1. Introduction and Business Problem

New York City is one of populous cities in United States where it provides a lot of business opportunities. So far this city has attracted numerous people into the area and is the centre for banking, finance, technology, tourism, etc. With the currency status of the market, it is a competitive market and it might be a challenge to start a proper business that is worth investing, therefore a careful analysis is required before deciding whether a business is worth invested and this will help reduce the risk of business failure and potentially having a decent amount of return on investment.

Business Problem

New York City is a multicultural place where different kinds of dishes can be found especially for western cuisine, and there seems to be a gap in the culinary industry where there might be a potential. Indonesian food is one of Asian food that has not been marketed really well compare to other Asian food such as Japanese, Korean, Chinese, Thai, etc. Starting a Indonesian food restaurant can be a great start to market another type of Asian food that might attract a certain group of people in New York, therefore opening the restaurant in a place where Asian restaurants are available should give enough attention for the restaurant to grow, with the purpose of attracting existing Asian cuisine lover to try out more of a similar delicacy.

2. Data

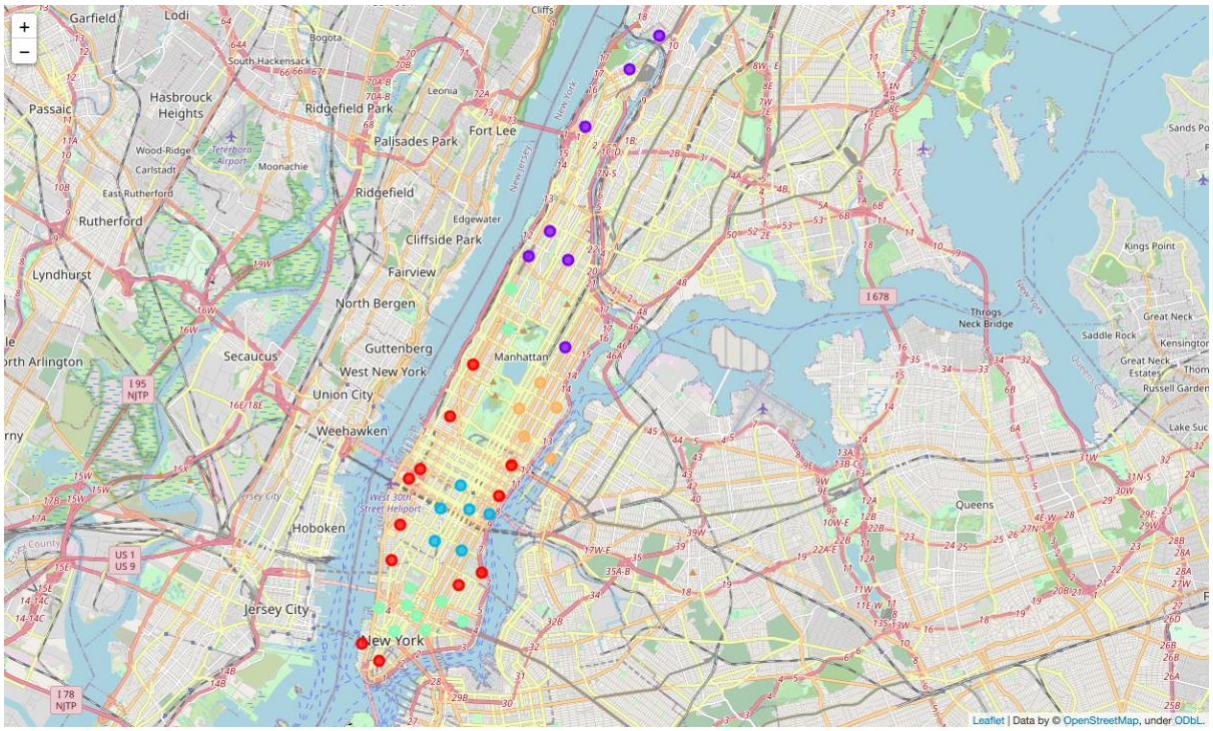
Neighbourhood has a total of 5 boroughs and 306 neighbourhoods. In order to segment the neighbourhoods and explore them, we will essentially need a dataset that contains the 5 boroughs and the neighbourhoods that exist in each borough as well as the latitude and longitude coordinates of each neighbourhood. This dataset exists for free on the web. Link to the dataset is: https://geo.nyu.edu/catalog/nyu_2451_34572. In addition, Sushi category Id 4bf58dd8d48988d142941735 is used for retrieving data from Foursquare API.

3. Methodology

In this project, I will use the basic methodology as taught in Week 3 lab. First, I convert addresses into their equivalent latitude and longitude values. Then we will use the Foursquare API to explore neighbourhoods in Manhattan, New York. After that, explore function to get sushi restaurant categories in each neighbourhood

Asian restaurants in Manhattan

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



Cluster

4. Results

K-mean Cluster Using K-mean to clustering data area with most diverse of common venue.

