Capstone Project

IBM Data Science Specialization

Opening a new Chinese Restaurant in Mumbai, India

By: Siddhartha Patra

July 2019

Business Problem

- Location of the Chinese restaurant is one of the most important decisions that will determine whether the restaurant will be a success or a failure
- Objective: To analyze and select the best locations in the city of Mumbai, India to open a new Chinese restaurant
- This project is timely as the city is currently suffering from oversupply of Chinese restaurants
- Business question
  - In the city of Mumbai, India, if a business owner is looking to open or invest in a new Chinese 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 chinese restuarants

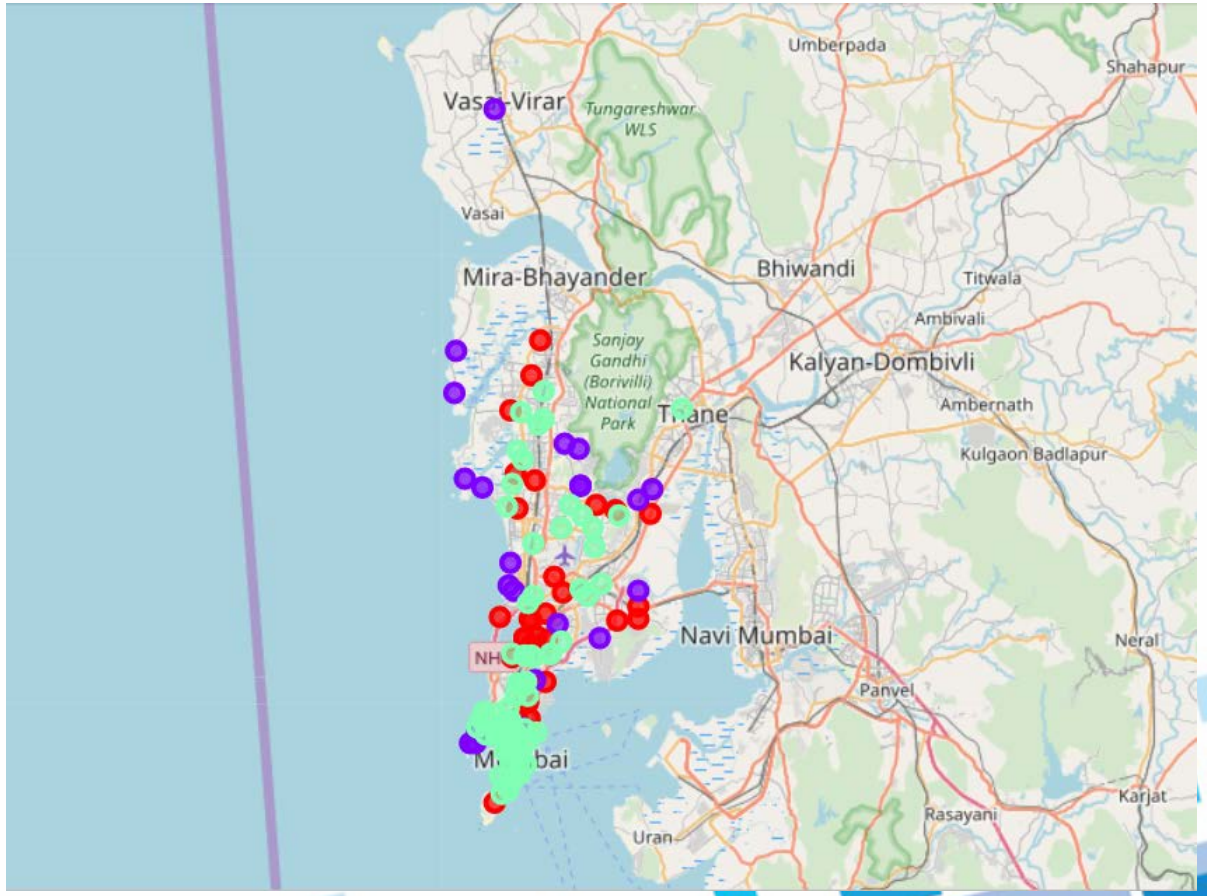
- Sources of data
  - Wikipedia page for neighbourhoods (https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Mumbai)
  - Geocoder package for latitude and longitude coordinates
  - Foursquare API for venue data

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 Shopping Mall
- 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 moderate number of Chinese restaurants
  - ➢Cluster 1: Neighborhoods with low number to no existence of Chinese restaurants
  - ➢Cluster 2: Neighborhoods with high concentration of Chinese restaurants



Discussion

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

Recommendations

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

Conclusion

- Answer to business question: The neighborhoods in cluster 1 are the most preferred locations to open a new Chinese restaurant
- 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 Chinese Restaurant

Thank you!