

Capstone Project - The Battle of the Neighborhoods

This report covers the following topics in succession:

- 1. Introduction**
- 2. Business Problem**
- 3. Data**
- 4. Methodology**
- 5. Results**
- 6. Discussions**
- 7. Conclusion**

Introduction

The City of New York, is the most populous city in the United States. It is diverse and is the financial capital of USA. It is multicultural. It provides lot of business opportunities and business friendly environment. It has attracted many different players into the market. It is a global hub of business and commerce. The city is a major center for banking and finance, retailing, world trade, transportation, tourism, real estate, new media, traditional media, advertising, legal services, accountancy, insurance, theater, fashion, and the arts in the United States. This also means that the market is highly competitive. As it is highly developed city so cost of doing business is also one of the highest. Thus, any new business venture or expansion needs to be analyzed carefully. The insights derived from analysis will give good understanding of the business environment which help in strategically targeting the market. This will help in reduction of risk. And the Return on Investment will be reasonable.

Business Problem

The City of New York is famous for its amazing cuisine. Its food culture includes an array of international cuisines influenced by the city's immigrant history. Thai restaurants have become so popular in the United States now it seems that there is one on every corner, not only in major cities but also in smaller cities. Starting a Thai restaurant can be a great business opportunity, but you need to distinguish yourself from others to enjoy long-term success.

If you plan a real restaurant that can demand higher prices for fresh sea food, focus on neighborhoods and outlets that already attract sophisticated Thai clients. If you plan a cheap buffet restaurant, points to the masses looking for affordable high-traffic locations with large shopping centers and other local points of interest.

The stakeholder wants to open his business in Manhattan area, so I focus on that borough during my analysis. We define potential neighborhood based on the number of Thai restaurants which are operating right in each neighborhood. Manhattan has full potential but also is a very challenging district to open a business because of high competition. New Thai restaurant should be opened in an area where the it can attract more customers. Therefore, in this analysis it is necessary to ensure that we have enough customers and that we are also close to other Thai outlets so that we get a good turnover of Thai food lovers.

Data

1: New York has a total of 5 boroughs and 306 neighborhoods. In order to segment the neighborhoods and explore them, we will essentially need a dataset that contains the 5 boroughs and the neighborhoods that exist in each borough as well as the latitude and

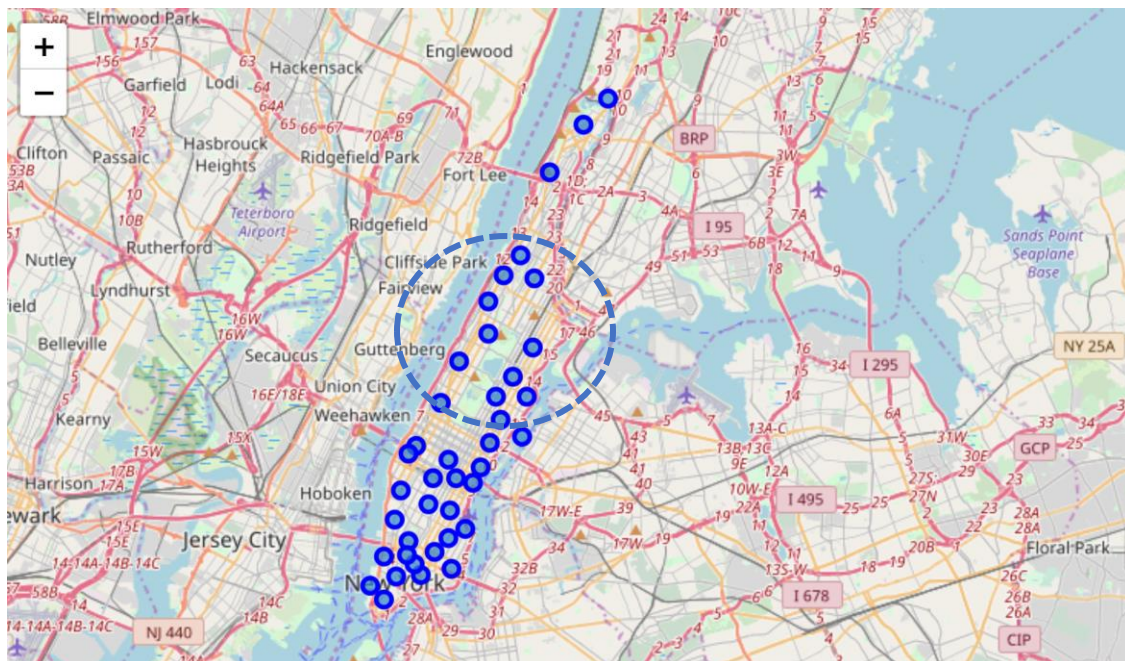
longitude coordinates of each neighborhood. This dataset exists for free on the web. Link to the dataset is: https://geo.nyu.edu/catalog/nyu_2451_34572

