PROSPECTIVE AREA TO SET-UP GYM IN NEW YORK

- Karthika Kumaran

BACKGROUND

- Setting up a Business (GYM) in a new city is a challenging task
- Different factors are considered, such as prime location for the business, financial factors, commutation , market factors etc. for the longevity of the business
- This project serves as a stepping stone for those who are planning to set-up any new business(GYM) in a new city

BUSINESS PROBLEM

- To find the best neighborhood in New York to open Gym
- Gym must be located in an area which is close to prominent Metro Stations so that will be easy for the customers to commute
- Gym should be set up in a neighborhood with few competitors so as to reduce the risk of competition.

DATA

Considering New York as the city to set-up the gym:

- The dataset for New York is readily available on the web and the link is: https://geo.nyu.edu/catalog/nyu_2451_34572
- Geospatial coordinates of New York is obtained by using **geopy library**
- Retrieve all the metro stations in a radius of 15 kms around New York using **Foursquare API**
- Data for all the gyms located in a radius of 750 meters for every metro stations are obtained using **Fourquare API**

Fig: Map of New York

DATA (CONTD…)

- Using the collected data, following are calculated:
  - ➢ the number of existing gyms near each station
  - ➢ the minimum distance of a gym for every metro station
- This data will be then used as input to K-means clustering algorithm to obtain the clusters of areas (metro stations)

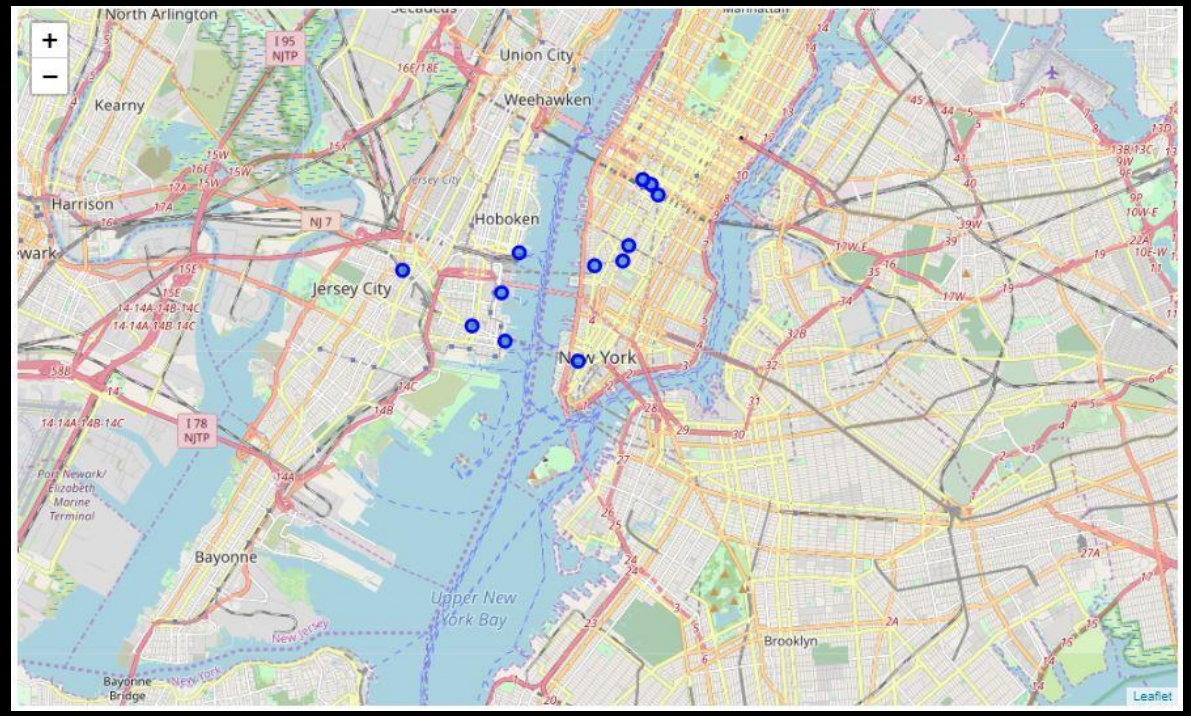

Fig : Map of New York with Metro Stations

METHODLOGY

- Machine Learning clustering algorithm was used to divide the stations and gyms data set into clusters of similar locations
- The elbow method was used to find the most suitable number of clusters (3 clusters)

