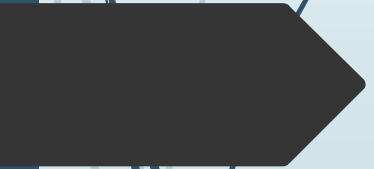


Open a new shopping mall in Kuala Lumpur

Capstone IBM Project Data Science

By VAMSI SANAPATHI





Business Problem

- To analyze and recommend best location in Kuala Lumpur City for opening a new shopping mall to the Property Developers and Real Estate Investors.
- Business Question – In the city of Kuala Lumpur which is the best neighborhood for opening a new shopping mall ?



Data

- List of neighborhoods in Kuala Lumpur City.
- Latitude and Longitude coordinates for each of the neighborhoods.
- Venue data i.e. shopping malls.

- Source of the data –
- This Wikipedia page
(https://en.wikipedia.org/wiki/Category:Suburbs_in_Kuala_Lumpur)
- Foursquare API data for venue data (shopping malls).

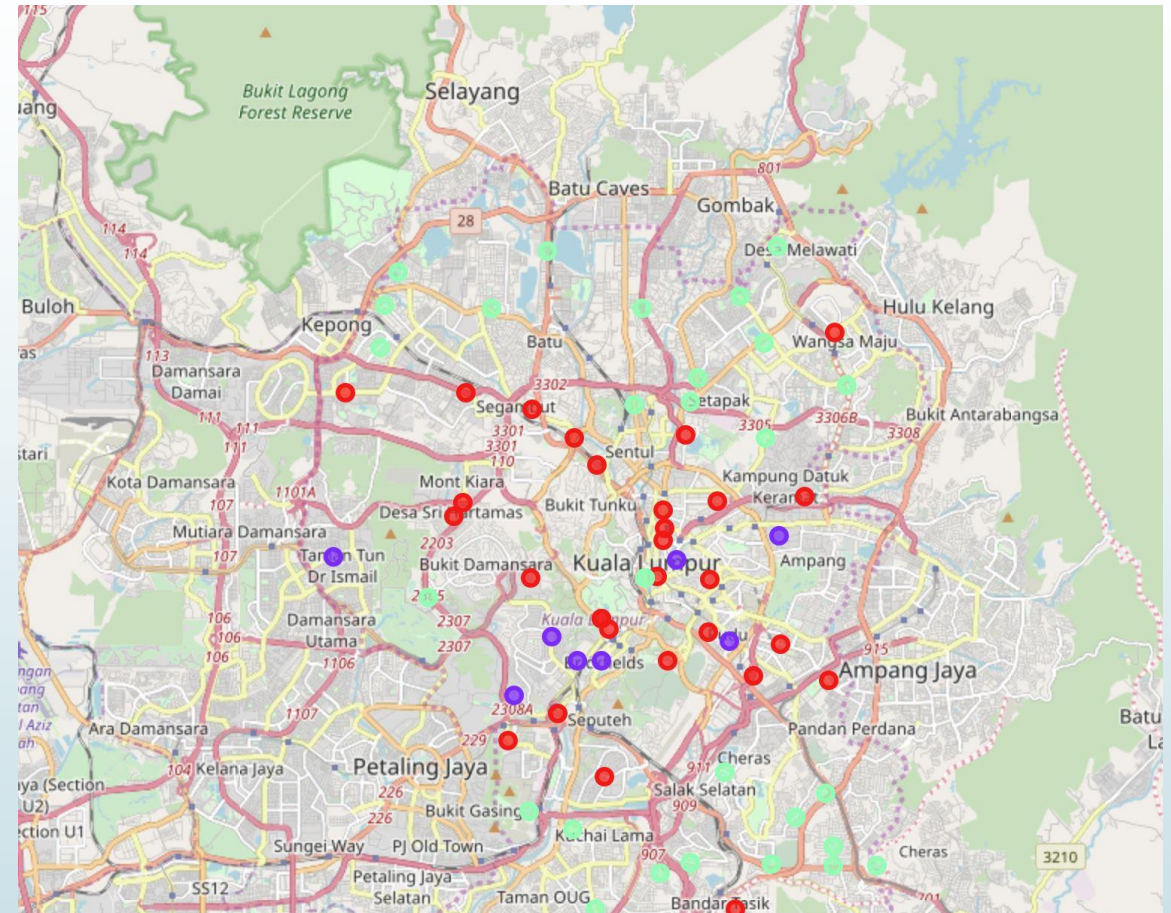


Methodology

- Web scraping of Wikipedia page for neighborhoods list.
- Get latitude and longitude values using Geocoder package.
- Use Foursquare API data for venue data (shopping malls).
- Group data by neighborhoods and taking the mean of frequency of occurrence of each venue category.
- Filter venue data by shopping malls.
- Perform K-Means Clustering.
- Visualize clusters in maps using folium.

Results

- Cluster 0 – Has moderate number of shopping malls.
- Cluster 1 – Has highest number of shopping malls.
- Cluster 2 – Has no major shopping malls.





Discussion

- Most of the shopping malls are located in the central are of the Kuala Lumpur City.
- Cluster 0 – Has moderate number of shopping malls.
- Cluster 1 – Has highest number of shopping malls.
- Cluster 2 – Has no major shopping malls, best neighborhood cluster for opening shopping malls.
- Central city has oversupply of shopping malls. Suburbs has lesser number of shopping malls.



Recommendation -

- Cluster 2 has no major shopping malls and hence it is the best available neighborhood cluster for opening new shopping malls because of less competition.
- Cluster 0 has moderate number of shopping malls so it can be considered as second alternative for opening new shopping malls.
- Cluster 1 has highest number of shopping malls so it is not recommended to open new shopping malls in this neighborhood cluster.



Conclusion

- Cluster 2 has is the best available neighborhood cluster for opening new shopping malls because of less competition.
- 