☰   🔍

Places

✎

SafeGraph's Places data provides baseline information for every record in the product suite, including location name, address, lat/long, category, brand, and more! With POI data for countries around the world, you can gain insights about any location that a person can visit aside from private residences.

Check out our [Summary Statistics](#) for detailed information on our coverage.

Contents:

- [Schema](#)
- [Key Concepts](#)
- [Column Details](#)
- [Data Issues or Artifacts](#)
- [Video Overview](#)

Places Schema

[core_poi.csv]

SafeGraph updates the Places dataset every month with the past month's openings and closings and maintains a persistent `placekey` across releases.

All SafeGraph datasets are formatted as delimited CSVs, and can be purchased independently or together. Please reference the [SafeGraph Attribute Matrix](#) for a breakdown of which columns are available per product, and reference [Column Ordering](#) for specific column orderings per product combination. Otherwise, stand-alone purchases will follow the column ordering listed in the schemas below:

Column Name	Description	Type	Example
placekey	Unique and persistent ID tied to this POI. See [Placekey](#) for details on placekey design.	String	
parent_placekey	If place is encompassed by a larger place (e.g. mall, airport), this lists the placekey of the parent place; otherwise `null`. See more on parent-child relationships in [Spatial Hierarchy](#).	String	
location_name	The name of the place of interest.	String	Salinas Valley Ford Lincoln
safegraph_brand_ids	Unique and consistent ID that represents this specific [brand](#).	List	SG_BRAND_59dcabd7cd2395a2, SG_BRAND_8310c2e3461b8b5a
brands	If this POI is an instance of a larger brand that we have explicitly identified, this column will contain that brand name. See more details in [brands](#).	List	Ford, Lincoln
† store_id	The unique ID associated with the store as provided and maintained by the store/brand itself. Please [contact sales](#) to learn more	String	36558

Column Name	Description	Type	Example
	about licensing this premium column.		
top_category	The label associated with the first 4 digits of the POI's [NAICS category](#).	String	Automobile Dealers
sub_category	The label associated with all 6 digits of the POI's [NAICS category](#). For POIs with a 4-digit NAICS category, this column is `null`	String	New Car Dealers
naics_code	4-digit or 6-digit [NAICS code](#) describing the business.	Integer	441110
latitude	[Latitude](#) coordinate of the place of interest.	Float	36.714767
longitude	[Longitude](#) coordinate of the place of interest.	Float	−121.662912
street_address	[Street address](#) of the place 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 place of interest. See [region](#) for more details.	String	CA
postal_code	The [postal code](#) of the place of interest.	String	92602
iso_country_code	The 2 letter [ISO 3166-1 alpha-2 country code](#). Expected values are `US`, `CA`, and `GB`.	String	US
phone_number	The [phone number](#) of this POI	String	+14151234567
† websites	The website(s) associated with the POI as provided and maintained by the POI itself. Please [contact sales](#) to learn more about licensing this premium column.	List	[ford.com,lincoln.com]
open_hours	A JSON string with days as keys and opening & closing times (in the POI's local time) as values. See [open_hours](#) for more details.	String	{ "Mon": [["8:00", "22:00"]], "Tue": [["8:00", "13:00"], ["18:00", "24:00"]], "Wed": [["0:00", "2:00"]], "Thu": [["0:00", "24:00"]], "Fri": [["23:00", "24:00"]], "Sat": [["0:00", "3:00"], ["15:00", "22:30"]], "Sun": [] }
category_tags	An array of descriptive tags indicating higher-resolution category information. See [category_tags](#) for more details.	List	[Mexican Food,Casual Dining,Lunch,Dinner]
opened_on	The outside year and month this POI opened in yyyy-mm format. If null, then we do not have enough metadata to determine an open date. See	String	2019−10

Column Name	Description	Type	Example
	the open_on logic for more details.		
closed_on	The outside year and month this POI closed in yyyy-mm format. If null, then this POI is open. See the closed_on logic for more details.	String	2020-03
tracking_closed_since	Indicates the year and month we started tracking "closed_on" for this POI. See the closed_on logic for more details.	String	2019-07
geometry_type	The geometric shape associated with this POI. Possible values are: 1) POLYGON : POI has a polygon and geometry metadata or is intended to have a polygon and geometry metadata once Geometry is available in the given country . 2) POINT : POI intentionally does not have a polygon nor Geometry metadata. See geometry_type for more details.	String	POINT

† `store_id` and `websites` are premium columns. Please Contact Sales for more details.

