# INVESTMENT IN RESTAURANT IN MANHATTAN, NEW YORK

## INTRODUCTION:

This is a capstone project for IBM Data Science Professional Certificate where I as an investor would like to search for the good locations to set up a restaurant in Manhattan, New York City. New York City is the city of dreams that houses highly diverse groups of people from around the globe and this is the reason that I chose it despite being the complete outsider. Moreover, Manhattan is the most happening and lively Borough within the perimeter of this beautiful city.

To decide on my investment, I will be looking into the demographics on different types of restaurants currently available in Manhattan and, technically, I will be running K-means clustering to segment the Neighborhoods in Manhattan based on the most visited venues to ensure that the locations has higher inflows of people within the Borough.

I will use geo locations of Neighborhoods in Manhattan which is downloaded and linked via path to my machine. Folium will be used to map out the coordinates of the Neighborhoods and to show the clusters after running K-Means clustering. FourSquare API is used to get the data on categories of restaurants, most visited venues in the Neighborhoods which are the key to my investment decisions-making process.

## 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 type of the restaurant that I would like to invest in by digging more information about other current 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.

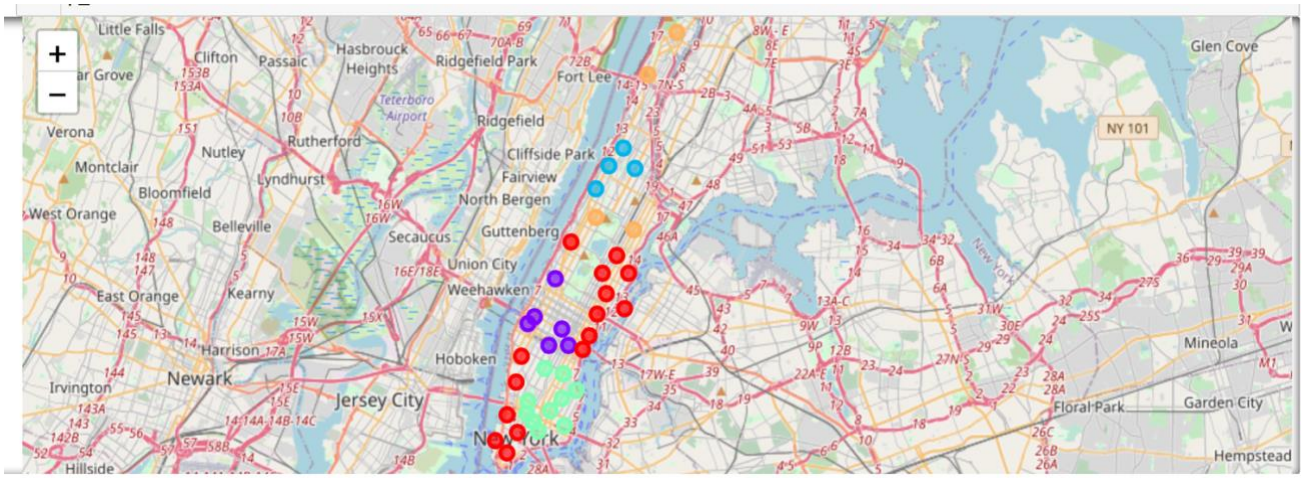
Further, I will be using Foursquare Restful API to scour more data on current amenities, types of running restaurant businesses in Manhattan and which in turns will be one of the parameters for me to cluster the Neighborhoods using K-Means Clustering and will be aid to decide on the area and type of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

## METHODOLOGY:

I have 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.

Next I need to get all the amenities situated in Manhattan and compare the category types of places based on the frequency of the visits. Such identification will help me to understand the demography of the people that influxes around in Manhattan and will definitely help me to select the nature of restaurant that I would like to start. Same feature will be used to segment the Neighborhoods using K-Means Clustering and I have created 5 clusters out of it. 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 '2' 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.