



Coursera Project

IBM Applied Data Science Capstone

THE BATTLE OF NEIGHBORHOODS

Introduction

- ▶ Find the suitable location to open a new restaurants in Brooklyn, NY
- ▶ Build a data frame of neighborhood in Brooklyn, NY by web scraping the data from website
- ▶ Use 'Foursquare' API to get the geographical coordinated of the neighborhoods and visualize the result using folium package.

Data Section:

- ▶ List of neighbourhoods in Brooklyn, New York has 5 boroughs and 306 neighbourhoods.
- ▶ Source of the data is obtained from https://geo.nyu.edu/catalog/nyu_2451_34572
- ▶ Use the Foursquare API to explore neighborhoods in New York City by using the ****explore**** function to get the most common venue categories in each neighborhood.

Methodology:

- ▶ Web scraping using the python and beautiful soup packages to the extract the list of neighborhood data with latitudes and longitude details.
- ▶ Use 'FourSquare' API to get the most common venues in the Brooklyn, NY using Foursquare credentials.
- ▶ Perform clustering on the data by using the k-means clustering.

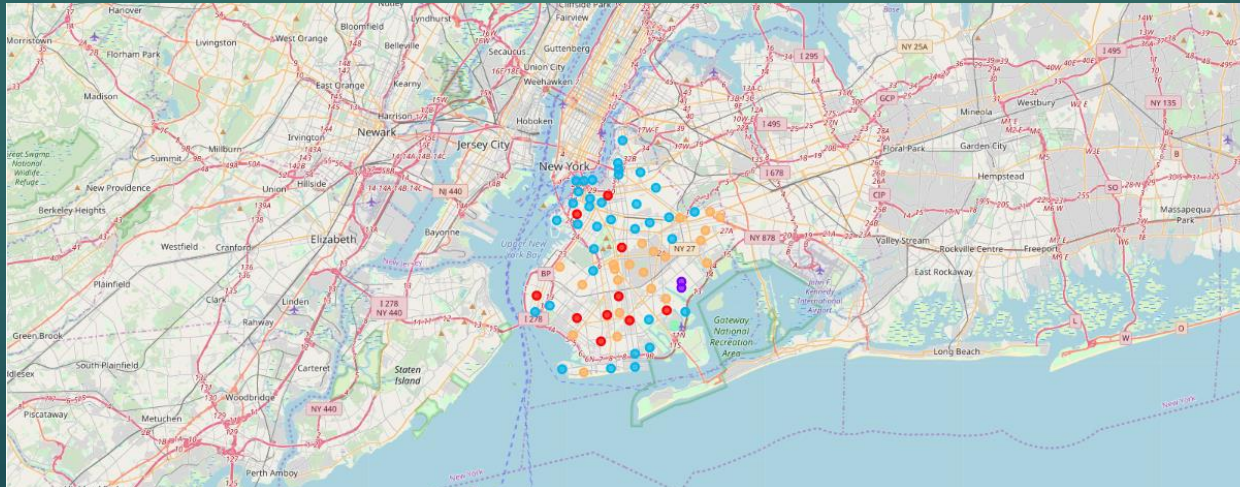
Result

- The result from the k-means clustering show that we can categorize the neighborhoods into 4 cluster based on the frequency of occurrence.

Cluster 0 : Neighborhoods with moderate frequency of operations.

Cluster 1, 2: Neighborhoods with low frequency of operations

Cluster 3, 4: Neighborhood with high frequency of operations



Conclusion

- ▶ The conclusion on the project is the best place to start restaurants in northern Brooklyn neighborhoods and starting with the fruit and salad restaurants with less junk foods will be the greatest combination to start new restaurants in the Brooklyn, NY.

Thank You