

1. Introduction & Business Problem :

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 analysed 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 excelllent cuisine. It's food culture includes an array of international cuisines influenced by the city's immigrant history. Sushi 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 sushi 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 fish, delivered daily from Japan, focus on neighborhoods and outlets that already attract a sophisticated Japanese client. 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.

My client 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 sushi bars 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 sushi bar should be open in an area that inadequate neighborhood in this way the bar can attract more customers. Therefore, this analysis necessary to ensure that we have enough customers and that we are not so close to other sushi places.

2. Data

Data 1: Neighborhood 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 the latitude and logitude 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

Data2: Newyork city 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, Sushi category Id 4bf58dd8d48988d1d2941735 is used for retrieving data from Foursquare API.

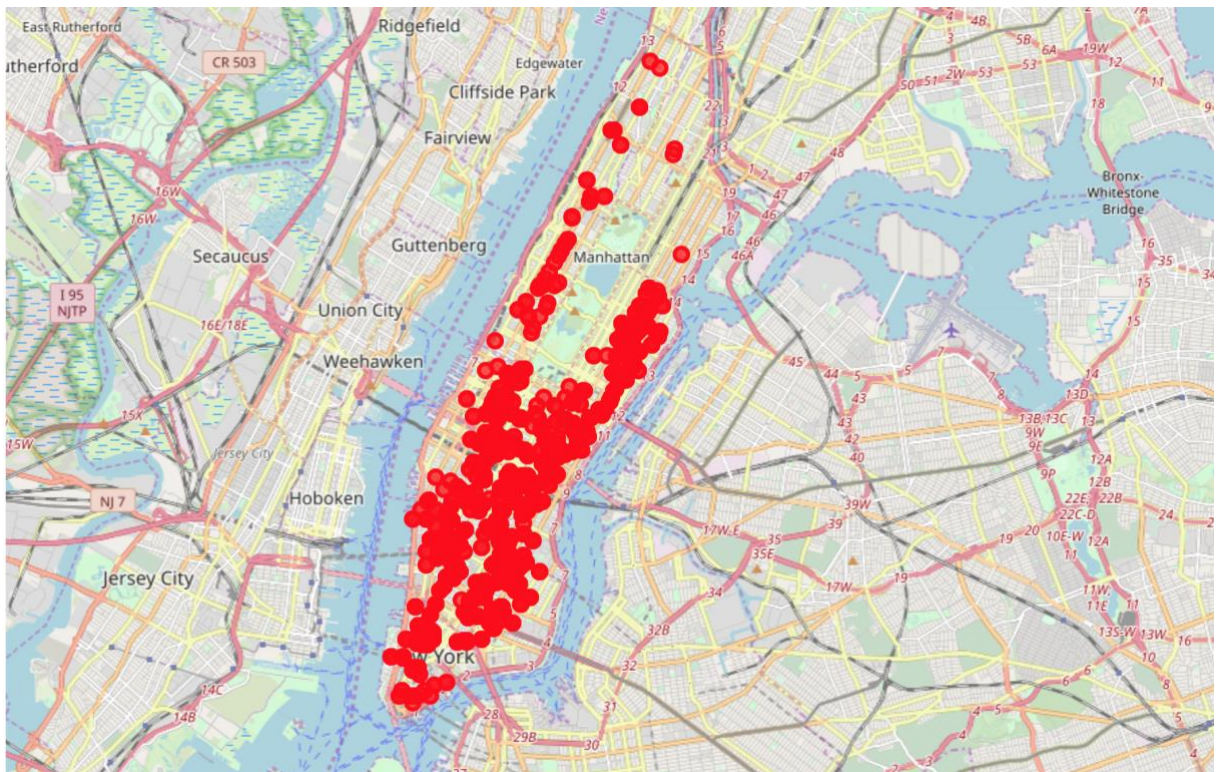
3. Methodology

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

	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 get sushi restaurant categories in each neighborhood.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.910660	Planet Tokyo	40.886233	-73.909479	Sushi Restaurant
1	Chinatown	40.715618	-73.994279	Shinsen	40.715608	-73.996611	Japanese Restaurant
2	Chinatown	40.715618	-73.994279	Sushumai Asian Fusion	40.721155	-73.987337	Sushi Restaurant
3	Chinatown	40.715618	-73.994279	Nakaji	40.715791	-73.996855	Sushi Restaurant
4	Chinatown	40.715618	-73.994279	Shōji	40.717582	-74.005592	Sushi Restaurant

