The Battle of Neighborhoods

Thai Restaurants in New York City

Chalothorn Chavalitcheevinkul

1.Introduction & Business Problem :

Problem Background:

The New York city is the one of populars city in the world. It is diverse and is the biggest city in USA. And this citi has multicultural so this is strong point, It can provides lot of business opportunities and business friendly environment. It has attracted many people into the market. New York is a hub of business and commerce of the world. With a major center for banking and finance, retailing, food, transportation and tourism in the United States.

Throughout its history, New York City has been a major point of entry for immigrants in last 200 years; the term "melting pot" was coined to describe densely populated immigrant neighborhoods on the Lower East Side. As many languages are spoken in New York, making it the most linguistically diverse city in the world.

From its diverse culture, comes diverse food menus. There are many restaurants in New york City, each belonging to different categories like Chinese , Japanese ,Thai, Indian , French, Italian etc.

Problem Description:

The goal for this is to list and visualize all major parts of New York City that has Thai restaurants.

Questions that can be asked.

- What is best location in New York City for Thai Cuisine ?
- Which areas have potential Thai Resturant Market ?
- Which all areas lack of Thai Resturants ?
- Which is the best place to stay if I prefer Thai Cuisine ?

2. Data

For this capstone project, following data :

New York City data that contains list Boroughs, Neighborhoods along with their latitude and longitude.

Source : https://cocl.us/new_york_dataset

Description : This data set contains the required information. And we will use this data set to explore various neighborhoods of new york city.

Thai resturants in each neighborhood of new york city.

Source : Fousquare API

Description : By using this api we will get all the venues in each neighborhood. We can filter these venues to get only Thai restaurants.

Geo Data.

Down from this Source :
https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm

Description : By using this geo space data we will get the New york Borough boundaries that will help us visualize choropleth map using folium.

Approach

1. Collect the new york city data from https://cocl.us/new_york_dataset
2. Using FourSquare API we will find all venues for each neighborhood.
3. Filter out all venues that are Thai Resturants.
4. Find rating , tips and like count for each Thai Resturants using FourSquare API.
5. Using rating for each resturant , then we will sort that data.
6. Visualize the Ranking of neighborhoods using folium library

Processing Dataset :

New York city neighbourhood has 5 boroughs and 306 neighborhoods. we will essentially need a dataset that contains the 5 boroughs and the neighborhoods that exist in each borough as well as the the latitude and longitude coordinates of each neighborhood.

the dataset can download here : https://cocl.us/new_york_dataset

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

Figure 1. Sample data in this dataset.

3.Methodology :

When we already put dataset into pandas dataframe. exploratory data analysis on this dataset show that Queens Borough contains highest number of neighbourhoods.

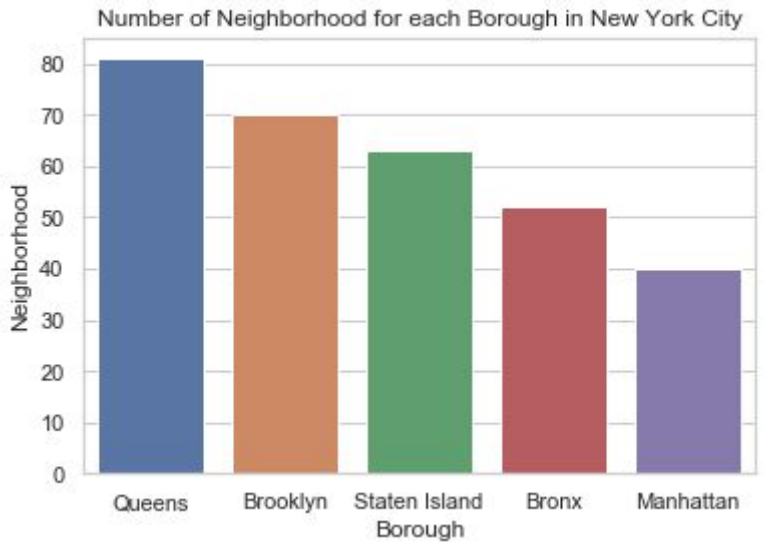

Figure 2. Number of Neighborhood for each Borough in New York City

Using Foursquare API to get Thai Restaurants in each borough

Next step is define a function to interact with FourSquare API and get Thai Restaurants within a radius of 1000 metres for a given latitude and longitude. function will return us the venue id , venue name and category.

