

# Weekly Patterns+

Weekly Patterns+ data provides the same foot traffic data insights as Monthly Patterns+ on a weekly basis, tracking data from Monday to the end of day on Sunday each week. Weekly Patterns+ are available starting from January 1st, 2018.

## Table of Contents


- Weekly Patterns+ Schema ..... 2**
- Optional: Panel Overview Data ..... 7**
  - Home Location Distributions by State/Census Block Group ..... 7**
  - Number of Visits/Visitors by State ..... 8**
  - Normalization Stats ..... 9**
- Key Concepts ..... 10**
- Column Name Detailed Descriptions ..... 11**
  - street\_address ..... 11
  - city ..... 11
  - region ..... 11
  - postal\_code ..... 12
  - visit\_counts ..... 12
  - visits\_by\_day ..... 12
  - visits\_by\_each\_hour ..... 12
  - visitor\_home\_cbgs ..... 12
  - visitor\_home\_aggregation ..... 12
  - visitor\_daytime\_cbgs ..... 13
  - visitor\_country\_of\_origin ..... 13
  - distance\_from\_home ..... 13
  - median\_dwell ..... 13
  - bucketed\_dwell\_times ..... 13
  - related\_same\_day\_brand ..... 14
  - related\_same\_week\_brand ..... 14
  - carrier\_name ..... 14





# Weekly Patterns+ Schema



File Names: [patterns\_plus\_\*.csv.gz]


Column Name	Description	Type	Example
id_store	Unique ID tied to this point of interest (" <b>POI</b> "). This ID <i>guaranteed to be persistent</i> across the lifespan of the company, excluding M&A events. In M&A situations, a different id_store is created and the pre/post merger locations are tracked separately. See field <code>persistent_id_store</code> below for how to handle M&A cases	Integer	6255608
ticker	The exchange-traded ticker, if available, or a custom ticker for private companies and subsidiaries.	String	COST
persistent_id	Unique, human-readable identifier that is <i>guaranteed to be persistent even across M&amp;A events</i> that represents this specific company.	String	A1_UPS
persistent_id_store	Persistent ID tied to this point of interest (" <b>POI</b> "). This ID <i>guaranteed to be persistent</i> across the lifespan of the location, <i>including</i> M&A events.	Integer	6255608
is_distributor	Whether this is a company owned location (the field is <code>null</code> ) or it is a distribution location for the company's products (the field is 1).	Integer	null
location_name	The name of the point of interest.	String	Salinas Valley Ford Lincoln
street_address	Street address of the point of interest.	String	1100 Auto Center Circle

city	The city of the point of interest.	String	Irvine
region	The state, province or county of the point of interest.	String	CA
postal_code	The postal code of the place of interest.	String	92602
iso_country_code	The 2 letter <a href="#">ISO 3166-1 alpha-2 country code</a> .	String	US
brand	If this POI is an instance of a company that we have explicitly identified, this column will contain the company DBA ("doing business as") name.	String	UPS
open_date	The date the store location opened. By convention, 1970-01-01 is the "start of time" (Unix Start Time) which in this context denotes any location that opened before 1/1/2010.	YYYY-MM-DD	1970-01-01
close_date	The date the store location closed. By convention, 2038-01-01 is the "end of time" (Unix End Time) which in this context denotes any location that is still open.	YYYY-MM-DD	2038-01-01
longitude	Longitude for POI's location	Float	-112.4556
latitude	Latitude for POI's location	Float	38.3441
naics_code	North American Industry Classification System Code	String	722511
top_category	4 digit NAICS description	String	Restaurants and Other Eating Places
sub_category	6 digit NAICS description	String	Full-Service Restaurants
poi_cbg	The <a href="#">census block group</a> (U.S.) or <a href="#">dissemination area</a> (Canada) the POI is located within.	String	560610112022
msa_code	5-digit Metropolitan Statistical Area	String	22300
date_range_start	Start time for measurement period in ISO 8601 format of YYYY-MM-	String	2020-03-02T00:00:00-06:00

