Coursera Capstone

IBM Applied Data Science Capstone

Opening Shopping Mall

Yashjeet Singh

August, 2020

THE PROBLEM

- Objective: To analyse and select the best locations in the city of Kuala Lumpur, Malaysia 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 Kuala Lumpur, Malaysia, if a property developer is looking to open a new shopping mall, where would you recommend that they open it?
- 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

DATA

- Data required
  - List of neighbourhoods in Kuala Lumpur
  - Latitude and longitude coordinates of the neighbourhoods
  - Venue data, particularly data related to shopping malls
- Sources of data
  - Wikipedia page for neighbourhoods (https://en.wikipedia.org/wiki/Category:Suburbs_in_Kuala_Lumpur)
  - Geocoder package for latitude and longitude coordinates
  - Foursquare API for venue data

IMPLEMENTATION

- **Collecting data & resources**
  - Web scraping Wikipedia page for neighbourhoods list
  - Get latitude and longitude coordinates using Geocoder

- **Processing Data**
  - Use Foursquare API to get venue data
  - Group data by neighbourhood and taking the mean of the frequency of occurrence of each venue category
  - Filter venue category by Shopping Mall

- **Clustering & visualizing**
  - Perform clustering on the data by using k-means clustering
  - Visualize the clusters in a map using Folium

RESULT

Categorized the neighbourhoods into 3 clusters :

- Cluster 0: Neighbourhoods with moderate number of shopping malls
- Cluster 1: Neighbourhoods with low number to no existence of shopping malls
- Cluster 2: Neighbourhoods with high concentration of shopping malls

DISCUSSION

- Most of the shopping malls are concentrated in the central area of the city
- Highest number in cluster 2 and moderate number in cluster 0
- Cluster 1 has very low number to no shopping mall in the neighbourhoods
- 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 neighbourhoods in cluster 1 with little to no competition
- Can also open in neighbourhoods in cluster 0 with moderate competition if have unique selling propositions to stand out from the competition
- Avoid neighbourhoods in cluster 2, already high concentration of shopping malls and intense competition

CONCLUSION

- Answer to business question: The neighbourhoods 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