vous pouvez l'utiliser, par exemple, pour supprimer de vos données des symboles de devises afin d'obtenir un simple nombre ou un nom. Par exemple, la syntaxe `stripchar(`$`, 'Cost')` renvoie un nouveau champ où toutes les valeurs apparaissent sans le symbole dollar. *Remarque* : veillez à encadrer le caractère spécifié de guillemets simples inversés.

Tableau 31. Fonctions sur chaînes CLEM (suite)

Fonction	Résultat	Description
`skipchar(CHAR, N, STRING)`	*Entier*	Recherche tout autre caractère que *CHAR* dans la chaîne *STRING*, en commençant par le caractère numéro *N*. Cette fonction renvoie un indice d'entier indiquant l'endroit auquel le caractère a été trouvé. Si tous les caractères qui suivent le caractère numéro *N* sont des caractères *CHAR*, la fonction renvoie 0. Si cette fonction a un décalage incorrect (par exemple, décalage dépassant la longueur de la chaîne), elle renvoie `$null$`. `locchar` est souvent utilisé en association avec les fonctions `skipchar` afin de déterminer la valeur de *N* (point de la chaîne à partir duquel commence la recherche) : par exemple, `skipchar(`s`, (locchar(`s`, 1, "MyString")), "MyString"))`.
`skipchar_back(CHAR, N, STRING)`	*Entier*	Semblable à `skipchar`, cette fonction effectue la recherche **vers l'arrière**, en commençant par le caractère numéro *N*.
`startstring(LENGTH, STRING)`	*Chaîne*	Extrait les *N* premiers caractères de la chaîne indiquée. Si la longueur de la chaîne est inférieure ou égale à la valeur spécifiée, aucune modification n'a lieu.
`strmember(CHAR, STRING)`	*Entier*	Equivaut à `locchar(CHAR, 1, STRING)`. Renvoie un indice entier qui indique l'endroit où *CHAR* s'est produit pour la première fois, ou 0. Si la fonction a un décalage non valide (par exemple, un décalage au-delà de la longueur de la chaîne), cette fonction renvoie `$null$`.
`subscrs(N, STRING)`	*CHAR*	Renvoie le caractère (*CHAR*) numéro *N* de la chaîne de saisie *STRING*. Cette fonction peut également être écrite en sténographie, `STRING(N)`. Par exemple, l'expression `lowertoupper("name"(1))` est correcte.
`substring(N, LEN, STRING)`	*Chaîne*	Renvoie la sous-chaîne *SUBSTRING*, constituée de *LEN* caractères de la chaîne *STRING*, à partir du caractère situé à l'indice *N*.
`substring_between(N1, N2, STRING)`	*Chaîne*	Renvoie la sous-chaîne de la chaîne *STRING* qui commence à l'indice *N1* et se termine à l'indice *N2*.
`trim(STRING)`	*Chaîne*	Supprime les espaces blancs situés au début et à la fin de la chaîne indiquée.
`trim_start(STRING)`	*Chaîne*	Supprime les espaces blancs situés au début de la chaîne indiquée.
`trimend(STRING)`	*Chaîne*	Supprime les espaces blancs situés à la fin de la chaîne indiquée.
`unicode_char(NUM)`	*CHAR*	L'entrée doit contenir des valeurs décimales et non hexadécimales. Renvoie le caractère avec la valeur Unicode *NUM*.
`unicode_value(CHAR)`	*NUM*	Renvoie la valeur Unicode de *CHAR*

Tableau 31. Fonctions sur chaînes CLEM (suite)

Fonction	Résultat	Description
`uppertolower(CHAR)` `uppertolower (STRING)`	*CHAR* ou *Chaîne*	L'entrée peut être soit une chaîne soit un caractère, qui est utilisé dans cette fonction pour renvoyer un nouvel élément de même type avec les caractères en majuscule convertis dans leurs équivalents en minuscule. *Remarque* : n'oubliez pas d'encadrer les chaînes spécifiées de guillemets doubles et les caractères de guillemets simples inversés. Les noms de champ simples ne doivent pas contenir de guillemets.

Fonctions SoundEx

SoundEx est une méthode qui permet de trouver des chaînes lorsque le son est connu mais pas l'orthographe exacte. Développée en 1918, elle recherche des mots présentant un son similaire, sur la base d'hypothèses phonétiques (modalités de prononciation de certaines lettres). Il est possible, par exemple, de l'utiliser pour rechercher des noms dans une base de données où l'orthographe et la prononciation de noms similaires peuvent varier. L'algorithme SoundEx de base est décrit dans diverses sources et, malgré un certain nombre de limitations connues (par exemple, les combinaisons de lettres de début telles que `ph` et `f` n'entraînent pas de correspondance même si le son est identique), il est pris en charge d'une certaine manière par la plupart des bases de données.

Tableau 32. Fonctions soundex CLEM

Fonction	Résultat	Description
`soundex(STRING)`	*Entier*	Renvoie le code SoundEx à quatre caractères pour la chaîne *STRING* indiquée.
`soundex_difference(STRING1, STRING2)`	*Entier*	Renvoie un entier entre 0 et 4 qui indique le nombre de caractères identiques dans le codage SoundEx pour les deux chaînes (0 = absence de similarité, 4 = similarité élevée ou chaînes identiques).

Fonctions date et heure

CLEM comprend une série de fonctions permettant de manipuler des champs avec un stockage datetime contenant des variables de chaîne représentant des dates et des heures. Les formats de date et d'heure utilisés sont propres à chaque flux et sont définis dans la boîte de dialogue Propriétés du flux. Les fonctions de date et d'heure analysent les chaînes de date et d'heure en fonction du format sélectionné.

Lorsque vous indiquez dans une date une année sur deux chiffres seulement (c'est-à-dire sans préciser le siècle), IBM SPSS Modeler utilise le siècle par défaut défini dans la boîte de dialogue des propriétés du flux.

Remarque : Si la fonction de données est retransmise à SQL ou IBM SPSS Analytic Server, dans une branche suivant une source de données Analytic Server, les chaînes de format de date (to_date) dans ces données doivent avoir le format de date indiqué dans le flux SPSS Modeler.

Tableau 33. Fonctions CLEM de date et d'heure.

Fonction	Résultat	Description
`@TODAY`	*Chaîne*	Si vous sélectionnez **Passer jours/minutes** dans la boîte de dialogue des propriétés du flux, cette fonction renvoie la date actuelle en tant que chaîne sous le format de date actuel. Si vous utilisez un format de date sur deux chiffres sans sélectionner l'option **Passer jours/minutes**, la fonction renvoie la valeur `$null$` sur le serveur en cours.
`to_time(ITEM)`	*Heure*	Convertit le stockage du champ spécifié en une heure.
`to_date(ITEM)`	*Date*	Convertit le stockage du champ spécifié en une date.
`to_timestamp(ITEM)`	*Horodatage*	Convertit le stockage du champ spécifié en un horodatage.
`to_datetime(ITEM)`	*Datetime*	Convertit le stockage du champ spécifié en une valeur d'heure, de date ou d'horodatage.
`datetime_date(ITEM)`	*Date*	Renvoie la valeur de date pour un *nombre*, une *chaîne*, ou un *horodatage*. Veuillez noter que cette fonction est la seule fonction vous permettant de reconvertir un nombre (en secondes) en date. Si `ITEM` est une chaîne, crée une date en analysant une chaîne au format de date sélectionné. Pour que cette fonction réussisse, vous devez spécifier un format de date correct dans la boîte de dialogue des propriétés du flux. Si `ITEM` est un nombre, il est interprété comme un nombre de secondes écoulées depuis la date de départ (ou période). Les fractions de jour sont supprimées. Si `ITEM` est un horodatage, la partie date de cet horodatage est renvoyée. Si `ITEM` est une date, il est renvoyé sans modification.
`date_before(DATE1, DATE2)`	*Booléen*	Renvoie une valeur true (vrai) si *DATE1* représente une date ou un horodatage antérieur à *DATE2*. Sinon, cette fonction renvoie la valeur 0.
`date_days_difference(DATE1, DATE2)`	*Entier*	Renvoie la durée en jours depuis la date ou l'horodatage qui sépare *DATE1* de *DATE2*, sous forme d'entier. Si *DATE2* est antérieure à *DATE1*, cette fonction renvoie un nombre négatif.
`date_in_days(DATE)`	*Entier*	Renvoie la durée en jours qui sépare la date ou l'horodatage de référence de l'élément *DATE* indiqué, sous la forme d'un entier. Si *DATE* est antérieure à la date de référence, cette fonction renvoie un nombre. Vous devez inclure une date réelle pour que le calcul fonctionne correctement. Par exemple, vous ne devez pas entrer la date : 29 février 2001. En effet, 2001 n'est pas une année bissextile et cette date n'existe pas.
`date_in_months(DATE)`	*Réel*	Renvoie la durée en mois qui sépare la date ou l'horodatage de référence de l'élément *DATE* indiqué, sous la forme d'un nombre réel. Ce chiffre est approximatif et utilise une base de 30.4375 jours par mois. Si *DATE* est antérieure à la date de référence, cette fonction renvoie un nombre. Vous devez inclure une date réelle pour que le calcul fonctionne correctement. Par exemple, vous ne devez pas entrer la date : 29 février 2001. En effet, 2001 n'est pas une année bissextile et cette date n'existe pas.

Tableau 33. Fonctions CLEM de date et d'heure (suite).

Fonction	Résultat	Description
`date_in_weeks(DATE)`	*Réel*	Renvoie la durée en semaines qui sépare la date ou l'horodatage de référence de l'élément *DATE* indiqué, sous la forme d'un nombre réel. La semaine de base utilisée comprend 7,0 jours. Si *DATE* est antérieure à la date de référence, cette fonction renvoie un nombre. Vous devez inclure une date réelle pour que le calcul fonctionne correctement. Par exemple, vous ne devez pas entrer la date : 29 février 2001. En effet, 2001 n'est pas une année bissextile et cette date n'existe pas.
`date_in_years(DATE)`	*Réel*	Renvoie la durée en années qui sépare la date ou l'horodatage de référence de l'élément *DATE* indiqué, sous la forme d'un nombre réel. Ce chiffre est approximatif et utilise une base de 365.25 jours. Si *DATE* est antérieure à la date de référence, cette fonction renvoie un nombre. Vous devez inclure une date réelle pour que le calcul fonctionne correctement. Par exemple, vous ne devez pas entrer la date : 29 février 2001. En effet, 2001 n'est pas une année bissextile et cette date n'existe pas.
`date_months_difference (DATE1, DATE2)`	*Réel*	Renvoie la durée en mois depuis la date ou l'horodatage qui sépare *DATE1* de *DATE2*, sous forme d'un nombre réel. Ce chiffre est approximatif et utilise une base de 30.4375 jours par mois. Si *DATE2* est antérieure à *DATE1*, cette fonction renvoie un nombre négatif.
`datetime_date(YEAR, MONTH, DAY)`	*Date*	Renvoie la valeur de date pour les éléments *YEAR*, *MONTH* et *DAY* (année, mois, jour) indiqués. Ces arguments doivent être des entiers.
`datetime_day(DATE)`	*Entier*	Renvoie le jour du mois à partir de la valeur *DATE* ou d'un horodatage. Le résultat est un entier compris entre 1 et 31.
`datetime_day_name(DAY)`	*Chaîne*	Renvoie le nom complet de l'élément *DAY* (jour) indiqué. L'argument doit être un entier compris entre 1 (dimanche) et 7 (samedi).
`datetime_hour(TIME)`	*Entier*	Renvoie l'heure à partir de la valeur *TIME* ou d'un horodatage. Le résultat est un entier compris entre 0 et 23.
`datetime_in_seconds(TIME)`	*Réel*	Renvoie la partie des secondes stockée dans *TIME*.
`datetime_in_seconds(DATE)`, `datetime_in_seconds(DATETIME)`	*Réel*	Renvoie le chiffre cumulé, converti en secondes, de la différence entre la date *DATE* ou *DATETIME* actuelle et la date de référence (1900-01-01).
`datetime_minute(TIME)`	*Entier*	Renvoie le nombre de minutes, à partir de *TIME* ou d'un horodatage. Le résultat est un entier compris entre 0 et 59.
`datetime_month(DATE)`	*Entier*	Renvoie le mois, à partir de *DATE* ou d'un horodatage. Le résultat est un entier compris entre 1 et 12.
`datetime_month_name (MONTH)`	*Chaîne*	Renvoie le nom complet de l'élément *MONTH* (mois) indiqué. L'argument doit être un entier compris entre 1 et 12.
`datetime_now`	*Horodatage*	Renvoie l'heure actuelle sous la forme d'un horodatage.
`datetime_second(TIME)`	*Entier*	Renvoie le nombre de secondes, à partir de *TIME* ou d'un horodatage. Le résultat est un entier compris entre 0 et 59.
`datetime_day_short_name (DAY)`	*Chaîne*	Renvoie le nom abrégé de l'élément *DAY* (jour) indiqué. L'argument doit être un entier compris entre 1 (dimanche) et 7 (samedi).

Tableau 33. Fonctions CLEM de date et d'heure (suite).

Fonction	Résultat	Description
`datetime_month_short_name (MONTH)`	*Chaîne*	Renvoie le nom abrégé de l'élément *MONTH* (mois) indiqué. L'argument doit être un entier compris entre 1 et 12.
`datetime_time(HOUR, MINUTE, SECOND)`	*Heure*	Renvoie la valeur d'heure pour les éléments *HOUR*, *MINUTE* et *SECOND* (heures, minutes, secondes) indiqués. Ces arguments doivent être des entiers.
`datetime_time(ITEM)`	*Heure*	Renvoie la valeur d'heure de l'élément *ITEM* indiqué.
`datetime_timestamp(YEAR, MONTH, DAY, HOUR, MINUTE, SECOND)`	*Horodatage*	Renvoie la valeur d'horodatage pour les éléments *YEAR*, *MONTH*, *DAY*, *HOUR*, *MINUTE* et *SECOND* (année, mois, jour, heures, minutes, secondes) indiqués.
`datetime_timestamp(DATE, TIME)`	*Horodatage*	Renvoie la valeur d'horodatage des éléments *DATE* et *TIME* (date, heure) indiqués.
`datetime_timestamp (NUMBER)`	*Horodatage*	Renvoie la valeur d'horodatage du nombre de secondes indiqué.
`datetime_weekday(DATE)`	*Entier*	Renvoie le jour de la semaine à partir de la valeur *DATE* ou d'un horodatage.
`datetime_year(DATE)`	*Entier*	Renvoie l'année, à partir de *DATE* ou d'un horodatage. Le résultat est un entier (comme 2002).
`date_weeks_difference (DATE1, DATE2)`	*Réel*	Renvoie la durée en semaines depuis la date ou l'horodatage qui sépare *DATE1* de *DATE2*, sous forme d'un nombre réel. La semaine de base utilisée comprend 7,0 jours. Si *DATE2* est antérieure à *DATE1*, cette fonction renvoie un nombre négatif.
`date_years_difference (DATE1, DATE2)`	*Réel*	Renvoie la durée en années depuis la date ou l'horodatage qui sépare *DATE1* de *DATE2*, sous forme d'un nombre réel. Ce chiffre est approximatif et utilise une base de 365.25 jours. Si *DATE2* est antérieure à *DATE1*, cette fonction renvoie un nombre négatif.
`date_from_ywd(YEAR, WEEK, DAY)`	*Entier*	Convertit l'année, la semaine dans l'année et le jour de la semaine, en une date au format ISO 8601.
`date_iso_day(DATE)`	*Entier*	Renvoie le jour de la semaine de la date au format ISO 8601.
`date_iso_week(DATE)`	*Entier*	Renvoie la semaine de l'année de la date au format ISO 8601.
`date_iso_year(DATE)`	*Entier*	Renvoie l'année de la date au format ISO 8601.
`time_before(TIME1, TIME2)`	*Booléen*	Renvoie une valeur true (vrai) si *TIME1* représente une heure ou un horodatage antérieur à *TIME2*. Sinon, cette fonction renvoie la valeur 0.
`time_hours_difference (TIME1, TIME2)`	*Réel*	Renvoie la durée en heures qui sépare les heures ou les horodatages représentés par *TIME1* et *TIME2*, sous la forme d'un nombre réel. Si vous sélectionnez l'option **Passer jours/minutes** dans la boîte de dialogue des propriétés du flux, une valeur plus élevée de *TIME1* fait référence au jour précédent. Si vous ne sélectionnez pas cette option, une valeur plus élevée de *TIME1* entraîne une valeur négative.
`time_in_hours(TIME)`	*Réel*	Renvoie le nombre d'heures représenté par l'heure *TIME*, sous la forme d'un nombre réel. Par exemple, pour le format d'heure `HHMM`, l'expression `time_in_hours('0130')` renvoie la valeur 1,5. *TIME* peut représenter une heure ou un horodatage.

Tableau 33. Fonctions CLEM de date et d'heure (suite).

Fonction	Résultat	Description
`time_in_mins(TIME)`	*Réel*	Renvoie le nombre de minutes représenté par l'heure *TIME*, sous la forme d'un nombre réel. *TIME* peut représenter une heure ou un horodatage.
`time_in_secs(TIME)`	*Entier*	Renvoie le nombre de secondes représenté par l'heure *TIME*, sous la forme d'un entier. *TIME* peut représenter une heure ou un horodatage.
`time_mins_difference(TIME1, TIME2)`	*Réel*	Renvoie la durée en minutes qui sépare les heures ou les horodatages représentés par *TIME1* et *TIME2*, sous la forme d'un nombre réel. Si vous sélectionnez l'option **Passer jours/minutes** dans la boîte de dialogue des propriétés du flux, une valeur plus élevée de *TIME1* est utilisée pour faire référence au jour précédent (ou à l'heure précédente, si les minutes et les secondes sont uniquement spécifiées dans le format actuel). Si vous ne sélectionnez pas cette option, une valeur plus élevée de *TIME1* entraîne une valeur négative.
`time_secs_difference(TIME1, TIME2)`	*Entier*	Renvoie la durée en secondes qui sépare l'heure ou l'horodatage représenté par *TIME1* de celle représentée par *TIME2*, sous la forme d'un entier. Si vous sélectionnez l'option **Passer jours/minutes** dans la boîte de dialogue des propriétés du flux, une valeur plus élevée de *TIME1* est utilisée pour faire référence au jour précédent (ou à l'heure précédente, si les minutes et les secondes sont uniquement spécifiées dans le format actuel). Si vous ne sélectionnez pas cette option, une valeur plus élevée de *TIME1* entraîne une valeur négative.

Conversion des valeurs date et heure.

Notez que les fonctions de conversion (et toutes les autres fonctions qui nécessitent un type spécifique d'entrée, par exemple une valeur de date ou d'heure) dépendent des formats actuels indiqués dans la boîte de dialogue des options de flux. Par exemple, si vous avez un champ *DATE* de type chaîne avec les valeurs *Jan 2003*, *Fev 2003*, etc., vous pouvez le convertir en stockage de date sous le format suivant :

```
to_date(DATE)
```

Pour que cette conversion fonctionne, sélectionnez le format de date correspondant **MON YYYY** comme format de date par défaut pour le flux.

Pour un exemple de conversion de valeurs de chaînes en dates à l'aide d'un noeud Remplacer, consultez le flux *broadband_create_models.str*, installé dans le dossier *\Demos* du sous-dossier *streams*.

Dates stockées sous forme de nombres. Veuillez noter que *DATE* dans l'exemple précédent est le nom d'un champ, alors que `to_date` est une fonction CLEM. Si vous avez des dates stockées sous la forme de nombres, vous pouvez les convertir en utilisant la fonction `datetime_date` où le nombre est interprété comme un nombre de secondes écoulées depuis la date de départ (ou période).

```
datetime_date(DATE)
```

En convertissant une date en un nombre de secondes (et vice-versa), vous pouvez effectuer des calculs tels que le calcul de la date actuelle plus ou moins un certain nombre de jours, par exemple :

```
datetime_date((date_in_days(DATE)-7)*60*60*24)
```

Fonctions séquentielles

Dans le cadre de certaines opérations, la séquence des événements est importante. Cette application vous permet de travailler avec les séquences d'enregistrements suivantes :

- Séquences et séries temporelles
- Fonctions séquentielles
- Indexation des enregistrements
- Moyenne, somme et valeurs comparatives
- Contrôle des changements - Différenciation
- `@SINCE`
- Valeurs de décalage
- Fonctions séquentielles supplémentaires

Pour nombre d'applications, chaque enregistrement circulant dans un flux peut être considéré comme un cas particulier, indépendant de tous les autres. Auxquels cas, l'ordre des enregistrements importe généralement peu.

Toutefois, pour certaines catégories de problèmes, la séquence des enregistrements est importante. Il s'agit en général de situations de séries temporelles dans lesquelles la séquence des enregistrements représente une séquence ordonnée d'événements ou d'occurrences. Chaque enregistrement représente un instantané d'un moment particulier. Cependant, une grande partie des informations les plus complexes peut être contenue non pas dans des valeurs instantanées, mais dans la façon dont de telles valeurs évoluent et se comportent dans le temps.

Bien entendu, le paramètre pertinent peut être autre que le temps. Par exemple, les enregistrements peuvent représenter des analyses effectuées à distance sur une ligne, mais les mêmes principes s'appliquent.

Les fonctions séquentielles et les fonctions spéciales sont reconnaissables aux caractéristiques suivantes.

- Elles sont toutes précédées du préfixe `@`.
- Leur nom apparaît en majuscules.

Les fonctions séquentielles font référence à l'enregistrement en cours traité par un noeud, aux enregistrements qui sont déjà passés dans un noeud et même, dans un cas, aux enregistrements qui ne sont pas encore passés dans un noeud. Les fonctions séquentielles peuvent être librement mélangées avec d'autres composants d'expressions CLEM, bien que certains aient des restrictions quant à leurs arguments.

Exemples

Vous pouvez avoir besoin de connaître le temps écoulé depuis un certain événement ou une condition vraie. Pour cela, utilisez la fonction `@SINCE`. Par exemple :

```
@SINCE(Income > Outgoings)
```

Cette fonction renvoie le décalage du dernier enregistrement où cette condition était vraie, c'est-à-dire le nombre d'enregistrements précédant celui où la condition était vraie. Si la condition n'a jamais été vraie, `@SINCE` renvoie `@INDEX + 1`.

Vous pouvez également devoir faire référence à une valeur de l'enregistrement en cours dans l'expression utilisée par `@SINCE`. Pour cela, utilisez la fonction `@THIS`, qui spécifie qu'un nom de champ s'applique toujours à l'enregistrement en cours. Pour déterminer le décalage du dernier enregistrement dont la valeur de champ `Concentration` est plus de deux fois supérieure à celle de l'enregistrement en cours, vous pouvez utiliser :

```
@SINCE(Concentration > 2 * @THIS(Concentration))
```

Dans certains cas, la condition attribuée à `@SINCE` est par définition vraie pour l'enregistrement en cours. Par exemple :

```
@SINCE(ID == @THIS(ID))
```

Pour cette raison, `@SINCE` n'évalue pas la condition de l'enregistrement en cours. Utilisez une fonction identique, `@SINCE0`, pour évaluer la condition de l'enregistrement en cours et des précédents. Si la condition est vraie dans l'enregistrement en cours, `@SINCE0` renvoie `0`.

Tableau 34. Fonctions séquentielles CLEM

Fonction	Résultat	Description
`MEAN(FIELD)`	*Réel*	Renvoie la moyenne des valeurs du ou des champs (*FIELD* ou *FIELDS*) spécifiés.
`@MEAN(FIELD, EXPR)`	*Réel*	Renvoie la moyenne des valeurs du champ *FIELD* correspondant aux *EXPR* (nombre représenté par une expression) derniers enregistrements reçus par le noeud en cours, y compris l'enregistrement en cours. *FIELD* doit être le nom d'un champ numérique. *EXPR* peut être toute expression donnant comme résultat un entier supérieur à 0. Si l'expression *EXPR* est omise ou donne un résultat supérieur au nombre d'enregistrements reçus jusqu'à présent, la moyenne des enregistrements reçus est renvoyée.
`@MEAN(FIELD, EXPR, INT)`	*Réel*	Renvoie la moyenne des valeurs du champ *FIELD* correspondant aux *EXPR* (nombre représenté par une expression) derniers enregistrements reçus par le noeud en cours, y compris l'enregistrement en cours. *FIELD* doit être le nom d'un champ numérique. *EXPR* peut être toute expression donnant comme résultat un entier supérieur à 0. Si l'expression *EXPR* est omise ou donne un résultat supérieur au nombre d'enregistrements reçus jusqu'à présent, la moyenne des enregistrements reçus est renvoyée. *INT* indique le nombre maximal de valeurs à prendre en compte. Cette méthode est beaucoup plus efficace que si vous vous limitez à deux arguments.
`@DIFF1(FIELD)`	*Réel*	Renvoie le premier différentiel du champ *FIELD*. La forme à un seul argument renvoie simplement la différence entre la valeur de champ en cours et la valeur de champ précédente. Renvoie **$null$** si les enregistrements précédents pertinents n'existent pas.
`@DIFF1(FIELD1, FIELD2)`	*Réel*	La forme à deux arguments renvoie le premier différentiel du champ *FIELD1* par rapport au champ *FIELD2*. Renvoie **$null$** si les enregistrements précédents pertinents n'existent pas. Calculé d'après la formule `@DIFF1(FIELD1)/@DIFF1(FIELD2)`.
`@DIFF2(FIELD)`	*Réel*	Renvoie le second différentiel du champ *FIELD*. La forme à un seul argument renvoie simplement la différence entre la valeur de champ en cours et la valeur de champ précédente. Renvoie **$null$** si les enregistrements précédents pertinents n'existent pas. `@DIFF2` est calculé d'après la formule `@DIFF(@DIFF(FIELD))`.
`@DIFF2(FIELD1, FIELD2)`	*Réel*	La forme à deux arguments renvoie le second différentiel de *FIELD1* par rapport à *FIELD2*. Renvoie **$null$** si les enregistrements précédents pertinents n'existent pas. Il s'agit d'un calcul complexe d'après la formule -- `@DIFF1(FIELD1)/@DIFF1(FIELD2) - @OFFSET(@DIFF1(FIELD1),1)/@OFFSET(@DIFF1(FIELD2))) / @DIFF1(FIELD2)`.

Tableau 34. Fonctions séquentielles CLEM (suite)

Fonction	Résultat	Description
`@INDEX`	*Entier*	Renvoie l'index de l'enregistrement en cours. Les index sont affectés aux enregistrements au fil de leur arrivée dans le noeud en cours. Le premier enregistrement reçoit l'index 1 et la valeur d'index augmente de 1 à chacun des enregistrements suivants.
`@LAST_NON_BLANK(FIELD)`	*Tout*	Renvoie la dernière valeur du champ *FIELD* qui était renseigné, tel que défini dans un noeud Typer ou source en amont. S'il n'existe aucune valeur renseignée pour le champ *FIELD* dans les enregistrements lus jusqu'à présent, la fonction renvoie `$null$`. Les valeurs non renseignées, également appelées valeurs manquantes utilisateur, peuvent être définies séparément pour chaque champ.
`@MAX(FIELD)`	*Nombre*	Renvoie la valeur maximale du champ *FIELD* spécifié.
`@MAX(FIELD, EXPR)`	*Nombre*	Renvoie la valeur maximale du champ *FIELD* correspondant aux *EXPR* (nombre représenté par une expression) derniers enregistrements reçus jusqu'à présent, enregistrement en cours compris. *FIELD* doit être le nom d'un champ numérique. *EXPR* peut être toute expression dont le résultat est un entier supérieur à 0.
`@MAX(FIELD, EXPR, INT)`	*Nombre*	Renvoie la valeur maximale du champ *FIELD* correspondant aux *EXPR* (nombre représenté par une expression) derniers enregistrements reçus jusqu'à présent, enregistrement en cours compris. *FIELD* doit être le nom d'un champ numérique. *EXPR* peut être toute expression donnant comme résultat un entier supérieur à 0. Si l'expression *EXPR* est omise ou donne un résultat supérieur au nombre d'enregistrements reçus jusqu'à présent, la valeur minimale des enregistrements reçus est renvoyée. *INT* indique le nombre maximal de valeurs à prendre en compte. Cette méthode est beaucoup plus efficace que si vous vous limitez à deux arguments.
`@MIN(FIELD)`	*Nombre*	Renvoie la valeur minimale du champ *FIELD* spécifié.
`@MIN(FIELD, EXPR)`	*Nombre*	Renvoie la valeur minimale du champ *FIELD* correspondant aux *EXPR* (nombre représenté par une expression) derniers enregistrements reçus jusqu'à présent, enregistrement en cours compris. *FIELD* doit être le nom d'un champ numérique. *EXPR* peut être toute expression dont le résultat est un entier supérieur à 0.
`@MIN(FIELD, EXPR, INT)`	*Nombre*	Renvoie la valeur minimale du champ *FIELD* correspondant aux *EXPR* (nombre représenté par une expression) derniers enregistrements reçus jusqu'à présent, enregistrement en cours compris. *FIELD* doit être le nom d'un champ numérique. *EXPR* peut être toute expression donnant comme résultat un entier supérieur à 0. Si l'expression *EXPR* est omise ou donne un résultat supérieur au nombre d'enregistrements reçus jusqu'à présent, la valeur minimale des enregistrements reçus est renvoyée. *INT* indique le nombre maximal de valeurs à prendre en compte. Cette méthode est beaucoup plus efficace que si vous vous limitez à deux arguments.

Tableau 34. Fonctions séquentielles CLEM (suite)

Fonction	Résultat	Description
`@OFFSET(FIELD, EXPR)`	*Tout*	Renvoie la valeur de *FIELD* pour l'enregistrement décalé de la valeur *EXPR* par rapport à l'enregistrement en cours. Un décalage positif se rapporte à un enregistrement déjà rencontré, tandis qu'un décalage négatif se rapporte à un enregistrement anticipé. Par exemple, `@OFFSET(Status, 1)` renvoie la valeur du champ `Etat` de l'enregistrement précédent, alors que `@OFFSET(Status, –4)` "anticipe" quatre enregistrements dans la séquence (enregistrements qui ne sont pas encore passés par ce noeud) pour obtenir la valeur. *Veuillez noter qu'un décalage négatif (fonction d'anticipation) doit être spécifié comme constante.* Pour les décalages positifs uniquement, *EXPR* peut également être une expression arbitraire CLEM, évaluée pour que l'enregistrement en cours calcule le décalage. Dans ce cas, l'utilisation de la version à trois arguments de cette fonction devrait améliorer les performances (consultez la fonction suivante). Si le résultat de cette expression n'est pas un entier non négatif, une erreur est générée. Il est en effet interdit de calculer des décalages anticipés. *Remarque* : une fonction `@OFFSET` autoréférentielle ne peut pas utiliser une anticipation littérale. Par exemple, dans un noeud Remplacer, vous ne pouvez pas remplacer la valeur de `field1` par une expression comme `@OFFSET(field1,-2)`. *Remarque* : Dans le noeud Remplacer, lors du remplacement d'un champ, existent en fait deux valeurs différentes de ce champ, à savoir la valeur pré-renseignée et la valeur qui l'a remplacée. Lorsque `@OFFSET` fait référence à lui-même, la valeur concernée est celle de remplacement. Cette valeur de remplacement n'existant que pour les lignes passées, une fonction `@OFFSET` se référençant elle-même ne peut se rapporter qu'à des lignes passées. Comme la fonction `@OFFSET` se référençant elle-même ne peut pas se rapporter au futur, elle effectue les vérifications suivantes pour le décalage : • Si le décalage est un littéral, et dans le futur, une erreur est renvoyée avant le début de l'exécution. • Si le décalage est une expression débouchant sur le futur, à l'exécution `@OFFSET` renvoie `$null$`. *Remarque* : L'utilisation à la fois d'un enregistrement anticipé et d'un enregistrement passé dans un noeud n'est pas prise en charge.