	DDTHH:mm:SS±hh:mm (local time with offset from GMT). The start time will be 12 a.m. Monday in local time.		
date_range_end	End time for measurement period in ISO 8601 format of YYYY-MM-DDTHH:mm:SS±hh:mm (local time with offset from GMT). The end time will be the following Monday at 12 a.m. local time.	String	2020-03-09T00:00:00-06:00
visit_counts	Number of estimated visits (i.e., the sum of each day's unique visitors across all the days in the date range) to this POI during the date range. The estimated visits are scaled using Advan's best current methodology for estimating actual visits.	Integer	1542
visitor_counts	Number of unique estimated visitors to this POI during the date range. The estimated visitors are scaled using Advan's best current methodology for estimating actual visitors.	Integer	1221
visits_by_day	The number of estimated visits to the POI each day (local time) over the covered time period.	JSON [Integer]	[33, 22, 33, 22, 33, 22, 22]
visits_by_each_hour	The number of estimated visits to the POI for each of the 168 hours of the week, starting at midnight on date_range_start.	JSON [Integer]	[33, 22, 33, 22, 33, 22, 22, 21, 23, 33, 22, 11, 44, 22, 22, 44, 11, 33, 44, 44, 44, 33, 34, 44, 22, 33, 44, 44, 34, 43, 43...]
 visitor_home_cbgs	The number of estimated visitors to the POI from each census block group or dissemination area based on the visitor's home location.	JSON {String: Integer}	{"360610112021": 603, "460610112021": 243, "560610112021": 106, "660610112021": 87, "660610112021": 51}

 visitor_home_aggregation	The number of estimated visitors to the POI from each census tract or aggregate dissemination area based on the visitor's home location.	JSON {String: Integer}	{ "17031440300": 1005, "18089021500": 522, "17197883516": 233, "17031826402": 5, "17031826301": 4, "04013115802": 4 }
 visitor_daytime_cbgs	The number of estimated visitors to the POI from each census block group or dissemination area based on visitor's primary daytime location on weekdays (M-F) between 8 am - 6 pm.	JSON {String: Integer}	{ "360610112030": 9872, "880610112021": 8441, "569610112020": 5671, "160610112041": 2296, "980610112021": 1985 }
 visitor_country_of_origin	The number of estimated visitors to the POI from each country based on visitor's home country code.	JSON {String: Integer}	{ "US": 98, "CA": 12 }
 distance_from_home	Median distance from home travelled by visitors (of visitors whose home we have identified) in meters.	Integer	1211
median_dwell	Median minimum dwell time in minutes.	Double	5
bucketed_dwell_times	The distribution of visit dwell times based on pre-specified buckets. Key is the range of dwell time in minutes and value is number of estimated visits that were within that range.	JSON {String: Integer}	{ "<5": 40, "5-20": 22, "21-60": 45, "61-240": 3, ">240": 5 }
related_same_day_brand	Other brands that the visitors to this POI visited on the same day as the visit to this POI. Limited to top 20. The field contains: "dba": estimated visitors	JSON {String: Integer}	{ "mcdonalds": 7, "amc": 5, "target": 3 }
related_same_week_brand	Other brands that the visitors to this POI visited in the same week as the visit to this POI. Limited to top 20. The field contains: "dba": estimated visitors	JSON {String: Integer}	{ "mcdonalds": 7, "amc": 5, "target": 3 }

 device_type	The number of estimated visitors to the POI that are using Android vs. iOS.	JSON {String: Integer}	{"android": 6, "ios": 8}
 <sup>1</sup> carrier_name	The number of estimated visitors to the POI based on the wireless carrier of the device.	JSON {String: Integer}	{"Verizon": 342, "T-Mobile": 288, "AT&T": 265}
Reserved1	Provided for backwards compatibility with legacy feeds. This field is currently set to NULL but Advan reserves the right to use it for other information in the future	Float	NULL
Reserved2	Provided for backwards compatibility with legacy feeds. This field is currently set to NULL but Advan reserves the right to use it for other information in the future	Float	NULL
Reserved3	Provided for backwards compatibility with legacy feeds. This field is currently set to NULL but Advan reserves the right to use it for other information in the future	Float	NULL
Reserved4	Provided for backwards compatibility with legacy feeds. This field is currently set to NULL but Advan reserves the right to use it for other information in the future	Float	NULL
Reserved5	Provided for backwards compatibility with legacy feeds. This field is currently set to NULL but Advan reserves the right to use it for other information in the future	Float	NULL

 We do not report data if less than 2 visitors are observed from that group. If there are between 2 and 4 visitors this is reported as 4. These numbers are before scaling to estimated visitors; the equivalent number of estimated visitors will be higher.

