

Coursera Capstone

IBM Applied Data Science Capstone

Buying a House in Kuala Lumpur

Muhammad Syahril

Business Problems

- House is one of the essential requirements for humans. Purchasing a house is one of the largest investments to be made.
- Objective: to analyze and select the best locations in the city of Kuala Lumpur, Malaysia to buy a house.
- Business question: In the city of Kuala Lumpur, Malaysia, if a homebuyer is looking to buy a house, where would you recommend that they buy it?

Data Acquisition

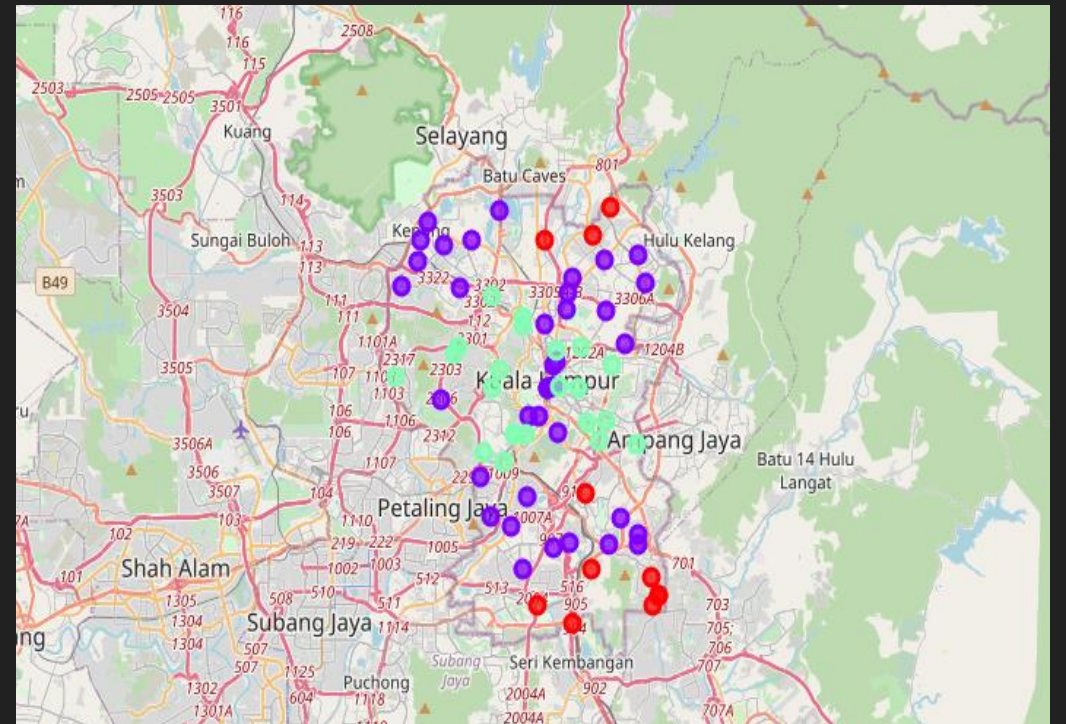
- Data required
 - List of neighbourhoods in Kuala Lumpur
 - Latitude and longitude coordinates of the neighbourhoods
 - Venue data, particularly data related to shopping malls
- Sources of data
 - Wikipedia page for neighbourhoods
(https://en.wikipedia.org/wiki/Category:Suburbs_in_Kuala_Lumpur)
 - Geocoder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

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 neighborhoods and taking the mean of the frequency of occurrence of each venue category
- Filter venue category by categories (Shopping Mall, Supermarket, Convenience Store, Halal Restaurant, Grocery Store, Gas Station)
- Perform clustering on the data by using k-means clustering
- Visualize the clusters in a map using Folium

Results

- Cluster 0: Neighbourhoods with no presence of Supermarket and majority of Shopping Mall is not exist.
- Cluster 1: Neighbourhoods with low number in majority of all the categories.
- Cluster 2: Neighbourhoods with high frequency of Shopping Mall and majority decent value of other categories.



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 1
- Cluster 0 has very low number to no super market in the neighborhoods
- **Recommendations**
 - This project **recommends** homebuyers to capitalize on these findings to buy a house in cluster 2 that could give benefits according to lifestyle and necessities.
 - Cluster 1 is also an option for homebuyers who prefer to avoid bustle place but still wanted to have necessities at reach.
 - However, avoid cluster 0 because of shortage on many things that only could be suitable for certain type of lifestyle that does not depends too much necessities.

Conclusion

- Answer to business question: The neighborhoods in cluster 2 are the most preferred locations to buy a house.
- The findings of this project will help the relevant homebuyers on high potential locations for houses by their needs and lifestyle.