1. Introduction and business problem

The City of New York, usually called either New York City (NYC) or simply New York (NY), is the most populous city in the United States. With an estimated 2018 population of 8,398,748 distributed over a land area of about 302.6 square miles (784 km^2), New York is also the most densely populated major city in the United States. Located at the southern tip of the state of New York, the city is the center of the New York metropolitan area, the largest metropolitan area in the world by urban landmass and one of the world's most populous megacities, with an estimated 19,979,477 people in its 2018 Metropolitan Statistical Area and 22,679,948 residents in its Combined Statistical Area. A global power city, New York City has been described as the cultural, financial, and media capital of the world, and exerts a significant impact upon commerce, entertainment, research, technology, education, politics, tourism, art, fashion, and sports. The city's fast pace has inspired the term *New York minute*. Home to the headquarters of the United Nations, New York is an important center for international diplomacy.

My client wants to open a new thai restaurant chain in Manhattan area, so I only focus on that borough during my analysis. The objective is to locate and recommend to the management which neighborhood of Newyork city will be best choice to start a restaurant. The Management also expects to understand the rationale of the recommendations made.

2. Data

To find the best location for our thai restaurant chain, we will use the following sources of information:

- Newyork 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

- From Foursquare Venues Categories
  - https://developer.foursquare.com/docs/resources/categories Thai restaurants Id - 4bf58dd8d48988d149941735

3. Methodology

In this project, I will use the basic methodology as taught in Week 3 lab.

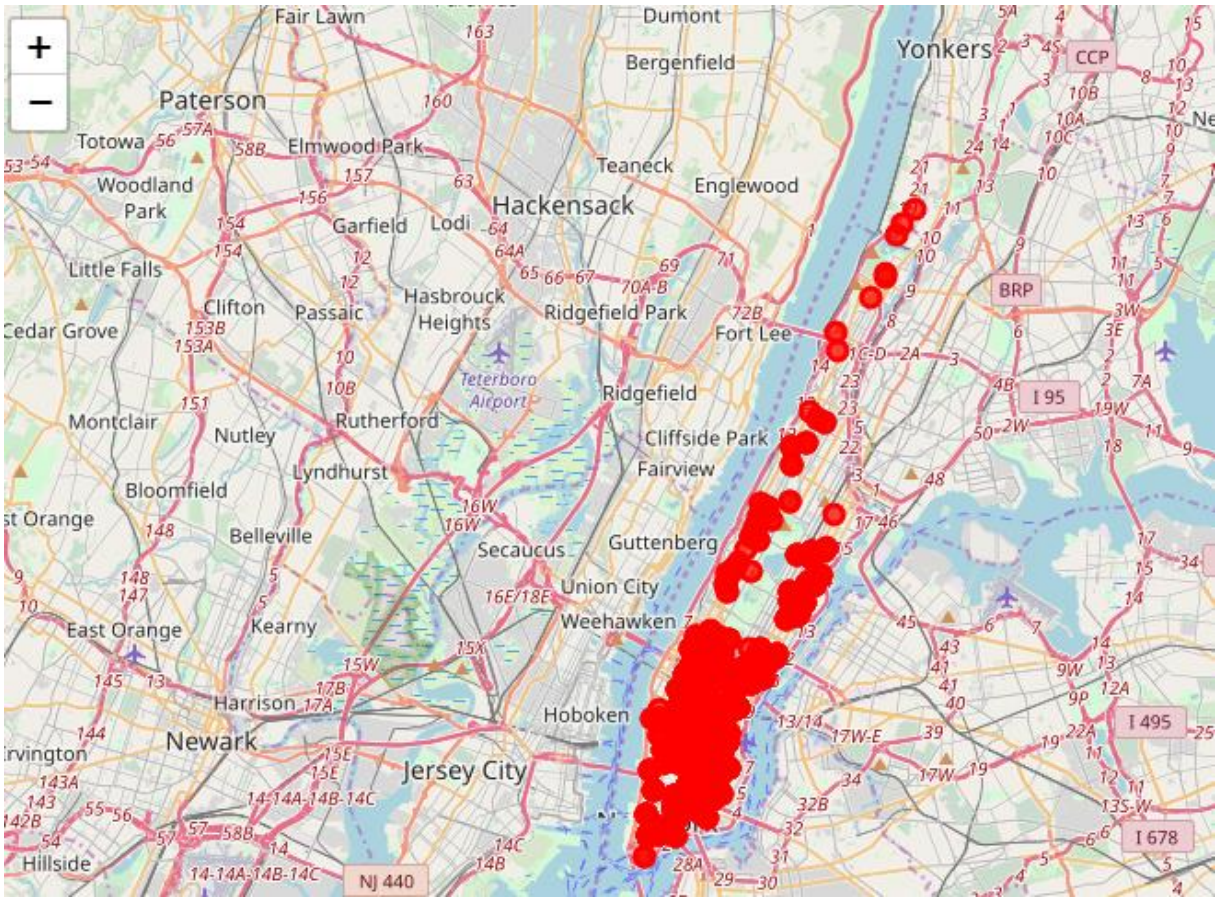
	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

