Coursera Capstone
IBM Applied Data Science Capstone

Setting up New Chain of Super Markets in Melbourne , Australia

By: Nisha Devendra Kumar

Business Problem

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

- Objective: To analyse and select the best locations in the city of Melbourne , Australia to open a new chain of Super Markets

- This project is timely as the city is currently suffering from oversupply of Super Markets

- Business question

- In the city of Melbourne , Australia, if a property developer is looking to open a new chain of Super Markets, where would you recommend that they start?

Data

- Data required:
  - List of neighbourhoods/Suburbs in Melbourne
  - Latitude and longitude coordinates of the neighbourhoods
  - Venue data, particularly data related to Super Markets
- Sources of data:
  Wikipedia page for neighbourhoods
  (https://en.wikipedia.org/wiki/Category:Suburbs_of_Melbourne)
 - Geocoder package for latitude and longitude coordinates Foursquare API for venue data

Methodology:

- Web scraping Wikipedia page for neighbourhoods/Suburbs 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 Super Markets

- 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 /Suburbs with moderate number of Super Markets
- ➢Cluster 1: Neighbourhoods/Suburbs with low number to no existence of Super Markets
- ➢Cluster 2: Neighbourhoods/Suburbs with high concentration of Super Markets
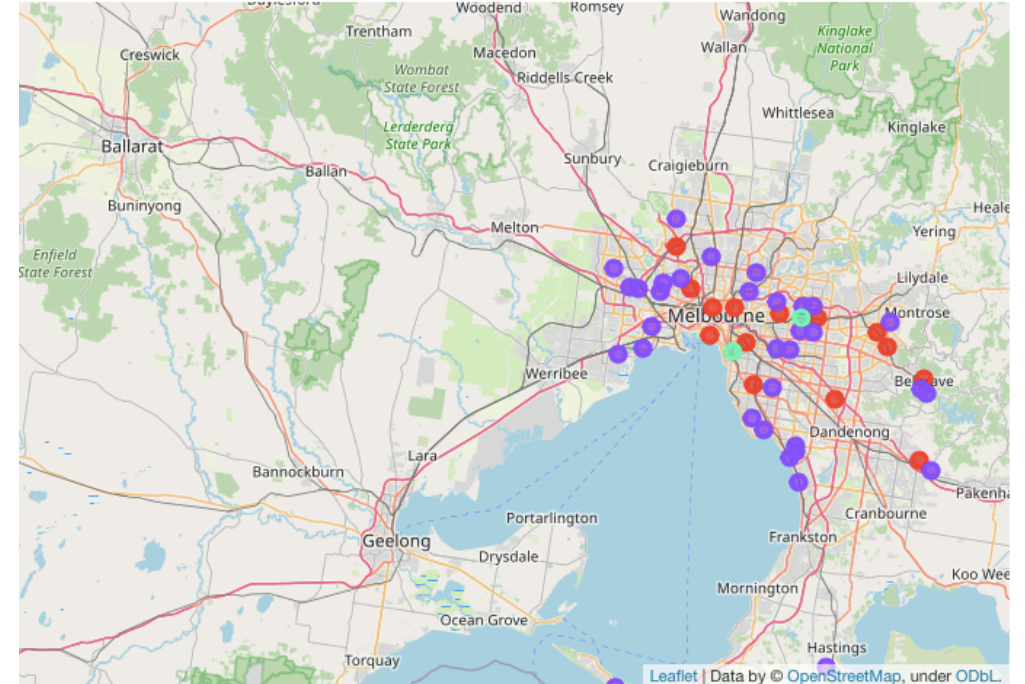
Discussion

- Most of the Super Markets are concentrated in the Northern & southern area of the city

- Highest number (per suburb) in cluster 2 (dense) and moderate number in cluster 0

- Cluster 1 has very low number of in the neighbourhoods

- Oversupply of Super Markets mostly happened in the central area of the city, with the suburb area still have very few Super Markets

Recommendations

- Open New Chain of Super Markets in Suburbs in cluster 1 with little less competition

- Can also open in neighbourhoods in cluster 0 with moderate competition if have unique selling propositions to stand out from the competition

- •Avoid neighbourhoods in cluster 2, already high concentration of Super Markets and intense competition