Coursera Capstone
IBM Applied Data Science Capstone

Opening a New Restaurant in Mumbai, India

Business Problem

- Location of the restaurant is one of the most important decisions that will determine whether the restaurant will be a success or a failure
- Objective: To analyse and select the best locations in the city of Mumbai, India to open a new restaurant
- Business question: In the city of Mumbai, India, if a person is looking to open a new restaurant, where would you recommend that they open it?

Data

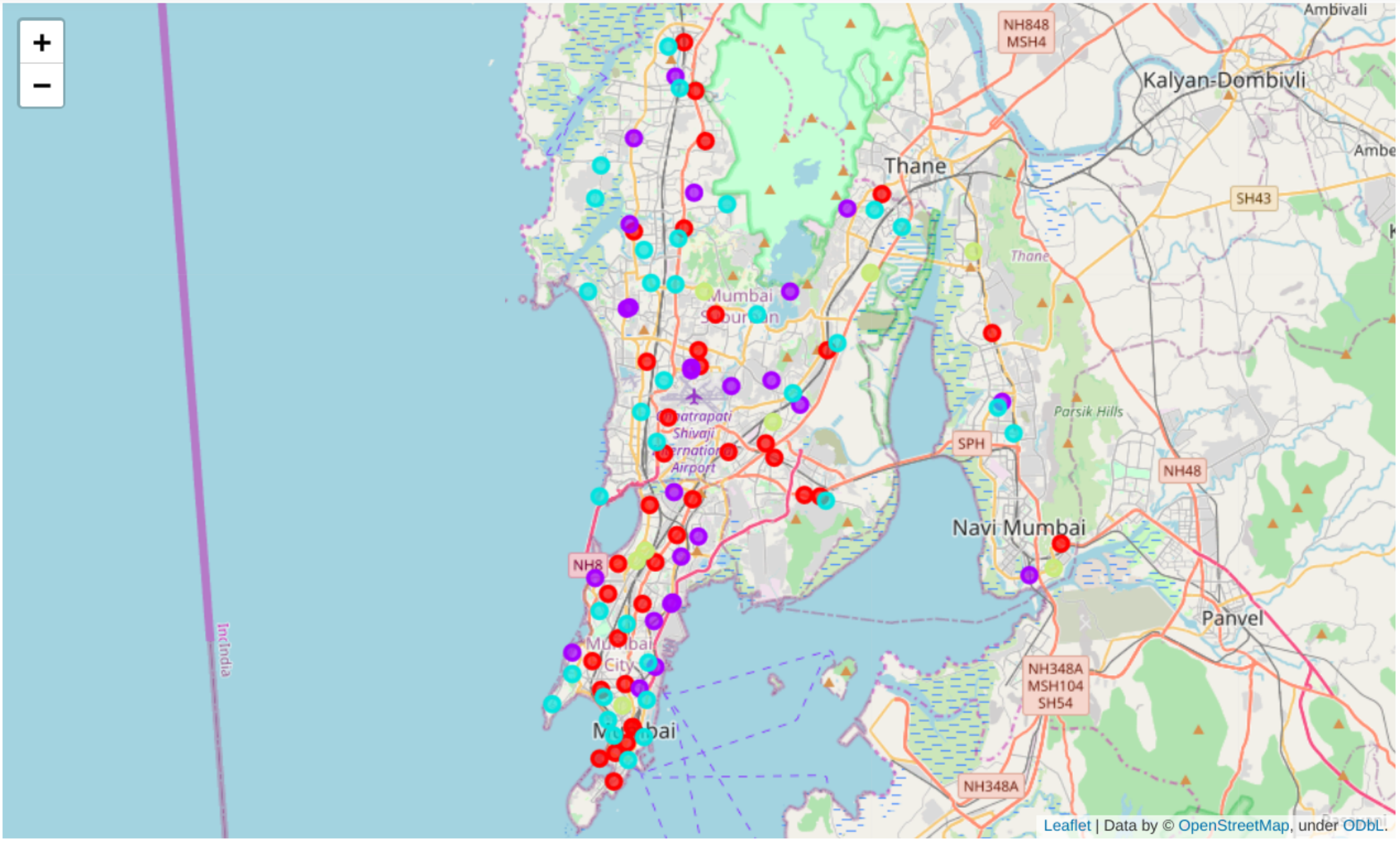
- Data required
  - List of neighbourhoods in Mumbai
  - Latitude and longitude coordinates of the neighbourhoods
  - Venue data, particularly data related to restaurants
- Sources of data
  - Html page for neighbourhoods ("https://mumbai7.com/postal-codes-in-mumbai/")
  - Geocoder package for latitude and longitude coordinates
  - Foursquare API for venue data

Methodology

- Web scraping html 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 Restaurant
- Perform clustering on the data by using k-means clustering
- Visualize the clusters in a map using Folium

Results

- Categorized the neighbourhoods into 4 clusters:
  - Cluster 0(Red): Neighbourhoods with 30 to 50 percent of venues being restaurants
  - Cluster 1(Purple): Neighbourhoods with low number to no existence of restaurants
  - Cluster 2(Light Blue): Neighbourhoods with 15 to 30 percent of venues being restaurants
  - Cluster 3(Dark Yellow): Neighbourhoods with high concentration of restaurants


Ambivali
NH848
MSH4
Kalyan-Dombivli
Thane
SH43
Mumbai
Suburban
Parsik Hills
SPH
NH48
Navi Mumbai
NH8
Panvel
NH348A
MSH104
SH54
NH348A
Leaflet | Data by © OpenStreetMap, under ODbL.

Discussion

- Most of the restaurants are concentrated in the central area as well as area near to the sea in the city
- Highest number in cluster 3 and moderate number in cluster 0
- Cluster 1 has very low number to no shopping mall in the neighbourhoods
- The results show that most of the restaurants are in the central area of the city, with the suburb area still have very few restaurants

Recommendations

- Open new shopping malls in neighbourhoods in cluster 1 with little to no competition
- Can also open restaurants in neighbourhoods in cluster 2 with low competition if have unique propositions to stand out from the competition
- Avoid neighbourhoods in cluster 3, already high concentration of restaurants and intense competition

Conclusion

- Answer to business question:
  - The neighbourhoods in cluster 1 are the most preferred locations to open a new shopping mall
- 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 restaurant

Thank You