Coursera Capstone

IBM APPLIED DATA SCIENCE CAPSTONE PROJECT

New Indian Restaurant in Kuala Lumpur, Malaysia

Vishal Gawai
Oct 2019

Business Problem:

- Kuala Lumpur, Malaysia is a great tourist attraction for many Indians. With rising India tourists, it provides great opportunity for new clients to open an Indian Restaurant to meet the rising demand of tourists.

- **Objective:**
  To analyze and identify new location where opening an Indian Restaurant can provide a great business opportunity to new business.

  What is the best place to open a new Indian Restaurant in Kuala Lumpur, Malaysia ?
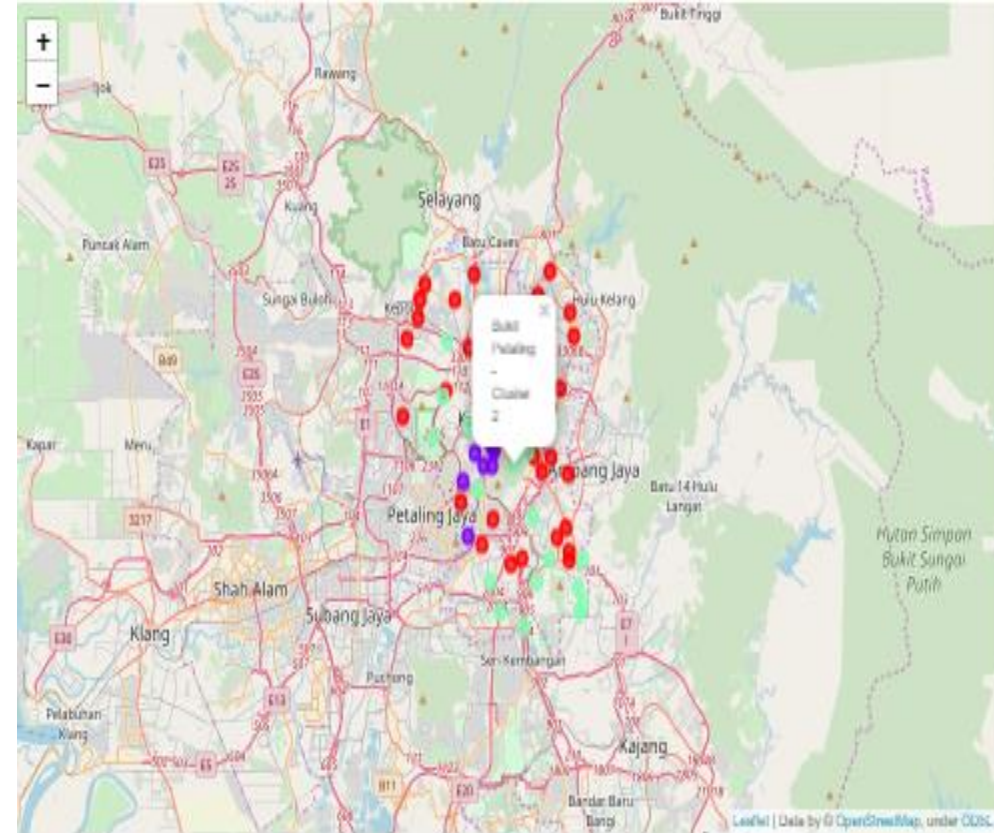
Data Information:

- **Data Gathering:**
  - ➤ Neighborhoods list in Kuala Lumpur.
  - ➤ Coordinates of Neighborhoods.
  - ➤ Venue details for Indian **Restaurant**.

- **Data Source:**
  - ➤ List of Neighborhoods from Wikipedia.
  - ➤ Get Coordinates of Neighborhoods using Geocode package.
  - ➤ Details for Indian Restaurant using Foursquare API.

Methodology:

- Get Neighborhood list using Beautiful Soup Package.

- Geocoder package to get Coordinate details.

- Venue Details of Indian Restaurant using Foursquare API.

- Group Neighborhood data and take mean frequency of occurrence of Venues categories.

- Sort Venues by Indian Restaurant.

- Using k-means algorithm, perform clustering.

- Using Folium library, visualize the data.

Results:

- **The Neighborhood has been categorized in 3 Clusters.**

  ☐ Cluster 0: Neighborhood with least concentration of Indian Restaurants.

  ☐ Cluster 1: Neighborhood with very high concentration of Indian Restaurants.

  ☐ Neighborhood with moderate concentration of Indian Restaurants.

Recommendation:

- Cluster 1 has the maximum number of Indian restaurants present and is hugely populated.

- Cluster 2 has moderate count of Indian restaurants which adequate opportunities.

- Cluster 0 has the least density of Indian restaurants providing the best neighborhoods in Kuala Lumpur to open new restaurant.

Conclusion:

❑ Through our analysis, we can conclude that neighborhoods in Cluster 1 is already concentrated with high number of Indian Restaurants and would provide bad opportunity to open new restaurants.

❑ Cluster 0 is least populated and Cluster 2 has moderate count of Indian restaurants providing the best chances for new Indian restaurants to thrive and grow.

❑ New Clients planning to open an Indian restaurants should focus on neighborhoods in cluster 0 and 1.

Thank You