	Borough	Neighborhood	ID	Name
0	Bronx	Riverdale	4b94a4a2f964a520787f34e3	Nam Thai
1	Bronx	Kingsbridge	4b94a4a2f964a520787f34e3	Nam Thai
2	Manhattan	Marble Hill	4a739e29f964a520f5dc1fe3	Siam Square
3	Bronx	Country Club	4b4c90d3f964a52046b626e3	Honey's Thai Pavilion
4	Bronx	Parkchester	5585c948498e0bd8a8a706f1	Thai No. 1

Figure 3. Dataframe show that Thai Restaurant data from radius in lat-long in dataset

	index	Neighborhood	count
0	20	Elmhurst	13
1	80	Woodside	6
2	40	Kensington	5
3	36	Jackson Heights	5
4	57	Ravenswood	4

Therefor the target is to find neighbourhood that contains top number of Thai Restaurants. and dataframe make us knew that Elmhurst have highest number of Thai Restaurants

Figure 4. Show dataframe that sorted by total of Thai Restaurants in each neighbourhood decending

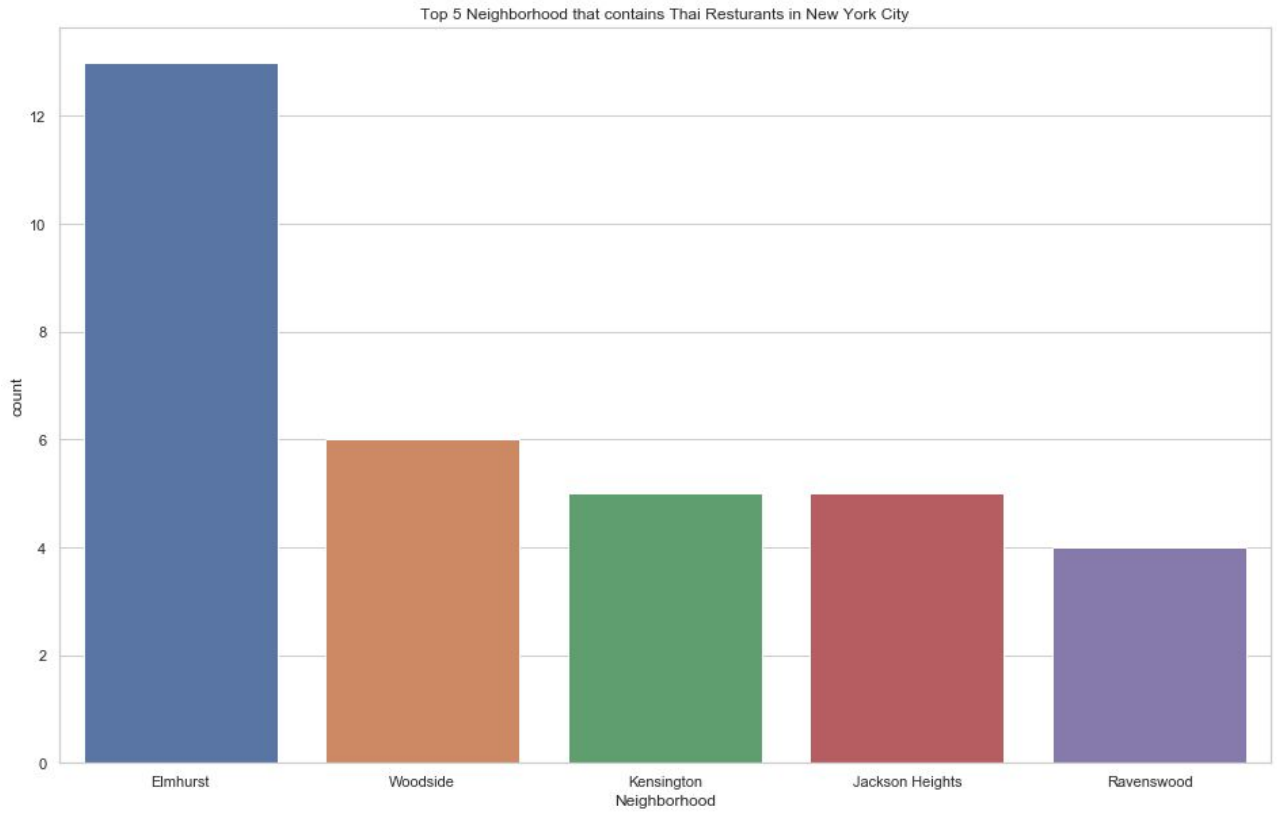

Figure 5. Show dataframe that contains TOP 5 number of Thai Restaurants in each neighbourhood