<sup>1</sup> carrier\_name is a premium column. Please [Contact Us](#) for more details.

<sup>2</sup> for trade area computations we utilize additional panels, which may in some cases result in total sum of visitors by CBG to be higher than visitor\_counts.

## Optional: Panel Overview Data

For conformity with legacy products, the Panel Overview Data is available upon request (see tables below). Since Patterns+ is already normalized for the panel, the device home distributions in this table is neither necessary nor recommended for normalization purposes. If you need the number of residents and number of people with a primary daytime location inside a census block group, a State, MSA, County, Zip code, etc., the recommended data to use are the **"Demographics by CBG"**, which is calibrated to estimate the actual population and also contains demographic information of the residents/visitors.

### Home Location Distributions by State/Census Block Group

**File Names:** [home\_panel\_summary.csv]

Column Name	Description	Type	Example
date_range_start	Start time for measurement period in ISO 8601 format of YYYY-MM-DDTHH:MM:SS±hh:mm (local time with offset from GMT). The start time will be 12 a.m. Sunday in local time.	String	2020-03-01T00:00:00-06:00
date_range_end	End time for measurement period in ISO 8601 format of YYYY-MM-DDTHH:MM:SS±hh:mm (local time with offset from GMT). The end time will be the following Sunday at 12 a.m. local time.	String	2020-03-08T00:00:00-06:00
region	Lowercase abbreviation of U.S. state or territory, or the Canadian Province or territory, depending on iso_country_code	String	ny

iso_country_code	The 2 letter ISO 3166-1 alpha-2 country code.	String	US
census_block_group	US <a href="#">FIPS code</a> for this <a href="#">Census block group</a> , or the Canadian <a href="#">Dissemination area</a> designation (Canadian units have CA: as a prefix)	String	530330080012
number_devices_residing	Number of distinct estimated devices (people) observed with a primary nighttime location in the specified census block group.	Integer	54481
number_devices_primary_daytime	Number of distinct estimated devices (people) observed with a primary daytime location in the specified census block group.	Integer	54482

## Number of Visits/Visitors by State

**File Names:** [visit\_panel\_summary.csv]

Note: Includes three rows designated ALL\_US, ALL\_CA, and ALL to provide total visitors seen during the week in the US, in Canada, and in both countries (might be less than sum of visitors by state due to same visitors having visits in multiple states).

Column Name	Description	Type	Example
starting_year	Calendar year that the first day of the weekly data is in	Integer	2018
starting_month	Calendar month that the first day of the weekly data is in	Integer	1
starting_day	Calendar day of the first day of the weekly data	Integer	1
region	Lowercase abbreviation of U.S. state or territory, or the Canadian	String	ny



	Province or territory, depending on <code>iso_country_code</code>		
<code>iso_country_code</code>	The 2 letter ISO 3166-1 alpha-2 country code.	String	US
<code>num_visits</code>	Number of point-of-interest visits observed in the specified state	Int	8900
<code>num_unique_visitors</code>	Number of unique visitors observed with at least 1 point-of-interest visit in the specified state	Integer	966

## Normalization Stats

**File Names:** [normalization\_stats.csv]

We also provide a file containing information on our panel for each day of the week to help with normalization. This file will have rows for each region for each day of the week.

