**Data science Capstone project**

**The Battle of Neighborhoods**

**Author: Viral Gorecha**

1. **INTRODUCTION:**

Welcome to the world's most exciting city. The city of New-York. The city that doesn't need introduction. The city with a diversity of people and culture spread throughout. The city that never sleeps. New-York is the financial capital of the United States of America - the biggest superpower in the world and hence one of the most expensive cities for small startups. The analysis of which neighborhoods are best for startup of a certain type of business will result in reduced costs and higher return on investment to any new entrepreneur.

**Problem:**

In the city of dreams the immigration from various countries has been so high that it has influenced in building its diverse culture hugely. The city is full of all kinds of restaurants and Sushi joints are no different. They can be found on every nook and corner of the city. The problem that prevails here is to find a neighborhood that makes you stand out in your business and give the client a definite success

**Find a suitable location:**

Finding a location depends on how huge or small a restaurant is being planned. if the target audience is a high-profile client and a large investment is being made then a sophisticated neighborhood can be selected. On the other hand, if the capital allows on a small joint then a crowded neighborhood with preferably a shopping area around the corner is the best choice for business.

**Target Audience:**

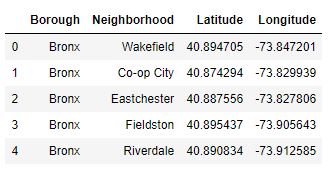
Here we are looking for a client who wants to set up a restaurant i.e. a Sushi joint in Manhattan so we shall analyses that borough for neighborhoods suitable for his business.

1. **Data:**

**Data :**

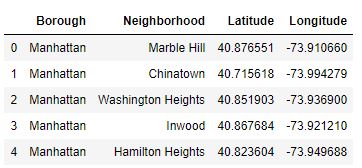
1. To find the best location for our Japanese cuisine joint, we will use the following sources of information:
2. We shall be using New-York city as our playground for this particular project. We need the dataset containing longitudes and latitudes of each neighborhood in all of the 5 boroughs of New-York City. We shall use the following free dataset for that matter: <https://geo.nyu.edu/catalog/nyu_2451_34572>

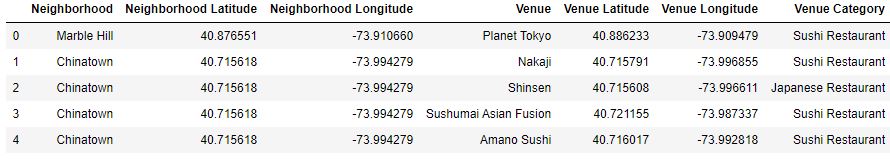
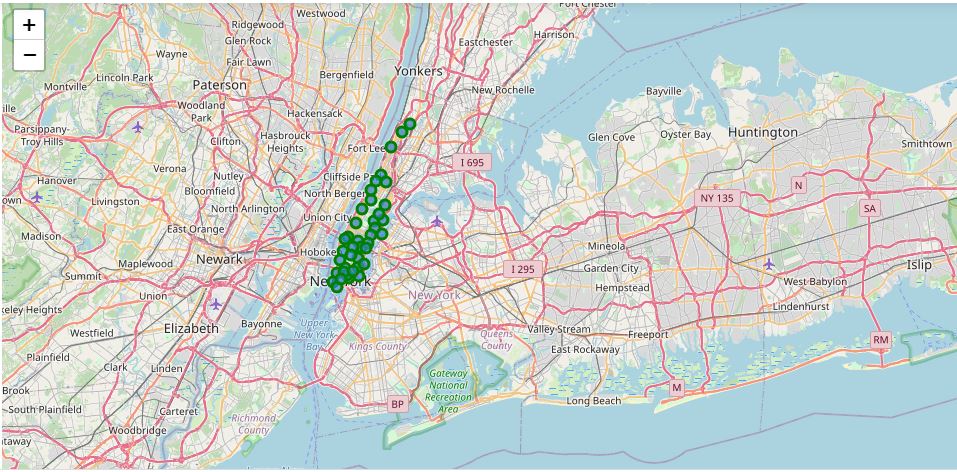
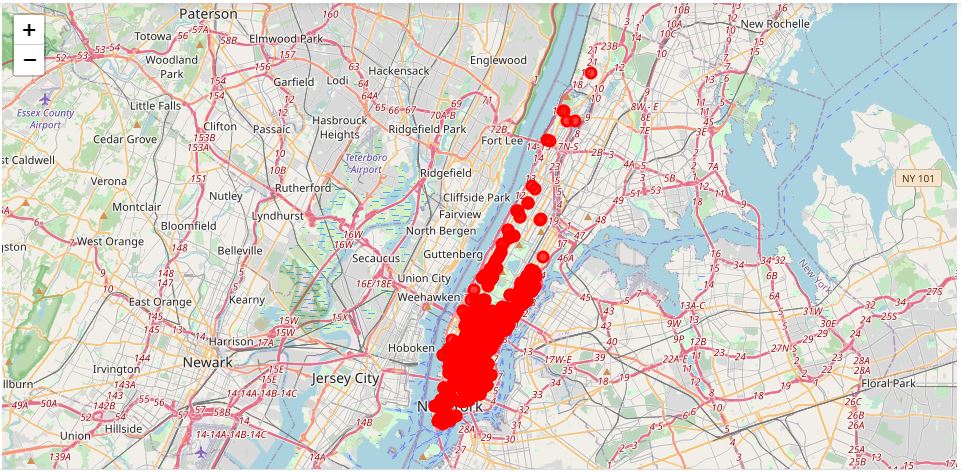
From Foursquare Venues Categories - https://developer.foursquare.com/docs/resources/categories Sushi category Id - 4bf58dd8d48988d1d2941735