Above, I have done convert addresses into their equivalent latitude and longitude values. Then we will use the Foursquare API to explore neighborhoods in Manhattan, New York. After that, explore function to Thai restaurants 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	Holy Basil	40.881678	-73.914376	Thai Restaurant
2	Marble Hill	40.876551	-73.910660	Nam Thai	40.886211	-73.909991	Thai Restaurant
3	Chinatown	40.715618	-73.994279	Noree Thai Bazaar	40.717900	-73.992966	Thai Restaurant
4	Chinatown	40.715618	-73.994279	Khao Man Gai NY	40.715514	-73.989680	Thai Restaurant

```
newyork_venues_thai.shape
```

```
(1000, 7)
```

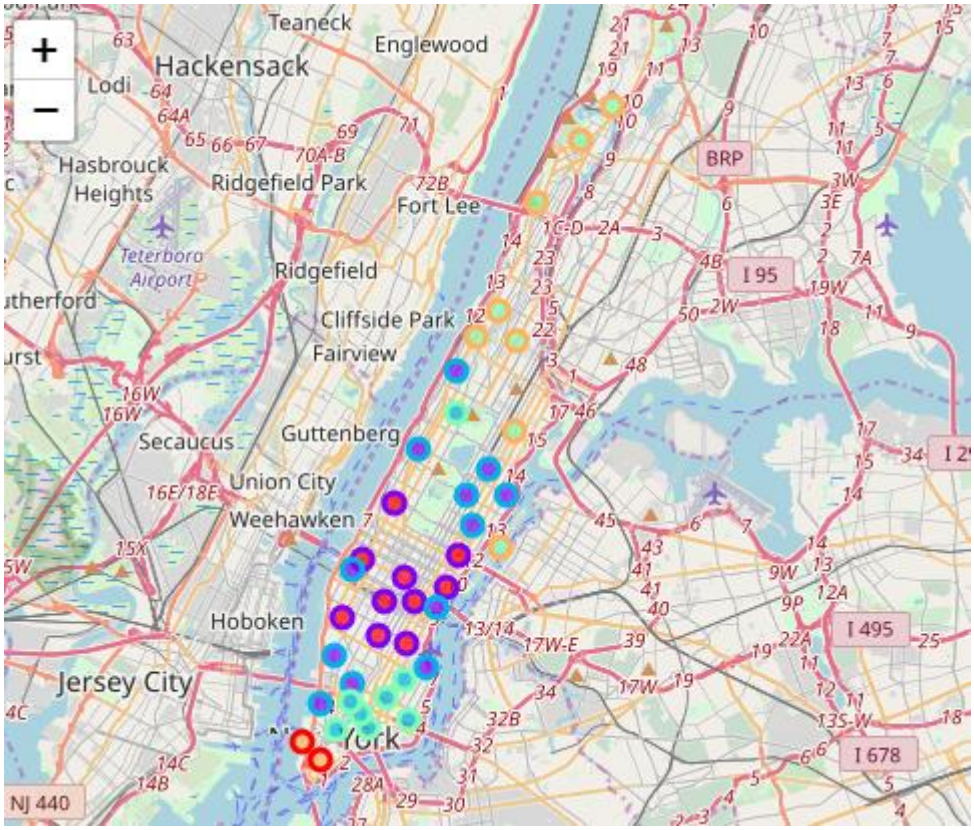


Thai restaurants in Manhattan

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

Then use this feature to group the neighborhoods into clusters K-means clustering algorithm will be use to complete this task. And also, the Folium library to visualize the neighborhoods in Manhattan and its emerging clusters.