Cluster 0 (Red)

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
12	Upper West Side	Sushi Restaurant	Chinese Restaurant	Japanese Restaurant	Thai Restaurant	Asian Restaurant
13	Lincoln Square	Thai Restaurant	Sushi Restaurant	Japanese Restaurant	Ramen Restaurant	Vietnamese Restaurant
14	Clinton	Thai Restaurant	Ramen Restaurant	Japanese Restaurant	Sushi Restaurant	Chinese Restaurant
17	Chelsea	Japanese Restaurant	Sushi Restaurant	Thai Restaurant	Chinese Restaurant	Ramen Restaurant
19	East Village	Japanese Restaurant	Korean Restaurant	Ramen Restaurant	Sushi Restaurant	Vietnamese Restaurant
24	West Village	Sushi Restaurant	Chinese Restaurant	Japanese Restaurant	Thai Restaurant	Noodle House
28	Battery Park City	Sushi Restaurant	Japanese Restaurant	Chinese Restaurant	Thai Restaurant	Korean Restaurant
29	Financial District	Sushi Restaurant	Japanese Restaurant	Chinese Restaurant	Thai Restaurant	Food Truck
34	Sutton Place	Japanese Restaurant	Chinese Restaurant	Sushi Restaurant	Asian Restaurant	Ramen Restaurant
35	Turtle Bay	Japanese Restaurant	Chinese Restaurant	Thai Restaurant	Ramen Restaurant	Sushi Restaurant
37	Stuyvesant Town	Sushi Restaurant	Ramen Restaurant	Korean Restaurant	Thai Restaurant	Japanese Restaurant
39	Hudson Yards	Chinese Restaurant	Thai Restaurant	Ramen Restaurant	Sushi Restaurant	Japanese Restaurant

Cluster 1 (Purple)

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Marble Hill	Chinese Restaurant	Asian Restaurant	Japanese Restaurant	Thai Restaurant	Korean Restaurant
2	Washington Heights	Chinese Restaurant	Asian Restaurant	Thai Restaurant	Sushi Restaurant	Ramen Restaurant
3	Inwood	Chinese Restaurant	Thai Restaurant	Sushi Restaurant	Asian Restaurant	Dumpling Restaurant
4	Hamilton Heights	Chinese Restaurant	Thai Restaurant	Sushi Restaurant	Japanese Restaurant	Cocktail Bar
5	Manhattanville	Chinese Restaurant	Japanese Restaurant	Sushi Restaurant	Thai Restaurant	Japanese Curry Restaurant
6	Central Harlem	Chinese Restaurant	Japanese Restaurant	Thai Restaurant	Sushi Restaurant	Asian Restaurant
7	East Harlem	Chinese Restaurant	Thai Restaurant	Asian Restaurant	Sushi Restaurant	Hawaiian Restaurant

Cluster 2 (Blue)

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
15	Midtown	Korean Restaurant	Japanese Restaurant	Sushi Restaurant	Ramen Restaurant	Thai Restaurant
16	Murray Hill	Korean Restaurant	Japanese Restaurant	Sushi Restaurant	Chinese Restaurant	Ramen Restaurant
27	Gramercy	Korean Restaurant	Japanese Restaurant	Ramen Restaurant	Chinese Restaurant	Asian Restaurant
33	Midtown South	Korean Restaurant	Japanese Restaurant	Sushi Restaurant	Asian Restaurant	Ramen Restaurant
36	Tudor City	Korean Restaurant	Japanese Restaurant	Asian Restaurant	Sushi Restaurant	Ramen Restaurant
38	Flatiron	Korean Restaurant	Japanese Restaurant	Sushi Restaurant	Thai Restaurant	Chinese Restaurant

Cluster 3 (Green)

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
1	Chinatown	Chinese Restaurant	Sushi Restaurant	Vietnamese Restaurant	Dim Sum Restaurant	Asian Restaurant
18	Greenwich Village	Sushi Restaurant	Japanese Restaurant	Thai Restaurant	Vietnamese Restaurant	Chinese Restaurant
20	Lower East Side	Korean Restaurant	Chinese Restaurant	Vietnamese Restaurant	Japanese Restaurant	Sushi Restaurant
21	Tribeca	Chinese Restaurant	Sushi Restaurant	Vietnamese Restaurant	Japanese Restaurant	Asian Restaurant
22	Little Italy	Chinese Restaurant	Sushi Restaurant	Asian Restaurant	Dim Sum Restaurant	Thai Restaurant
23	Soho	Sushi Restaurant	Chinese Restaurant	Asian Restaurant	Thai Restaurant	Vietnamese Restaurant
25	Manhattan Valley	Chinese Restaurant	Thai Restaurant	Sushi Restaurant	Asian Restaurant	Noodle House
26	Morningside Heights	Chinese Restaurant	Thai Restaurant	Asian Restaurant	Sushi Restaurant	Korean Restaurant
31	Noho	Sushi Restaurant	Chinese Restaurant	Vietnamese Restaurant	Korean Restaurant	Thai Restaurant
32	Civic Center	Chinese Restaurant	Asian Restaurant	Vietnamese Restaurant	Noodle House	Cha Chaan Teng

Cluster 4 (Yellow)

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
8	Upper East Side	Sushi Restaurant	Thai Restaurant	Chinese Restaurant	Japanese Restaurant	Vietnamese Restaurant
9	Yorkville	Sushi Restaurant	Chinese Restaurant	Thai Restaurant	Japanese Restaurant	Ramen Restaurant
10	Lenox Hill	Sushi Restaurant	Thai Restaurant	Japanese Restaurant	Chinese Restaurant	Asian Restaurant
11	Roosevelt Island	Sushi Restaurant	Chinese Restaurant	Japanese Restaurant	Thai Restaurant	Asian Restaurant
30	Carnegie Hill	Sushi Restaurant	Thai Restaurant	Chinese Restaurant	Japanese Restaurant	Vietnamese Restaurant

5. Discussion

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 concentration of Korean and Japanese food in Cluster 2 which is indicating that the place is possibly well known for both kinds of foods, therefore it might be risky to be opening Indonesian restaurant there
- Cluster 1 and 3 might be better choice for opening the restaurant, therefore further analysis in those areas might be required

6. Conclusion

- Classified areas to avoid and areas that can be further explore for opening Indonesian restaurant.
- The accuracy of the analysis might be improved with more data provided or different methods are used.
- Risk of opening Indonesian restaurant in New York is still unknown as there might be other unknown variable taking place such as general population food preference or population of residence living in chosen place.