



Coursera Capstone

OPENING A NEW BAR IN MUMBAI CITY, INDIA

Business Problem

1. Location of the Bar is one of the most important decisions that will determine whether the Bar will be a success or a failure
2. Objective: Use Machine Learning to find Optimal location to open a new Bar.
3. Business question :
 - In the city of Mumbai, if an investor wants to open a new Bar, where should he open it ?

Data

▶ Data required :

- List of neighborhoods in Mumbai.
- Latitude and longitude coordinates of the neighborhoods.
- Venue data, particularly data related to bars.

▶ Sources of Data:

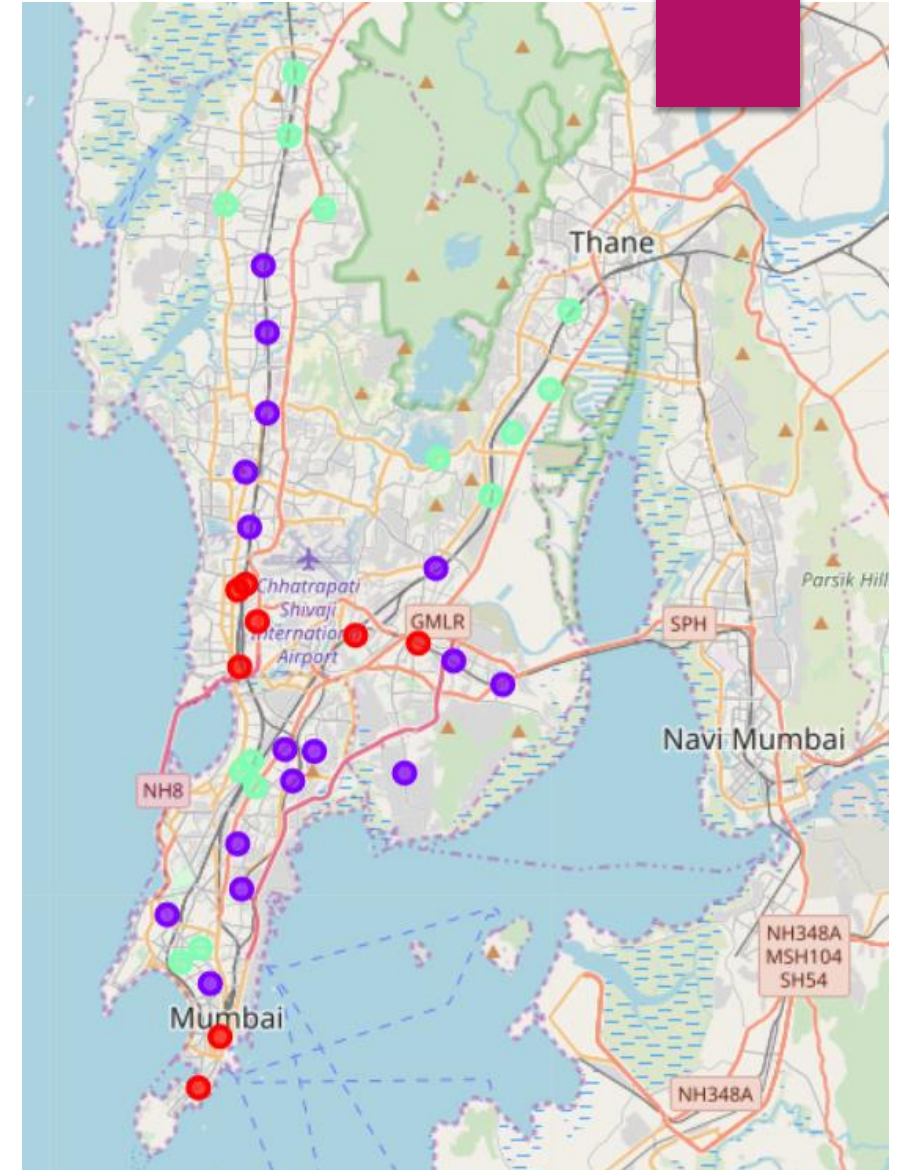
- Wikipedia - https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Mumbai
- Foursquare API for venue data

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 Bars
- ▶ 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: Neighborhoods with low number to no existence of bars.
 - Cluster 1: Neighborhoods with high concentration of bars.
 - Cluster 2: Neighborhoods with moderate number of bars.



Discussion

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

Recommendations

- ▶ Open new bars in neighborhoods in cluster 0 with little to no competition.
- ▶ Can also open in neighborhoods in cluster 2 with moderate competition if have unique selling propositions to stand out from the competition.
- ▶ Avoid neighborhoods in cluster 1, already high concentration of bars and intense competition

Conclusion

- ▶ Answer to business question: The neighborhoods in cluster 0 are the most preferred locations to open a new bar.
- ▶ 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 bar.