Coursera Capstone

IBM Applied Data Science Capstone

Opening a New supermarket in Kuala Lumpur, Malaysia

By: Sahil Aggarwal

Business Problem

- Location of the supermarket is one of the most important decisions that will determine whether it will be a success or a failure

- **Objective**: To analyse and select the best locations in the city of Kuala Lumpur, Malaysia to open a new supermarket

- **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?

Data

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

Methodology

- Web scraping Wikipedia page for neighbourhoods list

- Get latitude and longitude coordinates using Geocoder

- 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 Supermarket

- Perform clustering on the data by using k-means clustering

- Visualize the clusters in a map using Folium

Results

Categorized the neighbourhoods into 3 clusters :

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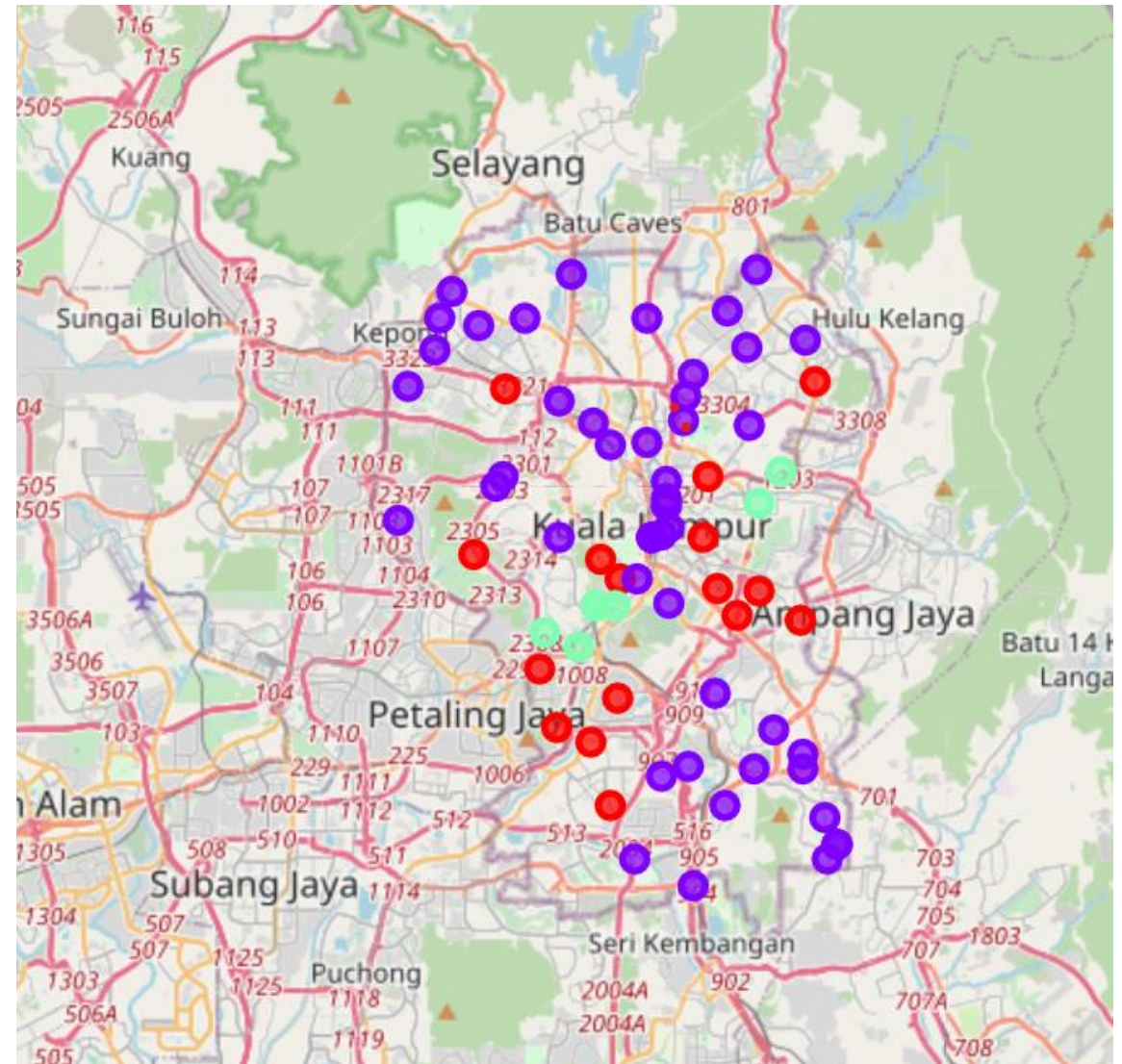
Discussion

Most of the supermarkets 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 supermarkets mostly happened in the central area of the city, with the suburb area still have very few supermarkets.

Recommendations

01

Open new supermarkets in neighbourhoods in cluster 1 with little to no competition

02

Can also open in neighbourhoods in cluster 0 with moderate competition if have unique selling propositions to stand out from the competition

03

Avoid neighbourhoods in cluster 2, already high concentration of supermarkets and intense competition

Conclusion

- Answer to business question: The neighbourhoods in cluster 1 are the most preferred locations to open a new supermarket

- 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 supermarket