IBM SPSS Modeler 18.4 - Guide de déploiement

Remarque

Avant d'utiliser le présent document et le produit associé, prenez connaissance des informations figurant à la section «Remarques», à la page 31.

Informations produit

Cette édition s'applique à la version 18, édition 4, modification 0 d' IBM® SPSS Modeler et à toutes les éditions et modifications ultérieures, sauf indication contraire dans les nouvelles éditions.

© Copyright International Business Machines Corporation .

Table des matières

Préface v

Chapitre 1. Présentation 1
- IBM SPSS Collaboration and Deployment Services 1
  - Collaboration 1
  - Déploiement 2
- Architecture du système 2
  - IBM SPSS Collaboration and Deployment Services Repository 3
  - IBM SPSS Modeler avec collaboration 4
  - IBM SPSS Deployment Manager 4
  - IBM SPSS Collaboration and Deployment Services Deployment Portal 5
  - IBM SPSS Deployment Manager basé sur un navigateur 5
  - Serveurs d'exécution 6
- Utilisation des fichiers IBM SPSS Modeler 6

Chapitre 2. Définitions de serveurs 7
- Ajout de nouvelles définitions de serveur 7
  - Paramètres de serveur IBM SPSS Modeler 7
- Modification des définitions de serveurs 8

Chapitre 3. Etapes de travaux IBM SPSS Modeler 9
- Utilisation des flux IBM SPSS Modeler 9
  - Configuration du serveur IBM SPSS Modeler 9
- Affichage des propriétés d'un travail IBM SPSS Modeler 9
  - Propriétés d'un travail IBM SPSS Modeler - Général 9
  - Propriétés d'un travail IBM SPSS Modeler - Fichiers de données 12
  - Propriétés d'un travail IBM SPSS Modeler - Vue de données 12
  - Propriétés d'un travail IBM SPSS Modeler - Sources de données ODBC 12
  - Propriétés d'un travail IBM SPSS Modeler - Geo Spatial 14
  - Propriétés d'un travail IBM SPSS Modeler - Paramètres 14
  - Propriétés d'un travail IBM SPSS Modeler - Résultats 14
  - Propriétés d'un travail IBM SPSS Modeler - Importation Cognos 15
  - Propriétés d'un travail IBM SPSS Modeler - Exportation Cognos 15
  - Propriétés d'un travail IBM SPSS Modeler - Importation TM1 existante 16
  - Propriétés d'un travail IBM SPSS Modeler - Exportation TM1 existante 16
  - Propriétés d'un travail IBM SPSS Modeler - Importation TM1 16
  - Propriétés d'un travail IBM SPSS Modeler - Exportation TM1 16
  - Propriétés d'un travail IBM SPSS Modeler - Analytic Server Import 17
  - Propriétés d'un travail IBM SPSS Modeler - Analytic Server Export 17
  - Propriétés d'un travail IBM SPSS Modeler - Notifications 17
- Affichage des flux dans IBM SPSS Modeler 17
- Codes d'achèvement IBM SPSS Modeler 18
- Limites des flux IBM SPSS Modeler 18
- Types de noeud 19
- Exemple de script : Comparaison de modèles champion et challenger 19

Chapitre 4. Service d'évaluation 21
- Limites des flux IBM SPSS Modeler 21

Chapitre 5. Etapes de travail champion/challenger 23

Présentation de champion/challenger....23
Mesures de scoring des modèles....23
Dépendance à l'ordre....23
Informations générales....24
Challengers....25
Sélection des challengers....25
Challengers non valides....26
Sélection des sources de données des challengers....26
Champion....26
Test du champion....27
Fichiers de données....28
Vue de données....29
sources de données ODBC....29
Importation de Cognos....30

Remarques....31
Marques commerciales....32
Dispositions relatives à la documentation du produit....32

Index....35

Préface

IBM SPSS Modeler est le plan de travail d'exploration de données d'entreprise d'IBM. SPSS Modeler aide les entreprises et les organismes à améliorer leurs relations avec les clients et les citoyens grâce à une compréhension approfondie des données. A l'aide des connaissances plus précises obtenues par le biais de SPSS Modeler, les entreprises et les organismes peuvent conserver les clients rentables, identifier les opportunités de vente croisée, attirer de nouveaux clients, détecter les éventuelles fraudes, réduire les risques et améliorer les services gouvernementaux.

L'interface visuelle de SPSS Modeler met à contribution les compétences professionnelles de l'utilisateur, ce qui permet d'obtenir des modèles prédictifs plus efficaces et de trouver des solutions plus rapidement. SPSS Modeler dispose de nombreuses techniques de modélisation, telles que les algorithmes de prévision, de classification, de segmentation et de détection d'association. Une fois les modèles créés, l'utilisateur peut utiliser IBM SPSS Modeler Solution Publisher pour les remettre aux responsables, où qu'ils se trouvent dans l'entreprise, ou pour les transférer vers une base de données.

A propos d'IBM Business Analytics

Le logiciel IBM Business Analytics propose des informations complètes, cohérentes et précises auxquelles les preneurs de décisions peuvent se fier pour améliorer les performances de leur entreprise. Un porte-feuilles étendu de veille économique, d'analyses prédictives, de gestion des performances et de stratégie financières et d'applications analytiques vous offre des informations claires, immédiates et décisionnelles sur les performances actuelles et vous permet de prévoir les résultats futurs. Ce logiciel allie des solutions dédiées à l'industrie, des pratiques ayant fait leur preuve et des services professionnels afin que les organisations de toute taille puissent obtenir la meilleure productivité possible, automatiser leurs décisions en toute confiance et améliorer leurs résultats.

Ce portefeuille intègre le logiciel IBM SPSS Predictive Analytics qui aide les organisations à prévoir les événements à venir et à réagir en fonction des informations afin d'améliorer leurs résultats. Les clients de l'industrie du commerce, de l'éducation et des administrations du monde entier font confiance à la technologie IBM SPSS qui offre un avantage concurrentiel en attirant et fidélisant les clients et en améliorant la base de données de la clientèle tout en diminuant la fraude et en réduisant les risques. En utilisant le logiciel IBM SPSS dans leurs opérations quotidiennes, les organisations deviennent des entreprises prédictives, capables de diriger et d'automatiser les décisions pour répondre aux objectifs commerciaux et obtenir un avantage concurrentiel mesurable. Pour plus d'informations ou pour contacter un interlocuteur IBM, visitez le site http://www.ibm.com/spss.

Assistance technique

L'assistance technique est disponible pour les clients du service de maintenance. Les clients peuvent contacter l'assistance technique pour obtenir de l'aide concernant l'utilisation des produits IBM ou l'installation dans l'un des environnements matériels pris en charge. Pour contacter l'assistance technique, rendez-vous sur le site Web IBM à l'adresse http://www.ibm.com/support. Votre nom, celui de votre société, ainsi que votre contrat de prise en charge vous seront demandés.

Chapitre 1. Présentation

IBM SPSS Collaboration and Deployment Services

IBM SPSS Collaboration and Deployment Services est une application d'entreprise qui permet une utilisation et un déploiement étendus d'un environnement d'analyse prédictive.

IBM SPSS Collaboration and Deployment Services offre un stockage centralisé, sûr et auditable des actifs analytiques, des capacités avancées de gestion et de contrôle des processus d'analyse prédictive, ainsi que des mécanismes perfectionnés de présentation des résultats du traitement analytique aux utilisateurs finals. IBM SPSS Collaboration and Deployment Services offre les avantages suivants :

- la protection de la valeur des actifs analytiques
- l'assurance de conformité avec les exigences réglementaires
- l'amélioration de la productivité des analystes
- la réduction des coûts informatiques pour la gestion des analyses

IBM SPSS Collaboration and Deployment Services vous permet de gérer en toute sécurité les actifs analytiques et de favoriser une plus grande collaboration entre les personnes qui les développent et les utilisateurs. En outre, les fonctions de déploiement assurent que les utilisateurs obtiennent les informations nécessaires pour prendre à temps les mesures appropriées.

Collaboration

La collaboration désigne la capacité à partager et à réutiliser les actifs analytiques de manière efficace. Elle est la clé du développement et de l'implémentation des analyses dans une entreprise.

Les analystes ont besoin d'un emplacement dans lequel placer les fichiers qui seront accessibles aux autres analystes ou aux autres utilisateurs. Cet emplacement a besoin d'une implémentation du contrôle de la version pour que les fichiers gèrent l'évolution de l'analyse. Des dispositifs de sécurité sont nécessaires pour contrôler l'accès aux fichiers et leurs modifications. Et pour finir, un mécanisme de restauration et de sauvegarde est nécessaire pour protéger l'entreprise et éviter la perte de ces ressources essentielles.

Pour répondre à ces besoins, IBM SPSS Collaboration and Deployment Services fournit un référentiel de stockage des ressources qui utilise une hiérarchie de dossiers similaires à celle de la plupart des systèmes de gestion de fichiers. Les fichiers stockés dans le IBM SPSS Collaboration and Deployment Services Repository sont disponibles aux utilisateurs de l'entreprise, si tant est que ces utilisateurs possèdent les droits d'accès appropriés. Afin d'aider les utilisateurs dans la recherche des ressources, le référentiel offre des fonctions de recherche.

Les analystes peuvent utiliser les fichiers du référentiel à partir des applications clientes qui exploitent l'interface de service d'IBM SPSS Collaboration and Deployment Services. Les produits tels que IBM SPSS Statistics et IBM SPSS Modeler permettent une interaction directe avec les fichiers du référentiel. Un analyste peut stocker la version d'un fichier en développement, récupérer cette version à un moment ultérieur et continuer à le modifier jusqu'à sa finalisation et son transfert dans un processus de production. Ces fichiers peuvent inclure des interfaces personnalisées qui exécutent des processus analytiques permettant aux utilisateurs de profiter du travail d'un analyste.

L'utilisation du référentiel protège l'entreprise en fournissant un emplacement central pour les actifs analytiques qui peuvent ainsi être facilement sauvegardés et restaurés. De plus, les autorisations au niveau de l'utilisateur, des fichiers et des libellés de version permettent le contrôle de l'accès aux ressources individuelles. Le contrôle de la version et les libellés de version des objets assurent que les versions appropriées des ressources sont utilisées dans les processus de production. Enfin, les fonctions de consignation offrent la capacité d'effectuer le suivi des modifications apportées aux fichiers et au système.

Déploiement

Pour comprendre pleinement les avantages des analyses prédictives, les actifs analytiques doivent fournir des informations destinées aux décisions d'ordre économique. Le déploiement comble le fossé entre les analyses et les actions prises en fournissant des résultats aux personnes et aux processus selon un planning ou en temps réel.

Dans IBM SPSS Collaboration and Deployment Services,les fichiers individuels stockés dans le référentiel peuvent être inclus dans les **travaux** d'exécution. Les travaux sont une séquence d'exécution des artefacts analytiques et peuvent être créés avec IBM SPSS Deployment Manager. Les résultats de l'exécution peuvent être stockés dans le référentiel, ou dans un système de gestion de fichiers ou être envoyés à des destinataires spécifiques. Les résultats stockés dans le référentiel sont accessibles à tout utilisateur disposant des autorisations nécessaires à l'aide de l'interface IBM SPSS Collaboration and Deployment Services Deployment Portal. Les travaux eux-mêmes peuvent être déclenchés selon un planning défini ou en réponse à des événements du système.

En outre, le service de scoring de IBM SPSS Collaboration and Deployment Services permet aux résultats analytiques provenant de modèles déployés d'être livrés en temps réel lors d'une interaction avec un client. Un modèle analytique configuré pour effectuer un scoring peut combiner des données collectées à partir d'une interaction en cours avec un client et des données historiques pour produire un score qui détermine le cours de l'interaction. Le service lui-même peut être exploité par toute application cliente, permettant la création d'interfaces personnalisées pour la définition du processus.

