Coursera Capstone

IBM Applied Data Science Capstone

Opening a New Fitness Center in Bay Area, California

Business Problem

- Location of the fitness center is one of the most important decisions that will determine whether the fitness center will be a success or a failure.
- **Objective:** To analyse and select the best locations in the Bay Area, California to open a new fitness center.
- **Business question**
  - In the Bay Area, if a property developer is looking to open a new fitness center, where would you recommend that they open it.

Data

- **Data required**

  ➢List of neighbourhoods in the Bay Area

  ➢Latitude and longitude coordinates of the neighbourhoods

  ➢Venue data, particularly data related to fitness centers

- **Sources of data**

  ➢Wikipedia page for neighbourhoods ("https://en.wikipedia.org/wiki/Category:Counties_in_the_San_Francisco_Bay_Area")

  ➢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 gyms,fitness centers and yoga studios
- Perform clustering on the data by using k-means clustering
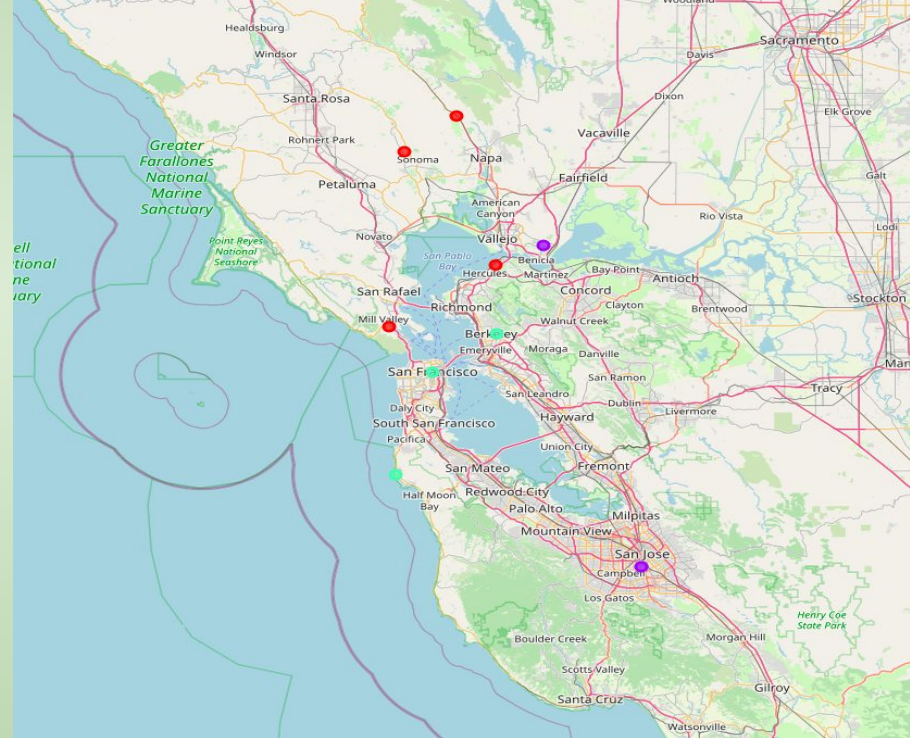- Visualize the clusters in a map using Folium

Results

• Categorized the neighbourhoods into 3 clusters :

➤**Cluster 0**: Neighbourhoods with low number to no existence of fitness centers

➤**Cluster 1**: Neighbourhoods with high concentration of fitness centers

➤**Cluster 2**: Neighbourhoods with moderate concentration of fitness centers

Discussion

- Highest number of fitness centers in cluster 1 and moderate number in cluster 2
- Cluster 0 has very low number to no fitness centers in the neighbourhoods
- Oversupply of fitness centers mostly happened in the cluster 1 neighborhoods , with the remaining neighborhoods still have very few fitness centers

Recommendations

- Open new fitness centers 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 fitness centers and intense competition

Conclusion

- **Answer to business question:** The neighbourhoods in cluster 0 are the most preferred locations to open a new fitness center

- 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 fitness center

Thank You!