The Battle of Neighborhoods Report

**Introduction**

This is the IBM Data science Capstone project. In this project, I will address the problem of a client who wants to open a restaurant in New York. The Clint who is an entrepreneur have lots of questions regarding his/her restaurant so have decided to consult a Data scientist to answer his/her questions based on the market data available. The main purpose of this project is to find the appropriate location to open a restaurant. So I am designing this project to help my client to find the most suitable location.

**Business problem**

A restaurant is a business which prepares and serves food and drink to customers in return for money, The City of New York is famous for its excellent cuisine. Its 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.

Opening a restaurant is very challenging, hence the main focus is given to neighborhoods and other outlets which are already operational.

**Target Audience:**

My Client who wants to open his business in Manhattan area, so I will only focus on that during my analysis. The objective is to locate and recommend to the management which neighborhood of New york city will be best choice to start a restaurant. The Management also expects to understand the rationale of the recommendations made.

**Data Requirement and Collection**

To solve the client's problem, we will need below data:

● List of neighborhoods in New york.

● Latitude and Longitude of these neighborhoods

● Venue data related to already existing sushi restaurants.

This will help us find the neighborhoods that are more suitable to open an sushi Restaurant.

**EXTRACTING THE DATA**

● The New York neighborhoods data is available at the given link.

https://geo.nyu.edu/catalog/nyu\_2451\_34572

● Using Foursquare API to get venue data related to these neighborhoods.

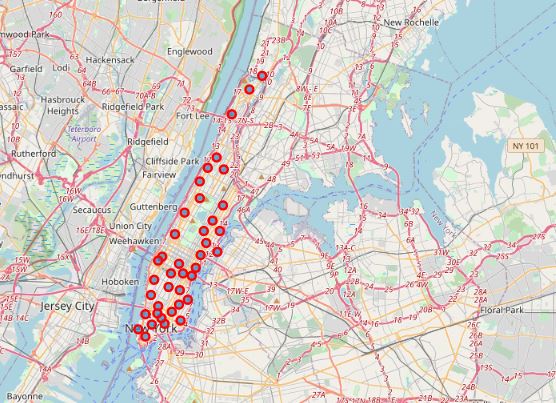
**Methodology**

For completing this project I will use the methodology taught us in week3.

The main and foremost step of completing any project is to understand the business requirement and what is the end result they are expecting from us.

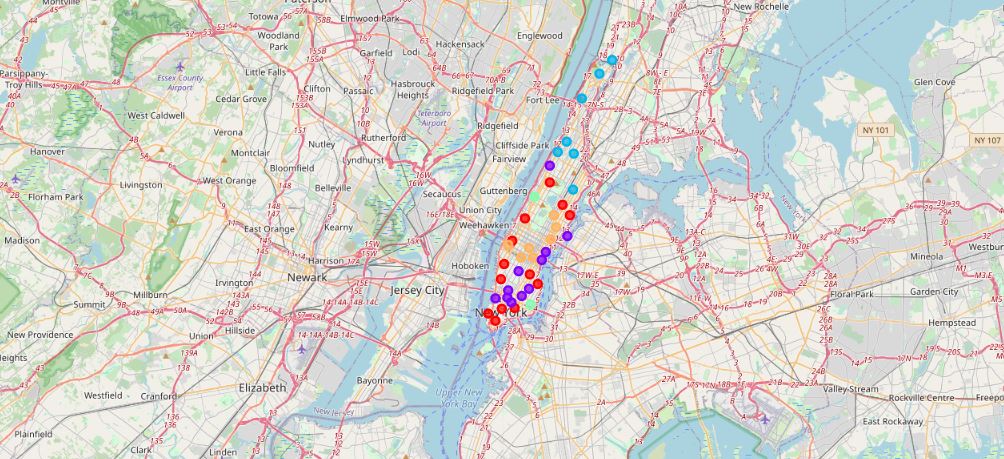
For the Analytic Approach

I have used Four Square API to know about the specific location that our client asked about us.



These are the sushi bars in Manhattan that are already up and running

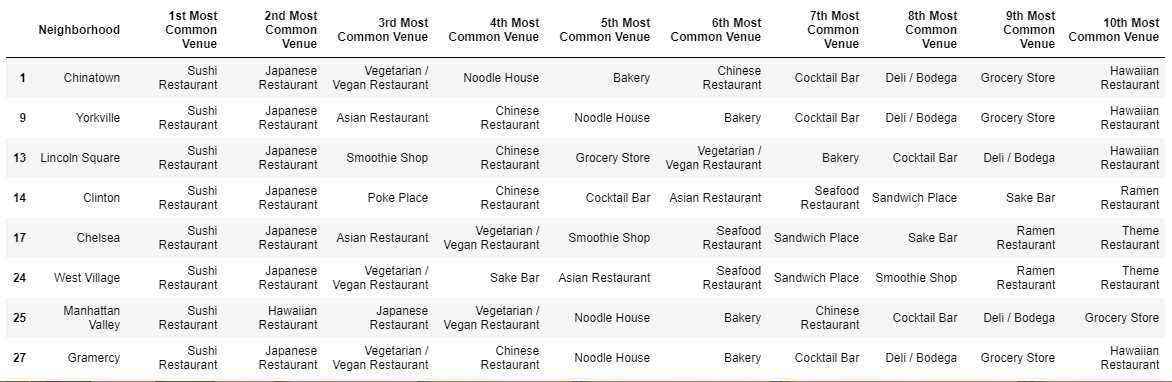
Then I am grouping the neighborhoods into clusters using K-means clustering algorithm. And also, the Folium library to visualize the neighborhoods in Manhattan and its emerging clusters.



**Results:-**

Using K-mean to cluster the data area with less number of sushi bars

**Cluster 0:-**

****

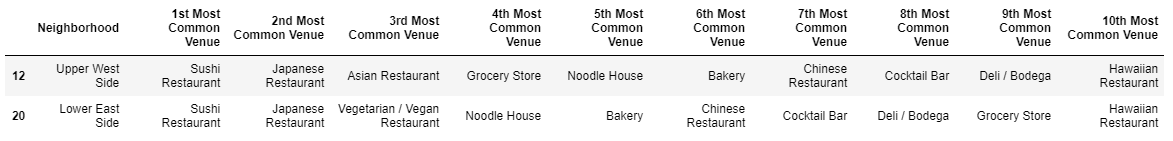
**Cluster 1:-**

****

**Cluster 2:-**

****

**Cluster 3:-**

****

Based on data frame analysis above Cluster 3 Upper West Side and Lower East Side areas are the best places to open a new sushi bar business.

**Discussion:-**

Here are the some observations that I noted while working on this project.

* Midtown and sohoso providesn a high risk competition for the new owners.
* Central Harlem has also potential for a business to grow.
* This project can be improved with adding more variables into the consideration as transportation, government rules and regulations, Availability of workers and raw materials.
* FourSquare proved to be a good source of data.

**Conclusion**

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.

Although all of the goals of this project were met but scope of expansion is still there by including more variables that can affect the restaurants performance.