1. Another data will be the data of longitudes and latitudes of neighborhoods fetched using the foursquare API

**3. Methodology:**

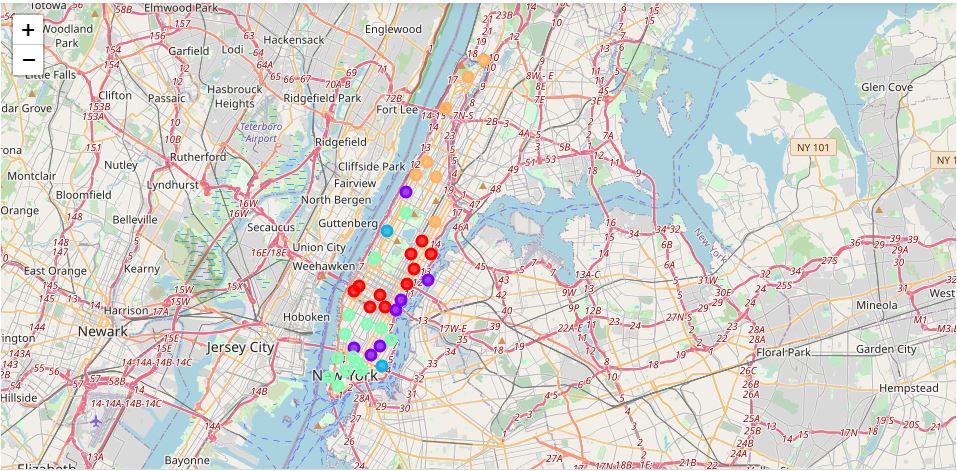
1. Converting the addresses in data into their respective longitudes and latitudes
2. Get neighborhoods of Manhattan using Foursquare API calls
3. Get a function to get neighborhoods with Sushi joints

Sushi joints in Manhattan

**A total view of sushi joints in Manhattan**

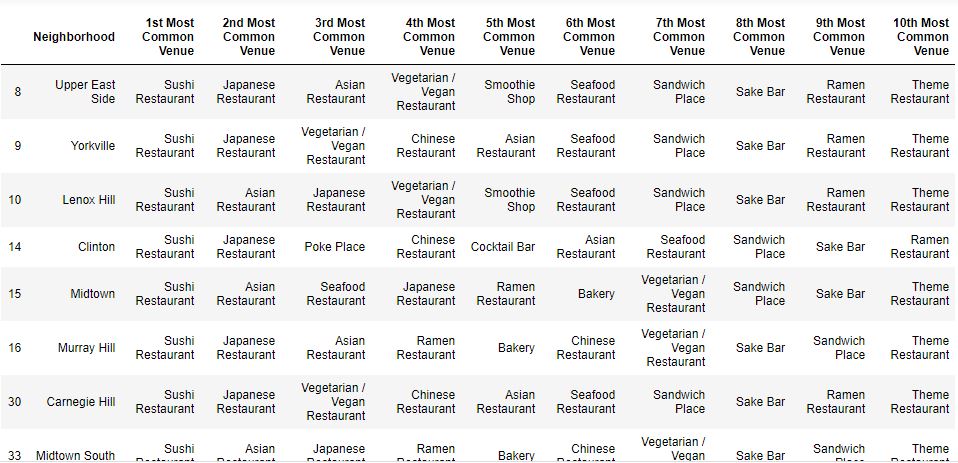
The next task is to perform clustering for which K-means clustering is used as we have the labelled data available making it easier with K-means clustering.