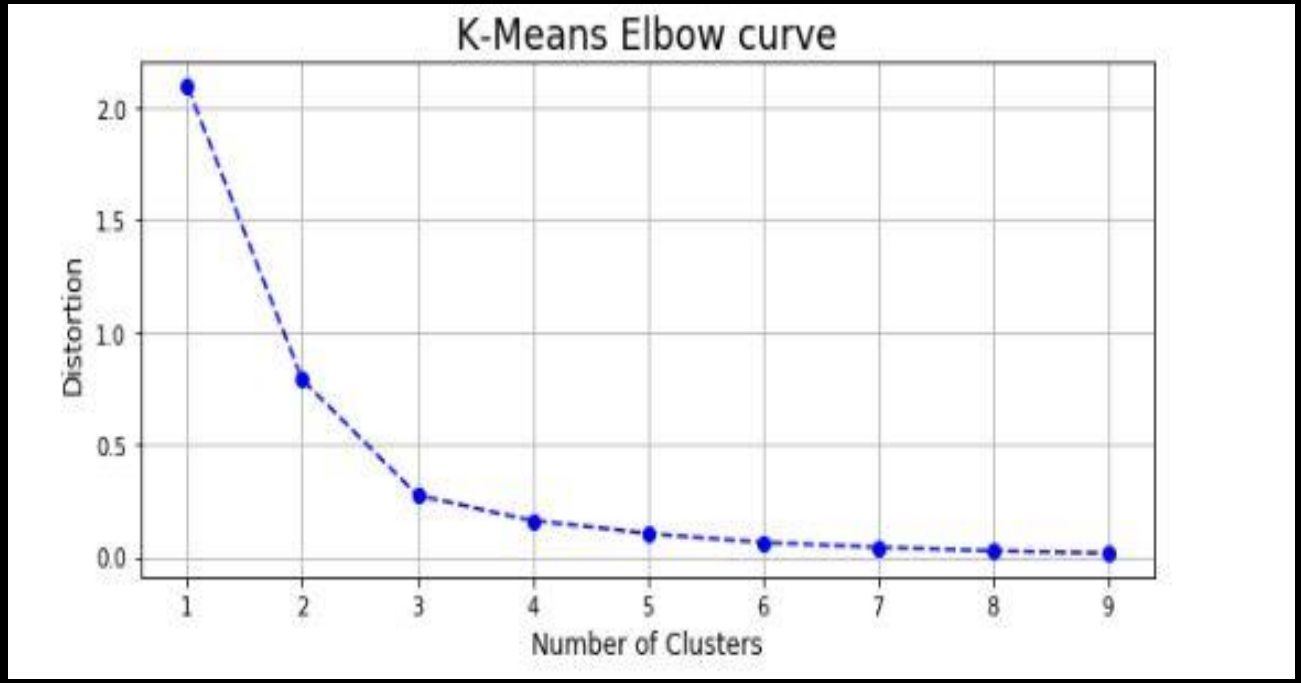

Fig : Elbow Curve

RESULT

After executing the K-Means clustering algorithm three clusters of Metro stations were created, identified by their respective colors on the following map

