Coursera Capstone

IBM Applied Data Science Capstone

Location identification for New Mall in Bengaluru, India.

Business Problem

- Objective: To analyse and select the best locations in Bengaluru city, Karnataka for a new shopping mall.
- Business Question: In Bengaluru, is a property developer is looking to open a new shopping mall, where would you recommend that they open?

Data

- Data required
  - List of neighbourhoods in Bengaluru.
  - Latitude and longitude of the neighbourhoods.
  - Venue data, particularly data related to shopping malls.

Data Source

- Wikipedia for neighbourhood list
- Geocoder for latitude and longitude coordinates
- Foursquare API for venue data

Methodology

- Web Scrapping Wikipedia page for neighbourhood list
- Get latitude and longitude coordinates using Geocoder
- Use Foursquare API for venue data
- Group data by neighbourhood and taking mean of the frequency of occurrence of each venue category
- Perform clustering on the data by using K-Means clustering
- Visualise the clusters in a map using Folium

Results

- Categorization of Neighbourhood into 3 clusters:
  - Cluster 0: Neighbourhood with the most number of shopping malls.
  - Cluster 1: Neighbourhood with only one shopping mall.
  - Cluster 2: Neighbourhood with moderate number of shopping malls.

Discussion

- Most of the shopping malls are concentrated in cluster 0.
- Cluster 1 has the least number of shopping malls.
- Oversupply of shopping malls mostly happened in the central area of the city.

Recommendations

- Open new shopping malls in neighbourhoods in cluster 1 which has just one mall
- Avoid neighbourhood in cluster 0 as it already has a high shopping mall concentration

Conclusion

- Answer to business question: The neighbourhoods cluster 0 should be the preferred location for a new mall.
- 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 shopping malls.

Thank you!