	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

2: New York City's geographical coordinates data will be utilized as input for the Foursquare API, that will be leveraged to provision venues information for each neighborhood. We will use the Foursquare API to explore neighborhoods in New York City. The below is image of the Foursquare API data.

In addition, [Thai Category Id 4bf58dd8d48988d1d2941735](#) is used for retrieving data from Foursquare API.

The below map shows all the neighborhoods of New York City. The blue circle roughly covers Manhattan.

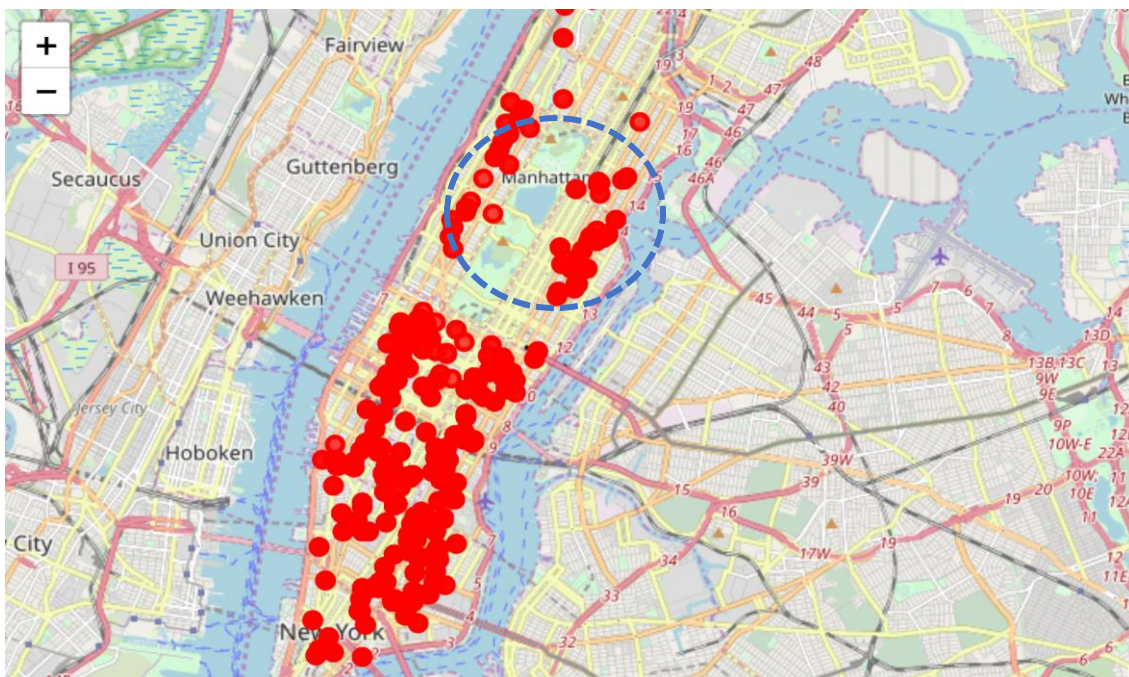


Methodology

The idea is to gather as much information as possible regarding Thai restaurants in New York City and try to categorize them based on their individual similarity. We converted addresses into their equivalent latitude and longitude values. Then we used the Foursquare API to explore neighborhoods in Manhattan, New York. After that, we called an explore function to get Thai restaurant categories in each neighborhood.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.910660	Siam Square	40.878796	-73.916701	Thai Restaurant
1	Marble Hill	40.876551	-73.910660	Nam Thai	40.886211	-73.909991	Thai Restaurant
2	Marble Hill	40.876551	-73.910660	Holy Basil	40.881678	-73.914376	Thai Restaurant
3	Chinatown	40.715618	-73.994279	Noree Thai Bazaar	40.717900	-73.992966	Thai Restaurant
4	Chinatown	40.715618	-73.994279	Wayla	40.718328	-73.993045	Thai Restaurant

The below map shows the various locations of Thai restaurant outlets in New York. From the dispersed distribution near Manhattan, we can see that the frequency of Thai restaurants here are quite low compared to the south of the city.



