Coursera Capstone
IBM Applied Data Science Capstone

Opening a New Shopping Mall in Trivandrum, India

By: Kevin Varghese
July2020

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 Trivandrum India to open a new shopping mall

▶ This project is timely as the city is currently suffering from oversupply of shopping malls

▶ Business question

▶ ➤In the city of Trivandrum, 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 Trivandrum

▶ 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:Suburbs_of_Thiruvananthapuram)

▶ 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 high number of shopping malls

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

- Cluster 2: Neighborhoods with moderate number of shopping malls

- Cluster 3: Neighborhoods with low number to no existence of shopping malls
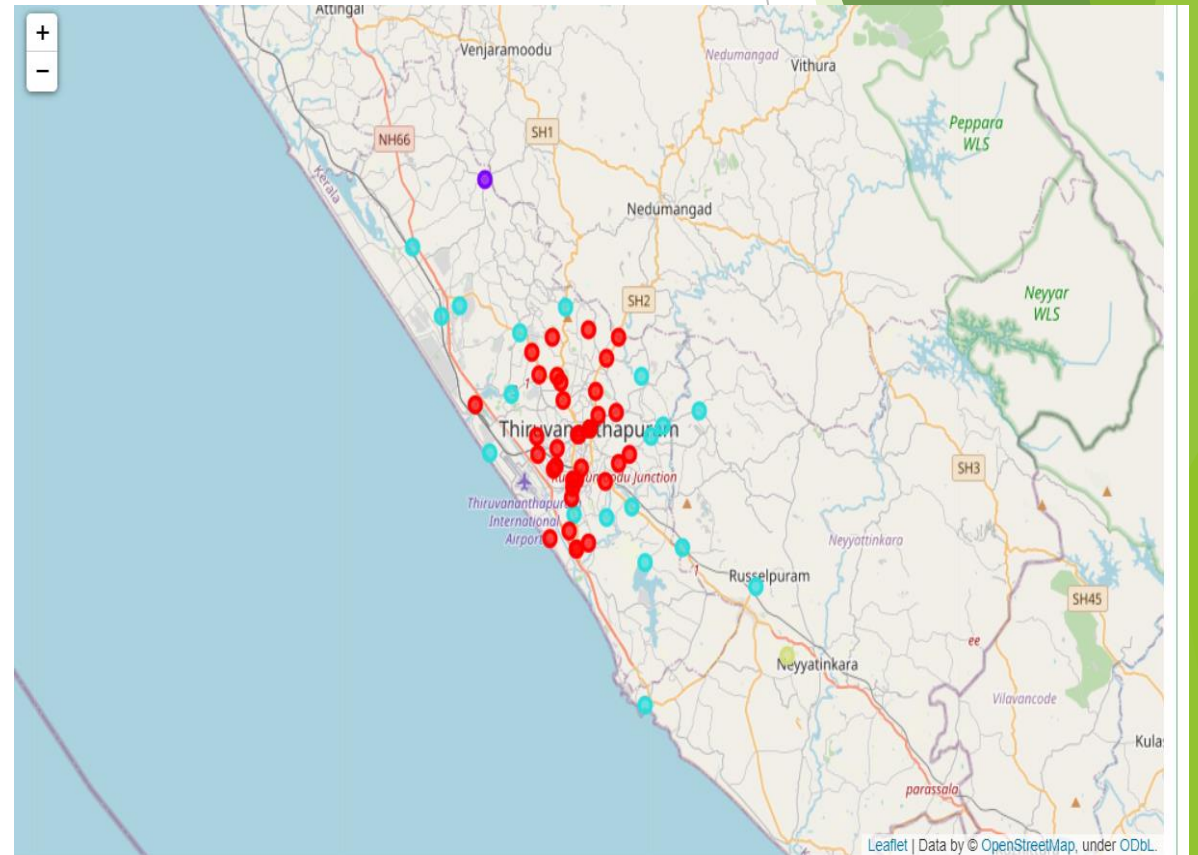
Discussion

► Most of the shopping malls are concentrated in the central area of the city

► Highest number in cluster 0 and moderate number in cluster 2

► Cluster 1 and cluster 3 has very low number to no shopping mall in the neighborhoods

► Oversupply of shopping malls mostly happened in the central area of the city, with the suburb area still have very few shopping malls

Recommendations

▶ Open new shopping malls in neighborhoods in cluster 1 and cluster 3 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 0, already high concentration of shopping malls and intense competition

Conclusion

- Answer to business question: The neighborhoods in cluster 1 and cluster 3 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