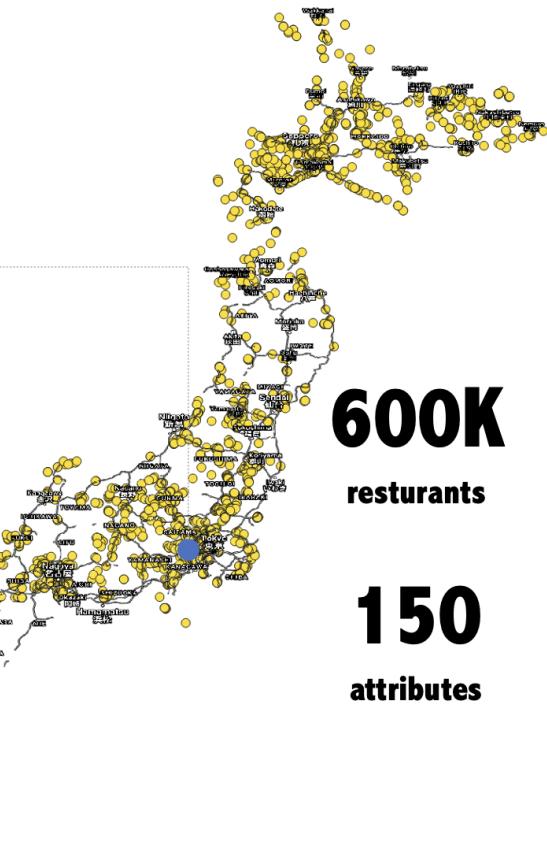


Restaurants data for Japan

Summary

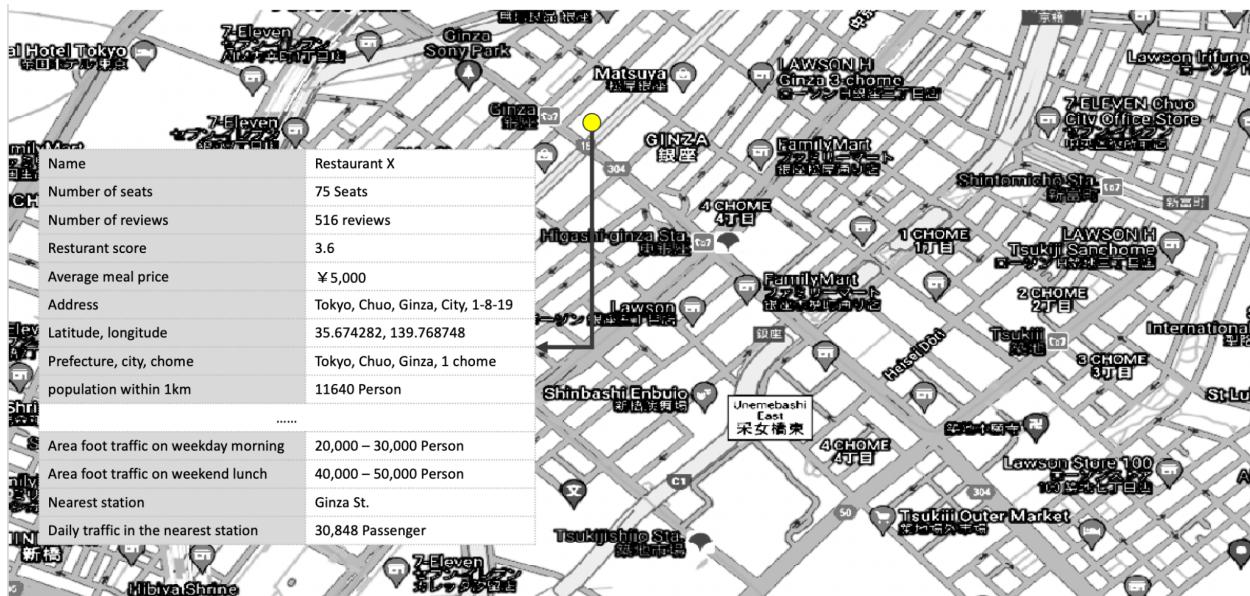
This is the most enriched Japan's restaurants dataset on the internet with more than 600,000 restaurants and over 150 attributes per restaurant. The dataset has been curated from multiple public domain and commercial sources and pre-processed and merged to form a solid information database as scale. (Sample data live view: [Link](#) – Password: xmap)

Name	MERCER BRUNCH Ginza
Number of seats	75 Seats
Number of reviews	516 reviews
Restaurant score	3.6
Average meal price	¥ 5,000
Address	Tokyo, Chuo, Ginza, City, 1-8-19
Latitude, longitude	35.674282, 139.768748
Prefecture, city, chome	Tokyo, Chuo, Ginza, 1 chome
population within 1km	11640 Person
<hr/>	
Area foot traffic on weekday morning	20,000 – 30,000 Person
Area foot traffic on weekend lunch	40,000 – 50,000 Person
Nearest station	Ginza ichome 
Daily traffic in the nearest station	30,848 Passenger



Data specs

Dataset summary	Restaurant information dataset
Country	Japan
Perfecures covered	47
Total records	600,000
Total Attributes per restaurant	Over 170 attributes
Update date	April, 2022



Sample data (50 restaurant per prefecture)

Number of records	1,732
Number of attributes per record	185
URL for download	Link – please request access
Live view of sample data	Link – Password: xmap
License for sample data	<ul style="list-style-type: none"> - You are allowed to use the data internally within the organization to test its usability - You are not allowed to publish the sample data online - You are not allowed to transfer the sample data to other companies without email written confirmation from the data owner using

Data collection methodology

Data has been curated from public available resources and merged together to enrich attributes.