Les fonctions de déploiement d'IBM SPSS Collaboration and Deployment Services sont conçues pour s'intégrer facilement à l'infrastructure de votre entreprise. L'authentification unique réduit le besoin de fournir manuellement des données d'identification aux différents stades du processus. De plus, le système peut être configuré pour être conforme à la norme FIPS PUB 140-2 (Federal Information Processing Standard Publication 140-2).

Remarque : Si un flux SPSS Modeler contient un noeud qui utilise un type de liste, la branche contenant le noeud ne prend pas en charge le service d'évaluation.

Architecture du système

De manière générale, IBM SPSS Collaboration and Deployment Services consiste en un IBM SPSS Collaboration and Deployment Services Repository unique et centralisé qui sert une variété de clients, et utilisent des serveurs d'exécution pour traiter les actifs analytiques.

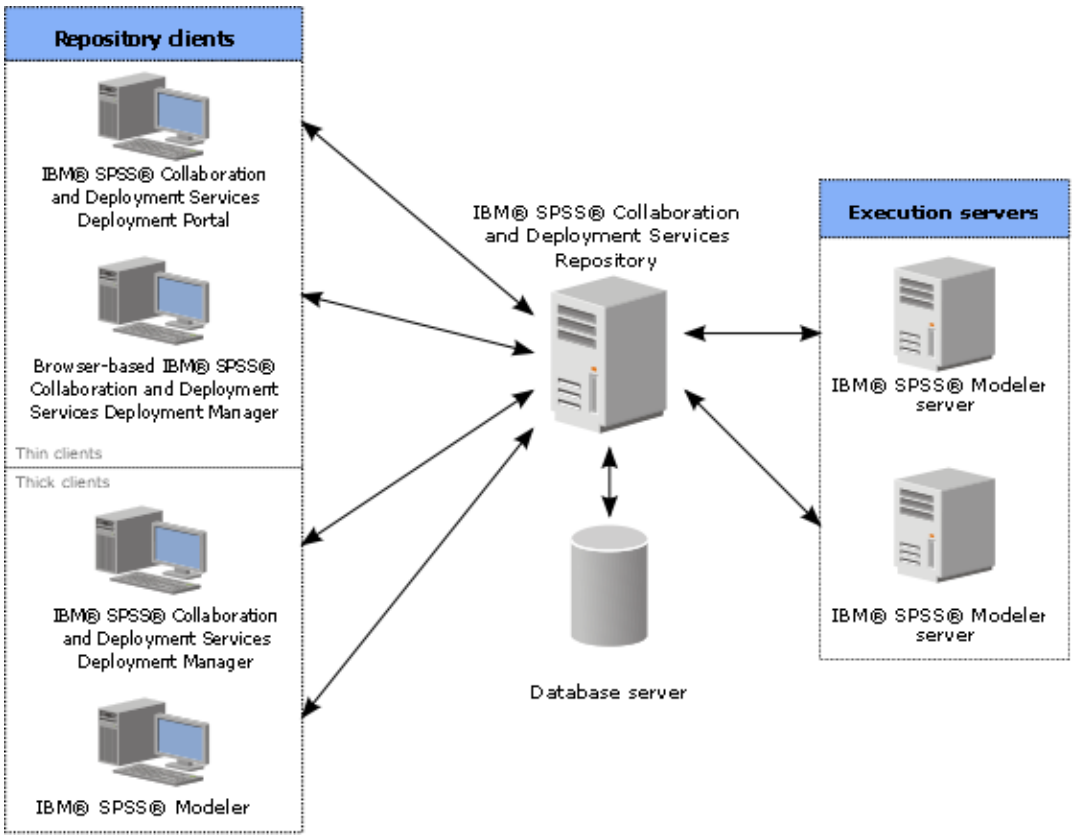

Figure 1. Architecture IBM SPSS Collaboration and Deployment Services

IBM SPSS Collaboration and Deployment Services est composé des composants suivants :

- IBM SPSS Collaboration and Deployment Services Repository pour les artefacts analytiques
- IBM SPSS Modeler
- IBM SPSS Deployment Manager
- IBM SPSS Collaboration and Deployment Services Deployment Portal
- IBM SPSS Deployment Manager (basé sur navigateur)

IBM SPSS Collaboration and Deployment Services Repository

Le référentiel offre un emplacement centralisé de stockage des actifs analytiques, tels que les modèles et les données. Le référentiel requiert l'installation d'une base de données relationnelle, telle qu'IBM Db2, Microsoft SQL Server ou Oracle.

Le référentiel comprend des fonctions destinées à :

- Sécurité
- le contrôle de la version
- la recherche
- l'audit

Les options de configuration du référentiel sont définies à l'aide du IBM SPSS Deployment Manager ou d'IBM SPSS Deployment Manager basé sur un navigateur. La gestion du contenu du référentiel s'effectue avec Deployment Manager, et l'accès au référentiel s'effectue avec IBM SPSS Collaboration and Deployment Services Deployment Portal.

IBM SPSS Modeler avec collaboration

IBM SPSS Modeler avec collaboration permet une interaction avec le IBM SPSS Collaboration and Deployment Services Repository depuis l'interface IBM SPSS Modeler . Les fichiers peuvent être stockés et extraits directement depuis IBM SPSS Modeler .

En outre, IBM SPSS Modeler les scénarios stockés dans le référentiel peuvent être exécutés en tant qu'étapes dans les travaux. Un travail peut contenir un nombre illimité d'étapes, chacune correspondant à un fichier distinct. Les relations définies entres les étapes déterminent le flux du traitement. Le travail peut être planifié afin de s'exécuter à un moment donné, en fonction d'un schéma de récurrence ou en réponse à des événements spécifiques. De plus, il est possible d'envoyer à des destinataires spécifiques des notifications sur le statut global de l'exécution du travail ou des étapes individuelles.

La collaboration entre IBM SPSS Collaboration and Deployment Services et IBM SPSS Modeler est assurée grâce à l'utilisation d'adaptateurs. Ces adaptateurs sont installés dans l'environnement IBM SPSS Collaboration and Deployment Services afin de fournir les fonctionnalités spécifiques à chaque produit. Pour plus d'informations, consultez la documentation IBM SPSS Modeler .

Remarque : Si un travail contenant un flux SPSS Modeler utilise un noeud source Analytic Server, vous devez autoriser la connexion directe entre le serveur IBM SPSS Collaboration and Deployment Services et le serveur Analytic Server. Sinon, le travail échouera si le pare-feu bloque la connexion entre les deux serveurs.

IBM SPSS Deployment Manager

IBM SPSS Deployment Manager est une application client pour IBM SPSS Collaboration and Deployment Services Repository qui permet aux utilisateurs de planifier, d'automatiser et d'exécuter des tâches analytiques comme la mise à jour de modèles ou la génération de scores.

L'application client permet à un utilisateur d'effectuer les tâches suivantes :

- Visualiser tous les fichiers existants du système, y compris les rapports IBM SPSS Modeler, les scénarios et les fichiers de données ;
- Importer des fichiers dans le référentiel
- Planifier une exécution répétée des travaux à l'aide d'un schéma de récurrence spécifique, par exemple tous les trimestres ou toutes les heures
- Modifier les propriétés des travaux existants
- Déterminer le statut d'un travail
- Spécifier une notification par courrier électronique du statut d'un travail

En outre, l'application client permet aux utilisateurs de réaliser des tâches administratives pour IBM SPSS Collaboration and Deployment Services, notamment :

- Gérer les utilisateurs
- Configurer des fournisseurs de sécurité
- Affecter des rôles et des actions

IBM SPSS Deployment Manager (basé sur navigateur)

IBM SPSS Deployment Manager basé sur un navigateur est une interface client léger permettant d'effectuer des tâches de configuration et de gestion du système, notamment :

- Définir des options de configuration système
- Configuration des fournisseurs de sécurité
- Gérer des types MIME

Les utilisateurs ne disposant pas de droits administratifs peuvent réaliser ces tâches, à condition que les actions appropriées soient associées à leurs informations d'identification et de connexion. Ces actions sont affectées par un administrateur.

En règle générale, vous accédez à IBM SPSS Deployment Manager basé sur un navigateur à l'adresse URL suivante :

```
http://<host IP address>:<port>/security/login
```

Remarque : Une adresse IPv6 doit être placée entre crochets, par exemple `[3ffe:2a00:100:7031::1]`.

Si votre environnement est configuré pour utiliser un chemin de contexte personnalisé pour les connexions de serveur, incluez ce chemin dans l'URL.

```
http://<host IP address>:<port>/<context path>/security/login
```

IBM SPSS Collaboration and Deployment Services Deployment Portal

IBM SPSS Collaboration and Deployment Services Deployment Portal est une interface pour client léger permettant d'accéder au référentiel. À la différence du IBM SPSS Deployment Manager basé sur un navigateur, lequel est destiné aux administrateurs, IBM SPSS Collaboration and Deployment Services Deployment Portal est un portail Web utile à une large gamme d'utilisateurs.

Le portail Web comprend les fonctionnalités suivantes :

- Navigation par dossier dans le référentiel de contenu
- Ouverture du contenu publié
- Exécution de travaux
- Génération de scores à l'aide de modèles stockés dans le référentiel
- Recherche dans le contenu du référentiel
- Affichage des propriétés du contenu
- Accès aux préférences des utilisateurs individuels, par exemple l'adresse email et le mot de passe, les options générales, les abonnements et les options des formats des fichiers de sortie

Vous accédez à la page d'accueil par le biais de l'URL suivante :

```
http://<host IP address>:<port>/peb
```

Remarque : Une adresse IPv6 doit être placée entre crochets, par exemple `[3ffe:2a00:100:7031::1]`.

Si votre environnement est configuré pour utiliser un chemin de contexte personnalisé pour les connexions de serveur, incluez ce chemin dans l'URL.

```
http://<host IP address>:<port>/<context path>/peb
```

IBM SPSS Deployment Manager basé sur un navigateur

IBM SPSS Deployment Manager basé sur un navigateur est une interface client léger permettant d'effectuer des tâches de configuration et de gestion du système, notamment :

- la configuration du système
- la configuration des fournisseurs de sécurité
- la gestion des types MIME

Les utilisateurs ne disposant pas de droits administratifs peuvent réaliser ces tâches, à condition que les actions appropriées soient associées à leurs informations d'identification et de connexion. Ces actions sont affectées par un administrateur.

Serveurs d'exécution

Les serveurs d'exécution permettent d'exécuter les ressources stockées dans le référentiel. Lorsqu'une ressource est incluse dans un travail à exécuter, la définition de l'étape du travail inclut la spécification du serveur d'exécution utilisé pour le traitement de l'étape. Le type du serveur d'exécution dépend de la ressource.

Le serveur d'exécution IBM SPSS Modeler est IBM SPSS Modeler Server, il permet de réaliser des analyses distribuées pour l'exploration de données et la création de modèles. Ce serveur d'exécution requiert la spécification des données d'identification utilisateur sous lesquelles le traitement se produit.

Pour permettre l'équilibre de charge, il est possible de regrouper plusieurs serveurs d'exécution ensemble dans un cluster. Lorsqu'une étape de travail utilise un cluster lors d'une exécution, IBM SPSS Collaboration and Deployment Services détermine quel serveur géré du cluster est le plus à même de traiter les requêtes à ce moment précis. Pour plus d'informations, consultez la documentation d'IBM SPSS Deployment Manager.

Au cours de la création du travail, affectez un serveur d'exécution ou un cluster de serveurs à chaque étape du travail. Lors de l'exécution du travail, le référentiel utilise les serveurs d'exécution spécifiés pour réaliser les analyses correspondantes.

Utilisation des fichiers IBM SPSS Modeler

Dans IBM SPSS Deployment Manager, le processus général d'utilisation des fichiers IBM SPSS Modeler implique :

1. Définition d'un serveur d'exécution pour le traitement du flux ou du scénario IBM SPSS Modeler si une définition appropriée n'existe pas déjà.
2. Ajout du du flux ou du scénario IBM SPSS Modeler à un travail en tant qu'étape du travail.
3. La spécification des propriétés du travail, notamment les planifications de l'exécution.