1. **Results:**

Lets move to the next page and get the results of the clusters obtained by K-means clustering

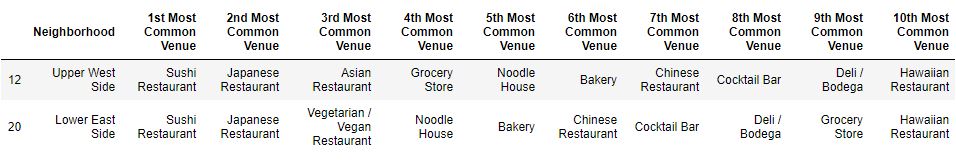
**Cluster 0:**

****

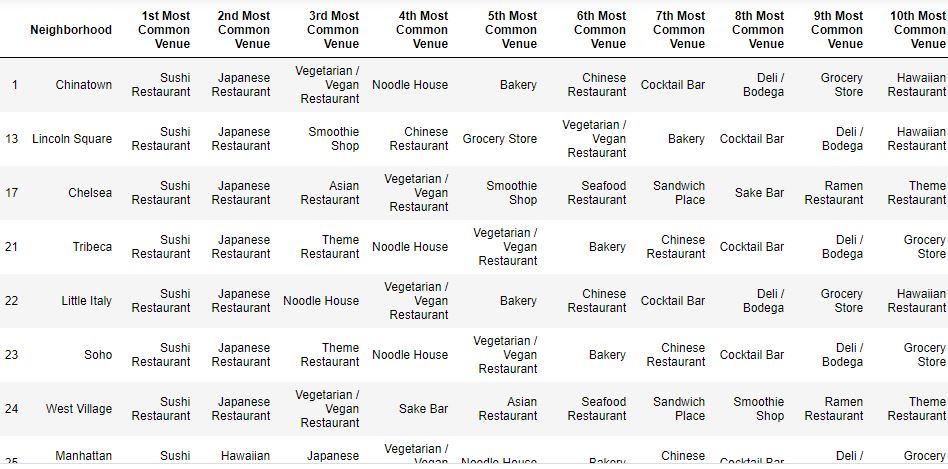
**Cluster 1:**

****

**Cluster 2:**

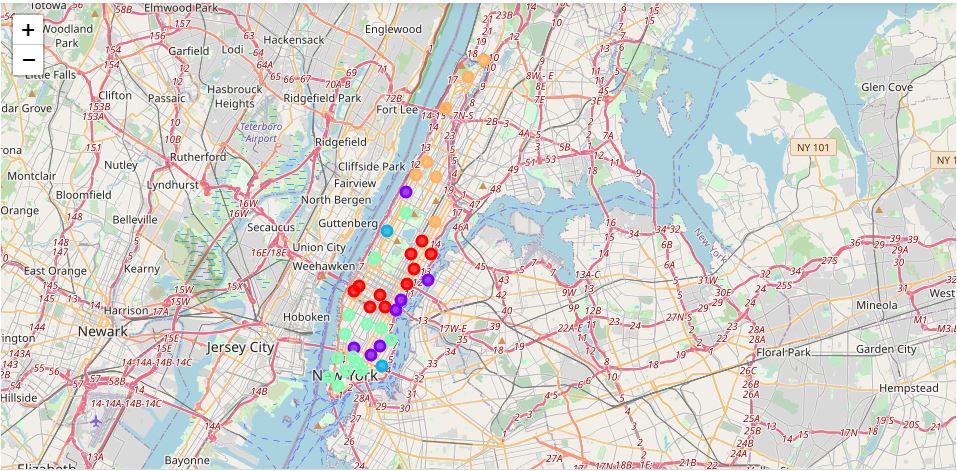
****

**Cluster 3:**

****

**Cluster 4:**

****



Based on the above clusters, Cluster 2 i.e. Upper West side and Lower East side are the best areas to start our Sushi Joint.

1. **Discussion:**

A rather wide and extensive data availability would make this project worth taking into consideration while starting up a new restaurant or a sushi joint in Manhattan and possibly New-York

1. **Conclusion:**

As discussed above, the Upper West side and Lower East side are the best areas for opening up a Sushi Bar as they have the least competition and crème customer availability. This project can be further Detailed into client requirements and can be processed as a full fledged AI application for data analysis and prediction using ML on Business startups in New-York

Reference: The idea for this project has been referenced from Mr.Ezgi Kaysikesler Github repository.