The Battle of Neighborhoods Report

1. Introduction:

1.1 Background

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 oppourtunities 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.

1.2 Business Problem:

The City of New York is famous for its excellent cuisine. It's food culture includes an array of international cuisines influenced by the city's immigrant history.

Thai food is quite diverse and mixed with various spices, It's make Thai food so popular in the United States, but only in Major cities. Therefore, the first thing that must be done before investing is to find the best location. It's a good opportunity to start investing in restaurants in a minor cities.

1.3 Target Audience

To recommend to the stakeholder or investor who was interested to invest Thai restaurant in Manhattan city at New York.

2. Data acquisition and cleaning

2.1 Data Source

New York Datasets has a total of 5 boroughs and 306 neighborhoods. In order to segement 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 the latitude and logitude coordinates of each neighborhood. Luckily, this dataset exists for free on the web. Feel free to try to find this dataset on your own, but here is the link to the dataset: 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
5	Bronx	Kingsbridge	40.881687	-73.902818
6	Manhattan	Marble Hill	40.876551	-73.910660
7	Bronx	Woodlawn	40.898273	-73.867315
8	Bronx	Norwood	40.877224	-73.879391
9	Bronx	Williamsbridge	40.881039	-73.857446

Foursquare API that will be leveraged to provision venues information for each neighborhood. We will use Forsquare API to explore neighborhoods in New York City with specific category for this case is Thai restaurant. The category id is "4bf58dd8d48988d149941735"

3. Methodology

3.1 Data Pre-processing and Understanding

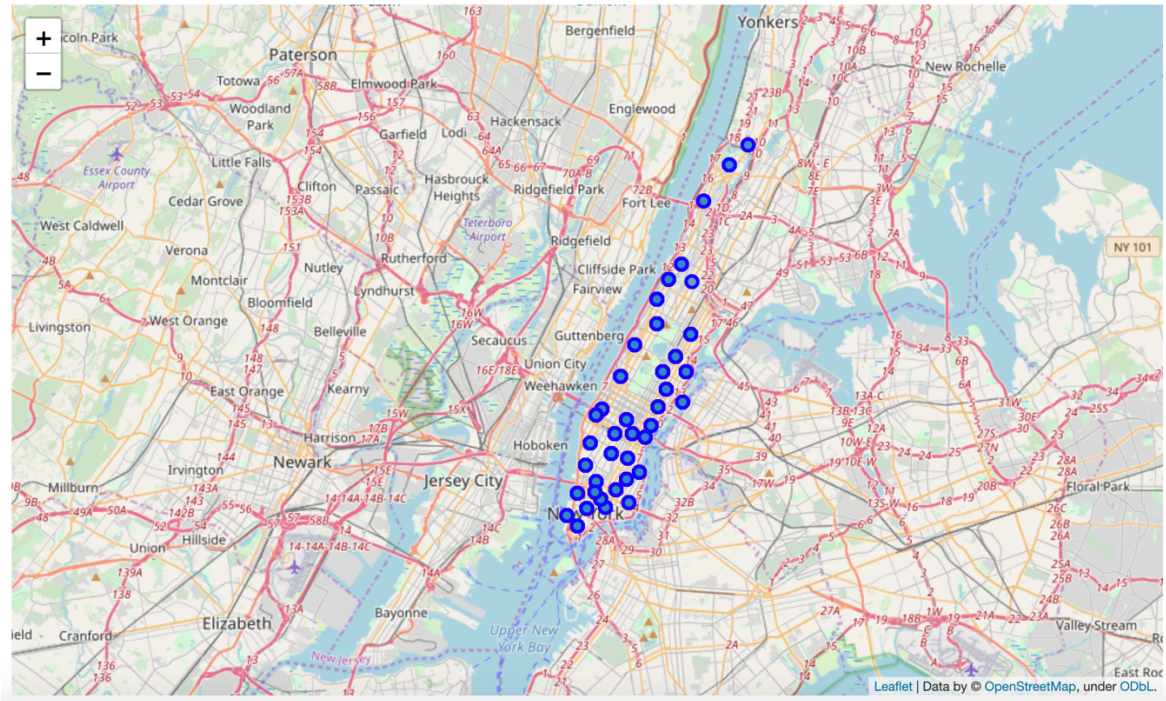
Convert New York datasets into tabular form and then select the datasets only Manhattan city and visualize to ensure the data is selected correctly. Then I will use the

Foursquare API to get Thai Restaurant in Manhattan city. The result from Foursquare API has 1095 records. Then visualize the neighborhood has few Thai restaurants.

