**Capstone Project**

**Introduction Section – Business Problem**

We eventually know that New York comprises of 5 boroughs and in this project, we will explore BROOKLYN and attractions in the city.

Problem Description and Analysis –

Whenever a new person who enter the New York has lot of choices to enjoy at cafe’s, parks, hotels, clubs or some come for businesses. Below are some points –

1. It is difficult to find these places for a new person even with smartphone technology and maps.
2. It is often quite difficult for start-up to find the suitable location for itself for better business growth.
3. Stakeholders – Businessman, Tourists.

**Data Section-**

The data used are –

1. New York city data from the Json file available in the course. This file comprises of details of all boroughs with their corresponding Neighborhoods. Also, the file contains coordinates of the New York city data set locations.

Four Square API is being used to get venues details for neighborhoods in New York.

**Methodology Section –**

In this project New York data is taken as a base which contains info of all borough and their corresponding neighborhoods. Some exploratory data analysis applied –

1. From the data frame we will subset the dataset to get only the records for BROOKLYN,
2. Data cleaning with removal of all NaN values plus any blank fields.
3. Visualize all neighborhood of BROOKLYN using FOLIUM library.
4. Fetching Latitude and Longitude of neighborhoods using Geopy Library
5. Normalizing the data fetched from Foursqare API.

**Machine learning technique** –

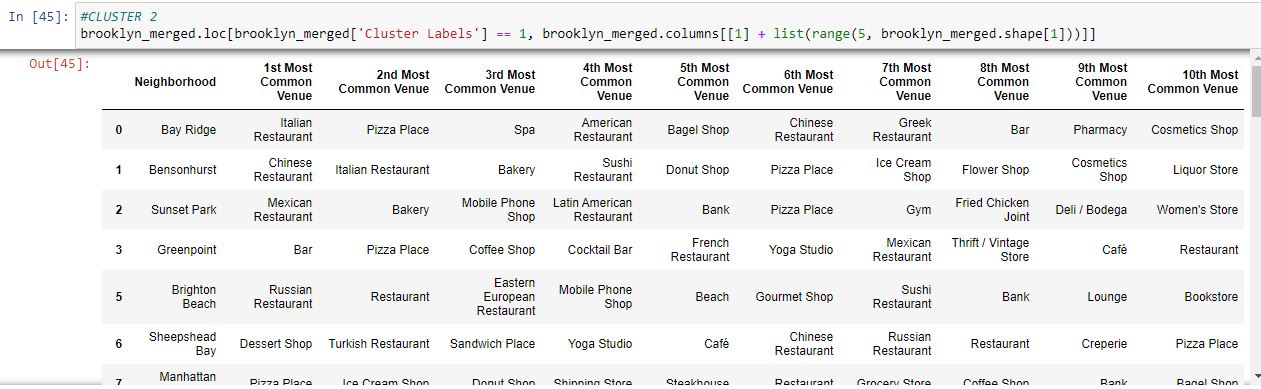
We will use Clustering technique of Unsupervised Learning to segment the neighborhoods and cluster them using K-MEANS clustering to get the clusters with similarities in their venues.

**Results Section** –

When we run this clustering on Brooklyn dataset we get the 5 clusters having neighborhood which are similar based on venues they have nearby. Below are top venues in each neighborhood.



**Sample of cluster formed-**



**Discussion Section –**

According to results we see that most common venues come out to be restaurants and Coffee shops, which means any new businessman can start restaurants provided that they need to compete with existing established chains but if he provides top class facilities to get to top.

**Conclusion –**

In this Project we have used Foursquare API with Machine learning techniques to provide the best results in segmenting the neighborhood according to their venues.