

Segmenting and Clustering Neighborhoods of Mumbai City to find Optimal Location for Opening an Indian Restaurant

Maitri A. Prajapati – Capstone Final project



Introduction

- Mumbai, formerly known as Bombay, is the capital city of Maharashtra. The Greater Mumbai area occupies a long, narrow peninsula in the Arabian Sea on the west coast of India. Mumbai is the most populous city in India and in the top five in the world. While the 2011 census estimated population at 12.4 million. Mumbai is the financial center, economic powerhouse, and industrial hub of India. In 2018, a report by New World Wealth ranked the city as the 12th wealthiest global city with a total private wealth of US Dollar 950 billion.
- The Mumbai city is one of the world's top centers of commerce in terms of financial flow. It is also home to important financial institutions, such as the Reserve Bank of India, the Bombay Stock Exchange, the National Stock Exchange of India, and corporate headquarters of many Indian companies and multinational corporations.

Description of problem

- Mumbai is rich in its Indian cuisine. It has huge variety of Indian foods from different parts of India. Mumbai seems to be a better place for setting up an Indian restaurant. Since there are lots of restaurants in Mumbai we will try to detect locations that are not already crowded with restaurants. We are also particularly interested in areas with no Indian restaurants in vicinity.

Target Audience

- Companies or Individuals looking into opening a restaurant would be interested in prediction of optimal location in Mumbai City. This project will provide an analysis whether the venture is feasible or not.

Data acquisition

- Web scraping link:

https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Mumbai

The city of **Mumbai** consists of two distinct regions: **Mumbai City** district and **Mumbai Suburban** district, which form two separate revenue districts of Maharashtra.^[1] The city district region is also commonly referred to as the *Island City* or *South Mumbai*.^[2] Mumbai Suburban district lies to the north of Mumbai City district and comprises all of Mumbai's suburbs. The western part of the Mumbai Suburban district forms the **Western Suburbs** and the eastern portion forms the **Eastern Suburbs**. The suburbs of **Chembur**, **Govandi**, **Mankhurd** and **Trombay** lie to the south-east of the Eastern Suburbs. These suburbs are generally not considered as part of the Eastern Suburbs and are sometimes referred to as the "Harbour Suburbs".

The total area of Mumbai is 603.4 km² (233 sq mi).^[3] Of this, the island city spans 67.79 km² (26 sq mi), while the suburban district spans 370 km² (143 sq mi), together accounting for 437.71 km² (169 sq mi) under the administration of **Brihanmumbai Municipal Corporation (BMC)**. The remaining area belongs to Defence, Mumbai Port Trust, Atomic Energy Commission and Borivali National Park, which are out of the jurisdiction of the BMC.^[4] Mumbai lies at the mouth of the **Ulhas River** on the western coast of India, in the coastal region known as the **Konkan**. It sits on **Salsette Island**, partially shared with the **Thane district**.^[5] Mumbai is bounded by the **Arabian Sea** to the west.^[6] Borivali National Park is located partly in the Mumbai suburban district, and partly in the Thane district, and it extends over an area of 103.09 km² (39.80 sq mi).^[7]

Contents [hide]

1 Western Suburbs

- 1.1 Andheri
- 1.2 Mira-Bhayandar
- 1.3 Bandra
- 1.4 Dharavi

Activate Windows
Go to Settings to activate Windows.
Show all

Capstone Final P....ipynb Maitri_Capstone....ipynb

Type here to search

11:14:24 PM 5/29/2021

Data cleaning

- Data downloaded from Wikipedia website contains Borough, their Neighborhoods and latitude and longitude coordinates of each neighborhood. The details of data cleaning methods are given below.
- The names of some neighborhoods/borough were found wrong and therefore, such names were corrected.
- I found some wrong coordinates and therefore I verified this information using geopy package of python. Most of the coordinates received using above python package differed from the coordinates given in the website. Therefore, I replaced the coordinates of the dataframe.