Chapitre 2. Définitions de serveurs

L'exécution d'une ressource d'IBM SPSS Collaboration and Deployment Services Repository en tant qu'étape de travail nécessite de spécifier un serveur correspondant approprié pour traiter les instructions contenues dans l'étape de travail. Les informations de connexion d'un tel serveur sont spécifiées dans une définition de serveur.

Les définitions de serveur, tout comme les serveurs de référentiel et les serveurs d'exécution, peuvent aussi être classifiées.

- Les serveurs d'exécution traitent le contenu d'une ressource du IBM SPSS Collaboration and Deployment Services Repository. Le type du serveur d'exécution doit correspondre au type de ressource étant traité. An IBM SPSS Modeler par exemple une définition de serveur an IBM SPSS Modeler .
- Un serveur de référentiel correspond à une installation de référentiel IBM SPSS Collaboration and Deployment Services. Un serveur de ce type est généralement utilisé par les étapes de travaux qui ont besoin de retourner des artefacts de résultats à un référentiel.

Les définitions de serveur sont contenues dans le dossier *Définitions de ressource* de l'explorateur de contenu. Plus précisément, elles sont définies dans le sous-dossier *Serveurs*.

Ajout de nouvelles définitions de serveur

Pour ajouter un nouveau serveur :

1. Dans l'explorateur de contenu, ouvrez le dossier *Définitions de ressources*.
2. Cliquez sur le dossier *Serveurs*.
3. Dans le menu Fichier, sélectionnez :

Nouveau > **Définition de serveur**

L'assistant Ajouter une nouvelle définition de serveur s'ouvre. Il est également possible d'accéder à la boîte de dialogue Nouvelle définition de serveur en cliquant sur **Nouveau** en regard d'un champ de serveur sur l'onglet Général pour certaines étapes. La procédure pour définir de nouveaux serveurs consiste en plusieurs actions :

1. La dénomination de la définition de serveur et la spécification de son type. Veuillez noter que les types de serveur disponibles dépendent des adaptateurs de produits installés sur le référentiel.
2. La sélection d'un emplacement dans le dossier *Serveurs* pour la définition.
3. La spécification des paramètres du serveur qui définissent les informations de connexion ou d'exécution. La définition des paramètres dépend du type de serveur.

Paramètres de serveur IBM SPSS Modeler

Une définition de serveur IBM SPSS Modeler spécifie les paramètres de connexion pour les serveurs IBM SPSS Modeler utilisés pour traitement des étapes de travail. Les flux IBM SPSS Modeler sont exécutés sur le serveur IBM SPSS Modeler.

1. Dans le champ Hôte, entrez le nom de l'hôte où se trouve le serveur. Par exemple, si vous créez une définition de serveur IBM SPSS Modeler, l'hôte est l'ordinateur qui contient votre serveur IBM SPSS Modeler.
2. Dans le champ Port, entrez le numéro de port à utiliser pour se connecter à l'hôte.
3. Dans le champ Chemin d’accès aux données par défaut, entrez le chemin sur lequel vous souhaitez placer les fichiers de données.
4. Si le protocole SSL (Secure Socket Layer) doit être utilisé pour la connexion au serveur, sélectionnez **Ceci est un port sécurisé**.
5. Cliquez sur **Terminer**. La nouvelle définition apparaît dans le dossier *Serveurs*.

Remarque : Lorsque vous exécutez un travail IBM SPSS Collaboration and Deployment Services dans un flux d'évaluation, 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.

Modification des définitions de serveurs

Pour modifier une définition de serveur :

1. Dans l'explorateur de contenu, ouvrez le dossier *Définitions de ressources*.
2. Ouvrez le dossier *Serveurs*.
3. Double-cliquez sur le serveur pour le modifier. La boîte de dialogue Modifier la définition du serveur s'ouvre.
4. Modifiez les paramètres de définition du serveur comme il se doit.
5. Cliquez sur Terminer pour enregistrer les changements.

Chapitre 3. Etapes de travaux IBM SPSS Modeler

Utilisation des flux IBM SPSS Modeler

Cette section décrit les flux IBM SPSS Modeler dans le contexte d'IBM SPSS Deployment Manager. Les flux IBM SPSS Modeler sont intégrés dans Deployment Manager entièrement formés. Pour obtenir des informations supplémentaires sur la création et l'utilisation des flux, reportez-vous à la documentation de IBM SPSS Modeler.

Les flux IBM SPSS Modeler peuvent être utilisés dans Deployment Manager. Comme toutes les autres étapes, un flux IBM SPSS Modeler doit d'abord être ajouté à un travail avant de pouvoir l'exécuter dans Deployment Manager. Vous pouvez en particulier effectuer les tâches suivantes :

- importation des flux
- modification des paramètres de flux
- Exécution des flux

Remarque : Si un travail contenant un flux SPSS Modeler utilise un noeud source Analytic Server, vous devez autoriser la connexion directe entre le serveur IBM SPSS Collaboration and Deployment Services et le serveur Analytic Server. Sinon, le travail échouera si le pare-feu bloque la connexion entre les deux serveurs.

Configuration du serveur IBM SPSS Modeler

Avant de commencer à utiliser les flux IBM SPSS Modeler dans Deployment Manager, vous devez effectuer les tâches de configuration suivantes :

- Créer une définition de serveur IBM SPSS Modeler. Pour plus d'informations, voir «Paramètres de serveur IBM SPSS Modeler», à la page 7.
- Définir les données d'identification du serveur.

Affichage des propriétés d'un travail IBM SPSS Modeler

Lorsque vous cliquez sur un flux IBM SPSS Modeler dans un travail, les propriétés de travail suivantes apparaissent :

- Général
- Fichiers de données
- Sources de données ODBC
- Paramètres
- Résultats
- Importation Cognos
- Exportation Cognos
- Notifications

Propriétés d'un travail IBM SPSS Modeler - Général

Par défaut, l'onglet des propriétés Général apparaît lorsque vous cliquez sur un flux IBM SPSS Modeler dans l'espace de travail.

L'onglet de propriétés Général comporte les informations suivantes :

Nom de l'étape de travail. Le nom de l'étape de travail. En général, le nom de l'étape de travail est le nom du flux IBM SPSS Modeler , ajouté au suffixe `_step`. Toutefois, ce nom peut être modifié.

Afin de modifier le nom de l'étape de travail, entrez le nom désiré dans le champ **Nom de l'étape de travail**. Vos modifications apparaissent dans l'espace de travail.

IBM SPSS Modeler Stream. Le nom et le chemin d'accès du flux IBM SPSS Modeler d'origine. Le flux source n'est pas modifiable.

Type de remplacement Utilisez cette option pour indiquer si les paramètres d'une étape de travail ou du noeud d'un flux sont prioritaires lors de l'exécution.

- **Le travail remplace le flux** Si vous sélectionnez cette option, lorsque vous exécutez le travail, les propriétés de noeud utilisent les valeurs que vous définissez dans l'étape du travail.
- **Le flux remplace le travail** Si vous sélectionnez cette option, les commandes des propriétés de noeud sont désactivées dans l'étape du travail, à l'exception de certaines commandes de données d'identification. Si vous modifiez les propriétés de noeud dans le flux, lorsque vous exécutez l'étape du travail, les propriétés de noeud modifiées sont utilisées.

Les options **Type de remplacement** affectent les paramètres des données d'identification comme suit :

- Si vous choisissez **Le travail remplace le flux**, vous pouvez éditer les commandes de sélection des données d'identification ; vos données d'identification sélectionnées sont appliquées lors de la phase d'exécution.
- Si vous choisissez **Le flux remplace le travail**, deux cas sont possibles :
  - Lorsque vous créez des noeuds dans SPSS Modeler Client qui utilisent des données d'identification, si le mode d'authentification est **Données d'identification stockées** et que le nom des données d'identification existe dans IBM SPSS Collaboration and Deployment Services, les commandes de sélection des données d'identification ne sont pas disponibles. Toutefois, si le nom des données d'identification n'existe pas dans IBM SPSS Collaboration and Deployment Services, vous pouvez utiliser la commande de sélection des données d'identification ; vos données d'identification sélectionnées sont appliquées lors de la phase d'exécution.
  - Lorsque vous créez des noeuds dans SPSS Modeler Client qui utilisent des données d'identification, si le mode d'authentification est **Nom d'utilisateur et mot de passe**, vous pouvez utiliser les commandes de sélection des données d'identification ; vos données d'identification sélectionnées sont appliquées lors de la phase d'exécution.

Version d'objet. La version avec libellés du flux IBM SPSS Modeler à utiliser. Dans la liste déroulante **Version d'objet**, sélectionnez la version étiquetée que vous voulez utiliser.

Liste des variables itératives. Si l'étape se comporte comme un consommateur itératif, identifiez la variable dans le producteur itératif fournissant les valeurs pour l'étape. Une étape de production itérative doit apparaître immédiatement avant l'étape dans le travail, afin que l'étape se comporte comme un consommateur itératif.

Serveur IBM SPSS Modeler. Le serveur ou grappe de serveurs IBM SPSS Modeler sur lequel le flux sera exécuté. La liste contient tous les serveurs et les clusters de serveurs actuellement configurés pour exécuter les étapes IBM SPSS Modeler. Pour modifier le serveur, procédez à une sélection dans la liste déroulante **IBM SPSS Modeler Server**. Pour créer une nouvelle définition de serveur, cliquez sur **Nouveau** pour lancer l'assistant de définition de serveur.

Connexion à IBM SPSS Modeler. Les informations d'identification utilisées pour accéder au serveur ou grappe de serveurs IBM SPSS Modeler. Pour modifier les données d'identification, procédez à une sélection dans la liste déroulante **Nom de connexion d'IBM SPSS Modeler**. Pour définir de nouvelles données d'identification, cliquez sur **Nouveau** pour lancer l'assistant de définition de données d'identification.

Serveur du référentiel de contenu. Le serveur du référentiel de contenu permet aux travaux d'enregistrer des fichiers dans un IBM SPSS Collaboration and Deployment Services Repository. Généralement, le serveur de référentiel de contenu est spécifié lors de l'actualisation de modèles à l'aide d'IBM SPSS Modeler. Pour spécifier un serveur de référentiel de contenu, sélectionnez un serveur dans la liste déroulante **Serveur de référentiel de contenu**. Pour créer une nouvelle définition de serveur, cliquez sur **Nouveau** pour lancer l'assistant de définition de serveur. Pour générer une définition de serveur de référentiel de contenu basée sur des informations de serveur actuelles, cliquez sur **Générer**. Une

définition de serveur est créée et automatiquement renseignée dans le champ *Serveur de référentiel de contenu*.

Connexion au référentiel de contenu. Les informations de connexion du serveur de référentiel de contenu. Pour spécifier un nom de connexion au référentiel de contenu, sélectionnez des données d'identification dans la liste déroulante **Connexion au référentiel de contenu**. Pour créer une nouvelle connexion, cliquez sur **Nouveau** pour lancer l'assistant de connexion au référentiel de contenu. Si l'authentification unique n'est pas utilisée pour vous connecter à IBM SPSS Collaboration and Deployment Services Repository, cliquez sur **Générer** pour générer une connexion au serveur de référentiel de contenu basée sur des paramètres de sécurité existants. Une connexion au référentiel de contenu est créée et automatiquement renseignée dans le champ *Connexion de référentiel de contenu*. La génération de connexion n'est pas disponible lors de l'utilisation d'une authentification unique.

Expression d'avertissement. Définissez des avertissements pour les étapes de travail connectées par un connecteur conditionnel. L'expression d'avertissement (par exemple, `code_achèvement`, `avertissement` ou `succès`) doit être en minuscules.

Pour utiliser des expressions d'avertissement :

1. Connectez deux étapes de travail avec un connecteur conditionnel. Dans le champ **Expression** du connecteur conditionnel, saisissez `avertissement==true (vrai)`.
2. Accédez à l'onglet Général de l'étape de travail parent.
3. Dans le champ **Expression d'avertissement**, spécifiez un code d'avertissement, par exemple, `code_achèvement==18`. Cette expression écrase le code d'avertissement par défaut, le cas échéant.