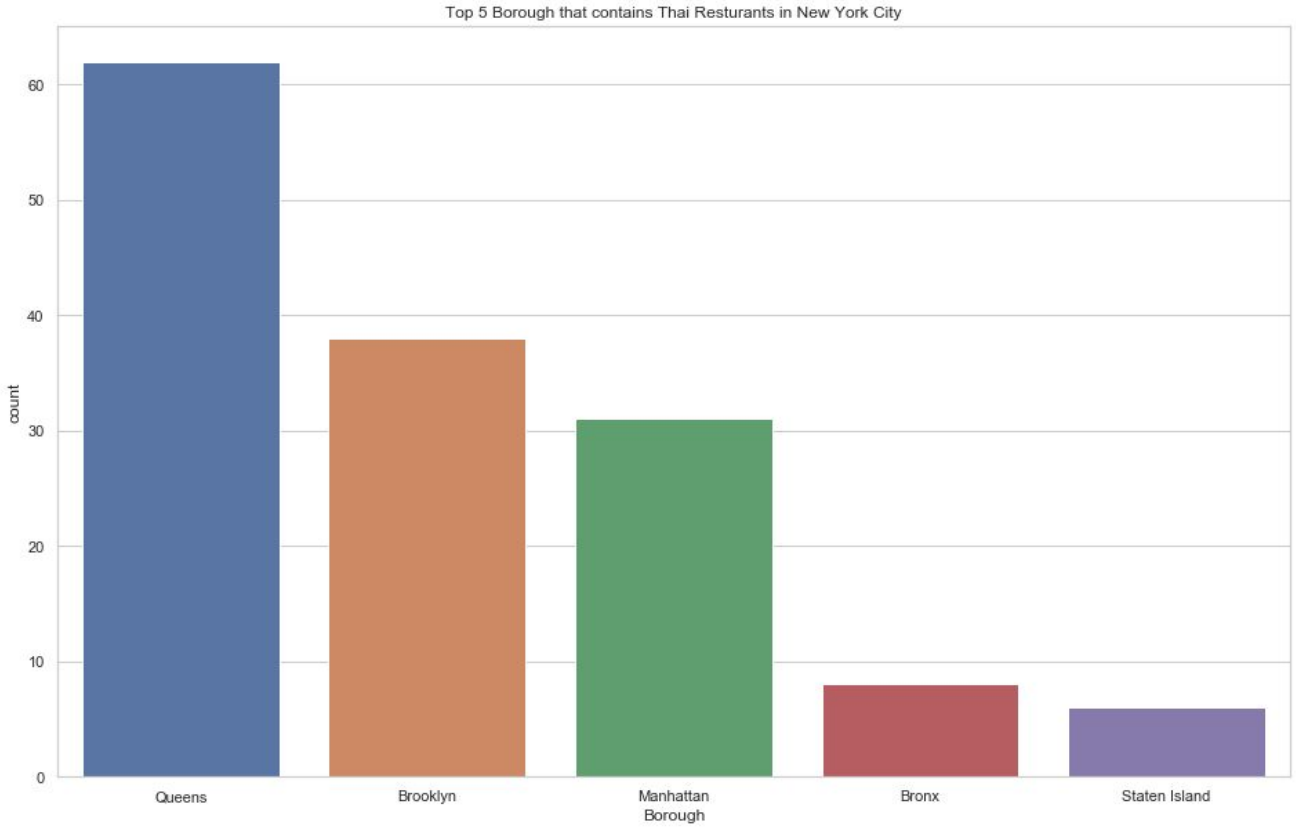

Figure 6. Show dataframe that contains TOP 5 number of Thai Restaurants in each Borough

Using Foursquare API to get Thai Restaurants details

Now we will define a function to get venue details like like count , rating , tip counts for a given venue id. This will be used for ranking. All details that we will show each ranking by different attribute that contained in Thai Restaurants. Foursquare API has limitations for using their data, For this reason I can get only 100 Thai Restaurants details

	Borough	Neighborhood	ID	Name	Likes	Rating	Tips
0	Bronx	Riverdale	4b94a4a2f964a520787f34e3	Nam Thai	3	7.0	5
1	Bronx	Kingsbridge	4b94a4a2f964a520787f34e3	Nam Thai	3	7.0	5
2	Manhattan	Marble Hill	4a739e29f964a520f5dc1fe3	Siam Square	15	7.2	13
3	Bronx	Country Club	4b4c90d3f964a52046b626e3	Honey's Thai Pavilion	19	7.4	28
4	Bronx	Parkchester	5585c948498e0bd8a8a706f1	Thai No. 1	17	7.9	1
5	Bronx	Van Nest	5585c948498e0bd8a8a706f1	Thai No. 1	17	7.9	1
6	Bronx	Spuyten Duyvil	4a739e29f964a520f5dc1fe3	Siam Square	15	7.2	13
7	Bronx	Spuyten Duyvil	4b94a4a2f964a520787f34e3	Nam Thai	3	7.0	5
8	Bronx	Pelham Bay	4b4c90d3f964a52046b626e3	Honey's Thai Pavilion	19	7.4	28
9	Brooklyn	Bay Ridge	50e8d45ae4b06202fbe39b08	Top Thai	14	7.2	11
10	Brooklyn	Sunset Park	4accd335f964a520a8c920e3	Mai Thai	10	7.2	9
11	Brooklyn	Greenpoint	4ac3a58ef964a5202d9c20e3	Ott	62	7.6	34
