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.

In this project we will try to find an optimal location for a restaurant. Specifically, this report will be targeted to stakeholders interested in opening an Japanese restaurant in New York.

Since there are lots of restaurants in Berlin we will try to detect locations that are not already crowded with restaurants. We are also particularly interested in areas with no Japanese restaurants in vicinity. We would also prefer locations as close to city center as possible, assuming that first two conditions are met.

We will use our data science powers to generate a few most promissing neighborhoods based on this criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

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.

1. **Data**

Based on definition of our problem, factors that will influence our decission are:

* number of existing restaurants in the neighborhood (any type of restaurant)
* number of and distance to Italian restaurants in the neighborhood, if any
* distance of neighborhood from city center

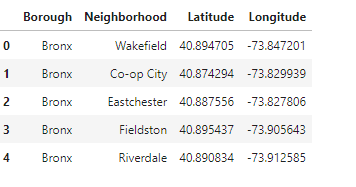
We decided to use regularly spaced grid of locations, centered around city center, to define our neighborhoods.

Following data sources will be needed to extract/generate the required information:

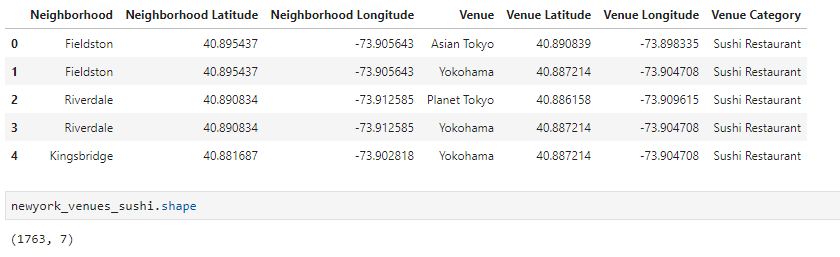
* Neighborhood has a total of 5 boroughs and 306 neighborhoods. In order to segement the neighborhoods and explore them, we will essentially need a dataset that contains the 5 boroughs and the neighborhoods that exist in each borough as well as the the latitude and logitude coordinates of each neighborhood. This dataset exists for free on the web. Link to the dataset is: <https://geo.nyu.edu/catalog/nyu_2451_34572>
* number of restaurants and their type and location in every neighborhood will be obtained using Foursquare API

