

Capstone Project

IBM Applied Data Science Capstone

- Opening a New Gym in the City of Hayward, California -

By: Ashneel Kumar

March 2020

Business Problem

.Where would you recommend an investor/developer in the city of Hayward to open a brand new Gym?

Data Required

- List of all the gyms in the city of Hayward
- Latitude and longitude coordinates of those gyms.
- More in depth data involving all the gyms in the city of Hayward This data will be used to preform the clustering on the neighborhood.

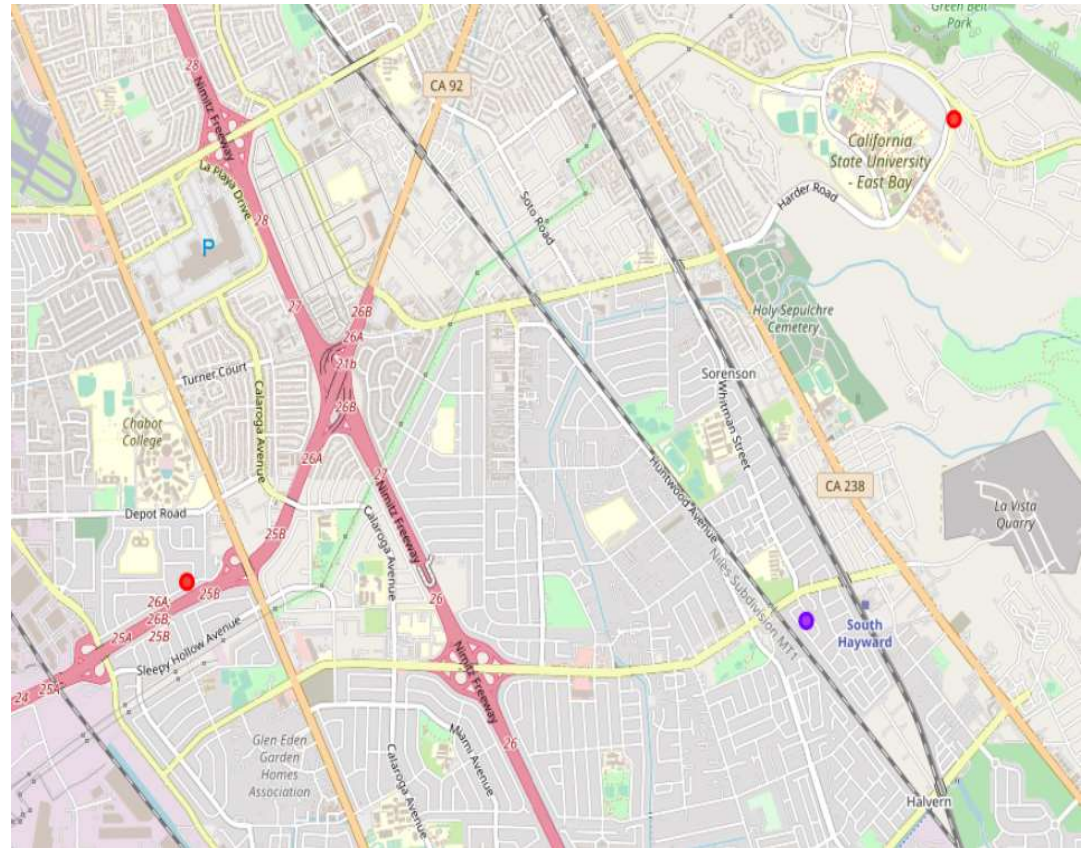
Methodology

1. Web scraping
2. Get latitude and longitude
3. Use Foursquare API
4. Group Data Together
5. Filter Data by category
6. Perform Clustering on data using K-Means Clustering
7. Visualize the clusters using folium.

Results

Categorized the neighborhood in 2 clusters:

- ❖ Cluster 0: Neighborhood with highest numbers of Gym's.
- ❖ Cluster 1: Neighborhood with smaller number of Gym's.



Observation

- Cluster 1 checks all the boxes
- Avoid Cluster 0 which already has high concentration of gyms and suffering from intense competition

Future Research

Two factors were considered in this project which includes location and frequency of occurrence of gym's in the different neighborhood in Hayward, CA.

For Future research I would recommend these extra steps to come to a different analysis and decision making.

Factor in:

- Income of Residents
- Population
- Crime Rate

Conclusion

- ❑ Business question: Where would you recommend an invest/developer in the city of hayward to open a brand-new gym?
- ❑ Answer to the business question: Cluster 1 is the most preferred location to open a new gym.

Thank You