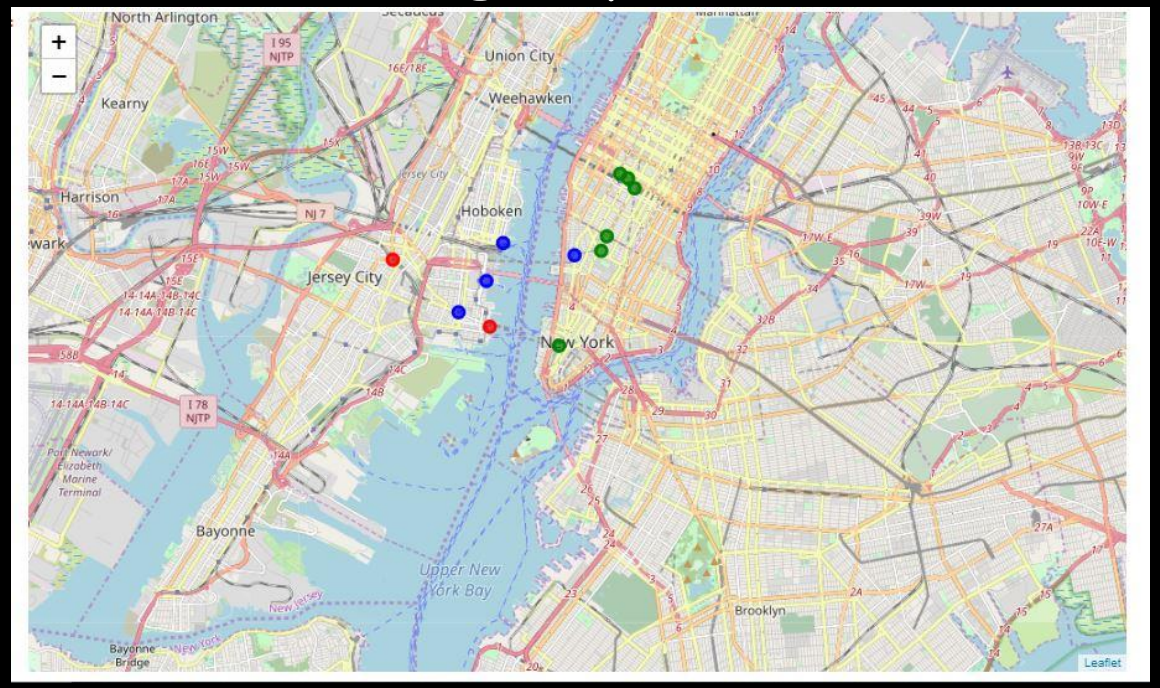

Fig : Three Clusters of Metro Stations

A graph of the distribution of clusters and their final centroids (center points) in black is shown below

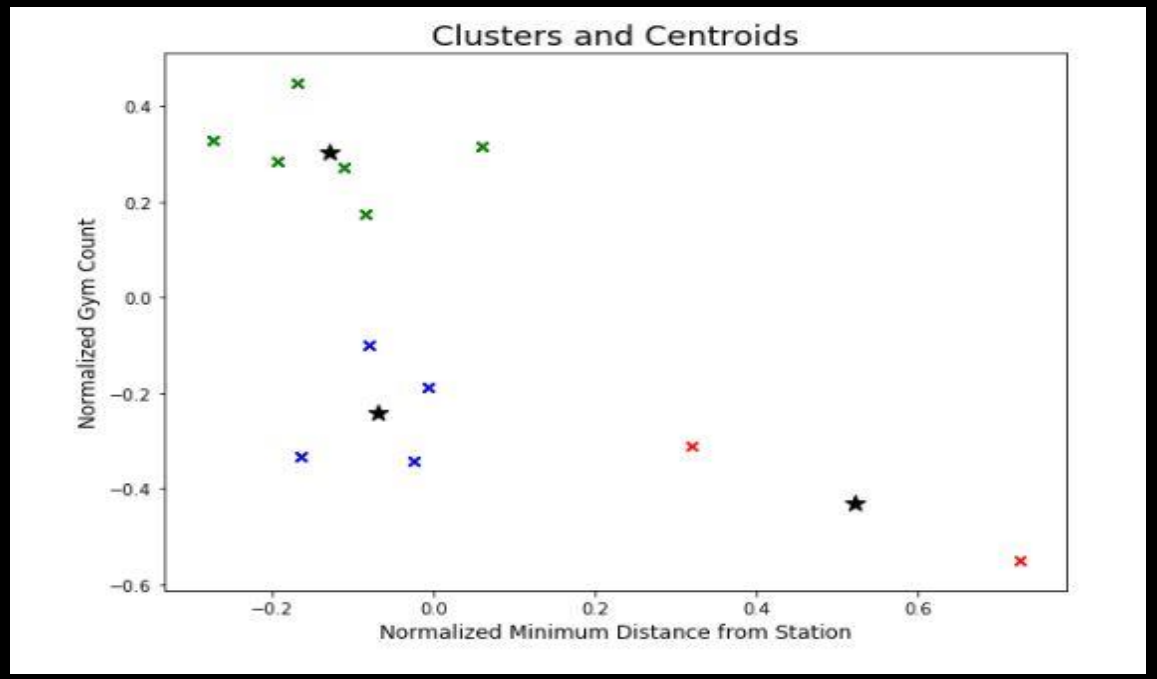

Fig : Clusters and Centroids

RESULT (CONTD..)