Lorsque le travail est exécuté, le système exécute l'étape de travail parent. Puis le système évalue la condition de `avertissement==true (vrai)`. Si l'avertissement est vrai, le système examine l'expression d'avertissement spécifiée et détermine si la condition est remplie. Si c'est la cas, le système passe à l'étape de travail suivante.

Type. La façon dont vous souhaitez exécuter le flux.

- **Exécuter le flux.** Avec cette option, vous pouvez exécuter le flux complet ou, si le flux contient des branches, choisir une ou plusieurs branches à exécuter. Dans ce cas, la liste **Exécuter les options** apparaît. Toutes les branches sont vérifiées par défaut, ce qui signifie que toutes les branches seront exécutées lorsque le flux sera exécuté. Désélectionnez les branches que vous ne souhaitez pas exécuter, ce qui pourrait améliorer les performances.

  Remarque : Si le flux contient un script et est configuré pour exécuter le script lors de l'exécution, aucune branche d'exécution ne s'affiche dans la liste de l'éditeur de travaux IBM SPSS Collaboration and Deployment Services.
- **Gestion des modèles.** Choisissez cette option si vous souhaitez que le flux utilise les fonctionnalités de gestion des modèles, telles que les fonctionnalités d'évaluation, de rafraîchissement et de scores. Choisissez la fonctionnalité dans la liste Types affichée. **Options d'évaluation**

  Performances. Spécifiez les pourcentages des seuils des catégories suivantes : Mauvais ou Bon. Le pourcentage correspondant à Meilleur n'est pas défini par l'utilisateur. Le système soustrait le Bon pourcentage de 100 pour obtenir le seuil Meilleur. Les valeurs de seuils sont obligatoires pour les flux d'évaluation.

  Mesure. Les critères de mesure permettant d'évaluer l'efficacité du flux. Les valeurs valides sont Accréditation, Exactitude et Gains. Si la valeur **Gains** est sélectionnée, un centile doit être également spécifié. Si la valeur **Accréditation** est sélectionnée, l'étape d'accréditation doit être également spécifiée. Les valeurs valides incluent « Collectionner les statistiques » et « Exécuter l'évaluation ».

  Options de rafraîchissement

  Répertorie toutes les relations modèle-nugget du flux pouvant être rafraîchies à l'aide des noms de noeuds définis dans le flux. Sélectionnez une relation spécifique à rafraîchir en cochant la case correspondante. Décochez les relations que vous ne souhaitez pas rafraîchir lors de l'exécution de l'étape. Pour rafraîchir toutes les relations, cliquez sur le bouton Tout sélectionner. Pour désélectionner toutes les relations, cliquez sur le bouton Tout désélectionner.

Options d'évaluation

Choisissez la branche ou les branches à désigner comme branches d'évaluation.

Instructions pour saisir les informations relatives à IBM SPSS Modeler et au référentiel de contenu

Toutes les étapes de travail IBM SPSS Modeler requièrent à la fois les informations relatives au serveur IBM SPSS Modeler et les informations de connexion à IBM SPSS Modeler.

Propriétés d'un travail IBM SPSS Modeler - Fichiers de données

Le tableau Fichiers de données comporte les informations suivantes :

Nom du noeud. Le nom du noeud d'entrée qui contient les données utilisées par le flux. Ce nom se voit ajouté en préfixe les noms des super noeuds contenant le noeud, séparés par des barres obliques. Par exemple, si le noeud *Monnoeud* est situé dans un super noeud nommé *Super noeud 1*, le nom qui apparaît est */Super noeud 1/Monnoeud*. Le nom du noeud n'est pas modifiable.

Type de noeud. Le type de noeud tel qu'il est défini dans le flux. Le type du noeud n'est pas modifiable.

Nom du fichier. Le nom du fichier de données d'entrée. Pour modifier le nom, cliquez dans la cellule Nom de fichier et modifiez le nom.

Format. Le format du fichier de sortie, par exemple, un fichier csv. Pour modifier le type du format de fichier, cliquez sur la cellule Format. Une flèche de menu déroulant apparaît. Sélectionnez le type de format.

Emplacement. L'emplacement des fichiers de données d'entrée. Pour modifier l'emplacement, cliquez dans la colonne puis cliquez sur le bouton des points de suspension. La boîte de dialogue Emplacement du fichier d'entrée apparaît. Modifiez l'emplacement.

Les noeuds situés dans les super noeuds verrouillés ne sont pas accessibles. Ils ne peuvent être ni affichés ni modifiés.

Propriétés d'un travail IBM SPSS Modeler - Vue de données

Si le flux contient un noeud Geo Spatial, les détails de connexion sont affichés ici.

Nom du noeud. Nom du noeud Vue de données.

Vue de données analytique. Vue de données analytique utilisée.

Étiquette. Libellé utilisé.

Nom de la table. Nom de la table de base de données utilisée.

Plan d'accès aux données. Sélectionnez un plan d'accès aux données dans la vue de données analytiques. Un plan d'accès aux données associe les tables du modèle de données figurant dans une vue de données analytiques à des sources de données physiques. Une vue de données analytiques contient généralement plusieurs plans d'accès aux données. Lorsque vous modifiez le plan d'accès aux données utilisé, vous modifiez également les données utilisées par votre flux. Par exemple, si la vue de données analytiques contient un plan d'accès aux données pour la formation d'un modèle et un plan d'accès aux données pour le test d'un modèle, vous pouvez passer des données de formation aux données de test en modifiant le plan d'accès aux données utilisé.

Propriétés d'un travail IBM SPSS Modeler - Sources de données ODBC

Le tableau *Sources de données ODBC* comporte les informations suivantes.

Nom du noeud. Le nom du nœud d'entrée qui contient les données utilisées par le flux. Ce nom est précédé des noms des super noeuds contenant le noeud, séparés par des barres obliques. Par exemple, si le noeud *MonNoeud* se trouve dans un super noeud intitulé *Supernoeud1*, le nom apparaît sous la forme */Supernoeud1/MonNoeud*.

Type du noeud. Le type de nœud tel qu'il est défini dans le flux.

Sources de données ODBC. Nom de la source de données ODBC (DSN). Pour passer à une autre source de données ODBC, cliquez sur la cellule contenant le nom de la source de données actuelle, puis cliquez sur le bouton « ... » qui s'affiche. Cette action entraîne l'apparition d'une boîte de dialogue sur laquelle vous pouvez choisir une DSN ou en créer une nouvelle. Notez que les paramètres de l'étape du travail ont toujours priorité sur le flux. Par conséquent, si vous modifiez le flux pour utiliser une source de données différente, vous devez également modifier le travail pour utiliser la même source, sinon le travail risque de ne pas s'exécuter.

Données d'authentification. Pour modifier le nom d'utilisateur et le mot de passe de la base de données lors de la modification de la source de données ODBC, cliquez sur la cellule contenant les données d'identification en cours, puis cliquez sur le bouton « ... » qui s'affiche. Cette action entraîne l'apparition d'une boîte de dialogue dans laquelle vous pouvez choisir une définition de données d'identification existante ou en créer une nouvelle.

Table de base de données. Table de base de données qui correspond au noeud.

Les noeuds situés dans des super noeuds verrouillés ne sont pas accessibles. Ils ne peuvent être ni affichés ni modifiés.

Modification de la connexion ODBC

Vous pouvez remplacer la connexion ODBC soit par une connexion déjà définie, soit par une connexion que vous créez ici.

Utiliser un DSN ODBC existant. Cliquez sur **Parcourir** pour afficher une liste des DSN ODBC existants parmi lesquels vous pouvez choisir.

Créer un nouveau DSN ODBC. Sélectionnez cette option et cliquez sur **Nouveau** pour définir une nouvelle connexion ODBC à utiliser avec les étapes de travaux de flux.

Recherche d'une connexion ODBC

Il s'agit de la liste de toutes les connexions ODBC définies sur cet hôte et auxquelles vous avez accès.

Choisissez-en une dans la liste et cliquez sur **OK**.

Modification des données d'identification pour l'accès à la base de données

Lorsque vous modifiez la connexion, vous pouvez opter pour l'utilisation du nom d'utilisateur et du mot de passe de la base de données ici. Vous pouvez choisir d'utiliser une définition de données d'identification existante dans le référentiel ou de créer une nouvelle définition.

Générer des données d'identification de référentiel à partir d'un nom d'utilisateur et d'un mot de passe existants. (ne s'affiche que lors de l'importation d'un travail à partir de la version 4.2.1 IBM SPSS Collaboration and Deployment Services) Cliquez sur le bouton **Générer** pour générer des données d'identification de référentiel à partir d'un nom d'utilisateur et d'un mot de passe existants si ces derniers ont été définis dans le travail importé. Par exemple, si le nom d'utilisateur `sa` existe déjà, le fait de cliquer plusieurs fois sur **Générer** va créer les noms d'utilisateur `sa1`, `sa2` et ainsi de suite.

Utiliser des données d'identification existantes. Cliquez sur le bouton **Parcourir** pour afficher une liste des définitions de données d'identification existantes parmi lesquelles vous pouvez choisir.

Créer de nouvelles données d'identification. Sélectionnez cette option et cliquez sur **Nouveau** pour créer une nouvelle définition de données d'identification à utiliser avec la connexion actuelle.

Recherche d'une définition de données d'identification

Il s'agit de la liste de toutes les définitions de données d'identification créées sur cet hôte et auxquelles vous avez accès.

Choisissez-en une dans la liste et cliquez sur **OK**.

Propriétés d'un travail IBM SPSS Modeler - Geo Spatial

Si le flux contient un noeud Geo Spatial, les détails de connexion sont affichés ici.

Type de source. Type de source de données.

Nom du fichier. Si vous utilisez un fichier de forme, indiquez le nom du fichier.

URL du service de carte. Si vous utilisez un service de carte, entrez l'adresse URL du service.

ID carte. Spécifiez l'ID carte.

Propriétés d'un travail IBM SPSS Modeler - Paramètres

Vous pouvez modifier les paramètres des flux IBM SPSS Modeler à l'aide de IBM SPSS Deployment Manager. Le tableau Paramètres comporte les informations suivantes :

Nom. Le nom du paramètre.

Stockage. Décrit la façon dont le paramètre est stocké, par exemple, sous forme de chaîne.

Valeur. La valeur du paramètre.

Type. Le type du paramètre.

Les paramètres situés dans les super noeuds verrouillés ne sont pas accessibles. Ils ne peuvent être ni affichés ni modifiés.

Propriétés d'un travail IBM SPSS Modeler - Résultats

Pour afficher les paramètres de sortie d'une étape de travail IBM SPSS Modeler, cliquez sur l'onglet Résultats.

Nom du noeud. Le nom du noeud qui contient la sortie provenant du traitement du flux. Ce nom se voit ajouté en préfixe les noms des super noeuds contenant le noeud, séparés par des barres obliques. Par exemple, si le noeud *MonNoeud* est situé dans un super noeud nommé*Super noeud 1*, le nom qui apparaît est */Super noeud 1/MonNoeud*. Le nom du noeud n'est pas modifiable.

Type de noeud. Le type de noeud tel qu'il est défini dans le flux. Le type du noeud n'est pas modifiable.

Nom du fichier. le nom du fichier correspondant. Pour modifier ce nom, cliquez sur la colonne Nom de fichier et entrez le nouveau nom.

Remarque : L'extension du fichier est déterminée par le format de fichier sélectionné et est automatiquement ajoutée au nom du fichier.

Format. le format du fichier de sortie. Pour modifier le type du format de fichier, cliquez dans la colonne Format. Une flèche de menu déroulant apparaît. Sélectionnez le type de format.

Emplacement. L'emplacement du fichier. Pour modifier l'emplacement, ouvrez la boîte de dialogue Emplacement des résultats en cliquant sur la colonne puis cliquez sur le bouton des points de suspension.

Droits d'accès. Droits d'accès des fichiers si ceux-ci sont enregistrés dans le référentiel. Pour modifier les droits d'accès, ouvrez la boîte de dialogue Droits d'accès de sortie en cliquant sur la colonne Droits d'accès puis cliquez sur le bouton des points de suspension.