Restaurant data has been collected from services such as tabelog, gurunavi, savorjapan, and others.

Traffic data is collected and stored and analyzed from various publicly available sources and services. Then pre-processed to produce average traffic in an area and mapped to each restaurant.

Population data has been collected from public government data and other data derived sources such as ciesen, Gridded Population of the World, High Resolution Population Density Maps by Facebook, WorldPop and others.

Points of interest data are collected from the Japanese government open data portals, the data has been pre-processed and converted into a format usable in processing and aggregations

Road network data, for specific applications related to area mapping on road network and density of road network estimation, road network data from open street maps has been pre-processed and converted into a format usable by routing engines.

Data preprocessing / processing methods

Format change

Data format has been changed and altered from different formats such as xml, json, TIFF and other shapefiles format to structured column format to be mapped and joined accurately.

Character encoding fixes

Different sources of data comes in different encoding, each dataset has been pre-processed to unify the language encoding

Data cleaning

Some datasets contain multiple missing values, missing values are either removed or imputed depending on the case. Some random text values or erroneous text are pre-processed to be normalized.

Data join / merge

Data from multiple datasets are joined using unique key join, spatial join and other join methods to add different attributes to the each data point to maximize its value

Zonal statistics

Zonal statistics are applied on raster data to extract the number of people living on a specific circular area either as a total population or a demographics split to represent the demographics structure of the area of people living in that area.

Geospatial encoding and decoding

Many datasets come in different encoding formats such as Japan grid system, geohashes, H3 encoding or no encoding latitude longitude. The data from different sources are mapped into a unified encoding system and decoded back to coordinates to be matched together accurately

Data columns sample

column name	sample	Column Explanation
restaurant_name	ビア カフェ バテレ	Restaurant name
restaurant_category	ビアバー、カフェ・喫茶(その他)	Category of the restaurant
review_count	65	Number of people that wrote a review about the restaurant on Tabelog
review_score	3.5	Restaurant star score out of 5
lunch_review_score	0	The score of the restaurant for lunch, if null or zero, there is no score for lunch
dinner_review_score	0	The score of the restaurant for dinner, if null or zero, there is no score for dinner
holiday	祝日関係なく土日のみの営業	The day of the holiday where the restaurant is closed
price_range_dinner	¥1,000～¥1,999	Range of price for dinner time
price_range_lunch	¥1,000～¥1,999	Range of price for lunch time
station_name	奥多摩	This represent the name of the nearest station name
image_url	https://tblq.k-img.com/restaurant/images/Rvw/46295/100x100_square_46295145.jpg	The image url of the restaurant
family_flag	0	This flag is binary and describe family / not family friendly
friends_flag	0	This flag is binary and describe if the restaurant is good for friends hangouts
alone_flag	1	This flag is binary and describe if the restaurant is good for going alone
date_flag	0	This flag is binary and describe if the restaurant is good for going for a date
restaurant_status	0	
party_flag	0	This flag is binary and describe if the restaurant can be used for parties
vac	0	
settai_flag	0	
rst_name	ビア カフェ バテレ(Beer Cafe VERTERE)	Restaurant name collected from another source and usually includes english name of the restaurant
categoy	ビアバー、カフェ・喫茶(その他)	Restaurant category collected from another source,

		usually the same category text as field 2
reservation_s		
reservation_availability	予約不可	This column explain if you can make a reservation
address	東京都西多摩郡奥多摩町氷川212 大きな地図を見る 周辺のお店を探す	The address of the restaurant detailed
transportation	JR青梅線 奥多摩駅を出て目の前に見える柳小路の右手側二件目。 奥多摩駅から80m	Access method to the restaurant
opening_hours	営業時間 [土・日]12:00~19:30 (フード L.O18:30 お酒L.O 19:00)※コロナウイルス感染症対策として営業時間短縮しております。 ※祝日関係なく土日のみの営業 日曜営業 定休日 祝日関係なく土日のみの営業	The hours where this restaurant is opened
budget	[夜]¥1,000～¥1,999 [昼]¥1,000～¥1,999	Budget number in text by the restaurant, this number is the same as price_range_lunch and price_range_dinner.
budget_based_on_review	[夜]¥2,000～¥2,999 [昼]¥1,000～¥1,999 予算分布を見る	Budget number in text, by the reviewers or visitors of the restaurant
payment_method	カード不可 電子マネー不可	Payment methods that are allowed
service_charge	チャージ料なし	If there are any service charge included
number_of_seats	30席 (テラス席あり。)	The total number of seats in the restaurant
private_rooms	無	Does this room contain private room or not
private_parties	不可	If this restaurant can be used for private parties
smoking	全席禁煙 店外に喫煙スペースあり	Smoking rules written in text format
parking	無 店舗の近くにコインパーキングあり。	If the restaurant offer parking service