Tableau 34. Fonctions séquentielles CLEM (suite)

Fonction	Résultat	Description
`@OFFSET(FIELD, EXPR, INT)`	*Tout*	Effectue la même opération que la fonction `@OFFSET`, avec un troisième argument, *INT*, qui indique le nombre maximal de valeurs à prendre en compte. Si le décalage est calculé à partir d'une expression, ce troisième argument devrait améliorer les performances. Par exemple, dans une formule telle que `@OFFSET(Foo, Month, 12)`, le système sait comment conserver uniquement les douze dernières valeurs de `Foo`. Sinon, il doit stocker chaque valeur à tout hasard. Pour les décalages dont la valeur est constante y compris les décalages anticipés négatifs qui doivent être constants le troisième argument n'est d'aucune utilité. Il est recommandé d'utiliser la version à deux arguments de cette fonction. Consultez également la note sur les fonctions autoréférentielles dans la version à deux arguments décrite plus tôt. *Remarque* : L'utilisation à la fois d'un enregistrement anticipé et d'un enregistrement passé dans un noeud n'est pas prise en charge.
`@SDEV(FIELD)`	*Réel*	Renvoie la somme des valeurs du ou des champs (*FIELD* ou *FIELDS*) spécifiés.
`@SDEV(FIELD, EXPR)`	*Réel*	Renvoie l'écart-type des valeurs du champ *FIELD* pour les *EXPR* derniers enregistrements reçus par le noeud en cours, enregistrement en cours compris. *FIELD* doit être le nom d'un champ numérique. *EXPR* peut être toute expression donnant comme résultat un entier supérieur à 0. Si l'expression *EXPR* est omise ou donne un résultat supérieur au nombre d'enregistrements reçus jusqu'à présent, l'écart-type des enregistrements reçus est renvoyé.
`@SDEV(FIELD, EXPR, INT)`	*Réel*	Renvoie l'écart-type des valeurs du champ *FIELD* pour les *EXPR* derniers enregistrements reçus par le noeud en cours, enregistrement en cours compris. *FIELD* doit être le nom d'un champ numérique. *EXPR* peut être toute expression donnant comme résultat un entier supérieur à 0. Si l'expression *EXPR* est omise ou donne un résultat supérieur au nombre d'enregistrements reçus jusqu'à présent, l'écart-type des enregistrements reçus est renvoyé. *INT* indique le nombre maximal de valeurs à prendre en compte. Cette méthode est beaucoup plus efficace que si vous vous limitez à deux arguments.
`@SINCE(EXPR)`	*Tout*	Renvoie le nombre d'enregistrements qui ont été transmis depuis que l'expression CLEM arbitraire *EXPR* est devenue vraie.
`@SINCE(EXPR, INT)`	*Tout*	Lorsque vous ajoutez le second argument, *INT*, vous indiquez le nombre maximal d'enregistrements à prendre en compte. Si *EXPR* n'a jamais été vraie, *INT* est équivalent à `@INDEX+1`.
`@SINCE0(EXPR)`	*Tout*	Considère l'enregistrement en cours, contrairement à `@SINCE`. La fonction `@SINCE0` renvoie la valeur 0 si l'expression *EXPR* est vraie pour l'enregistrement en cours.
`@SINCE0(EXPR, INT)`	*Tout*	Lorsque vous ajoutez le second argument, *INT* indique le nombre maximal d'enregistrements à prendre en compte.

Tableau 34. Fonctions séquentielles CLEM (suite)

Fonction	Résultat	Description
`@SUM(FIELD)`	*Nombre*	Renvoie la somme des valeurs du ou des champs (*FIELD* ou *FIELDS*) spécifiés.
`@SUM(FIELD, EXPR)`	*Nombre*	Renvoie la somme des valeurs du champ *FIELD* pour les *EXPR* (nombre représenté par une expression) derniers enregistrements reçus par le noeud en cours, enregistrement en cours compris. *FIELD* doit être le nom d'un champ numérique. *EXPR* peut être toute expression donnant comme résultat un entier supérieur à 0. Si l'expression *EXPR* est omise ou donne un résultat supérieur au nombre d'enregistrements reçus jusqu'à présent, la somme des enregistrements reçus est renvoyée.
`@SUM(FIELD, EXPR, INT)`	*Nombre*	Renvoie la somme des valeurs du champ *FIELD* pour les *EXPR* (nombre représenté par une expression) derniers enregistrements reçus par le noeud en cours, enregistrement en cours compris. *FIELD* doit être le nom d'un champ numérique. *EXPR* peut être toute expression donnant comme résultat un entier supérieur à 0. Si l'expression *EXPR* est omise ou donne un résultat supérieur au nombre d'enregistrements reçus jusqu'à présent, la somme des enregistrements reçus est renvoyée. *INT* indique le nombre maximal de valeurs à prendre en compte. Cette méthode est beaucoup plus efficace que si vous vous limitez à deux arguments.
`@THIS(FIELD)`	*Tout*	Renvoie la valeur du champ *FIELD* dans l'enregistrement en cours. Utilisée uniquement dans les expressions `@SINCE`.

Fonctions globales

Les fonctions `@MEAN`, `@SUM`, `@MIN`, `@MAX` et `@SDEV` portent sur tous les enregistrements lus jusqu'à l'enregistrement en cours, ce dernier inclus. Toutefois, dans certains cas, il est utile de comparer les valeurs de l'enregistrement en cours à celles de tout le jeu de données. Lorsque vous utilisez un noeud V. globales (Valeurs globales) pour générer des valeurs dans tout le jeu de données, vous pouvez accéder à ces valeurs dans une expression CLEM en utilisant les fonctions globales.

Par exemple,

```
@GLOBAL_MAX(Age)
```

renvoie la valeur maximale du champ `Age` dans le jeu de données, tandis que l'expression

```
(Value - @GLOBAL_MEAN(Value)) / @GLOBAL_SDEV(Value)
```

exprime la différence entre la valeur du champ `Valeur` dans l'enregistrement en cours et la moyenne globale, sous la forme d'un écart-type. Vous ne pouvez utiliser les valeurs globales qu'après leur calcul par un noeud V. globales (Valeurs globales). Vous pouvez annuler toutes les valeurs globales actuelles en cliquant sur le bouton **Effacer les valeurs globales**, dans l'onglet Valeurs globales de la boîte de dialogue des propriétés du flux.

Tableau 35. Fonctions globales CLEM

Fonction	Résultat	Description
`@GLOBAL_MAX(FIELD)`	*Nombre*	Renvoie la valeur maximale du champ *FIELD* sur tout le jeu de données, tel qu'il a été précédemment généré par un noeud V. globales (Valeurs globales). *FIELD* doit être le nom d'un champ numérique, de date/heure/date et heure, ou chaîne. Une erreur se produit si la valeur globale correspondante n'a pas été définie.
`@GLOBAL_MIN(FIELD)`	*Nombre*	Renvoie la valeur minimale du champ *FIELD* sur tout le jeu de données, tel qu'il a été précédemment généré par un noeud V. globales (Valeurs globales). *FIELD* doit être le nom d'un champ numérique, de date/heure/date et heure, ou chaîne. Une erreur se produit si la valeur globale correspondante n'a pas été définie.
`@GLOBAL_SDEV(FIELD)`	*Nombre*	Renvoie l'écart-type des valeurs du champ *FIELD* sur tout le jeu de données, tel qu'il a été précédemment généré par un noeud V. globales (Valeurs globales). *FIELD* doit être le nom d'un champ numérique. Une erreur se produit si la valeur globale correspondante n'a pas été définie.
`@GLOBAL_MEAN(FIELD)`	*Nombre*	Renvoie la moyenne des valeurs du champ *FIELD* sur tout le jeu de données, tel qu'il a été précédemment généré par un noeud V. globales (Valeurs globales). *FIELD* doit être le nom d'un champ numérique. Une erreur se produit si la valeur globale correspondante n'a pas été définie.
`@GLOBAL_SUM(FIELD)`	*Nombre*	Renvoie la somme des valeurs du champ *FIELD* sur tout le jeu de données, tel qu'il a été précédemment généré par un noeud V. globales (Valeurs globales). *FIELD* doit être le nom d'un champ numérique. Une erreur se produit si la valeur globale correspondante n'a pas été définie.

Traitement des valeurs nulles et non renseignées

Le langage CLEM permet d'indiquer que certaines valeurs d'un champ doivent être considérées comme "blancs", ou manquantes. Vous pouvez utiliser les fonctions suivantes pour travailler avec des valeurs non renseignées.

Tableau 36. Fonctions des valeurs nulles et non renseignées CLEM

Fonction	Résultat	Description
`@BLANK(FIELD)`	*Booléen*	Renvoie la valeur true (vrai) pour tous les enregistrements dont les valeurs sont vides, d'après les règles de traitement des blancs définies dans un noeud Typer ou un noeud source en amont (onglet Types).
`@LAST_NON_BLANK(FIELD)`	*Tout*	Renvoie la dernière valeur du champ *FIELD* qui était renseigné, tel que défini dans un noeud Typer ou source en amont. S'il n'existe aucune valeur renseignée pour le champ *FIELD* dans les enregistrements lus jusqu'à présent, la fonction renvoie `$null$`. Les valeurs non renseignées, également appelées valeurs manquantes utilisateur, peuvent être définies séparément pour chaque champ.

Tableau 36. Fonctions des valeurs nulles et non renseignées CLEM (suite)

Fonction	Résultat	Description
`@NULL(FIELD)`	*Booléen*	Renvoie la valeur true (vrai) si la valeur du champ *FIELD* est la valeur manquante système `$null$`. Renvoie une valeur fausse (False) pour toutes les autres valeurs, notamment les blancs définis par l'utilisateur. Pour rechercher les deux types d'élément, utilisez `@BLANK(FIELD)` et `@NULL(FIELD)`.
`undef`	*Tout*	Cette fonction est généralement utilisée dans CLEM pour entrer une valeur `$null$` (par exemple, pour remplacer des valeurs non renseignées par des valeurs nulles dans le noeud Remplacer).

Les champs non renseignés peuvent être « remplis » avec le noeud Remplacer. Dans les noeuds Filler et Derive (mode Multiple uniquement), la fonction CLEM spéciale `@FIELD` fait référence aux champs en cours.

Champs spéciaux

Les fonctions spéciales servent à signaler les champs spéciaux en cours d'analyse ou à générer une liste de champs en tant qu'entrée. Par exemple, lors du calcul simultané de plusieurs champs, utilisez `@FIELD` pour que le programme comprenne que "l'action de calcul doit porter sur les champs sélectionnés". L'expression `log(@FIELD)` calcule un nouveau champ de logarithme pour chaque champ sélectionné.

Tableau 37. Champs spéciaux CLEM

Fonction	Résultat	Description
`@FIELD`	*Tout*	Effectue une action sur tous les champs spécifiés dans le contexte de l'expression.
`@TARGET`	*Tout*	Quand une expression CLEM est utilisée dans une fonction d'analyse définie par l'utilisateur, `@TARGET` représente le champ cible ou la "valeur correcte" de la paire cible/prédite analysée. Cette fonction est fréquemment utilisée dans les noeuds Analyse.
`@PREDICTED`	*Tout*	Quand une expression CLEM est utilisée dans une fonction d'analyse définie par l'utilisateur, `@PREDICTED` représente la valeur prédite de la paire cible/prédite analysée. Cette fonction est fréquemment utilisée dans les noeuds Analyse.
`@PARTITION_FIELD`	*Tout*	Remplace le nom du champ de partition en cours.

Tableau 37. Champs spéciaux CLEM (suite)

Fonction	Résultat	Description
`@TRAINING_PARTITION`	*Tout*	Renvoie la valeur de la partition d'apprentissage en cours. Par exemple, pour sélectionner des noeuds d'apprentissage utilisant un noeud Sélectionner, utilisez l'expression CLEM suivante : `@PARTITION_FIELD = @TRAINING_PARTITION`. Ainsi, le noeud Sélectionner sera toujours fonctionnel, quelles que soient les valeurs utilisées pour représenter chaque partition des données.
`@TESTING_PARTITION`	*Tout*	Renvoie la valeur de la partition de test en cours.
`@VALIDATION_PARTITION`	*Tout*	Renvoie la valeur de la partition de validation en cours.
`@FIELDS_BETWEEN(start, end)`	*Tout*	Renvoie la liste de noms des champs situés entre les champs de début et de fin spécifiés (ces champs étant inclus), en fonction de l'ordre naturel (c'est-à-dire l'ordre d'insertion) des champs dans les données.
`@FIELDS_MATCHING(pattern)`	*Tout*	Renvoie la liste des noms de champ qui correspondent à un motif spécifié. Il est possible d'inclure un point d'interrogation (?) dans le motif pour obtenir une correspondance exacte avec un caractère ; un astérisque (*) implique une correspondance avec zéro ou plusieurs caractères. Pour obtenir une correspondance avec un point d'interrogation ou un astérisque de littéral (plutôt que de les utiliser comme caractères génériques), il est possible d'utiliser une barre oblique inverse (\) en tant que caractère d'échappement. **Remarque :** Cela requiert un littéral chaîne comme argument ; il n'est pas possible d'utiliser une expression imbriquée pour générer l'argument.
`@MULTI_RESPONSE_SET`	*Tout*	Renvoie la liste des champs dans l'ensemble à réponses multiples nommé.

Chapitre 11. Utilisation d'IBM SPSS Modeler avec un référentiel

A propos du IBM SPSS Collaboration and Deployment Services Repository

SPSS Modeler peut être utilisé avec le référentiel IBM SPSS Collaboration and Deployment Services, ce qui vous permet de gérer le cycle de vie des modèles d'exploration de données et des objets prédictifs associés qui pourront être utilisés par des applications, outils et solutions d'entreprise. Les objets IBM SPSS Modeler qui peuvent être partagés de cette manière incluent les flux, les noeuds, les sorties de flux, les projets et les modèles. Les objets sont stockés dans le référentiel central, dans lequel ils peuvent être partagés avec d'autres applications et suivis à l'aide de capacités étendues de gestion des versions, des métadonnées et de recherche.

Avant que vous ne puissiez utiliser SPSS Modeler avec le référentiel, vous devez installer un adaptateur sur l'hôte du référentiel. Sans cet adaptateur, il est possible que le message suivant apparaisse lors d'une tentative d'accès aux objets du référentiel de certains noeuds ou modèles SPSS Modeler :

```
The repository may need updating to support new node, model and output types.
```

Pour des instructions sur l'installation de l'adaptateur, reportez-vous au guide *SPSS Modeler - Installation du déploiement*, disponible sous forme de fichier PDF avec le téléchargement de votre produit. Vous pouvez obtenir des informations sur l'accès aux objets du référentiel d'IBM SPSS Modeler depuis IBM SPSS Deployment Manager dans le guide de déploiement d'*SPSS Modeler.*

Les sections suivantes fournissent des informations sur l'accès au référentiel depuis SPSS Modeler.

Prise en charge étendue de la gestion des versions et de la recherche

Le référentiel contient des versions d'objet et des capacités de recherche étendues. Par exemple, supposons que vous créiez un flux et que vous le stockiez dans le référentiel (à partir duquel il peut être partagé avec les chercheurs d'autres services). Si vous mettez ensuite le flux à jour dans SPSS Modeler, vous pouvez ajouter la version mise à jour au référentiel sans remplacer la version précédente. Toutes les versions restent accessibles et peuvent faire l'objet d'une recherche sur la base du nom, du libellé, et des champs utilisés ou d'autres attributs. Par exemple, vous pouvez rechercher toutes les versions de modèle utilisant les revenus nets comme entrée ou tous les modèles créés par un auteur en particulier. (Pour effectuer cette recherche dans un système de fichiers traditionnel, vous devriez enregistrer chaque version sous un nom de fichier différent et le logiciel ne reconnaîtrait pas les relations entre les versions.)

Connexion unique

La fonction de connexion unique permet à l'utilisateur de se connecter au référentiel sans devoir entrer son nom d'utilisateur et son mot de passe à chaque fois. Les informations de connexion du réseau local existant fournissent l'authentification nécessaire à IBM SPSS Collaboration and Deployment Services. Cette fonction dépend des paramètres suivants :

- IBM SPSS Collaboration and Deployment Services doit être configuré pour utiliser un fournisseur de connexion unique.
- L'utilisateur doit être connecté à un hôte compatible avec le fournisseur.

Pour plus d'informations, voir «Connexion au référentiel», à la page 192.

Stockage et déploiement des objets du référentiel

Les flux créés dans IBM SPSS Modeler peuvent être **stockés** tels quels dans le référentiel, tout comme les fichiers avec l'extension `.str`. De cette façon, plusieurs utilisateurs de la même entreprise peuvent accéder à un même flux. Pour plus d'informations, voir la rubrique «Stockage des objets dans le référentiel», à la page 193.

Il est également possible de déployer un flux dans le référentiel. Un flux déployé est stocké en tant que fichier avec des métadonnées supplémentaires. Un flux déployé peut utiliser les fonctions au niveau de l'entreprise d'IBM SPSS Collaboration and Deployment Services, telles que le scoring et le rafraîchissement de modèle automatisés. Par exemple, un modèle peut être automatiquement mis à jour à intervalles réguliers lorsque de nouvelles données sont disponibles. Il est aussi possible de déployer un ensemble de flux pour l'analyse Champion Challenger dans laquelle des flux sont comparés afin de déterminer celui qui contient le modèle prédictif le plus efficace.

Remarque : Les noeuds de modélisation Règles d'association, STP et TCM ne prennent pas en charge les étapes Evaluation de modèle ou Champion Challenger dans IBM SPSS Collaboration and Deployment Services.

Vous pouvez déployer un flux (avec l'extension `.str`). Le déploiement en tant que flux permet l'utilisation du flux par IBM SPSS Modeler Advantage de l'application client léger. Pour plus d'informations, voir la rubrique «Ouverture d'un flux dans IBM SPSS Modeler Advantage», à la page 211.

Pour obtenir plus d'informations, reportez-vous au «Options de déploiement des flux», à la page 205.

Autres options de déploiement

L'application IBM SPSS Collaboration and Deployment Services offre des fonctions très étendues pour la gestion du contenu d'entreprise, mais d'autres mécanismes sont également disponibles pour le déploiement ou l'exportation de flux :

- Exportez le flux et le modèle afin de pouvoir les utiliser ultérieurement avec IBM SPSS Modeler Solution Publisher Runtime.
- Exportez un ou plusieurs modèles au format PMML. Ce format XML permet de coder les informations sur le modèle. Pour plus d'informations, voir la rubrique «Importation et exportation de modèles au format PMML», à la page 212.

Connexion au référentiel

1. Pour vous connecter au référentiel, à partir du menu IBM SPSS Modeler principal, cliquez sur :
   Outils > **Référentiel** > **Options...**
2. Dans la zone **URL du référentiel** , entrez ou sélectionnez le chemin d'accès, ou l'URL, de l'installation du référentiel auquel vous voulez accéder. Vous pouvez vous connecter à un seul référentiel à la fois.
   Les paramètres sont propres à chaque site ou installation. Pour les détails de connexion spécifiques, contactez votre administrateur système local.

Définir les données d'identification. Laissez cette case décochée pour permettre à la fonction de *connexion unique* de se connecter à l'ordinateur local à l'aide de vos informations de nom d'utilisateur et de mot de passe. Si la connexion unique n'est pas disponible, ou si vous sélectionnez cette option pour désactiver la connexion unique (par exemple pour vous connecter à un compte administrateur), un écran s'affiche dans lequel vous entrez vos données d'identification.

Saisie des informations d'identification pour le référentiel

Selon vos paramètres, vous devrez peut-être renseigner les champs suivants dans la boîte de dialogue Référentiel : identifiants :

ID utilisateur et mot de passe. Indiquez un nom d'utilisateur et un mot de passe valides lors de la connexion. Si nécessaire, contactez votre administrateur local pour plus d'informations.

Fournisseur. Choisir un fournisseur de sécurité pour l'authentification. Le référentiel peut être configuré pour utiliser différents fournisseurs de sécurité, si nécessaire, contactez votre administrateur local pour plus d'informations.

Enregistrer le nom de référentiel et l'ID utilisateur. Enregistre les paramètres en cours en tant que paramètres par défaut pour vous éviter de les entrer à nouveau à chaque connexion.

Parcours des données d'identification de référentiel

Lorsque vous vous connectez à un référentiel à partir d'un noeud source Analytic Server, Cognos, ODBC ou TM1, vous pouvez sélectionner les données d'identification préalablement enregistrées pour la connexion au référentiel. Ces données sont répertoriées dans la boîte de dialogue Sélectionner les données d'identification de référentiel. Pour sélectionner cette boîte de dialogue, cliquez sur Parcourir en regard du champ **Données d'identification**.

Dans la boîte de dialogue Sélectionner les données d'identification de référentiel, mettez en évidence les données d'identification de référentiel dans la liste qui s'affiche et cliquez sur OK. Si la liste est trop longue, entrez dans le champ **Filtrer** le nom ou une partie du nom correspondant aux données recherchées.

Accès au contenu du référentiel

Le référentiel vous permet de parcourir le contenu stocké de la même manière que l'Explorateur Windows ; vous pouvez également accéder aux différentes *versions* de chaque objet stocké.

1. Pour ouvrir la fenêtre IBM SPSS Collaboration and Deployment Services Repository à partir des menus SPSS Modeler, cliquez sur :

Outils > **Référentiel** > **Explorer...**

1. Indiquez, si nécessaire, des paramètres de connexion au référentiel. Pour plus d'informations, voir la rubrique «Connexion au référentiel», à la page 192. Pour les informations de port, de mot de passe et d'autres détails concernant la connexion, contactez votre administrateur système local.

La fenêtre de l'explorateur affiche à l'origine une vue de l'arbre de la hiérarchie des dossiers. Cliquez sur le nom du dossier pour afficher son contenu.

Les objets correspondant à la sélection ou aux critères de recherche en cours sont répertoriés dans le panneau de droite, et les informations détaillées concernant la version sélectionnée apparaissent dans le panneau inférieur droit. Les attributs affichés s'appliquent à la version la plus récente.

Stockage des objets dans le référentiel

Vous pouvez stocker les flux, les noeuds, les modèles, les palettes de modèles, les projets et les objets de sortie dans le référentiel, à partir duquel les applications et les autres utilisateurs peuvent accéder.

Vous pouvez également publier la sortie d'un flux dans le référentiel dans un format permettant aux autres utilisateurs de la consulter sur Internet grâce à IBM SPSS Collaboration and Deployment Services Deployment Portal.

Définition des propriétés d'un objet

Lorsque vous stockez un objet, la boîte de dialogue Référentiel : Stocker apparaît, pour que vous y définissiez les valeurs de certaines propriétés de l'objet. Vous pouvez :

- cliquer sur le nom et le dossier du référentiel sous lequel l'objet doit être stocké

- Ajouter des informations sur l'objet telle que le libellé de version et d'autres propriétés pouvant être recherchées
- Affecter une ou plusieurs rubriques de classification à l'objet
- Définir les options de sécurité de l'objet

Les sections suivantes décrivent les propriétés que vous pouvez définir.

Choix de l'emplacement de stockage des objets

Dans la boîte de dialogue Référentiel : Stocker, entrez les informations suivantes :

Enregistrer dans. Indique le dossier en cours : l'emplacement où l'objet sera stocké. Faites un double clic sur le nom du dossier dans la liste pour définir ce dossier comme le dossier en cours. Utilisez le bouton Dossier supérieur pour naviguer vers le dossier parent. Utilisez le bouton Nouveau dossier pour créer un dossier au niveau en cours.

Nom du fichier. Le nom sous lequel l'objet sera stocké.

Stocker. Stocke l'objet dans l'emplacement en cours.

Ajout d'informations sur les objets stockés

Tous les champs de l'onglet Informations de la boîte de dialogue Référentiel : Stocker sont facultatives.

Auteur. Le nom d'utilisateur de l'utilisateur créant l'objet dans le référentiel. Par défaut, cela affiche le nom d'utilisateur utilisé pour la connexion du référentiel mais vous pouvez modifier ce nom ici.

Libellé de version. Sélectionnez un libellé dans la liste pour indiquer la version de l'objet ou cliquez sur **Ajouter** pour créer un nouveau libellé. Evitez d'utiliser le caractère "[" dans le libellé. Vérifiez qu'aucune case n'est sélectionnée si vous ne souhaitez pas assigner un libellé à la version de cet objet. Pour plus d'informations, voir la rubrique «Affichage et modification des propriétés des objets», à la page 202.

Description. Une description de l'objet. Les utilisateurs peuvent rechercher les objets par description (reportez-vous à la remarque).

Mots-clés. Un ou plusieurs mots-clés associés à l'objet et qui peuvent être utilisés pour la recherche (reportez-vous à la remarque).

Expiration. Une date à partir de laquelle l'objet n'est plus visible aux utilisateurs réguliers, bien qu'elle puisse encore être vue par son propriétaire et par l'administrateur du référentiel. Pour définir une date d'expiration, sélectionnez l'option **Date** et saisissez la date ou choisissez-en une à l'aide du bouton du calendrier.

Stocker. Stocke l'objet dans l'emplacement en cours.

Remarque : Les informations dans les champs **Description** et **Mots-clés** sont traitées séparément de tout ce qui peut être saisi dans SPSS Modeler dans l'onglet Annotations de l'objet. Une recherche dans le référentiel par description ou mot-clé ne renvoie pas d'informations de l'onglet Annotations. Pour plus d'informations, voir la rubrique «Recherche d'objets dans le référentiel», à la page 199.

Affectation des rubriques à un objet stocké

Les rubriques sont un système de classification hiérarchique du contenu stocké dans le référentiel. Vous pouvez choisir parmi les rubriques disponibles lors du stockage d'objets et les utilisateurs peuvent également rechercher des objets par rubrique. La liste des rubriques disponibles est définie par des utilisateurs du référentiel avec les privilèges appropriés (pour plus d'informations, consultez le *Guide de l'utilisateur de Deployment Manager*).

Pour affecter une rubrique à l'objet, à partir de l'onglet Rubriques de la boîte de dialogue Référentiel : Stocker :

1. Cliquez sur le bouton **Ajouter**.
2. Cliquez sur le nom d'une rubrique dans la liste des rubriques disponibles.
3. Cliquez sur **OK**.

   Pour supprimer l'affectation d'une rubrique :

4. Sélectionnez la rubrique dans la liste de rubriques affectées.
5. Cliquez sur **Supprimer**.

Définition des options de sécurité des objets stockés

Vous pouvez définir ou modifier un certain nombre d'options de sécurité pour un objet stocké à partir de l'onglet Sécurité de la boîte de dialogue Référentiel : Stocker. Pour une ou plusieurs **entité(s)** (c'est-à-dire des utilisateurs ou des groupes d'utilisateurs), vous pouvez :

- Affecter les droits d'accès à l'objet
- Modifier les droits d'accès à l'objet
- Supprimer les droits d'accès à l'objet

Principal. Le nom d'utilisateur du référentiel d'un utilisateur ou d'un groupe qui a les droits d'accès à cet objet.

Droits d'accès. Les droits d'accès que cet utilisateur ou ce groupe ont pour cet objet.

Ajouter. Permet d'ajouter un ou plusieurs utilisateurs ou groupes à la liste de ceux ayant des droits d'accès à cet objet. Pour plus d'informations, voir la rubrique «Ajout d'un utilisateur à la liste des droits d'accès».

Modifier. Permet de modifier les droits d'accès de l'utilisateur ou du groupe à cet objet. L'accès en lecture est accordé par défaut. Cette option permet d'accorder des droits d'accès supplémentaires qui sont les droits de Propriétaire, d'Ecriture, de Supprimer et de Modifier.

Supprimer. Supprime l'utilisateur ou le groupe sélectionné de la liste des droits d'accès à cet objet.

Ajout d'un utilisateur à la liste des droits d'accès : Les options suivantes sont disponibles lorsque vous sélectionnez **Ajouter** sur l'onglet Sécurité de la boîte de dialogue Référentiel : Stocker :

Sélectionner un fournisseur. Choisir un fournisseur de sécurité pour l'authentification. Le référentiel peut être configuré pour utiliser différents fournisseurs de sécurité, si nécessaire, contactez votre administrateur local pour plus d'informations.

Rechercher. Saisissez le nom d'utilisateur du référentiel de l'utilisateur ou du groupe à ajouter puis cliquez sur **Rechercher** pour afficher le nom dans la liste d'utilisateurs. Pour ajouter plusieurs noms d'utilisateur en une fois, laissez ce champ vide et cliquez simplement sur **Rechercher** pour afficher une liste de tous les noms d'utilisateur du référentiel.

Liste des utilisateurs. Sélectionnez un ou plusieurs noms d'utilisateur et cliquez sur OK pour les ajouter à la liste des droits d'accès.

Modification des droits d'accès à un objet : Les options suivantes sont disponibles lorsque vous sélectionnez **Modifier** sur l'onglet Sécurité de la boîte de dialogue Référentiel : Stocker :

Propriétaire. Sélectionnez cette option pour accorder à cet utilisateur ou à ce groupe les droits d'accès à cet objet. Le propriétaire a le contrôle entier sur l'objet, y compris les droits d'accès Supprimer et Modifier.

Lecture. Par défaut, un utilisateur ou un groupe qui n'est pas le propriétaire de l'objet ne possède que les droits d'accès en lecture à l'objet. Sélectionnez les cases appropriées pour ajouter les droits d'accès Ecriture, Supprimer et Modifier pour cet utilisateur ou ce groupe.

Stockage des flux

Vous pouvez stocker un flux en tant que fichier *.str* dans le référentiel où il sera accessible aux autres utilisateurs.

Remarque : Pour des informations sur le déploiement d'un flux, pour bénéficier de fonctionnalités supplémentaires du référentiel, consultez «Déploiement de flux», à la page 204.

Pour stocker le flux en cours :

1. Dans le menu principal, cliquez sur :

   Fichier > **Stocker** > **Stocker en tant que flux...**
2. Indiquez, si nécessaire, des paramètres de connexion au référentiel. Pour plus d'informations, voir la rubrique «Connexion au référentiel», à la page 192. Pour les informations de port, de mot de passe et d'autres détails concernant la connexion, contactez votre administrateur système local.
3. Dans la boîte de dialogue Référentiel : Stocker, sélectionnez le dossier où vous souhaitez enregistrer l'objet, spécifiez toutes les autres informations que vous souhaitez enregistrer et cliquez sur le bouton **Stocker**. Pour plus d'informations, voir la rubrique «Définition des propriétés d'un objet», à la page 193.

Stockage de projets

Vous pouvez stocker un projet IBM SPSS Modeler entier en tant que fichier *.cpj* dans le référentiel où il sera accessible aux autres utilisateurs.

Parce qu'un fichier de projet est un conteneur pour d'autres objets IBM SPSS Modeler, vous devez informer IBM SPSS Modeler de stocker les objets du projet dans le référentiel. Vous utilisez pour cela un paramètre de la boîte de dialogue Propriétés du projet. Pour plus d'informations, voir la rubrique «Définition des propriétés d'un projet», à la page 218.

Une fois que vous avez configuré un projet en vue de stocker des objets dans le référentiel, IBM SPSS Modeler vous invite automatiquement à stocker les objets dès que vous les ajoutez au projet.

Lorsque la session IBM SPSS Modeler est terminée, vous devez stocker une nouvelle version du fichier de projet de façon à ce que les ajouts soient mémorisés. Le fichier de projet inclut automatiquement (et extrait) les versions les plus récentes des objets qu'il contient. Si vous n'avez pas ajouté d'objets à un projet au cours d'une session IBM SPSS Modeler, il est inutile de stocker de nouveau le fichier de projet. Vous devez toutefois stocker la nouvelle version des objets modifiés (flux, sortie, etc.).

Pour stocker un projet

1. Sélectionnez le projet dans l'onglet CRISP-DM ou Classes dans le panneau des gestionnaires dans IBM SPSS Modeler, et dans le menu principal, cliquez sur :

   Fichier > **Projet** > **Stocker le projet...**
2. Indiquez, si nécessaire, des paramètres de connexion au référentiel. Pour plus d'informations, voir la rubrique «Connexion au référentiel», à la page 192. Pour les informations de port, de mot de passe et d'autres détails concernant la connexion, contactez votre administrateur système local.
3. Dans la boîte de dialogue Référentiel : Stocker, sélectionnez le dossier où vous souhaitez enregistrer l'objet, spécifiez toutes les autres informations que vous souhaitez enregistrer et cliquez sur le bouton **Stocker**. Pour plus d'informations, voir la rubrique «Définition des propriétés d'un objet», à la page 193.

Stockage des noeuds

Vous pouvez stocker une définition de noeud individuel du flux en cours en tant que fichier *.nod* dans le référentiel où il sera accessible aux autres utilisateurs.

Pour stocker un noeud :

1. Cliquez avec le bouton droit de la souris sur le noeud dans l'espace de travail de flux et cliquez sur **Stocker le noeud**.
2. Indiquez, si nécessaire, des paramètres de connexion au référentiel. Pour plus d'informations, voir la rubrique «Connexion au référentiel», à la page 192. Pour les informations de port, de mot de passe et d'autres détails concernant la connexion, contactez votre administrateur système local.
3. Dans la boîte de dialogue Référentiel : Stocker, sélectionnez le dossier où vous souhaitez enregistrer l'objet, spécifiez toutes les autres informations que vous souhaitez enregistrer et cliquez sur le bouton **Stocker**. Pour plus d'informations, voir la rubrique «Définition des propriétés d'un objet», à la page 193.

Stockage des objets de sortie

Vous pouvez stocker un objet de sortie du flux en cours en tant que fichier *.cou* dans le référentiel où il sera accessible aux autres utilisateurs.

Pour stocker un objet de sortie :

1. Cliquez sur l'objet dans l'onglet Sorties dans le panneau des gestionnaires dans SPSS Modeler, et dans le menu principal, cliquez sur :

   Fichier > **Sorties** > **Stocker la sortie...**
2. Vous pouvez également cliquer avec le bouton droit de la souris sur un objet de l'onglet Sorties et cliquer sur **Stocker**.
3. Indiquez, si nécessaire, des paramètres de connexion au référentiel. Pour plus d'informations, voir la rubrique «Connexion au référentiel», à la page 192. Pour les informations de port, de mot de passe et d'autres détails concernant la connexion, contactez votre administrateur système local.
4. Dans la boîte de dialogue Référentiel : Stocker, sélectionnez le dossier où vous souhaitez enregistrer l'objet, spécifiez toutes les autres informations que vous souhaitez enregistrer et cliquez sur le bouton **Stocker**. Pour plus d'informations, voir la rubrique «Définition des propriétés d'un objet», à la page 193.

