



PlaceIQ Usage Manual:

Place Intelligence Metrics (PIM)

Q3 2021 | Version 1.0

Table of Contents

PlaceIQ Overview	3
Summary	3
Value Proposition & Differentiators	3
Movement Data Sourcing	5
Basemap Creation	6
Data Processing	6
Product Overview: Place Intelligence Metrics	8
Key Facts	8
Data Structure	8
Schema	8
Usage and Best Practices	10
Metrics Descriptions	11
Data Delivery: Additional Details	13
Data Structure	13
Metadata	13
Appendix	14
Closed Hours	14

PlaceIQ Overview

Summary

PlaceIQ is a leading data and technology provider that powers critical business and marketing decisions with location data, analytics and insights. An early industry pioneer, PlaceIQ has become the standard for fueling better decisions for marketers, analysts and publishers through powerful location-based consumer insights, real-world measurement and attribution.

With PlaceIQ, companies can uncover opportunities within the consumer journey by learning about and connecting with location-based audiences, measuring real-world ROI, and applying insights that drive intelligent marketing and successful business outcomes.

The Enhanced Visits data set allows brands to use prepared location data to develop smarter marketing strategies and drive stronger real-world consumer actions.

PlaceIQ's location-intelligence platform ingests and processes billions of movement data points to understand the places consumer visit. This refined data is the core of PlaceIQ's premiere data sets.

Value Proposition & Differentiators

Scale, Quality & Expertise Make the Difference

Harnessing location intelligence enables you to drive measurable results that impact your entire business. Working with PlaceIQ means you'll benefit from our expertise, data quality, and the sheer scale of our data set.

PlaceIQ's core differentiators include:

- A peerless, patented location data processing platform: PlaceIQ Location Intelligence Platform is a robust processing and cleansing platform that performs all of our ingesting, processing, cleansing and clustering algorithms. Honed over the past 10 years, it is flexible enough to handle all data sets and is portable across geographies.
- Proprietary Basemap: PlaceIQ has curated a proprietary basemap utilizing 1st party hand drawn polygons verified utilizing satellite imagery, 3rd party POIs and additional spatial data. This precise basemap allows us to contextualize visits with confidence and at scale, whereas

simply licensing a 3rd party basemap built for routing/navigation introduces more error and imprecision

- **Proven Data Quality:** PlaceIQ has participated in a wide variety of 3rd party data evaluations competing with location companies from across the ecosystem. In all of these data evaluations we have fared incredibly well which has landed PlaceIQ as the only location partner in the Oracle Premier Provider Program, the preferred location provider for comScore, the sole provider into IRIs Liquid Data Platform, and the only location data provider to Experian (who has also invested strategically in PlaceIQ).
- **Coverage/Scale:** With over 140 MM+ MAUs in the US and an ever growing movement data portfolio, PlaceIQ's addressable universe is amongst the largest, if not the largest, in the space.
- **Diversity:** PlaceIQ's portfolio approach to movement data has allowed us to acquire a diverse data set of data that is nationally represented, and has been tested by 3rd parties to confirm its representativeness. Single source apps, SDKs and check-in/survey based panels introduce inherent biases which a multi-sourced portfolio eliminates.
- **Portfolio Optimization:** As a result of PlaceIQ's multi-sourced movement data approach, we've built an intense evaluation rubric that all data providers must go through prior to entering our production pipelines. This ensures that all our data sets are best in breed and no bad actors are allowed into our co-op. Additionally, this allows us to add, remove and optimize our partner set with relative ease.
- **Data Transparency:** Because we take pride in our data quality and integrity we provide a greater level of transparency into our dataset than competitors. All PlaceIQ data is provided down to the individual provider so clients have the ability to filter out and toggle between movement data providers as needed
- **Investment & Expertise:** PlaceIQ has invested considerably in equity and time to build out a scalable location analytics platform that can ingest and process a wide array of data. Customers have been ingesting and leveraging PlaceIQ's data asset for 4+ years for various analytics use cases.

Movement Data Sourcing

PlaceIQ's SDK-based location/movement data is sourced both directly (app publishers) and indirectly (reputable SDK data aggregators) at scale.

Having a variety of different data sources (known as “multi-sourcing”) has many benefits including maximizing coverage of high quality observable devices, neutralizing the data bias associated with a single app or SDK, eliminating any “single point of failure” or unmanaged churn in collection of data, persistence of devices observed across device life, and cross-validating accuracy of different types of data collected by app developers and operating systems.

