**Capstone Project - The Battle of Neighborhoods Research Report**

**Author: Yan Haixiang**

***Introduction/Business Problem***

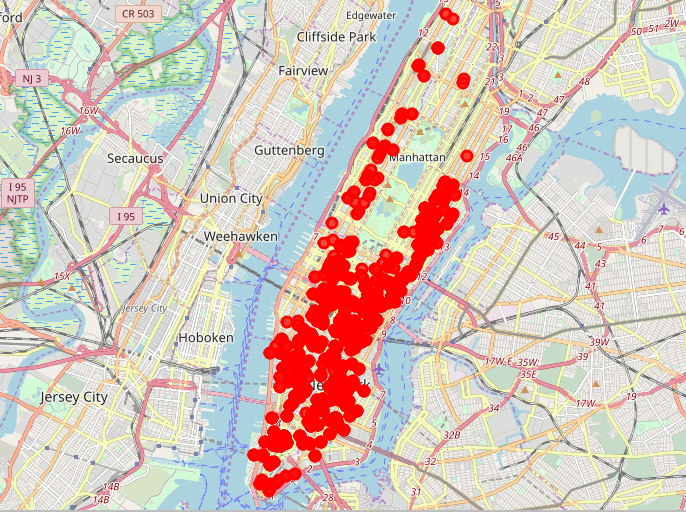
* The City of New York is famous for its excelllent cuisine. It's food culture includes an array of international cuisines influenced by the city's immigrant history.
* Sushi restaurants have become so popular in the United States now it seems that there is one on every corner, not only in major cities 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.
* 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.

***Data Selection***

* To identify the characteristics of our competitors' venues in Manhattan, we would first need to find out the number of sushi bars in Manhattan currently and their location.
* We then used Google Map API to find their geographic coordinates based on their postal code addresses.
* In Manhattan, there is 1763 sushi bars are currently operating.
* Next, we also used Google Map API to find their geographic coordinates of the 5 locations shortlisted for our sushi bar:

***Methodologies***

* Addresses are converted into their equivalent latitude and longitude values.
* Foursquare API is used to explore neighborhoods in Manhattan, New York.
* After that, explore function to get sushi restaurant categories in each neighborhood.



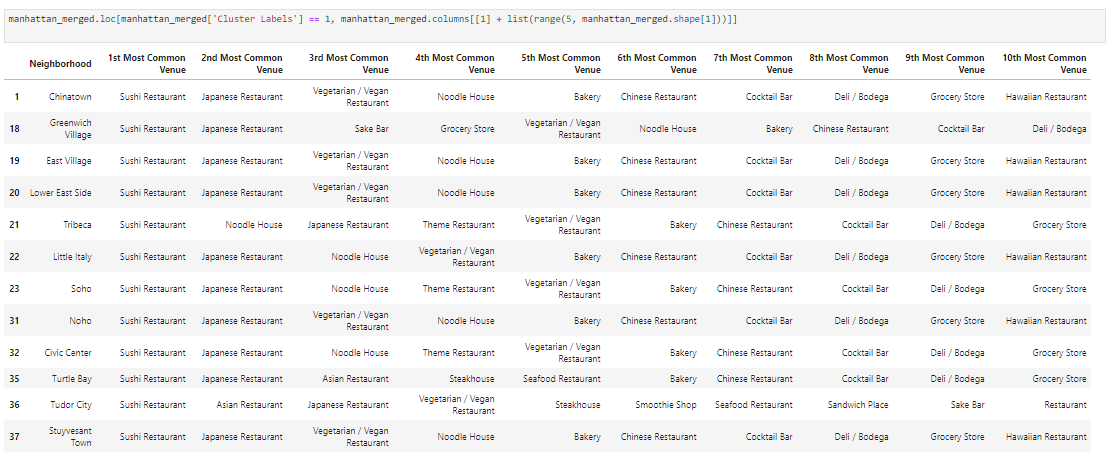
* Then using this feature to group the neighborhoods into clusters K-means clustering algorithm will be use to complete this task. And also, the Folium library to visualize the neighborhoods in Manhattan and its emerging clusters.