Stockage des modèles et des palettes de modèles

Vous pouvez stocker un modèle individuel en tant que fichier *.gm* dans le référentiel où il sera accessible aux autres utilisateurs. Vous pouvez également stocker le contenu complet de la palette des modèles en tant que fichier *.gen* dans le référentiel.

Stockage d'un modèle :

1. Cliquez sur l'objet de la palette Modèles dans SPSS Modeler, et dans le menu principal, cliquez sur :

   Fichier > **Modèles** > **Stocker le modèle...**
2. Vous pouvez également cliquer avec le bouton droit de la souris sur un objet de la palette des modèles et cliquer sur **Stocker le modèle**.
3. Continuez depuis la rubrique « Terminer la procédure de stockage ».

Stockage d'une palette de modèles :

1. Faites un clic droit sur l'arrière-plan de la palette des modèles.
2. Dans le menu contextuel, cliquez sur **Stocker la palette**.
3. Continuez depuis la rubrique « Terminer la procédure de stockage ».

Terminer la procédure de stockage :

1. Indiquez, si nécessaire, des paramètres de connexion au référentiel. Pour plus d'informations, voir la rubrique «Connexion au référentiel», à la page 192. Pour les informations de port, de mot de passe et d'autres détails concernant la connexion, contactez votre administrateur système local.
2. Dans la boîte de dialogue Référentiel : Stocker, sélectionnez le dossier où vous souhaitez enregistrer l'objet, spécifiez toutes les autres informations que vous souhaitez enregistrer et cliquez sur le bouton **Stocker**. Pour plus d'informations, voir la rubrique «Définition des propriétés d'un objet», à la page 193.

Extraction d'objets du référentiel

Vous pouvez extraire des flux, des modèles, des palettes de modèles, des noeuds, des projets et des objets de sortie qui ont été stockés dans le référentiel.

Remarque : en plus d'utiliser les options de menu décrites ici, vous pouvez également extraire des flux, des objets de sortie, des modèles et des palettes de modèles en faisant un clic droit sur l'onglet approprié du panneau des gestionnaires en haut à droite de la fenêtre SPSS Modeler.

1. Pour extraire un flux, dans le menu IBM SPSS Modeler principal, cliquez sur :

   Fichier > **Extraire le flux...**
2. Pour extraire un modèle, une palette de modèles, un projet ou un objet de sortie, dans le menu IBM SPSS Modeler principal, cliquez sur :

   Fichier > **Modèles** > **Extraire le modèle...**

   ou

   Fichier > **Modèles** > **Extraire la palette Modèles...**

   ou

   Fichier > **Projets** > **Extraire un projet...**

   ou

   Fichier > **Sorties** > **Extraire la sortie...**
3. Vous pouvez également cliquer avec le bouton droit de la souris dans le panneau du projet ou des gestionnaires, puis cliquer sur **Extraire** dans le menu contextuel.
4. Pour extraire un noeud, dans le menu IBM SPSS Modeler principal, cliquez sur :

   Insertion > **Noeud depuis le référentiel/Super noeud à partir du référentiel...**

   a. Indiquez, si nécessaire, des paramètres de connexion au référentiel. Pour plus d'informations, voir la rubrique «Connexion au référentiel», à la page 192. Pour les informations de port, de mot de passe et d'autres détails concernant la connexion, contactez votre administrateur système local.
5. Dans la boîte de dialogue Référentiel : Récupérer, recherchez l'objet, sélectionnez-le et cliquez sur le bouton **Extraire**. Pour plus d'informations, voir la rubrique .

Choix d'un objet à extraire

Les options suivantes sont disponibles dans la boîte de dialogue Référentiel : Récupérer/Rechercher :

Rechercher dans. Affiche la hiérarchie des dossiers pour le dossier en cours. Pour naviguer vers un autre dossier, sélectionnez-en un dans cette liste pour vous y rendre directement, ou naviguez à l'aide de la liste des objets sous ce champ.

Bouton Dossier supérieur. Passe à un niveau au-dessus du dossier en cours dans la hiérarchie.

Bouton Nouveau dossier. Crée un nouveau dossier au niveau en cours dans la hiérarchie.

Nom du fichier. Le nom de fichier du référentiel de l'objet sélectionné. Pour extraire cet objet, cliquez sur **Extraire**.

Fichiers de type. Le type d'objet que vous avez choisi d'extraire. Seuls les objets de ce type, en plus des dossiers, apparaissent dans la liste d'objets. Pour afficher les objets d'un autre type d'extraction, sélectionnez le type d'objet dans la liste.

Ouvrir comme verrouillé. Par défaut, lorsqu'un objet est extrait, il est verrouillé dans le référentiel pour que les autres personnes ne puissent pas le mettre à jour. Si vous ne souhaitez pas que l'objet soit verrouillé lors de l'extraction, décochez cette case.

Description, mots-clés. Si des informations supplémentaires sur l'objet ont été définies lors de son stockage, ces informations apparaissent ici. Pour plus d'informations, voir la rubrique «Ajout d'informations sur les objets stockés», à la page 194.

Version. Pour extraire une version de l'objet autre que la plus récente, cliquez sur le bouton. Les informations détaillées de toutes les versions qui apparaissent vous permettent de sélectionner la version de votre choix.

Sélection de la version d'objet

Pour sélectionner la version spécifique d'un objet du référentiel, dans la boîte de dialogue Référentiel : Sélectionner une version :

1. (Facultatif) Pour trier la liste par version, libellé, taille, date de création ou créateur, double-cliquez sur l'en-tête de colonne approprié.
2. Sélectionnez la version d'objet à utiliser.
3. Cliquez sur Poursuivre.

Recherche d'objets dans le référentiel

Vous pouvez rechercher des objets en fonction du nom, du dossier, du type, du libellé, de la date ou d'autres critères.

Recherche d'objets en fonction du nom

1. Dans le menu IBM SPSS Modeler principal, cliquez sur :

   Outils > **Référentiel** > **Explorer...**

   a. Indiquez, si nécessaire, des paramètres de connexion au référentiel. Pour plus d'informations, voir la rubrique «Connexion au référentiel», à la page 192. Pour les informations de port, de mot de passe et d'autres détails concernant la connexion, contactez votre administrateur système local.
2. Cliquez sur l'onglet **Rechercher**.
3. Dans le champ **Rechercher des objets nommés**, spécifiez le nom de l'objet à rechercher.

Lorsque vous recherchez des objets en fonction de leur nom, vous pouvez utiliser un astérisque (`*`) comme caractère générique remplaçant toute chaîne de caractères et un point d'interrogation (`?`) qui, lui, remplace un caractère unique. Par exemple, une recherche portant sur `*cluster*` renvoie tous les objets dont le nom inclut la chaîne `cluster`. La recherche de la chaîne `m0?_*` renvoie `M01_cluster.str` et `M02_cluster.str` mais pas `M01a_cluster.str`. Les recherches ne distinguent pas les majuscules des minuscules (la recherche de la chaîne `cluster` renvoie aussi bien `Cluster` que `CLASSE`).

Remarque : Si le nombre d'objets est élevé, les recherches peuvent prendre un certain temps.

Recherche en fonction d'autres critères

Vous pouvez effectuer une recherche en fonction du titre, du libellé, de la date, de l'auteur, de mots-clés, d'un contenu indexé ou de la description d'un objet. Seuls les objets correspondant à *tous* les critères de

recherche indiqués sont détectés. Par exemple, vous pouvez localiser tous les flux qui contiennent un ou plusieurs modèles de classification, présentant un libellé spécifique et ayant été modifiés après une date définie.

Types d'objets. Vous pouvez restreindre la recherche aux modèles, aux flux, aux sorties, aux noeuds, aux super noeuds, aux projets, aux palettes de modèles ou à d'autres types d'objets.

- **Modèles.** Vous pouvez rechercher des modèles par catégorie (classification, approximation, classification non supervisée, etc.) ou par algorithme de modélisation particulier tel que Kohonen.

  Vous pouvez également effectuer une recherche selon les champs utilisés, comme tous les modèles qui utilisent un champ *revenu* en tant que champ d'entrée ou de sortie (cible).
- **Flux.** Pour les flux, vous pouvez restreindre la recherche aux champs utilisés ou au type de modèle (catégorie ou algorithme) contenu dans le flux.

Rubriques. Vous pouvez rechercher les modèles associés à des rubriques particulières dans une liste définie par les utilisateurs du référentiel qui ont les droits d'accès appropriés (pour plus d'informations, consultez le *guide d'utilisation de Deployment Manager*). Pour obtenir la liste, cochez cette case puis cliquez sur le bouton « Ajouter des rubriques » qui apparaît, sélectionnez une ou plusieurs rubriques dans la liste puis cliquez sur OK.

Libellé. Limite la recherche à des libellés de version d'objet spécifiques.

Dates. Vous pouvez indiquer une date de création ou de modification, et rechercher les objets antérieurs, ultérieurs ou appartenant à l'intervalle de dates indiqué.

Auteur. Limite la recherche à des objets créés par un utilisateur particulier.

Mots-clés. Recherche reposant sur des mots-clés précis. Dans IBM SPSS Modeler, les mots-clés sont indiqués dans l'onglet Annotation d'un flux, d'un modèle ou d'un objet de sortie.

Description. Recherche des termes précis figurant dans le champ de description. Dans IBM SPSS Modeler, la description figure dans l'onglet Annotation d'un flux, d'un modèle ou d'un objet de sortie. Vous pouvez séparer les expressions recherchées par des points-virgules, comme l'illustre l'exemple suivant : `revenu ; type de culture ; valeur de la demande`. (Au sein d'une expression recherchée, les espaces ont une importance. Par exemple, l'occurrence `type de culture` dotée d'un espace et l'occurrence `type de culture` incluant deux espaces sont différentes.)

Modification des objets du référentiel

Vous pouvez modifier des objets existants dans le référentiel directement depuis SPSS Modeler. Vous pouvez :

- Créer, renommer ou supprimer des dossiers
- Verrouiller ou déverrouiller des objets
- Supprimer des objets

Création, changement de nom et suppression de dossiers

1. Pour effectuer des opérations sur des dossiers dans le référentiel, dans le menu SPSS Modeler principal, cliquez sur :

   Outils > **Référentiel** > **Explorer...**

   a. Indiquez, si nécessaire, des paramètres de connexion au référentiel. Pour plus d'informations, voir la rubrique «Connexion au référentiel», à la page 192. Pour les informations de port, de mot de passe et d'autres détails concernant la connexion, contactez votre administrateur système local.
2. Vérifiez que l'onglet **Dossiers** est actif.

3. Pour créer un nouveau dossier, cliquez avec le bouton droit de la souris sur le dossier parent et cliquez sur **Nouveau dossier**.
4. Pour renommer un dossier, cliquez dessus avec le bouton droit de la souris et cliquez sur **Renommer le dossier**.
5. Pour supprimer un dossier, cliquez dessus avec le bouton droit de la souris et cliquez sur **Supprimer le dossier**.

Verrouillage et déverrouillage d'objets de référentiel

Vous pouvez verrouiller un objet pour empêcher les autres utilisateurs de mettre à jour ses versions existantes ou de créer de nouvelles versions. Un objet verrouillé est indiqué par un symbole de cadenas sur l'icône de l'objet.

Figure 17. Objet verrouillé

Pour verrouiller un objet

1. Dans la fenêtre de l'explorateur du référentiel, cliquez avec le bouton droit de la souris sur l'objet souhaité.
2. Cliquez sur **Verrouiller**.

Pour déverrouiller un objet

1. Dans la fenêtre de l'explorateur du référentiel, cliquez avec le bouton droit de la souris sur l'objet souhaité.
2. Cliquez sur **Déverrouiller**.

Suppression d'objets du référentiel

Avant de supprimer un objet du référentiel, vous devez choisir de supprimer toutes les versions ou seulement une version particulière de l'objet.

Pour supprimer toutes les versions d'un objet :

1. Dans la fenêtre de l'explorateur du référentiel, cliquez avec le bouton droit de la souris sur l'objet souhaité.
2. Cliquez sur **Supprimer les objets**.

Pour supprimer la version la plus récente d'un objet :

1. Dans la fenêtre de l'explorateur du référentiel, cliquez avec le bouton droit de la souris sur l'objet souhaité.
2. Cliquez sur **Supprimer**.

Pour supprimer une version antérieure d'un objet :

1. Dans la fenêtre de l'explorateur du référentiel, cliquez avec le bouton droit de la souris sur l'objet souhaité.
2. Cliquez sur **Supprimer les versions**.
3. Sélectionnez les versions à supprimer et cliquez sur **OK**.

Gestion des propriétés des objets du référentiel

Vous pouvez contrôler différentes propriétés d'objets depuis SPSS Modeler. Vous pouvez :

- Afficher les propriétés d'un dossier
- Afficher et modifier les propriétés d'un objet
- Créer, appliquer et supprimer les libellés de version d'un objet

Affichage des propriétés d'un dossier

Pour visualiser les propriétés d'un dossier dans la fenêtre du référentiel, cliquez avec le bouton droit de la souris sur le dossier requis. Cliquez sur **Propriétés du dossier**.

Onglet Général

Cet onglet affiche le nom du dossier et ses dates de création et de modification.

Onglet Autorisations

Cet onglet permet de définir les droits d'accès en lecture et en écriture pour le dossier. Tous les utilisateurs et les groupes ayant accès au dossier parent sont répertoriés. Les autorisations respectent une hiérarchie. Par exemple, si vous ne disposez pas d'un droit d'accès en lecture, vous ne pouvez pas bénéficier d'un droit d'accès en écriture. De même, si vous ne disposez pas d'un droit d'accès en écriture, vous ne pouvez pas bénéficier d'une autorisation de suppression.

Utilisateurs et groupes. Répertorie les utilisateurs et groupes d'utilisateurs du référentiel qui ont au moins les droits d'accès en lecture à ce dossier. Sélectionnez les cases Ecriture et Supprimer pour ajouter à un utilisateur ou à un groupe spécifique les droits d'accès à ce dossier. Cliquez sur l'icône **Ajouter les utilisateurs / les groupes** située à droite de l'onglet Autorisations pour attribuer l'accès à des utilisateurs et à des groupes supplémentaires. La liste des utilisateurs et des groupes disponibles est contrôlée par l'administrateur.

Droits d'accès en cascade. Choisissez une option pour contrôler la façon dont les modifications effectuées sur le dossier en cours sont appliquées à ses dossiers enfants, s'il en existe.

- **Tous les droits d'accès en cascade.** Met en cascade les paramètres de droits d'accès du dossier en cours pour tous les dossiers enfant et inférieurs. C'est un moyen rapide de définir les autorisations de plusieurs dossiers simultanément. Définissez les autorisations requises pour le dossier parent, puis mettez-le en cascade le cas échéant.
- **Uniquement les modifications en cascade.** Met uniquement en cascade les modifications apportées depuis la dernière validation des modifications. Par exemple, si un nouveau groupe a été ajouté et que vous voulez lui donner accès à tous les dossiers de la branche Ventes, vous pouvez lui donner accès au dossier racine Ventes, puis mettre cette modification en cascade pour tous les sous-dossiers. Tous les autres droits d'accès aux sous-dossiers existants restent inchangés.
- **Ne pas mettre en cascade.** Toutes les modifications apportées s'appliquent uniquement au dossier en cours et ne sont pas répercutées en cascade sur les dossiers enfant.

Affichage et modification des propriétés des objets

Dans la boîte de dialogue Propriétés de l'objet, vous pouvez afficher et modifier les propriétés. Même si certaines propriétés ne sont pas modifiables, vous pouvez mettre un objet à jour en ajoutant une nouvelle version.

1. Dans la fenêtre du référentiel, cliquez avec le bouton droit de la souris sur l'objet requis.
2. Cliquez sur **Propriétés de l'objet**.

Onglet Général

Nom. Nom de l'objet tel qu'il apparaît dans le référentiel.

Créé le. Date de création de l'objet (et non de la version).

Dernière modification. Date de modification de la version la plus récente.

Auteur. Nom de connexion de l'utilisateur.

Description. Par défaut, ce champ contient la description indiquée dans l'onglet Annotation de l'objet dans SPSS Modeler.

Rubriques connexes. Le référentiel permet l'organisation des modèles et des objets liés en rubriques le cas échéant. La liste des rubriques disponibles est définie par des utilisateurs du référentiel avec les privilèges appropriés (pour plus d'informations, consultez le *Guide de l'utilisateur de Deployment Manager*).

Mots-clés. Définissez les mots-clés dans l'onglet Annotation d'un flux, d'un modèle ou d'un objet de sortie. Les mots-clés sont séparés par des espaces et sont limités à 255 caractères. (Si les mots-clés comportent des espaces, utilisez des guillemets pour les séparer.)

Onglet Versions

Les objets stockés dans le référentiel peuvent avoir plusieurs versions. L'onglet Versions affiche des informations sur chaque version.

Les propriétés suivantes peuvent être spécifiées ou modifiées pour certaines versions spécifiques d'un objet stocké :

Version. Identificateur unique de la version, généré selon l'heure à laquelle la version a été stockée.

Libellé. Libellé en cours de la version. Contrairement à l'identificateur de la version, vous pouvez déplacer les libellés d'une version d'un objet à une autre.

La taille de fichier, la date de création et l'auteur sont également indiqués pour chaque version.

Modifier les libellés. Cliquez sur l'icône **Modifier les libellés** en haut à droite de l'onglet Versions pour définir, appliquer ou supprimer des libellés pour des objets stockés. Pour plus d'informations, voir la rubrique «Gestion des libellés de version d'objet», à la page 204.

Onglet Autorisations

L'onglet Autorisations vous permet de définir les droits d'accès en lecture et en écriture de l'objet. Tous les utilisateurs et les groupes ayant accès à l'objet en cours sont répertoriés. Les autorisations respectent une hiérarchie. Par exemple, si vous ne disposez pas d'un droit d'accès en lecture, vous ne pouvez pas bénéficier d'un droit d'accès en écriture. De même, si vous ne disposez pas d'un droit d'accès en écriture, vous ne pouvez pas bénéficier d'une autorisation de suppression.

Utilisateurs et groupes. Répertorie les utilisateurs et groupes d'utilisateurs du référentiel qui ont au moins les droits d'accès en lecture à cet objet. Sélectionnez les cases Ecriture et Supprimer pour ajouter à un utilisateur ou à un groupe spécifique les droits d'accès à cet objet. Cliquez sur l'icône **Ajouter les utilisateurs / les groupes** située à droite de l'onglet Autorisations pour attribuer l'accès à des utilisateurs et à des groupes supplémentaires. La liste des utilisateurs et des groupes disponibles est contrôlée par l'administrateur.

Gestion des libellés de version d'objet

La boîte de dialogue Modifier les libellés de version permet d'effectuer les opérations suivantes :

- Appliquer des libellés à l'objet sélectionné
- Supprimer des libellés de l'objet sélectionné
- Définir un nouveau libellé et l'appliquer à l'objet

Pour appliquer des libellés à l'objet sélectionné

1. Sélectionnez un ou plusieurs libellés dans la liste **Libellés disponibles**.
2. Cliquez sur la flèche droite pour déplacer les libellés sélectionnées vers la liste **Libellés appliquées**.
3. Cliquez sur **OK**.

Pour supprimer des libellés de l'objet

1. Sélectionnez un ou plusieurs libellés dans la liste **Libellés appliquées**.
2. Cliquez sur la flèche gauche pour déplacer les libellés sélectionnées vers la liste **Libellés disponibles**.
3. Cliquez sur **OK**.

Pour définir un nouveau libellé et l'appliquer à l'objet

1. Saisissez le nom du libellé dans le champ **Nouveau libellé**.
2. Cliquez sur la flèche droite pour déplacer le nouveau libellé vers la liste **Libellés appliqués**.
3. Cliquez sur **OK**.

Déploiement de flux

Pour qu'un flux puisse être utilisé avec l'application client léger IBM SPSS Modeler Advantage, il doit être déployé en tant que flux (fichier `.str`) dans le référentiel.

Remarque : Vous ne pouvez pas déployer un flux ayant plus d'un noeud source dans la branche de scoring.

Le flux peut profiter des fonctionnalités de niveau entreprise d'IBM SPSS Collaboration and Deployment Services. Pour plus d'informations, voir la rubrique «Stockage et déploiement des objets du référentiel», à la page 192.

Pour déployer le flux en cours (méthode du menu Fichier)

1. Dans le menu principal, cliquez sur :
   Fichier > **Stocker** > **Déployer**
2. Choisissez le type de déploiement et suivez les autres instructions de la boîte de dialogue si nécessaire.
3. Cliquez sur **Déployer en tant que flux** pour déployer le flux à utiliser avec IBM SPSS Modeler Advantage ou IBM SPSS Collaboration and Deployment Services.
4. Cliquez sur **Stocker**. Pour plus d'informations, cliquez sur **Aide**.
5. Continuez depuis la rubrique Terminer le processus de déploiement.

Pour déployer le flux en cours (méthode du menu Outils)

1. Dans le menu principal, cliquez sur :
   Outils > **Propriétés de flux** > **Déploiement**
2. Choisissez le type de déploiement, remplissez le reste de l'onglet Déploiement en fonction des besoins et cliquez sur **Stocker**. Pour plus d'informations, voir la rubrique «Options de déploiement des flux», à la page 205.

Terminer le processus de déploiement

1. Indiquez, si nécessaire, des paramètres de connexion au référentiel. Pour plus d'informations, voir la rubrique «Connexion au référentiel», à la page 192. Pour les informations de port, de mot de passe et d'autres détails concernant la connexion, contactez votre administrateur système local.
2. Dans la boîte de dialogue Référentiel : Stocker, sélectionnez le dossier où vous souhaitez enregistrer l'objet, spécifiez toutes les autres informations que vous souhaitez enregistrer et cliquez sur **Stocker**. Pour plus d'informations, voir la rubrique «Définition des propriétés d'un objet», à la page 193.

Options de déploiement des flux

L'onglet Déploiement dans la boîte de dialogue Options du flux vous permet de spécifier les options de déploiement du flux.

Lorsque vous déployez un flux, vous pouvez ouvrir et modifier le flux dans l'application client léger IBM SPSS Modeler Advantage. Le flux est stocké dans le référentiel en tant que fichier avec l'extension `.str`.

Le déploiement d'un flux vous permet de profiter des fonctionnalités supplémentaires disponibles avec IBM SPSS Collaboration and Deployment Services, comme l'accès multi-utilisateur, le scoring automatique, le rafraîchissement de modèle et l'analyse Champion Challenger.

Remarque : Les noeuds de modélisation Règles d'association, STP et TCM ne prennent pas en charge les étapes Evaluation de modèle ou Champion Challenger dans IBM SPSS Collaboration and Deployment Services.

Dans l'onglet Déploiement, vous pouvez également prévisualiser la description du flux qu'IBM SPSS Modeler crée pour le flux. Pour plus d'informations, voir la rubrique «Descriptions de flux», à la page 57.

Type de déploiement. Sélectionnez la façon dont vous souhaitez déployer le flux. Tous les flux nécessitent un noeud de scoring avant de pouvoir être déployés. Les exigences et options supplémentaires dépendent du type de déploiement.

Remarque : Les noeuds de modélisation Règles d'association, STP et TCM ne prennent pas en charge les étapes Evaluation de modèle ou Champion Challenger dans IBM SPSS Collaboration and Deployment Services.

- **<aucun>**. Le flux ne sera pas déployé dans le référentiel. Toutes les options sont désactivées à l'exception de l'aperçu de la description du flux.
- **Uniquement le scoring**. Le flux est déployé dans le référentiel lorsque vous cliquez sur le bouton **Stocker**. Il est possible de stocker des données à l'aide du noeud que vous avez désigné dans le champ **Noeud de scoring**.
- **Rafraîchissement de modèle**. Identique à Uniquement le scoring, mais en outre le modèle peut être mis à jour dans le référentiel à l'aide des objets que vous avez désignés dans les champs **Noeud de modélisation** et **Nugget de modèle**. Notez que l'actualisation automatique de modèle n'est pas prise en charge par défaut dans IBM SPSS Collaboration and Deployment Services ; vous devez donc choisir ce type de déploiement si vous voulez utiliser cette fonction au cours de l'exécution d'un flux à partir du référentiel. Pour plus d'informations, voir «Rafraîchissement du modèle», à la page 207.

Noeud de scoring. Sélectionnez un graphique, un noeud de sortie ou d'exportation pour identifier la branche du flux à utiliser pour le scoring des données. Bien que le flux puisse en réalité contenir n'importe quel nombre de branches, de modèles et de noeuds terminaux valides, une seule et unique branche de scoring doit être indiquée à des fins de déploiement. C'est la condition de base requise pour déployer un flux.

Paramètres de scoring. Permet d'indiquer des paramètres pouvant être modifiés lors de l'exécution de la branche de scoring. Pour plus d'informations, voir «Paramètres de scoring et de modélisation», à la page 206.

Noeud de modélisation. Pour le rafraîchissement de modèle, spécifie le noeud de modélisation utilisé pour régénérer ou mettre à jour le modèle dans le référentiel. Doit être un noeud de modélisation du même type que celui spécifié pour le **Nugget de modèle**.

Paramètres du concepteur de modèle. Permet d'indiquer des paramètres pouvant être modifiés lors de l'exécution du noeud de modélisation. Pour plus d'informations, voir «Paramètres de scoring et de modélisation».

Nugget de modèle. Dans le cas d'un rafraîchissement de modèle, indique le nugget de modèle à mettre à jour ou à régénérer à chaque mise à jour du flux dans le référentiel (généralement dans le cadre d'un travail planifié). Le modèle doit se trouver sur la branche de scoring. La branche de scoring peut comporter plusieurs modèles, mais vous ne pouvez en désigner qu'un seul. Au moment de la création du flux, il peut s'agir en fait d'un modèle substituable mis à jour ou régénéré chaque fois que des nouvelles données sont disponibles.

Déployer comme flux. Cliquez sur cette option si vous souhaitez utiliser le flux avec IBM SPSS Modeler Advantage ou IBM SPSS Collaboration and Deployment Services.

Vérifier. Cliquez sur ce bouton pour vérifier s'il s'agit d'un flux valide pour le déploiement. Tous les flux doivent avoir un noeud de scoring désigné avant de pouvoir être déployés. Des messages d'erreur apparaissent si ces conditions ne sont pas satisfaites.

Stocker. Déploie le flux s'il est valide. Dans le cas contraire, un message d'erreur apparaît. Cliquez sur le bouton **Réparer**, corrigez l'erreur, puis réessayez.

Prévisualiser la description du flux. Vous permet d'afficher le contenu de la description du flux créé par IBM SPSS Modeler pour le flux. Pour plus d'informations, voir «Descriptions de flux», à la page 57.

Remarque : Les noeuds de modélisation Règles d'association, STP et TCM ne prennent pas en charge les étapes Evaluation de modèle ou Champion Challenger dans IBM SPSS Collaboration and Deployment Services.

Paramètres de scoring et de modélisation

Lorsque vous déployez un flux dans IBM SPSS Collaboration and Deployment Services, vous pouvez choisir les paramètres pouvant être affichés et modifiés chaque fois que le modèle est mis à jour ou évalué. Par exemple, il est possible d'indiquer des valeurs minimum ou maximum ou une autre valeur susceptible de varier à chaque exécution d'un travail.

1. Pour rendre un paramètre visible afin que vous puissiez l'afficher ou le modifier une fois le flux déployé, sélectionnez-le dans la liste de la boîte de dialogue Paramètres de scoring.

La liste des paramètres disponibles figure dans l'onglet Paramètres de la boîte de dialogue Propriétés du flux. Pour plus d'informations, voir la rubrique «Définition des paramètres de flux et de session», à la page 53.

La branche de scoring

Si vous déployez un flux, une branche de ce flux doit être désignée comme **branche de scoring** (c'est-à-dire celle contenant le noeud de scoring). Lorsque vous désignez une branche comme branche de scoring, cette branche est mise en évidence dans l'espace de travail, tout comme le lien du modèle vers le nugget de la branche de scoring. Cette représentation visuelle est particulièrement utile dans les flux complexes avec plusieurs branches, où la branche de scoring n'est pas immédiatement évidente.

Remarque : un seul flux peut être désigné comme branche de scoring.

Si le flux contient déjà une branche de scoring définie, la nouvelle branche désignée la remplace comme branche de scoring. Vous pouvez définir la couleur de l'indication de la branche de scoring à l'aide de l'option Couleur personnalisée. Pour plus d'informations, voir la rubrique «Définition des options d'affichage», à la page 226.

Vous pouvez afficher ou masquer l'indication de la branche de scoring à l'aide du bouton de la barre d'outils Afficher/Masquer le balisage des flux.

Figure 18. Bouton de barre d'outils Afficher/Masquer le balisage des flux

Identification de la branche de scoring pour le déploiement

Vous pouvez désigner la branche de scoring depuis le menu contextuel d'un noeud terminal ou depuis le menu Outils. Si vous utilisez le menu contextuel, le noeud de scoring est défini automatiquement dans l'onglet Déploiement des propriétés du flux.

Pour désigner une branche comme branche de scoring (menu contextuel)

1. Connectez le nugget de modèle à un noeud terminal (un noeud d'exécution ou de sortie en aval du nugget).
2. Faites un clic droit sur le noeud terminal.
3. Dans le menu, cliquez sur **Utiliser en tant que branche de scoring**.

Pour désigner une branche comme branche de scoring (menu Outils)

1. Connectez le nugget de modèle à un noeud terminal (un noeud d'exécution ou de sortie en aval du nugget).
2. Dans le menu principal, cliquez sur :

   Outils > **Propriétés de flux** > **Déploiement**
3. Dans la liste des **Type de déploiement**, cliquez sur **Scoring uniquement** ou **Rafraîchissement de modèle** selon les besoins. Pour plus d'informations, voir la rubrique «Options de déploiement des flux», à la page 205.
4. Cliquez sur le champ **Noeud de scoring** et sélectionnez un noeud terminal dans la liste.
5. Cliquez sur **OK**.

Rafraîchissement du modèle

Le rafraîchissement de modèle est le processus qui permet de reconstruire un modèle existant dans un flux à l'aide de données plus récentes. Le flux lui-même n'est pas modifié dans le référentiel. Par exemple, le type d'algorithme et les paramètres spécifiques au flux restent les mêmes mais le modèle est de nouveau formé à partir des nouvelles données et mis à jour si la nouvelle version du modèle fonctionne mieux que l'ancienne.

Un seul nugget de modèle dans un flux peut être défini pour être rafraîchi : il s'appelle le **modèle de rafraîchissement**. Si vous cliquez sur l'option **Rafraîchissement de modèle** dans l'onglet Déploiement des propriétés du flux (voir «Options de déploiement des flux», à la page 205), le nugget de modèle que vous désignez à ce moment devient le modèle de rafraîchissement. Vous pouvez également désigner un modèle comme modèle de rafraîchissement à partir du menu contextuel d'un nugget de modèle. Pour ce faire, le nugget doit déjà être sur la branche de scoring.

Si vous désactivez le statut « modèle de rafraîchissement » d'un nugget, cela revient à définir le type de déploiement du flux sur Scoring uniquement et l'onglet Déploiement de la boîte de dialogue des

propriétés du flux est mis à jour en conséquence. Vous pouvez activer et désactiver ce statut au moyen de l'option **Utiliser comme modèle de rafraîchissement** dans le menu contextuel du nugget dans la branche de scoring en cours.

Supprimer le lien de modèle d'un nugget dans la branche de scoring supprime également le statut de « modèle de rafraîchissement » du nugget. Vous pouvez annuler la suppression du lien de modèle avec le menu Edition ou la barre d'outils ce qui restaure le statut du « modèle de rafraîchissement » du nugget.

Sélection du modèle de rafraîchissement

Tout comme la branche de scoring, le lien vers le modèle de rafraîchissement est également mis en évidence dans le flux. Le nugget de modèle choisi comme modèle de rafraîchissement, et par conséquent le lien mis en évidence, dépend du nombre de nuggets dans le flux.

Modèle unique dans le flux

Si un unique nugget de modèle lié se trouve dans la branche de scoring lorsqu'il est identifié comme tel, ce nugget devient le modèle de rafraîchissement pour le flux.

Modèles multiples dans le flux

S'il existe plusieurs nuggets liés dans le flux, le modèle de rafraîchissement est choisi comme suit :

Si un nugget de modèle a été défini dans l'onglet Déploiement de la boîte de dialogue des propriétés du flux et se trouve également dans le flux, alors le nugget devient le modèle de rafraîchissement.

Si aucun nugget n'a été défini dans l'onglet Déploiement, ou si un nugget a été défini mais ne se trouve pas dans la branche de scoring, alors le nugget le plus proche du noeud terminal devient le modèle de rafraîchissement.