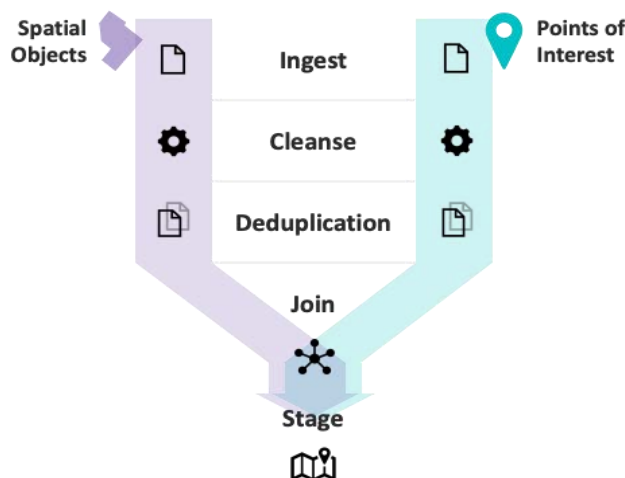
PlaceIQ maintains a record of where all data is sourced, so that they can be isolated and removed from different downstream products and analyses, according to our customers' business rules.

Our data sources are contractually obligated to require app developers they partner with to include appropriate consent and disclosure about their data collection practices. These developer's privacy policy and data collection practices are reviewed by PlaceIQ prior to completing an agreement with them (PlaceIQ only partners directly with developers that have appropriate consents and disclosures about their data collection practices).

Additionally, PlaceIQ movement data supplier ecosystem is CCPA compliant. Not only have we rolled CCPA in California, where it is legally required, but we've extended the force of that legislation across our entire business.

Basemap Creation

PlaceIQ's basemap consists of 1st party hand drawn polygons verified utilizing satellite imagery as well as 3rd party POIs, spatial objects and more. We've developed a proprietary algorithm that analyzes all the geospatial objects for a given location, de-dupes and selects the ideal polygon/parcel. The basemap is best thought of as a pipeline for managing spatial data and POI, similar to the below:



There are four distinct steps in the pipeline, outlined below:

- **Ingest:** Data is acquired and staged for loading into the pipeline.
- **Cleanse:** Data is standardized to a single format, verified, and collapsed into single records.
- **Deduplication:** Places and shapes are conflated into a single record based on their spatial or linguistic overlap. Overly complex spatial elements are joined or clustered as needed.
- **Join:** POI and Spatial Object datasets are associated with each other by comparing the diverse ingested datasets for geographic and contextual similarity.

Data Processing

Movement data ingested by PlaceIQ is fused with our propriety basemap in a six-stage process that starts with the ingestion of raw data and ends the output of visit data which is fed into our downstream products (Audiences, Measurement, & Analytics). These stages are:

- **Ingest:** In the ingest process PlaceIQ acquires, stages, and stores raw movement data from diverse partners. We also adapt external data schema into our own standardized format.

- **Time Fusion:** In the time fusion stage we compress the data set by fusing movement observations by time. This is done by identifying and joining identical or near-identical lat/longs.
- **Cleansing:** In the cleansing process we employ a proprietary algorithm which removes incorrect places, fraudulent devices, and spammy movement data that doesn't accurately reflect real world behaviors due to infrastructural or malicious artifacts.
- **Clustering:** The clustering phase looks to create two data sets: Visit Clusters (devices at rest) and Trajectories (devices in motion).
- **Visit Attribution:** The visit attribution phase determines real-world behaviors by associating the aforementioned visit clusters with POI by joining visitation clusters with our basemap dataset. The output of this step is visits to commercial locations like retailers, restaurants, hotels, airports, etc.
- **Dwell Attribution:** The dwell attribution phase detects locations repeatedly visited by a device. Once identified these are scored to create different HomeDwell, WorkDwell, TertiaryDwell, etc. locations for devices.

Product Overview: Place Intelligence Metrics

PlaceIQ's delivers aggregated metrics for up to 4+ million POIs allowing organizations to leverage PlaceIQ insights without ingest raw ID/event level data.