Brand Info

[brand_info.csv]

A SafeGraph `brand` is defined as a logo or branded store which has multiple locations all under the same logo or store banner. For a deep dive on how we think about brands, see our November 2018 Release Notes.

The brand_info file is a separate csv that is complimentary with a core places purchase. See brands for more details.

Column Name	Description	Type	Example
safegraph_brand_id	Unique and persistent ID that represents this specific brand.	String	SG_BRAND_59dcabd7cd2395a2
brand_name	This is the brand_name corresponding to the safegraph_brand_id .	String	Ford Motor Company
parent_safegraph_brand_id	There are 2 possible values: 1) If this brand has a parent, this will list the ID of the parent brand. 2) If this brand has no parent, this will be null .	String	SG_BRAND_8310c2e3461b8b5a
naics_code	4-digit or 6-digit NAICS code describing the business.	Integer	441110
top_category	The label associated with the first 4 digits of the POI's NAICS category.	String	Automobile Dealers
subcategory	The label associated with all 6 digits of the POI's NAICS category. For POIs with a 4-digit NAICS category, this column is null	String	New Car Dealers

Column Name	Description	Type	Example
stock_symbol	The stock ticker (if the corporation is traded publicly)	String	F
stock_exchange	The stock exchange on which this corporation is listed (if the corporation is traded publicly).	String	NYSE
iso_country_codes_open	A list of all 2 letter ISO 3166-1 alpha-2 country codes for each country this brand has at least 1 open POI (closed_on is null).	String	["US", "GB"]
iso_country_codes_closed	A list of all 2 letter ISO 3166-1 alpha-2 country codes for each country this brand has at least 1 closed POI (closed_on is not null).	String	["US", "CA"]

Key Concepts

`Places Scope`

Places provides baseline information for every record in the SafeGraph product suite via the Places schema and polygon information when applicable via the [Geometry schema]. The current scope of a **place** is defined as any location humans can visit with the exception of single-family homes. This definition encompasses a diverse set of places ranging from restaurants, grocery stores, and malls; to parks, hospitals, museums, offices, and industrial parks. Premium sets of Places include apartment buildings (`naics_code` = 531110), Parking Lots (coming soon), and Point POIs (Places schema only).

SafeGraph Core Places is a global offering with varying coverage depending on the country. Note that address conventions and formatting vary across countries. SafeGraph has coalesced these fields into the Core Places schema.

Geometry is currently offered in the US, Canada, and Great Britain. 🇺🇸 🇨🇦 🇬🇧 Geometry for key brands and categories outside of US, CA, and GB will begin a phased roll-out starting in Q1 of 2022 🎉🌎 .

👀 Are we missing a brand or country? 👀 Please let us know here! We are adding hundreds of new brands across all countries every month, and we prioritize our brand queue based on your feedback. If you have more general product feedback or questions, please contact us here.

`placekey`

Placekey is a unique and persistent identifier for any physical place in the world that intelligently partitions the ID into meaningful encodings. So how does Placekey work?

When both parts of a placekey come together, the final result reads as This is a unique way of shedding light on both the descriptive element of a place as well as its geospatial position in the physical world via a single identifier.

What: Address Encoding
The first three characters refer to the Address Encoding, creating a unique identifier for a given address. An address at "555 Main Street Suite 105" will have a different Address Encoding than "555 Main Street Suite 106." However, "444 Second Street, Suite 4" will have the same address encoding as "444 2nd St. #4" to adjust for common address formats.

What: POI Encoding
The second set of three characters in the 'What Part' refers to the POI Encoding. If a specific place has a location name (like "Central Park") and is already included in the Placekey reference datasets, these characters will be present. The benefit of the POI Encoding is that it can point to a specific point of interest that may have existed at a certain address at a given point in time.

Where: H3 Encoding
The 'Where Part,' on the other hand, is made up of three unique character sequences, built upon Uber's open source H3 grid system. This information in the 'Where Part' is based on the centroid of that place. In other words, we take the latitude and longitude of a specific place and then use a conversion function to determine a hexagon in the physical world, representing about 15,000 sq. meters, containing the centroid of that place. The 'Where Part' of the Placekey is, therefore, the full encoding of that hexagon.