Si ensuite, vous désélectionnez tous les liens de modèle définis comme liens de rafraîchissement, seule la branche de scoring est mise en évidence et pas les liens. Le type de déploiement est défini sur Scoring uniquement.

Remarque : vous pouvez choisir de définir un des liens sur le statut Remplacer mais pas les deux. Dans ce cas, le nugget de modèle choisi comme modèle de rafraîchissement est celui qui contient un lien de rafraîchissement et qui est le plus proche du noeud terminal lorsque la branche de scoring est désignée.

Aucun modèle dans le flux

S'il n'existe pas de modèles dans le flux, ou uniquement des modèles sans lien de modèle, le type de déploiement est défini sur Scoring uniquement.

Vérification des erreurs dans une branche de scoring

Lorsque vous désignez la branche de scoring, des erreurs y sont recherchées.

Si une erreur est trouvée, la branche de scoring est mise en évidence dans la couleur d'erreur des branches de scoring et un message d'erreur apparaît. Vous pouvez définir la couleur de l'erreur à l'aide de l'option Couleur personnalisée. Pour plus d'informations, voir la rubrique «Définition des options d'affichage», à la page 226.

Si une erreur est trouvée, suivez les instructions suivantes :

1. Corrigez l'erreur en fonction du contenu du message d'erreur.
2. Dans le menu principal, cliquez sur :
   Outils > **Propriétés de flux** > **Déploiement**
   et cliquez sur **Vérifier**.

3. Si nécessaire, répétez ce processus jusqu'à ne plus trouver d'erreurs.

Chapitre 12. Exportation vers des applications externes

A propos de l'exportation vers des applications externes

IBM SPSS Modeler offre divers mécanismes pour exporter la totalité du processus d'exploration de données vers des applications externes. Ainsi, les tâches de préparation des données et de création des modèles que vous effectuez dans IBM SPSS Modeler peuvent également être utilisées dans une application externe.

La section précédente a montré comment déployer des flux dans un référentiel IBM SPSS Collaboration and Deployment Services pour bénéficier de l'accès multi-utilisateur, de la planification de travaux et d'autres fonctionnalités. De même, les flux IBM SPSS Modeler peuvent aussi être utilisés en conjonction avec :

- IBM SPSS Modeler Advantage
- Les applications qui peuvent importer et exporter des fichiers au format PMML

Pour des informations supplémentaires sur l'utilisation des flux avec IBM SPSS Modeler Advantage, consultez «Ouverture d'un flux dans IBM SPSS Modeler Advantage».

Pour de plus amples informations sur les modèles d'exportation et d'importation comme fichiers PMML, la possibilité de partager des modèles avec toutes les autres applications qui prennent en charge ce format, reportez-vous à «Importation et exportation de modèles au format PMML», à la page 212.

Ouverture d'un flux dans IBM SPSS Modeler Advantage

Les flux IBM SPSS Modeler peuvent s'utiliser avec l'application client léger IBM SPSS Modeler Advantage. Bien qu'il soit possible de créer des applications personnalisées complètes au sein d'IBM SPSS Modeler Advantage, vous pouvez aussi utiliser un flux déjà créé dans IBM SPSS Modeler comme base d'un flux de travail de l'application.

Pour ouvrir un flux dans IBM SPSS Modeler Advantage :

1. Déployez le flux dans le référentiel IBM SPSS Collaboration and Deployment Services et veillez à cliquer sur l'option **Déployer en tant que flux**. Pour plus d'informations, voir la rubrique «Déploiement de flux», à la page 204.
2. Cliquez sur le bouton de la barre d'outils Ouvrir dans IBM SPSS Modeler Advantage ou, dans le menu principal, cliquez sur :

Fichier > **Ouvrir dans IBM SPSS Modeler Advantage**

1. Indiquez, si nécessaire, des paramètres de connexion au référentiel. Pour plus d'informations, voir la rubrique «Connexion au référentiel», à la page 192. Pour les informations de port, de mot de passe et d'autres détails concernant la connexion, contactez votre administrateur système local.

Remarque : le logiciel IBM SPSS Modeler Advantage doit également être installé sur le serveur du référentiel.

1. Dans la boîte de dialogue Référentiel : Stocker, sélectionnez le dossier où vous souhaitez enregistrer l'objet, spécifiez toutes les autres informations que vous souhaitez enregistrer et cliquez sur le bouton **Stocker**. Pour plus d'informations, voir la rubrique «Définition des propriétés d'un objet», à la page 193.

Ainsi, IBM SPSS Modeler Advantage sera lancé et le flux sera déjà ouvert. Le flux est fermé dans IBM SPSS Modeler.

Importation et exportation de modèles au format PMML

PMML (Predictive Model Markup Language - langage de balisage de modèle de prévision) est un format XML permettant de décrire les modèles d'exploration de données et de statistiques, notamment les entrées des modèles, les transformations utilisées pour la préparation des données pour l'exploration de données et les paramètres définissant les modèles eux-mêmes. IBM SPSS Modeler peut importer et exporter le format PMML, et permet ainsi de partager des modèles avec d'autres applications qui prennent en charge ce format, par exemple IBM SPSS Statistics.

Pour plus d'informations sur le langage PMML, reportez-vous au site Web de l'exploration de données group (*http://www.dmg.org*).

Pour exporter un modèle

L'exportation PMML est prise en charge pour la plupart des types de modèle générés dans IBM SPSS Modeler. Pour plus d'informations, voir la rubrique «Types de modèle prenant en charge PMML», à la page 213.

1. Cliquez avec le bouton droit de la souris sur un nugget de modèle dans la palette des modèles. (Vous pouvez aussi double-cliquer sur un nugget de modèles dans l'espace de travail et sélectionner le menu Fichier.)
2. Dans le menu, cliquez sur **Exporter PMML**.
3. Dans la boîte de dialogue Exporter (ou Enregistrer), choisissez un répertoire cible et un nom unique pour le modèle.

Remarque :

Vous pouvez modifier les options d'exportation PMML dans la boîte de dialogue Options utilisateur. Dans le menu principal, cliquez sur :

Outils > **Options** > **Options utilisateur**

et cliquez sur l'onglet PMML.

Pour plus d'informations, voir la rubrique «Définition des options d'exportation PMML», à la page 227.

Pour importer un modèle enregistré au format PMML

Les modèles exportés au format PMML à partir d'IBM SPSS Modeler ou d'une autre application peuvent être importés dans la palette de modèles. Pour plus d'informations, voir la rubrique «Types de modèle prenant en charge PMML», à la page 213.

1. Cliquez avec le bouton droit de la souris sur cette palette, puis sélectionnez **Importer PMML** dans le menu.
2. Sélectionnez le fichier à importer et indiquez les options requises pour les libellés de variables.
3. Cliquez sur **Ouvrir**.

Utilisez des libellés de variables s'ils sont présents dans le modèle. Le langage PMML (Predictive Model Markup Language) permet de spécifier les noms et les libellés (par exemple, ID Référence pour *IDRéf*) des variables du dictionnaire de données. Sélectionnez cette option pour utiliser des libellés de variables s'ils apparaissent dans le langage PMML initialement exporté.

Si vous avez sélectionné l'options de libellé de variable alors que le langage PMML ne contient pas de libellé de variable, le nom des variables est utilisé.

Types de modèle prenant en charge PMML

Exportation en PMML

Modèles IBM SPSS Modelers. Les modèles suivants créés dans IBM SPSS Modeler peuvent être exportés au format PMML 4.0 :

- Arbre C&RT
- QUEST
- CHAID
- Réseau de neurones
- C5.0
- Régression logistique
- Genlin
- SVM
- Apriori
- Carma
- k moyenne
- Kohonen
- TwoStep
- TwoStep-AS
- GLMM (PMML est exporté pour tous les modèles GLMM, mais le modèle PMML a uniquement des effets fixes)
- Liste de décisions
- Cox
- Séquence (l'évaluation pour les modèles Sequence PMML n'est pas prise en charge)
- Random Trees
- Tree-AS
- Linéaire
- Linear-AS
- Régression
- Logistique
- GLE
- LSVM
- Détection des anomalies
- KNN
- Règles d'association

Modèles natifs de base de données. Pour les modèles générés à l'aide d'algorithmes natifs de base de données, l'exportation PMML n'est pas disponible. Les modèles créés avec Analysis Services de Microsoft ou Oracle Data Miner ne peuvent pas être exportés.

Importation de PMML

IBM SPSS Modeler peut importer et évaluer les modèles PMML générés par les versions actuelles de tous les produits IBM SPSS Statistics, y compris les modèles exportés depuis IBM SPSS Modeler, et les modèles ou transformations PMML générés par IBM SPSS Statistics17.0 ou une version ultérieure. En d'autres termes, il s'agit de tout PMML que le moteur de scoring peut évaluer, à l'exception des éléments suivants :

- Les modèles Apriori, CARMA, Détection des anomalies, Séquence et Règles d'association ne peuvent pas être importés.
- Les modèles PMML ne seront peut-être pas parcourus après importation dans IBM SPSS Modeler, même s'ils peuvent être utilisés pour le scoring. (Notez que cela inclut les modèles qui ont été exportés à l'origine d'IBM SPSS Modeler. Pour contourner cet obstacle, exportez le modèle au format de fichier de modèle généré [`*.gm`] plutôt qu'au format PMML.)
- La validation limitée se produit lors de l'importation, mais la validation complète s'effectue lors du scoring du modèle. Par conséquent, il est possible que l'importation réussisse mais que le scoring échoue ou produise des résultats incorrects.

Remarque : Dans le cas de PMML tiers importé dans IBM SPSS Modeler, IBM SPSS Modeler tente d'affecter un score au PMML valide pouvant être reconnu et recevoir un score. Il n'est cependant pas garanti que tout le PMML reçoive un score et que celui-ci soit le même que l'application qui l'a généré.

Chapitre 13. Projets et rapports

Introduction aux projets

Un **projet** est un groupe de fichiers lié à un travail d'exploration de données. Les projets incluent des flux de données, des graphiques, des modèles générés, des rapports, ainsi que tout élément créé dans IBM SPSS Modeler. Au premier coup d'oeil, il peut sembler que les projets IBM SPSS Modeler ne sont qu'un simple moyen d'organiser les sorties, mais ils sont en fait capables de bien plus. A l'aide des projets, vous pouvez notamment :

- Annoter chacun des objets du fichier de projet.
- Utiliser la méthodologie CRISP-DM dans vos tâches d'exploration de données. Les projets contiennent également le système d'aide CRISP-DM qui fournit des détails et des exemples réels relatifs à l'exploration de données combiné à la méthodologie CRISP-DM.
- Ajouter au projet des objets créés hors d'IBM SPSS Modeler, tels qu'un diaporama PowerPoint utilisé pour présenter vos objectifs d'exploration de données ou vos livres blancs sur les algorithmes que vous envisagez d'utiliser.
- Générer à la fois des rapports de mise à jour simples et détaillés, sur la base de vos annotations. Il est possible de générer ces rapports au format HTML afin d'en faciliter la publication sur l'intranet de votre entreprise.

Remarque : si le panneau de projet n'est pas visible dans la fenêtre IBM SPSS Modeler, cliquez sur **Project** dans le menu Vue.

Vous pouvez visualiser les objets ajoutés à un projet de deux façons : **Vue Classes** et **Vue CRISP-DM**. Tous les éléments que vous ajoutez à un projet sont ajoutés aux deux modes et vous pouvez basculer entre ces modes pour créer l'environnement de travail le plus efficace.

Vue CRISP-DM

Grâce à la prise en charge de CRISP-DM (Cross-Industry Standard Process for Data Mining), les projets IBM SPSS Modeler fournissent une méthodologie confirmée et non exclusive permettant d'organiser les éléments de vos travaux d'exploration de données. CRISP-DM utilise six phases pour décrire le processus du début (collecte des exigences métier) à la fin (déploiement de vos résultats). Même si certaines phases n'impliquent généralement pas d'actions dans IBM SPSS Modeler, le panneau de projet intègre les six phases de sorte que vous disposiez d'un emplacement central dédié au stockage et au suivi de tous les éléments associés au projet. Par exemple, la phase de compréhension de l'entreprise implique généralement la collecte des exigences et une réunion entre collègues afin de déterminer les objectifs plutôt que le traitement des données dans IBM SPSS Modeler. Le panneau de projet vous permet de stocker vos notes résultant de ce type de réunion dans le dossier *Compréhension de l'entreprise* pour vous y reporter ultérieurement et pour l'inclure dans les rapports.

La vue CRISP-DM dans le panneau de projet comporte son propre système d'aide afin de vous guider à travers le cycle de vie de l'exploration de données. Dans IBM SPSS Modeler, vous pouvez accéder à ce système d'aide en cliquant sur **Aide CRISP-DM** dans le menu Aide.

Remarque : Si le panneau de projet n'est pas visible dans la fenêtre, cliquez sur **Projet** dans le menu Vue.

Définition de la phase par défaut du projet

Les objets ajoutés à un projet le sont dans une phase par défaut de CRISP-DM. En d'autres termes, il est nécessaire d'organiser les objets manuellement selon la phase d'exploration de données dans laquelle vous les utilisez. Il est judicieux de définir le dossier par défaut en fonction de la phase dans laquelle vous travaillez actuellement.

Pour sélectionner la phase à utiliser par défaut :

1. Dans la vue CRISP-DM, cliquez avec le bouton droit de la souris sur le dossier de la phase à définir comme phase par défaut.
2. Dans le menu, cliquez sur **Définir par défaut**.

Le dossier par défaut est affiché en caractères gras.

Vue Classes

La vue Classes du panneau de projet organise votre travail dans IBM SPSS Modeler en catégories, selon les types d'objet créés. Vous pouvez ajouter des objets enregistrés à l'une des catégories suivantes :

- Flux
- Noeuds
- Modèles
- Tableaux, graphiques, rapports
- Autres (fichiers n'appartenant pas à IBM SPSS Modeler, tels que des diaporamas ou des livres blancs applicables à votre travail d'exploration de données)

L'ajout d'objets à la vue Classes entraîne également leur ajout dans le dossier de la phase par défaut de la vue CRISP-DM.

Remarque : Si le panneau de projet n'est pas visible dans la fenêtre, cliquez sur **Projet** dans le menu Vue.

Création d'un projet

Un projet est avant tout un fichier qui contient des références à tous les fichiers que vous associez au projet. En d'autres termes, les éléments du projet sont enregistrés à la fois individuellement et en tant que référence dans le fichier de projet (*.cpj*). Etant donné cette structure référentielle, il est nécessaire de prendre en considération les points suivants :

- Les éléments de projet doivent d'abord être enregistrés individuellement avant d'être ajoutés à un projet. Si un élément n'est pas enregistré, vous serez invité à l'enregistrer avant de l'ajouter au projet en cours.
- Les objets mis à jour individuellement, tels que les flux, sont également mis à jour dans le fichier de projet.
- Lorsque vous déplacez ou supprimez manuellement des objets (tels que des flux, des noeuds ou des objets de sortie) du système de fichiers, les liens du fichier de projet ne sont plus valides.

Création d'un projet

Vous pouvez facilement créer des projets dans la fenêtre IBM SPSS Modeler. Vous pouvez créer un projet à partir d'un projet ouvert, ou fermer le projet existant et créer entièrement un projet.

Dans le menu principal, cliquez sur :

Fichier > **Projet** > **Nouveau projet...**

Ajout à un projet

Après avoir créé ou ouvert un projet, vous pouvez ajouter des objets, tels que des flux de données, des noeuds et des rapports, à l'aide de plusieurs méthodes.

Ajout d'objets à partir des gestionnaires

A l'aide des gestionnaires situés dans la partie supérieure droite de la fenêtre IBM SPSS Modeler, vous pouvez ajouter des flux ou des sorties.

1. Sélectionnez un objet, tel qu'un tableau ou un flux, dans l'un des onglets des gestionnaires.
2. Cliquez avec le bouton droit de la souris sur l'élément et cliquez sur **Ajouter au projet**.

   Si l'objet a été préalablement enregistré, il sera automatiquement ajouté au dossier d'objets approprié (dans la vue Classes) ou au dossier de la phase par défaut (dans la vue CRISP-DM).
3. Vous pouvez également faire glisser les objets des gestionnaires vers le panneau des projets.

Remarque : vous pouvez être invité à enregistrer d'abord l'objet. Au moment de l'enregistrement, veillez à ce que l'option **Ajouter le fichier au projet** soit sélectionnée dans la boîte de dialogue Enregistrer. Cette sélection entraînera l'ajout automatique de l'objet au projet après l'enregistrement.

Ajout de noeuds à partir de l'espace de travail

Vous pouvez ajouter des noeuds individuels à partir de l'espace de travail de flux à l'aide de la boîte de dialogue Enregistrer.

1. Sélectionnez un noeud dans l'espace de travail.
2. Cliquez avec le bouton droit de la souris sur l'élément et cliquez sur **Enregistrer le noeud**. Sinon, dans le menu principal, cliquez sur :

   Edition > **Noeud** > **Enregistrer le noeud...**
3. Dans la boîte de dialogue Enregistrer, sélectionnez **Ajouter le fichier au projet**.
4. Nommez le noeud et cliquez sur **Enregistrer**.

Cette action enregistre le fichier et l'ajoute au projet. Les noeuds sont ajoutés au dossier *Noeuds* dans la vue Classes et au dossier de la phase par défaut dans la vue CRISP-DM.

Ajout de fichiers externes

Vous pouvez ajouter à un projet de nombreux types d'objet n'appartenant pas à IBM SPSS Modeler. Cette possibilité est utile lorsque vous gérez le processus global d'exploration de données dans IBM SPSS Modeler. Vous pouvez, par exemple, stocker des liens vers des données, des notes, des présentations et des graphiques dans un projet. Dans la vue CRISP-DM, vous pouvez ajouter des fichiers externes au dossier de votre choix. Dans la vue Classes, les fichiers externes peuvent uniquement être enregistrés dans le dossier *Autre*.

Pour ajouter des fichiers externes à un projet :

1. Faites glisser les fichiers du bureau vers le projet.

   ou
2. Cliquez avec le bouton droit de la souris sur le dossier cible dans la vue CRISP-DM ou dans la vue Classes.
3. Dans le menu, cliquez sur **Ajouter au dossier**.
4. Sélectionnez un fichier dans la boîte de dialogue et cliquez sur **Ouvrir**.

Cette opération permet d'ajouter une référence à l'objet sélectionné au sein des projets IBM SPSS Modeler.

Transfert de projets vers le IBM SPSS Collaboration and Deployment Services Repository

Vous pouvez transférer un projet entier, avec tous ses fichiers constitutifs, dans IBM SPSS Collaboration and Deployment Services Repository en une seule opération. Les objets déjà présents dans l'emplacement cible ne sont pas déplacés. Cette fonctionnalité est également disponible en sens inverse : vous pouvez transférer des projets entiers du référentiel IBM SPSS Collaboration and Deployment Services Repository vers votre système de fichiers local.

Transfert d'un projet

Vérifiez que le projet à transférer est ouvert dans le panneau de projet.

Pour transférer un projet :

1. Cliquez avec le bouton droit de la souris sur le dossier racine du projet et cliquez sur l'option **Transférer le projet**.
2. Si un message vous le demande, connectez-vous à IBM SPSS Collaboration and Deployment Services Repository.
3. Spécifiez le nouvel emplacement du projet et cliquez sur **OK**.

Définition des propriétés d'un projet

Vous pouvez personnaliser le contenu et la documentation d'un projet en utilisant la boîte de dialogue des propriétés du projet. Pour accéder aux propriétés du projet :

1. Cliquez avec le bouton droit de la souris sur un objet ou un dossier dans le panneau de projet et choisissez **Propriétés du projet**.
2. Cliquez sur l'onglet **Projet** pour spécifier les informations élémentaires sur le projet.

Créé(e). Indique la date de création du projet (non modifiable).

Récapitulatif. Vous pouvez entrer un récapitulatif de votre projet d'exploration de données. Il sera affiché dans le rapport du projet.

Contenu. Répertorie le type des composants référencés par le fichier de projet, ainsi que leur nombre (non modifiable).

Enregistrer les objets en tant que. Indique si les objets non enregistrés doivent être sauvegardés dans le système de fichiers local ou stockés dans le référentiel. Pour plus d'informations, voir la rubrique «A propos du IBM SPSS Collaboration and Deployment Services Repository», à la page 191.

Mettre à jour les références d'objet lors du chargement du projet. Sélectionnez cette option pour mettre à jour les références du projet sur ses composants. *Remarque* : les fichiers ajoutés à un projet ne sont pas enregistrés dans le fichier de projet lui-même. Une référence aux fichiers est en fait stockée dans le projet. En d'autres termes, le fait de déplacer ou de supprimer un fichier entraîne la suppression de cet objet du projet.

Annotation d'un projet

Le panneau de projet met à votre disposition différentes méthodes d'annotation de vos travaux d'exploration de données. Les annotations relatives au projet sont souvent utilisées pour assurer le suivi des objectifs et décisions de façon globale, tandis que les annotations relatives au dossier ou au noeud fournissent des détails supplémentaires. L'onglet Annotations offre assez d'espace pour consigner des détails relatifs au projet, tels que l'exclusion de données contenant des éléments manquants irrécupérables, ou d'hypothèses formulées durant l'exploration des données et pouvant s'avérer utiles.

Pour annoter un projet :

1. Sélectionnez le dossier de projet dans la vue CRISP-DM ou dans la vue Classes.
2. Cliquez avec le bouton droit de la souris sur le dossier et sélectionnez **Propriétés du projet**.
3. Cliquez sur l'onglet **Annotations**.
4. Entrez des mots-clés et un texte décrivant le projet.

Annotations et propriétés du dossier

Vous pouvez annoter des dossiers de projet individuels (à la fois dans la vue CRISP-DM et dans la vue Classes). Dans la vue CRISP-DM, ces annotations peuvent s'avérer extrêmement efficaces pour documenter les objectifs de votre entreprise pour chaque phase de l'exploration de données. Par exemple, en utilisant l'outil Annotation pour le dossier *Compréhension de l'entreprise*, vous pouvez inclure une

formulation telle que "L'objectif commercial de cette étude est de réduire le score d'attrition parmi les clients les plus importants." Ce texte pourra alors être automatiquement inclus dans le rapport du projet si vous sélectionnez **Inclure dans le rapport**.

Pour annoter un dossier :

1. Sélectionnez un dossier dans le panneau de projet.
2. Cliquez avec le bouton droit de la souris sur le dossier et cliquez sur **Propriétés du dossier**.

Dans la vue CRISP-DM, les dossiers sont annotés avec un résumé de l'objectif de chaque phase, ainsi que des lignes directrices sur l'exécution des travaux d'exploration de données. Vous pouvez éditer ou supprimer ces annotations.

Nom. Cette zone affiche le nom du champ sélectionné.

Texte de l'info-bulle. Permet de créer des info-bulles personnalisées qui apparaissent lorsque vous positionnez le pointeur de la souris sur un dossier de projet. Ces info-bulles sont utiles dans la vue CRISP-DM, par exemple, pour fournir un aperçu des objectifs de chaque phase ou pour marquer le statut d'une phase, tel que "En cours" ou "Terminé".

Champ Annotation. Utilisez ce champ pour les annotations plus longues pouvant être rassemblées dans le rapport du projet. La vue CRISP-DM inclut une description de chaque phase d'exploration de données dans l'annotation, mais vous pouvez évidemment personnaliser cet élément selon votre projet.

Inclure dans le rapport. Pour inclure l'annotation dans les rapports, sélectionnez **Inclure dans le rapport**.

Propriétés de l'objet

Vous pouvez visualiser les propriétés des objets et choisir d'inclure des objets individuels dans le rapport du projet. Pour accéder aux propriétés des objets :

1. Cliquez avec le bouton droit de la souris sur un objet dans le panneau du projet.
2. Dans le menu, cliquez sur **Propriétés de l'objet**.

Nom. Cette zone indique le nom de l'objet enregistré.

Chemin. Cette zone indique l'emplacement de l'objet enregistré.

Inclure dans le rapport. Cochez cette case si vous souhaitez que les détails de l'objet soient intégrés dans un rapport.

Fermeture d'un projet

Lorsque vous quittez IBM SPSS Modeler ou que vous ouvrez un nouveau projet, le fichier du projet existant (fichier *.cpj*) est fermé.

Certains fichiers associés à ce projet (tels que les flux, les noeuds ou les graphiques) peuvent rester ouverts. Si vous souhaitez garder ces fichiers ouverts, choisissez `Non` à l'affichage du message **... Souhaitez-vous enregistrer et fermer ces fichiers ?**

Si vous modifiez et enregistrez des fichiers associés après la fermeture d'un projet, ces nouvelles versions seront incluses dans le projet lors de sa prochaine ouverture. Pour empêcher cette action, retirez le fichier du projet ou enregistrez-le sous un nom différent.

Génération d'un rapport

L'une des fonctions les plus utiles des projets est la capacité à générer des rapports sur la base des éléments et des annotations du projet. Ce composant est essentiel pour un travail d'exploration de données efficace, comme l'explique la section relative à la méthodologie CRISP-DM. Vous pouvez générer un rapport directement dans l'un des divers types de fichier disponibles ou dans une fenêtre de sortie à l'écran, pour consultation immédiate. Vous pouvez alors imprimer, enregistrer ou afficher le rapport dans un navigateur Web. Vous avez également la possibilité de remettre les rapports enregistrés à d'autres personnes de votre entreprise.

Un même projet donne souvent lieu à la création de plusieurs rapports au cours du processus d'exploration de données ; ils sont distribués à toutes les personnes impliquées dans le projet. Le rapport sélectionne des informations sur les objets référencés à partir du fichier de projet et à partir de toutes les annotations créées. Vous pouvez créer des rapports dans la vue Classes ou dans la vue CRISP-DM.

Pour générer un rapport :

1. Sélectionnez le dossier de projet dans la vue CRISP-DM ou dans la vue Classes.
2. Cliquez avec le bouton droit de la souris sur le dossier et cliquez sur **Rapport du projet**.
3. Spécifiez les options du rapport et cliquez sur **Générer le rapport**.

Les options de la boîte de dialogue du rapport offrent différentes façons de générer le type de rapport souhaité :

Nom de la sortie. Spécifiez le nom de la fenêtre de sortie si vous souhaitez afficher les données du rapport à l'écran. Vous pouvez indiquer un nom personnalisé ou autoriser IBM SPSS Modeler à attribuer automatiquement un nom à la fenêtre.

Sortie à l'écran. Sélectionnez cette option pour générer et afficher le rapport dans une fenêtre de sortie. Notez que vous avez la possibilité d'exporter le rapport vers différents types de fichier à partir de la fenêtre de sortie.

Sortie dans le fichier. Sélectionnez cette option pour générer et enregistrer le rapport au format indiqué dans la liste Type de fichier.

Fichier. Indiquez le nom de fichier du rapport généré. Les fichiers sont enregistrés par défaut dans le répertoire d'IBM SPSS Modeler *\bin*. Utilisez le bouton **...** pour indiquer un autre emplacement.

Type de fichier. Les types de fichier disponibles sont les suivants :

- **Document HTML.** Le rapport est enregistré sous la forme d'un fichier HTML unique. S'il contient des graphiques, ils sont enregistrés comme des fichiers PNG référencés par le fichier HTML. Lors de la publication du rapport sur Internet, veillez à envoyer le fichier HTML et les images référencées.
- **Document texte.** Le rapport est enregistré sous la forme d'un fichier texte unique. S'il contient des graphiques, seuls le nom de fichier et le chemin d'accès sont inclus dans le rapport.
- **Document MS Word.** Le rapport est enregistré sous la forme d'un document unique, les graphiques étant intégrés directement dans le document.
- **Document MS Excel.** Le rapport est enregistré sous la forme d'une feuille de calcul unique, les graphiques étant intégrés directement dans la feuille de calcul.
- **Document MS PowerPoint.** Chaque étape apparaît sur une nouvelle diapositive. Les graphiques sont intégrés directement dans les diapositives PowerPoint.
- **Objet de sortie.** Quand il est ouvert dans IBM SPSS Modeler, ce fichier (*.cou*) est le même que celui obtenu avec l'option **Sortie à l'écran** dans le groupe **Format de rapport**.

Remarque : pour effectuer une exportation vers un fichier Microsoft Office, l'application correspondante doit être installée.

Titre. Spécifiez le titre du rapport.

Structure du rapport. Sélectionnez **CRISP-DM** ou **Classes**. La vue CRISP-DM fournit un rapport de situation contenant des résumés de la situation globale, ainsi que des détails sur chaque phase de l'exploration de données. La vue Classes est une vue orientée objet qui est davantage appropriée au suivi interne des données et des flux.

Auteur. Le nom d'utilisateur par défaut est affiché, mais vous pouvez le modifier.

Le rapport comprend. Sélectionnez la méthode d'introduction des objets dans le rapport. Sélectionnez **tous les dossiers et objets** pour inclure tous les éléments ajoutés au fichier de projet. Vous pouvez également inclure des éléments en fonction de la sélection ou non de l'option **Inclure dans le rapport** dans les propriétés de l'objet. Pour vérifier les éléments non consignés, vous pouvez également choisir d'inclure uniquement les éléments identifiés comme devant être exclus (l'option **Inclure dans le rapport** n'est pas sélectionnée).

Sélectionner. Cette option vous permet de fournir des mises à jour du projet en sélectionnant uniquement les **éléments récents** du rapport. Vous pouvez également assurer un suivi des questions plus anciennes et non résolues en définissant des paramètres pour les **éléments anciens**. Sélectionnez **tous les éléments** pour exclure le paramètre temps du rapport.

Classé par. Vous pouvez sélectionner une combinaison des caractéristiques d'objet suivantes afin de les classer au sein d'un dossier :

- **Type.** Permet de grouper les objets par type.
- **Nom.** Permet de classer les objets par ordre alphabétique.
- **Date d'ajout.** Permet de trier les objets en fonction de la date d'ajout au projet.

Enregistrement et exportation de rapports générés

Les rapports générés à l'écran sont affichés dans une nouvelle fenêtre de sortie. Les graphiques qu'ils contiennent apparaissent sous la forme d'images en ligne.

Terminologie du rapport

Le nombre total de noeuds dans chaque flux est présenté dans une liste du rapport. Ces nombres se trouvent sous chacun des en-têtes suivants, qui utilisent la terminologie IBM SPSS Modeler et non la terminologie CRISP-DM.

- **Lecteurs de données.** Noeuds source.
- **Rédacteurs de données.** noeuds d'exportation.
- **Concepteurs du modèle.** Concevoir ou modéliser des noeuds.
- **Utilisateurs du modèle.** Modèles générés, également appelés nuggets.
- **Concepteurs de sortie.** Noeuds de graphique ou de sortie.
- **Autre.** Tous les autres noeuds associés au projet. Par exemple, ceux disponibles dans l'onglet Ops sur champs ou l'onglet Ops sur lignes de la palette des noeuds.

Pour enregistrer un rapport :

1. Dans le menu Fichier, cliquez sur **Enregistrer**.
2. Spécifiez un nom de fichier.

   Le rapport est enregistré sous la forme d'un objet de sortie.

   Pour exporter un rapport :
3. Dans le menu Fichier, cliquez sur **Exporter** et indiquez le type de fichier dans lequel effectuer l'exportation.
4. Spécifiez un nom de fichier.

Le rapport est enregistré dans le format choisi.

Vous pouvez effectuer l'exportation vers les types de fichier suivants :

- HTML
- Texte
- Microsoft Word
- Microsoft Excel
- Microsoft PowerPoint

Remarque : pour effectuer une exportation vers un fichier Microsoft Office, l'application correspondante doit être installée.

Utilisez les boutons situés en haut de la fenêtre pour :

- Imprimer le rapport.
- Afficher le rapport au format HTML dans un navigateur Web externe.

Chapitre 14. Personnalisation d'IBM SPSS Modeler

Personnalisation des options d'IBM SPSS Modeler

Vous pouvez effectuer un certain nombre d'opérations pour personnaliser IBM SPSS Modeler en fonction de vos besoins. Cette personnalisation consiste avant tout à définir des options utilisateur spécifiques telles que l'allocation de mémoire, les répertoires par défaut et l'utilisation du son et de la couleur. Vous pouvez aussi personnaliser la palette Noeuds située au bas de la fenêtre IBM SPSS Modeler.

Définition des options d'IBM SPSS Modeler

Vous pouvez personnaliser et définir les options d'IBM SPSS Modeler de plusieurs façons :

- Définissez les options système, telles que l'utilisation de la mémoire et les paramètres régionaux, en cliquant sur **Options système** dans le menu **Outils > Options**.
- Définissez les options utilisateur, telles que les polices et couleurs affichées, en cliquant sur **Options utilisateur** dans le menu **Outils > Options**.
- Indiquez l'emplacement des applications utilisées avec IBM SPSS Modeler en cliquant sur **Programmes externes** dans le menu **Outils > Options**.
- Indiquez les répertoires par défaut utilisés dans IBM SPSS Modeler en cliquant sur **Définir le répertoire** ou **Définir le répertoire du serveur** dans le menu Fichier.