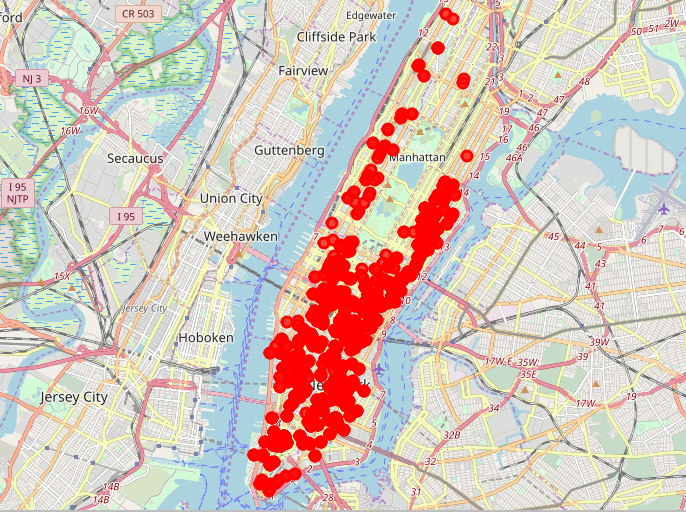
1. **Methodology**

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

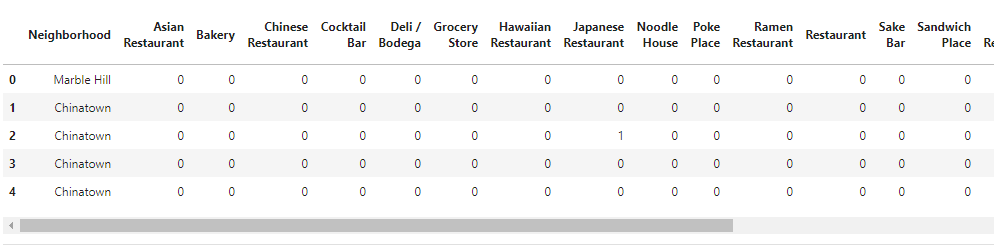


**Above, I have done convert addresses into their equivalent latitude and longitude values. Then we will use the Foursquare API to explore neighborhoods in Manhattan, New York. After that, explore function to get sushi restaurant categories in each neighborhood.**



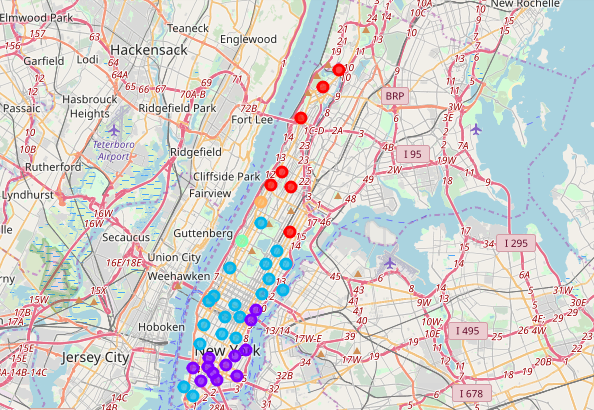


Sushi bars in Manhattan



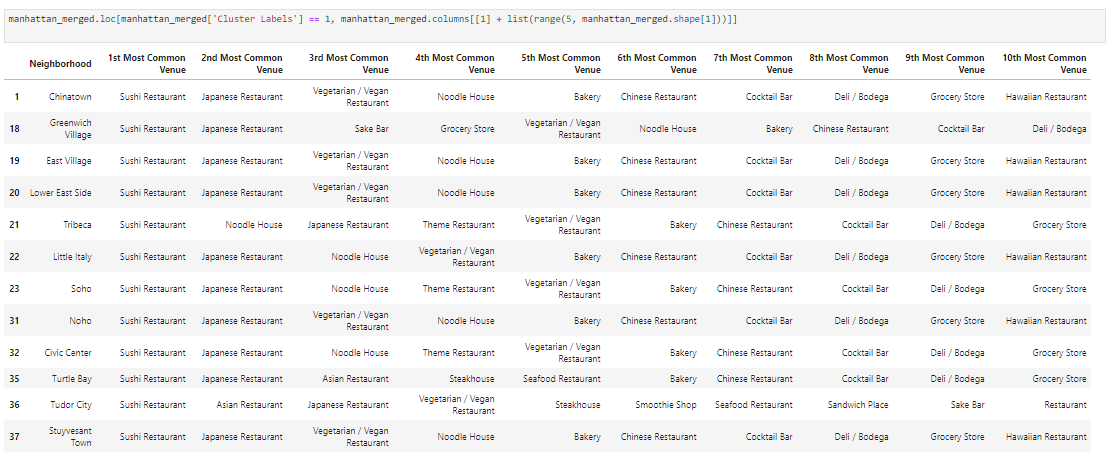
Then use this feature to group the neighborhoods into clusters K-means clustering algorithm will be use to complete this task. And also, the Folium library to visualize the neighborhoods in Manhattan and its emerging clusters.

