INTRODUCTION:

This is capstone project for IBM Data Science Professional Certificate where I as an investor would like to search for the locations to set up a restaurant in Manhattan, New York City. New York City is the city of dreams that houses highly diverse group of people globally. This is the reason that I want to dive directly here despite being the outsider Moreover, the reason I choose Manhattan is due to the fact that it is the most happening and lively borough within the perimeter of this fabulous city, New York City. Besides the location, I will also be looking into the demographics of restaurants to decide on the type that I should longed for.

BUSINESS PROBLEM:

The primary objective of this capstone project is to find the location to set up my restaurant based on the types of amenities that site offers making it suitable for the investment. This will further help me to define the nature of the restaurant that I would like to invest into digging up more information in other running eatery businesses there.

DATA:

I will be using following link to download data on neighborhoods in Manhattan, New York City from the link https://geo.nyu.edu/catalog/nyu_2451_34572. This gives a data on the geo positions of all the Neighborhoods of Manhattan Borough which later will be used generate and visualize the maps of the site.

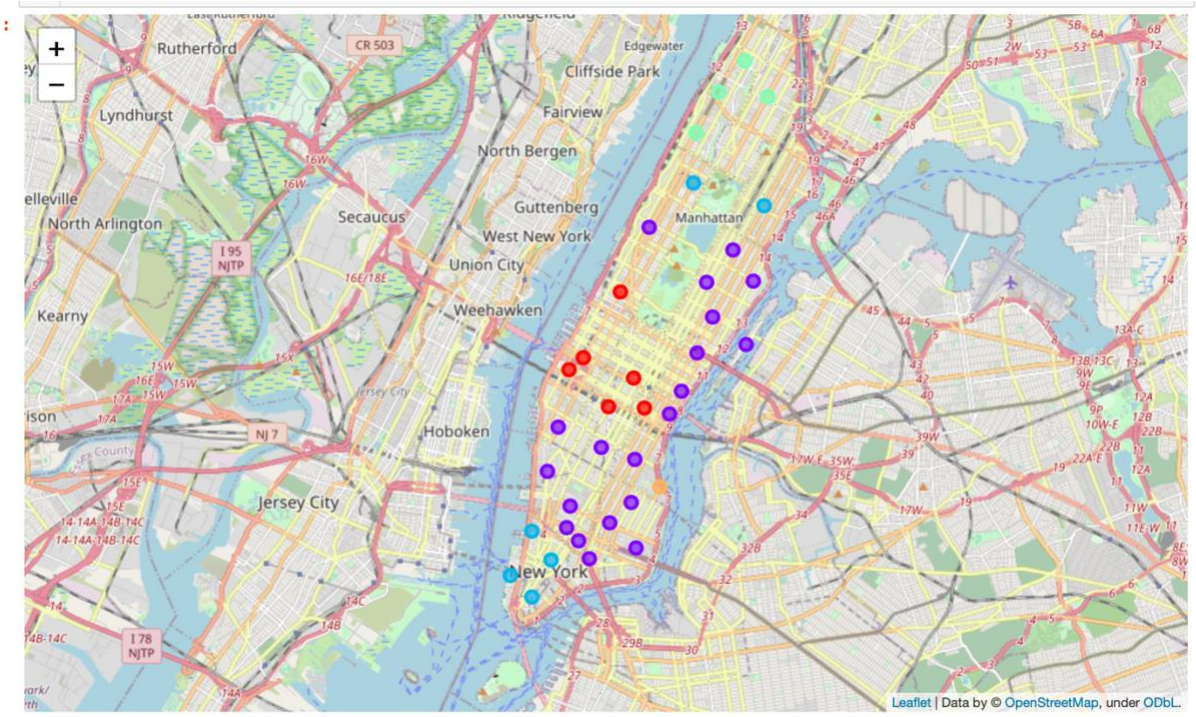
Further, I will be using Foursquare Restful API to scour more data on current amenities, types of running restaurant businesses in Manhattan and which in turns will be one of the parameters for me to cluster the Neighborhoods using K-Means Clustering and will be aid to decide on the area and type of restaurant to open within Manhattan.

METHODOLOGY:

I used the above mentioned link https://geo.nyu.edu/catalog/nyu_2451_34572 to download the geo position data on Neighborhoods of Manhattan Borough in New York City and in turn will be used to generate the maps of the site later using a folium library.

Then with the latitudes and longitudes of the Neighborhoods and use of restful API FourSquare I will generate the quired data on the types of restaurants in Manhattan and also other amenities present with the perimeter, the frequently visited places in the Neighborhood are later calculated based on which the clusters are created using K-Means Clustering method to divided the neighborhood into 5 clusters. This further simplifies my decision making on the prospective investment in Manhattan as a restaurant entrepreneur.

RESULT:



Cluster 1: red
Cluster 2: Purple
Cluster 3: Blue
Cluster 4: Parrot
Cluster 5: Orange

CONCLUSION:

Cluster '1' has the Neighborhoods with a very few or no restaurants than other clusters so it is the prime location of our restaurant opening and the type would be either Italian or Mexican restaurant based on the absence of this type in the cluster.

The most visited places in the clusters are mostly entertaining places like Theater, Park and we believe that most of the people would like to be there with the near and dear ones for longer period of time which in turn will increase the chances of their visit in our restaurant. Moreover, fast food outlets presence is high in number in the cluster and I believe this further improves our unique selling point and increases our competitiveness of serving the fresh and hot foods to the customers.