Vous pouvez également définir des options qui s'appliquent à tout ou partie de vos flux. Pour plus d'informations, voir la rubrique «Définition des options pour des flux», à la page 44.

Options du système

Vous pouvez indiquer la langue ou les paramètres régionaux à utiliser avec IBM SPSS Modeler en cliquant sur **Options système** dans le menu **Outils > Options**. Vous pouvez également définir la quantité maximale de mémoire de SPSS Modeler et spécifier la fréquence de sauvegarde automatique des flux. Notez que les changements effectués dans cette boîte de dialogue ne prennent effet qu'après avoir redémarré SPSS Modeler.

Mémoire maximale. Cette option permet d'imposer une limite en mégaoctets à la mémoire utilisée par IBM SPSS Modeler. Sur certaines plates-formes, SPSS Modeler limite la taille de son traitement pour réduire l'impact sur les ordinateurs surchargés ou disposant de ressources limitées. Si vous travaillez avec de grandes quantités de données, une erreur de « mémoire insuffisante » risque d'être générée. Vous pouvez alléger la charge de la mémoire en sélectionnant un nouveau seuil.

Par exemple, une tentative d'affichage d'un arbre de décision très volumineux peut entraîner une erreur mémoire. Dans ce cas, il est recommandé d'accroître la mémoire jusqu'à la valeur maximum, à savoir 4096 Mo. Dans de tels cas où vous traitez probablement de grandes quantités de données, après avoir augmenté l'allocation de mémoire, arrêtez SPSS Modeler et redémarrez-le depuis une ligne de commande pour garantir que le volume de mémoire maximal soit utilisé lors du traitement des données.

Pour son lancement depuis une ligne de commande (en supposant que SPSS Modeler est installé dans l'emplacement par défaut), entrez la commande suivante depuis une invite de commande :

```
C:\Program Files\IBM\SPSS\Modeler\18.2.1\bin\modelerclient.exe" -J-Xss4096M
```

Utiliser les paramètres régionaux du système. Cette option est sélectionnée par défaut et paramétrée sur Anglais (Etats-Unis). Désélectionnez-la pour choisir une autre langue dans la liste des langues et des paramètres régionaux disponibles.

Intervalle d'enregistrement automatique du flux (minutes). Spécifiez la fréquence à laquelle vous souhaitez que SPSS Modeler sauvegarde automatiquement les flux. La valeur maximale est de 60 minutes, la valeur minimale, de 1 minute, et la valeur par défaut, de 5 minutes.

Gestion de la mémoire

Outre le paramètre **Mémoire maximale** indiqué dans la boîte de dialogue Options système, vous pouvez optimiser l'utilisation de la mémoire de plusieurs manières :

- Vous pouvez ajuster l'option **Nombre maximal des membres pour les champs nominaux** de la boîte de dialogue des propriétés du flux. Cette option permet de spécifier le nombre maximal de membres pour les champs nominaux. Au-delà de ce nombre, le niveau de mesure du champ devient *Sans type*. Pour plus d'informations, voir la rubrique «Définition d'options générales pour les flux», à la page 44.
- Vous pouvez forcer IBM SPSS Modeler à libérer de la mémoire en cliquant dans l'angle inférieur droit de la fenêtre, où figurent la mémoire utilisée par IBM SPSS Modeler et la quantité allouée (*xx*Mo / *xx*Mo). Lorsque vous cliquez sur cette zone, elle prend une couleur plus foncée et la quantité de mémoire allouée diminue. La région reprend sa couleur normale lorsque IBM SPSS Modeler a libéré toute la mémoire possible.

Définition des répertoires par défaut

Vous pouvez indiquer le répertoire par défaut des navigateurs de fichiers et de la sortie en sélectionnant **Définir le répertoire** ou **Définir le répertoire du serveur** dans le menu Fichier.

- **Définir le répertoire.** Vous pouvez utiliser cette option pour définir le répertoire de travail. Le répertoire de travail par défaut est basé sur le chemin d'installation de votre version d'IBM SPSS Modeler ou sur le chemin de ligne de commande utilisé pour lancer IBM SPSS Modeler. En mode local, le répertoire de travail est le chemin d'accès utilisé pour toutes les opérations effectuées côté client et pour les fichiers de sortie (s'ils sont référencés avec des chemins d'accès relatifs).
- **Définir le répertoire du serveur.** L'option Définir le répertoire du serveur du menu Fichier est activée lorsqu'il existe une connexion à un serveur distant. Utilisez cette option pour définir le répertoire par défaut de tous les fichiers serveur et fichiers de données indiqués pour l'entrée ou la sortie. Le répertoire du serveur par défaut est *$CLEO/data*, où *$CLEO* est le répertoire dans lequel la version serveur d'IBM SPSS Modeler est installée. Vous pouvez également utiliser la ligne de commande pour remplacer ce répertoire par défaut : utilisez l'indicateur `-server_directory` avec l'argument de ligne de commande `modelerclient`.

Définition des options utilisateur

Vous pouvez définir les options générales d'IBM SPSS Modeler en sélectionnant **Options utilisateur** dans le menu **Outils > Options**. Ces options s'appliquent à tous les flux utilisés dans IBM SPSS Modeler.

Les types d'option suivants peuvent être définis en cliquant sur l'onglet correspondant :

- Options de notification, telles que le remplacement d'un modèle et les messages d'erreur.
- Options d'affichage, telles que les couleurs des graphiques et des arrière-plans.
- Options d'affichage de la couleur de syntaxe.
- Options d'exportation PMML utilisées lors de l'exportation de modèles au format PMML (Predictive Model Markup Language).
- Informations relatives à l'utilisateur et à l'auteur (vos nom, initiales et adresse électronique, par exemple). Ces informations peuvent être affichées dans l'onglet Annotations pour les noeuds et les autres objets créés.
- Permutation entre le mode traditionnel et le mode Analytic Server.

Pour définir des options propres à un flux, telles que des séparateurs décimaux, des formats de date et d'heure, l'optimisation, la présentation du flux et des scripts de flux, utilisez la boîte de dialogue Propriétés du flux, accessible à partir des menus Fichier et Outils.

Définition d'options de notification

L'onglet Notifications de la boîte de dialogue Options utilisateur vous permet de définir différentes options concernant le type et l'occurrence des avertissements et des fenêtres de confirmation dans IBM SPSS Modeler. Vous pouvez également définir le comportement des onglets Sorties et Modèles dans le panneau des gestionnaires lorsque vous générez de nouveaux modèles et sorties.

Afficher la boîte de dialogue des commentaires sur l'exécution de flux Sélectionnez cette option pour afficher une boîte de dialogue comprenant un indicateur de progression lorsque l'exécution du flux est en cours depuis trois secondes. La boîte de dialogue contient également des détails sur les objets de sortie créés par le flux.

- **Fermer la boîte de dialogue une fois terminé** Par défaut, la boîte de dialogue se ferme lorsque l'exécution du flux est terminée. Désélectionnez cette case si vous voulez que la boîte de dialogue reste visible lorsque l'exécution du flux se termine.

Avertir lorsqu'un noeud écrase un fichier Sélectionnez cette option pour être averti par un message d'erreur lorsque des opérations de noeud écrasent un fichier existant.

Avertir lorsqu'un noeud écrase une table de base de données Sélectionnez cette option pour être averti par un message d'erreur lorsque des opérations de noeud écrasent une table de base de données existante.

Notifications sonores

Utilisez la liste pour indiquer si vous souhaitez être averti par un son lorsqu'un événement ou une erreur se produit. Il existe un grand nombre de sons disponibles. Utilisez le bouton Lecture (haut-parleur) pour émettre le son sélectionné. Utilisez le bouton **(...)** pour sélectionner un son.

Remarque : les fichiers `.wav` utilisés pour créer des sons dans IBM SPSS Modeler sont stockés dans le répertoire `/media/sounds` du dossier d'installation.

- **Désactiver tous les sons** Sélectionnez cette option pour désactiver la notification sonore pour tous les événements.

Notifications visuelles

Les options de ce groupe sont utilisées pour indiquer le comportement des onglets Sorties et Modèles du panneau des gestionnaires en haut à droite de l'affichage lorsque vous générez de nouveaux éléments. Sélectionnez **Nouveau modèle** ou **Nouvelle sortie** dans la liste pour indiquer le comportement de l'onglet correspondant.

L'option suivante est disponible pour **Nouveau modèle** :

Remplacer le modèle précédent Si cette option est sélectionnée (par défaut), un modèle existant de ce flux est remplacé dans l'onglet Modèles et sur le canevas du flux. Si cette case est décochée, le modèle est ajouté aux modèles existants sur l'onglet et sur l'espace de travail. Remarque : ce paramètre est remplacé par le paramètre de remplacement du modèle sur un lien de modèle.

L'option suivante est disponible pour **Nouvelle sortie** :

Avertir lorsque les sorties dépassent [n] Indiquez si un avertissement doit être affiché lorsque le nombre d'éléments sur l'onglet Sorties dépasse une quantité prédéfinie. La quantité par défaut est 20 ; toutefois, vous pouvez la modifier au besoin.

Les options suivantes sont disponibles dans tous les cas :

Sélectionner l'onglet Indiquez si vous désirez basculer vers l'onglet Sorties ou vers l'onglet Modèles lorsque l'objet correspondant est généré lors de l'exécution du flux.

- Sélectionnez **Toujours** pour accéder à l'onglet correspondant dans le panneau des gestionnaires.
- Sélectionnez **Si généré par le flux actuel** pour accéder à l'onglet correspondant uniquement pour les objets générés par le flux actuellement visible dans l'espace de travail.
- Sélectionnez **Jamais** pour empêcher que le logiciel accède à l'onglet correspondant afin de vous avertir des sorties ou des modèles générés.

Mettre l'onglet en évidence Indiquez si vous souhaitez mettre en évidence l'onglet Sorties ou Modèles dans le panneau des gestionnaires lorsque de nouvelles sorties ou de nouveaux modèles sont générés.

- Sélectionnez **Si non sélectionné** pour mettre en évidence l'onglet correspondant (s'il n'est pas déjà sélectionné) lorsque de nouveaux objets sont générés dans le panneau des gestionnaires.
- Sélectionnez **Jamais** pour empêcher que le logiciel mette en évidence l'onglet correspondant afin de vous avertir d'objets générés.

Faire défiler la palette pour rendre visible (**Nouveau modèle** uniquement). Indiquez si vous souhaitez faire défiler automatiquement l'onglet Modèles du panneau des gestionnaires pour afficher le modèle le plus récent.

- Sélectionnez **Toujours** pour activer le défilement.
- Sélectionnez **Si généré par le flux actuel** pour activer le défilement uniquement pour les objets générés par le flux actuellement visible dans l'espace de travail.
- Sélectionnez **Jamais** pour empêcher que le logiciel n'active le défilement automatique de l'onglet Modèles.

Ouvrir une fenêtre (**Nouvelle sortie** uniquement). Sélectionnez s'il faut ouvrir automatiquement une fenêtre de sortie lors de la génération.

- Sélectionnez **Toujours** pour qu'une nouvelle fenêtre de sortie soit automatiquement ouverte.
- Sélectionnez **Si généré par le flux actuel** pour ouvrir une nouvelle fenêtre pour la sortie générée par le flux actuellement visible dans l'espace de travail.
- Sélectionnez **Jamais** pour empêcher que le logiciel ouvre automatiquement une nouvelle fenêtre pour la sortie générée.

Pour rétablir les paramètres système par défaut pour cet onglet, cliquez sur **Valeurs par défaut**.

Définition des options d'affichage

L'onglet Afficher de la boîte de dialogue Options utilisateur vous permet de définir les options d'affichage des polices et des couleurs dans IBM SPSS Modeler.

Afficher la boîte de dialogue de bienvenue au démarrage. Sélectionnez cette option pour que la boîte de dialogue de bienvenue s'affiche au démarrage. Cette boîte de dialogue contient des options permettant de lancer le tutoriel d'exemples de l'application, d'ouvrir un flux de démonstration ou un flux ou un projet existant, ou encore pour créer un nouveau flux.

Afficher les balisages de flux et de super noeud. Si cette option est sélectionnée, le balisage (s'il en existe) des flux et des super noeuds est affiché par défaut. Le balisage inclut les commentaires de flux, les liens de modèle et la mise en évidence des branches de scoring.

Polices & couleurs standard (activées après redémarrage). Les options de cette boîte de dialogue permettent de définir le dessin de l'écran d'IBM SPSS Modeler, le modèle de couleur ainsi que la taille des polices affichées. Les options que vous sélectionnez ici ne sont appliquées qu'après la fermeture et le redémarrage d'IBM SPSS Modeler.

- **Aspect.** Sélectionnez un schéma de couleurs et une conception d'écran standard. Vous pouvez choisir entre :

  - **SPSS Standard**, la conception par défaut.
  - **SPSS Classic**, une conception familière pour les utilisateurs des versions précédentes d'SPSS Modeler.
  - **Windows**, conception Windows permettant d'obtenir un contraste plus élevé dans le canevas de flux et les palettes.
  - **Analytics Carbon**, conception moderne avec des icônes et des couleurs élégantes.
- **Taille de police par défaut des noeuds.** Indiquez la taille de police à utiliser dans les palettes de noeuds et pour les noeuds affichés dans l'espace de travail de flux.
- **Spécifiez la police de largeur fixe.** Cochez cette case pour sélectionner une police de largeur fixe et la **taille** de police associée à utiliser dans les contrôles de scripts et d'expressions CLEM. Par défaut, il s'agit de `Monospace plain` ; cliquez sur **Changer...** pour afficher une liste des autres polices que vous pouvez sélectionner.

Remarque : vous pouvez définir la taille des icônes des noeuds d'un flux sur le panneau Présentation de l'onglet Options dans la boîte de dialogue des propriétés du flux. Dans le menu principal, sélectionnez **Outils** > **Propriétés du flux** > **Options** > **Présentation**.

Couleurs personnalisées. Ce tableau répertorie les couleurs actuellement sélectionnées et utilisées pour les différents éléments d'affichage. Vous pouvez modifier la couleur en cours de chaque élément répertorié dans la liste en double-cliquant sur la ligne correspondante dans la colonne **Couleur** et en sélectionnant la couleur voulue dans la liste. Pour définir une couleur personnalisée, faites défiler la liste jusqu'en bas et cliquez sur l'entrée **Couleur...**.

Ordre des couleurs des graphiques. Ce tableau répertorie les couleurs actuellement sélectionnées et utilisées dans de nouveaux graphiques. L'ordre des couleurs reflète celui dans lequel elles sont utilisées dans le graphique. Par exemple, si un champ nominal utilisé comme superposition de couleurs contient quatre valeurs uniques, seules les quatre premières couleurs répertoriées ici seront utilisées. Vous pouvez modifier la couleur en cours de chaque élément répertorié dans la liste en double-cliquant sur la ligne correspondante dans la colonne **Couleur** et en sélectionnant la couleur voulue dans la liste. Pour définir une couleur personnalisée, faites défiler la liste jusqu'en bas et cliquez sur l'entrée **Couleur...**. Les changements apportés ici n'affectent pas les graphiques créés auparavant.

Pour rétablir les paramètres système par défaut pour cet onglet, cliquez sur **Valeurs par défaut**.

Définition des options d'affichage de syntaxe

L'onglet Syntaxe de la boîte de dialogue Options utilisateur vous permet de définir les options des attributs de police et des couleurs d'affichage dans les scripts que vous créez dans IBM SPSS Modeler.

Mise en évidence de syntaxe. Ce tableau répertorie les couleurs actuellement sélectionnées et utilisées pour les différents éléments de syntaxe, notamment la police et la fenêtre dans lesquelles ils sont affichés. Vous pouvez modifier la couleur en cours de chaque élément répertorié dans la liste en cliquant sur la ligne correspondante et en sélectionnant une couleur de la liste. En outre, pour les éléments de police, vous pouvez choisir d'ajouter une mise en évidence en gras et en italique.

Aperçu. Ce tableau comporte un exemple de syntaxe affichée qui utilise les couleurs et les attributs de police que vous sélectionnez dans le tableau de **mise en évidence de la syntaxe**. Cet aperçu se met à jour dès que vous modifiez une sélection.

Cliquez sur **Valeurs par défaut** pour revenir aux paramètres système par défaut de cet onglet.

Définition des options d'exportation PMML

Dans l'onglet PMML, vous pouvez déterminer la manière dont IBM SPSS Modeler exporte des modèles au format PMML. Pour plus d'informations, voir la rubrique «Importation et exportation de modèles au format PMML», à la page 212.

Exporter PMML. Vous pouvez configurer ici les variations du langage PMML les plus adaptées à votre application cible.

- Sélectionnez l'option **avec les extensions** pour autoriser les extensions PMML dans les cas de figure spéciaux où aucun équivalent PMML standard n'existe. Dans la plupart des cas, cela produira le même résultat qu'avec le langage PMML standard.
- Sélectionnez **PMML standard...** pour exporter les données PMML qui se conforment au maximum à la norme PMML.

Options PMML standard. Lorsque vous sélectionnez l'option **PMML standard...**, vous avez le choix entre deux méthodes valides pour exporter des modèles de régression linéaire et logistique :

- En tant que modèles **PMML <GeneralRegression>**
- En tant que modèles **PMML <Regression>**

Pour plus d'informations sur le langage PMML, reportez-vous au site Web de l'exploration de données group (http://www.dmg.org).

Définition des informations utilisateur

Informations sur l'utilisateur/auteur. Les informations que vous saisissez ici peuvent être affichées dans l'onglet Annotations des noeuds et des autres objets que vous créez.

Définition du mode

Paramètres du mode de modélisateur. Dans l'onglet **Mode**, vous pouvez choisir l'un des modes suivants :

- **Mode SPSS Modeler traditionnel** : affiche tous les noeuds et expressions disponibles dans l'interface utilisateur.
- **Mode Analytic Server** : affiche uniquement les noeuds et les expressions pris en charge par Analytic Server. Notez toutefois que certains noeuds et certaines expressions CLEM restent affichés bien qu'ils ne soient pas *entièrement* pris en charge par Analytic Server. Le tableau ci-après fournit des informations générales sur les noeuds pris en charge, partiellement pris en charge et non pris en charge par Analytic Server.

Pour plus d'informations, voir Noeuds pris en charge.

Pour plus d'informations sur Analytic Server, voir la documentation d'Analytic Server.

Si vous configurez SPSS Modeler pour afficher des noeuds de base de données dans la palette de modélisation de base de données, ils ne seront pas impactés lors d'un changement de modes. Les noeuds de base de données seront toujours affichés. Si vous utilisez l'intégration d'IBM Db2 for z/OS ou d'IBM Netezza, dans certains cas, ces noeuds peuvent disparaître de la palette de modélisation de base de données après le passage au mode Analytic Server. Si c'est le cas, sélectionnez **Outils** > **Options** > **Applications externes** et réinitialisez les cases à cocher.

Tableau 38. Prise en charge des noeuds

Type de noeud (nom de la palette)	Pris en charge par Analytic Server	Partiellement pris en charge par Analytic Server	Non pris en charge par Analytic Server
Sources	• Noeud source Analytic Server		• Base de données • Fichier variable • Fichier fixe • Fichier de statistiques • Collection de données • IBM Cognos • Importation TWC • Importation TM1 • Fichier SAS • Excel • Entrée utilisateur XML • Génération de simulation • Vue de données • Géospatial • Stockage d'objets • Importation d'extension • Importation R • SNA (Analyse de diffusion et Analyse de groupe)
Ops sur lignes	• Sélectionner • Trier • Equilibre • Distinct • Agréger RFM • Ajouter • Streaming TS • Transformation d'extension • Streaming TCM	• Echantillon (ne prend en charge Random% que pour la méthode simple. La méthode complexe n'est pas prise en charge.) • Fusionner (ne prend en charge que les clés et conditions de jointure) • Agréger (le 1er quantile, le 3e quantile et la médiane ne sont pas pris en charge par Analytic Server)	• Boîtes espace-temps • Optimisation CPLEX

Tableau 38. Prise en charge des noeuds (suite)

Type de noeud (nom de la palette)	Pris en charge par Analytic Server	Partiellement pris en charge par Analytic Server	Non pris en charge par Analytic Server
Ops sur champs	• Type • Filtrer • Dériver • Remplacer • Rechiffrer • Ensemble • Binariser • Restructurer • Re-trier • Reprojeter • Intervalles de temps	• Prép. auto. des données (ne prend en charge que les transformations) • Discrétiser (la méthode de regroupement par casiers `equalFreq` n'est pas prise en charge si le paramètre Ex-aequo **Conserver dans l'élément actuel** est sélectionné) • Analyse RFM (la méthode de regroupement par casiers `Tiles` n'est pas prise en charge si le paramètre Ex-aequo **Conserver dans l'élément actuel** est sélectionné) • Partition (non pris en charge par Analytic Server à moins qu'un champ unique ne soit utilisé pour affecter les lignes aux partitions de manière répétée)	• Anonymiser • Historique • Transposition
Graphiques	• Nuage • Courbes • Tracé horaire • Distribution • Histogramme • Collection • Relations • Evaluation • Visualisation de carte • Tracé E (Bêta) • t-SNE	• Représentation graphique (ne prend en charge la fonction de mode d'agrégation que pour les champs dont le niveau de mesure est discret, nominal , ordinal ou indicateur)	

Tableau 38. Prise en charge des noeuds (suite)

Type de noeud (nom de la palette)	Pris en charge par Analytic Server	Partiellement pris en charge par Analytic Server	Non pris en charge par Analytic Server
Modélisation	• Série temporelle • TCM • Isotonique sur AS • Random Trees • Tree-AS • Linear-AS • GLE • LSVM • STP • TwoStep-AS • Règles d'association • XGBoost-AS • K moyenne AS	• Discriminant automatique • Numérisation automatique (les deux noeuds ne prennent en charge que la scission et un champ avec un rôle de scission doit être fourni si l'option **Exécuter sur Analytic Server (divisions activées)** de Discriminant automatique est utilisée) • Extension (le modèle de génération de syntaxe R n'est pas pris en charge par Analytic Server) • Les noeuds suivants ne prennent en charge que la scission et PSM : – Arbre C&RT – Linéaire – Réseau de neurones – CHAID – Quest	• Classification non supervisée automatique • Liste de décisions • C5.0 • Régression • ACP/Facteur • Sélection de fonction • Discriminant • Logistique • LinéaireGénéralisé • GLMM • Réseau Bayésien • Apriori • Carma • Séquence • k moyenne • Kohonen • TwoStep • Anomalie • KNN • R • Forêt aléatoire • Les noeuds suivants ont asl, mais uniquement en lecture/écriture : – Cox – SVM – SLRM
Sortie	• Table • Matrice • Analyse • Audit des données • Transformation • Statistiques • Moyennes • Rapport • Valeurs globales	• Sortie d'extension • Sortie de syntaxe R	• Evaluation de simulation • Ajustement de simulation • Sortie R

Tableau 38. Prise en charge des noeuds (suite)

Type de noeud (nom de la palette)	Pris en charge par Analytic Server	Partiellement pris en charge par Analytic Server	Non pris en charge par Analytic Server
Exportation	• Noeud d'exportation Analytic Server	• Exportation d'extension • Exportation de syntaxe R	• Base de données • Fichier à plat • Export Statistics • Collection de données • Excel • Exportation IBM Cognos • Exportation TM1 • SAS • Exportation XML • Stockage d'objets • Exportation R
IBM SPSS Statistics			• Fichier de statistiques • Transformation Statistics • Modèle Statistics • Sortie Statistics • Export Statistics
IBM SPSS Text Analytics	• Analyse des liens du texte • Exploration de texte • Noeud Langue		• Liste de fichiers • Fil de nouvelles • Analyse des liens du texte • Convertir • Exploration de texte • Visualiseur de fichiers
Python			• SMOTE • SVM à classe unique • XGBoost Tree • XGBoost Linear • t-SNE • Forêt aléatoire • HDBSCAN
Spark	Tous pris en charge		

Personnalisation de la palette Noeuds

Les flux sont créés à l'aide de noeuds. La palette Noeuds située au bas de la fenêtre IBM SPSS Modeler contient tous les noeuds pouvant être utilisés pour la création des flux. Pour plus d'informations, voir la rubrique «Palette de noeuds», à la page 15.

Vous pouvez réorganiser la palette Noeuds de deux façons :

- Personnaliser le gestionnaire de palette. Pour plus d'informations, voir la rubrique «Personnalisation du gestionnaire de palette.», à la page 233.

- Modifier l'affichage des onglets de la palette contenant des sous-palettes dans la palette Noeuds. Pour plus d'informations, voir la rubrique «Création d'une sous-palette», à la page 235.

Personnalisation du gestionnaire de palette.

Le gestionnaire de palette peut être personnalisé en fonction de votre utilisation d'IBM SPSS Modeler. Par exemple, si vous analysez fréquemment des données de séries temporelles à partir d'une base de données, vous souhaiterez peut-être vous assurer que le noeud source Base de données, le noeud des intervalles temporels, le noeud Séries temporelles et le noeud du graphique Tracé horaire soient disponibles ensemble à partir d'un seul onglet de la palette. Le gestionnaire de palette vous permet d'apporter facilement ces modifications en créant vos onglets de palette personnalisés dans la palette de noeuds.

Le gestionnaire de palette vous permet d'effectuer différentes tâches :

- Contrôler quels onglets de la palette apparaissent dans la palette de noeuds sous l'espace de travail du flux.
- Modifier l'ordre d'affichage des onglets de la palette dans la palette de noeuds.
- Créer et modifier vos propres onglets de palette et des éventuelles sous-palettes associées.
- Editer les sélections de noeud par défaut dans vos onglets.

Pour accéder au gestionnaire de palette, cliquez sur **Favoris** dans le menu Outils.

Nom de palette. Chaque onglet de palette disponible est répertorié, qu'il apparaisse ou non dans la palette de noeuds. Cela comprend les éventuels onglets de palette que vous avez créés. Pour plus d'informations, voir la rubrique «Création d'un onglet Palette».

Nbre de noeuds. Le nombre de noeuds affichés dans chaque onglet de la palette. Un nombre élevé ici signifie qu'il vous sera peut-être plus pratique de créer des sous-palettes afin de diviser les noeuds dans l'onglet. Pour plus d'informations, voir la rubrique «Création d'une sous-palette», à la page 235.

Affiché ? Sélectionnez ce champ pour afficher l'onglet de palette dans la Palette de noeuds. Pour plus d'informations, voir la rubrique «Affichage des onglets Palette dans la palette de noeuds», à la page 234.

Sous-palettes. Pour sélectionner des sous-palettes à afficher dans un onglet de palette, mettez en surbrillance le **Nom de la palette** requis et cliquez sur ce bouton pour afficher la boîte de dialogue Sous-palettes. Pour plus d'informations, voir la rubrique «Création d'une sous-palette», à la page 235.

Restaurer les paramètres par défaut. Pour supprimer complètement tous les changements et les ajouts apportés aux palettes et aux sous-palettes et revenir aux paramètres de palette par défaut, cliquez sur ce bouton.

Création d'un onglet Palette

Pour créer un onglet de palette personnalisé :

1. Dans le menu Outils, ouvrez le gestionnaire de la palette.
2. A droite de la colonne *Affiché ?*, cliquez sur le bouton Ajouter une palette ; la boîte de dialogue Créer/Editer une palette s'affiche.
3. Saisissez un **Nom de la palette** unique.
4. Dans la zone **Noeuds disponibles**, sélectionnez le noeud à ajouter à l'onglet palette.
5. Cliquez sur la flèche droite Ajouter un noeud pour déplacer un noeud mis en évidence dans la zone **Noeuds sélectionnés**. Répétez l'opération jusqu'à ce que vous ayez ajouté tous les noeuds de votre choix.

   Une fois que vous avez ajouté tous les noeuds requis, vous pouvez modifier l'ordre de leur apparition dans l'onglet palette :

6. Utilisez les boutons représentant une simple flèche pour déplacer un noeud d'une ligne vers le haut ou vers le bas.
7. Utilisez les boutons représentant une flèche et une ligne pour placer un noeud tout en bas ou tout en haut de la liste.
8. Pour supprimer un noeud d'une palette, mettez le noeud en surbrillance et cliquez sur le bouton Supprimer à droite de la zone **Noeuds sélectionnés**.

Affichage des onglets Palette dans la palette de noeuds

Il peut exister dans IBM SPSS Modeler des options que vous n'utilisez jamais ; dans ce cas, vous pouvez utiliser le gestionnaire de palette pour masquer les onglets contenant ces noeuds.

Pour sélectionner les onglets à afficher dans la palette de noeuds :

1. Dans le menu Outils, ouvrez le gestionnaire de la palette.
2. A l'aide des cases à cocher figurant dans la colonne *Affiché ?*, choisissez s'il faut inclure ou masquer chaque onglet de la palette.

Pour supprimer définitivement un onglet de palette de la palette de noeuds, mettez le noeud en évidence et cliquez sur le bouton Supprimer à droite de la colonne *Affiché ?*. Une fois supprimé, un onglet de palette ne peut pas être récupéré.

Remarque : vous ne pouvez pas supprimer les onglets par défaut de la palette fournis avec IBM SPSS Modeler, à l'exception de l'onglet Favoris.

Modification de l'ordre d'affichage dans la palette de noeuds

Une fois que vous avez sélectionné les onglets de la palette à afficher, vous pouvez modifier leur ordre d'affichage dans la palette de noeuds :

1. Utilisez les boutons représentant une simple flèche pour déplacer un noeud de la palette d'une ligne vers le haut ou vers le bas. Un déplacement vers le haut les déplace vers la gauche de la palette de noeuds, et vice versa.
2. Utilisez les boutons représentant une flèche et une ligne pour placer un onglet de la palette tout en bas ou tout en haut de la liste. Ceux qui apparaissent en haut de la liste figureront à gauche de la palette de noeuds.

Affichage des sous-palettes sur un onglet Palette

De même que vous pouvez contrôler les onglets de palette qui s'affichent dans la palette de noeuds, vous pouvez contrôler les sous-palettes disponibles à partir de leur onglet de palette parent.

Pour sélectionner les sous-palettes à afficher dans un onglet palette :

1. Dans le menu Outils, ouvrez le gestionnaire de la palette.
2. Sélectionnez la palette de votre choix.
3. Cliquez sur le bouton Sous-palettes ; la boîte de dialogue Sous-palettes s'affiche.
4. A l'aide des cases à cocher figurant dans la colonne *Affiché ?*, choisissez s'il faut inclure chaque sous-palette dans l'onglet palette. La sous-palette **Tous** est toujours affichée et ne peut pas être supprimée.
5. Pour supprimer définitivement une sous-palette de l'onglet de palette, mettez la sous-palette en évidence et cliquez sur le bouton Supprimer à droite de la colonne *Affiché ?*.

Remarque : vous ne pouvez pas supprimer les sous-palettes par défaut fournies avec l'onglet Modélisation de la palette.

Modification de l'ordre d'affichage dans l'onglet Palette

Une fois que vous avez sélectionné les onglets de la palette à afficher, vous pouvez modifier leur ordre d'affichage dans l'onglet de la palette parent :

1. Utilisez les boutons représentant une simple flèche pour déplacer une sous-palette d'une ligne vers le haut ou vers le bas.
2. Utilisez les boutons représentant une flèche et une ligne pour placer une sous-palette tout en bas ou tout en haut de la liste.

Les sous-palettes que vous créez sont affichées dans la palette de noeuds lorsque vous sélectionnez l'onglet palette parent. Pour plus d'informations, voir la rubrique «Modification d'une vue de l'onglet Palette».

Création d'une sous-palette

Comme vous pouvez ajouter tout noeud existant aux onglets de palette personnalisés que vous créez, vous sélectionnerez peut-être davantage de noeuds qu'on ne peut en afficher facilement à l'écran sans défilement. Pour éviter d'avoir à défiler, vous pouvez créer des sous-palettes dans lesquelles vous placerez les noeuds choisis pour l'onglet palette. Par exemple, si vous avez créé un onglet de palette contenant les noeuds que vous utilisez le plus fréquemment pour créer vos flux, vous pourriez créer quatre sous-palettes divisant les sélections par noeud source, opérations de champ, modélisation et sortie.

Remarque : Vous ne pouvez sélectionnez la sous-palette que parmi celles qui sont ajoutées à l'onglet palette parent.

Pour créer une sous-palette :

1. Dans le menu Outils, ouvrez le gestionnaire de la palette.
2. Sélectionnez la palette à laquelle vous souhaitez ajouter les sous-palettes.
3. Cliquez sur le bouton Sous-palettes ; la boîte de dialogue Sous-palettes s'affiche.
4. A droite de la colonne *Affiché ?*, cliquez sur le bouton Ajouter une sous-palette ; la boîte de dialogue Créer/Editer une sous-palette s'affiche.
5. Saisissez un **nom de sous-palette** unique.
6. Dans la zone **Noeuds disponibles**, sélectionnez le noeud à ajouter à l'onglet sous-palette.
7. Cliquez sur la flèche droite Ajouter un noeud pour déplacer un noeud sélectionné dans la zone **Noeuds sélectionnés**.
8. Lorsque vous avez ajouté les noeuds requis, cliquez sur **OK** pour revenir à la boîte de dialogue Sous-palettes.