Open access to your own datasets using the FREE Placekey API.

Point POIs

Some places are small and not well defined by a geometric shape. We refer to these places as "Point POIs" and intentionally do not offer a polygon nor Patterns data. Places like transit stops, ATMs, kiosks, and electric vehicle charging stations are examples of Point POIs found in our data, and we flag these by setting the `geometry_type` column = "POINT." Point POIs are a premium portion of the Core places offering, and we are continually adding new types and brands.

Brands

SafeGraph curates over 7,700 distinct brands and growing. These are chains of commercial POIs that include all major brands in the United States, Canada, and Great Britain (McDonald's, AMC, Macy's, Chevrolet, Whole Foods Market, etc.).

Note that ~80% of POIs have no brand associated as they are single commercial locations (local restaurants, museums, etc.). SafeGraph is continually improving the fill rate of brands with each release - please contact us if you notice a brand missing.

Some POIs include multiple brands. For example, a car dealership may sell multiple car brands, or branded POIs may be co-located (Ex: Taco Bell and KFC in the same space; IMAX and AMC cinemas in the same building). In these cases, the `brands` and `safegraph_brand_ids` are listed as an array that is alphabetized by brand name (the order does not specify any importance).

Brands provide an easy way to isolate major stores. If you know you are searching for a brand that we cover, we advise searching by the `brands` column instead of the `location_name` column. For even better specificity, search the brand_info file by `brands` and build your workflows around `safegraph_brand_id`.

Every place has a `location_name`, but only POIs belonging to a chain will have a `brand`. In some cases, `location_name` and `brand` will be the same, but in other cases they are intentionally different. For example, the most common name for an individual Starbucks store is its brand name, so it is also reflected in the `location_name` column. However, the most common name for the Bellagio Hotel & Casino is not its brand name "MGM Resorts." In this case, the `location_name` shows "Bellagio Hotel & Casino" and `brands` shows "MGM Resorts."

If you are having difficulty matching location names or brand names to your own POI listings, we offer a matching service that will provide you with the `placekeys` of of our locations mapped to your existing POI data.

Categorization of POI

SafeGraph Places uses the North American Industry Classification System (NAICS) developed by the US Census Bureau, which consists of a numeric NAICS code up to 6 digits in length. Although this taxonomy was developed in the US, we have found it just as useful for categorizing POIs in other countries as well and will continue to use it until a better alternative presents itself. We currently reference the 2017 version of NAICS. We will provide an update if and when we ultimately update to reflect the 2022 changes.

The NAICS code itself is hierarchical; in other words, the first 2 digits describe a very general category, and additional digits describe more and more specific categories. For example:

- `72` is the general category `Accommodation and Food Services`.
- `722` is the more specific category `Food Services and Drinking Places`.
- `7225` is the even more specific category `Restaurants and Other Eating Places`.
- `722513` is the most specific category `Limited-Service Restaurants` (i.e. quick-serve or fast-food restaurants).

We strive to assign a best fitting `naics_code` for all of our POIs. Our goal is to assign a full six digits for maximum granularity wherever possible, but our category algorithm cannot always infer a high confidence six digit `naics_code` based on POI name and other descriptive metadata. In these cases, we provide a shorter `naics_code` where we do have high confidence in the assignment (i.e. 3, 4 , 5 digits). In these circumstances, we choose to sacrifice the extra digits of precision in exchange for high veracity predictions and also because the extra precision is not always meaningfully different (i.e. some adjacent 6 digit NAICS are extremely similar).

See our Places Summary Statistics for the latest details on counts and coverage.

Also see our use of Category Tags to provide more flexibility and granularity where the NAICS code classification falls short.

Determining when POI Open and Close

`opened_on` and `closed_on` dates are determined from metadata at the source level. If a new POI from an existing source repeatedly appears in our build pipeline, it is flagged as `opened_on` during the month in which it first appears.

Similarly, if a POI from an existing source repeatedly disappears in our build pipeline, it is flagged as `closed_on` during the month in which it first disappears. These flags are added to the Places product permitting final QA checks and overall data hygiene.