Key Facts

- Delivery Cadence: Weekly
- Data Latency: 11 Day Lag
- Data Size: 5-10GBs/week
- Preferred File Format: Parquet (csv supported)
- Preferred Delivery Method: s3, GCP, Azure (sFTP supported)

Data Structure

PlaceIQ's Place Intelligence Metrics data set is delivered as six tables which can be joined via a variety of key based joins. These tables are described below.

Schema

Table: METRIC

Name	Field Type	Description	Nullable?
spatial_id	Long	ID for the spatial feature the place is associated with	No
category	String	Category of the associated metric (CBG, Cross Shopping, Device Attribute, Measure, etc.)	No
sub_category	String	Sub-category of the associated metric (Home CBG, Frequency, Hourly Visitation, etc.)	Yes
metric	String	Count of devices visiting the spatial_ID broken down by demographics (HH Income, Age, HH Size, Education, Ethnicity)	Yes
value	String	Other brands devices who visited the spatial_ID visited in the same week	Yes
value_units	String	Unit of measure pertaining to the value (Device Count, Device Share, Visit Count, Miles, Minutes, etc.)	Yes
start_date	Date	First date of the week	No
end_date	Date	Last date of the week	No

Table: PLACE

Name	Field Type	Description	Nullable?
place_id	Long	Unique id for an individual place (point of interest)	No
spatial_id	Long	ID for the spatial feature the place is associated with	No
chain_id	String	ID for the commercial chain the place belongs to (if any)	No
category_ids	String	Comma separated strings with category ids	Yes
latitude	Double	Latitude of the place centroid, 5 decimal places	No
longitude	Double	Longitude of the place centroid, 5 decimal places	No
names	String	Comma separated string	Yes
address	String	Full address of the place	Yes
city	String	City the place is located in	Yes
state	String	State the place is located in	No
zip_code	String	Zip code the place is located in	No
census_block_group	String	Census Block Group Id the place is located in	No
dma_id	Integer	Designated Market Area Id the place is located in	No
cbsa_id	Integer	Core-based Statistical Area Id the place is located in	No
local_timezone	String	Time zone the place is located in	No
opening_date	Date	Date the place opened at the spatial feature	Yes
closing_date	Date	Date the place closed at the spatial feature	Yes
is_open	Boolean	Is the place currently active at this spatial feature	No
is_high_confidence	Boolean	Does the place support analytics grade use cases	Yes
is_multiuse	Boolean	Are multiple places occupying the same spatial feature (e.g. is this a mall or similar)	No
closed_hours	String	Comma separated integers representing the hours the place is closed	Yes

Table: CHAIN

Name	Field Type	Description	Nullable?
chain_id	String	Unique identifier for a chain	No
chain_name	String	Readable name of the chain	No
ticker_symbol	String	Financial ticker symbol if the chain is owned by a publicly traded entity	No
figi	String	Parent level Financial Instrument Global Identifier if the chain is owned by a publicly traded entity	No

Table: CATEGORY

Name	Field Type	Description	Nullable?
category_id	String	Unique identifier for a category	No
description	String	Description of what the category represents	No
category_type	String	Overall taxonomy the category is in (ex: SIC8 Code)	No

Table: DIM_DMA

Name	Field Type	Description	Nullable?
dma_code	Integer	The DMA code of the location where the visit occurred	No
dma_name	String	The DMA name of the location where the visit occurred	No

Table: DIM_MOVEMENT_SOURCE

Name	Field Type	Description	Nullable?
movement_source_key	Integer	Unique anonymized identifier representing the source of movement data from which PlaceIQ derived this visit event	No
movement_source_type	String	Tell if the source is "foreground", "background", "ad request" or "mixed"	No

Usage and Best Practices

- Metrics are associated directly with a spatial object (a polygon). One or more places can be associated with the spatial object for a period of time. In this way, if a given location changes ownership over time, the same spatial object will persist, but a new Place will be associated with the spatial object.
- In the case where a new polygon is drawn for a given location, a new spatial object will be created and the Place entry will be updated with the new spatial ID. The place ID will remain the same as long as the brand associated remains the same. A delta backfill will include the historical visitation to the new spatial object.
- In general, any entity with two or more locations will be represented with a chain. Single location entities (i.e. a "Mom and Pop" shop) will have a null chain. Every Place delivered in Firehose will have at least a chain or category (if not both). By using these fields in conjunction with Place metadata (e.g. City, State, Zip, etc) will enable you to narrow down to specific places of interest.
- IS_HIGH_CONFIDENCE is a boolean flag that denotes whether PlaceIQ has algorithmically deemed the given Place as having clean, analytical grade signal. Generally this means that the Place's polygon footprint and visitation behavior is reasonably in line with the chain's other locations.
- IS_MULTIIUSE is a boolean flag that will be true for a Place that lies within a spatial feature occupied by other Places. Examples include malls, airport terminals, and university common buildings. Because visitation can only be tied to a spatial feature, it is recommended that these are typically excluded from most use cases. If you opt to include these Places, be aware that the volume of visitation may not reflect actual foot traffic to the Place specifically