Primary Dataframe after data cleaning

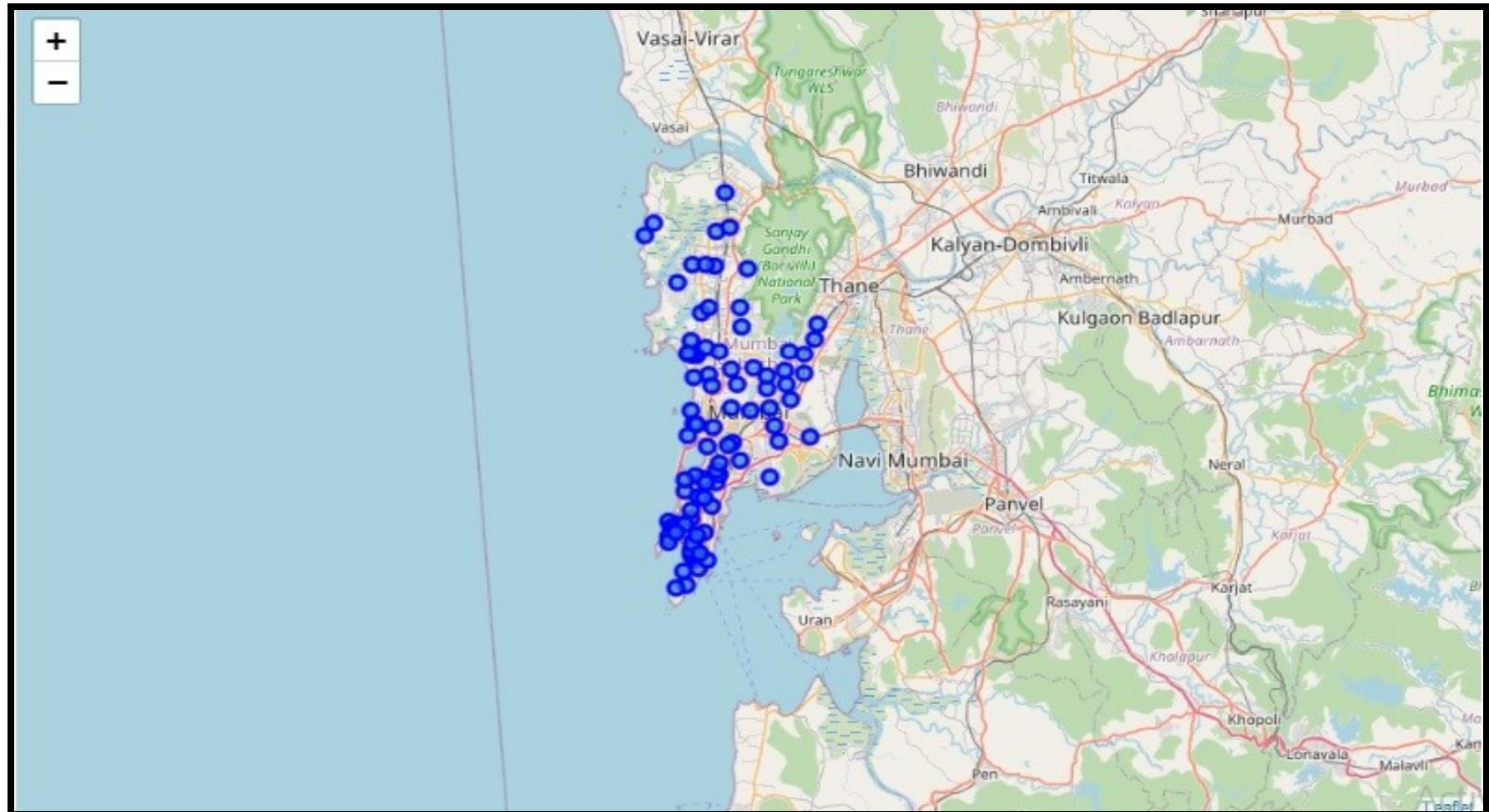
```
In [87]: Mumbai = Mumbai.assign(Latitude=latitude, Longitude=longitude)
          print(Mumbai.shape)
          Mumbai.head()
```

