Opening New Café in Kuala Lumpur

IBM Applied Data Science Capstone
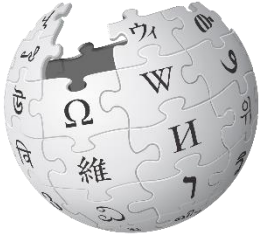
Ahmad Syahmi Adnan

7/8/2020

Business Problem

- Problem statement:
  - Which area/neighbourhood in Kuala Lumpur is the best to open a new café?

- Objective:
  - To analyse and select the best locations in the city of Kuala Lumpur, Malaysia to open a new coffee house business.

Data Required & Data Source

- List of neighbourhood in Kuala Lumpur.
  - Web scrapping from (https://en.wikipedia.org/wiki/Category:Suburbs_in_Kuala_Lumpur)

- Venue data in each neighbourhood
  - From Foursquare API

- Coordinate of neighbourhoods in Kuala Lumpur.
 - Using Geocoder

Methodology

- Get the list of neighbourhoods in Kuala Lumpur by using Python requests and beautifulsoup packages

- Use the Geocoder package to convert address into geographical coordinates in the form of latitude and longitude.

- Mapping each neighbourhoods on map.

- Use Foursquare API to extract top 100 venues from each neighbourhoods in radius of 2 km.

- Analyse each neighbourhood by grouping the rows by neighbourhood.

- Create new dataframe to determine the number of café in each neighbourhood.

- Run K-means to cluster the neighborhoods in Kuala Lumpur into 3 clusters.

- Analyse and visualise each cluster via Autoplotter and Geocoder packages.

Result & Discussion:
Cluster Map
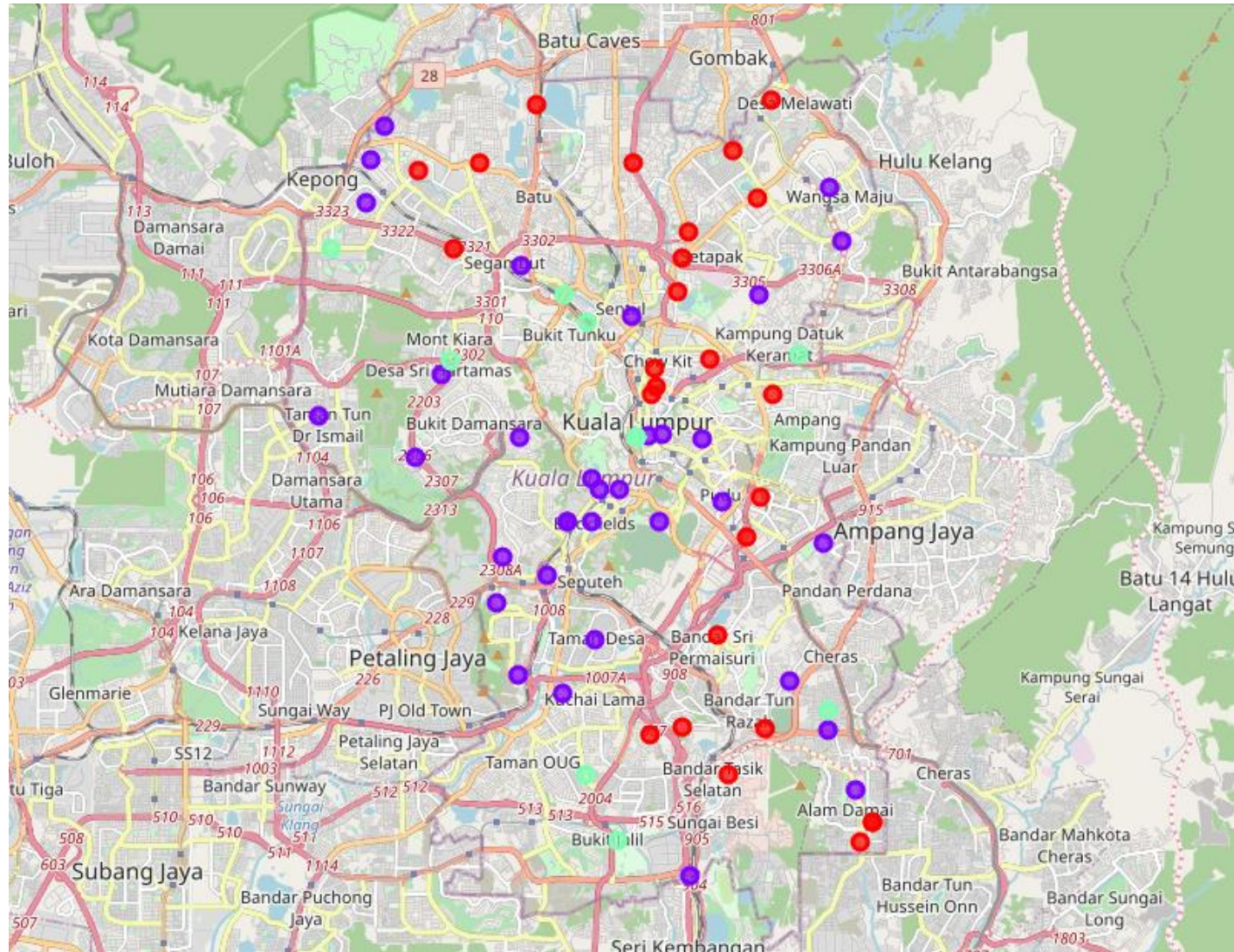
- 🔴 Cluster 0
- 🟣 Cluster 1
- 🟢 Cluster 2

Result & Discussion:
Cluster 0

- Low density of coffee shops

- 1 to 6 cafés in each neighbourhood.

- Consists of 22 neighbourhoods

- Most of the the neighbourhoods in this cluster located at northern and southeaster part of Kuala Lumpur.

- This area has high potential for opening of a new cafe since there is less competitor compared to neighbourhoods in cluster 1 and 2.

Result & Discussion:
Cluster 1

- Second highest density of cafés in the neighborhoods.

- Consist of 6 to 9 cafés.

- Biggest cluster with total of 33 neighborhoods and consist of almost half of all neighborhoods in Kuala Lumpur (47%).

- Most of the neighborhoods in this cluster are located at the central area of Kuala Lumpur which also consists a lot of commercial buildings and offices.

- Maybe the target market of cafés in that area is the worker from those commercial buildings.

Result & Discussion:
Cluster 2

- Highest count of café in each neighbourhood.

- Consists of more than 10 cafés.

- This cluster consists of 12 neigbourhoods out of 70 neighbourhoods which equivalet of only 17% of neighbourhoods in Kuala Lumpur.

- Located in area of Mont Kiara, Bukit Jalil, Bukit Tunku and Damansara which are categories as high end neighbourhood in Kuala Lumpur.

Recommendation

- This project recommends investor to capitalize on these findings to open new café in neighbourhoods in cluster 0 with little competition.

- Investor with unique selling propositions to stand out from the competition can also open new cafés in neighbourhoods in cluster 1 with moderate competition and also nearby to commercial buidings and office in Kuala Lumpur which is a good market for café.

- Investors are advised to avoid neighbourhoods in cluster 2 which already have high concentration of cafés and suffering from intense competition.

Conclusion

- The neighbourhoods in cluster 1 are the most preferred locations to open a new café. The 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 café.