Coursera Capstone

Opening a New Shopping Mall in Mumbai, India.

IBM Applied Data Science Capstone

By Ruchit Bagade
May 2020

Business Problem

• Location of the shopping mall is one of the most important decisions that will determine whether the mall will be a success or a failure

• Objective: To analyze and select the best locations in the city of Mumbai, India to open a new shopping mall

• This project is particularly useful to property developers and investors looking to open or invest in new shopping malls in Mumbai.

Business Question

In Mumbai, if a property developer, builders or real-estate businessmen is looking to open a new shopping mall then which location will be the best location for the mall to open ? What would you recommend?

Data

- Data required

➢List of neighborhoods in Mumbai, India.
➢Latitude and longitude coordinates of the neighborhoods
➢Venue data, particularly data related to shopping malls
- Sources of data

➢Wikipedia page for neighbourhoods
(https://en.wikipedia.org/wiki/Category:Suburbs_in_Mumbai)
➢Geocoder package for latitude and longitude 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

Categorized the neighborhoods into 3 clusters :

• Cluster 0: Neighborhoods with low number to no existence of shopping malls.
• Cluster 1: Neighborhood with high concentration of shopping malls
• Cluster 2: Neighborhoods with moderate number of shopping malls.
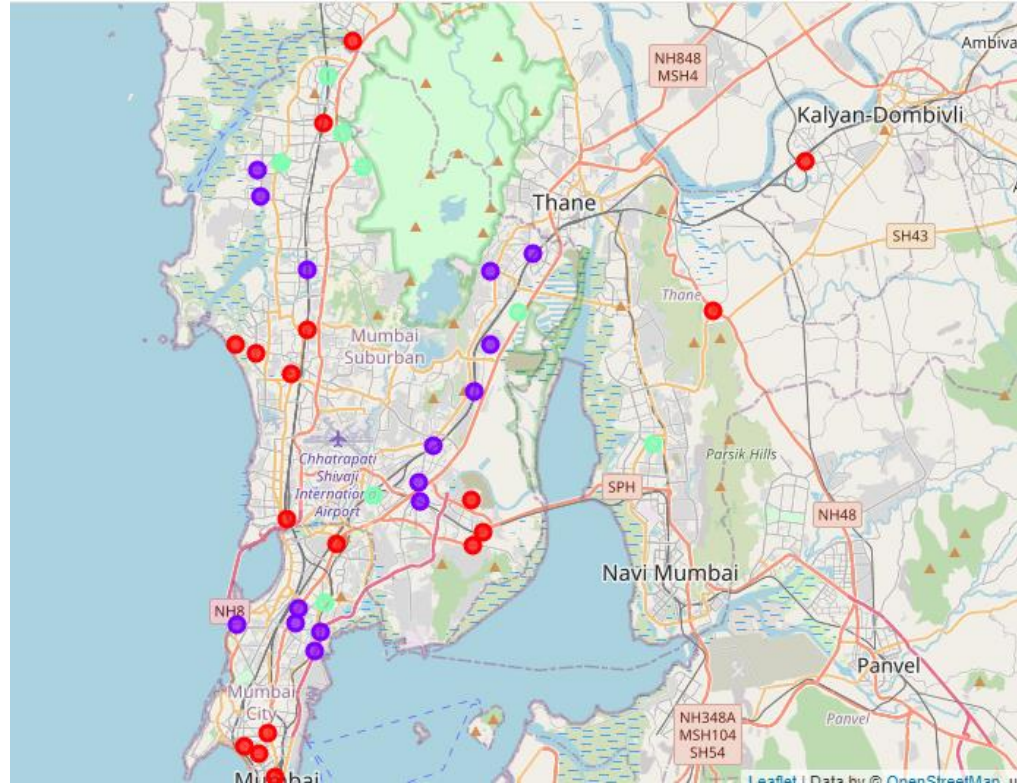
The results of the clustering are visualized in the map below with
cluster 0 in red color,
cluster 1 in purple color,
cluster 2 in mint green color.

Discussion

Categorized the neighborhoods into 3 clusters :

Most of the shopping malls are concentrated in the southern and central area of the city
- Highest number in cluster 1 and moderate number in cluster 2
- Cluster 0 has very low number to no shopping mall in the neighborhoods
- Oversupply of shopping malls mostly happened in the southern area of the city, with the suburb area still have very few shopping malls

Recommendations

• Open new shopping malls in neighborhoods in cluster 0 with little to no competition

• Can also open in neighborhoods in cluster 2 with moderate competition if have unique selling propositions to stand out from the competition

• Avoid neighborhoods in cluster 1, already high concentration of shopping malls and intense competition

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

Thank You