(93, 4)

Out[87]:

	Neighbourhood	Borough	Latitude	Longitude
0	Amboli	Andheri,Western Suburbs	19.07599	72.877393
1	Chakala	Andheri,Western Suburbs	19.07599	72.877393
2	D.N. Nagar	Andheri,Western Suburbs	19.07599	72.877393
3	Four Bungalows	Andheri,Western Suburbs	19.07599	72.877393
4	Lokhandwala	Andheri,Western Suburbs	19.07599	72.877393

Visualization of Mumbai Neighborhoods



Used python folium library

Mumbai Neighborhoods Segmentation

Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	
0	Amboli	19.132010	72.849864	5 Spice , Bandra	19.130421	72.847206	Chinese Restaurant
1	Amboli	19.132010	72.849864	Domino's Pizza	19.131000	72.848000	Pizza Place
2	Amboli	19.132010	72.849864	Cafe Arfa	19.128930	72.847140	Indian Restaurant
3	Amboli	19.132010	72.849864	Bostan Restaurant	19.135898	72.847581	Asian Restaurant
4	Chakala	19.115287	72.861808	Courtyard Mumbai International Airport	19.114167	72.864131	Hotel
...	
301	Thane	19.076000	72.878700	Sahara Restaurant	19.079532	72.880152	Mughlai Restaurant
302	Thane	19.076000	72.878700	Nawab Sheek Corner	19.076933	72.878260	Middle Eastern Restaurant
303	Thane	19.076000	72.878700	Vigour gym	19.075981	72.877656	Gym
304	Thane	19.076000	72.878700	Mithi Nadi	19.076005	72.874680	River
305	Thane	19.076000	72.878700	Sahara	19.079696	72.880148	Activate Vindu Diner

Top five most common venues of Mumbai Neighborhoods

Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0 Aarey Milk Colony	Fast Food Restaurant	Women's Store	French Restaurant	Food Truck	Food Court
1 Agripada	Coffee Shop	Soccer Field	Platform	Athletics & Sports	Bakery
2 Altamount Road	Café	Coffee Shop	Restaurant	Salon / Barbershop	Bookstore
3 Amboli	Indian Restaurant	Asian Restaurant	Pizza Place	Chinese Restaurant	Field
4 Amrut Nagar	Indian Restaurant	Electronics Store	Steakhouse	Restaurant	Fast Food Restaurant
...
86 Vidyavihar	Fast Food Restaurant	Restaurant	Cricket Ground	Juice Bar	Comedy Club
87 Vikhroli	Café	Restaurant	Racetrack	Vegetarian / Vegan Restaurant	Seafood Restaurant
88 Vile Parle	Indian Restaurant	Fast Food Restaurant	Sandwich Place	Café	Tea Room
89 Walkeshwar	Ice Cream Shop	Restaurant	Food & Drink Shop	Park	Food Truck
90 Worli	Scenic Lookout	Yoga Studio	Paper / Office Supplies Store	Garden	Sports Club