Les sous-palettes que vous créez sont affichées dans la palette de noeuds lorsque vous sélectionnez l'onglet palette parent. Pour plus d'informations, voir la rubrique «Modification d'une vue de l'onglet Palette».

Modification d'une vue de l'onglet Palette

En raison du grand nombre de noeuds disponibles dans IBM SPSS Modeler, ils peuvent ne pas être tous visibles dans les petits écrans sans défilement vers la gauche ou vers la droite de la palette de noeuds ; cela est particulièrement remarquable dans l'onglet de la palette Modélisation. Pour réduire les besoins de défilement, vous pouvez choisir d'afficher uniquement les noeuds figurant dans une sous-palette (le cas échéant). Pour plus d'informations, voir la rubrique «Création d'une sous-palette».

Pour modifier les noeuds affichés dans l'onglet palette ; sélectionnez l'onglet palette puis, dans le menu de gauche, choisissez d'afficher soit tous les noeuds, soit seulement les noeuds figurant dans une sous-palette spécifique.

Chapitre 15. Remarques relatives aux performances des flux et des noeuds

Vous pouvez concevoir vos flux pour maximiser les performances en organisant les noeuds selon la configuration la plus efficace, en activant les caches de noeud le cas échéant et en tenant compte des autres critères décrits dans cette section.

En plus des critères décrits ici, des améliorations supplémentaires et plus importantes des performances peuvent généralement être obtenues en utilisant de manière efficace votre base de données, en particulier à travers l'optimisation SQL.

Ordre des noeuds

Même si vous n'utilisez pas l'optimisation SQL, l'ordre des noeuds d'un flux peut avoir une incidence sur les performances. L'objectif principal consiste à réduire au maximum le traitement en aval ; ainsi, lorsque vous disposez de noeuds qui permettent de réduire la quantité de données, placez-les au début du flux. IBM SPSS Modeler Server peut appliquer automatiquement plusieurs règles de réorganisation au cours de la compilation et faire avancer certains noeuds lorsque cette opération s'avère sûre. (Cette fonctionnalité est activée par défaut. Assurez-vous auprès de l'administrateur système que c'est le cas pour votre installation.)

Lorsque vous utilisez l'optimisation SQL, vous souhaitez augmenter au maximum sa disponibilité et son efficacité. L'optimisation s'arrête lorsque le flux contient une opération qui ne peut pas être exécutée dans la base de données ; par conséquent, il s'avère plus judicieux de regrouper les opérations optimisées SQL au début du flux. Grâce à cette méthode, une plus grande partie du traitement est confinée dans la base de données et une moindre quantité de données est ainsi transmise dans IBM SPSS Modeler.

Les opérations ci-dessous peuvent être réalisées dans la plupart des bases de données. Essayez de les regrouper au *début* du flux :

- Fusion par clé (jointure)
- Sélectionner
- Agréger
- Trier
- Echantillonner
- Ajouter
- Opérations de distinction en mode *inclure* (tous les champs sont sélectionnés)
- Opérations de remplacement
- Opérations de calcul de base qui utilisent une manipulation arithmétique ou une manipulation de chaîne standard (en fonction des opérations prises en charge par la base de données)
- Binarisation

Les opérations ci-dessous ne peuvent pas être réalisées dans la plupart des bases de données. Elles doivent donc être placées dans le flux *après* celles mentionnées dans la liste précédente :

- Opérations effectuées sur des données non incluses dans une base de données (les fichiers plats, par exemple)
- Fusion par ordre
- Equilibre
- Opérations de distinction en mode *supprimer* (seul un sous-ensemble de champs est sélectionné comme étant distinct)

© Copyright IBM Corp. 1994, 2019

- Toute opération qui requiert un accès aux données à partir d'enregistrements autres que celui en cours de traitement
- Calculs de champ (état et comptage)
- Opérations relatives au noeud Historiser
- Opérations impliquant des fonctions "@" (séries temporelles)
- Modes de vérification du type *Avertir* et *Stopper*
- Construction, application et analyse de modèle

  Remarque : les modèles d'arbre décision, d'ensemble de règles et de régression linéaire, ainsi que ceux générés par facteur, peuvent tous générer du code SQL et donc être renvoyés vers la base de données.
- Sortie de données vers tout emplacement autre que la base de données qui traite les données

Caches de noeud

Pour optimiser l'exécution du flux, vous pouvez définir un *cache* sur n'importe quel nœud non terminal. Lorsque vous définissez un cache sur un noeud, le cache est alimenté par les données passant par le noeud lors de l'exécution suivante du flux de données. Les données sont alors lues à partir du cache (stocké sur le disque dans un répertoire temporaire) et non à partir de la source de données.

La mise en cache est plus utile après une opération de longue durée telle que le tri, la fusion ou l'agrégation. Par exemple, supposez que vous ayez un noeud source configuré pour la lecture des données relatives aux ventes à partir de la base de données et un noeud Agréger qui fournit des résultats de ventes par zone géographique. Vous pouvez définir un cache sur le noeud Agréger plutôt que sur le noeud source pour que le cache stocke uniquement les données agrégées et non le jeu de données intégral.

Remarque : La mise en cache au niveau des noeuds source, qui stocke simplement une copie des données originales telle qu'elle est lue dans IBM SPSS Modeler, n'améliorera pas les performances dans la plupart des cas.

Lorsque la mise en cache d'un noeud est activée, une petite icône en forme de document apparaît dans l'angle supérieur droit. Une fois les données mises en cache au niveau du noeud, l'icône de document devient verte.

Pour activer un cache

1. Dans l'espace de travail de flux, cliquez avec le bouton droit de la souris sur le noeud, puis cliquez sur **Cache** dans le menu.
2. Dans le sous-menu de mise en cache, cliquez sur **Activer**.
3. Vous pouvez désactiver le cache en cliquant sur le noeud avec le bouton droit de la souris et en cliquant sur **Désactiver** dans le sous-menu de mise en cache.

Mise en mémoire cache des noeuds dans une base de données

Dans le cas des flux exécutés dans une base de données, les données peuvent être mises en mémoire cache en milieu de flux dans une table temporaire de la base de données, plutôt que dans le système de fichiers. Vous pouvez combiner cela à la fonction d'optimisation SQL et accroître ainsi les performances de manière significative. Par exemple, la sortie d'un flux qui fusionne plusieurs tables afin de créer une vue d'exploration de données peut être mise en mémoire cache et réutilisée selon les besoins. La génération automatique du code SQL pour tous les noeuds en aval peut améliorer encore les performances.

Pour tirer profit de la mise en mémoire cache de base de données, vous devez activer les fonctions d'optimisation SQL et de mise en mémoire cache de base de données. Les paramètres d'optimisation du serveur annulent ceux définis pour le client. Pour plus d'informations, voir la rubrique «Définition des options d'optimisation pour les flux», à la page 47.

Une fois la mise en mémoire cache de base de données activée, il vous suffit de cliquer à l'aide du bouton droit de la souris sur un noeud non terminal afin de mettre les données en mémoire cache à ce stade du flux ; le cache est automatiquement créé dans la base de données lors de l'exécution suivante du flux. Si la mise en mémoire cache de base de données ou l'optimisation SQL n'est pas activée, le cache est écrit dans le système de fichiers.

Remarque : Les bases de données suivantes prennent en charge les tables temporaires à des fins de mise en mémoire cache : Db2, Oracle, SQL Server et Teradata. D'autres bases de données, comme Netezza, utilisent une table standard pour la mise en mémoire cache de base de données. Si nécessaire, le code SQL peut être personnalisé pour des bases de données spécifiques : contactez l'assistance technique.

Performances: noeuds d'exécution

Trier. Le noeud Trier doit lire le jeu de données d'entrée dans son intégralité pour pouvoir le trier. Les données sont stockées en mémoire jusqu'à une certaine limite, et l'excédent de données est placé sur le disque. L'algorithme de tri consiste en un algorithme combinatoire : les données sont lues en mémoire jusqu'à ce que la limite soit atteinte, puis elles sont triées à l'aide d'un algorithme hybride de tri rapide. Si toutes les données tiennent en mémoire, le tri est terminé. Sinon, un algorithme de fusion/tri est appliqué. Les données triées sont écrites dans un fichier et le fragment de données suivant est lu en mémoire, trié, puis écrit sur le disque. Cette opération est répétée jusqu'à ce que toutes les données soient lues, suite à quoi les fragments triés sont fusionnés. L'opération de fusion peut nécessiter plusieurs passages sur les données stockées sur le disque. Lors d'une utilisation intensive du système, le noeud Trier place deux copies complètes des données sur le disque : l'une triée et l'autre non.

Le temps d'exécution global de l'algorithme est de l'ordre de `N*log(N)`, où *N* correspond au nombre d'enregistrements. Le tri en mémoire s'avère plus rapide que la fusion à partir du disque ; vous pouvez, par conséquent, réduire le temps d'exécution réel en allouant davantage de mémoire à l'opération de tri. L'algorithme s'alloue une partie de la mémoire RAM physique contrôlée par l'option de configuration *Multiplicateur d'utilisation de la mémoire* d'IBM SPSS Modeler Server. Pour accroître la quantité de mémoire utilisée pour le tri, fournissez davantage de mémoire RAM physique ou augmentez cette valeur. Lorsque la quantité de mémoire utilisée dépasse l'ensemble de travail du processus de telle sorte qu'une partie de la mémoire est paginée sur le disque, les performances sont réduites ; en effet, le modèle d'accès mémoire de l'algorithme de tri en mémoire est aléatoire et peut entraîner une pagination excessive. L'algorithme de tri est utilisé par plusieurs autres noeuds que Trier, mais les mêmes règles de performances s'appliquent.

Discrétiser. Le noeud Discrétiser lit le jeu de données d'entrée dans son intégralité afin de calculer les limites d'intervalle et d'allouer des enregistrements à ces intervalles. Le jeu de données est mis en mémoire cache pendant le calcul des limites ; il est ensuite analysé à nouveau pour procéder à l'allocation. Lorsque vous choisissez la méthode de regroupement par casiers *Largeur fixe* ou *Moyenne et écart-type*, le jeu de données est mis en mémoire cache directement sur le disque. Ces méthodes présentent un temps d'exécution linéaire et requièrent suffisamment d'espace disque pour stocker l'intégralité du jeu de données. Si vous utilisez la méthode de regroupement par casiers *Rangs* ou *Quantiles*, le jeu de données est trié à l'aide de l'algorithme de tri décrit précédemment, puis le jeu de données trié est utilisé comme cache. Le tri confère à ces méthodes un temps d'exécution de `M*N*log(N)`, où *M* correspond au nombre de champs mis en intervalles et *N* au nombre d'enregistrements ; cette opération nécessite une quantité d'espace disque équivalant à deux fois la taille du jeu de données.

La génération d'un noeud Calculer basée sur des intervalles générés permet d'améliorer les performances lors des passages suivants. Les opérations de calcul sont bien plus rapides que la création d'intervalles.

Fusion par clé (jointure). Le noeud Fusionner, lorsque vous utilisez la méthode *Clés* (qui équivaut à une jointure de base de données), trie chaque jeu de données d'entrée en fonction des champs-clés. Cette partie de la procédure présente un temps d'exécution de `M*N*log(N)`, où *M* correspond au nombre d'entrées et *N* au nombre d'enregistrements contenus dans le jeu de données d'entrée le plus volumineux ; elle nécessite une quantité suffisante d'espace disque pour pouvoir stocker tous les jeux de données d'entrée plus une deuxième copie du jeu de données le plus volumineux. Le temps d'exécution de la fusion à proprement parler est proportionnel à la taille du jeu de données de sortie, qui dépend de la fréquence des clés correspondantes. Dans le cas le plus extrême, où la sortie équivaut au produit cartésien des entrées, le temps d'exécution peut avoisiner `NM`. Ce cas est rare ; la plupart des jointures comportent bien moins de clés correspondantes. Si un jeu de données est sensiblement plus volumineux que les autres ou si les données entrantes sont déjà triées en fonction d'un champ-clé, vous pouvez améliorer les performances du noeud par le biais de l'onglet Optimisation.

Agréger. Lorsque l'option *Les clés sont adjacentes* n'est pas définie, ce noeud lit (mais ne stocke pas) l'intégralité du jeu de données d'entrée avant de générer une sortie agrégée. Dans les cas les plus extrêmes, où la taille des données agrégées atteint une limite (déterminée par l'option de configuration *Multiplicateur d'utilisation de la mémoire* de IBM SPSS Modeler Server), le reste du jeu de données est trié et traité comme si l'option *Les clés sont adjacentes* était activée. Lorsque cette option est définie, aucune donnée n'est stockée car les enregistrements de sortie agrégés sont générés au fur et à mesure de la lecture des données d'entrée.

Distinguer. Le noeud Distinguer stocke tous les champs-clés uniques dans le jeu de données d'entrée ; dans les cas où tous les champs sont des champs-clés et tous les enregistrements sont uniques, il stocke le jeu de données complet. Par défaut, le noeud Distinguer trie les données dans les champs-clés puis sélectionne (ou supprime) le premier enregistrement de chaque groupe. Pour les jeux de données d'une taille plus restreinte avec un faible nombre de clés distinctes ou ceux qui ont été prétriés, vous pouvez choisir des options pour améliorer la vitesse et l'efficacité du traitement.

Type. Dans certaines instances, le noeud Typer met en mémoire cache les données d'entrée lors de la lecture des valeurs ; le cache est utilisé pour les traitements en aval. Un cache requiert une quantité suffisante d'espace disque pour pouvoir stocker l'intégralité du jeu de données, mais il accélère le traitement.

Evaluation. Le noeud Evaluation doit trier les données d'entrée pour pouvoir calculer les quantiles. L'opération de tri est répétée pour chaque modèle évalué car les scores et l'ordre des enregistrements qui en découle diffèrent dans chaque cas. Le temps d'exécution équivaut à `M*N*log(N)`, où *M* correspond au nombre de modèles et *N* au nombre d'enregistrements.

Performances : noeuds de modélisation

Réseau de neurones et Kohonen. Les algorithmes d'apprentissage du réseau de neurones (y compris l'algorithme de Kohonen) effectuent plusieurs passages sur les données d'apprentissage. Les données sont stockées en mémoire jusqu'à une certaine limite, et l'excédent de données est placé sur le disque. Il n'est pas judicieux d'accéder aux données d'apprentissage à partir du disque ; en effet, la méthode d'accès est aléatoire et peut donc entraîner une activité excessive sur le disque. Vous pouvez désactiver l'utilisation du stockage sur le disque pour ces algorithmes et forcer le stockage en mémoire de toutes les données ; pour ce faire, sélectionnez l'option d'**optimisation de la vitesse** dans l'onglet Modèle de la boîte de dialogue du noeud. Si la quantité de mémoire requise pour le stockage des données est supérieure à la taille de l'ensemble de travail du processus serveur, une partie sera paginée sur le disque et les performances s'en trouveront réduites.

Lorsque l'option d'**optimisation de la mémoire** est activée, un pourcentage de mémoire RAM physique est alloué à l'algorithme en fonction de la valeur affectée à l'option de configuration IBM SPSS Modeler Server relative à la *limite (en pourcentage) de la mémoire de modélisation*. Pour consacrer davantage de

mémoire aux réseaux de neurones d'apprentissage, allouez davantage de mémoire RAM ou augmentez la valeur de cette option ; notez toutefois que si vous définissez une valeur trop élevée, une pagination sera réalisée.

Le temps d'exécution des algorithmes de réseau de neurones dépend du niveau d'exactitude requis. Vous pouvez gérer le temps d'exécution en définissant une condition d'arrêt dans la boîte de dialogue du noeud.

k moyenne. L'algorithme de classification non supervisée k moyenne comporte les mêmes options de gestion de l'utilisation de la mémoire que les algorithmes de réseau de neurones. Les performances relatives aux données stockées sur le disque sont toutefois supérieures car l'accès aux données s'effectue de manière séquentielle.

Performances : expressions CLEM

Les fonctions séquentielles CLEM (“fonctions `@`”) qui se réfèrent à nouveau au flux de données doivent stocker suffisamment de données pour permettre la consultation en aval la plus longue possible. Pour les opérations dont le degré de consultation en aval n'est pas limité, toutes les valeurs du champ doivent être stockées. Une opération est dite illimitée si la valeur de décalage n'est pas un entier littéral ; par exemple, `@OFFSET(Ventes, Mois)`. La valeur de décalage correspond au nom de champ *Mois* dont la valeur est inconnue avant exécution. Le serveur doit enregistrer toutes les valeurs du champ *Ventes* afin de garantir l'exactitude des résultats. Lorsqu'une limite supérieure est connue, fournissez-la en tant qu'argument supplémentaire ; par exemple, `@OFFSET(Ventes, Mois, 12)`. Cette opération indique au serveur de ne stocker au maximum que les 12 valeurs les plus récentes de *Ventes*. Les fonctions séquentielles, limitées ou non, entravent quasi systématiquement la génération SQL.

Chapitre 16. Accessibilité dans IBM SPSS Modeler

Présentation de l'accessibilité dans IBM SPSS Modeler

IBM SPSS Modeler est accessible par tous les utilisateurs et offre également une assistance spécifique pour les utilisateurs souffrant de déficiences visuelles ou autres. La présente section décrit les fonctions et méthodes de travail à l'aide des améliorations d'accessibilité, telles que les lecteurs d'écran et les raccourcis clavier.

Types de prise en charge pour l'accessibilité

Que vous souffriez de déficience visuelle ou que vous utilisiez systématiquement le clavier pour effectuer toute manipulation, il existe diverses méthodes d'utilisation de cet outil d'exploration de données. Par exemple, vous pouvez créer des flux, spécifier des options et lire des résultats sans jamais vous servir de la souris. Les raccourcis clavier disponibles sont répertoriés dans les rubriques suivantes. En outre, IBM SPSS Modeler fournit une prise en charge quasi complète des lecteurs d'écran, tels que JAWS pour Windows. Vous pouvez également optimiser le modèle de couleurs pour obtenir un contraste supplémentaire. Ces prises en charge sont abordées dans les rubriques suivantes.

Accessibilité pour les utilisateurs ayant une déficience visuelle

De nombreuses propriétés sont à votre disposition dans IBM SPSS Modeler pour vous permettre d'optimiser l'utilisation du logiciel.

Options d'affichage

Vous pouvez choisir les couleurs d'affichage des graphiques. Vous pouvez également utiliser vos propres paramètres Windows dans Clementine. Vous pourrez ainsi intensifier le contraste visuel.

1. Pour définir les options d'affichage, cliquez sur **Options utilisateur** dans le menu Outils.
2. Cliquez sur l'onglet **Affichage**. Les options de cet onglet vous permettent de choisir le modèle de couleurs du logiciel, les couleurs des graphiques, ainsi que la taille de police des noeuds.

Remarque : Le lecteur d'écran n'est pas en mesure de lire les graphiques, de sorte que ces derniers ne sont pas accessibles pour les utilisateurs malvoyants.

Utilisation de l'avertissement sonore

Vous pouvez contrôler le type d'avertissement utilisé pour certaines opérations dans le logiciel en activant ou désactivant les sons. Par exemple, vous pouvez activer les sons pour des événements, tels que la création ou la suppression d'un noeud, ou la génération de nouveaux résultats ou modèles.

1. Pour définir les options de notification, cliquez sur **Options utilisateur** dans le menu Outils.
2. Cliquez sur l'onglet **Notifications**.

Contrôle du lancement automatique de nouvelles fenêtres

L'onglet Notifications de la boîte de dialogue Options utilisateur permet également de choisir si les nouvelles sorties, comme les tableaux et les graphiques, doivent être générées dans une fenêtre séparée. Vous pouvez désactiver ces options et ouvrir une fenêtre de résultats uniquement lorsque vous le souhaitez.

1. Pour définir ces options, cliquez sur **Options utilisateur** dans le menu Outils.
2. Cliquez sur l'onglet **Notifications**.

© Copyright IBM Corp. 1994, 2019

3. Dans la boîte de dialogue, sélectionnez **Nouvelle sortie** dans la liste du groupe **Notifications visuelles**.
4. Sous **Ouvrir une fenêtre**, sélectionnez **Jamais**.

Taille des noeuds

La taille d'affichage des noeuds peut être standard ou petite. Vous pouvez choisir la taille en fonction de vos besoins.

1. Pour définir les options de taille de noeud, cliquez sur **Propriétés du flux** dans le menu Fichier.
2. Cliquez sur l'onglet **Présentation**.
3. Dans la liste **Taille des noeuds**, sélectionnez **Standard**.

Accessibilité pour les utilisateurs aveugles

La prise en charge proposée aux personnes aveugles dépend essentiellement de l'utilisation d'un lecteur d'écran, tel que JAWS pour Windows. Pour optimiser l'utilisation d'un lecteur d'écran avec IBM SPSS Modeler, vous pouvez spécifier plusieurs paramètres.

Options d'affichage

De manière générale, les lecteurs d'écran fonctionnent mieux lorsque le contraste visuel est élevé. Si vos paramètres Windows offrent un contraste élevé, utilisez-les dans Clementine.

1. Pour définir les options d'affichage, cliquez sur **Options utilisateur** dans le menu Outils.
2. Cliquez sur l'onglet **Affichage**.

Remarque : Le lecteur d'écran n'est pas en mesure de lire les graphiques, de sorte que ces derniers ne sont pas accessibles pour les utilisateurs non-voyants.

Utilisation de l'avertissement sonore

Vous pouvez contrôler le type d'avertissement utilisé pour certaines opérations dans le logiciel en activant ou désactivant les sons. Par exemple, vous pouvez activer les sons pour des événements, tels que la création ou la suppression d'un noeud, ou la génération de nouveaux résultats ou modèles.

1. Pour définir les options de notification, cliquez sur **Options utilisateur** dans le menu Outils.
2. Cliquez sur l'onglet **Notifications**.

Contrôle du lancement automatique de nouvelles fenêtres

L'onglet Notifications de la boîte de dialogue Options utilisateur permet également de choisir si les nouvelles sorties doivent être générées dans une fenêtre séparée. Vous pouvez désactiver cette option et ouvrir une fenêtre de résultats uniquement lorsque vous en avez besoin.

1. Pour définir ces options, cliquez sur **Options utilisateur** dans le menu Outils.
2. Cliquez sur l'onglet **Notifications**.
3. Dans la boîte de dialogue, sélectionnez **Nouvelle sortie** dans la liste du groupe **Notifications visuelles**.
4. Sous **Ouvrir une fenêtre**, sélectionnez **Jamais**.

Accessibilité via le clavier

La fonctionnalité du produit est accessible à partir du clavier. Vous pouvez appuyer simultanément sur la touche Alt + la touche appropriée pour activer les menus de la fenêtre (par exemple, Alt+F pour accéder au menu Fichier) ou sur la touche Tab pour passer d'une commande à l'autre dans une boîte de dialogue. Cependant, vous découvrirez ici des problèmes spécifiques liés à chaque fenêtre principale du produit, ainsi que des astuces utiles pour naviguer dans les boîtes de dialogue.

Cette section traite de l'utilisation du clavier pour effectuer des opérations, telles que l'ouverture d'un flux, l'utilisation des boîtes de dialogue des noeuds ou l'utilisation des résultats. En outre, des listes de raccourcis clavier sont fournies pour vous offrir une navigation encore plus efficace.

Raccourcis de navigation dans la fenêtre principale

Vous réalisez la plupart de votre travail d'exploration de données dans la fenêtre principale d'IBM SPSS Modeler. La zone principale est appelée l'**espace de travail de flux**. Elle permet de créer et d'exécuter les flux de données. La partie inférieure de la fenêtre contient les **palettes de noeuds** qui répertorient tous les noeuds disponibles. Les palettes sont organisées en onglets correspondant au type d'opération d'exploration de données pour chaque groupe de noeuds. Par exemple, les noeuds permettant d'importer les données dans IBM SPSS Modeler sont regroupés dans l'onglet Sources. Ceux permettant de calculer, de filtrer ou de classer les champs sont regroupés dans l'onglet Ops sur champs (abréviation d'Opérations sur les champs).

La partie droite de la fenêtre contient plusieurs outils de gestion des flux, des sorties et des projets. La partie supérieure à droite contient les **gestionnaires**. Vous y trouverez trois onglets permettant de gérer les flux, les sorties, ainsi que les modèles générés. Vous pouvez accéder à ces objets en sélectionnant l'onglet et l'objet voulu dans la liste. La partie inférieure à droite contient le **panneau du projet** qui vous permet d'organiser votre travail en projets. Cette zone contient deux onglets, qui correspondent à deux vues d'un projet. Le **vue Classes** trie les objets du projet par type, alors que le **mode CRISP-DM** trie les objets en fonction de la phase d'exploration de données, telle que la préparation des données ou la modélisation. Ces différents aspects de la fenêtre d'IBM SPSS Modeler sont traités dans le système d'aide, ainsi que dans le guide de l'utilisateur.

Le tableau suivant contient les raccourcis utilisés pour vous déplacer dans la fenêtre principale d'IBM SPSS Modeler et pour créer des flux. Les raccourcis des boîtes de dialogue et des sorties sont répertoriés dans les rubriques suivantes. Ces touches de raccourci sont disponibles uniquement dans la fenêtre principale.

Tableau 39. Raccourcis de la fenêtre principale

Touche de raccourci	Fonction
Ctrl+F5	Active la palette de noeuds.
Ctrl+F6	Active l'espace de travail de flux.
Ctrl+F7	Active le panneau des gestionnaires.
Ctrl+F8	Active le panneau des projets.

Tableau 40. Raccourcis des noeuds et des flux

Touche de raccourci	Fonction
Ctrl+N	Crée un espace de travail de flux vide.
Ctrl+O	Affiche la boîte de dialogue Ouvrir dans laquelle vous pouvez sélectionner et ouvrir un flux existant.
Ctrl+touches numériques	Permet d'activer l'onglet correspondant d'une fenêtre ou d'un panneau. Par exemple, dans une fenêtre à onglet, Ctrl+1 affiche le premier onglet à partir de la gauche, Ctrl+2 le deuxième, etc.
Ctrl+Flèche bas	S'utilise dans la palette de noeuds pour passer d'un onglet de la palette au premier noeud situé sous cet onglet.
Ctrl+Flèche haut	S'utilise dans la palette de noeuds pour se déplacer d'un noeud à l'onglet de sa palette.

Tableau 40. Raccourcis des noeuds et des flux (suite)

Touche de raccourci	Fonction
Entrée	Lorsqu'un noeud est sélectionné dans la palette de noeuds (y compris les modèles affinés de la palette de modèles générés), cette touche permet d'ajouter le noeud à l'espace de travail de flux. Appuyer sur **Entrée** lorsqu'un noeud est sélectionné dans l'espace de travail ouvre la boîte de dialogue de ce noeud.
Ctrl+Entrée	Lorsqu'un noeud est sélectionné dans la palette, ce raccourci permet d'ajouter le noeud à l'espace de travail de flux sans avoir à le sélectionner et d'activer le premier noeud de la palette.
Alt+Entrée	Lorsqu'un noeud est sélectionné dans la palette, ce raccourci permet d'ajouter le noeud à l'espace de travail de flux et de le sélectionner tout en activant le premier noeud de la palette.
Maj.+Barre d'espace	Lorsqu'un noeud ou un commentaire est activé dans la palette, ce raccourci permet de basculer entre la sélection ou la désélection de ce noeud ou de ce commentaire. Si d'autres noeuds ou commentaires sont également sélectionnés, cela a pour effet de les désélectionner.
Ctrl+Maj.+Barre d'espace	Lorsqu'un noeud ou un commentaire est activé dans le flux, ou qu'un noeud ou un commentaire est activé dans la palette, ce raccourci permet de basculer entre la sélection ou la désélection de ce noeud ou de ce commentaire. Ceci n'a aucune influence sur les autres noeuds ou commentaires sélectionnés.
Flèche gauche/droite	Si l'espace de travail de flux est sélectionné, déplace horizontalement le flux entier sur l'écran. Si l'onglet d'une palette est sélectionné, permet de se déplacer entre les onglets. Si un noeud de palette est activé, permet de se déplacer entre les noeuds de cette palette.
Flèche haut/bas	Si l'espace de travail de flux est sélectionné, déplace verticalement le flux entier sur l'écran. Si un noeud de palette est activé, permet de se déplacer entre les noeuds de cette palette. Si une sous-palette est activée, permet de se déplacer entre les autres sous-palettes de l'onglet de cette palette.
Alt+Flèche gauche/droite	Permet de déplacer horizontalement les noeuds et les commentaires sélectionnés dans l'espace de travail de flux dans le sens indiqué par la flèche.
Alt+Flèche haut/bas	Permet de déplacer verticalement les noeuds et les commentaires sélectionnés dans l'espace de travail de flux dans le sens indiqué par la flèche.
Ctrl+A	Sélectionne tous les noeuds d'un flux.
Ctrl+Q	Lorsqu'un noeud est activé, permet de le sélectionner, ainsi que tous les noeuds en aval et désélectionne tous les noeuds en amont.
Ctrl+W	Lorsqu'un noeud est activé, permet de le désélectionner ainsi que tous les noeuds sélectionnés en aval.
Ctrl+Alt+D	Duplique un noeud sélectionné.
Ctrl+Alt+L	Lorsqu'un nugget de modèle est sélectionné dans le flux, ouvre une boîte de dialogue Insérer pour vous permettre de charger un modèle enregistré d'un fichier `.nod` dans le flux.
Ctrl+Alt+R	Affiche l'onglet Annotations d'un noeud sélectionné et vous permet de renommer ce noeud.
Ctrl+Alt+U	Crée un noeud source Utilisateur.

Tableau 40. Raccourcis des noeuds et des flux (suite)

Touche de raccourci	Fonction
Ctrl+Alt+C	Permet d'activer ou de désactiver le cache d'un noeud.
Ctrl+Alt+F	Vide le cache d'un noeud
Tabulation	Dans l'espace de travail de flux, permet de se déplacer entre tous les noeuds source et les commentaires dans le flux en cours. Sur une palette de noeuds, permet de se déplacer entre les noeuds de la palette. Sur une sous-palette sélectionnée, permet de se déplacer vers le premier noeud de la sous-palette.
Maj+Tab	Effectue la même opération que la touche Tab mais dans l'ordre inverse.
Ctrl+Tab	Lorsque le panneau des gestionnaires ou celui des projets est activé, active l'espace de travail de flux. Lorsqu'une palette de noeuds est active, permet de basculer l'activation entre un noeud et l'onglet de la palette.
Une touche alphabétique quelconque	Dans le flux en cours, lorsque l'activation est sur un noeud, permet d'activer le noeud suivant dont le nom commence par la lettre de la touche sur laquelle vous avez appuyée.
F1	Ouvre le système d'Aide à une rubrique relative à l'élément activé.
F2	Permet de lancer le processus de connexion du noeud sélectionné dans l'espace de travail. Servez-vous de la touche Tab pour passer au noeud voulu dans l'espace de travail, puis appuyez sur Maj+barre d'espacement pour terminer la connexion.
F3	Permet de supprimer toutes les connexions du noeud sélectionné dans l'espace de travail.
F6	Déplace l'activation entre le panneau des gestionnaires, le panneau des projets et les palettes des noeuds.
F10	Permet d'ouvrir le menu Fichier.
Maj+F10	Ouvre le menu contextuel pour le noeud ou le flux.
Suppr	Supprime un noeud sélectionné de l'espace de travail.
Echap	Ferme le menu contextuel ou la boîte de dialogue.
Ctrl+Alt+X	Permet de développer un super noeud.
Ctrl+Alt+Z	Effectue un zoom avant dans un super noeud.
Ctrl+Alt+Maj+Z	Effectue un zoom arrière dans un super noeud.
Ctrl+E	Lorsque l'élément activé se trouve dans l'espace de travail de flux, permet d'exécuter le flux en cours.

Il existe également certaines touches de raccourci standard dans IBM SPSS Modeler, comme Ctrl+C pour copier un élément. Pour plus d'informations, voir la rubrique «Utilisation des touches de raccourci», à la page 22.

Raccourcis des boîtes de dialogue et des tableaux

Il existe plusieurs raccourcis et touches de lecteur d'écran très pratiques à utiliser dans les boîtes de dialogue et les tableaux, ainsi que dans les tableaux situés dans les boîtes de dialogue. La liste exhaustive des raccourcis clavier et de lecteur d'écran spéciaux est fournie ci-dessous.

Tableau 41. Raccourcis des boîtes de dialogue et du Générateur de formules

Touche de raccourci	Fonction
Alt+4	Permet de fermer toutes les boîtes de dialogue ou fenêtres de sortie actives. Les sorties peuvent être récupérées dans l'onglet Sorties du panneau des gestionnaires.

