IBM Applied Data Science Capstone

Opening a New Gym or Fitness Center in Bangalore, India

BY: KISHORE CHANDRA DASH

Business Problem

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

- **Objective** → To analyze and select the best locations in the city of Bangalore, India to open a new Gym.

- **Business question** → In the city of Bangalore, India, if a property developer is looking to open a new Gym, where would you recommend that they open it?

Data

- Data required :

  - List of neighborhoods in Bangalore.

  - Latitude and longitude coordinates of the neighborhoods.

  - Venue data, particularly data related to Gym.

- Sources of data:

  - Wikipedia page for neighborhoods (https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Bangalore)

  - Geocoder package for latitude and longitude coordinates
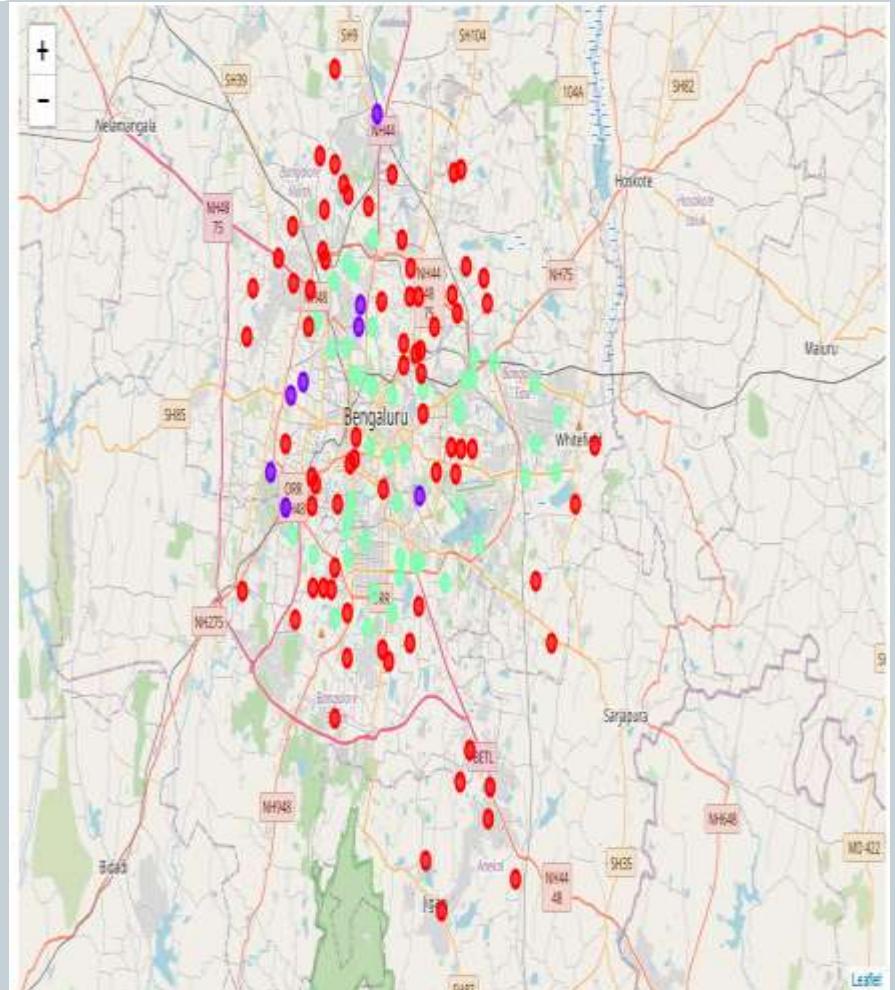
  - Foursquare API for venue data

Methodology

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

Results

- Categorized the neighborhoods into 3 clusters :

  - Cluster 0 (purple color): Neighborhoods with high concentration of Gyms.

  - Cluster 1 (mini green color): Neighborhoods with moderate number of Gyms.

  - Cluster 2 (red color.): Neighborhoods with low number to no existence of Gyms.

Discussion

- Most of the Gyms are concentrated in the central area of the city
- Highest number in cluster 0 and moderate number in cluster 1
- Cluster 2 has very low number to no Gym in the neighborhoods
- Oversupply of Gyms mostly happened in the central area of the city, with the suburb area still have very few Gyms

Recommendations

- Open new Gym in neighborhoods in cluster 2 with little to no competition

- Can also open in neighborhoods in cluster 1 with moderate competition if have unique selling propositions to stand out from the competition

- Avoid neighborhoods in cluster 0, already high concentration of Gyms and intense competition

Conclusion

- Answer to business question: The neighborhoods in cluster 2 are the most preferred locations to open a new Gym

- 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 Gym