Note: This includes the ALL\_US, ALL\_CA, and ALL rows similar to `visit_panel_summary.csv`

Column Name	Description	Type	Example
<code>year</code>	Calendar year	Integer	2019
<code>month</code>	Calendar month starting from 1 as January	Integer	1
<code>day</code>	Calendar day	Integer	1
<code>region</code>	When <code>iso_country_code == US</code> , then this is the USA state or territory. When <code>iso_country_code == CA</code> , then this is the Canadian Province or territory.	String	CA
<code>iso_country_code</code>	The 2 letter ISO 3166-1 alpha-2 country code.	String	US
<code>total_visits</code>	All visits we saw on the given day in local time (includes visits to POI and visits to homes)	Integer	200

<code>total_devices_seen</code>	Total devices in our panel which we saw on the given day with any visit in local time (POI or home visit)	Integer	50
<code>total_home_visits</code>	Visits we saw on the given day in local time to the device's home geohash-7	Integer	120
<code>total_home_visitors</code>	Total devices we saw on the given day with at least 1 visit to the device's home geohash-7	Integer	35

## SUCCESS and HISTORY RESTATED Flags

The following empty files are placed in the bucket(s) periodically:

- `_SUCCESS_PPLUS_WEEKLY_YYYYMMDD`: placed in the associated GCS or AWS bucket as soon as all the weekly files are ready.
- `_HISTORY_RESTATED_PPLUS_WEEKLY_YYYYMMDD`: placed in the associated GCS or AWS bucket whenever the historical data is restated, signaling the restatement is complete and ready to be loaded.

## Key Concepts

Weekly Patterns+ uses the same logic as our Monthly Patterns+ product for key concepts such as:

- **Visit Attribution:** we compute the visits/visitors and other metrics inside a POI using the POI's geometry. We do not apply any dwell time or any concept of "stops"; we rely on the polygon for accuracy. We have tested our data on 1,500 publicly traded tickers versus (a) top line revenue as reported from the companies and (b) credit card transaction counts on physical locations, and we have determined consistently that in the vast majority of cases filtering for dwell time reduces the signal and makes the correlation/forecasting worse.
- **Determining Home Location:** we compute a device's home/work (night/day) location by computing the time a device spent in each building in the country; then taking the most frequented building.
- **Backfills:** Backfill is when we take our most recent locations (i.e., addresses + geofences) and run our visit attribution algorithm backward in time to generate a new history of "backfilled" Patterns+. Backfills are typically generated every time new Advan POIs are added (typically monthly, with the exception of August and December).

The best way to think about the two products is that the underlying visits and algorithms are the same, but they are aggregated at different timescales and delivered at difference frequencies.

Below are some differences between Weekly Patterns+ and Monthly Patterns+:

1. Each delivery of Weekly Patterns+ covers one week starting Monday and ending end of day on Sunday. The data is available three days later on Wednesday of each week, providing more frequent actionable data.
2. In Weekly Patterns+, we include a `visits_by_each_hour` column to enable you to get a more detailed view of the week.
3. Weekly Patterns+ does not include `popularity_by_hour` (covered by `visits_by_each_hour`) and `popularity_by_day` (covered by `visits_by_each_hour`).
4. Our list of POIs are updated monthly. When an updated POI list is used for generating the Weekly or Monthly Patterns+, the traffic is being computed for the full history of any new or updated POIs (a “backfill”, or a “release”). The majority of the changes pertain to new POIs, but there are a number of POIs (on the order of 1% of the total) that are updated on each release; such updates typically modify the date the POI may have closed, i.e., we mark any store location that closed since the prior POI update as closed in the latest POI version, or improve the accuracy of the polygons delineating the POIs. If you require a very high level of accuracy for your historical visits comparisons, you may want to load the historical back-filled data on each release; for most use cases however, you do not need to reload the historical data. You can also employ an intermediate approach where you reload the history only on quarterly, semi-annual, or annual basis.
5. Weekly Patterns+ includes Canada foot traffic data, unlike Monthly Patterns+ which contains only US data. This means there are also some subtle changes to how columns behave and names of columns that make it distinct from the Monthly Patterns+ product (e.g., the `state` column in Supplemental Files is named `region` in Weekly Patterns+).

## Column Name Detailed Descriptions

### `street_address`

- We implement a number of steps to clean, validate and standardize `street_address`.
- You should expect `street_address` to be title-cased, consistent, and friendly for human reading. Please send us your feedback if you see otherwise.

### `city`

- City names are the output of normalized address strings from POI sources.

### `region`

- When `iso_country_code == US`, then this is the US state or territory.
- When `iso_country_code == CA`, then this is the Canadian Province or territory.

- When `iso_country_code == GB`, then this is the [United Kingdom country](#).

