Cyber-capital

deriving insight on the use of **space** from the **digital** layer of cities and territories
for **location intelligence**

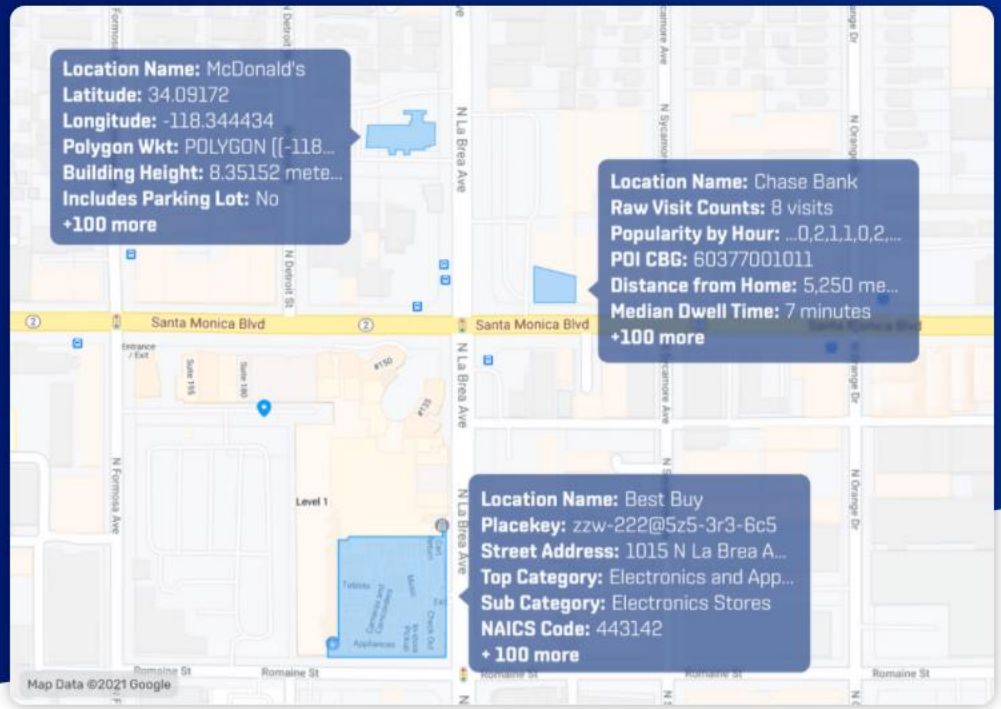
Massimo Izzo
massimo.izzo@polimi.it

SAFEGRAPH Product Docs https://www.safegraph.com/ PREVIEW DATA CONTACT SALES

The Source of Truth for Places Data

Organizations trust SafeGraph data to drive their business forward. Access the most accurate point of interest (POI) and foot traffic data on the market.

Preview Data Contact Sales



INDUSTRY LEADERS RUN ON SAFEGRAPH

verizon media esri Sysco mapbox CHOICE HOTELS Tripadvisor ARES

SAFEGRAPH https://www.safegraph.com/

Core Places

Over 6,400,000 point-of-interest locations where consumers spend time or money.

View the data schema.

Geometry

Building footprints with spatial hierarchy for all POIs in the Core dataset.

View the data schema.

Patterns

Foot-traffic insights for places derived from anonymized mobile devices.

View the data schema.

SafeGraph Neighborhood Patterns

the dataset contains user-generated data from mobile devices (Android & iOS systems)

representing **pedestrian** occurrences by **US Census Block Group (CBG)** (~600 to 3000 residents per CBG; source: US Census Bureau)

selected area

State of New York (15461 CBG's)

time span of the available free sample

June 1st, 2020 – July 1st, 2020

Schema

15461 rows, 35 columns

note: State of NY only (whole dataset has 220684 rows, 35 columns)

CBG id's & overall counts

- `area`: CBG id, it can be joined with **census data** to match the population structure
- `raw_device_counts`: # unique users or visitors in the CBG (unique mobile devices: Android & iOS)
- `raw_stop_counts`: # stops (stays lasting for > 1 minute)
- `stops_by_day`; `stops_by_each_hour`: # stops by day (31 ordered counts); # stops by hour (744 ordered counts)

visitors by origin CBG

- `device_home_areas`: # visitors by home origin CBG (6 pm – 7 am)
- `weekday_device_home_areas`; `weekend_device_home_areas`; # visitors by home origin CBG (weekdays & weekends)
- `breakfast` / `lunch` / `afternoon_tea` / `dinner` / `nightlife` / `work_hours_device_home_areas`; # visitors by home origin CBG (times of the day)
- `device_daytime_areas`; # visitors by primary daytime origin CBG (9 am – 5 pm)

distances
proximity / space impedance

- `distance_from_home`: median distance from home origin CBG's (meters)
- `distance_from_primary_daytime_location`; median distance from primary daytime origin CBG's (meters)

visits to Points of Interest (PoI's)