Tableau 41. Raccourcis des boîtes de dialogue et du Générateur de formules (suite)

Touche de raccourci	Fonction
Ctrl+Fin	Lorsqu'une commande quelconque du Générateur de formules est activée, cette combinaison déplace le point d'insertion à la fin de l'expression.
Ctrl+1	Dans le Générateur de formules, permet d'activer la commande d'édition de l'expression.
Ctrl+2	Dans le Générateur de formules, permet d'activer la liste des fonctions.
Ctrl+3	Dans le Générateur de formules, permet d'activer la liste des champs.

Raccourcis de tableaux

Les raccourcis de tableaux sont utilisés pour les tableaux de sortie, mais également pour les commandes de tableau dans les boîtes de dialogue de noeuds tels que type, Filtrer et Fusionner. Vous utilisez généralement la touche Tab pour vous déplacer d'une cellule du tableau à l'autre et les touches Ctrl+Tab pour quitter la commande de tableau. *Remarque* : certains lecteurs d'écran peuvent parfois ne pas commencer immédiatement la lecture du contenu d'une cellule. Appuyez une fois ou deux sur les touches de direction pour réinitialiser le logiciel et commencer la lecture.

Tableau 42. Raccourcis de tableaux

Touche de raccourci	Fonction
Ctrl+W	Dans les tableaux, permet de lire la brève description de la **ligne** sélectionnée. Par exemple, « Les valeurs de la ligne 2 sélectionnée sont sexe, indicateur, m/f, etc. »
Ctrl+Alt+W	Dans les tableaux, permet de lire la longue description de la **ligne** sélectionnée. Par exemple, "Les valeurs de la ligne 2 sélectionnée sont champs = sexe, type = indicateur, sexe = m/f, etc."
Ctrl+D	Dans les tableaux, permet de lire la brève **description** de la zone sélectionnée. Par exemple, « Sélection : une ligne pour six colonnes ».
Ctrl+Alt+D	Dans les tableaux, fournit la longue **description** de la zone sélectionnée. Par exemple, "Sélection : une ligne pour six colonnes. Les colonnes sélectionnées sont Champs, Type, Manquant. La ligne 1 est sélectionnée."
Ctrl+T	Dans les tableaux, permet de lire la brève description des colonnes sélectionnées. Par exemple, "Champs, Type, Manquant".
Ctrl+Alt+T	Dans les tableaux, permet de lire la longue description des colonnes sélectionnées. Par exemple, "Les colonnes sélectionnées sont Champs, Type, Manquant".
Ctrl+R	Indique le nombre d'**enregistrements** contenus dans le tableau.
Ctrl+Alt+R	Indique le nombre d'**enregistrements** contenus dans le tableau, ainsi que le nom des colonnes.
Ctrl+I	Dans les tableaux, lit les **informations** ou le contenu de la cellule activée.
Ctrl+Alt+I	Dans les tableaux, lit la longue description d'**informations** (nom de la colonne et contenu de la cellule) sur la cellule active.
Ctrl+G	Dans les tableaux, fournit de brèves informations **générales** sur la sélection.
Ctrl+Alt+G	Dans les tableaux, fournit de longues informations **générales** sur la sélection.

Tableau 42. Raccourcis de tableaux (suite)

Touche de raccourci	Fonction
Ctrl+Q	Dans les tableaux, permet de passer **rapidement** d'une cellule à l'autre. Ctrl+Q permet de lire les descriptions longues, comme "Sexe=Féminin," lorsque vous vous déplacez dans le tableau à l'aide des touches de direction. Si vous sélectionnez de nouveau Ctrl+Q, vous passez aux descriptions courtes (contenu de la cellule).
F8	Pour les tableaux, lorsque le tableau a le focus, affecte le focus à l'en-tête de colonne.
Barre d'espace	Pour les tableaux, lorsque l'en-tête de colonne a le focus, active le tri de colonne.

Raccourcis pour les commentaires

Lorsque vous utilisez des commentaires à l'écran, vous pouvez utiliser les raccourcis suivants.

Tableau 43. Raccourcis de commentaires

Touche de raccourci	Fonction
Alt+C	Permet d'afficher ou de masquer les commentaires.
Alt+M	Insère un nouveau commentaire si des commentaires sont actuellement affichés ; affiche des commentaires s'ils sont actuellement masqués.
Tabulation	Dans l'espace de travail de flux, permet de se déplacer entre tous les noeuds source et les commentaires dans le flux en cours.
Entrée	Lorsqu'un commentaire est activé, indique le début de la modification.
Alt+Entrée ou Ctrl+Tab	Termine la modification et enregistre les modifications.
Echap	Annule les modifications. Les modifications effectuées au cours de l'édition sont perdues.
Alt+Maj+Flèche haut	Réduit la hauteur de la zone de texte d'une cellule de grille (ou d'un pixel) si l'ajustement à la grille est activé (ou désactivé).
Alt+Maj+Flèche bas	Augmente la hauteur de la zone de texte d'une cellule de grille (ou d'un pixel) si l'ajustement à la grille est activé (ou désactivé).
Alt+Maj+Flèche gauche	Réduit la largeur de la zone de texte d'une cellule de grille (ou d'un pixel) si l'ajustement à la grille est activé (ou désactivé).
Alt+Maj+Flèche droite	Augmente la largeur de la zone de texte d'une cellule de grille (ou d'un pixel) si l'ajustement à la grille est activé (ou désactivé).

Raccourcis pour le Visualiseur de clusters et le Visualiseur de modèles

Les touches de raccourcis sont disponibles pour la navigation dans les fenêtres du Visualiseur de clusters et du Visualiseur de modèles.

Tableau 44. Raccourcis pour le Visualiseur de clusters et le Visualiseur de modèles

Touche de raccourci	Fonction
Tabulation	Active la commande de l'écran suivant.
Maj+Tab	Active la commande de l'écran précédent.

Tableau 44. Raccourcis pour le Visualiseur de clusters et le Visualiseur de modèles (suite)

Touche de raccourci	Fonction
Flèche Bas	Si une liste déroulante est activée, ouvre la liste ou passe à l'élément suivant de la liste. Si un menu est activé, passe à l'élément suivant du menu. Si un graphique en miniature est activé, déplace l'activation au graphique suivant dans l'ensemble (ou au premier graphique de l'ensemble si le dernier graphique est activé).
Flèche Haut	Si une liste déroulante est ouverte, passe à l'élément précédent de la liste. Si un menu est activé, passe à l'élément précédent du menu. Si un graphique en miniature est activé, déplace l'activation au graphique précédent dans l'ensemble (ou au dernier graphique de l'ensemble si le premier graphique miniature est activé).
Entrée	Ferme une liste déroulante ou exécute une sélection dans un menu ouvert.
F6	Bascule l'activation entre les panneaux de gauche et de droite de la fenêtre.
Flèches de gauche et de droite	Si un onglet est activé, passe à l'onglet précédent ou suivant. Si un menu est activé, passe au menu précédent ou suivant.
Alt+*lettre*	Sélectionne le bouton ou le menu dont cette lettre est soulignée dans le nom.
Echap	Ferme une liste déroulante ou un menu ouvert.

Visualiseur de clusters uniquement

Le Visualiseur de clusters possède une vue de clusters qui contient une grille de clusters par caractéristiques.

Pour sélectionner la vue des clusters au lieu de la vue récapitulative du modèle :

1. Appuyez à plusieurs sur la touche Tab jusqu'à ce que le bouton **Vue** soit sélectionné.
2. Appuyez deux fois sur la flèche bas pour sélectionner **Clusters**.
   A ce stade, vous pouvez sélectionner une cellule individuelle dans la grille :
3. Appuyez à plusieurs reprises sur la touche Tab jusqu'à ce que vous atteigniez la dernière icône de la barre d'outils de visualisation.

Figure 19. Icône Afficher l'arbre de visualisation

4. Appuyez une fois de plus sur la touche Tab, puis sur la barre d'espacement et sur une touche fléchée.

Les raccourcis clavier suivants sont désormais disponibles :

Tableau 45. Raccourcis pour le Visualiseur de clusters

Touche de raccourci	Fonction
Touche fléchée	Permet de passer d'une cellule individuelle à l'autre dans la grille. L'affichage Distribution des cellules dans le panneau de droite change avec l'activation.
Ctrl+, (virgule)	Sélectionne ou désélectionne la totalité de la colonne de la grille dont une cellule est activée. Pour ajouter une colonne à la sélection, utilisez les touches fléchées pour accéder à une cellule de cette colonne et appuyez de nouveau sur Ctrl+.
Tabulation	Désactive la grille et active la commande de l'écran suivant.
Maj+Tab	Désactive la grille et revient à la commande de l'écran précédent.
F2	Passe en mode d'édition (cellules libellé et description uniquement).
Entrée	Enregistre les modifications d'édition et quitte le mode d'édition (cellules libellé et description uniquement).
Echap	Quitte le mode d'édition sans enregistrer les modifications (cellules libellé et description uniquement).

Exemple de touches de raccourci : création de flux

Afin de clarifier le processus de création de flux pour les utilisateurs qui utilisent le clavier ou un lecteur d'écran, suivez l'exemple de création de flux sans souris ci-dessous. Dans cet exemple, vous allez créer un flux contenant un noeud Délimité, un noeud Calculer et un noeud Histogramme en procédant comme suit :

1. **Démarrez IBM SPSS Modeler.** Lorsqu'IBM SPSS Modeler démarre pour la première fois, l'onglet Favoris de la palette de noeuds est activé.
2. **Ctrl+Flèche bas.** Lorsque l'onglet est activé, permet d'activer son contenu.
3. **Flèche droite.** Permet d'activer le noeud Délimité.
4. **Barre d'espacement.** Permet de sélectionner le noeud Délimité.
5. **Ctrl+Entrée.** Ajoute le noeud Délimité dans l'espace de travail du flux. Cette combinaison de touches permet également de conserver l'activation du noeud Délimité afin d'établir une connexion avec le noeud ajouté suivant.
6. **Tabulation.** Active à nouveau la palette de noeuds.
7. **Flèche droite (4 fois).** Permet d'accéder au noeud Calculer.
8. **Barre d'espacement.** Permet de sélectionner le noeud Calculer.
9. **Alt+Entrée.** Permet d'ajouter le noeud Calculer à l'espace de travail et d'activer le noeud Calculer. Ce noeud est maintenant prêt à être connecté au noeud ajouté suivant.
10. **Tabulation.** Active à nouveau la palette de noeuds.
11. **Flèche droite (5 fois).** Permet d'activer le noeud Histogramme dans la palette.
12. **Barre d'espacement.** Permet de sélectionner le noeud Histogramme.
13. **Entrée.** Permet d'ajouter le noeud au flux et d'activer l'espace de travail de flux.

Passez à l'exemple suivant ou enregistrez le flux si vous souhaitez essayer ultérieurement l'exemple suivant.

Exemple de touches de raccourci : édition de noeuds

Dans cet exemple, vous utiliserez le flux créé dans l'exemple précédent. Le flux se compose d'un noeud Délimité, d'un noeud Calculer et d'un noeud Histogramme. Les instructions commencent avec l'activation du troisième noeud du flux, le noeud Histogramme.

1. **Ctrl+Flèche gauche deux fois.** Permet d'activer à nouveau le noeud Délimité.

2. **Entrée.** Permet d'ouvrir la boîte de dialogue Délimité. A l'aide de la touche Tab, accédez au champ Fichier et saisissez le chemin et le nom du fichier pour sélectionner ce fichier. Appuyez sur Ctrl+Tab pour accéder à la partie inférieure de la boîte de dialogue, puis au bouton OK et appuyez sur Entrée pour fermer la boîte de dialogue.
3. **Ctrl+Flèche droite.** Permet d'activer le deuxième noeud, le noeud Calculer.
4. **Entrée.** Permet d'ouvrir la boîte de dialogue du noeud Calculer. A l'aide de la touche Tab, sélectionnez les champs et spécifiez les conditions de calcul. Appuyez sur Ctrl+Tab pour accéder au bouton OK et appuyez sur Entrée pour fermer la boîte de dialogue.
5. **Ctrl+Flèche droite.** Permet d'activer le troisième noeud, le noeud Histogramme.
6. **Entrée.** Permet d'ouvrir la boîte de dialogue du noeud Histogramme. A l'aide de la touche Tab, sélectionnez les champs et spécifiez les options de graphique. Pour les listes déroulantes, appuyez sur la flèche bas pour ouvrir la liste et pour mettre en surbrillance un élément de la liste, puis appuyez sur Entrée pour sélectionner l'élément de la liste. A l'aide de la touche Tab, accédez au bouton OK et appuyez sur Entrée pour fermer la boîte de dialogue.

Vous pouvez maintenant ajouter d'autres noeuds ou exécuter le flux en cours. Gardez à l'esprit les conseils suivants lorsque vous créez des flux :

- Lorsque vous reliez manuellement des noeuds, utilisez F2 pour créer le point de départ d'une connexion, appuyez sur la touche Tab pour accéder au point de fin, puis utilisez Maj+Espace pour finaliser la connexion.
- Utilisez la touche F3 pour supprimer toutes les connexions du noeud sélectionné dans l'espace de travail.
- Lorsque le flux est créé, utilisez Ctrl+E pour exécuter le flux en cours.

La liste complète des touches de raccourci est disponible. Pour plus d'informations, voir la rubrique «Raccourcis de navigation dans la fenêtre principale», à la page 245.

Utilisation d'un lecteur d'écran

Divers lecteurs d'écran sont disponibles sur le marché. IBM SPSS Modeler est configuré pour la prise en charge de JAWS pour Windows en utilisant Java Access Bridge installé avec IBM SPSS Modeler. Si JAWS est installé, démarrez-le simplement avant de lancer IBM SPSS Modeler pour utiliser ce produit.

Remarque : Il est recommandé d'affecter au moins 6 Go d'espace pour l'exécution de JAWS avec SPSS Modeler.

Compte tenu de la nature de la représentation graphique unique du processus d'exploration de données d'IBM SPSS Modeler, une analyse visuelle des graphiques et des graphiques est la solution conseillée. Il est toutefois possible de comprendre les éléments et de prendre des décisions en fonction des résultats et modèles affichés textuellement avec un lecteur d'écran.

Remarque : Avec des machines clientes 64 bits, certaines caractéristiques technologiques d'assistance ne fonctionnent pas. Ceci est dû au fait que Java Access Bridge n'est pas conçu pour une utilisation en 64 bits.

Utilisation du fichier de dictionnaire dans IBM SPSS Modeler

Un fichier de dictionnaire d'IBM SPSS Modeler (*Awt.JDF*) peut être inclus dans JAWS. Pour utiliser ce fichier :

1. Accédez au sous-répertoire */accessibility* de votre installation IBM SPSS Modeler et copiez le fichier de dictionnaire (*Awt.JDF*).
2. Copiez-le dans le répertoire où se trouvent vos scripts JAWS.

Un fichier appelé *Awt.JDF* existe peut-être déjà sur votre ordinateur si d'autres applications JAVA sont en cours d'exécution. Dans ce cas, vous aurez peut-être à éditer manuellement le fichier de dictionnaire pour pouvoir l'utiliser.

Utilisation d'un lecteur d'écran avec une sortie au format HTML

Lorsque vous consultez un résultat affiché au format HTML dans IBM SPSS Modeler à l'aide d'un lecteur d'écran, vous risquez de rencontrer des problèmes. Plusieurs types de sortie sont concernés, notamment :

- Sortie affichée dans l'onglet Options avancées pour les noeuds Régression, Régression logistique et Analyse factorielle/ACP
- Sortie du noeud Rapport

Dans chacune de ces fenêtres ou boîtes de dialogue, la barre d'outils contient un outil qui permet de lancer la sortie dans votre navigateur par défaut et qui fournit la prise en charge de lecteur d'écran standard. Vous pouvez ensuite utiliser le lecteur d'écran pour consulter les informations de sortie.

Accessibilité dans la fenêtre Arbre interactif

L'affichage standard des modèles d'arbre décision dans la fenêtre Arbre interactif peut poser problème aux lecteurs d'écran. Pour ouvrir une version accessible, dans les menus de l'arbre interactif, cliquez sur :

Vue > **Fenêtre accessible**

Très semblable à la carte d'arbre standard, l'affichage obtenu peut être lu par JAWS. Vous pouvez vous déplacer vers le haut, le bas, la droite ou la gauche à l'aide des touches de direction standard. Au fur et à mesure que vous parcourez la fenêtre accessible, l'activation dans la fenêtre Arbre interactif change. Utilisez la barre d'espacement pour modifier la sélection, ou les touches Ctrl+Espace pour agrandir la sélection en cours.

Conseils d'utilisation

Il existe plusieurs astuces pour faciliter l'accès à l'environnement IBM SPSS Modeler. Voici des conseils généraux pour travailler dans IBM SPSS Modeler.

- **Fermeture des zones de texte étendues.** Utilisez Ctrl+Tab pour quitter les zones de texte étendues. Remarque : Ctrl+Tab permet également de quitter les commandes de tableaux.
- **Utilisation de la touche Tab à la place des touches de direction.** Pour sélectionner des options dans une boîte de dialogue, utilisez la touche Tab pour accéder à ces options. Les touches de direction ne fonctionnent pas dans ce contexte.
- **Listes déroulantes.** Dans les listes déroulantes des boîtes de dialogue, vous pouvez utiliser la touche Echap ou la barre d'espacement pour sélectionner un élément, puis fermer la liste. Vous pouvez également utiliser la touche Echap pour fermer une liste déroulante qui est restée active alors que vous avez accédé à une autre commande.
- **Statut de l'exécution**. Lorsque vous exécutez un flux dans une base de données volumineuse, JAWS risque de lire le statut du flux avec un peu de retard. Appuyez régulièrement sur la touche Ctrl pour actualiser le rapport sur le statut.
- **Utilisation des palettes de noeuds.** Lorsque vous accédez pour la première fois à un onglet des palettes de noeuds, JAWS lit parfois "groupbox" (zone de groupe) au lieu du nom du noeud. Dans ce cas, utilisez Ctrl+Flèche droite, puis Ctrl+Flèche gauche pour réinitialiser le lecteur d'écran et écouter le nom du noeud.
- **Lecture des menus.** Lors de la première ouverture d'un menu, il arrive que JAWS ne puisse pas lire le premier élément de menu. Si vous pensez que cette situation s'est produite, utilisez la flèche vers le bas, puis la flèche vers le haut pour entendre le nom du premier élément du menu.
- **Menus en cascade.** JAWS ne lit pas le premier niveau d'un menu en cascade. Si vous entendez une rupture dans l'énoncé pendant que vous vous déplacez dans un menu, appuyez sur la flèche droite pour entendre le nom des éléments du menu enfant.

En outre, si vous avez installé IBM SPSS Modeler Text Analytics, les conseils suivants peuvent vous rendre l'interface utilitaire interactive plus accessible.

- **Entrée dans une boîte de dialogue.** Vous devrez peut-être appuyer sur la touche Tab pour activer la première commande lorsque vous accédez à une boîte de dialogue.
- **Fermeture des zones de texte étendues.** Utilisez Ctrl+Tab pour fermer les zones de texte étendues et accéder à la commande suivante. Remarque : Ctrl+Tab permet également de quitter les commandes de tableaux.
- **Saisie de la première lettre pour trouver un élément dans la liste de l'arbre.** Lorsque vous recherchez un élément dans les panneaux des catégories ou des résultats extraits, ou dans l'arbre de bibliothèque, vous pouvez taper la première lettre de l'élément si le panneau est actif. Cela permet de sélectionner l'élément suivant commençant par la lettre saisie.
- **Listes déroulantes.** Dans les listes déroulantes des boîtes de dialogue, vous pouvez utiliser la barre d'espacement pour sélectionner un élément, puis fermer la liste.

Interférence avec d'autres logiciels

Lors des tests d'IBM SPSS Modeler avec des lecteurs d'écran comme JAWS, notre équipe de développement a découvert que l'emploi de SMS (Systems Management Server) dans votre entreprise risquait de provoquer des interférences lorsque JAWS lit des applications Java, telles que IBM SPSS Modeler. Ce problème est résolu si vous désactivez SMS. Pour plus d'informations sur le SMS, consultez le site Web de Microsoft.

JAWS et Java

Les différentes versions de JAWS offrent divers niveaux de prise en charge des applications logicielles Java. Bien qu'IBM SPSS Modeler fonctionne avec toutes les versions récentes de JAWS, certaines versions risquent tout de même de rencontrer de légers problèmes lorsqu'elles sont utilisées avec des systèmes Java. Visitez le site Web de JAWS pour Windows à l'adresse http://www.FreedomScientific.com.

Utilisation des graphiques dans IBM SPSS Modeler

Il est difficile pour un lecteur d'écran d'interpréter l'affichage visuel des informations, comme les histogrammes, les graphiques d'évaluation, les courbes et les nuages de points. Remarquez cependant que les graphiques Relations et les distributions peuvent être affichés en utilisant un récapitulatif textuel disponible dans la fenêtre de sortie.

Chapitre 17. prise en charge d'Unicode

Prise en charge d'Unicode dans IBM SPSS Modeler

IBM SPSS Modeler prend totalement en charge Unicode pour IBM SPSS Modeler et IBM SPSS Modeler Server. Ainsi, il est possible d'échanger des données avec d'autres applications Unicode, y compris les bases de données multilingues et ce, sans perdre d'informations suite à une conversion vers ou depuis un code propre aux paramètres régionaux.

- IBM SPSS Modeler stocke les données Unicode en interne, et peut lire et créer les données multilingues stockées au format Unicode dans les bases de données, sans aucune perte.
- IBM SPSS Modeler peut lire et écrire les fichiers texte ayant fait l'objet d'un codage UTF-8. Les fonctions d'importation/exportation de fichiers texte utilisent par défaut le code régional, mais prennent en charge UTF-8 en guise d'alternative. Ce paramètre peut être spécifié dans les noeuds d'importation/exportation de fichiers. Le codage par défaut peut également être modifié dans la boîte de dialogue Propriétés du flux. Pour plus d'informations, voir la rubrique «Définition d'options générales pour les flux», à la page 44.
- Les fichiers de données Statistiques, SAS et de texte stockés dans le code régional sont convertis au format UTF-8 au cours de l'importation, puis une nouvelle fois durant l'exportation. Lors de l'écriture d'un fichier, si certains caractères Unicode ne figurent pas dans le jeu de caractères régional, ils sont remplacés et un message d'avertissement apparaît. Cela ne doit se produire que si les données sont importées depuis une source de données qui prend en charge Unicode (une base de données ou un fichier texte UTF-8), et qui inclut les caractères d'autres paramètres régionaux ou jeux de caractères.
- Les images IBM SPSS Modeler Solution Publisher font l'objet d'un codage UTF-8 et peuvent être littéralement transférées entre plates-formes, quels que soient les paramètres régionaux.

A propos d'Unicode

La norme Unicode a pour objectif d'assurer le codage cohérent d'un texte multilingue pour qu'il puisse être partagé facilement entre les frontières, quels que soient les paramètres régionaux et les applications. La norme Unicode, dont la version actuelle est 4.0.1, définit un jeu de caractères constituant un sur-ensemble de tous les jeux de caractères utilisés aujourd'hui à l'échelle internationale, et attribue à chaque caractère un nom et un code caractère uniques. Les caractères et les codes caractère correspondants sont identiques à ceux du jeu universel de caractères défini par la norme ISO-10646. Pour plus d'informations, consultez la page d'accueil d'Unicode.

© Copyright IBM Corp. 1994, 2019

Remarques

Le présent document a été développé pour des produits et des services proposés aux Etats-Unis et peut être mis à disposition par IBM dans d'autres langues. Toutefois, il peut être nécessaire de posséder une copie du produit ou de la version du produit dans cette langue pour pouvoir y accéder.

Le présent document peut contenir des informations ou des références concernant certains produits, logiciels ou services IBM non annoncés dans ce pays. Pour plus de détails, référez-vous aux documents d'annonce disponibles dans votre pays, ou adressez-vous à votre partenaire commercial IBM. Toute référence à un produit, programme ou service IBM n'implique pas que seul ce produit, programme ou service IBM puisse être utilisé. Tout produit, programme ou service fonctionnellement équivalent peut être utilisé s'il n'enfreint aucun droit de propriété intellectuelle d'IBM. Cependant l'utilisateur doit évaluer et vérifier l'utilisation d'un produit, programme ou service non IBM.

IBM peut détenir des brevets ou des demandes de brevet couvrant les produits mentionnés dans le présent document. L'octroi de ce document n'équivaut aucunement à celui d'une licence pour ces brevets. Vous pouvez envoyer par écrit des questions concernant la licence à :

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

Pour toute demande au sujet des licences concernant les jeux de caractères codés sur deux octets (DBCS), contactez le service Propriété intellectuelle IBM de votre pays ou adressez vos questions par écrit à :

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan

LE PRESENT DOCUMENT EST LIVRE "EN L'ETAT" SANS AUCUNE GARANTIE EXPLICITE OU IMPLICITE. IBM DECLINE NOTAMMENT TOUTE RESPONSABILITE RELATIVE A CES INFORMATIONS EN CAS DE CONTREFACON AINSI QU'EN CAS DE DEFAUT D'APTITUDE A L'EXECUTION D'UN TRAVAIL DONNE. Certaines juridictions n'autorisent pas l'exclusion des garanties implicites, auquel cas l'exclusion ci-dessus ne vous sera pas applicable.

Ces informations peuvent contenir des erreurs techniques ou des erreurs typographiques. Ce document est mis à jour périodiquement. Chaque nouvelle édition inclut les mises à jour. IBM peut, à tout moment et sans préavis, modifier les produits et logiciels décrits dans ce document.

Les références à des sites Web non IBM sont fournies à titre d'information uniquement et n'impliquent en aucun cas une adhésion aux données qu'ils contiennent. Les éléments figurant sur ces sites Web ne font pas partie des éléments du présent produit IBM et l'utilisation de ces sites relève de votre seule responsabilité.

IBM pourra utiliser ou diffuser, de toute manière qu'elle jugera appropriée et sans aucune obligation de sa part, tout ou partie des informations qui lui seront fournies.

Les licenciés souhaitant obtenir des informations permettant : (i) l'échange des données entre des logiciels créés de façon indépendante et d'autres logiciels (dont celui-ci), et (ii) l'utilisation mutuelle des données ainsi échangées, doivent adresser leur demande à :

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

Ces informations peuvent être disponibles, soumises à des conditions générales, et dans certains cas payantes.

Le programme sous licence décrit dans le présent document et tous les éléments sous licence disponibles s'y rapportant sont fournis par IBM conformément aux dispositions du Livret Contractuel IBM, des Conditions internationales d'utilisation des Logiciels IBM ou de tout autre contrat équivalent.

Les données de performances et les exemples de clients ne sont présentés qu'à des fins d'illustration. Les performances réelles peuvent varier en fonction des configurations et des conditions d'exploitation spécifiques.

Les informations concernant des produits non IBM ont été obtenues auprès des fournisseurs de ces produits, par l'intermédiaire d'annonces publiques ou via d'autres sources disponibles. IBM n'a pas testé ces produits et ne peut confirmer l'exactitude de leurs performances ni leur compatibilité. Elle ne peut recevoir aucune réclamation concernant des produits non IBM. Toute question concernant les performances de produits non IBM doit être adressée aux fournisseurs de ces produits.

Toute instruction relative aux intentions d'IBM pour ses opérations à venir est susceptible d'être modifiée ou annulée sans préavis, et doit être considérée uniquement comme un objectif.

Ces informations contiennent des exemples de données et de rapports utilisés au cours d'opérations quotidiennes standard. Pour les illustrer le mieux possible, ces exemples contiennent des noms d'individus, d'entreprises, de marques et de produits. Toute ressemblance avec des noms de personnes, de sociétés ou des données réelles serait purement fortuite.

Marques

IBM, le logo IBM et ibm.com sont des marques d'International Business Machines dans de nombreux pays. Les autres noms de produits et de services peuvent être des marques d'IBM ou d'autres sociétés. La liste actualisée de toutes les marques d'IBM est disponible sur la page Web "Copyright and trademark information" à www.ibm.com/legal/copytrade.shtml.

Adobe, le logo Adobe, PostScript et le logo PostScript sont des marques déposées ou des marques commerciales de Adobe Systems Incorporated aux Etats-Unis et/ou dans d'autres pays.

Intel, le logo Intel, Intel Inside, le logo Intel Inside, Intel Centrino, le logo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium, et Pentium sont des marques commerciales ou des marques déposées de Intel Corporation ou de ses filiales aux Etats-Unis et dans d'autres pays.

Linux est une marque déposée de Linus Torvalds aux Etats-Unis et/ou dans d'autres pays.

Microsoft, Windows, Windows NT et le logo Windows sont des marques commerciales de Microsoft Corporation aux Etats-Unis et/ou dans d'autres pays.

UNIX est une marque déposée de The Open Group aux Etats-Unis et dans d'autres pays.

Java ainsi que tous les logos et toutes les marques incluant Java sont des marques d'Oracle et/ou de ses sociétés affiliées.

Dispositions applicables à la documentation du produit

Les droits d'utilisation relatifs à ces publications sont soumis aux dispositions suivantes.

Applicabilité

Les présentes dispositions viennent s'ajouter à toute autre condition d'utilisation applicable au site Web IBM.

Usage personnel

Vous pouvez reproduire ces publications pour votre usage personnel, non commercial, sous réserve que toutes les mentions de propriété soient conservées. Vous ne pouvez distribuer ou publier tout ou partie de ces publications ou en faire des oeuvres dérivées sans le consentement exprès d'IBM.

Usage commercial

Vous pouvez reproduire, distribuer et afficher ces publications uniquement au sein de votre entreprise, sous réserve que toutes les mentions de propriété soient conservées. Vous ne pouvez reproduire, distribuer, afficher ou publier tout ou partie de ces publications en dehors de votre entreprise, ou en faire des oeuvres dérivées, sans le consentement exprès d'IBM.

Droits

Excepté les droits d'utilisation expressément accordés dans ce document, aucun autre droit, licence ou autorisation, implicite ou explicite, n'est accordé pour ces informations ou autres données, logiciels ou droits de propriété intellectuelle contenus dans ces informations.

IBM se réserve le droit de retirer les autorisations accordées ici si, à sa discrétion, l'utilisation des informations s'avère préjudiciable à ses intérêts ou que, selon son appréciation, les instructions susmentionnées n'ont pas été respectées.

Vous ne pouvez télécharger, exporter ou réexporter ces informations qu'en total accord avec toutes les lois et règlements applicables dans votre pays, y compris les lois et règlements américains relatifs à l'exportation.

IBM N'OCTROIE AUCUNE GARANTIE SUR LE CONTENU DE CES PUBLICATIONS. LES PUBLICATIONS SONT LIVREES EN L'ETAT SANS AUCUNE GARANTIE EXPLICITE OU IMPLICITE. IBM DECLINE NOTAMMENT TOUTE RESPONSABILITE RELATIVE A CES PUBLICATIONS EN CAS DE CONTREFAÇON AINSI QU'EN CAS DE DEFAUT D'APTITUDE A L'EXECUTION D'UN TRAVAIL DONNE.