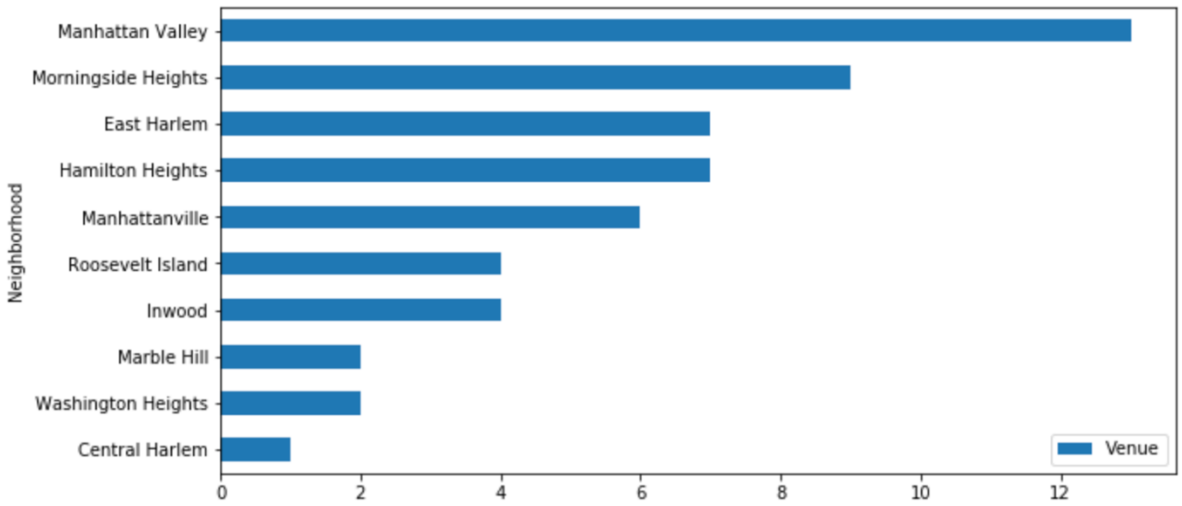
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.910660	Siam Square	40.878796	-73.916701	Thai Restaurant
1	Marble Hill	40.876551	-73.910660	Indian Road Café	40.872922	-73.918459	Café
2	Chinatown	40.715618	-73.994279	Uncle Boons	40.721200	-73.994463	Thai Restaurant
3	Chinatown	40.715618	-73.994279	Wayla	40.718291	-73.992584	Thai Restaurant
4	Chinatown	40.715618	-73.994279	Uncle Boons Sister	40.721360	-73.995592	Thai Restaurant

```
manhattan_venues.shape
```

```
(1095, 7)
```



Thai Restaurant in Manhattan



The neighborhood that has the least Thai restaurants

3.2 Feature Engineering

I use the one-hot encoding method to encode category data into a binary. then group the output from one-hot encoding to represents ratio Thai restaurant each neighborhood.