space_facilities	オシャレな空間、落ち着いた空間、カウンター席あり、オーブンテラスあり	The design of the space and the associated facilities
mobile_phones	docomo、au、SoftBank、Y!mobile	
go_to_eat		If this restaurant support go to eat campaigns vouchers
images		Restaurant images links collected from reviews
usage	一人で入りやすい こんな時によく使われます。	The usage of the restaurant of advice on who can use the restaurant
location	景色がきれい、隠れ家レストラン、一軒家レストラン	Description of the location of the restaurant
website	http://verterebrew.com/	The URL of the website of the restaurant
open_date	2015年7月6日	The opening date of the restaurant
first_contributor	とみイ(2319)	The name of the person contributed most on tabelog and his contribution overall the platform
maximum_number_of_bookings		Maximum number of bookings
bringing_children		Flag indicating if you can bring children
remarks		
contact_us	0428-85-8590	The contact phone number of the restaurant
dress_code		The dress code if it exists
lat	35.80879524	Latitude using WGS 84 system
lng	139.0963081	Longitude using WGS 84 system
admin_area_code	13308002001	The administrative boundary area code based on the Japanese government administrative boundary definition
prefecture	13	The number of the prefecture from 1 - 47
city	308	The code of the city
chome	2001	The code of the neighborhood
prefecture_name	東京都	The name of the prefecture
city_name	奥多摩町	The name of the city
chome_name	氷川	The name of the neighborhood
geohash_level_11	xn6gtnyc6gh	The geohash code level 11
h3_lvl_15	8f2f5b460535482	The h3 uber indexing level 15
footfall_holiday_breakfast	1	Footfall data scraped from online resources between 1 and 10 where that number represents the number of people in a 500 meter mesh around the restaurant. Exact table is below. Please refer to mobility data explanation for exact details
footfall_holiday_dinner	1	

r		
footfall_holiday_lunch	1	
footfall_holiday_night	1	
footfall_weekend_breakfast	1	
footfall_weekend_dinner	1	
footfall_weekend_lunch	1	
footfall_weekend_night	1	
pop_1km	1709	The number of people living in a radius of 1 km around the restaurant
pop_5km	2539	The number of people living in a radius of 5 km around the restaurant
pop_500m	966	The number of people living in a radius of 500 meters around the restaurant
500m_jpn_f_0_2020	0	The number of Females aged 0 years old living within 500 meter radius from the restaurant location
500m_jpn_f_25_2020	6	The number of Females aged 20-25 years old living within 500 meter radius from the restaurant location
500m_jpn_f_30_2020	6	The number of Females aged 25-30 years old living within 500 meter radius from the restaurant location
500m_jpn_f_50_2020	6	~
500m_jpn_f_60_2020	5	~
500m_jpn_f_45_2020	8	~
500m_jpn_f_5_2020	3	~
500m_jpn_f_10_2020	3	~
500m_jpn_f_15_2020	4	~
500m_jpn_f_1_2020	2	~
500m_jpn_f_35_2020	7	~
500m_jpn_f_55_2020	5	~
500m_jpn_	5	~

f_75_2020		
500m_jpn_m_10_2020	3	The number of Male aged 5-10 years old living within 500 meter radius from the restaurant location
500m_jpn_m_15_2020	4	~
500m_jpn_m_1_2020	3	~
500m_jpn_m_20_2020	6	~
500m_jpn_m_0_2020	0	~
500m_jpn_m_65_2020	5	~
500m_jpn_m_60_2020	5	~
500m_jpn_m_45_2020	9	~
500m_jpn_m_25_2020	6	~
500m_jpn_m_55_2020	5	~
500m_jpn_m_5_2020	3	~
500m_jpn_m_70_2020	6	~
500m_jpn_m_40_2020	8	~
500m_jpn_m_35_2020	7	~
500m_jpn_m_50_2020	6	~
500m_jpn_m_30_2020	6	~
500m_jpn_f_65_2020	6	~
500m_jpn_f_70_2020	7	~
500m_jpn_f_40_2020	7	~
500m_jpn_f_20_2020	5	~
500m_jpn_m_75_2020	4	~
500m_jpn_m_80_2020	5	~
500m_jpn_f_80_2020	9	~

1km_jpn_f_55_2020	15	The number of Female aged 50-55 years old living within 1 kilometer radius from the restaurant location
1km_jpn_f_65_2020	19	~
1km_jpn_m_45_2020	27	~
1km_jpn_m_20_2020	17	~
1km_jpn_f_15_2020	11	~
1km_jpn_m_15_2020	12	~
1km_jpn_m_10_2020	11	~
1km_jpn_f_20_2020	16	~
1km_jpn_m_30_2020	20	~
1km_jpn_f_60_2020	15	~
1km_jpn_f_0_2020	2	~
1km_jpn_f_30_2020	19	~
1km_jpn_m_80_2020	14	~
1km_jpn_f_5_2020	10	~
1km_jpn_f_50_2020	18	~
1km_jpn_f_45_2020	25	~
1km_jpn_m_75_2020	13	~
1km_jpn_f_25_2020	18	~
1km_jpn_f_70_2020	21	~
1km_jpn_m_35_2020	21	~
1km_jpn_f_80_2020	28	~
1km_jpn_m_50_2020	20	~
1km_jpn_f_35_2020	20	~
1km_jpn_	18	~