12	Brooklyn	Flatbush	5903f2d9c21cb14daa03993f	Corthaiyou	26	7.4	7
13	Brooklyn	Flatbush	593b1eabfc9e940adf8a5da1	Mondayoff By Plant Love House	23	7.3	9

Figure 7. Sample of dataframe that contains each Thai Restaurants details

Analysis from Thai Restaurants details

Thai Resturant with maximum Likes is **Uncle Boons**

```
Borough                    Manhattan
Neighborhood               Chinatown
ID            516863b1011ca6684e37a88f
Name                     Uncle Boons
Likes                            989
Rating                           9.2
Tips                             318
```

Thai Resturant with maximum Rating is **Up Thai**

```
Borough                    Manhattan
Neighborhood              Lenox Hill
ID            52be229211d2aabfe6f24d20
Name                         Up Thai
Likes                            593
Rating                           9.3
Tips                             153
```

Thai Resturant with maximum Tips is **SriPraPhai**

Borough	Queens
Neighborhood	Woodside
ID	3fd66200f964a5204ef11ee3
Name	SriPraPhai
Likes	755
Rating	8.9
Tips	354

Neighborhood with Average Rating

Visualize neighborhood with maximum average rating of Thai restaurants to show that what neighborhood have high average rating from all Thai restaurants in that neighborhood. Lenox neighborhood got highest rank of average rating that mean overall of Thai restaurants become popular.