(for example, the Place may be a small store in a large mall, and the Place itself only sees a portion of the foot traffic in reality).

- **CLOSED_HOURS** is a string with comma separated integers corresponding to half hour increments for which the store is closed. The sequence begins with 1 representing Monday 00:00-00:29:59 and ends with 336 representing Sunday, 11:30-11:59:59. The [appendix](#) includes the full sequence.

Metrics Descriptions

- **Visit Frequency:** Visit Frequency represents the average amount of times a device visits a given spatial object in a given week.
- **Median Dwell Time:** Median Dwell Time represents the median amount of time a device spends at a given spatial object in a given week. This is presented in minutes.
- **Total Weekly Visitors:** Total weekly visitors represents the total number of devices (deduped) PlaceIQ observed at this spatial object in a given week from the panel used for PIM.
- **Visitation by Hour:** Visitation by Hour represents the total number of visits observed by PlaceIQ at each hour (local time) of the day, throughout the week.
 - Total Visitation by Hour for a given spatial object should be comparable, although not identical, to Total Weekly Visitors.
- **Visitation by Day of Week:** Visitation by Day of Week is a normalized visits dataset representing the normalized visit count PlaceIQ sees to a spatial object on a given day in given week.
 - This metric leverages the same normalization approach used in Daily Visit Counts.
 - Because this dataset is normalized, the counts will be noticeably different from those in Total Weekly Visitors and Visitation by Hour.
- **Weekly Cross Shopping:** Weekly Cross Shopping represents the total number of devices who visited the given spatial object AND another chain in the same week.
 - A device can only cross shop at a given chain once a week.
 - A device cannot cross shop at the same spatial object but can cross shop at the same brand (so long as it is a different spatial object).

- **Same Day Cross Shopping:** Same Day Cross Shopping represents the total number of devices who visited the given spatial object AND another chain in the same day.
 - A device can only cross shop at a given chain once a day, but can cross shop at that chain multiple days per week. The maximum number of cross shopping days by a device in a week is 7. Because a device can cross shop on 7 days in a given week, the total number of Same Day Cross Shopping events will often exceed the Weekly Cross Shopping count for the same spatial object and chain pair.
 - A device cannot cross shop at the same spatial object but can cross shop at the same brand (so long as it is a different spatial object).
- **Home CBG of Visitors:** Home CBG of Visitors represents the # of devices who've visited the spatial object from a specific Home Census Block Group in a given week.
- **Daytime CBG of Visitors:** Home CBG of Visitors represents the # of devices who've visited the spatial object from a specific Daytime Census Block Group in a given week.
- **Median Distance From Home CBG:** Median Distance From Home CBG represents the median number of miles from the Home Census Block Group to the spatial object.
- **Median Distance From Daytime CBG:** Median Distance From Daytime CBG represents the median number of miles from the Daytime Census Block Group to the spatial object.
- **Device Attributes:** Device attributes represents the fraction of visitors to the given spatial object who have a specific attribute (Carrier & Device Type).
- **Demographics:** Demographics represents the fraction of visitors to the given spatial object who have a specific set of demographics (Age, Ethnicity, Income, Education Level).

Data Delivery: Additional Details

Data Structure

All PlaceIQ deliveries adhere to a structured upload spec which is outlined below. A sample path indicating how part files may be delivered is also included.

