**REPORT**

**The Battle of Neighborhoods**

1. **Introduction**
   1. **Background**

The City of New York is the most crowded and popular US. It is a harmony of cultures and cuisines. Opening a new business can be a good opportunity in New York. However, it is also risky because of the competitive market and high operating costs. Making right choices take the most important factor. Selecting the right place for the business is going to be helpful to increase chances of being successful.

* 1. **Problem**

An investor wants to open a new business in Manhattan, the budget for this business is limited. This business will focus on restaurant supplies and food for Mediterranean restaurants. I will help my client to analyze the provided data and help him to find the best place for this business. Planning before investment will be crucial. This place should be close to most of the target restaurants.

* 1. **Target Audience**

This project will be helpful anyone interested in opening restaurant business in Manhattan.

1. **Data**

New York 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 latitude and longitude coordinates of each neighborhood. Luckily, this dataset exists for free on the web. Here is the link to the dataset: <https://geo.nyu.edu/catalog/nyu_2451_34572>

To analyze venue categories and their locations:

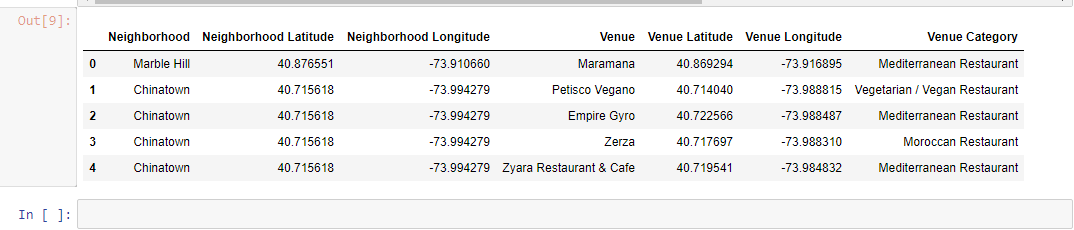
Foursquare developer portal

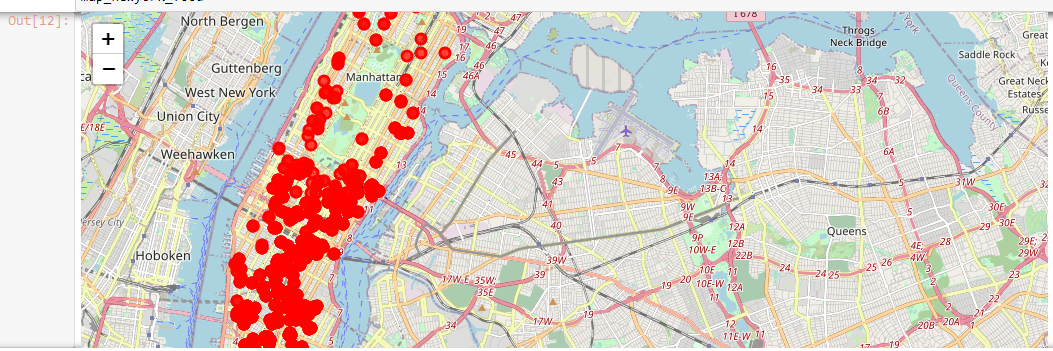
Venues Categories: <https://developer.foursquare.com/docs/resources/categories>