	Neighborhood	American Restaurant	Asian Restaurant	Bagel Shop	Bakery	Bar	Belgian Restaurant	Breakfast Spot	Café	Chinese Restaurant	Cocktail Bar	Coffee Shop	Deli / Bodega	Dessert Shop	Dim Sum Restaurant	Diner
0	Marble Hill	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	Marble Hill	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
2	Chinatown	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	Chinatown	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Chinatown	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

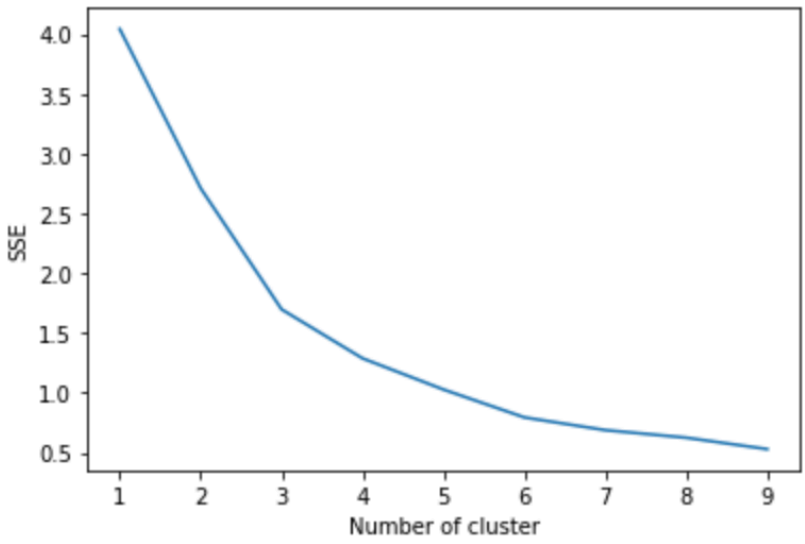
One-hot Encoding

	Neighborhood	American Restaurant	Asian Restaurant	Bagel Shop	Bakery	Bar	Belgian Restaurant	Breakfast Spot	Café	Chinese Restaurant	Cocktail Bar	Coffee Shop	Deli / Bodega	Dessert Shop
0	Battery Park City	0.000000	0.176471	0.000000	0.000000	0.000000	0.000000	0.000000	0.058824	0.000000	0.000000	0.000000	0.000000	0.000000
1	Carnegie Hill	0.000000	0.037037	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	Central Harlem	0.000000	0.000000	0.000000	0.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3	Chelsea	0.000000	0.090909	0.000000	0.000000	0.000000	0.030303	0.030303	0.030303	0.030303	0.000000	0.000000	0.000000	0.000000
4	Chinatown	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.025641	0.025641	0.000000	0.025641	0.000000	0.000000
5	Civic Center	0.000000	0.113636	0.000000	0.000000	0.000000	0.000000	0.000000	0.022727	0.045455	0.022727	0.022727	0.000000	0.000000
6	Clinton	0.000000	0.023256	0.000000	0.000000	0.046512	0.000000	0.000000	0.046512	0.000000	0.000000	0.000000	0.000000	0.000000
7	East Harlem	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
8	East Village	0.000000	0.022727	0.022727	0.000000	0.000000	0.000000	0.000000	0.022727	0.022727	0.022727	0.000000	0.000000	0.022727
9	Financial District	0.000000	0.095238	0.000000	0.000000	0.000000	0.000000	0.000000	0.047619	0.000000	0.000000	0.000000	0.000000	0.000000
10	Flatiron	0.000000	0.050847	0.000000	0.016949	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.033898

Ratio Thai Restaurant in Manhattan

3.3 Modeling

I use K-means clustering algorithms to cluster neighborhood at Manhattan. The evaluation of the model using the elbow method so the optimal K is 6 ks.