91 rows × 6 columns

Activate Wind
Go to Settings to ↗

176 unique categories returned by Foursquare API

Clustering of Mumbai Neighborhoods Venue Categories

	Neighbourhood	Borough	Latitude	Longitude	Labels	Common Venue	Common Venue	Common Venue	Common Venue	Common Venue
0	Amboli	Andheri,Western Suburbs	19.132010	72.849864	0.0	Indian Restaurant	Asian Restaurant	Pizza Place	Chinese Restaurant	Field
1	Chakala	Andheri,Western Suburbs	19.115287	72.861808	4.0	Hotel	Restaurant	Café	Multiplex	Fast Food Restaurant
2	D.N. Nagar	Andheri,Western Suburbs	19.128292	72.830193	4.0	Indian Restaurant	Coffee Shop	Women's Store	Gym / Fitness Center	Liquor Store
3	Four Bungalows	Andheri,Western Suburbs	19.128794	72.825554	4.0	Department Store	Indian Restaurant	Bar	Women's Store	Juice Bar
4	Lokhandwala	Andheri,Western Suburbs	19.143189	72.824081	0.0	Indian Restaurant	Café	Bakery	Pizza Place	Cupcake Shop
...
88	Parel	South Mumbai	19.009482	72.837661	4.0	Bakery	Coffee Shop	Restaurant	Dessert Shop	Movie Theater
89	Gowalia Tank	Tardeo,South Mumbai	18.962095	72.810098	4.0	Café	Bookstore	Gift Shop	Salon / Barbershop	Gastropub
90	Cama Hospital	South Mumbai	18.942041	72.832259	4.0	Café	Bar	Coffee Shop	Indian Restaurant	Multiplex
91	Dharavi	Mumbai	19.044463	72.858618	4.0	Indian Restaurant	Snack Place	Café	Gym / Fitness Center	Fast Food Restaurant
92	Thane	Mumbai	19.076000	72.878700	0.0	Indian Restaurant	Diner	Middle Eastern Restaurant	River	Pizza Place

93 rows x 10 columns

Using K-Means Clustering Method

Examining Five Clusters

Cluster 0

Neighbourhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0 Amboli	0.0	Indian Restaurant	Asian Restaurant	Pizza Place	Chinese Restaurant	Field
4 Lokhandwala	0.0	Indian Restaurant	Café	Bakery	Pizza Place	Cupcake Shop
5 Marol	0.0	Indian Restaurant	Hotel	Diner	Bus Station	Snack Place
6 Sahar	0.0	Indian Restaurant	Hotel Bar	Hotel	Gym	Farmers Market
3 Kherwadi	0.0	Indian Restaurant	Pizza Place	Café	Tunnel	Gym / Fitness Center
5 I.C. Colony	0.0	Indian Restaurant	Chinese Restaurant	Bar	Liquor Store	Food Truck
7 Dahisar	0.0	Indian Restaurant	Diner	Restaurant	Pizza Place	Train Station
5 Thakur village	0.0	Ice Cream Shop	Indian Restaurant	Pizza Place	Gym	General Entertainment
1 Naigaon	0.0	Indian Restaurant	Gym	Historic Site	Grocery Store	Food Court
5 Vile Parle	0.0	Indian Restaurant	Fast Food Restaurant	Sandwich Place	Café	Tea Room
6 Bhandup	0.0	Indian Restaurant	Train Station	Jewelry Store	Track Stadium	Dessert Shop
7 Amrut Nagar	0.0	Indian Restaurant	Electronics Store	Steakhouse	Restaurant	Fast Food Restaurant
1 Nehru Nagar	0.0	Ice Cream Shop	Indian Restaurant	Playground	Fast Food Restaurant	Seafood Restaurant
2 Nahur	0.0	Indian Restaurant	Bus Station	Restaurant	Ice Cream Shop	Antique Shop
4 Bhuleshwar	0.0	Indian Restaurant	Fast Food Restaurant	Hotel	Arcade	Vegetarian / Vegan Restaurant
0 Cuffe Parade	0.0	Indian Restaurant	Spa	Italian Restaurant	Thai Restaurant	Hotel
4 Dongri	0.0	Indian Restaurant	Smoke Shop	Bar	Field	Food Truck
9 Mahim	0.0	Indian Restaurant	Dessert Shop	Fast Food Restaurant	Intersection	Arcade
9 CGS Colony	0.0	Pizza Place	Yoga Studio	Fast Food Restaurant	Food Truck	Food Court
4 Bhuleshwar	0.0	Indian Restaurant	Fast Food Restaurant	Hotel	Arcade	Vegetarian / Vegan Restaurant
5 Fanas Wadi	0.0	Indian Restaurant	Fast Food Restaurant	Food	Vegetarian / Vegan Restaurant	Jewelry Store
6 Chor Bazaar	0.0	Indian Restaurant	Dessert Shop	BBQ Joint	Market	Restaurant
2 Thane	0.0	Indian Restaurant	Diner	Middle Eastern Restaurant	River	Pizza Place