### **postal\_code**

- When `iso_country_code == US`, then this is the US 5 digit zip code.
- When `iso_country_code == CA`, then this is the [Canadian postal code](#) in the form of a 3 digit Forward Sortation Area (FSA), a space, and the 3 digit Local Delivery Unit (LDU).
- When `iso_country_code == GB`, then this is the British postal code. [Learn more about Great Britain postal code precision here](#).

### **visit\_counts**

Number of estimated visits (i.e., the sum of each day's unique visitors across all the days in the date range) to this POI during the date range. The estimated visits are scaled using Advan's best current methodology for estimating actual visits.

### **visits\_by\_day**

- This is an array of visits on each day in the week, Monday through Sunday.
- We are breaking up days based on local time.

### **visits\_by\_each\_hour**

- This is an array of visits for each hour that exists in the week.
- Only the start of a visit is used for this column so multi-hour visits will only be counted once unless the visit crosses a UTC day boundary. Note that this is a subtle difference compared to how `popularity_by_hour` is calculated in Monthly Patterns+.
- We are breaking up days based on local time.

### **visitor\_home\_cbgs**

- These are the home census block groups (U.S.) or dissemination areas (Canada) of the visitors to the POI.
- For each census block group, we show the number of associated *visitors* (as opposed to the number of *visits*). If visits by home cbg is desired, we recommend taking the *visitors* from each CBG and multiplying by the average visits/visitor (i.e., `raw_visit_counts / raw_visitor_counts`) as an approximation.
- We do not have a home census block group for each visitor and not each visitor originates from the US. The number of US visitors listed in the `visitor_country_of_origin` column represents the total number of visitors which we have determined originate from the US versus Canada.

### **visitor\_home\_aggregation**

- This is similar to `visitor_home_cbgs` except they represent the home *census tracts* (U.S.) or *aggregate dissemination areas* (Canada) of the visitors to the POI.
- We recommend using this column when you do not need to know visitor homes areas at such a fine level (CBGs represent 600-3000 people while DAs represent 400-700), but can aggregate to the next-level-up geographic unit (CTs represent 2,500 to 8,000 people while ADAs represent 5,000 to 15,000).

#### **`visitor_daytime_cbgs`**

- These are the daytime census block groups of the visitors to the POI, namely the most frequent building that the visitors spend 8am-6pm M-F during the calendar month of the data.
- For each census block group, we show the number of associated *visitors* (as opposed to the number of *visits*).

#### **`visitor_country_of_origin`**

- These are the countries of origin of the visitors to the POI.

#### **`distance_from_home`**

- This is the median distance from home to the POI in meters for the visitors we have identified a home location.
- This is calculated by taking the haversine distance between the visitor's home geohash-7 and the location of the POI for each visit. We then take the median of all of the home-POI distance pairs.
- If we have fewer than 5 visitors to a POI, the value will be null.
- We do not adjust for visits - each visitor is counted equally.

#### **`median_dwell`**

- This is the median of the minimum dwell times we have calculated for each of the visits to the POI.
- We determine the minimum dwell time by looking at the first and last ping we see from a device during a visit. This is a minimum dwell because it is possible the device was at the POI longer than the time of the last ping.
- It is possible to have a minimum dwell of 0 if we only saw 1 ping and determined the visit based on factors such as wifi.

#### **`bucketed_dwell_times`**

- This is a dictionary of different time spans and the number of visits that were of each duration.

- The time spans are in minutes.
- Data contains the following bins: { "<5", "5-10", "11-20", "21-60", "61-120", "121-240", ">240" }

#### **related\_same\_day\_brand**

- These are the brands that the visitors to this POI also visit, on the same day that they visit the POI. The number mapped to each brand is an indicator of how highly correlated a POI is to a certain brand. The value is a simple percent of POI visitors that visited the other brand on the same day.
- Only the first 20 brands are returned.

#### **related\_same\_week\_brand**

These are the brands that the visitors to this POI visit over the course of the week. Interpreted and calculated in the same way as `related_same_day_brand`.

#### **carrier\_name**

- This is a premium column that maps wireless carrier names to the number of visitors to the POI whose device uses that wireless carrier.