	Neighborhood	Average Rating
25	Lenox Hill	9.150000
38	Soho	9.050000
26	Little Italy	9.050000
5	Chinatown	9.050000
14	East Williamsburg	8.900000
39	South Ozone Park	8.900000
20	Gowanus	8.850000
49	Yorkville	8.800000
6	Clinton	8.750000
4	Carroll Gardens	8.733333
0	Astoria	8.700000
18	Fort Greene	8.600000
31	Park Slope	8.600000
8	Cobble Hill	8.500000
11	Downtown	8.400000
22	Jackson Heights	8.380000
15	Elmhurst	8.338462

Figure 8. dataframe show neighborhood ranking that having high average rating of Thai resturants

And we zoom out to Borough that average rating of Thai restaurants is decreasing from neighborhood view and Manhattan got high rating

	Borough	Rating
2	Manhattan	8.443478
3	Queens	8.164706
1	Brooklyn	8.068571
0	Bronx	7.350000

Figure 9. dataframe show Borough ranking that having high average rating of Thai resturants

We want to get neighborhood with Average Rating to plot on Folium map, But now we don't have Lat-Lon so we need to merge dataframe from Neighborhood with Average Rating and dataframe from New York City Dataset on Neighborhood, Finally we got dataframe that contains Neighborhood, Average Rating and Lat-Long

	Neighborhood	Average Rating	Borough	Latitude	Longitude
0	Lenox Hill	9.150000	Manhattan	40.768113	-73.958860
1	Soho	9.050000	Manhattan	40.722184	-74.000657
2	Little Italy	9.050000	Manhattan	40.719324	-73.997305
3	Chinatown	9.050000	Manhattan	40.715618	-73.994279
4	East Williamsburg	8.900000	Brooklyn	40.708492	-73.938858
5	South Ozone Park	8.900000	Queens	40.668550	-73.809865
6	Gowanus	8.850000	Brooklyn	40.673931	-73.994441
7	Yorkville	8.800000	Manhattan	40.775930	-73.947118
8	Clinton	8.750000	Manhattan	40.759101	-73.996119
9	Carroll Gardens	8.733333	Brooklyn	40.680540	-73.994654
10	Astoria	8.700000	Queens	40.768509	-73.915654
11	Fort Greene	8.600000	Brooklyn	40.688527	-73.972906
12	Park Slope	8.600000	Brooklyn	40.672321	-73.977050
13	Cobble Hill	8.500000	Brooklyn	40.687920	-73.998561

Figure 10. dataframe show contains Neighborhood, Average Rating and Lat-Long of Thai resturants

For fully function on Folium map it need to label every pin that must contains detail about restaurants and its rating,And display marker only that have rating more than 8.5 because 8.5 is a higher score than top rating in Borough view

	Neighborhood	Average Rating	Borough	Latitude	Longitude	Labels
0	Lenox Hill	9.150000	Manhattan	40.768113	-73.958860	Lenox Hill, Manhattan (Rating: 9.15)
1	Soho	9.050000	Manhattan	40.722184	-74.000657	Soho, Manhattan (Rating: 9.05)
2	Little Italy	9.050000	Manhattan	40.719324	-73.997305	Little Italy, Manhattan (Rating: 9.05)
3	Chinatown	9.050000	Manhattan	40.715618	-73.994279	Chinatown, Manhattan (Rating: 9.05)
4	East Williamsburg	8.900000	Brooklyn	40.708492	-73.938858	East Williamsburg, Brooklyn (Rating: 8.9)
5	South Ozone Park	8.900000	Queens	40.668550	-73.809865	South Ozone Park, Queens (Rating: 8.9)
6	Gowanus	8.850000	Brooklyn	40.673931	-73.994441	Gowanus, Brooklyn (Rating: 8.85)
7	Yorkville	8.800000	Manhattan	40.775930	-73.947118	Yorkville, Manhattan (Rating: 8.8)
8	Clinton	8.750000	Manhattan	40.759101	-73.996119	Clinton, Manhattan (Rating: 8.75)
9	Carroll Gardens	8.733333	Brooklyn	40.680540	-73.994654	Carroll Gardens, Brooklyn (Rating: 8.733333333...
10	Astoria	8.700000	Queens	40.768509	-73.915654	Astoria, Queens (Rating: 8.7)
11	Fort Greene	8.600000	Brooklyn	40.688527	-73.972906	Fort Greene, Brooklyn (Rating: 8.6)
12	Park Slope	8.600000	Brooklyn	40.672321	-73.977050	Park Slope, Brooklyn (Rating: 8.6)
13	Cobble Hill	8.500000	Brooklyn	40.687920	-73.998561	Cobble Hill, Brooklyn (Rating: 8.5)