Examining Five Clusters

Cluster 1

Neighbourhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
51	Mahul	1.0	ATM	Women's Store	French Restaurant	Food Truck

Cluster 2

Neighbourhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
18	Aarey Milk Colony	2.0	Fast Food Restaurant	Women's Store	French Restaurant	Food Truck
46	Vidyavihar	2.0	Fast Food Restaurant	Restaurant	Cricket Ground	Juice Bar

Cluster 3

Neighbourhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
81	Navy Nagar	3.0	Garden	Women's Store	French Restaurant	Food Truck

Examining Five Clusters

Cluster 4

Neighbourhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
1 Chakala	4.0	Hotel	Restaurant	Café	Multiplex	Fast Food Restaurant
2 D.N. Nagar	4.0	Indian Restaurant	Coffee Shop	Women's Store	Gym / Fitness Center	Liquor Store
3 Four Bungalows	4.0	Department Store	Indian Restaurant	Bar	Women's Store	Juice Bar
7 Seven Bungalows	4.0	Pub	Café	Chinese Restaurant	Coffee Shop	Bistro
8 Versova	4.0	Pub	Chinese Restaurant	Café	Coffee Shop	Bar
...
87 Matunga	4.0	Indian Restaurant	Vegetarian / Vegan Restaurant	Bar	Snack Place	Café
88 Parel	4.0	Bakery	Coffee Shop	Restaurant	Dessert Shop	Movie Theater
89 Gowalia Tank	4.0	Café	Bookstore	Gift Shop	Salon / Barbershop	Gastropub
90 Cama Hospital	4.0	Café	Bar	Coffee Shop	Indian Restaurant	Multiplex
91 Dharavi	4.0	Indian Restaurant	Snack Place	Café	Gym / Fitness Center	Fast Food Restaurant