	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	Battery Park City	Thai Restaurant	Japanese Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Indian Restaurant	Food Truck	Chinese Restaurant
1	Carnegie Hill	Thai Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck
2	Central Harlem	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck	Chinese Restaurant
3	Chelsea	Thai Restaurant	Malay Restaurant	Japanese Restaurant	Food Truck	Chinese Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Indian Restaurant
4	Chinatown	Thai Restaurant	Asian Restaurant	Vietnamese Restaurant	Sushi Restaurant	Malay Restaurant	Wine Bar	Ramen Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck



4. Results

K-mean Cluster Using K-mean to clustering data area with less number of sushi bars

Cluster 0

```
manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 0, manhattan_merged.columns[[1] + list(range(5, manhattan_merged
```

	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
13	Lincoln Square	Thai Restaurant	Ramen Restaurant	Food Truck	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant
14	Clinton	Thai Restaurant	Ramen Restaurant	Food Truck	Asian Restaurant	Chinese Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Malay Restaurant	Japanese Restaurant
15	Midtown	Thai Restaurant	Food Truck	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant
16	Murray Hill	Thai Restaurant	Food Truck	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant
17	Chelsea	Thai Restaurant	Malay Restaurant	Japanese Restaurant	Food Truck	Chinese Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Indian Restaurant
27	Gramercy	Thai Restaurant	Food Truck	Chinese Restaurant	Malay Restaurant	Japanese Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant
33	Midtown South	Thai Restaurant	Food Truck	Japanese Restaurant	Chinese Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant
34	Sutton Place	Thai Restaurant	Food Truck	Chinese Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant

Cluster 1

```
manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 1, manhattan_merged.columns[[1] + list(range(5, manhattan_merged
```

	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
8	Upper East Side	Thai Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck
9	Yorkville	Thai Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck
10	Lenox Hill	Thai Restaurant	Food Truck	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Chinese Restaurant
12	Upper West Side	Thai Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck
18	Greenwich Village	Thai Restaurant	Asian Restaurant	Sushi Restaurant	Chinese Restaurant	Wine Bar	Vietnamese Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant
21	Tribeca	Thai Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck
24	West Village	Thai Restaurant	Chinese Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck
26	Morningside Heights	Thai Restaurant	Indian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Food Truck	Chinese Restaurant

Cluster 2

```
manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 2, manhattan_merged.columns[[1] + list(range(5, manhattan_merged
```

	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	Asian Restaurant	Vietnamese Restaurant	Sushi Restaurant	Malay Restaurant	Wine Bar	Ramen Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck
19	East Village	Thai Restaurant	Sushi Restaurant	Chinese Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Food Truck	Ramen Restaurant	Malay Restaurant	Japanese Restaurant
20	Lower East Side	Thai Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck
22	Little Italy	Thai Restaurant	Asian Restaurant	Vietnamese Restaurant	Sushi Restaurant	Malay Restaurant	Wine Bar	Ramen Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck
23	Soho	Thai Restaurant	Asian Restaurant	Sushi Restaurant	Malay Restaurant	Wine Bar	Vietnamese Restaurant	Ramen Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck
25	Manhattan Valley	Thai Restaurant	Indian Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Food Truck
31	Noho	Thai Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Chinese Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant
32	Civic Center	Thai Restaurant	Malay Restaurant	Japanese Restaurant	Asian Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Indian Restaurant	Food Truck