***Results***

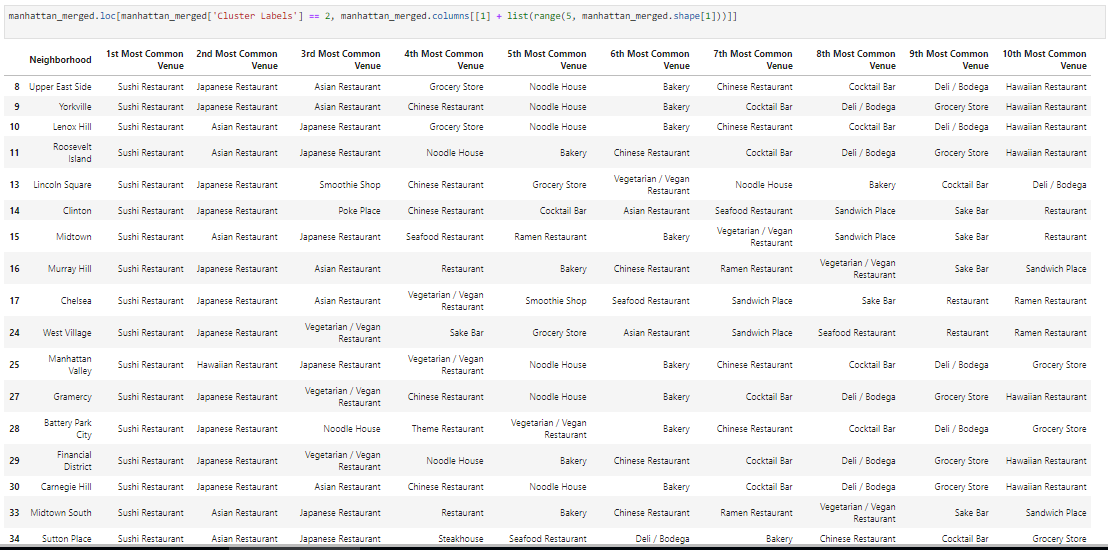
* Using K-mean to clustering data area with less number of sushi bars



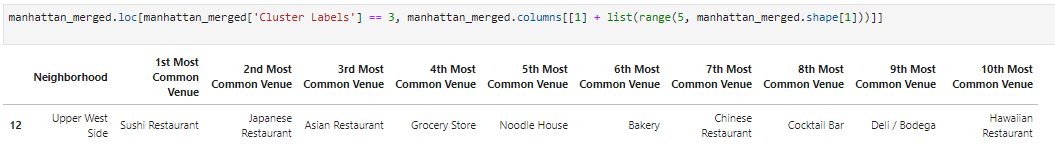
Cluster 0



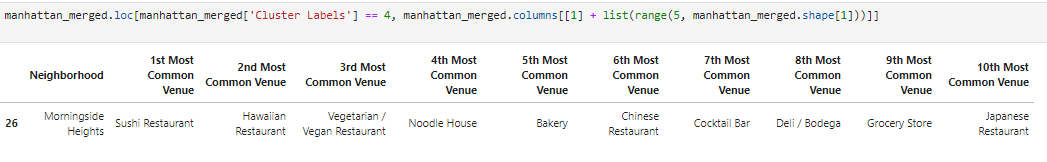
Cluster 1



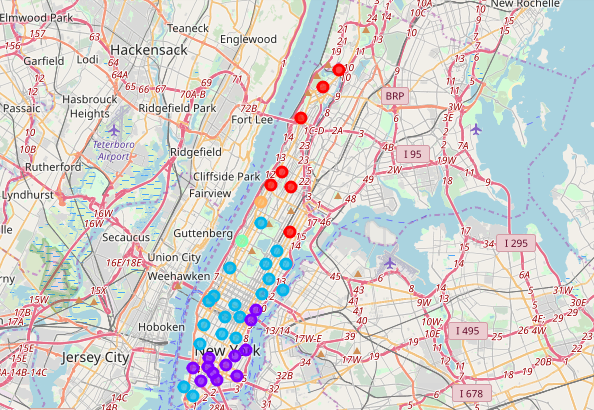
Cluster 2



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.

***Discussion***

* 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.

***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.

(End of Report)