Propriétés. Les propriétés (métadonnées) du fichier. Pour définir les propriétés, ouvrez la boîte de dialogue Propriétés de sortie en cliquant sur la colonne Propriétés puis cliquez sur le bouton des points de suspension.

Les noeuds situés dans les super noeuds verrouillés ne sont pas accessibles. Ils ne peuvent être ni affichés ni modifiés.

Affichage des résultats de sortie

Pour afficher les résultats de sortie de votre flux, cliquez deux fois sur les résultats que vous souhaitez afficher dans la colonne Résultats du tableau Historique d'étape de travail. Les résultats s'affichent dans une fenêtre distincte.

Par exemple, supposons que l'un de vos résultats de sortie est un histogramme. Un histogramme peut ressembler à ceci.

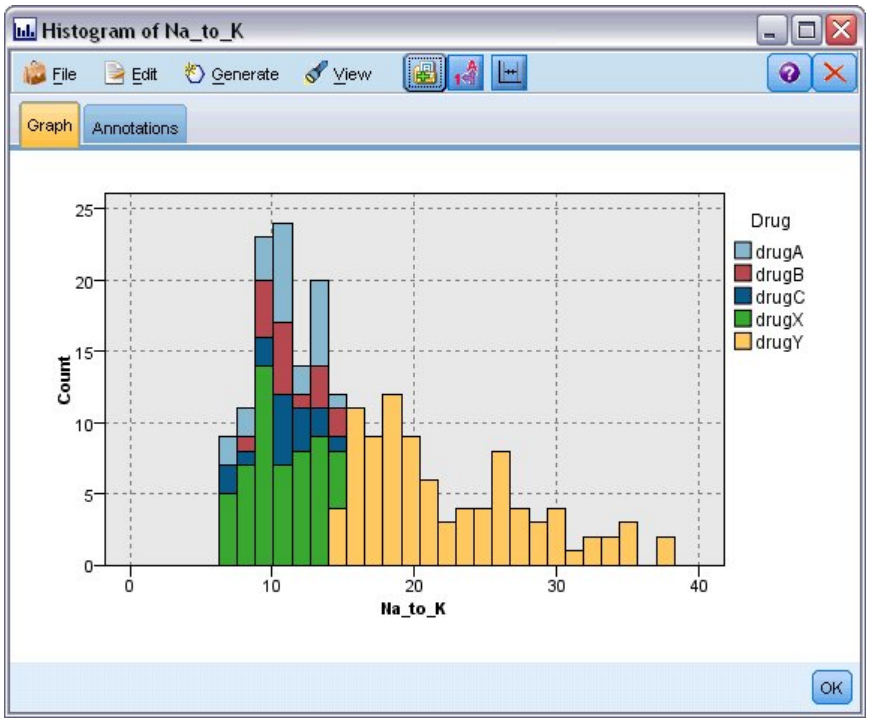

Figure 2. Exemple d'histogramme

Propriétés d'un travail IBM SPSS Modeler - Importation Cognos

Si le flux contient un noeud source IBM Cognos, les détails de la connexion Cognos s'affichent ici.

Nom du noeud. Le nom du noeud source Cognos.

URL de connexion. L'URL du serveur Cognos auquel se connecter.

Nom du package. Le nom du package Cognos depuis lequel les métadonnées sont importées.

Anonyme. Contient **Anonyme** si une connexion anonyme est utilisée pour la connexion au serveur Cognos, ou **Informations d'identification** si un nom d'utilisateur et un mot de passe Cognos spécifiques sont utilisés.

Données d'authentification. Le nom d'utilisateur et le mot de passe (si requis) sur le serveur Cognos.

Remarque : Les données d'identification Cognos doivent être créées dans un domaine, qui représente l'ID d'espace-noms Cognos.

Propriétés d'un travail IBM SPSS Modeler - Exportation Cognos

Si le flux contient un noeud Exportation IBM Cognos, les détails de la connexion Cognos et ODBC s'affichent ici.

Nom du noeud. Le nom du noeud d'exportation Cognos.

URL de connexion. L'URL du serveur Cognos à laquelle la connexion est établie.

Nom du package. Le nom du package Cognos utilisé pour exporter les métadonnées.

Source de données. Le nom de la base de données Cognos utilisée pour exporter les données.

Dossier. Le chemin d'accès et le nom du dossier du serveur Cognos où le package d'exportation est créé.

Anonyme. Contient **Anonyme** si une connexion anonyme est utilisée pour la connexion au serveur Cognos, ou **Données d'identification** si un nom d'utilisateur et un mot de passe Cognos spécifiques sont utilisés.

Données d'identification. Le nom d'utilisateur et le mot de passe (si nécessaire) sur le serveur Cognos.

DSN. Le nom source de base de données (DSN) de la base de données Cognos.

Nom de la table. Le nom de la table de la base de données utilisée pour l'exportation.

Données d'identification. Le nom d'utilisateur et le mot de passe utilisés pour la connexion à la base de données utilisée pour l'exportation.

Remarque : Les données d'identification Cognos doivent être créées dans un domaine qui représente l'ID de l'espace-noms Cognos.

Propriétés d'un travail IBM SPSS Modeler - Importation TM1 existante

Si le flux contient un noeud Importation TM1 existante, les détails de connexion sont affichés ici.

Nom du noeud. Nom du noeud Importation TM1 existante.

Serveur TM1 . Nom du serveur Cognos TM1.

Cube. Cube TM1 depuis lequel les données seront importées.

Vue. Vue à importer du cube TM1.

Données d'identification. Données d'identification du serveur Cognos TM1.

Propriétés d'un travail IBM SPSS Modeler - Exportation TM1 existante

Si le flux contient un noeud Exportation TM1 existante, les détails de connexion sont affichés ici.

Nom du noeud. Nom du noeud Exportation TM1 existante.

Serveur TM1 . Nom du serveur Cognos TM1.

Cube. Cube TM1 vers lequel les données seront exportées.

Données d'identification. Données d'identification du serveur Cognos TM1.

Propriétés d'un travail IBM SPSS Modeler - Importation TM1

Si le flux contient un noeud Importation TM1, les détails de connexion sont affichés ici.

Nom du noeud. Nom du noeud Importation TM1.

Hôte d'administration. Hôte d'administration de Cognos TM1.

Serveur TM1 . Nom du serveur Cognos TM1.

Cube. Cube TM1 depuis lequel les données seront importées.

Vue. Vue à importer du cube TM1. Vous pouvez importer une vue publique ou privée.

Lignes. Lignes à importer du cube TM1.

Colonnes. Colonnes à importer du cube TM1.

Données d'identification. Données d'identification du serveur Cognos TM1.

Propriétés d'un travail IBM SPSS Modeler - Exportation TM1

Si le flux contient un noeud Exportation TM1, les détails de connexion sont affichés ici.

Nom du noeud. Nom du noeud Exportation TM1.

Hôte d'administration. Hôte d'administration de Cognos TM1.

Serveur TM1 . Nom du serveur Cognos TM1.

Cube. Cube TM1 vers lequel les données seront exportées.

Mesure. Mesure qui sera exportée.

Données d'identification. Données d'identification du serveur Cognos TM1.

Propriétés d'un travail IBM SPSS Modeler - Analytic Server Import

Si le flux contient un noeud source Analytic Server, les détails de connexion sont affichés ici.

Use Default Analytic Server. La valeur **True** ou **False** indique si la connexion Analytic Server par défaut définie par un administrateur dans le fichier `options.cfg` est utilisée (**True**) ou si un autre serveur Analytic Server est utilisé, celui spécifié dans le flux/travail (**False**).

URL. URL d'Analytic Server au format `https://nom_hôte:port/racine_contexte`, où `nom_hôte` est l'adresse IP ou le nom d'hôte d'Analytic Server, `port` est son numéro de port et `racine_contexte` est la racine de contexte d'Analytic Server.

Locataire. Nom du locataire auquel appartient SPSS Modeler Server.

Données d'identification. Données d'identification utilisées pour la connexion à Analytic Server.

Nom principal du service. Nom principal du service Kerberos.

Chemin du fichier de configuration. Chemin du fichier de configuration du service Kerberos.

Source de données. Nom de la source de données sur Analytic Server.

Propriétés d'un travail IBM SPSS Modeler - Analytic Server Export

Si le flux contient un noeud Analytic Server Export, les détails de connexion sont affichés ici.

Use Default Analytic Server. La valeur **True** ou **False** indique si la connexion Analytic Server par défaut définie par un administrateur dans le fichier `options.cfg` est utilisée (**True**) ou si un autre serveur Analytic Server est utilisé, celui spécifié dans le flux/travail (**False**).

URL. URL d'Analytic Server au format `https://nom_hôte:port/racine_contexte`, où `nom_hôte` est l'adresse IP ou le nom d'hôte d'Analytic Server, `port` est son numéro de port et `racine_contexte` est la racine de contexte d'Analytic Server.

Locataire. Nom du locataire auquel appartient SPSS Modeler Server.

Données d'identification. Données d'identification utilisées pour la connexion à Analytic Server.

Nom principal du service. Nom principal du service Kerberos.

Chemin du fichier de configuration. Chemin du fichier de configuration du service Kerberos.

Source de données. Nom de la source de données sur Analytic Server.

Propriétés d'un travail IBM SPSS Modeler - Notifications

Vous permet de spécifier des notifications par courrier électronique lorsqu'une étape de travail échoue ou réussit.

Cliquez sur le bouton **Mise à jour** dans chaque cas pour ajouter ou supprimer les destinataires des notifications.

Affichage des flux dans IBM SPSS Modeler

Vous pouvez lancer l'application IBM SPSS Modeler directement depuis IBM SPSS Deployment Manager.

Pour afficher le flux dans l'application IBM SPSS Modeler, cliquez deux fois sur le flux dans l'explorateur de contenu. Le système lance l'application IBM SPSS Modeler et affiche le flux dans IBM SPSS Modeler.

Il est important de noter que si des modifications sont apportées aux fichiers d'un travail, par exemple, un fichier de flux IBM SPSS Modeler (`.str`), chaque travail qui contient le fichier est affecté. Lorsque des changements sont apportés au fichier, une nouvelle version du fichier est enregistrée dans le référentiel. Toutefois, le travail contenant le fichier n'est pas automatiquement mise à jour avec le fichier modifié. Pour incorporer les mises à jour du fichier dans le travail affecté :

1. Rouvrez le travail. Lorsque le travail est rouvert, un astérisque apparaît à côté du nom du travail dans l'espace de travail des travaux, indiquant que le travail contient des changements non enregistrés.
2. Enregistrez à nouveau le travail.

Codes d'achèvement IBM SPSS Modeler

Les codes d'achèvement des travaux IBM SPSS Modeler sont décrits dans le tableau suivant. Utilisez ces codes d'achèvement pour toute relation conditionnelle qui implique des flux IBM SPSS Modeler.

Tableau 1. Codes d'achèvement des travaux IBM SPSS Modeler

CODE	Description
0	succès
1	erreur d'exécution de flux
2	erreur de publication
8	erreur inconnue

Limites des flux IBM SPSS Modeler

Lorsque vous utilisez des flux dans IBM SPSS Deployment Manager, les contraintes suivantes s'appliquent au système :

- **Conventions de nom.** Si le nom de noeud, le libellé et le type sont identiques, vous ne pouvez pas planifier l'exécution d'un travail en raison de la présence d'un conflit.
- **Scripts.** Si un script de flux contient des valeurs de remplacement pour un noeud spécifique, par exemple, le script définit l'emplacement de sortie d'un graphique, le script remplace les valeurs en conflit définies par l'utilisateur dans Deployment Manager.
- **Supernoeuds.** L'exécution d'une étape de travail IBM SPSS Modeler correspond au traitement et à l'exécution de tous les noeuds terminaux de niveau supérieur dans le flux. Si un des noeuds terminaux est un super noeud terminal, les noeuds terminaux de ce super noeud s'exécutent de manière récursive. En revanche, si des super noeuds source ou de processus, par définition non terminaux, contiennent des noeuds terminaux, ces noeuds ne s'exécutent pas. Les noeuds terminaux situés dans les super noeuds non terminaux apparaissent dans Deployment Manager mais ne sont pas traités lors de l'exécution de l'étape.
- **Paramètres.** Les paramètres définis pour les super noeuds ne peuvent être nommés de la même manière que les paramètres définis pour un flux spécifique. Les noms doivent être uniques.
- **Restrictions du noeud Text Analytics.** Vous ne pouvez pas utiliser les noeuds Flux Web et Liste de fichiers SPSS Modeler Text Analytics pour l'évaluation dans une configuration IBM SPSS Collaboration and Deployment Services - Scoring.

