

**BEST
LOCATIONS
FOR NEW
MALL IN
KUALA
LUMPUR,
MALAYSIA**

IBM CAPSTONE PROJECT
BY
SURBHI CHATURVEDI



Business Problem

- Location is the most important part of decision-making process when building a new shopping
- City of Kuala Lumpur in Malaysia is currently suffering with the oversupply of shopping mall
- The location of existing malls help determine the success and failure of new mall in the same area



Objective

Data analysis for the selection of best locations in the city of Kuala Lumpur, Malaysia to open a new shopping mall




Business question

Recommendation backed by Data for real estate developers which advices which location is best suited for new shopping mall in the city of Kuala Lumpur, Malaysia.




Data

- List of neighborhoods in Kuala Lumpur
 - Latitude and longitude coordinates of the neighborhoods
 - Venue data, particularly data related to shopping malls
- 



Data Sources

- Wikipedia page for neighborhoods - 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 neighborhood and get mean frequency of occurrence of each venue category
- Filter venue category by Shopping Mall
- K-means for data clustering
- Folium for cluster map

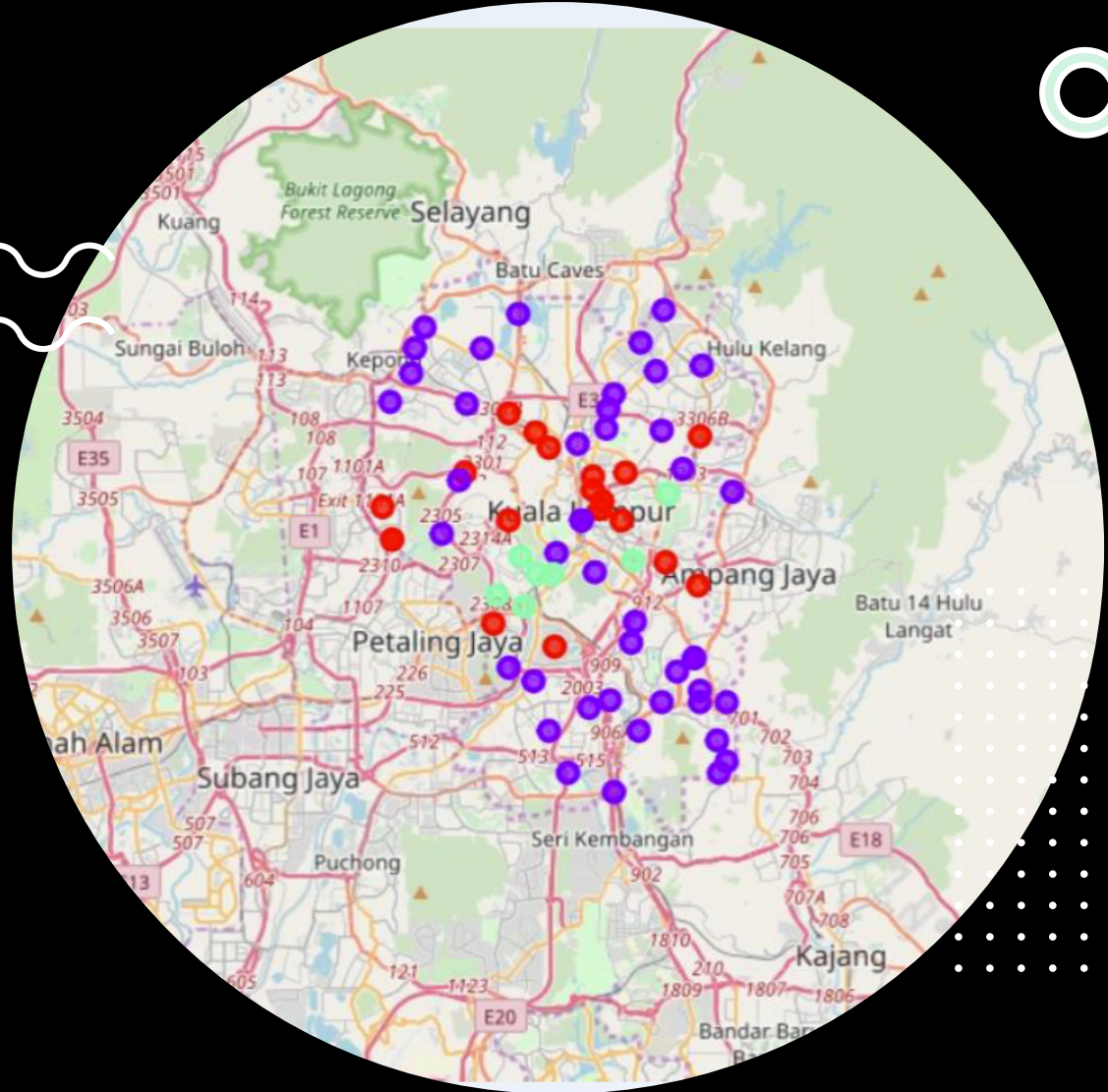


Results

Categorized the neighborhoods into 3 clusters :

- Cluster 0: Moderate number of shopping malls
- Cluster 1: Low number to zero shopping malls
- Cluster 2: High concentration of shopping malls

RESULTING MAP





Discussions

- Most of the shopping malls concentrated in the center of the city
- Oversupply of shopping malls mostly happened in the center of the city, while suburban areas still have very few shopping malls



Recommendations

- Neighborhoods in Cluster 1 has least competition for the new shopping mall
- Cluster 0 with moderate number can be good if the business is unique and has the potential to standout
- Avoid neighborhoods in cluster 2 due to higher concentration of malls