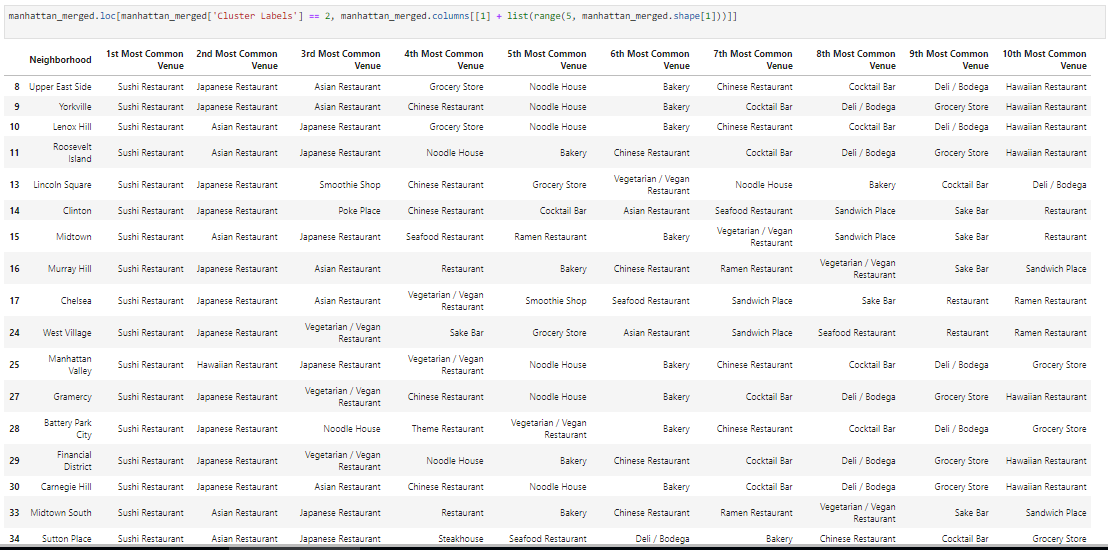
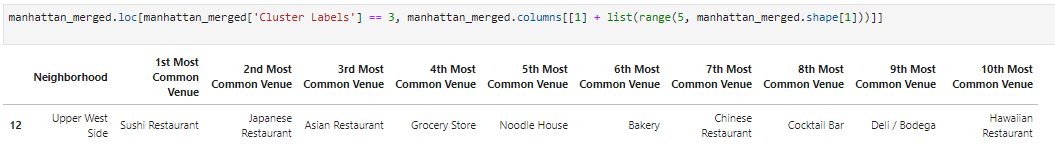


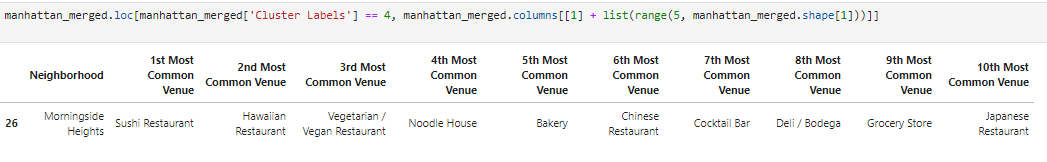


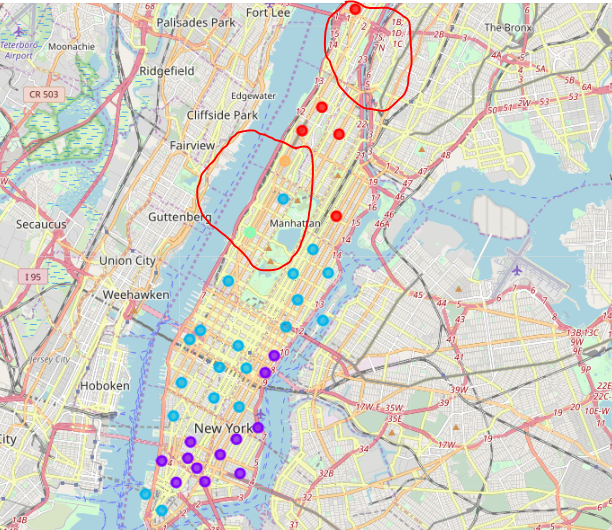
1. **Results**

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

**Cluster 0****Cluster 1**

**Cluster 2****Cluster 3****Cluster 4**





Based on dataframe analysis above Cluster 3 (Upper West Side ) and Cluster 4 (Morningside Heights) areas are the best places to open a new sushi bar business.

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. 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.

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

Final decission on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.