Thai Restaurants in Manhattan

	Neighborhood	Asian Restaurant	Chinese Restaurant	Food Truck	Indian Restaurant	Japanese Restaurant	Malay Restaurant	Ramen Restaurant	Sushi Restaurant	Thai Restaurant
0	Marble Hill	0	0	0	0	0	0	0	0	1
1	Marble Hill	0	0	0	0	0	0	0	0	1
2	Marble Hill	0	0	0	0	0	0	0	0	1
3	Chinatown	0	0	0	0	0	0	0	0	1
4	Chinatown	0	0	0	0	0	0	0	0	1

Then use this feature to group the neighborhoods into clusters. K-means clustering algorithm was used to cluster the restaurants. And also, the Folium library to visualize the neighborhoods in Manhattan and its emerging clusters.

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Manhattan	Marble Hill	40.876551	-73.910660	3	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant
1	Manhattan	Chinatown	40.715618	-73.994279	0	Thai Restaurant	Asian Restaurant	Vietnamese Restaurant	Sushi Restaurant	Malay Restaurant
2	Manhattan	Washington Heights	40.851903	-73.936900	3	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant
3	Manhattan	Inwood	40.867684	-73.921210	3	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant
4	Manhattan	Hamilton Heights	40.823604	-73.949688	3	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant

Results

Let us first visualize the cluster data as returned by the algorithm.

Cluster 0

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
1	Chinatown	Thai Restaurant	Asian Restaurant	Vietnamese Restaurant	Sushi Restaurant	Malay Restaurant	Wine Bar	Ramen Restaurant	Japanese Restaurant	Indian Restaurant
19	East Village	Thai Restaurant	Sushi Restaurant	Chinese Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Food Truck	Ramen Restaurant	Malay Restaurant
20	Lower East Side	Thai Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant
22	Little Italy	Thai Restaurant	Asian Restaurant	Vietnamese Restaurant	Sushi Restaurant	Malay Restaurant	Wine Bar	Ramen Restaurant	Japanese Restaurant	Indian Restaurant
23	Soho	Thai Restaurant	Asian Restaurant	Sushi Restaurant	Malay Restaurant	Wine Bar	Vietnamese Restaurant	Ramen Restaurant	Japanese Restaurant	Indian Restaurant
25	Manhattan Valley	Thai Restaurant	Indian Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant
31	Noho	Thai Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Chinese Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant

Cluster 1

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
8	Upper East Side	Thai Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant
9	Yorkville	Thai Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant
12	Upper West Side	Thai Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant
18	Greenwich Village	Thai Restaurant	Asian Restaurant	Sushi Restaurant	Chinese Restaurant	Wine Bar	Vietnamese Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant
21	Tribeca	Thai Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant
24	West Village	Thai Restaurant	Chinese Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant
26	Morningside Heights	Thai Restaurant	Indian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Food Truck
30	Carnegie Hill	Thai Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant
37	Stuyvesant Town	Thai Restaurant	Sushi Restaurant	Wine Bar	Vietnamese Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck

Cluster 2

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
13	Lincoln Square	Thai Restaurant	Ramen Restaurant	Food Truck	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Malay Restaurant	Japanese Restaurant
14	Clinton	Thai Restaurant	Ramen Restaurant	Food Truck	Asian Restaurant	Chinese Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Malay Restaurant
15	Midtown	Thai Restaurant	Food Truck	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant
16	Murray Hill	Thai Restaurant	Food Truck	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant
17	Chelsea	Thai Restaurant	Malay Restaurant	Japanese Restaurant	Food Truck	Chinese Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant
27	Gramercy	Thai Restaurant	Chinese Restaurant	Sushi Restaurant	Malay Restaurant	Japanese Restaurant	Food Truck	Asian Restaurant	Wine Bar	Vietnamese Restaurant
33	Midtown South	Thai Restaurant	Food Truck	Japanese Restaurant	Chinese Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant
34	Sutton Place	Thai Restaurant	Food Truck	Chinese Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant
35	Turtle Bay	Thai Restaurant	Food Truck	Chinese Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant

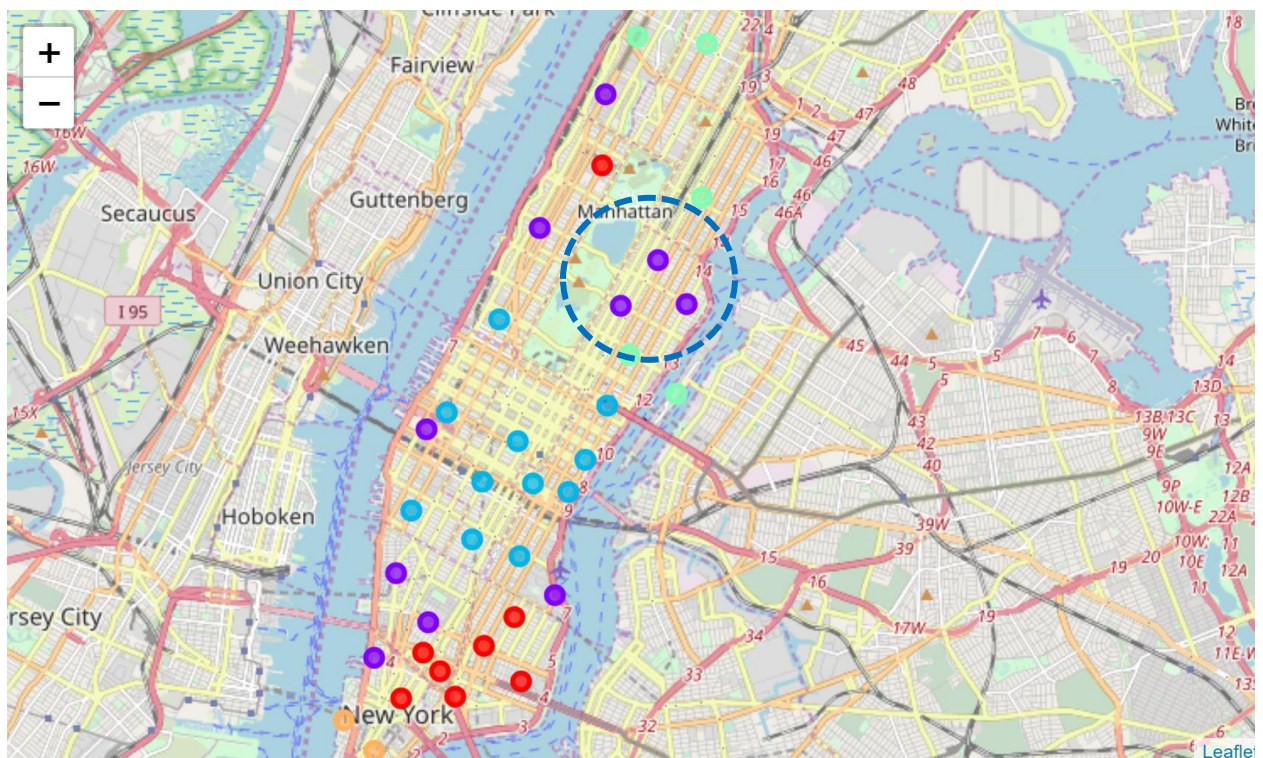
Cluster 3

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
0	Marble Hill	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck
2	Washington Heights	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck
3	Inwood	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck
4	Hamilton Heights	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck
5	Manhattanville	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck
6	Central Harlem	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck
7	East Harlem	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck
10	Lenox Hill	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck

Cluster 4

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
28	Battery Park City	Thai Restaurant	Japanese Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Indian Restaurant	Food Truck
29	Financial District	Thai Restaurant	Japanese Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Indian Restaurant	Food Truck

Now lets carefully review the cluster map.



After studying the cluster data from the dataframes and reviewing the cluster map, we come to the conclusion that most optimum area to open a new Thai Restaurant would be near **Yorkville and Carnegie north of Upper East Side in Cluster 1.**

Discussion

In this section, I would be discussing the observations I have noted and the recommendation that I can make based on the results.

This analysis is performed on limited data. Most optimum results may vary based on input parameters and also the initial cluster value input.

Also, few key highlights:

- Washington Heights would be least favorable location to open a Thai restaurant pertaining to the sparsity in the region which may be as a result of people's preferences.
- More detailed analysis can be done and better results can be achieved by considering other factors such as transportation, demographics of inhabitants.

Conclusion

Although all of the goals of this project were met, there is definitely room for further improvement and development. However, this analysis can be used as a dependable guideline for any stakeholder planning to venture having this use case.

As per the neighborhood or restaurant type mentioned, this analysis can be referred to. A venue with lowest risk and competition can be identified.

The location parameters can be tweaked to match any requirements around the world provided proper location specific data is present.