Temporary closures are not captured in open/close tracking, and it became difficult to distinguish permanent closures from temporary closures at the onset of COVID-19. This resulted in a relatively low count of POIs with `closed_on` values between "2020-03" and "2020-06" as we erred towards the side of caution to not mistakenly mark temporarily closed businesses as permanently closed.

If a POI has not yet been sourced consistently enough to provide the metadata needed to determine `closed_on` dates, then it will have a null value in the `tracking_closed_since` column. In general, the SafeGraph Places product tracks `opened_on` and `closed_on` dates from as early as 2019-07 onward, and therefore, the majority of POIs that have a `tracking_closed_since` date will show a value of "2019-07."

Please note that `closed_on` values are over-indexed on "2020-01" as January 2020 was the first Places release featuring the open/close columns . At this time, only branded POIs (POIs with a `safegraph_brand_id`) contained enough metadata to determine a true store closure during that month. Non-branded POIs with a "2020-01" `closed_on` value implies that the POI *closed sometime before January 2020,* but we do not have enough metadata history to determine the exact yyyy-mm.

A second spike in `closed_on` values occurred in October of 2021 thanks to new information about more than 130k "longtail" POIs. Like the January 2020 anomaly, the "2021-10" `closed_on` value implies that the POI *closed sometime before October 2021,* but we cannot determine the exact yyyy-mm.

For countries outside of the US, CA and GB, we anticipate non-null values for `tracking_closed_since`, `closed_on`, and `opened_on` beginning in the November 2021 release which will be the first time where we have a long enough track record to support these columns.

All other `closed_on` values are precise within a < 60 day margin of error.

The `opened_on`, `closed_on` and `tracking_closed_since` columns are specific to Core Places. These are not available in stand-alone Geometry or Patterns purchases. If Core is purchased in combination with Geometry and/or Patterns, the Geometry and Patterns specific fields will be null for any POIs with a `closed_on` date. Please reference Column Ordering for details on where these columns exist per product combination.

Column Name Detailed Descriptions

`placekey`

Placekey is a unique and persistent identifier for any physical place in the US that intelligently partitions the ID into meaningful encodings. See the Placekey key concept for a detailed description.

`parent_placekey`

This Placekey column will identify a larger place that may encompass a given POI, which we refer to as the "Parent". Think of an indoor shopping mall as the parent of the individual stores inside. For any place without an assigned polygon, the `parent_placekey` column will be null because we rely on geometric relationships to identify parent/child hierarchy. So for example, any of our Point POI will not have an assigned parent because they do not have defined polygons. You can find out more about our process for defining these relationships in our Spatial Hierarchy section where we also include a list of all the types of places that can serve as "Parents".

`top_category`, `sub_category`, `naics_code`

`top_category` and `sub_category` are the string labels associated with the first 4 digits and 6 digits of `naics_code`, respectively. See Categorization of POI section above.

`latitude`, `longitude`

- In general, latitude and longitude are defined by our best knowledge of the POI location. It is not designed to specifically locate the front door of the business, but rather defines the general center of the business.
- Latitude and longitude still attempt to identify the individual business even if that business and others have the same polygon (e.g. strip mall).

`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.

- If you care about street addresses as much as we do, we also have more specific address columns to split out address components. These are optional and available upon request for future deliveries.
  - `primary_number`
  - `street_predirection`
  - `street_name`
  - `street_postdirection`
  - `street_suffix`

`city`

- In the US, all centroids (latitudes/longitudes) are referenced against a geospatial file of city boundaries as defined by the US Census Bureau ([browse the boundaries here](#)). In edge cases, the preferred city name in the address line reflects a [pre-annexed](#) city name, and we try our best to preserve those city names where possible.
- In Canada, city names are the output of normalized address strings from POI sources.
- In Great Britain, city names are the output of normalized address strings from POI sources, but in edge cases, we allow POIs to have a null city name as long as `region` is populated. The `region` column in Great Britain refers to county boundaries, and counties are a decent alternative to cities for geographic filtering.

`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 county](#).

`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](#).

`phone_number`

This is a 10 digit phone number in the US and Canada or a 12 digit phone number in Great Britain. We filter out toll-free numbers (e.g. 1-800) and strive to have POI-specific numbers (not franchise-level or corporate-level numbers).

`open_hours`

The new format for open hours is a JSON string with days as keys and opening & closing times (**in the POI's local time**) as values.