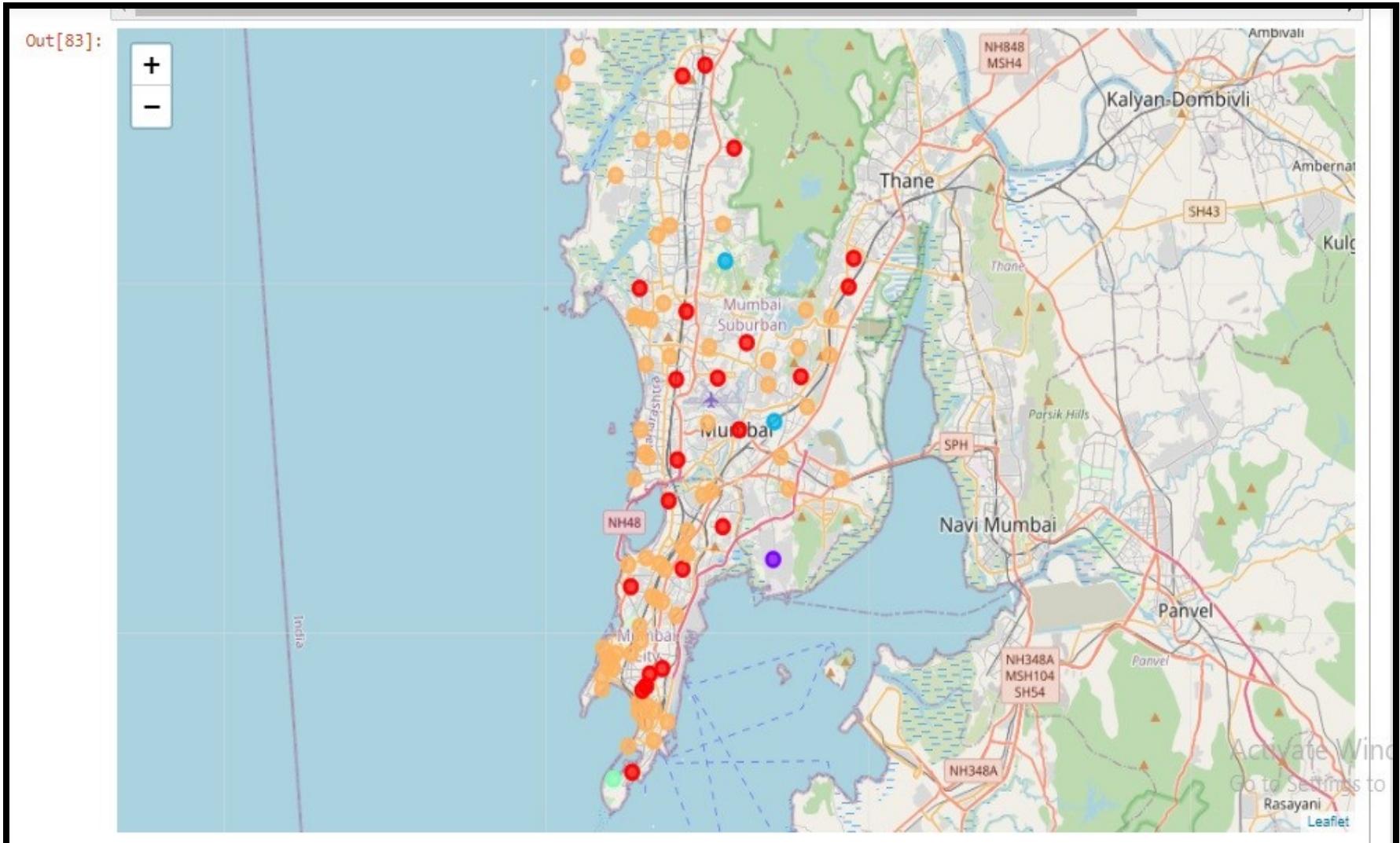
65 rows × 7 columns

Activate Windows

Examining Five Clusters

- We can clearly see that Cluster 4 contains highest number of neighborhoods (65), It contains restaurants in large numbers. In cluster 1, There is only one neighborhood, but it does not contain Indian restaurant top 5 most common venues. Cluster 2 contains only two neighborhoods but Indian Restaurant is not in top most five common venues. Cluster 3 contains only one neighborhood but Indian Restaurant is not in top most five common venues. Finally, cluster 0 contains 23 neighborhoods and Indian Restaurant is the first most common venue in most of the neighborhoods i.e. 20 out of 23 (87%).