Cluster 3

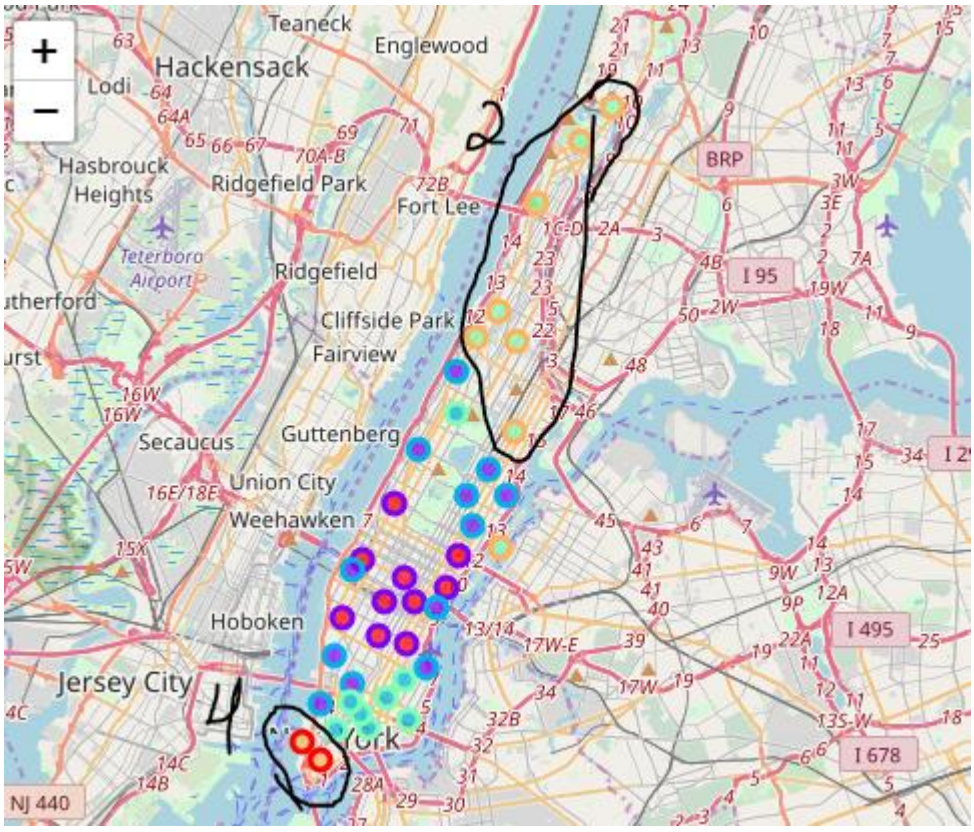
```
manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 3, manhattan_merged.columns[[1] + list(range(5, manhattan_merged
```

	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	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck	Chinese Restaurant
2	Washington Heights	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck	Chinese Restaurant
3	Inwood	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck	Chinese Restaurant
4	Hamilton Heights	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck	Chinese Restaurant
5	Manhattanville	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck	Chinese Restaurant
6	Central Harlem	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck	Chinese Restaurant
7	East Harlem	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck	Chinese Restaurant
11	Roosevelt Island	Thai Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Japanese Restaurant	Indian Restaurant	Food Truck	Chinese Restaurant

Cluster 4

```
manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 4, manhattan_merged.columns[[1] + list(range(5, manhattan_merge
```

	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
28	Battery Park City	Thai Restaurant	Japanese Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Indian Restaurant	Food Truck	Chinese Restaurant
29	Financial District	Thai Restaurant	Japanese Restaurant	Wine Bar	Vietnamese Restaurant	Sushi Restaurant	Ramen Restaurant	Malay Restaurant	Indian Restaurant	Food Truck	Chinese Restaurant



Based on dataframe analysis above Cluster 2 and Cluster 4 areas are the best places to open a new sushi bar business.

5. 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. This may be right or may be wrong. But if good amount of data is available there is scope to come up with better results.

- There is high competition in Midtown and Soho so it is very risky to open business in these areas.
- Central Harlem has also potential where closes to Morningside Heights area.
- It can be done more detailed analysis by adding other factors such as transportation, demographics of inhabitants.

Finally, FourSquare proved to be a good source of data but frustrating at times. Despite having a Developer account I regularly exceeded my hourly limit locking me out for the day.

6. Conclusion

Although all of the goals of this project were met there is definitely room for further improvement and development as noted below. However, the goals of the project were met and, with some more work, could easily be devleoped into a fully phledged application that could support the opening a business idea in an unknown location.

As per the neighborhood or restaurant type mentioned like Thai restaurants analysis can be checked. A venue with lowest risk and competition can be identified.