Mediterranean Restaurant id:4bf58dd8d48988d1c0941735

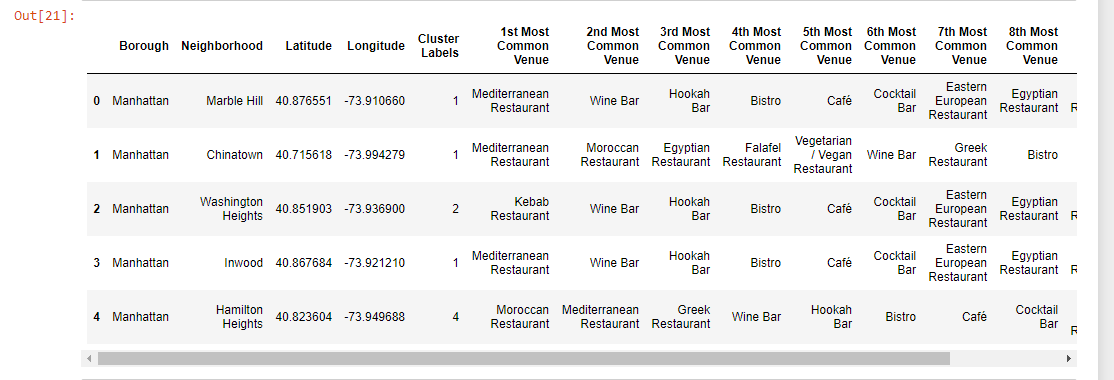
1. **Methodology**

I used pandas to generate data frames from New York Json file. Using the generated data frames and Foursquare developer API I was able to find venues in each neighborhood.

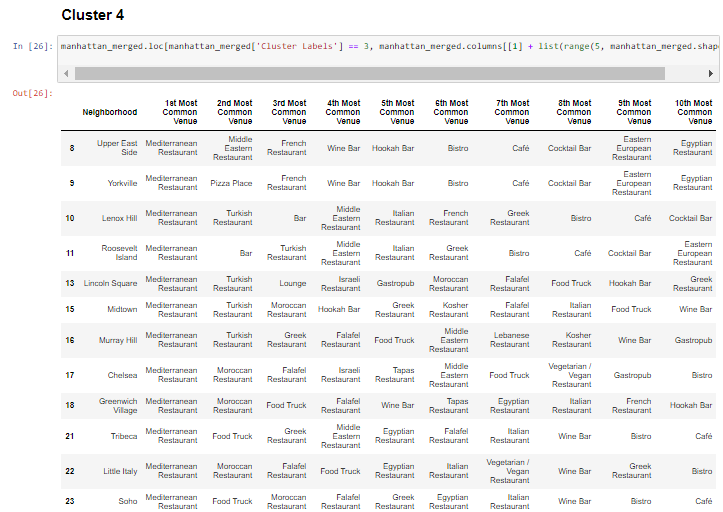


 I used Folium library to visualize the geographic coordinate of the neighborhoods.

I used Clustering (k-means clustering) was used to find the clusters of similar neighborhoods.

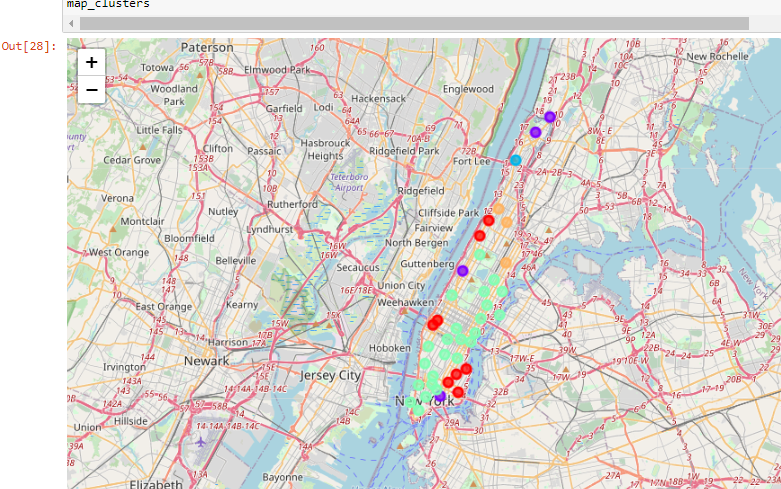
1. **Result**

Neighborhoods with the higher number of Mediterranean restaurants





Clusters on the map

 Cluster 1 Cluster 2 Cluster 3 Cluster 4 Cluster 5

1. **Discussion**

5 clusters found.

Potential clusters suitable for the business are Cluster 4 and Cluster 1.

1. **Conclusion**

There are many clusters with the target restaurants, among these clusters locating the business in the area of cluster4 will be more reasonable