Outre les éléments répertoriés ci-dessus, il existe des contraintes spécifiques.

- Aucune case à cocher de déverrouillage n'est disponible lorsque vous stockez un flux dans le référentiel IBM SPSS Collaboration and Deployment Services. Déverrouiller est la valeur par défaut lors du stockage des flux. Pour verrouiller ou déverrouiller un objet, sélectionnez **Outils** > **Référentiel** > **Explorer**, puis accédez à l'objet et cliquez avec le bouton droit de la souris sur son nom pour afficher le menu contextuel.

- Dans Deployment Manager, lorsque vous exécutez un travail contenant un flux avec un noeud Evaluation configuré pour générer un graphique Gains, la sortie de ce graphique risque d'être incomplète si le système fonctionne sous Oracle Weblogic 11g en utilisant l'environnement d'exécution Java d'Oracle JRockit. Pour éviter ce problème, utilisez l'environnement d'exécution Java IBM.

Types de noeud

Lorsque vous ouvrez un flux IBM SPSS Modeler à partir de IBM SPSS Deployment Manager, vous pouvez voir que les noeuds de flux sont représentés par différentes formes d'icônes. Les icônes en forme de cercle représentent des noeuds de sources de données, alors que les noeuds de forme hexagonale représentent des opérations de traitement d'enregistrements de données et de champs. Les triangles indiquent un champs de sortie graphique, tandis que les pentagones représentent des noeuds modélisation. Les champs de sortie (autres que graphiques) et les exportations sont indiqués par des noeuds rectangulaires.

Les différents noeuds sont décrits de manière complète dans les guides *Noeuds source, noeuds d'exécution et noeuds de sortie d'IBM SPSS Modeler* et *Noeuds de modélisation d'IBM SPSS Modeler*.

Exemple de script : Comparaison de modèles champion et challenger

L'extrait de script IBM SPSS Modeler suivant est un exemple de comparaison champion et challenger. Le modèle champion est comparé au modèle challenger. Si le modèle challenger est meilleur que le modèle champion, le modèle champion est remplacé. Ainsi le modèle challenger devient le nouveau modèle champion.

```
# -----------------------------------
# Before rebuilding model, examine performance of existing model
# -----------------------------------
#--- LOAD Model
set working_file = ^SPSS_SMS_INT_PERSIST_PATH >< ^slash >< "ResolutionStatus.gm"
load model ^working_file

#--- First do the evaluation of the current champion model
delete 'ResolutionStatusModel'
insert model 'ResolutionStatus' at 558 652
rename 'ResolutionStatus':applyc50node as 'ResolutionStatusModel'
connect 'ValidationSet' to 'ResolutionStatusModel'
connect 'ResolutionStatusModel' to 'PredictQuality'

set 'ModelEval'.full_filename = ^SPSS_SMS_INT_TEMP_PATH >< ^slash >< "ModelEvalChamp.txt"
execute 'ModelEval':tablenode
var missRateChamp
set missRateChamp = value 'ModelEval':tablenode.output at 1 6
delete 'ResolutionStatusModel'
clear generated palette

#--- Execute c50node 'ResStatusModelBuild'
if ^useBoost = 1
    set 'ResStatusModelBuild'.use_boost = T
    set 'ResStatusModelBuild'.boost_num_trials = ^numBoostTrials
else
    set 'ResStatusModelBuild'.use_boost = F
endif
if ^useMisCosts = 1
    set 'ResStatusModelBuild'.use_costs = T
    set 'ResStatusModelBuild'.costs = [{'Not fraud' 'Not fraud' 0}{'Not fraud' 'Fraud'
^falsePosCost}
    {'Fraud' 'Not fraud' ^falseNegCost}{'Fraud' 'Fraud' 0}]
endif
execute 'ResStatusModelBuild'

#--- Now evaluate the new model
insert model 'ResolutionStatus' at 558 652
rename 'ResolutionStatus':applyc50node as 'ResolutionStatusModel'
connect 'ValidationSet' to 'ResolutionStatusModel'
connect 'ResolutionStatusModel' to 'PredictQuality'

set 'ModelEval'.full_filename = ^SPSS_SMS_INT_TEMP_PATH >< ^slash >< "ModelEvalChallenge.txt"
execute 'ModelEval':tablenode
```

```
var missRateChallenge
set missRateChallenge = value 'ModelEval':tablenode.output at 1 6

#--- Determine which one should be the champion
#--- Overwrite existing champion if challenger is better

if ^missRateChamp > ^missRateChallenge
    #--- SAVE Model
    set working_file = ^SPSS_SMS_INT_PERSIST_PATH >< ^slash >< "ResolutionStatus.gm"
    save model 'ResolutionStatus':applyc50node as ^working_file

    #--- EXPORT Model
    set working_file = ^SPSS_SMS_INT_PERSIST_PATH >< ^slash >< "ResolutionStatus.xml"
    export model 'ResolutionStatus' as ^working_file format pmml

    #--- EXPORT Model
    set working_file = ^SPSS_SMS_INT_PERSIST_PATH >< ^slash >< "ResolutionStatus_summary.html"
    export model 'ResolutionStatus' summary  as ^working_file format html
endif
```

Chapitre 4. Service d'évaluation

Service d'évaluation permet aux applications clientes d'utiliser des scores en temps réel dérivés des modèles prédictifs développés dans IBM SPSS Modeler. Le service extrait le modèle indiqué, le charge, appelle l'implémentation d'évaluation correcte et renvoie le résultat au client.

L'évaluation est le processus consistant à générer des valeurs en temps réel en fournissant des données d'entrée aux modèles prédictifs. Un modèle d'évaluation est un artefact permettant de générer des valeurs de sortie si des données d'entrée lui sont fournies. En général, pour utiliser un modèle pour générer des scores, procédez comme suit :

1. Sélectionnez un modèle à utiliser pour l'évaluation dans IBM SPSS Collaboration and Deployment Services Repository.
2. Définissez la configuration d'évaluation du modèle.
3. Fournissez des données au modèle configuré et générez des scores.

Pour plus d'informations sur l'évaluation et Service d'évaluation, consultez la documentation de IBM SPSS Collaboration and Deployment Services.

Limites des flux IBM SPSS Modeler

Lorsque vous utilisez des flux dans IBM SPSS Deployment Manager, les contraintes suivantes s'appliquent au Service d'évaluation :

- **Supernoeuds.** Les noeuds source dans les super noeuds ne sont pas pris en charge, et les noeuds terminaux dans les super noeuds ne sont pas pris en charge.
- **Noeuds géospatiaux.** Les noeuds géospatiaux ne sont pas pris en charge.
- **Noeuds de créateur de modèle.** Le noeud de créateur de modèle en tant que noeud terminal n'est pas pris en charge.
- **Exploration des bases de données.** Les flux contenant des nuggets d'exploration des bases de données ne sont pas pris en charge.
- **UDF, UDA et WUDA.** Les flux utilisant les fonctions de base de données (UDF), les agrégations de base de données (UDA) ou les agrégats de fenêtre de base de données (WUDA) ne sont pas pris en charge.
- **Noeuds source.** Les noeuds source avec un modèle de données de sortie contenant un type de liste ne sont pas pris en charge.
- **Noeuds terminaux.** Les noeuds terminaux avec un modèle de données d'entrée contenant un type de liste ne sont pas pris en charge.
- **Restrictions du noeud Text Analytics.** Vous ne pouvez pas utiliser les noeuds Flux Web et Liste de fichiers SPSS Modeler Text Analytics pour l'évaluation dans une configuration IBM SPSS Collaboration and Deployment Services - Scoring.
- **Noeuds source Analytic Server.** Si un travail contenant un flux SPSS Modeler utilise un noeud source Analytic Server, vous devez autoriser la connexion directe entre le serveur IBM SPSS Collaboration and Deployment Services et le serveur Analytic Server. Sinon, le travail échouera si le pare-feu bloque la connexion entre les deux serveurs.

Chapitre 5. Etapes de travail champion/challenger

Présentation de champion/challenger

Avec IBM SPSS Deployment Manager, il est possible de comparer des fichiers de modèles générés par IBM SPSS Modeler afin de déterminer quel fichier contient le modèle prédictif le plus efficace. L'étape de travail champion/challenger évalue un modèle et le compare à un ou plusieurs challengers.

Après que le système ait comparé les résultats, le meilleur modèle devient le nouveau champion.

Champion. Le champion correspond au modèle le plus efficace. Pour la première exécution de l'étape de travail challenger/champion, il n'y a pas de champion, mais uniquement le premier challenger et la liste des challengers correspondante. Pour les autres exécutions de l'étape de travail, le système détermine le champion.

Challenger. Les challengers sont comparés les uns avec les autres. Le challenger générant les meilleurs résultats devient alors le nouveau champion.

Processus de sélection du champion

Le processus de comparaison des challenger/champion est composé des tâches suivantes :

1. Scoring de chacun des modèles concurrents.
2. Evaluation des scores.
3. Comparaison des résultats de scoring et sélection du challenger champion.
4. Enregistrement du nouveau champion dans le référentiel (facultatif).

Ajout d'un travail Challenger/Champion à un autre travail

Pour ajouter une étape de travail challenger/champion à un autre travail, sélectionnez l'outil Challenger/Champion dans la palette des travaux et cliquez sur l'espace de travail des travaux.

Mesures de scoring des modèles

L'évaluation et la comparaison des modèles peut porter sur la précision, les gains ou l'accréditation.

- **Précision**. La précision d'un modèle reflète le pourcentage de réponses cible qui sont correctement prédites. Les modèles ayant un pourcentage élevé de prédictions correctes sont favorisés par rapport à ceux ayant un faible pourcentage.
- **Gains.** La statistique sur les gains est un indicateur des performances d'un modèle. Cette mesure compare les résultats d'un modèle aux résultats obtenus sans utiliser de modèle. L'amélioration des résultats lorsque le modèle est utilisé est appelée « gains ». Lors de la comparaison de deux modèles, le modèle ayant la valeur de gains la plus élevée à un centile donné est favorisé.
- **Accréditation.** L'accréditation d'un modèle reflète la crédibilité d'un modèle. Cette approche examine la similarité entre de nouvelles données et les données de formation sur lesquelles un modèle est basé. Les valeurs d'accréditation varient de 0 à 1, les valeurs élevées indiquant une similarité plus importante entre les prédicteurs dans les deux jeux de données. Lors de la comparaison de deux modèles, le modèle ayant la valeur d'accréditation la plus élevée est basé sur les données de formation les plus similaires aux nouvelles données, le rendant ainsi plus crédible et lui donnant l'avantage.

Dépendance à l'ordre

Contrairement aux autres types d'étapes de travaux, les onglets dans l'étape de travail challenger/champion sont sensibles à l'ordre.

Par exemple, un challenger doit être sélectionné dans l'onglet Challengers avant que des informations puissent être modifiées dans l'onglet Champion. De plus, les informations qui apparaissent dans certains onglets dépendent des challengers sélectionnés dans le tableau des challengers.

Le processus d'exécution de la comparaison des challenger/champion est composé des étapes suivantes :

1. Informations générales sur le travail.
2. Identification des challengers.
3. Spécification d'informations sur le champion.
4. Affichage des informations relatives aux paramètres.
5. Spécification des notifications.