- Each JSON string is guaranteed to have all 7 days as keys
- We indicate that a POI is closed for the day by giving it a value of "[]"
- We indicate that a POI is open the entire day by using a format like: `
  - "Thu": [["0:00", "24:00"]]`
- For POI that open and close multiple times throughout the day (e.g. a restaurant open in the morning and evening but not midday), we list multiple opening/closing pairs. For example:
  - `"Sat": [["8:00", "13:00"], ["15:00", "22:30"]]`
  - This indicates that a POI is open from 8 am to 1 pm and also from 3 pm to 10:30 pm on Saturday.
- For POI that open and close on different days (e.g. a bar which opens on Tuesday at 6 pm and closes on Wednesday at 2 am), we use a format like:
  - `"Tue": [["18:00", "24:00"]], "Wed": [["0:00", "2:00"]]`

`category_tags`

- `category_tags` provide higher-granularity category information beyond what can be found in a NAICS code label. So instead of just "Full-Service Restaurant" (NAICS 722511), we'll also provide tags like 'Pizza', 'Lunch', 'Dinner', 'Drive Through', and 'Late Night' so that you can glean more meaningful details about that specific restaurant. Trying to find a place that serves coffee? Our 'coffee shop' tag spans more than 30 NAICS categorizations, giving you a much easier method to pinpoint specific places based on granular inputs.
- Category information is conveyed is a list of descriptive words about the POI. e.g `['Mexican Food, Dinner]`
- Category tags are broadly available across NAICS codes though not every NAICS code will have them. [Here is the full list of possible tags for each NAICS code](#). SafeGraph strives to label all relevant tags and will include up to 13 tags for any given POI.

`geometry_type`

This is the geometric shape associated with the POI where possible values are: "POLYGON" or "POINT." This is meant to distinguish traditional SafeGraph places which currently have Geometry or will eventually have Geometry ("POLYGON") from places that intentionally do not have Geometry ("POINT.") Some POIs have `geometry_type` = "POLYGON" but no corresponding WKT because we have not yet built our Geometry product in that region of the world. See our summary stats page for details on Geometry coverage.

`store_id`

Store_id is the unique ID associated with a store as provided and maintained by the store/brand itself. This is a premium column only applicable to places with a `safegraph_brand_id` . Most `store_ids` can be found directly on store locators, but in some cases, the `store_id` is embedded within the store locator URL for the specific store. Note that there is no single source of truth for `store_ids` , and some first party datasets may not define `store_id` in the same manner which SafeGraph does, but we strive to provide the most widely used concept of `store_id` .

Typically, `store_ids` are alphanumeric codes unique to each location. However, they are not always alphanumeric. For example the store number for this store is **1615.**

`store_id` is especially useful as join key when working with transaction data. For example, "TJ256Y8" may be the only location specific information within a transaction dataset. A Places dataset which also contains "TJ256Y8" as a `store_id` enables a join to contextualize transaction data (or other internal, store-level data) with SafeGraph places information.

Column Ordering

Reference this sheet for specific column orders when licensing various product combinations.

Known Data Issues or Artifacts

`closed_on` First Featured

Please note that `closed_on` values are over-indexed on "2020-01" as January 2020 was the first Places release featuring the open/close columns. At this time, only branded POIs (POIs with a `safegraph_brand_id`) contained enough metadata to determine a true store closure during that month. Non-branded POIs with a "2020-01" `closed_on` value implies that the POI *closed sometime before January 2020,* but we do not have enough metadata history to determine the exact yyyy-mm. All other `closed_on` values are precise within a < 60 day margin of error.

`opened_on` for "non-branded" POIs

Please note that `opened_on` dates are only inferred for POIs with a `safegraph_brand_id` . We are working towards sourcing more robust metadata for "non-branded" POIs to close this gap.

`safegraph_place_id` , `parent_safegraph_place_id` , `tracking_opened_since` columns dropped

As of July-2021 Release Notes, `safegraph_place_id` and `parent_safegraph_place_id` were dropped and `placekey` and `parent_placekey` are referenced moving forward. `tracking_opened_since` was dropped, due to the column being redundant. If a POI has an `opened_on` value, it implies we've been tracking it since that date. If a POI does not have an `opened_on` value, it implies we were not able to track the exact date it opened.

Video Overview