

New York City Food Venues Analysis

Peiwen Liu

June 29, 2020

Data Description

- The list of all community boards in New York City with relevant information is fetched from the main table of [Neighborhoods in New York City](#) Wikipedia page.^[4] The main libraries I used to build the function for fetching the table are **urllib.request** and **BeautifulSoup**.

The database has main components of *Community Boards*, *Area*, *Population* and *Population per Area* data in New York City. To combine it with the following geometric database, an extra column of *boro_cd* is added as the codes of community boards, which serves as the primary key for the two databases.

	Community Board	Area km ²	Pop.	Pop. /km ²	Neighborhood	boro_cd
0	Bronx CB 1	7.17	91,497	12,761	Melrose, Mott Haven, Port Morris	201
1	Bronx CB 2	5.54	52,246	9,792	Hunts Point, Longwood	202
2	Bronx CB 3	4.07	79,762	19,598	Claremont, Concourse Village, Crotona Park, Mo...	203
3	Bronx CB 4	5.28	146,441	27,735	Concourse, Highbridge	204
4	Bronx CB 5	3.55	128,200	36,145	Fordham, Morris Heights, Mount Hope, Universit...	205

- The geometric and boundary information of all the community boards is found in the source geojson file of [The NYC Boundaries Map](#), which is an open source tool for viewing and querying overlapping administrative boundaries in New York City.^[5]

The key information of the database is in the *geometry* column, where the polygon or multipolygon data sets are presented for individual *boro_cd*. The geometric information is necessary for the function `folium.GeoJson` that is incorporated in the choropleth map.

	boro_cd	shape_area	shape_leng	geometry	Latitude	Longitude
0	311	103177785.365	51549.5578567	MULTIPOLYGON (((-73.97299 40.60881, -73.97259 ...	40.607216	-73.993636
1	313	88195686.2748	65821.875577	MULTIPOLYGON (((-73.98372 40.59582, -73.98305 ...	40.580649	-73.982140
2	312	99525500.0655	52245.8304843	MULTIPOLYGON (((-73.97140 40.64826, -73.97121 ...	40.631227	-73.983260
3	304	56663217.0191	37008.1004432	MULTIPOLYGON (((-73.89647 40.68234, -73.89653 ...	40.694549	-73.916836
4	206	42664311.28	35875.7111716	MULTIPOLYGON (((-73.87185 40.84376, -73.87192 ...	40.849604	-73.887530

- The search for food venues near each community board and the acquisition of venue details are performed with **Foursquare Places API**.^[6] In the process of data collection and cleaning, several issues are worth stressing. First, it only allows user to get the venues near specific locations rather than administrative districts. To make sure certain venue obtained belongs to the right community board, it is important to filter out venues by checking their coordination with the boundaries of community boards during search.

Second, since there is a limit of 50 for number of venues to return for each community board, the resulting dataset is incomplete in terms of counting all food venues in New York City. However, the dataset can be regarded as evenly random samples for these administrative districts and have significant value of statistics.

Third, it is found in the later investigation that venues corresponding to some category such as Flower Shop and Clothing Store are included in the database. To avoid misleading information in categorical analysis, category list is manually verified for the purpose of cleaning.

Last, there are venues which don't deliver either rating scores or price tiers message (None value). During the analysis of rating- or price-based information, database is wrangled accordingly to be ready for use.

The final database which contains the information of **venue id, name, latitude, longitude and category, rating, rating signal, price tier, tips, likes and listed count** is shown below:

boro_cd	Community Board	Community Board Latitude	Community Board Longitude	Venue Id	Venue Name	Venue Latitude	Venue Longitude	Venue Category	Rating	Rating Signal	Price	Tips Number	Likes	Listed Number	
0	201	Bronx CB 1	40.810909	-73.91647	4dae69211e7207bbeb106d82	Cypress Corner Deli	40.807903	-73.913446	Café	None	None	None	0	0	0
1	201	Bronx CB 1	40.810909	-73.91647	4c0c07676071a593902be232	Primos	40.808725	-73.915402	Deli / Bodega	None	None	None	1	1	0
2	201	Bronx CB 1	40.810909	-73.91647	4c0952f0bbc676b0b30148d5	La Morada	40.810670	-73.921758	Mexican Restaurant	8.9	84	None	29	57	223
3	201	Bronx CB 1	40.810909	-73.91647	4cf2eb9c1c158cfa95b9dab5	Fresco Pizza & Pasta	40.814358	-73.913183	Pizza Place	7	20	None	9	10	5
4	201	Bronx CB 1	40.810909	-73.91647	4bf1b1ef3fa220a150991820	Mexicocina	40.811837	-73.909512	Mexican Restaurant	9	84	2	24	55	71
5	201	Bronx CB 1	40.810909	-73.91647	4e4e242dbd4101d0d7a333fc	KFC	40.816524	-73.918461	Fried Chicken Joint	6.2	8	1	2	4	2
6	201	Bronx CB 1	40.810909	-73.91647	4d2e531979dd6ea8795588d3	Checkers	40.815628	-73.917990	Fast Food Restaurant	4.8	12	None	11	0	2
7	201	Bronx CB 1	40.810909	-73.91647	5db1caab98c98d0008f537f9	Starbucks	40.816376	-73.918473	Coffee Shop	7.1	0	None	0	0	0
8	201	Bronx CB 1	40.810909	-73.91647	56eaeefc4498e0c20d8bda9f1	Capri Cakes	40.816880	-73.920876	Bakery	None	None	None	1	3	0
9	201	Bronx CB 1	40.810909	-73.91647	6b54e70182a750002c111916	Taco Bell	40.818109	-73.923674	Taco Place	6	0	None	0	0	0