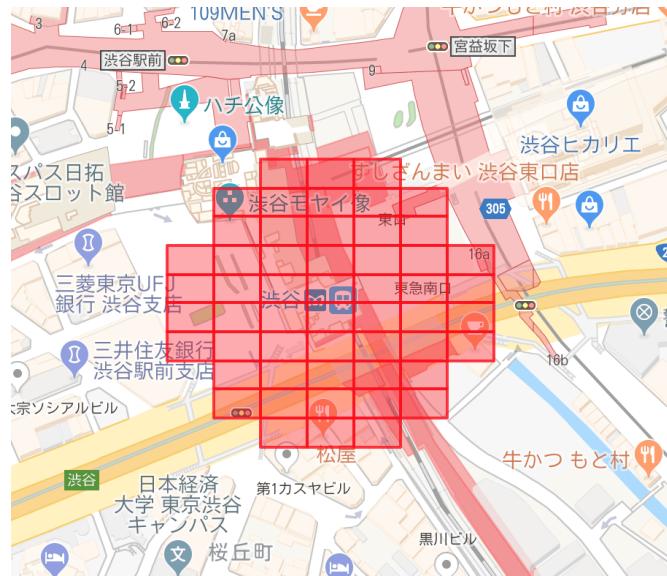
m_70_2020		
1km_jpn_f_10_2020	10	~
1km_jpn_f_1_2020	8	~
1km_jpn_m_25_2020	19	~
1km_jpn_m_40_2020	24	~
1km_jpn_m_0_2020	2	~
1km_jpn_m_1_2020	9	~
1km_jpn_f_40_2020	23	~
1km_jpn_f_75_2020	16	~
1km_jpn_m_65_2020	17	~
1km_jpn_m_55_2020	16	~
1km_jpn_m_5_2020	11	~
1km_jpn_m_60_2020	15	~
5km_jpn_f_55_2020	78	~
5km_jpn_f_65_2020	95	~
5km_jpn_m_45_2020	135	~
5km_jpn_m_20_2020	87	~
5km_jpn_f_15_2020	58	~
5km_jpn_m_15_2020	60	~
5km_jpn_m_10_2020	54	~
5km_jpn_f_20_2020	83	~
5km_jpn_m_30_2020	100	~
5km_jpn_f_60_2020	78	~
5km_jpn_f_0_2020	11	~

5km_jpn_f_30_2020	95	~
5km_jpn_m_80_2020	74	~
5km_jpn_f_5_2020	52	~
5km_jpn_f_50_2020	94	~
5km_jpn_f_45_2020	126	~
5km_jpn_m_75_2020	65	~
5km_jpn_f_25_2020	90	~
5km_jpn_f_70_2020	107	~
5km_jpn_m_35_2020	107	~
5km_jpn_f_80_2020	140	~
5km_jpn_m_50_2020	99	~
5km_jpn_f_35_2020	102	~
5km_jpn_m_70_2020	90	~
5km_jpn_f_10_2020	52	~
5km_jpn_f_1_2020	43	~
5km_jpn_m_25_2020	95	~
5km_jpn_m_40_2020	121	~
5km_jpn_m_0_2020	11	~
5km_jpn_m_1_2020	45	~
5km_jpn_f_40_2020	114	~
5km_jpn_f_75_2020	84	~
5km_jpn_m_65_2020	86	~
5km_jpn_m_55_2020	81	~
5km_jpn_	55	~

m_5_2020		
5km_jpn_m_60_2020	77	~
gid	533950772	This grid id represent where the restaurant is located in the standard gridding system in Japan
update_date	2022-05-01	The data when data is collected and updated

Geohash explanation

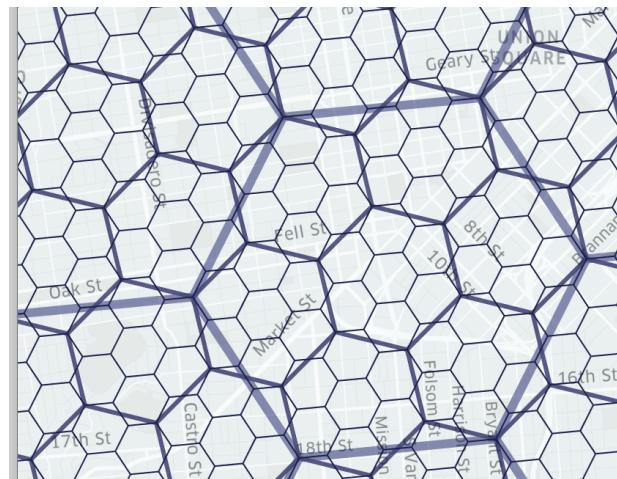
Geohashing is a geocoding method used to encode geographic coordinates (latitude and longitude) into a short string of digits and letters delineating an area on a map, which is called a cell, with varying resolutions. The more characters in the string, the more precise the location.



GeoHash length	Grid Area width x height
1	5,009.4km x 4,992.6km
2	1,252.3km x 624.1km
3	156.5km x 156km
4	39.1km x 19.5km
5	4.9km x 4.9km
6	1.2km x 609.4m
7	152.9m x 152.4m
8	38.2m x 19m
9	4.8m x 4.8m
10	1.2m x 59.5m
11	14.9cm x 14.9cm
12	3.7cm x 1.9cm

Uber H3 explanation

H3 is a geospatial indexing system using a hexagonal grid that can be (approximately) subdivided into finer and finer hexagonal grids, combining the benefits of a hexagonal grid with S2's hierarchical subdivisions.



