Coursera Capstone

IBM Applied Data Science Capstone

Opening a New Shopping Mall in Mumbai, India

BY: HARSH CHANDAK
APRIL 2020

Business Problem

❖ Location of 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

❖ Business question
  ➢ In the city of Mumbai, India, if a property developer is looking to open a new shopping mall, where would you recommend that they open it?

Data

❖ Data required
- ➢ List of neighborhoods in Mumbai
- ➢ Latitude and longitude coordinates of the neighborhoods
- ➢ Venue data, particularly data related to shopping malls

❖ Sources of data
- ➢ Wikipedia page for neighborhoods (https://en.wikipedia.org/wiki/Category:Neighborhoods_in_Mumbai)
- ➢ Geopy (geopy.geocoders) 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 Geopy

❖ 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 moderate number of shopping malls

➢ Cluster 1: Neighborhoods with low number to no existence of shopping malls

➢ Cluster 2: Neighborhoods with high concentration of shopping malls
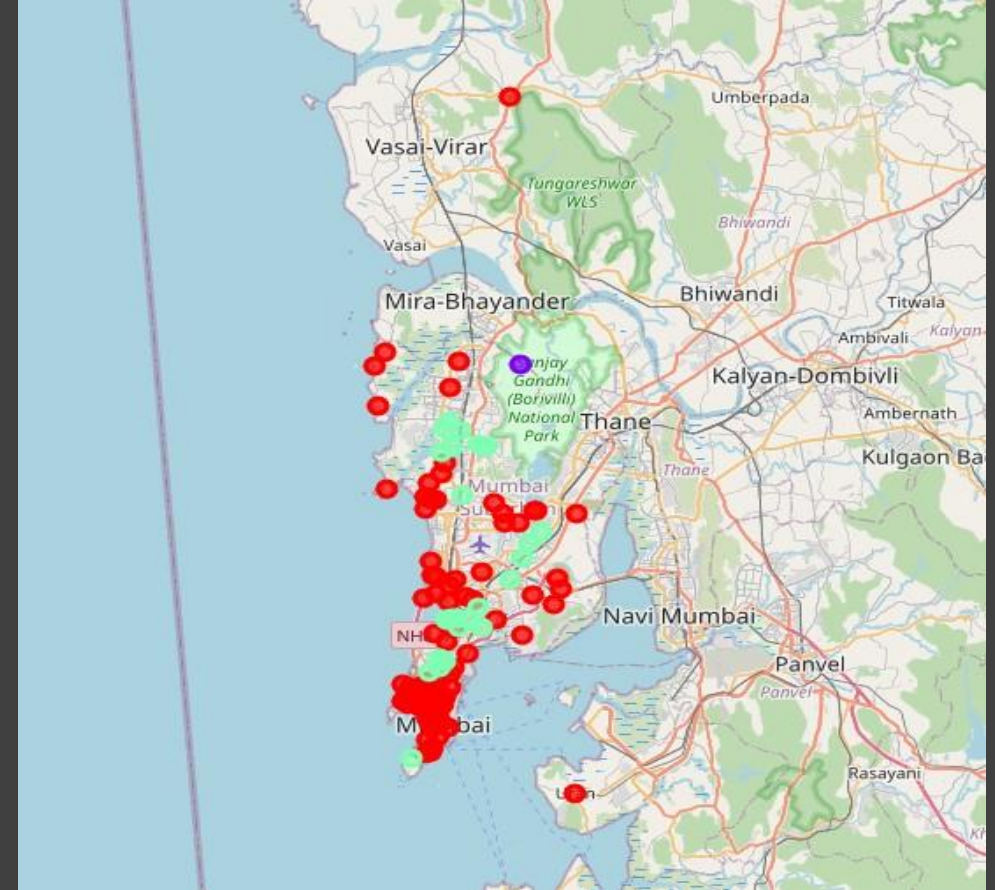
Discussion

❖ Most of the shopping malls are concentrated in the southern and central area of Mumbai

❖ Highest number of shopping malls in cluster 2 and moderate number in cluster 0

❖ Cluster 1 has very low number to no shopping mall in the neighborhoods

❖ There is opportunity to make shopping malls in the northern area and central area of Mumbai

Conclusion

❖ Answer to business question: The neighborhoods in cluster 1 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