Root	Version	File Format	Table Name	Date	Part
CLIENT SPECIFIC	v4_1	csv	visit	dt=20200915	part-00037-9d8dd29d-7c69-4c4c-b227-f188ed7e1c1b-c000.csv.gz
CLIENT SPECIFIC	v4_1	parquet	place	dt=20200907	part-00050-934cd2d5-5fd5-412f-ba7c-121785252f29-c000.csv.gz
CLIENT SPECIFIC	v4_1	parquet	category	dt=20200907	part-00041-1be182af-cd19-4ed2-8878-9c2194506189-c000.csv.gz
CLIENT SPECIFIC	v4_1	csv	chain	dt=20200907	part-00099-a8253e8f-bde3-4a9e-9148-48a254000041-c000.csv.gz

...../landmark_v4_1/csv/visit/dt=20200915/part-00037-9d8dd29d-7c69-4c4c-b227-f188ed7e1c1b-c000.csv.gz

Metadata

All PlaceIQ deliveries have two metadata tables included so that customers can verify that a full delivery has been ingested. These tables are called `dataset_metadata` and `file_metadata`. Schemas can be found below:

Table: DATASET_METADATA

Name	Field Type	Description	Nullable?
dataset_name	String	The name of the dataset	No
number_of_part_files	Long	Total number of part files for the dt	No
row_count	Long	Total number of rows in all part files for the dt	No

Table: FILE_METADATA

Name	Field Type	Description	Nullable?
file_name	String	The part file name	No
row_count	Long	Total number of rows in part file	No
byte_count	Long	Size of part file in bytes	No
md5_checksum	String	md5 checksum of part file	No
dataset_name	String	The name of the dataset part file belongs to	No

Appendix

Closed Hours

	1 - Mon	2 - Tue	3 - Wed	4 - Thu	5 - Fri	6 - Sat	7 - Sun
00:00-00:30 AM	1	49	97	145	193	241	289
00:30-01:00 AM	2	50	98	146	194	242	290
01:00-01:30 AM	3	51	99	147	195	243	291
01:30-02:00 AM	4	52	100	148	196	244	292
02:00-02:30 AM	5	53	101	149	197	245	293
02:30-03:00 AM	6	54	102	150	198	246	294
03:00-03:30 AM	7	55	103	151	199	247	295
03:30-04:00 AM	8	56	104	152	200	248	296
04:00-04:30 AM	9	57	105	153	201	249	297
04:30-05:00 AM	10	58	106	154	202	250	298
05:00-05:30 AM	11	59	107	155	203	251	299
05:30-06:00 AM	12	60	108	156	204	252	300
06:00-06:30 AM	13	61	109	157	205	253	301
06:30-07:00 AM	14	62	110	158	206	254	302
07:00-07:30 AM	15	63	111	159	207	255	303
07:30-08:00 AM	16	64	112	160	208	256	304
08:00-08:30 AM	17	65	113	161	209	257	305
08:30-09:00 AM	18	66	114	162	210	258	306
09:00-09:30 AM	19	67	115	163	211	259	307
09:30-10:00 AM	20	68	116	164	212	260	308
10:00-10:30 AM	21	69	117	165	213	261	309
10:30-11:00 AM	22	70	118	166	214	262	310
11:00-11:30 AM	23	71	119	167	215	263	311
11:30-Noon	24	72	120	168	216	264	312
12:00-12:30 PM	25	73	121	169	217	265	313
12:30-01:00 PM	26	74	122	170	218	266	314
01:00-01:30 PM	27	75	123	171	219	267	315
01:30-02:00 PM	28	76	124	172	220	268	316
02:00-02:30 PM	29	77	125	173	221	269	317
02:30-03:00 PM	30	78	126	174	222	270	318
03:00-03:30 PM	31	79	127	175	223	271	319
03:30-04:00 PM	32	80	128	176	224	272	320
04:00-04:30 PM	33	81	129	177	225	273	321
04:30-05:00 PM	34	82	130	178	226	274	322
05:00-05:30 PM	35	83	131	179	227	275	323
05:30-06:00 PM	36	84	132	180	228	276	324
06:00-06:30 PM	37	85	133	181	229	277	325
06:30-07:00 PM	38	86	134	182	230	278	326
07:00-07:30 PM	39	87	135	183	231	279	327
07:30-08:00 PM	40	88	136	184	232	280	328
08:00-08:30 PM	41	89	137	185	233	281	329
08:30-09:00 PM	42	90	138	186	234	282	330
09:00-09:30 PM	43	91	139	187	235	283	331
09:30-10:00 PM	44	92	140	188	236	284	332
10:00-10:30 PM	45	93	141	189	237	285	333
10:30-11:00 PM	46	94	142	190	238	286	334
11:00-11:30 PM	47	95	143	191	239	287	335
11:30-Midnight	48	96	144	192	240	288	336