

Coursera Capstone IBM Applied Data Science Capstone

New Shopping Mall Opening in Brooklyn, NY, United States of America

Done by: Hasan Ali Mohamed Taresh

April, 2020

Business Problem

- **Objective:** To analyze and choose the best locations in Brooklyn (a borough of New York City, USA) to open a new shopping mall.
- **Business Question:** If a property developer is planning to open a new shopping mall in Brooklyn, in New York City, USA, where is the best recommended area that should be chosen to build the mall on?

Data

- **Data Required:**

- A list of neighborhoods in Brooklyn.
- Coordinates of the neighborhoods.
- Venue data, particularly data related to shopping malls.

- **Data Sources:**

- Wikipedia Page: [https://en.wikipedia.org/wiki/Category:Neighborhoods in Brooklyn](https://en.wikipedia.org/wiki/Category:Neighborhoods_in_Brooklyn).
- Geocoder package for coordinates.
- Foursquare API for venue data.

Methodology

- Web scraping Wikipedia page for neighborhoods list.
- Get latitude and longitude coordinates using Geocoder.
- Use Foursquare API to get venue data.
- Group data by neighborhood and taking the mean of the frequency of occurrence of each venue category.
- Filter venue category by Shopping Mall.
- Perform clustering on the data by using k-means clustering.
- Visualize the clusters in a map using Folium.

Results

- Neighborhoods are categorized into 2 clusters:
 - Cluster 0: Neighborhoods with no existence of shopping malls.
 - Cluster 1: Neighborhoods with low to moderate number of shopping malls.



Discussion

- After observation of data, we noticed the following:
 - Most of shopping malls are concentrated on cluster 1, which is in the lower side of Brooklyn close to the beach.
 - The upper part of Brooklyn has few to no shopping malls in the neighborhoods.
 - This represents a great opportunity and high potential areas to open new shopping malls as there is very little to no competition from existing malls.
 - Meanwhile, shopping malls in cluster 1 are likely suffering from intense competition due to oversupply and high concentration of shopping malls.

Conclusion

- **Answer to Business Question:** The neighborhoods in cluster 0 are the most preferred locations to open a new shopping mall.
- Findings of this project will help the relevant stakeholders to capitalize on the opportunities on high potential locations while avoiding overcrowded areas in their decisions to open a new shopping mall.

Thanks!