- `top_same_day_brand`: top 20 PoI's by % attracted visitors on the total visitors to the CBG (one day peak)
- `top_same_month_brand`; top 20 PoI's by % attracted visitors on the total visitors to the CBG (whole month)

popularity

- `popularity_by_hour_monday/tuesday/wednesday/thursday/friday/saturday/sunday`: # stops by hour on any Monday to Sunday in the month

duration of stay

- `median_dwell`: median duration of stays within the CBG (minutes)

Schema

detail

comma-separated, row wise dictionary

24 hours * 31 days = 744 stop counts

	area	raw_stop_counts	raw_device_counts	stops_by_day	stops_by_each_hour	device_home_areas		weekday_device_home_areas	weekend_device_home_areas
1	361031592032	8389	987	[275,277,252,273,253,262,25...	[12,0,0,0,1,1,1,8,6,4,15,9,12,...	{"361031592032":151,"361031592031":42,...		{"361031592032":145,"361031593002":30...	{"361031592032":136,"361031592031":21...
2	360290052013	4446	1107	[151,189,225,192,167,105,71,...	[3,0,1,1,1,4,10,12,11,11,10,8,...	{"360290052013":33,"360290047003":11,"...		{"360290052013":34,"360290047003":10,"...	{"360290052013":25,"360290138003":10,...
3	360470782001	2965	592	[84,92,84,102,83,88,91,95,96,...	[4,1,0,0,1,0,2,4,2,1,0,3,6,4,2,...	{"360470782001":73,"360470782002":19,"...		{"360470782001":69,"360470782002":14,"...	{"360470782001":64,"360470782002":12,...
4	360470253001	1673	447	[58,39,47,55,52,57,63,51,59,6...	[3,0,0,0,0,0,0,3,2,3,1,1,2,9,2,...	{"360470253001":39,"360470255002":30,"...		{"360470253001":36,"360470255002":23,"...	{"360470253001":31,"360470255002":14,...
5	361119545002	3159	533	[97,87,95,109,116,91,121,93,...	[4,0,1,0,1,2,2,6,6,4,2,5,4,5,9,...	{"361119545002":36,"361119545003":22,"...		{"361119545002":38,"361119548001":17,"...	{"361119545002":35,"361119548001":13,...
6	361190034002	1551	265	[39,46,48,52,39,67,52,55,47,4...	[3,1,0,0,1,0,1,0,3,1,1,2,1,3,1,...	{"361190034002":45,"361190034004":8,"3...		{"361190034002":43,"361190034004":7,"3...	{"361190034002":36,"361190034004":5,"...
7	360595197033	6451	601	[207,217,203,211,203,213,23...	[2,1,2,0,0,1,0,2,6,4,9,1,4,8,7,...	{"360595197033":121,"360595197032":17,...		{"360595197033":124,"360595197032":12...	{"360595197033":115,"360595197032":10...
8	360811193002	2037	423	[80,83,66,66,66,70,70,69,54,7...	[5,1,0,2,0,1,0,0,0,3,2,0,5,5,7,...	{"360811193002":64,"360811191001":18,"...		{"360811193002":67,"360811191001":14,"...	{"360811193002":52,"360811191001":12,...
9	360594093001	3968	430	[107,124,119,122,117,111,14...	[10,1,0,1,0,0,0,0,1,1,2,4,4,5,2...	{"360594093001":78,"360594093002":14,"...		{"360594093001":72,"360594093002":9,"3...	{"360594093001":70,"360594092002":8,"...
10	360811010021	3294	589	[114,103,89,114,80,54,101,11...	[6,3,1,0,1,1,5,4,6,3,3,4,4,3,4,...	{"360811010021":82,"360811010012":22,"...		{"360811010021":75,"360811010012":22,"...	{"360811010021":72,"360811010012":8,"...

rows = unique CBG's

1st CBG = current row GBG

2nd to last CBG = CBG's of other rows / extra NY areas

Possible questions

descriptive

- which day of the week is a CBG busiest?
- when during the day is a CBG busy?
- where do the devices that stop in a CBG during breakfast, lunch or dinner time travel from?
- how do the weekday demographics of a CBG compare to the weekend demographics?

predittori: **predictive**

support to location intelligence

predict the **time of stay** (`median_dwell`) for a suitable allocation of services / businesses to catch pedestrian flows, based on:

- # stops (> 1 minute) (`raw_stop_counts`), # unique visitors (`raw_device_counts`)
- average travel distance from home / from the main daytime location (`distance_from_home` / `primary_daytime_location`)
- diversity of origin areas during the weekdays / weekends (`weekday_` / `weekend_` / `device_home_areas`) metrica di entropia per origini diversi
- attraction of visitors from the same area / other areas during the day to night time (`breakfast` / `lunch` / `afternoon_tea` / `dinner` / `nightlife` / `work_hours_device_home_areas` / `device_daytime_areas`)
- locally recurring top PoI's by attraction (`top_same_day_brand` / `top_same_day_brand`); average hourly variations of stops (`stops_by_day` / `stops_by_each_hour`) punti di interesse ricorrenti (`popularity_by_hour_monday/tuesday/wednesday/thursday/friday/saturday/sunday`)

Thank you

Massimo Izzo
massimo.izzo@polimi.it