Index

A

accessibilité 243, 254
  astuces dans IBM SPSS Modeler 253
  exemple 251
  fonctions dans IBM SPSS Modeler 243
activer des noeuds 40
affichage 99, 119
  légendes 124
  libellés de dimension 119
  lignes et colonnes 119
  notes de bas de page 124
  résultats 99
  titres 120
ajout
  à un projet 216
ajout de connexions IBM SPSS Modeler Server 11
ajout de libellés de groupe 117
ajuster les flux à la vue 21
alignement 100, 129
  sortie 100, 129
Analyse Champion/Challenger 192, 205
analyse en arbre
  applications standard 25
annotation
  flux 59, 64
  noeuds 59, 64
annotations
  dossier 218
  projet 218
Annotations
  conversion en commentaires 64
annuler 18
aperçu
  données de noeud 43
applications 25
applications d'exploration de données 26
applications standard 25
apprentissage automatique 25
arbres décision
  accessibilité 253
arrêter l'exécution 18
association de lignes ou de colonnes 117
astuces
  pour l'accessibilité 253
  utilisation générale 69
attribute 25
attribution d'un nom à des noeuds et à des flux 64
automatisation 139
avertissements 51
  définition des options 225

B

barre d'outils 18
base de données
  fonctions 147, 148
blanc
  suppression dans les chaînes 143, 171
blancs 132, 144
boîte de dialogue d'accueil 226
boîte de dialogue de démarrage 226
boîtes à moustaches 82
Boîtes de dialogue 128
  Affichage des libellés de variable 128
  Affichage des noms de variable 128
  Ordre d'affichage des variables 128
bordures 123, 126
  affichage des bordures masquées 126
bouton central de la souris
  simulation 22, 38
branches, modélisation et scoring 59, 206, 207, 208

C

cache
  activation 224
  définition d'un cache 40
  enregistrement 41
  vidage 41, 44
caractère barre oblique inversée dans les expressions CLEM 156
caractères 155, 156
caractères spéciaux
  suppression dans les chaînes 143
cartes thermiques 86
cellules dans les tableaux croisés dynamiques 119, 122, 126
  affichage 119
  Formats 122
  largeur 126
  masquage 119
  sélection 126
Centrage de sortie 100, 129
chaînes 155, 156
  manipulation dans les expressions CLEM 143
  mise en correspondance 143
  remplacement 143
champs 25, 155, 157
  affichage des valeurs 151
  dans les expressions CLEM 150
champs du modèle 69
champs essentiels 67, 69
champs non mappés 67
champs obligatoires 69
champs Temps
  conversion 180
chargement
  états 67
  noeuds 67
classes 18, 215, 216
CLEM 147
  création d'expressions 147
  exemples 139
  expressions 142, 155
CLEM *(suite)*
  fonctions 147, 148
  Introduction 23, 139
  langage 155
  types de données 155, 156, 157
  vérification des expressions 151
client
  répertoire par défaut 224
close_to
  fonctions spatiales 168
codage 44, 255
codage de texte 44
codage UTF-8 44, 255
codage par défaut du flux 44
collage dans d'autres applications 103
coller 18
columns 126
  modification de la largeur dans les tableaux croisés dynamiques 126
  sélection dans les tableaux croisés dynamiques 126
commentaires
  lister tout sur un flux 63
  raccourcis clavier 249
  sur des noeuds et des flux 59
compatibilité 508 243
concaténation de chaînes 162
conditions 142
connexion à IBM SPSS Modeler Server 10
connexion unique 10
connexion unique, IBM SPSS Collaboration and Deployment Services Repository 191, 192
connexions
  à IBM SPSS Analytic Server 12
  à IBM SPSS Collaboration and Deployment Services Repository 192
  à IBM SPSS Modeler Server 10, 11
  cluster de serveurs 11
conseils
  utilisation générale 69
contrôle du nombre de lignes à afficher 121
conventions 161
coordinateur de processus 11
coordonnées géospatiales
  Format d'affichage 51
  sélection de systèmes 51
COP 11
copie et collage dans d'autres applications 103
copie spéciale 103
copier 18
couches 112, 119, 121, 123
  affichage 119
  création 119
  dans les tableaux croisés dynamiques 119
  impression 112, 121, 123

couleur d'arrière-plan 123
couleurs
définition 226
couleurs dans les tableaux croisés dynamiques 123
bordures 123
couleurs de ligne en alternance
tableaux croisés dynamiques 122
couleurs de script
définition 227
couper 18
création d'une palette personnalisée 233
création de sous-palettes 235
CRISP-DM 18, 215
vue des projets 215
crosses
fonctions spatiales 168

D

dates
conversion 180
Manipulation 180
décimales
formats d'affichage 47
découverte de connaissances 25
défilement
définition des options 49
degrés
unités de mesure 47
déplacement des lignes et des colonnes 116
déploiement 192
désactiver des noeuds 40
descriptions de flux 57, 58
déverrouillage d'objets IBM SPSS Collaboration and Deployment Services Repository 201
distance
fonctions spatiales 168
distribution f
fonctions de probabilité 168
distribution Khi-deux
fonctions de probabilité 168
distribution normale
fonctions de probabilité 168
distribution t
fonctions de probabilité 168
documentation 3
Données
aperçu 43
données d'identification
pour IBM SPSS Collaboration and Deployment Services Repository 193
données parasites 26
dossiers, IBM SPSS Collaboration and Deployment Services Repository 200, 202
DTD 212
durées d'exécution, affichage 52

E

Editeur de données
options relatives aux statistiques descriptives 130
plusieurs fichiers de données ouverts 128
En-têtes 113
enregistrement
états 65
flux 65
noeuds 65
objets de sortie 66
objets multiples 66
Enregistrement de graphiques 105, 111, 112
Fichiers BMP 105, 112
Fichiers EMF 105
Fichiers EPS 105, 112
Fichiers JPEG 105, 111
Fichiers PICT 105
Fichiers PNG 112
Fichiers PostScript 112
Fichiers TIFF 112
Métafichiers 105
enregistrement de sortie 105, 109, 110
au format PDF. 105, 110
Format Excel 105, 108
Format HTML 105
Format PowerPoint 105, 109
Format texte 105, 110
Format Word 105, 107
HTML 105, 106
rapport Web 107
enregistrements 25
Valeurs manquantes 132
valeurs manquantes système 133
ensembles 44
ensembles à catégories multiples
dans les expressions CLEM 146
ensembles à réponses multiples
dans les expressions CLEM 146, 150
ensembles de dichotomies multiples
dans les expressions CLEM 146
ensembles de règles
évaluation 44
entiers 155
erreur dans le rendu de la vue
mémoire insuffisante 223
erreur de dépassement de la capacité de pile 223
espace de travail 14
espace de travail de flux
paramètres 49
espaces
suppression dans les chaînes 143, 171
états
chargement 67
enregistrement 65
exécution des flux 58
exemples
Guide des applications 3
vue d'ensemble 4
exemples d'application 3
exploration de données 25
exemples d'application 35
stratégie 27
exportation
descriptions de flux 58
PMML 212, 213
Exportation de graphiques 105, 111, 112
Exportation de sortie 105, 109, 110
au format PDF. 105, 110
Format Excel 105, 108
Format HTML 105
Format PowerPoint 105
Format Word 105, 107
HTML 106
rapport Web 107
expressions 155
expressions CLEM
performances 241
extraction d'objets à partir de IBM SPSS Collaboration and Deployment Services Repository 198

F

facteur 253
Fenêtre Arbre interactif
accessibilité 253
fenêtre principale 14
fichier de dictionnaire 252
fichiers 102
ajout d'un fichier texte au Visualiseur 102
Fichiers BMP 105, 112
Exportation de graphiques 105, 112
fichiers de données
plusieurs fichiers de données ouverts 128
fichiers de données de texte
codage 255
fichiers de flux de sauvegarde
restauration 65
Fichiers de résultat
enregistrement 66
fichiers de statistiques
codage 255
Fichiers EPS 105, 112
Exportation de graphiques 105, 112
fichiers journaux
affichage du code SQL généré 49
Fichiers JPEG 105, 111
Exportation de graphiques 105, 111
Fichiers PNG 105, 112
Exportation de graphiques 105, 112
Fichiers PostScript (encapsulés) 105, 112
Exportation de graphiques 105, 112
fichiers PowerPoint 216
fichiers SAS
codage 255
Fichiers TIFF 112
Exportation de graphiques 105, 112
flux 9, 14, 204
affichage des durées d'exécution 52
ajout aux projets 216
ajout de commentaires 59
ajout de noeuds 38, 40
ajuster à la vue 21
annotation 59, 64
chargement 67
connexion de noeuds 38
contournement de noeuds 39

flux *(suite)*
création 37
désactivation de noeuds 40
enregistrement 65
Exécution 58
fichiers de sauvegarde 65
options 44, 46, 47, 49, 51
options de déploiement 205
renommer 57, 64
stockage dans IBM SPSS Collaboration and Deployment Services Repository 196
système de coordonnées géospatiales 51
flux de données
création 37
fonction @BLANK 136, 162, 187
fonction @DIFF 180
fonction @FIELD 136, 188
fonction @FIELDS_BETWEEN 136, 145, 188
fonction @FIELDS_MATCHING 136, 145, 188
fonction @INDEX 180
fonction @LAST_NON_BLANK 180, 187
fonction @MAX 180
fonction @MEAN 180
fonction @MIN 180
fonction @MULTI_RESPONSE_SET 146, 188
fonction @NULL 136, 162, 187
fonction @OFFSET 180
remarques sur les performances 241
fonction @PARTITION_FIELD 188
fonction @PREDICTED 188
fonction @SDEV 180
fonction @SINCE 180
fonction @SUM 180
fonction @TARGET 188
fonction @TESTING_PARTITION 188
fonction @THIS 180
fonction @TODAY 176
fonction @TRAINING_PARTITION 188
fonction @VALIDATION_PARTITION 188
fonction (exponentielle) de puissance 166
fonction abs 166
fonction aléatoire 171
fonction allbutfirst 171
fonction allbutlast 171
fonction alphabefore 171
fonction arccos 167
fonction arccosh 167
fonction arcsin 167
fonction arcsinh 167
fonction arctan 167
fonction arctan2 167
fonction arctanh 167
fonction area 168
fonction cdf_chisq 168
fonction cdf_f 168
fonction cdf_normal 168
fonction cdf_t 168
fonction close_to 168
fonction cos 167
fonction cosh 167
fonction count_equal 145, 164
fonction count_greater_than 145, 164
fonction count_less_than 145, 164
fonction count_non_nulls 164
fonction count_not_equal 145, 164
fonction count_nulls 136, 145, 164
fonction count_substring 171
fonction crosses 168
fonction date_before 164
fonction datetime_date 162
fonction DIFF 180
fonction distance 168
fonction div 166
fonction endstring 171
fonction exponentielle 166
fonction first_index 146, 164
fonction first_non_null 146, 164
fonction first_non_null_index 146, 164
fonction fracof 166
fonction hasendstring 171
fonction hasmidstring 171
fonction hasstartstring 171
fonction hassubstring 171
fonction INDEX 180
fonction integer_bitcount 169
fonction integer_leastbit 169
fonction integer_length 169
fonction intof 166
fonction is_date 162
fonction is_datetime 162
fonction is_integer 162
fonction is_number 162
fonction is_real 162
fonction is_string 162
fonction is_time 162
fonction is_timestamp 162
fonction isalphacode 171
fonction isendstring 171
fonction islowercode 171
fonction ismidstring 171
fonction isnumbercode 171
fonction isstartstring 171
fonction issubstring 171
fonction issubstring_count 171
fonction issubstring_lim 171
fonction isuppercode 171
fonction last_index 146, 164
fonction LAST_NON_BLANK 180
fonction last_non_null 146, 164
fonction last_non_null_index 146, 164
fonction length 171
fonction locchar 171
fonction locchar_back 171
fonction log 166
fonction log10 166
fonction lowertoupper 171
fonction matches 171
fonction max 164
fonction MAX 180
fonction max_index 146, 164
fonction max_n 145, 164
fonction MEAN 180
fonction mean_n 145, 166
fonction member 164
fonction min 164
fonction MIN 180
fonction min_index 146, 164
fonction min_n 145, 164
fonction mod 166
fonction negate 166
fonction num_points 168
fonction OFFSET 180
fonction oneof 171
fonction overlap 168
fonction pi 167
fonction random0 171
fonction rem 166
fonction replace 171
fonction replicate 171
fonction round 166
fonction SDEV 180
fonction sdev_n 145, 166
fonction sign 166
fonction sin 167
fonction SINCE 180
fonction sinh 167
fonction skipchar 171
fonction skipchar_back 171
fonction soundex 176
fonction soundex_difference 176
fonction sqrt 166
fonction startstring 171
fonction stripchar 171
fonction strmember 171
fonction subscrs 171
fonction substring 171
fonction substring_between 171
fonction SUM 180
fonction sum_n 145, 166
fonction tan 167
fonction tanh 167
fonction testbit 169
fonction THIS 180
fonction time_before 164
fonction to_date 162, 176
fonction to_dateline 176
fonction to_datetime 162
fonction to_integer 162
fonction to_number 162
fonction to_real 162
fonction to_string 162
fonction to_time 162, 176
fonction to_timestamp 162, 176
fonction trim 171
fonction trim_start 171
fonction trimend 171
fonction undef 187
fonction unicode_char 171
fonction unicode_value 171
fonction uppertolower 171
fonction value_at 146, 164
fonction within 168
fonctions 157, 158, 161, 162, 180
@BLANK 135
@FIELD 147, 188
@GLOBAL_MAX 186
@GLOBAL_MEAN 186
@GLOBAL_MIN 186
@GLOBAL_SDEV 186
@GLOBAL_SUM 186
@PARTITION 188
@PREDICTED 147, 188
@TARGET 147, 188
base de données 147, 148

fonctions *(suite)*
dans les expressions CLEM 147
exemples 139
fonctions définies par l'utilisateur (FDU) 147
traitement des valeurs manquantes 136
fonctions CLEM
aléatoire 171
chaîne 171
comparaison 164
datetime 176
fonctions spéciales 188
global 186
informations 162
logique 166
numérique 166
probabilité 168
séquence 180
sur les bits 169
trigonométrique 167
Valeurs manquantes 136
valeurs nulles et non renseignées 187
fonctions d'heure 157, 158
time_before 164
fonctions d'heure et de date 157, 158
fonctions d'informations 162
fonctions date
date_before 176
date_days_difference 176
date_in_days 176
date_in_months 176
date_in_weeks 176
date_in_years 176
date_months_difference 176
date_weeks_difference 176
date_years_difference 176
fonction @TODAY 176
fonctions datetime
datetime_date 176
datetime_day 176
datetime_day_name 176
datetime_day_short_name 176
datetime_hour 176
datetime_in_seconds 176
datetime_minute 176
datetime_month 176
datetime_month_name 176
datetime_month_short_name 176
datetime_now datetime_second 176
datetime_time 176
datetime_timestamp 176
datetime_weekday 176
datetime_year 176
fonctions de base de données
dans les expressions CLEM 148
fonctions définies par l'utilisateur (FDU) 148
fonctions de comparaison 164
fonctions de conversion 162
fonctions de date 157, 158
date_before 164
Fonctions de distribution 168
fonctions de probabilité 168
fonctions définies par l'utilisateur (FDU) 147, 148
fonctions globales 186
fonctions if, then, else 166
Fonctions logiques 166
fonctions CLEM
Conversion 162
liste des fonctions disponibles 160
spatiales 168
fonctions numériques 166
fonctions séquentielles 180
fonctions spatiales 168
fonctions spéciales 188
fonctions sur chaînes 171
fonctions sur les bits 169
fonctions time
time_before 176
time_hours_difference 176
time_in_hours 176
time_in_mins 176
time_in_secs 176
time_mins_difference 176
time_secs_difference 176
fonctions trigonométriques 167
format d'affichage monétaire 47
Format Excel
Exportation de sortie 105, 108
Format PowerPoint
Exportation de sortie 105
Format Word
Exportation de sortie 105, 107
tableaux larges 105
formats d'affichage
coordonnées géospatiales 51
décimales 47
Monnaie 47
nombres 47
scientifique 47
symbole de regroupement 47
formats d'affichage des nombres 47
formats d'heure 46, 157, 158
formats de coordonnées géospatiales 51
formats de date 46, 157, 158
fractionnement des tableaux 127
contrôle des ruptures de tableau 127

G

générateur de formules 247
accès 147
utilisation 147
vue d'ensemble 147
générateur de graphiques
galerie 72
présentation 71
terminologie 71
génération de scripts 23, 139
génération SQL
consignation 49
prévisualisation 49
gestionnaires 16
graphique à bulles 83
graphique à courbes groupées 75
graphique de tableau 128
graphique de type rivière 93
graphique Q-Q 73
graphiques 73, 99, 105, 128
3D 90
ajout aux projets 216
barre 79
graphiques *(suite)*
barre d'erreur 79, 89
boîte à moustaches 82
bulle 83
carte 84, 85
carte thermique 86
cercles imbriqués 83
circulaires 78
colonne 79
création à partir de tableaux croisés dynamiques 128
en chandeliers 92
enregistrement de sortie 66
exportation 105
graphique Q-Q 73
histogramme 77
ligne 75
ligne de chute 74
masquage 99
Modèles 96
nuage de mots 88
nuage de points 74
parallèle 80
personnalisé 95
polygone des effectifs 77
pyramide de population 77
relation 81
séries multiples 76
t-SNE 87
tableau de bord 95
thème rivière 93
tracé de points 74
tracé de points récapitulatifs 74
tracés horaires 94
graphiques 3D 90
graphiques à barres 79
graphiques à barres 3D 79
graphiques à barres empilées 79
graphiques à barres en cluster 79
graphiques à cercles imbriqués 83
graphiques à colonnes 79
graphiques à courbes 75
ligne de chute 91
lignes de chute 74
tracés horaires 94
graphiques à lignes de chute 74, 91
graphiques à plusieurs lignes 75
graphiques circulaires 78
graphiques de barre d'erreur 79, 89
graphiques de nuages de mots 88
graphiques de parallèles 80
graphiques de séries multiples 76
graphiques de type Carte 84
graphiques en chandeliers 92
graphiques personnalisés 95
graphiques t-SNE 87

H

histogrammes 77
HTML 105, 106
Exportation de sortie 105, 106

I

IBM SPSS Analytic Server
  connexion 12
  connexions multiples 12
IBM SPSS Collaboration and Deployment Services 192
IBM SPSS Collaboration and Deployment Services Repository 191, 192
  accès 193
  connexion 192
  connexion unique 192
  données d'identification 193
  dossiers 200, 202
  extraction d'objets 198
  propriétés de l'objet 202
  recherche 199
  stockage d'objets 193
  suppression d'objets et de versions 201
  transfert de projets 217
  verrouillage et déverrouillage d'objets 201
IBM SPSS Modeler 1, 14
  astuces et raccourcis 69
  démarrage 9
  démarrage à partir de la ligne de commande 9
  documentation 3
  fonctions d'accessibilité 243
  options 223
  vue d'ensemble 9, 223
IBM SPSS Modeler Advantage 192
IBM SPSS Modeler Server 1
  ID utilisateur 10
  nom d'hôte 10, 11
  nom de domaine (Windows) 10
  numéro de port 10, 11
  password 10
icônes
  définition des options 21, 49
ID utilisateur
  IBM SPSS Modeler Server 10
importation
  PMML 212, 213
impression 23, 112, 113, 114, 121, 123, 127
  Aperçu avant impression 113
  contrôle des ruptures de tableau 127
  couches 112, 121, 123
  En-têtes et pieds de page 113
  Espace entre les éléments de sortie 114
  flux 21, 40
  graphiques 112
  mise à l'échelle de tableaux 121, 123
  Numéros de page 114
  sortie de texte 112
  tableaux croisés dynamiques 112
  Taille des graphiques 114
Info-bulles
  annotation de noeuds 64
insertion de libellés de groupe 117
introduction
  IBM SPSS Modeler 9, 223
Introduction 155
invites, exécution 54

J

Java 254
JAWS 243, 252, 254
Justification 100, 129
  sortie 100, 129

K

keywords
  annotation de noeuds 64

L

Langage
  modification de la langue des sorties 118
langue
  options 223
Largeur de colonne 121, 126, 129
  contrôle de la largeur maximale 121
  Contrôle de la largeur par défaut 129
  contrôle de la largeur pour le texte avec renvoi à la ligne 121
  tableaux croisés dynamiques 126
lecteurs d'écran 245, 247, 252, 253
  exemple 251
Légende 101
  Développement 101
  Modification des niveaux 101
  Réduction 101
légendes 124
libellés
  value 212
  variable 212
Libellés 117
  affichage 44
  insertion de libellés de groupe 117
  Suppression 117
libellés, objet IBM SPSS Collaboration and Deployment Services Repository 204
libellés de groupe 117
libellés de valeurs 118, 130
  dans le panneau de légendes 130
  dans les tableaux croisés dynamiques 130
libellés de variable 118, 128, 130
  dans le panneau de légendes 130
  dans les boîtes de dialogue 128
  dans les tableaux croisés dynamiques 130
libellés de version, objet IBM SPSS Collaboration and Deployment Services Repository 204
ligne de commande
  démarrage de IBM SPSS Modeler 9
liste de tous les commentaires sur un flux 63
listes 155, 156
locataire
  IBM SPSS Analytic Server 12

M

mappage de champs 67
mappage des données 69
masquage 99, 100, 119
  légendes 124
  libellés de dimension 119
  Lignes et colonnes 119
  notes de bas de page 124
  Résultats d'une procédure 100
  titres 120
matrices de nuage de points 91
matrices de nuages de points (SPLOM) 74
mémoire
  erreur de dépassement de la capacité de pile 223
  gestion 223, 224
Mémoire 128
mémoire insuffisante 223
messages
  affichage du code SQL généré 49
messages d'erreur 51
Métafichiers 105
  Exportation de graphiques 105
méthodes de sélection 126
  sélection de lignes et de colonnes dans les tableaux croisés dynamiques 126
méthodologie CRISP-DM 27, 28
mise à l'échelle
  tableaux croisés dynamiques 121, 123
Mise en page 113, 114
  En-têtes et pieds de page 113
  Taille des graphiques 114
modèles
  ajout aux projets 216
Modèles 59, 68, 96
  exportation 227
  graphiques 96
  rafraîchissement 207, 208
  remplacement 225
  stockage dans IBM SPSS Collaboration and Deployment Services Repository 197
modèles de rafraîchissement 207, 208
modèles PMML
  régression linéaire 227
  régression logistique 227
modélisation
  branche 59
mot de passe
  IBM SPSS Analytic Server 12

N

navigation
  raccourcis clavier 244
noeud Agréger
  performances 239
noeud Audit données
  utilisation au cours de l'exploration 25
  utilisation dans l'exploration de données 26
noeud Bâtir règle
  chargement 67
noeud Discrétiser
  performances 239
noeud Distinguer
  performances 239

noeud Evaluation
performances 239
noeud fichier cache
chargement 67
noeud Fusionner
performances 239
noeud K moyenne
grands ensembles 44
performances 240
noeud Kohonen
grands ensembles 44
performances 240
noeud Qualité
Valeurs manquantes 132
noeud Remplacer
Valeurs manquantes 135
noeud Réseau de neurones
grands ensembles 44
performances 240
noeud Sélection de fonction
Valeurs manquantes 132
noeud Trier
performances 239
noeud Typer
performances 239
Valeurs manquantes 135
noeuds 9
activation 40
affichage sur la palette 234
ajout 38, 40
ajout aux projets 216
ajout de commentaires sur 59
annotation 59, 64
aperçu des données 43
chargement 67
connexion dans un flux 38
contournement dans un flux 39
création d'une palette personnalisée 233
création de sous-palettes personnalisées 235
définition des options 40
désactivation 40
désactivation dans un flux 40
duplication 40
durées d'exécution 52
enregistrement 65
Introduction 38
modification 40
ordre 237
performances 239, 240
personnalisation d'un onglet palette 235
recherche 56
stockage dans IBM SPSS Collaboration and Deployment Services Repository 197
Suppression 38
suppression de connexions 40
suppression de la palette 234
verrouillage 43
noeuds d'exécution 38
performances 239
noeuds de modélisation 38
performances 240
personnalisation d'un onglet palette de modélisation 235
noeuds de sortie 38
noeuds source 38
mappage des données 68
rafraîchissement 44
noeuds terminaux 38
nom d'hôte
IBM SPSS Modeler Server 10, 11
nom de domaine (Windows)
IBM SPSS Modeler Server 10
nombres 144, 155, 156
noms de flux 64
noms de noeud 64
noms de variable 128
dans les boîtes de dialogue 128
notation scientifique
Format d'affichage 47
Notation scientifique 128
Suppression dans la sortie 128
notes de bas de page 121, 124, 125, 126
marqueurs 121
renumérotation 125
notifications
définition des options 225
nouvelles fonctionnalités 7
nuages de points 74, 91
1D 74
3D 74
groupés 74
matriciels 74
Matriciels 91
simple 74
superposition 74
tracés de points 74
nuages de points en 3D 74
nuages de points groupés 74
nuages de points matriciels 74
nuages de points superposés 74
nuggets 59
défini 16
nuggets de modèle 59
num_points
fonctions spatiales 168
numéro de port
IBM SPSS Modeler Server 10, 11
Numérotation des pages 114

O

objets
propriétés 219
objets de sortie
stockage dans IBM SPSS Collaboration and Deployment Services Repository 197
observation 25
opérateur and 166
opérateur Différent de 164
opérateur Egal à 164
opérateur Inférieur à 164
opérateur not 166
opérateur or 166
opérateur Supérieur à 164
opérateurs
dans les expressions CLEM 147
jonction de chaînes 162
options 128, 129, 130, 223
affichage 226
options *(suite)*
générales 128
Libellés de sortie 130
Modèle de tableau croisé dynamique 129
PMML 227
pour IBM SPSS Modeler 223
propriétés du flux 44, 46, 47, 49, 51, 52
Statistiques descriptives dans l'éditeur de données 130
syntaxe 227
utilisateur 224
Visualiseur 129
options de déploiement 205
options utilisateur 224
ordre d'affichage 116
outil de mappage de données 67, 68
ouverture
états 67
flux 67
Modèles 67
noeuds 67
projets 216
sortie 67
overlap
fonctions spatiales 168

P

palette de modèles 197
palette de modèles générés 16
palettes 14
personnalisation 233
panneau de légende 100
dans le Visualiseur 100
par défaut
phase de projet 215
paramètres
création de modèles 206
dans les expressions CLEM 150
flux 54
invites d'exécution 54
scoring 206
session 54
type 54
paramètres de flux 54
paramètres de session 54
paramètres régionaux
options 223
passage de jours 46
password
IBM SPSS Modeler Server 10
PDF
Exportation de sortie 105, 110
performances
des noeuds d'exécution 239
expressions CLEM 241
noeuds de modélisation 240
personnalisation d'un onglet palette 235
Pieds de page 113
plusieurs fichiers de données ouverts 128
PMML
exportation de modèles 212, 213
importation de modèles 212, 213
options d'exportation 227

point 44
Polices 101, 123, 226, 227
  dans le panneau de légendes 101
polygones des effectifs 77
PowerPoint 109
  exportation de la sortie au format PowerPoint 109
priorité 158
priorité des opérateurs 158
prise en charge d'Unicode 255
programmation visuelle 14
projets 18, 215
  ajout d'objets 216
  annotation 218
  création 216
  dans IBM SPSS Collaboration and Deployment Services Repository 217
  définition d'un dossier par défaut 215
  définition de propriétés 218
  fermeture 219
  génération de rapports 220
  propriétés de l'objet 219
  propriétés du dossier 218
  stockage dans IBM SPSS Collaboration and Deployment Services Repository 196
  Vue Classes 216
  Vue CRISP-DM 215
propriétés 121
  dossier de projet 218
  flux de données 44
  phases dans le rapport 220
  Tableaux 121
  tableaux croisés dynamiques 121
propriétés d'objet, IBM SPSS Collaboration and Deployment Services Repository 202
propriétés de cellule 123, 124
propriétés du flux
  Analytic Server 50
pyramides de population 77

Q

quadrillages 126
  tableaux croisés dynamiques 126

R

raccourcis
  clavier 22, 244, 245, 247, 249
  utilisation générale 69
raccourcis clavier 244, 245, 247, 249
radians
  unités de mesure 47
rafraîchir
  noeuds source 44
rafraîchissement de modèle 205
rapport actif Cognos 107
rapport Web 107
  Exportation de sortie 107
rapports
  ajout aux projets 216
  définition de propriétés 220
recherche
  de noeuds dans un flux 56
recherche d'objets dans IBM SPSS Collaboration and Deployment Services Repository 199
recherche de connexions dans COP 11
rechercher et remplacer
  documents du Visualiseur 102
redimensionnement 20
réduction 20
réels 155, 156
régression 253
régression linéaire
  exportation au format PMML 227
régression logistique 253
  exportation au format PMML 227
remplacement de modèles 225
renommer
  flux 57
  noeuds 64
renvoi à la ligne 121
  contrôle de la largeur des colonnes pour le texte avec renvoi à la ligne 121
réorganisation des lignes et des colonnes 116
répertoire
  par défaut 224
répertoire temporaire 13
rotation des libellés 117
rows 126
  sélection dans les tableaux croisés dynamiques 126
ruptures de tableau 127

S

scoring
  branche 59, 206, 207, 208
sélection de la palette de noeud 234
sélection de système de coordonnées géospatiales 51
serveur
  ajout de connexions 11
  connexion 10
  recherche de serveurs dans COP 11
  répertoire par défaut 224
service de carte 85
sessions multiples IBM SPSS Modeler 13
sortie 16, 99, 100, 103, 105, 114, 129
  affichage 99
  alignement 100, 129
  Centrage 100, 129
  chiffrement 114
  collage dans d'autres applications 103
  Copie 100
  déplacement 100
  enregistrement 114
  exportation 105
  Interactive 104
  masquage 99
  modification de la langue des sorties 118
  suppression 100
  Suppression 100
  Visualiseur 99
sortie HTML
  Lecteur d'écran 253
sortie interactive 104
souris
  utilisation dans IBM SPSS Modeler 22, 38
sous-palette
  affichage dans l'onglet palette 234
  création 235
  suppression de l'onglet palette 234
stockage d'objets dans IBM SPSS Collaboration and Deployment Services Repository 193
suppression de libellés de groupe 117
suppression de sorties 100
Surface
  fonctions spatiales 168
symbole de regroupement
  formats d'affichage des nombres 44
symbole décimal
  formats d'affichage des nombres 44
système
  options 223
Système de mesure 128
Système de production 128
  Utilisation de la syntaxe de commande à partir d'un fichier-journal 128

T

tableau de bord 95
  graphiques 95
tableaux
  ajout aux projets 216
Tableaux 127, 247
  alignement 124
  contrôle des ruptures de tableau 127
  couleur d'arrière-plan 123
  enregistrement de sortie 66
  marges 124
  Polices 123
  propriétés de cellule 123, 124
tableaux classiques 128
tableaux croisés dynamiques 99, 103, 105, 112, 116, 117, 118, 119, 120, 121, 122, 123, 126, 127, 128, 129
  adapter à la page 121, 123
  affichage des bordures masquées 126
  affichage et masquage de cellules 119
  Ajustement de la largeur de colonne par défaut 129
  alignement 124
  annulation des modifications 118
  association de lignes ou de colonnes 117
  bordures 123
  collage dans d'autres applications 103
  collage sous forme de tableaux 103
  contrôle des ruptures de tableau 127
  contrôle du nombre de lignes à afficher 121
  couches 119
  couleur d'arrière-plan 123
  couleurs de ligne en alternance 122

tableaux croisés dynamiques *(suite)*
création de graphiques à partir de tableaux 128
croisement 116
déplacement des lignes et des colonnes 116
dissociation de lignes ou de colonnes 117
Exportation au format HTML 105
formats de cellule 122
Impression de couches 112
impression des grands tableaux 127
insertion de libellés de groupe 117
insertion de lignes et de colonnes 117
langage 118
largeur des cellules 126
légendes 124
libellés de valeurs 118
libellés de variable 118
Manipulation 116
marges 124
masquage 99
Modèle de tableau par défaut des nouveaux tableaux 129
modification 116
modification de l'aspect 120
modification de l'ordre d'affichage 116
notes de bas de page 124, 125, 126
Polices 123
propriétés 121
propriétés de cellule 123, 124
Propriétés des notes de bas de page 121
propriétés générales 121
quadrillages 126
représenter plus rapidement les tableaux 129
rotation des libellés 117
sélection de lignes et de colonnes 126
suppression des libellés de groupe 117
tableaux classiques 128
tableaux croisés dynamiques rapides 129
texte de lignes incomplètes 123
transposition des lignes et des colonnes 116
tri des lignes 117
utilisation d'icônes 116
tableaux croisés dynamiques rapides 129
tableaux de bord
enregistrement de sortie 66
génération 220
tableaux de relation 81
tableaux larges
collage dans Microsoft Word 103
TableLooks 120
application 120
création 120
Tailles 101
Dans la légende 101
texte 101, 102, 105, 110
ajout au Visualiseur 101
texte *(suite)*
ajout d'un fichier texte au Visualiseur 102
Exportation de sortie en tant que texte 105, 110
texte de libellé vertical 117
texte de lignes incomplètes 123
pour les tableaux croisés dynamiques 123
titres 101
ajout au Visualiseur 101
touches de raccourci 22
tracé de points récapitulatifs 74
tracés horaires 94
traitement des blancs
fonctions CLEM 187
transposition des lignes et des colonnes 116
tri
lignes de tableau croisé dynamique 117
type de déploiement 205
types de données 141
dans les paramètres 54

U

URL
IBM SPSS Analytic Server 12

V

valeurs 141
affichage à partir d'un audit de données 151
ajout dans des expressions CLEM 151
valeurs de date/heure 144
valeurs globales
dans les expressions CLEM 150
valeurs manquantes
dans les enregistrements 132
expressions CLEM 136
système 133
Valeurs manquantes 132, 144
valeurs manquantes système
dans les enregistrements 133
valeurs nulles 144
variables 25, 128
Variables
Ordre d'affichage dans les boîtes de dialogue 128
vérification des expressions CLEM 151
verrouillage d'objets IBM SPSS Collaboration and Deployment Services Repository 201
verrouillage des noeuds 43
virgule 44
Visualiseur 99, 100, 101, 114, 129, 130
Affichage des libellés de valeurs 130
Affichage des libellés de variable 130
Affichage des noms de variable 130
Affichage des valeurs de données 130
Déplacement de la sortie 100
Visualiseur *(suite)*
Développement de la ligne de légende 101
Enregistrement du document 114
Espace entre les éléments de sortie 114
Masquage de résultats 99
modification de la police de la légende 101
Modification des niveaux de la ligne de légende 101
Modification des tailles de la légende 101
Options d'affichage 129
panneau de légende 99, 100
panneau de résultats 99
rechercher et remplacer des informations 102
Réduction de la ligne de légende 101
suppression de sorties 100

W

within
fonctions spatiales 168

Z

zoom 18

Imprimé en France