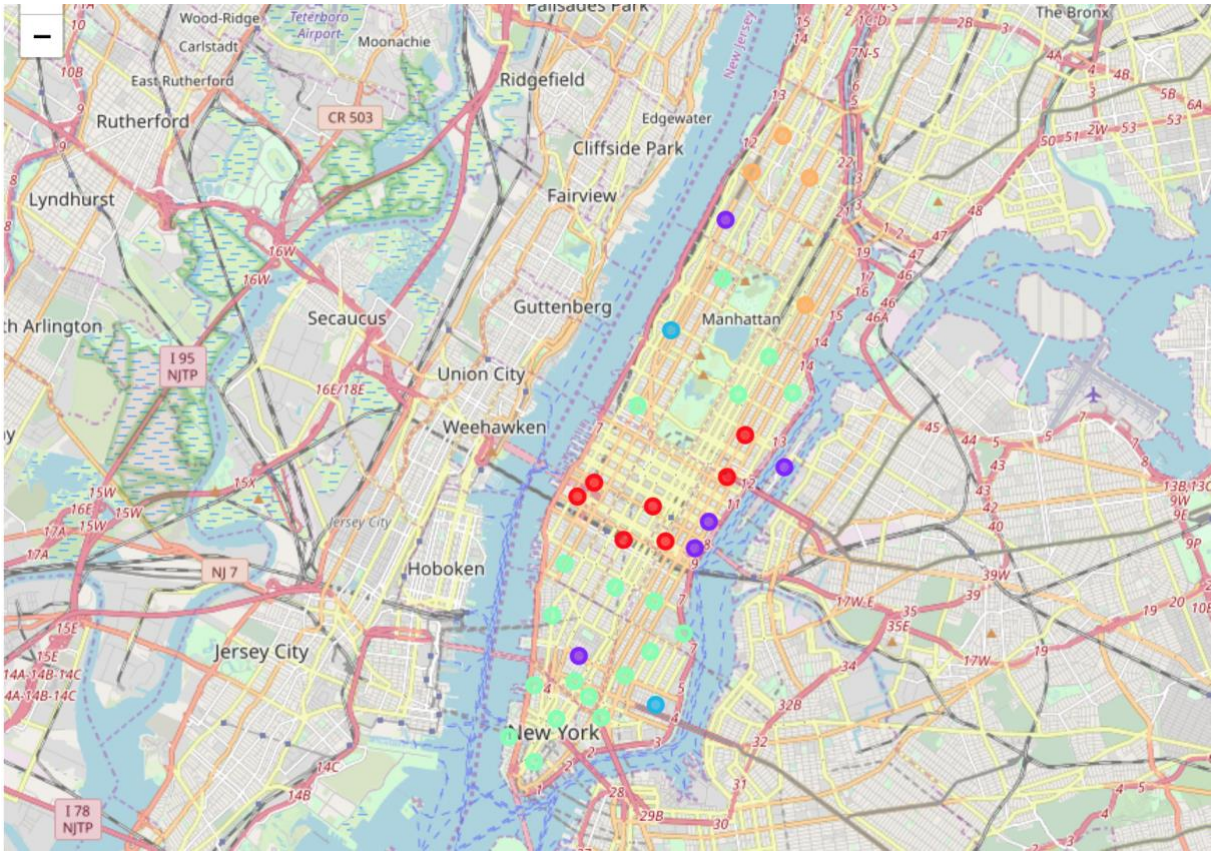


Sushi bars in Manhattan

[illegible]