H3 Resolution	Average Hexagon Area (km2)	Average Hexagon Edge Length (km)
0	4,250,546.8477000	1,107.712591000
1	607,220.9782429	418.676005500
2	86,745.8540347	158.244655800
3	12,392.2648621	59.810857940
4	1,770.3235517	22.606379400
5	252.9033645	8.544408276
6	36.1290521	3.229482772
7	5.1612932	1.220629759
8	0.7373276	0.461354684
9	0.1053325	0.174375668
10	0.0150475	0.065907807
11	0.0021496	0.024910561
12	0.0003071	0.009415526
13	0.0000439	0.003559893
14	0.0000063	0.001348575

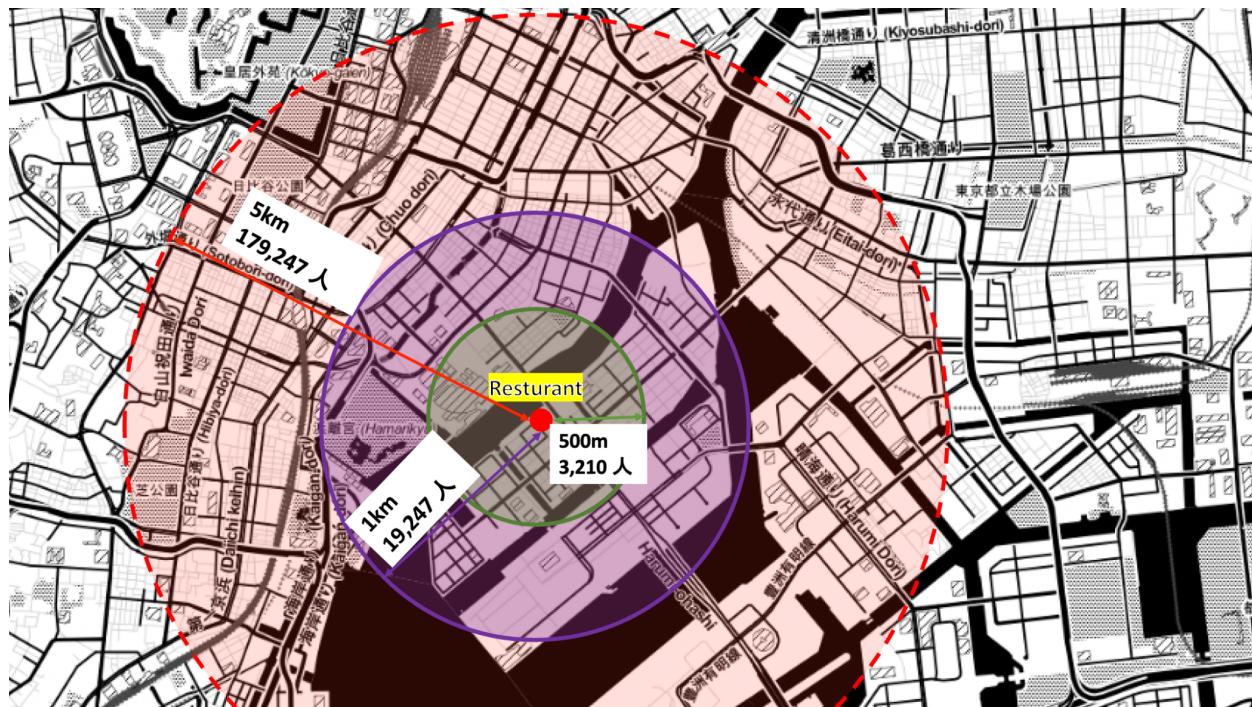
15	0.0000009	0.000509713
----	-----------	-------------

Population data explanation

The number of people living around any facility contributes to understanding the facility performance. Population data are calculated from one of the most recent population statistics public domain datasets. The population datasets are then disaggregated to a high resolution mesh by a third party institution. Using disaggregated high resolution dataset we estimate population data around each restaurant on the following categories

