

Coursera Capstone

IBM Applied Data Science Capstone

Opening a new café in Mumbai

By: Adwait Gore

Business Problem

- Location of the café is one of the most important decisions that will determine whether the cafe will be a success or a failure
- Objective: To analyse and select the best locations in the city of Mumbai, to open a new café
- This project is timely as the city is currently suffering from oversupply of cafes.
- Business question
 - In the city of Mumbai, if businessman is looking to open a new café, where would you recommend that they open it?

Data

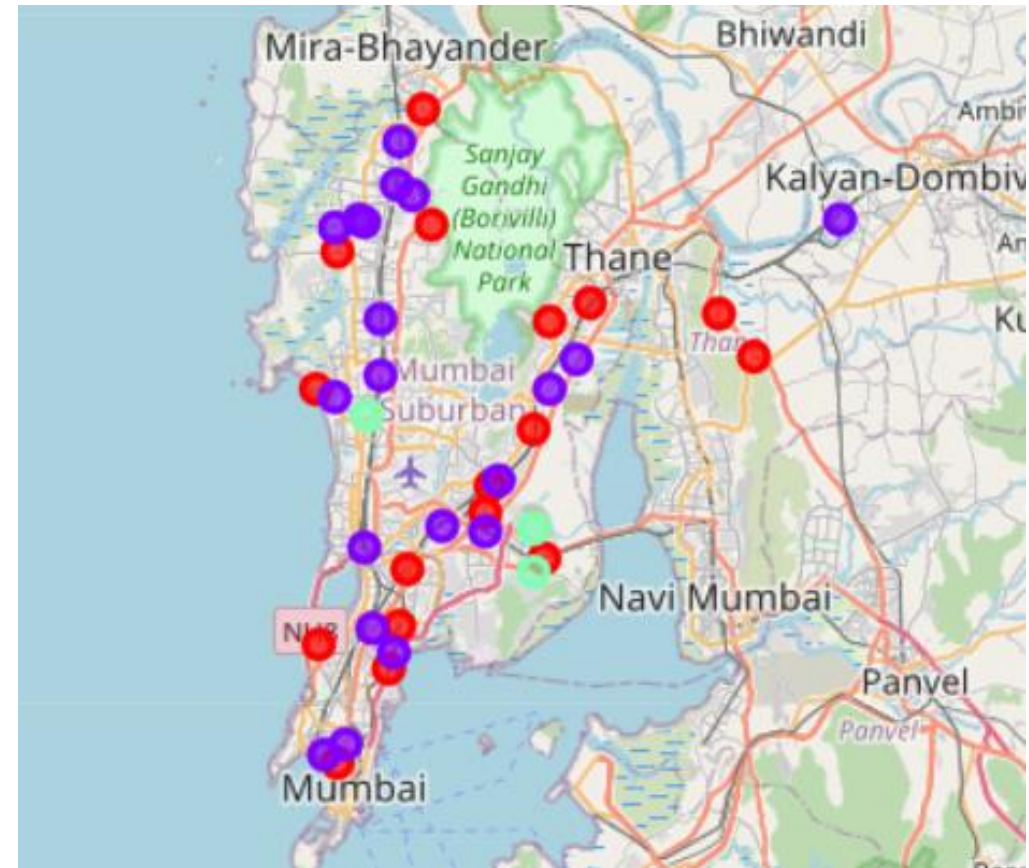
- Data required
 - List of neighbourhoods in Mumbai
 - Latitude and longitude coordinates of the neighbourhoods
 - Venue data, particularly data related to cafes
- Sources of data
 - Wikipedia page for neighbourhoods
(https://en.wikipedia.org/wiki/Category:Suburbs_of_Mumbai)
Geocoder package for latitude 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 cafe
- 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 moderate number of cafes
 - Cluster 1: Neighbourhoods with highest number of cafes
 - Cluster 2: Neighbourhoods with low number to zero existence of cafes



Discussion

- Most of the cafes are concentrated in the central area of the city.
- Highest number in cluster 1 and moderate number in cluster 0.
- Cluster 2 has very low number to no cafes in the neighbourhoods.

Recommendations

- Open new cafes in neighbourhoods in cluster 2 with little to no 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 1, already high concentration of cafes and intense competition

Conclusion

- Answer to business question: The neighbourhoods in cluster 2 are the most preferred locations to open a new cafe.
- 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 cafe.

[illegible]