Figure 11. dataframe show contains Labels for markers

After we create map using Folium that will display Thai restaurants that got rating over 8.5. So we saw Thai restaurants with high rating bunched together in Manhattan

Figure 12. Map using Folium with markers

So now we need to create choropleth map using folium, With New York Geodata we will get boundaries materials for create map. Then we create map and fill color shade comparing Average rating

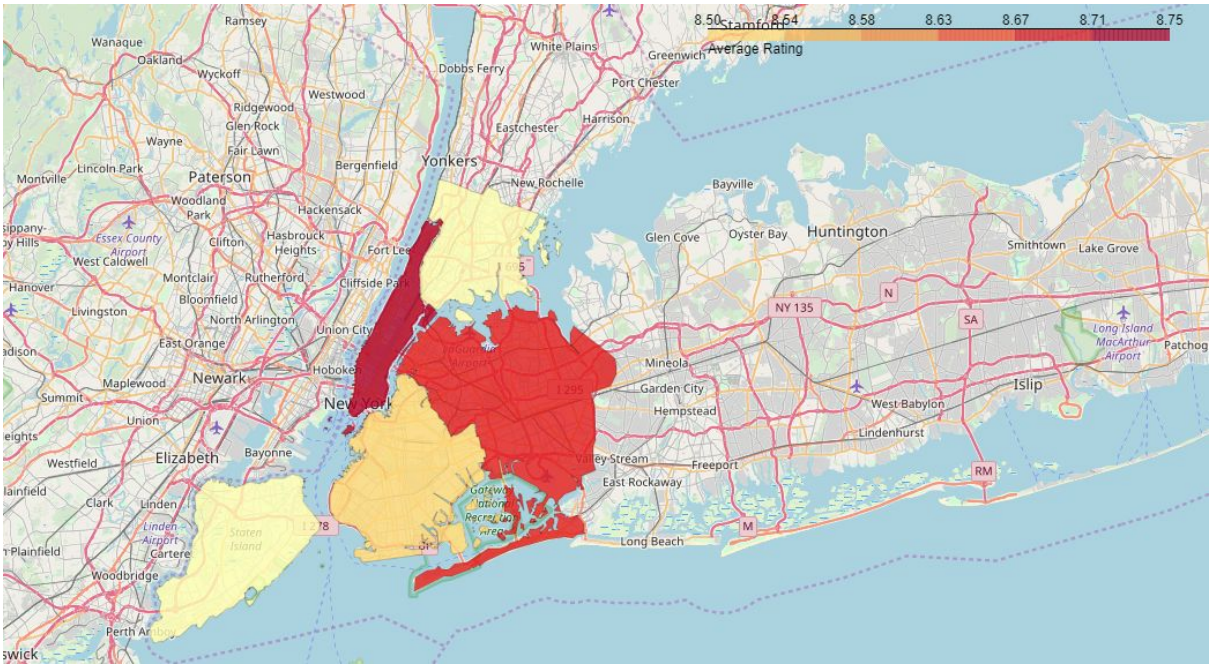

Figure 13. choropleth map using Folium show color shade compare with Average rating of Thai restaurants