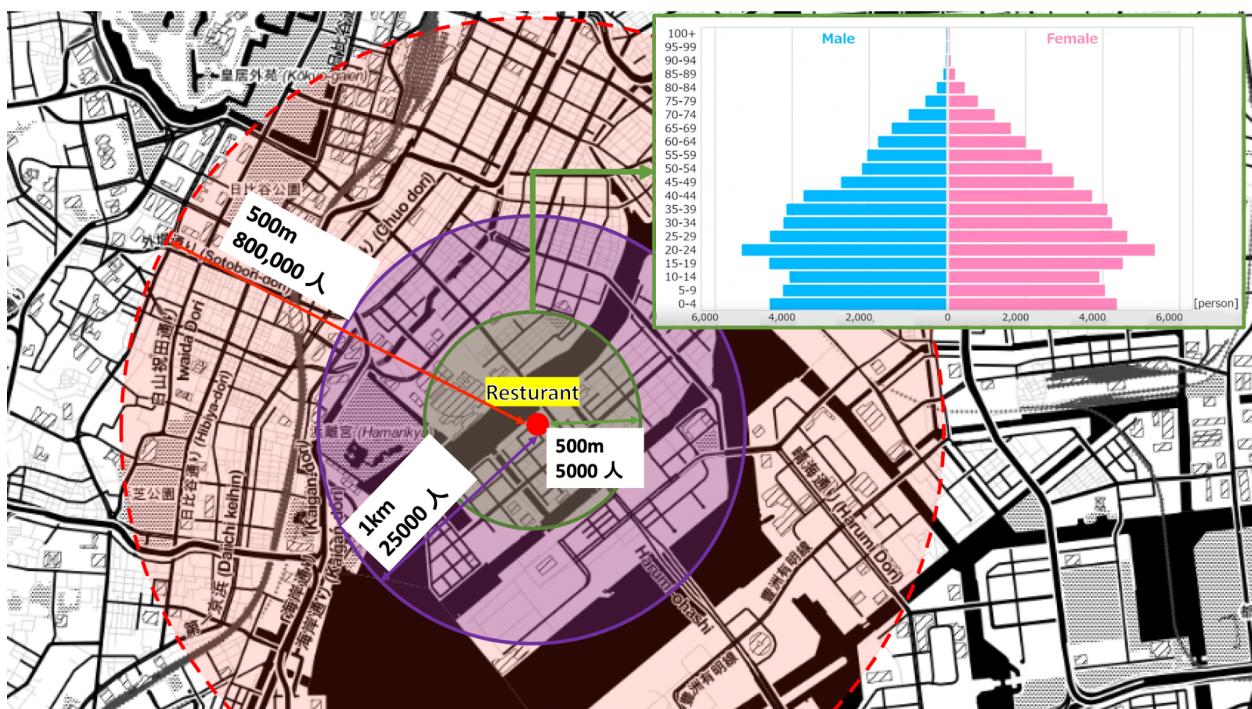
Restaurant data example

Attribute name	Sample data	Explanation
500m radius population	1500 Person	Total number of people living in the area of 500 meters around the restaurant.
1km radius population	5200 person	Total number of people living in the area of 1 kilometer meters around the restaurant.
5km radius population	8731 person	Total number of people living in the area of 5 kilometers around the restaurant.



Demographics data explanation

The number of people living around any facility disaggregated by gender and age contributes to understanding the facility performance. Demographics data are calculated from one of the most recent demographics settlement statistics public domain datasets. The demographics datasets are then disaggregated to a high resolution mesh by a third party institution. Using a disaggregated high resolution dataset we estimate demographics distribution around each restaurant.



Example demographic data explanation

Attribute name	Sample data
500m radius male age 0	80
500m radius male age 0-5	122
500m radius male age 5-10	110
500m radius male age 10-15	102
500m radius male age 15-20	150
500m radius male age 20-25	130
500m radius male age 25-30	122

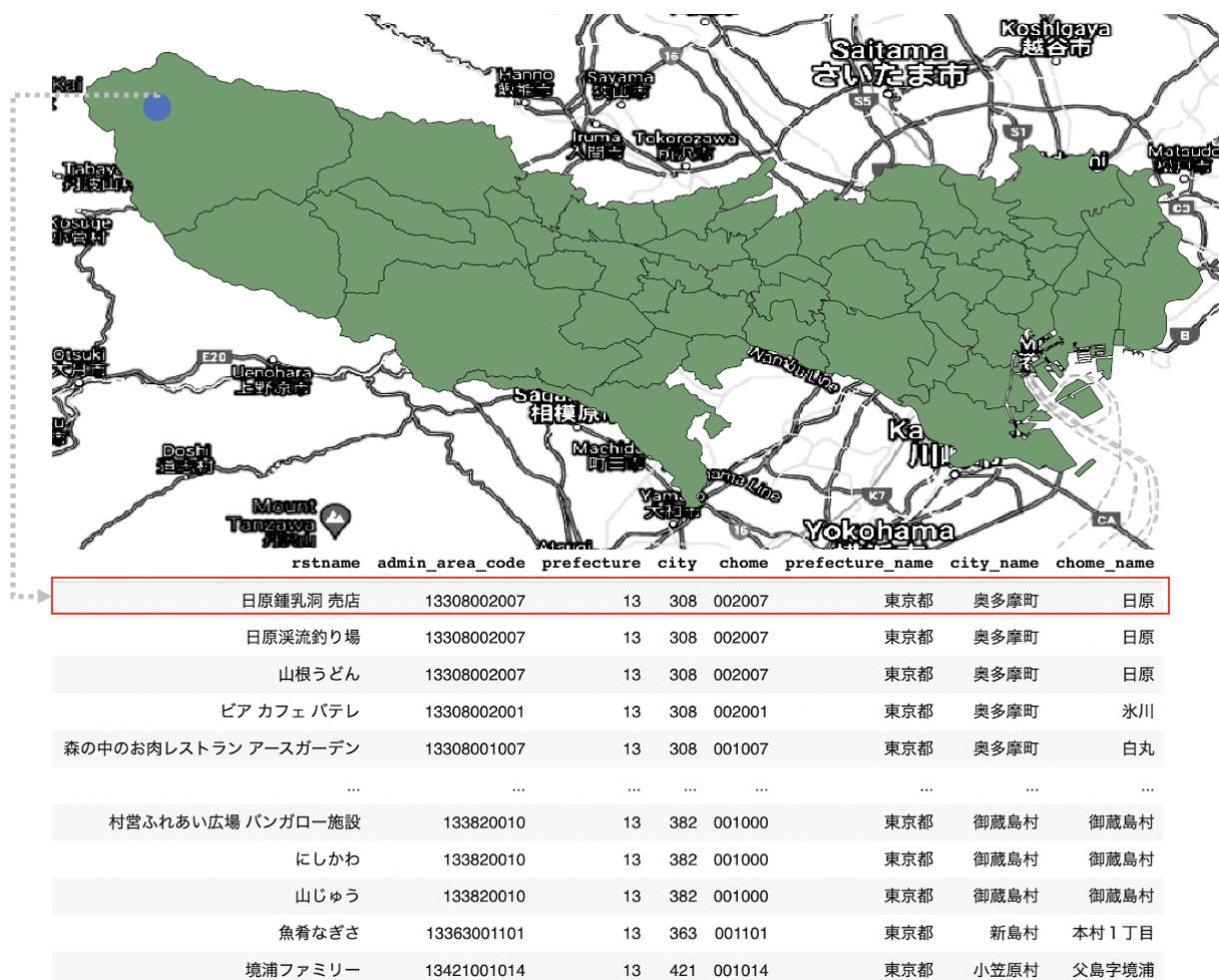
500m radius male age 30-35	150
500m radius male age 35-40	125
500m radius male age 40-45	122
500m radius male age 45-50	120
500m radius male age 55-60	110
500m radius male age >60	115
500m radius female age 0-5	117
500m radius female age 0-5	112
500m radius female age 5-10	110
500m radius female age 10-15	100
500m radius female age 15-20	200
500m radius female age 20-25	110
500m radius female age 25-30	102
500m radius female age 30-35	95
500m radius female age 35-40	90
500m radius female age 40-45	113
500m radius female age 45-50	120
500m radius female age 55-60	122
500m radius female age >60	124
1km radius total female	320
1km radius male age 0-5	488
1km radius male age 0-5	440
1km radius male age 5-10	408
1km radius male age 10-15	600
1km radius male age 15-20	520
1km radius male age 20-25	488
1km radius male age 25-30	600
1km radius male age 30-35	500
1km radius male age 35-40	488
1km radius male age 40-45	480
1km radius male age 45-50	440
1km radius male age 55-60	460
1km radius male age >60	468
1km radius female age 0-5	448
1km radius female age 0-5	440
1km radius female age 5-10	400