Informations générales

L'onglet Général contient des informations qui s'appliquent à l'étape de travail challenger/champion globale.

Le serveur IBM SPSS Modeler et ses informations de connexion sont nécessaires à l'exécution d'une étape de travail challenger/champion. Le serveur du référentiel de contenu et ses informations de connexion sont nécessaires à l'exécution du travail et à l'enregistrement des nouveaux champions dans IBM SPSS Collaboration and Deployment Services Repository. (Le serveur du référentiel de contenu et ses informations de connexion sont nécessaires même si vous ignorez les résultats de l'analyse.) Les informations de connexion sont basées sur l'utilisateur actuellement connecté au système.

Nom de l'étape du travail. Le nom de l'étape de travail. Par défaut, le nom de la première étape de travail est *Événement 1*. Les étapes de travail suivantes sont nommées *Evénement 2*, *Evénement 3*, etc. Le nom spécifié ici apparaît dans le tableau d'historique des travaux après l'exécution de l'étape de travail.

Serveur IBM SPSS Modeler. Le serveur ou grappe de serveurs IBM SPSS Modeler sur lequel le flux sera exécuté. La liste contient tous les serveurs et les clusters de serveurs actuellement configurés pour exécuter les étapes IBM SPSS Modeler. Pour modifier le serveur, procédez à une sélection dans la liste déroulante **IBM SPSS Modeler Server**. Pour créer une nouvelle définition de serveur, cliquez sur **Nouveau** pour lancer l'assistant de définition de serveur.

Connexion à IBM SPSS Modeler. Les informations d'identification utilisées pour accéder au serveur ou grappe de serveurs IBM SPSS Modeler. Pour modifier les données d'identification, procédez à une sélection dans la liste déroulante **Nom de connexion d'IBM SPSS Modeler**. Pour définir de nouvelles données d'identification, cliquez sur **Nouveau** pour lancer l'assistant de définition de données d'identification.

Serveur du référentiel de contenu. Le serveur du référentiel de contenu permet aux travaux d'enregistrer des fichiers dans un IBM SPSS Collaboration and Deployment Services Repository. Généralement, le serveur de référentiel de contenu est spécifié lors de l'actualisation de modèles à l'aide d'IBM SPSS Modeler. Pour spécifier un serveur de référentiel de contenu, sélectionnez un serveur dans la liste déroulante **Serveur de référentiel de contenu**. Pour créer une nouvelle définition de serveur, cliquez sur **Nouveau** pour lancer l'assistant de définition de serveur. Pour générer une définition de serveur de référentiel de contenu basée sur des informations de serveur actuelles, cliquez sur **Générer**. Une définition de serveur est créée et automatiquement renseignée dans le champ *Serveur de référentiel de contenu*.

Connexion au référentiel de contenu. Les informations de connexion du serveur de référentiel de contenu. Pour spécifier un nom de connexion au référentiel de contenu, sélectionnez des données d'identification dans la liste déroulante **Connexion au référentiel de contenu**. Pour créer une nouvelle connexion, cliquez sur **Nouveau** pour lancer l'assistant de connexion au référentiel de contenu. Si l'authentification unique n'est pas utilisée pour vous connecter à IBM SPSS Collaboration and Deployment Services Repository, cliquez sur **Générer** pour générer une connexion au serveur de référentiel de contenu basée sur des paramètres de sécurité existants. Une connexion au référentiel de contenu est créée et automatiquement renseignée dans le champ *Connexion de référentiel de contenu*. La génération de connexion n'est pas disponible lors de l'utilisation d'une authentification unique.

Challengers

Au moins un premier challenger doit être sélectionné pour exécuter une étape de travail challenger/champion. Il est important de noter que le premier challenger sélectionné n'implique aucun ordre de comparaison ni aucune priorité dans le processus de scoring. Le premier challenger est simplement la base de référence.

La source de données et les libellés utilisées pour déterminer les autres challengers sont établies par le premier challenger. Après la sélection ou mise à jour du premier challenger, les champs restants dans cet onglet sont mis à jour avec les informations qui correspondent au premier challenger.

Premier challenger. Nom du premier challenger. Pour parcourir le référentiel, cliquez sur **Parcourir**.

Libellé du premier challenger. Le libellé associé au fichier de modèle contenant le premier challenger. Spécifiez cette valeur lors de la sélection du premier challenger.

Challenger de source de données. Le challenger fournissant le nœud source de données utilisé pour l'étape de travail. Cliquez sur **Parcourir** pour choisir ce challenger dans la liste des entrées sélectionnées dans le tableau Challengers.

Mesure. Critères de mesure selon lesquels les challengers sont comparés. Les valeurs valides comprennent *accréditation*, *précision* et *gains*. Si *gains* est sélectionné, un centile doit aussi être spécifié. Pour plus d'informations, voir la rubrique «Mesures de scoring des modèles», à la page 23.

Tableau Challengers

Le tableau Challengers répertorie les branches de scoring par défaut pour les challengers qui correspondent à la source de données et au libellé associés au premier challenger. Seuls les challengers sélectionnés dans ce tableau seront comparés au premier challenger lorsque l'étape de travail est exécutée. En cas de sélection (ou de désélection) d'un challenger de la liste, le système met à jour les informations correspondantes dans les autres onglets des étapes de travaux.

Chaque fois qu'un travail qui contient une étape challenger/champion enregistrée précédemment est ouvert, la liste des challengers est automatiquement mise à jour. De nouveaux challengers peuvent être ajoutés à la liste s'ils correspondent aux critères de source de données et de libellés du premier challenger. Inversement, les challengers qui ne correspondent plus à ces critères peuvent être supprimés de la liste des challengers. Si un challenger sélectionné a été supprimé du référentiel, le système génère un message indiquant que ce challenger n'est plus disponible.

Bien que les challengers puissent être sélectionnés et désélectionnés pour être comparés, les informations du tableau des challengers ne sont pas modifiables. Le tableau Challengers contient les informations suivantes spécifiques.

Nom. Nom du challenger.

Libellé. Libellé associé au challenger.

Description. Description du challenger.

Modifications du premier challenger

Les changements effectués sur le premier challenger après l'enregistrement du travail peuvent avoir une influence sur l'analyse du challenger/champion. Par exemple, imaginons que le premier challenger est supprimé du référentiel ou que le libellé associé au premier challenger est supprimé. Lorsque vous accédez à l'onglet Challengers, le système génère un message indiquant que le premier challenger n'est plus utilisable. Dans ce cas, un nouveau premier challenger doit être spécifié.

Sélection des challengers

Pour sélectionner les modèles challenger et les inclure dans les analyses de challenger/champion, exécutez les étapes suivantes :

1. Dans le premier onglet d'une étape challenger/champion, cliquez sur **Parcourir** pour le premier challenger. Si vous ajoutez des challengers manuellement, cliquez sur **Ajouter** pour le tableau Challengers.
2. Sélectionnez le fichier de modèle en cliquant sur **Parcourir**. Le fichier de modèle est un flux IBM SPSS Modeler contenant une branche de scoring par défaut avec un nugget de modèle valide.
3. Sélectionnez le libellé désignant la version du fichier de modèle sélectionné à utiliser.
4. Dans le tableau Challengers, sélectionnez la branche de scoring à utiliser.
5. Cliquez sur **OK**.

Challengers non valides

Pour pouvoir être comparés, les fichiers de modèle doivent contenir des branches de scoring qui utilisent une structure de données commune.

Les caractéristiques des données qui doivent correspondre dans les challengers incluent les éléments suivants :

- les sources de données doivent avoir le même nombre de champs.
- Les noms des champs doivent être les mêmes dans toutes les sources de données.
- Les niveaux de mesure des champs doivent être les mêmes dans toutes les sources de données.

Si le système ne trouve pas de challengers pouvant être comparés au premier challenger sélectionné, la boîte de dialogue Challenger non valide apparaît. Pour sélectionner un nouveau challenger :

1. Cliquez sur **OK** pour retourner à l'onglet Challengers.
2. Sélectionnez un nouveau challenger.

Sélection des sources de données des challengers

Pour sélectionner la source de données utilisée dans les analyses champion/challenger, exécutez les étapes suivantes :

1. Dans le premier onglet d'une étape challenger/champion, cliquez sur **Parcourir** pour le challenger de la source de données.
2. Dans la liste des challengers inclus dans l'analyse, sélectionnez la branche de scoring qui contient la source de données à utiliser.
3. Cliquez sur **OK**.

Champion

Avant de pouvoir spécifier les informations pour un champion, au moins un challenger doit être sélectionné. Si vous accédez à l'onglet Champion avant d'avoir sélectionné un challenger, la boîte de dialogue Premier challenger non sélectionné apparaît pour indiquer qu'un challenger doit être sélectionné.

Ne pas créer une nouvelle version du champion. Sélectionnez cette option pour empêcher la création d'une nouvelle version du champion. Dans ce cas, la version libellée sélection du champion est modifiée. Désélectionnez cette option pour créer une nouvelle version du champion au lieu de modifier la version libellée.

Nom du fichier. Le nom à utiliser pour la copie du challenger identifié comme champion.

Emplacement. L'emplacement dans lequel est stockée la copie du fichier du champion.

Privilèges. Les permissions associées à la copie du champion.

Métadonnées. Les propriétés associées à la copie du champion. La spécification de métadonnées pour une sortie de champion se fait de la même façon que la spécification de métadonnées pour d'autres sorties de travail.

Utilisation du champion dans d'autres travaux

Après l'exécution du travail du challenger/champion, le champion sélectionné peut être utilisé dans d'autres travaux. Pour inclure le champion dans un autre travail, les informations suivantes sont nécessaires :

- Nom du champion.
- Emplacement du champion.

Lorsque le champion est utilisé dans un autre travail, le libellé *LATEST* est appliqué. Ce libellé ne peut pas être modifié.

Test du champion

Par défaut, le système crée une nouvelle copie du champion à chaque fois que l'étape de travail challenger/champion est exécutée, il stocke cette copie dans l'emplacement de sortie spécifié et écrit les résultats dans le fichier journal de l'historique des travaux.

Cependant, il peut exister des instances dans lesquelles il est préférable de ne pas enregistrer de copie du champion. Par exemple, imaginons que vous souhaitiez seulement tester l'étape de travail challenger/champion.

Pour désactiver la création d'une copie, sélectionnez la case à cocher **Ne pas créer de nouvelle version du champion**. Si cette case à cocher est sélectionnée, les autres options de l'onglet sont désactivées. Le système utilise les mêmes informations que celles du champion actuel.

Le système exécute l'étape de travail challenger/champion et détermine un nouveau champion. Cependant, aucune nouvelle version du champion ne sera créée ou enregistrée dans le référentiel. Les résultats seront simplement écrits dans le fichier journal de l'historique des travaux et indiqueront quel challenger aurait été choisi comme champion.