- **Cluster 1 [Cluster Label 0] (Green color on the map):**
  There are already many existing gyms in the area and the nearest gym is in most cases in a relatively short distance from the station

- **Cluster 2 [Cluster Label 1] (Red color on the map):**
  There are not many already existing gyms in the area and the nearest gym is in most cases relatively not in a short distance to the metro station

Cluster 3 [Cluster Label 2] (Blue color on the map):
Although not a prohibitive metro station to open a gym in its vicinity, there is already a fair number of gyms in the area and the nearest one is not far from the metro station.

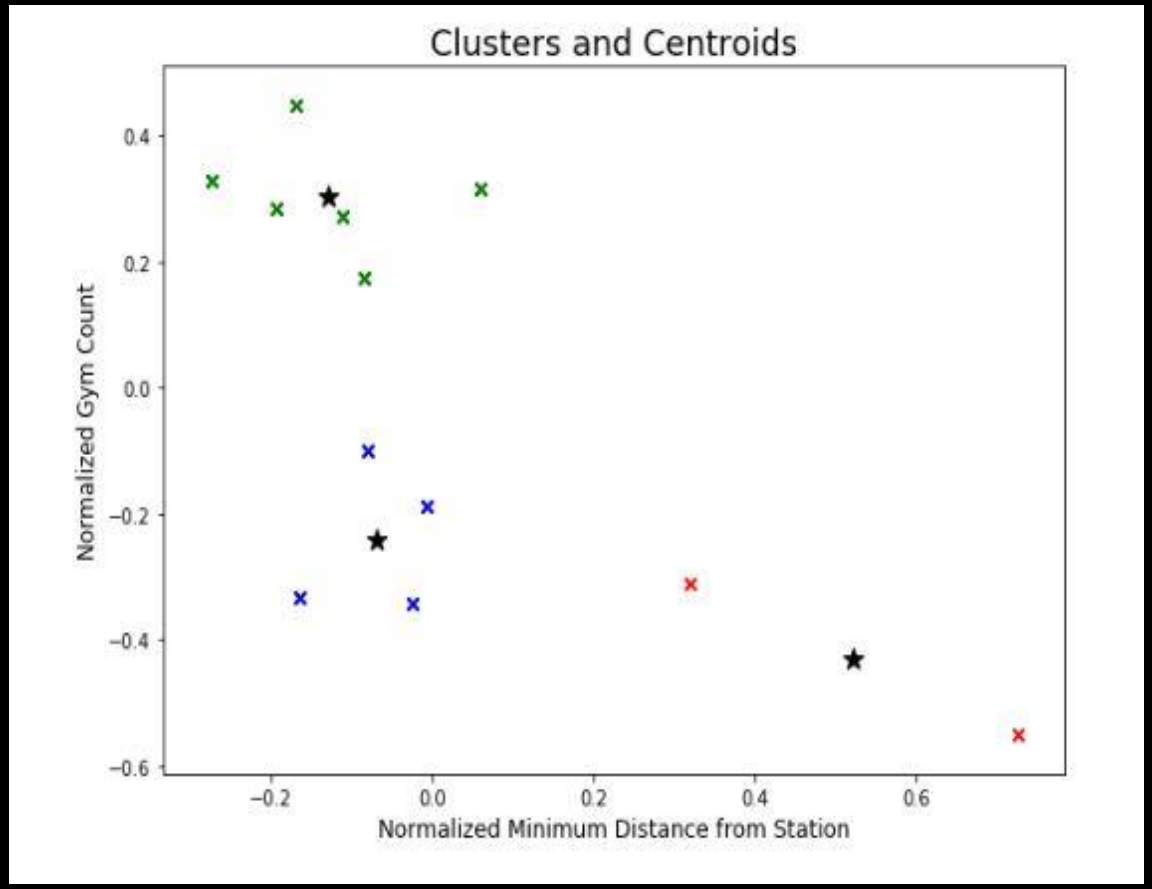

Fig : Clusters and Centroids

DISCUSSION

- Clusters of areas (in our case Metro stations) were identified as groups of similar in their potential locations for opening a gym
- Some examples of extra factors can be :
  - Population density in that neighborhood
  - Number of business operating in that neighborhood (people might want their workplace close to the gym)
  - Age
  - Average income of the people in that neighborhood
  - Property prices in that neighborhood

CONCLUSION

The above results can be a good starting point for a businessman who wants to set-up their gym in a new and happening city like New York. Similar methodology can be used for setting up any business, probably with customized criteria

THANK YOU