1km radius female age 10-15	800
1km radius female age 15-20	440
1km radius female age 20-25	408
1km radius female age 25-30	380
1km radius female age 30-35	360
1km radius female age 35-40	452
1km radius female age 40-45	480
1km radius female age 45-50	488
1km radius female age 55-60	496
1km radius female age >60	320
5km radius total female	4160
5km radius male age 0	6344
5km radius male age 0-5	5720
5km radius male age 5-10	5304
5km radius male age 10-15	7800
5km radius male age 15-20	6760
5km radius male age 20-25	6344
5km radius male age 25-30	7800
5km radius male age 30-35	6500
5km radius male age 35-40	6344
5km radius male age 40-45	6240
5km radius male age 45-50	5720
5km radius male age 55-60	5980
5km radius male age >60	6084
5km radius female age 0	5824
5km radius female age 0-5	5720
5km radius female age 5-10	5200
5km radius female age 10-15	10400
5km radius female age 15-20	5720
5km radius female age 20-25	5304
5km radius female age 25-30	4940
5km radius female age 30-35	4680
5km radius female age 35-40	5876
5km radius female age 40-45	6240

5km radius female age 45-50	6344
5km radius female age 55-60	6448
5km radius female age >60	4160

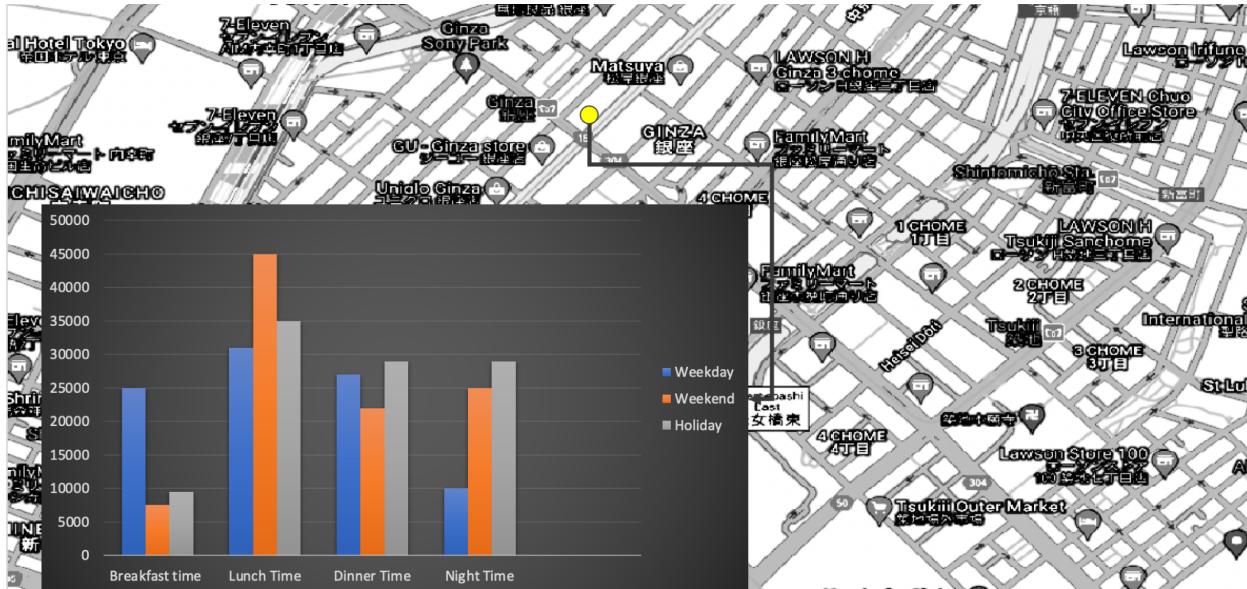
Administrative data explanations

Administrative boundary level 3 data published by the Japanese government on official channels are collected, curated and pre-processed in a format usable along other datasets. Data format is changed from shapefile into tabular format and merged with existing restaurant data to extract administrative boundary information accurately for every restaurant. Using this method, we can associate every restaurant with the finest level of administrative boundary data in both English and Japanese.



