Coursera Capstone

IBM Applied Data Science Capstone

Opening a New Shopping Mall in Cape Town, South Africa

By: Rishi Ghosh

June 2020

Business Problem

- Location of the shopping mall is one of the most important decisions that will determine whether the mall will be a success or a failure

- Objective: To analyse and select the best locations in the city of Cape Town to open a new shopping mall

- This project is timely as the city is currently suffering from oversupply of shopping malls

- Business question
   - In the city of Cape Town, if a property developer is looking to open a new shopping mall, where would you recommend that they open it?

Data

- Data required
  - List of neighbourhoods in Cape Town
  - Latitude and longitude coordinates of the neighbourhoods
  - Venue data, particularly data related to shopping malls

- Sources of data
  - Wikipedia page for neighbourhoods
    (https://en.wikipedia.org/wiki/Category:Suburbs_of_Cape_Town)
  - Geocoder package for latitude and longitude coordinates
  - Foursquare API for venue data

Methodology

- Web scraping Wikipedia page for neighbourhoods list
- Get latitude and longitude coordinates using Geocoder
- Use Foursquare API to get venue data
- Group data by neighbourhood and taking the mean of the frequency of occurrence of each venue category
- Filter venue category by Shopping Mall
- Perform clustering on the data by using k-means clustering
- Visualize the clusters in a map using Folium

Results

- Categorized the neighbourhoods into 3 clusters :
  - Cluster 0: Neighbourhoods with no shopping malls
  - Cluster 1: Neighbourhoods with high concentration of shopping malls
  - Cluster 2: Neighbourhoods with a single shopping mall

Discussion

- Most of the shopping malls are concentrated in the central area of the city
- Highest number in cluster 1.
- Cluster 0 has very no shopping mall in the neighbourhoods

Recommendations

- Open new shopping malls in neighbourhoods in cluster 0 with little to no competition

- Can also open in neighbourhoods in cluster 2 with moderate competition if have unique selling propositions to stand out from the competition

- Avoid neighbourhoods in cluster 1, already high concentration of shopping malls and intense competition

Conclusion

- Answer to business question: The neighbourhoods in cluster 0 are the most preferred locations to open a new shopping 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 mall