SSE using Elbow Method

4. Results

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	10th Most Common Venue
5	Manhattanville	Thai Restaurant	Japanese Restaurant	Bar	Cocktail Bar	Asian Restaurant	Bagel Shop	Dessert Shop	Fried Chicken Joint	Food Truck	Food Court
23	Soho	Thai Restaurant	Vietnamese Restaurant	Asian Restaurant	Malay Restaurant	Café	Salad Place	Sandwich Place	Shanghai Restaurant	Chinese Restaurant	Coffee Shop
24	West Village	Thai Restaurant	Vegetarian / Vegan Restaurant	Asian Restaurant	Sushi Restaurant	Chinese Restaurant	Coffee Shop	Soup Place	Sandwich Place	Malay Restaurant	New American Restaurant
28	Battery Park City	Thai Restaurant	Asian Restaurant	Wings Joint	Snack Place	Vietnamese Restaurant	Food Truck	Food Court	Café	Fried Chicken Joint	Breakfast Spot
29	Financial District	Thai Restaurant	Food Truck	Asian Restaurant	Snack Place	Food Court	Vietnamese Restaurant	Fried Chicken Joint	Café	Wings Joint	Vegetarian / Vegan Restaurant
32	Civic Center	Thai Restaurant	Vietnamese Restaurant	Asian Restaurant	Malay Restaurant	Chinese Restaurant	Vegetarian / Vegan Restaurant	Wings Joint	Food Truck	Fried Chicken Joint	Grocery Store

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	10th Most Common Venue
1	Chinatown	Thai Restaurant	Vietnamese Restaurant	Malay Restaurant	Salad Place	Vegetarian / Vegan Restaurant	Sushi Restaurant	Dim Sum Restaurant	Shanghai Restaurant	Seafood Restaurant	Coffee Shop
4	Hamilton Heights	Thai Restaurant	Japanese Restaurant	Bar	Cocktail Bar	Bagel Shop	Bakery	Fried Chicken Joint	Food Truck	Food Court	Dumpling Restaurant
8	Upper East Side	Thai Restaurant	Salad Place	Asian Restaurant	Sushi Restaurant	Sports Bar	Vietnamese Restaurant	Seafood Restaurant	Chinese Restaurant	Cocktail Bar	Dumpling Restaurant
9	Yorkville	Thai Restaurant	Vietnamese Restaurant	Salad Place	Asian Restaurant	Sushi Restaurant	Sports Bar	Seafood Restaurant	Chinese Restaurant	Cocktail Bar	Dumpling Restaurant
10	Lenox Hill	Thai Restaurant	Asian Restaurant	Japanese Restaurant	Seafood Restaurant	Chinese Restaurant	Bakery	Deli / Bodega	Food Truck	Food Court	Dumpling Restaurant
12	Upper West Side	Thai Restaurant	Chinese Restaurant	Salad Place	Pub	Vegetarian / Vegan Restaurant	Asian Restaurant	Sushi Restaurant	Ramen Restaurant	Dumpling Restaurant	Diner
13	Lincoln Square	Thai Restaurant	Bar	Sushi Restaurant	Café	Salad Place	Wings Joint	Coffee Shop	Food Truck	Food Court	Dumpling Restaurant
15	Midtown	Thai Restaurant	Vegetarian / Vegan Restaurant	Food Court	Asian Restaurant	Café	Japanese Restaurant	Vietnamese Restaurant	Dessert Shop	Latin American Restaurant	Noodle House

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	10th Most Common Venue
2	Washington Heights	Thai Restaurant	Wings Joint	Coffee Shop	Fried Chicken Joint	Food Truck	Food Court	Dumpling Restaurant	Diner	Dim Sum Restaurant	Dessert Shop
3	Inwood	Thai Restaurant	Café	Wings Joint	Coffee Shop	Fried Chicken Joint	Food Truck	Food Court	Dumpling Restaurant	Diner	Dim Sum Restaurant
7	East Harlem	Thai Restaurant	Wings Joint	Coffee Shop	Fried Chicken Joint	Food Truck	Food Court	Dumpling Restaurant	Diner	Dim Sum Restaurant	Dessert Shop
14	Clinton	Thai Restaurant	Bar	Café	Asian Restaurant	Noodle House	Gay Bar	Wings Joint	Dessert Shop	Food Truck	Food Court
19	East Village	Thai Restaurant	Chinese Restaurant	Asian Restaurant	Taco Place	Bagel Shop	Dumpling Restaurant	Café	Cocktail Bar	Dessert Shop	Deli / Bodega
20	Lower East Side	Thai Restaurant	Cocktail Bar	Chinese Restaurant	Asian Restaurant	Sushi Restaurant	Coffee Shop	Food Truck	Food Court	Dumpling Restaurant	Diner
25	Manhattan Valley	Thai Restaurant	Dessert Shop	Wings Joint	Coffee Shop	Fried Chicken Joint	Food Truck	Food Court	Dumpling Restaurant	Diner	Dim Sum Restaurant
27	Gramercy	Thai Restaurant	Food Truck	Vegetarian / Vegan Restaurant	Asian Restaurant	Bagel Shop	Bakery	Bar	Juice Bar	Dumpling Restaurant	Diner