Mobility data explanation

Mobility data explain the average footfall traffic in the area where the restaurant is located. The data is collected from multiple sources and preprocessed in a way to represent actual people flow in an area of approximately 500m x 500m where the restaurant is located. We have the following category of time periods:



Attribute name	Sample data	Explanation
Weekday breakfast time	25000	Foot fall on weekday On breakfast time (06:00 am - 10:00 am)
Weekday lunch time	31000	Foot fall on weekday On lunch time (11:00 am - 3:00 pm)
Weekday dinner time	27000	Foot fall on weekday On dinner time (05:00 pm - 10:00 pm)
Weekday night time	10000	Foot fall on weekday On night time (11:00 pm - 05:00 am)
Weekend breakfast time	7500	Foot fall on weekend On breakfast time (06:00 am - 10:00 am)
Weekend lunch time	45000	Foot fall on weekend On lunch time (11:00 am - 3:00 pm)
Weekend dinner time	22000	Foot fall on weekend On dinner time (05:00 pm - 10:00 pm)
Weekend night time	25000	Foot fall on weekend On night time (11:00 pm - 05:00 am)
Holiday breakfast time	9500	Foot fall on a holiday On breakfast time (06:00 am - 10:00 am)
Holiday lunch time	35000	Foot fall on a holiday On lunch time (11:00 am - 3:00 pm)

Holiday dinner time	29000	Foot fall on a holiday On dinner time (05:00 pm - 10:00 pm)
Holiday night time	29000	Foot fall on a holiday On night time (11:00 pm - 05:00 am)

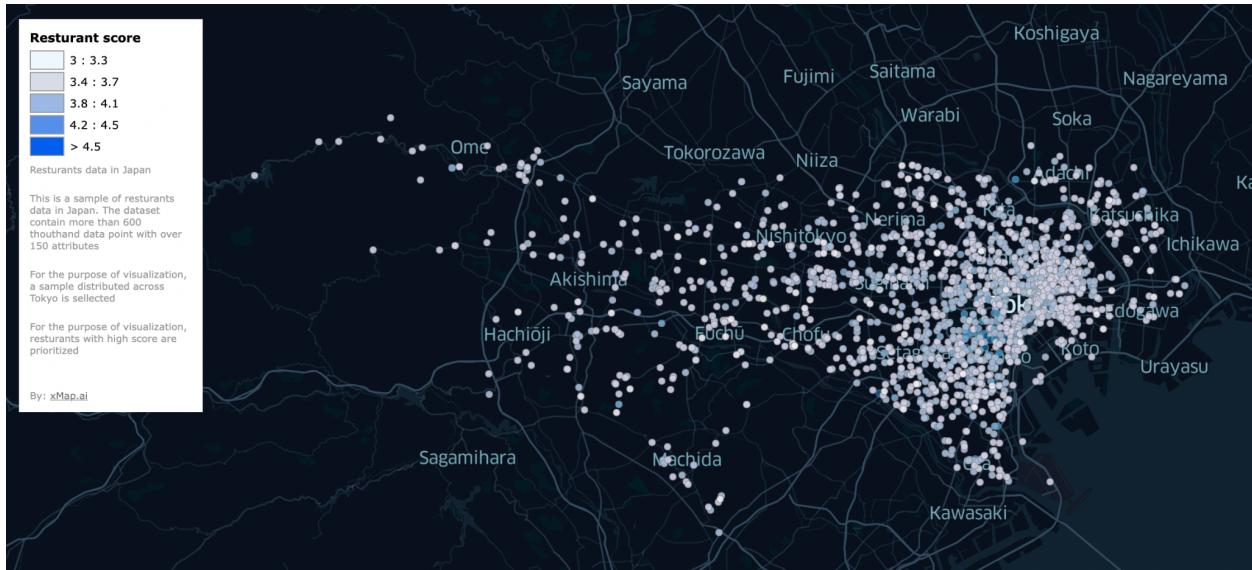
Code explanation

Code explanation table	Number of people
10	51,200~
9	25,600~
8	12,800~
7	6,400~
6	3,200~
5	1,600~
4	800~
3	400~
2	200~
1	100~

Sample data live visualization

Sample data is collected across Tokyo prefecture in a distributed manner. The data is visualized on a map to be visible and accessible by interested buyers. Please hover over each element to see sample attributes.

(Sample data live view: [Link](#) – Password: xmap)



Data update method

Restaurant data attributes change such as, review, review count, pictures, population flow around the restaurant, other restaurant attributes.

New restaurants are added every three months

Category	Sample columns	Update frequency
Restaurant basic data	restaurant_name restaurant_category Restaurants address Restaurant latitude Restaurant longitude	Annual
Restaurant online information	review_count review_score lunch_review_score dinner_review_score price_range_dinner price_range_lunch	1 months
Static population data	pop_1km pop_5km pop_500m	Annual
Static demographics data	500m_jpn_f_0_2020 500m_jpn_f_25_2020 500m_jpn_f_30_2020	Annual
People flow data around the restaurant	Weekday breakfast time Weekday lunch time Weekday dinner time Weekday night time Weekend breakfast time Weekend lunch time	Every 15 days.

Data delivery method

Data will be delivered through file sharing systems such as Amazon S3 bucket, Google GCS or Google Drive folder.