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	Sushi Restaurant	Japanese Restaurant	Noodle House	Theme Restaurant	Vegetarian / Vegan Restaurant	Bakery	Chinese Restaurant	Cocktail Bar	Deli / Bodega	Grocery Store
1	Carnegie Hill	Sushi Restaurant	Japanese Restaurant	Asian Restaurant	Chinese Restaurant	Noodle House	Bakery	Cocktail Bar	Deli / Bodega	Grocery Store	Hawaiian Restaurant
2	Central Harlem	Sushi Restaurant	Vegetarian / Vegan Restaurant	Noodle House	Bakery	Chinese Restaurant	Cocktail Bar	Deli / Bodega	Grocery Store	Hawaiian Restaurant	Japanese Restaurant
3	Chelsea	Sushi Restaurant	Japanese Restaurant	Asian Restaurant	Vegetarian / Vegan Restaurant	Smoothie Shop	Seafood Restaurant	Sandwich Place	Sake Bar	Restaurant	Ramen Restaurant
4	Chinatown	Sushi Restaurant	Japanese Restaurant	Vegetarian / Vegan Restaurant	Noodle House	Bakery	Chinese Restaurant	Cocktail Bar	Deli / Bodega	Grocery Store	Hawaiian Restaurant



4. Results

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

Cluster 0

	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	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Manhattan	Marble Hill	40.876551	-73.910660	4	Sushi Restaurant	Vegetarian / Vegan Restaurant	Noodle House	Bakery	Chinese Restaurant	Cocktail Bar	Deli / Bodega	Grocery Store	Hawaiian Restaurant	Japanese Restaurant
1	Manhattan	Chinatown	40.715618	-73.994279	3	Sushi Restaurant	Japanese Restaurant	Vegetarian / Vegan Restaurant	Noodle House	Bakery	Chinese Restaurant	Cocktail Bar	Deli / Bodega	Grocery Store	Hawaiian Restaurant
2	Manhattan	Washington Heights	40.851903	-73.936900	4	Sushi Restaurant	Vegetarian / Vegan Restaurant	Noodle House	Bakery	Chinese Restaurant	Cocktail Bar	Deli / Bodega	Grocery Store	Hawaiian Restaurant	Japanese Restaurant
3	Manhattan	Inwood	40.867684	-73.921210	4	Sushi Restaurant	Vegetarian / Vegan Restaurant	Noodle House	Bakery	Chinese Restaurant	Cocktail Bar	Deli / Bodega	Grocery Store	Hawaiian Restaurant	Japanese Restaurant
4	Manhattan	Hamilton Heights	40.823604	-73.949688	4	Sushi Restaurant	Vegetarian / Vegan Restaurant	Noodle House	Bakery	Chinese Restaurant	Cocktail Bar	Deli / Bodega	Grocery Store	Hawaiian Restaurant	Japanese Restaurant

Cluster 1

```
manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 1, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[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
11	Roosevelt Island	Sushi Restaurant	Asian Restaurant	Noodle House	Bakery	Chinese Restaurant	Cocktail Bar	Deli / Bodega	Grocery Store	Hawaiian Restaurant	Japanese Restaurant
18	Greenwich Village	Sushi Restaurant	Japanese Restaurant	Sake Bar	Vegetarian / Vegan Restaurant	Noodle House	Bakery	Chinese Restaurant	Cocktail Bar	Deli / Bodega	Grocery Store
26	Morningside Heights	Sushi Restaurant	Hawaiian Restaurant	Vegetarian / Vegan Restaurant	Noodle House	Bakery	Chinese Restaurant	Cocktail Bar	Deli / Bodega	Grocery Store	Japanese Restaurant
35	Turtle Bay	Sushi Restaurant	Japanese Restaurant	Asian Restaurant	Steakhouse	Seafood Restaurant	Bakery	Chinese Restaurant	Cocktail Bar	Deli / Bodega	Grocery Store
36	Tudor City	Sushi Restaurant	Asian Restaurant	Japanese Restaurant	Vegetarian / Vegan Restaurant	Steakhouse	Smoothie Shop	Seafood Restaurant	Sandwich Place	Sake Bar	Restaurant

Cluster 2

```
manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 2, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[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
12	Upper West Side	Sushi Restaurant	Japanese Restaurant	Asian Restaurant	Grocery Store	Noodle House	Bakery	Chinese Restaurant	Cocktail Bar	Deli / Bodega	Hawaiian Restaurant
20	Lower East Side	Sushi Restaurant	Japanese Restaurant	Vegetarian / Vegan Restaurant	Noodle House	Bakery	Chinese Restaurant	Cocktail Bar	Deli / Bodega	Grocery Store	Hawaiian Restaurant

Based on data frame analysis above Cluster 2 Upper West Side and Lower East side 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.
- 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 developed into a fully fledged application that could support the opening a business idea in an unknown location.

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