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	10th Most Common Venue
6	Central Harlem	Bar	Wings Joint	Coffee Shop	Fried Chicken Joint	Food Truck	Food Court	Dumpling Restaurant	Diner	Dim Sum Restaurant	Dessert Shop

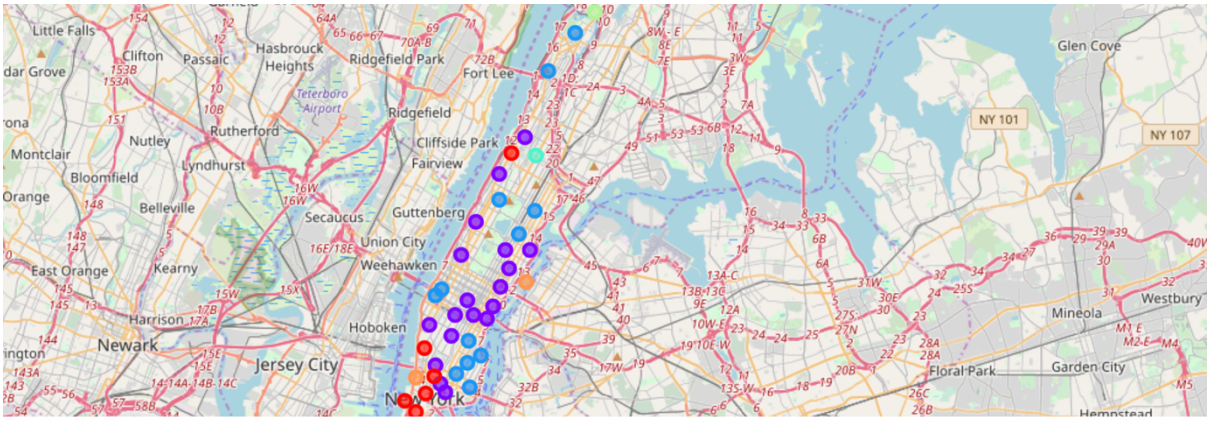
Cluster 4

	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	10th Most Common Venue
0	Marble Hill	Thai Restaurant	Café	Wings Joint	Coffee Shop	Fried Chicken Joint	Food Truck	Food Court	Dumpling Restaurant	Diner	Dim Sum Restaurant

Cluster 5

	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	10th Most Common Venue
11	Roosevelt Island	Asian Restaurant	Thai Restaurant	Chinese Restaurant	Wings Joint	Deli / Bodega	Fried Chicken Joint	Food Truck	Food Court	Dumpling Restaurant	Diner
21	Tribeca	Asian Restaurant	Sandwich Place	Thai Restaurant	Soup Place	Café	Snack Place	Wings Joint	Food Truck	Bar	Bakery

Visualize 6 Clusters



Cluster 0 – Red point ( )

Cluster 1 – Purple point ( )

Cluster 2 – Blue point ( )

Cluster 3 – Aquamarine point ( )

Cluster 4 – Mint Green point ( )

Cluster 5 – Orange point ( )

5. Discussion

5.1. This analysis perform on limited data. This may be correct or maybe not but if massive of data is available there will be better result.

5.2. There is scope to explore restaurant in Manhattan only

6. Conclusion

The best place to investing the Thai restaurant is Central Harlem. Inferior are Roosevelt Island and Tribeca have a nice Asian Restaurant environment. Thus, the people who want to eat Asian food usually go to those places.