Par exemple, imaginons qu'un travail champion/challenger est exécuté et que le créateur du travail ne doive pas créer de nouvelle version du champion. Le fichier journal de l'historique des travaux pourrait ressembler à cela :

```
Stream execution started
500 500
1000 1000
1500 1500
2000 2000
2500 2500
Stream execution complete, Elapsed=26.22 sec, CPU=18.97 sec
Stream execution started
1000 0
2000 0
Field 'Correct_Sum' has only one value
Field 'Count' has only one value
Field 'Traffic Light Result' has only one value
2855 145
2855 1145
2855 2145
Field 'campaign' has only one value
Field 'gold_card' has only one value
Field 'response' has only one value
Stream execution complete, Elapsed=0.39 sec, CPU=0.2 sec
Stream execution started
500 500
1000 1000
1500 1500
2000 2000
2500 2500
Stream execution complete, Elapsed=26.06 sec, CPU=17.75 sec
Stream execution started
1000 0
```

```
2000 0
Field 'Correct_Sum' has only one value
Field 'Count' has only one value
Field 'Traffic Light Result' has only one value
2855 145
2855 1145
2855 2145
Field 'campaign' has only one value
Field 'gold_card' has only one value
Field 'response' has only one value
Stream execution complete, Elapsed=0.48 sec, CPU=0.19 sec
Stream execution started
500 500
1000 1000
1500 1500
2000 2000
2500 2500
Stream execution complete, Elapsed=21.48 sec, CPU=17.34 sec
Stream execution started
1000 0
2000 0
Field 'Correct_Sum' has only one value
Field 'Count' has only one value
Field 'Traffic Light Result' has only one value
2855 145
2855 1145
2855 2145
Field 'campaign' has only one value
Field 'gold_card' has only one value
Field 'response' has only one value
Stream execution complete, Elapsed=0.39 sec, CPU=0.17 sec
The result for challenger cc_cartresponse.str is 98.809.
The result for challenger cc_neuralnetresponse.str is 98.844.
The result for challenger cc_c51response.str is 98.809.
The declared Champion is cc_neuralnetresponse.str.
```

Veuillez remarquer la dernière ligne du fichier journal :

```
The declared Champion is cc_neuralnetresponse.str.
```

Cette ligne indique que le flux *cc_neuralnetresponse* aurait été le champion. Toutefois, aucune copie de ce flux n'a été enregistrée dans le référentiel car le système n'a pas créé de nouvelle version du flux. Si le flux avait été enregistré dans le référentiel, le fichier journal aurait contenu une ligne supplémentaire pour l'indiquer. Par exemple :

```
Adding artifact spsstr:/PMDemo/ModelManagement/cc_neuralnetresponse.str.
```

Fichiers de données

Des informations sur les fichiers de données apparaissent pour les challengers sélectionnés dans la liste de challengers de l'onglet Challengers.

Tout changement effectué dans cet onglet s'applique uniquement à l'étape de travail de challenger/champion. Les modifications apportées aux informations du fichier de données ne sont pas répercutées dans le challenger enregistré dans IBM SPSS Collaboration and Deployment Services Repository. Le tableau Fichiers de données contient les informations suivantes.

Nom du noeud. Le nom du nœud d'entrée qui contient les données utilisées par le flux. Le nom du noeud n'est pas modifiable.

Type du noeud. Le type de nœud tel qu'il est défini dans le flux. Le type du noeud n'est pas modifiable.

Nom du fichier. Le nom du fichier de données d'entrée. Pour modifier ce nom, cliquez sur la cellule Nom du fichier et modifiez le nom.

Format. Le format du fichier de sortie, par exemple, un fichier csv. Pour modifier le type de format du fichier, cliquez sur la cellule Format. Une flèche de liste déroulante apparaît. Sélectionnez le type de format.

Emplacement. L'emplacement des fichiers de données d'entrée. Pour modifier cet emplacement, cliquez sur la colonne puis cliquez sur le bouton des points de suspension (...) qui apparaît. La boîte de dialogue Emplacement de fichier d'entrée apparaît. Modifiez l'emplacement.

Vue de données

Les informations de la vue de données analytiques apparaissent pour les challengers sélectionnés dans la liste des challengers de l'onglet Challengers.

Tout changement effectué dans cet onglet s'applique uniquement à l'étape de travail de challenger/champion. Les modifications apportées aux informations de la vue de données ne sont pas appliquées au challenger enregistré dans IBM SPSS Collaboration and Deployment Services Repository. La table de la vue de données contient les informations suivantes.

Nom du noeud. Nom du noeud de vue de données qui contient les données utilisées par le flux. Le nom du noeud n'est pas modifiable.

Vue de données analytiques. Vue de données analytiques référencée par le noeud de vue de données.

Libellé. Libellé identifiant la version de la vue de données analytiques utilisée.

Nom du tableau. Table contenant les champs de données d'entrée.

Plan d'accès aux données. Plan fournissant les enregistrements de données pour les champ de données d'entrée. Pour modifier le plan d'accès aux données utilisé pour un noeud, sélectionnez la cellule contenant ce plan et cliquez sur le bouton points de suspension (...) résultant.

sources de données ODBC

Des informations sur les sources de données ODBC apparaissent pour les challengers sélectionnés dans la liste des challengers de l'onglet Challengers.

Tout changement effectué dans cet onglet s'applique uniquement à l'étape de travail de challenger/champion. Les modifications des informations des sources de données ODBC ne sont pas renvoyées au challenger enregistré dans IBM SPSS Collaboration and Deployment Services Repository. Le tableau Sources de données ODBC contient les informations suivantes.

Nom du noeud. Le nom du nœud d'entrée qui contient les données utilisées par le flux. Ce nom est précédé des noms des super noeuds contenant le noeud, séparés par des barres obliques. Par exemple, si le noeud *MonNoeud* se trouve dans un super noeud intitulé *Supernoeud1*, le nom apparaît sous la forme */Supernoeud1/MonNoeud*.

Type du noeud. Le type de nœud tel qu'il est défini dans le flux.

Sources de données ODBC. Nom de la source de données ODBC (DSN). Pour passer à une autre source de données ODBC, cliquez sur la cellule contenant le nom de la source de données actuelle, puis cliquez sur le bouton « ... » qui s'affiche. Cette action entraîne l'apparition d'une boîte de dialogue sur laquelle vous pouvez choisir une DSN ou en créer une nouvelle. Notez que les paramètres de l'étape du travail ont toujours priorité sur le flux. Par conséquent, si vous modifiez le flux pour utiliser une source de données différente, vous devez également modifier le travail pour utiliser la même source, sinon le travail risque de ne pas s'exécuter.

Données d'authentification. Pour modifier le nom d'utilisateur et le mot de passe de la base de données lors de la modification de la source de données ODBC, cliquez sur la cellule contenant les données d'identification en cours, puis cliquez sur le bouton « ... » qui s'affiche. Cette action entraîne l'apparition d'une boîte de dialogue dans laquelle vous pouvez choisir une définition de données d'identification existante ou en créer une nouvelle.

Table de base de données. Table de base de données qui correspond au noeud.

Les noeuds situés dans des super noeuds verrouillés ne sont pas accessibles. Ils ne peuvent être ni affichés ni modifiés.

Importation de Cognos

Si les fichiers de modèles contiennent des noeuds source IBM Cognos BI, les détails de la connexion Cognos sont affichés ici.

Nom du noeud. Le nom du noeud source Cognos.

URL de connexion. L'URL du serveur Cognos auquel se connecter.

Nom du package. Le nom du package Cognos depuis lequel les métadonnées sont importées.

Anonyme. Contient **Anonyme** si une connexion anonyme est utilisée pour la connexion au serveur Cognos, ou **Informations d'identification** si un nom d'utilisateur et un mot de passe Cognos spécifiques sont utilisés.

Données d'authentification. Le nom d'utilisateur et le mot de passe (si requis) sur le serveur Cognos.

Remarque : Les données d'identification Cognos doivent être créées dans un domaine, qui représente l'ID d'espace-noms Cognos.

Remarques

Le présent document peut contenir des informations ou des références concernant certains produits, logiciels ou services IBM non annoncés dans ce pays. Cette documentation peut être proposée 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, logiciel ou service IBM n'implique pas que seul ce produit, logiciel ou service IBM puisse être utilisé. Tout autre élément fonctionnellement équivalent peut être utilisé, s'il n'enfreint aucun droit d'IBM. Il est de la responsabilité de l'utilisateur d'évaluer et de vérifier lui-même les installations et applications réalisées avec des produits, logiciels ou services non expressément référencés par IBM.

IBM peut détenir des brevets ou des demandes de brevet couvrant les produits mentionnés dans le présent document. La remise de ce document ne vous donne aucun droit de licence sur ces brevets ou demandes de brevet. Si vous désirez recevoir des informations concernant l'acquisition de licences, veuillez en faire la demande par écrit à l'adresse suivante :

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
USA

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". IBM DECLINE TOUTE RESPONSABILITE, EXPLICITE OU IMPLICITE, RELATIVE AUX INFORMATIONS QUI Y SONT CONTENUES, Y COMPRIS EN CE QUI CONCERNE LES GARANTIES DE VALEUR MARCHANDE OU D'ADAPTATION A VOS BESOINS. Certaines juridictions n'autorisent pas l'exclusion des garanties implicites, auquel cas l'exclusion ci-dessus ne vous sera pas applicable.

Le présent document peut contenir des inexactitudes ou des coquilles. 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
USA

Ces informations peuvent être soumises à des conditions particulières, prévoyant notamment le paiement d'une redevance.

Le logiciel 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 Contrat sur les produits et services IBM, aux Conditions Internationales d'Utilisation de Logiciels IBM ou de tout autre accord 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.

Le présent document peut contenir des exemples de données et de rapports utilisés couramment dans l'environnement professionnel. Ces exemples mentionnent des noms fictifs de personnes, de sociétés, de marques ou de produits à des fins illustratives ou explicatives uniquement. Toute ressemblance avec des noms de personnes ou de sociétés réelles serait purement fortuite.

Marques commerciales

IBM, le logo IBM et ibm.com sont des marques d'International Business Machines Corporation enregistré dans de nombreuses juridictions dans le monde entier. 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 l'adresse www.ibm.com/legal/copytrade.shtml.

Adobe, le logo Adobe, PostScript et le logo PostScript sont des marques d'Adobe Systems Incorporated aux Etats-Unis et/ou dans certains 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 d'Intel Corporation ou de ses filiales aux Etats-Unis et dans certains autres pays.

Linux est une marque de Linus Torvalds aux Etats-Unis et/ou dans certains autres pays.

Microsoft, Windows, Windows NT et le logo Windows sont des marques de Microsoft Corporation aux Etats-Unis et/ou dans certains autres pays.

UNIX est une marque enregistrée de The Open Group aux Etats-Unis et/ou dans certains 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 relatives 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 publications ou autres informations, données, logiciels ou droits de propriété intellectuelle contenus dans ces publications.

IBM se réserve le droit de retirer les autorisations accordées ici si, à sa discrétion, l'utilisation des publications 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

C

champion/challenger 23–30
clusters de serveurs 2, 6
Cognos
  champion/challenger 30
collaboration 1
comparaison
  modèles 23
  Modèles 19
comparaison champion et challenger 19

D

définition de serveur
  IBM SPSS Modeler 7
définition de serveurs
  ajout 7
  modification 8
définitions de ressources
  définition de serveur 7
  définition de serveurs 7
déploiement 2
dossier
  définitions de ressources 7

E

étapes
  IBM SPSS Modeler 9, 12, 14, 15, 17, 18, 21

F

fichiers
  IBM SPSS Modeler 9, 12, 14, 15, 17, 18
Fichiers
  IBM SPSS Modeler 21
fichiers de données
  champion/challenger 28
flux
  IBM SPSS Modeler 9, 12, 14, 15, 17, 18

I

IBM SPSS Collaboration and Deployment Services Deployment Manager 2, 4
IBM SPSS Collaboration and Deployment Services Deployment Portal 2, 5
IBM SPSS Collaboration and Deployment Services Repository 2, 3
IBM SPSS Modeler
  définition de serveur 7
  flux 9, 12, 14, 15, 17, 18
  serveur d'exécution 7
IBM SPSS Modeler *(suite)*
  serveurs d'exécution 2, 6
  streams 21

M

modèles
  champion/challenger 23
  comparaison 23
  IBM SPSS Modeler 23
  scoring 23
  surveillance 23
Modèles
  comparaison 19
  comparaison champion et challenger 19
  évaluation 19
  IBM SPSS Modeler 19
  surveillance 19
modèles challenger
  sources de données 26
modèles champion 19, 23, 26, 27

N

nom de l'étape de travail
  champion/challenger 24

P

paramètres 18
préférences de l'utilisateur 5

S

scriptage
  champion/challenger 23
  comparaison champion et challenger 19
  IBM SPSS Modeler 19, 23
serveurs d'exécution
  IBM SPSS Modeler 2, 6
service d'évaluation 21
sources de données ODBC
  champion/challenger 29
streams
  IBM SPSS Modeler 21
super noeuds 18

T

travaux
  IBM SPSS Modeler 9, 12, 14, 15, 17, 18, 21

V

vue de données analytiques
champion/challenger 29