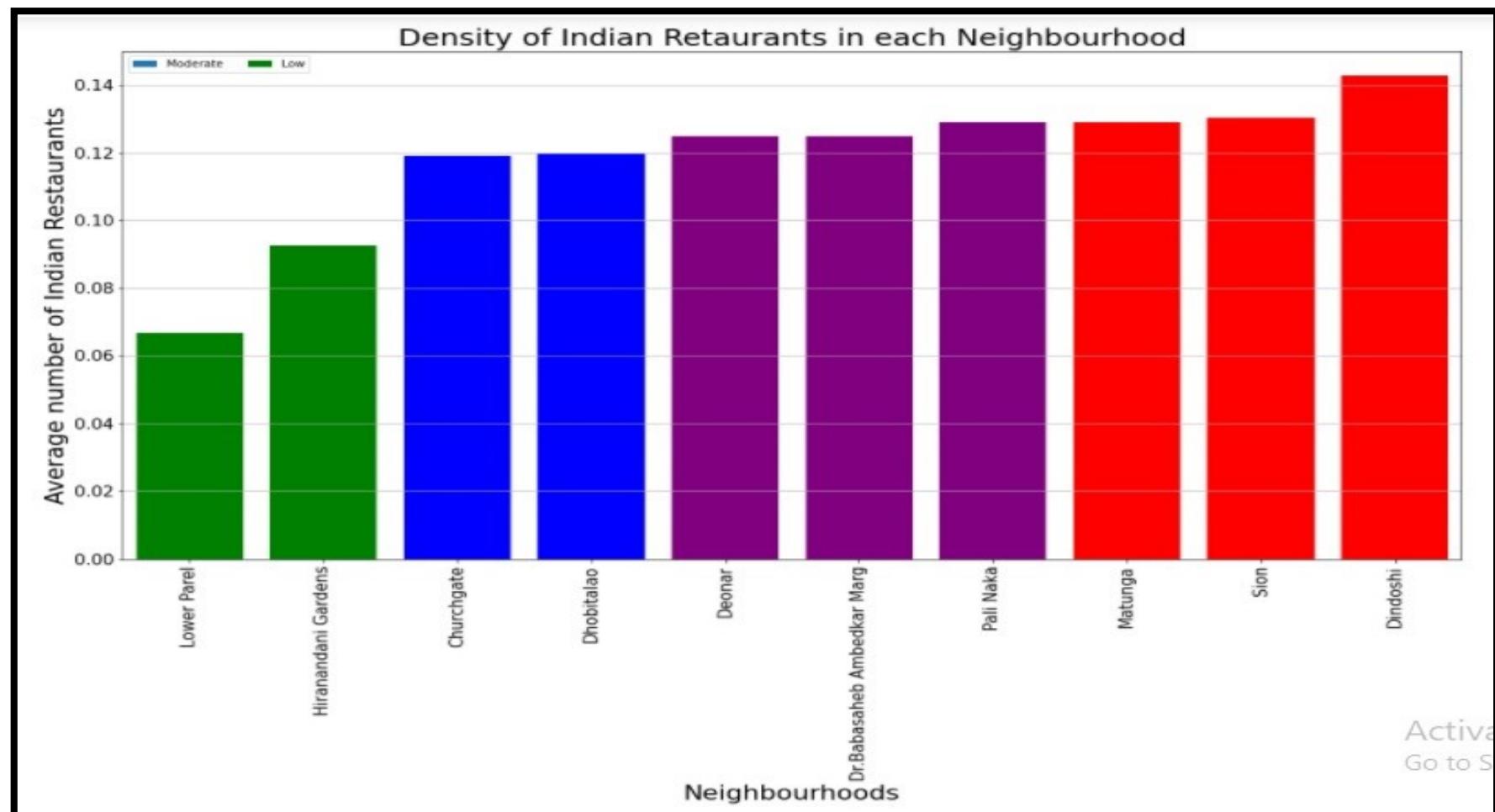
Visualized resulting clusters



10 selected neighborhoods with lowest average number of Indian Restaurants (excluding 0)

Neighbourhood	INDIAN RESTAURANT
Lower Parel	0.066667
Hiranandani Gardens	0.092593
Churchgate	0.119048
Dhobitalao	0.120000
Deonar	0.125000
Dr.Babasaheb Ambedkar Marg	0.125000
Pali Naka	0.129032
Matunga	0.129032
Sion	0.130435
Dindoshi	0.142857

Density of Indian Restaurants in each selected neighborhood



Discussion

From the graph, we can say that the neighborhoods Dindoshi, Sion and Matunga have the highest competition within the 10 selected neighborhoods for opening an Indian Restaurants in Mumbai. This indicating that it has the greatest obstacles in opening a new restaurant among the selected neighborhoods. Dhobitalao and Churchgate; etc. have moderate competition. Following neighborhoods have lowest competition for opening an Indian Restaurant in Mumbai, which will enable a new business to establish easily:

- Lower Parel
- Hiranandani Gardens
- Churchgate
- Dhobitalao

Conclusion

- This project recommends some of the ideal places to open an Indian restaurant in Mumbai, India. The analysis shows there are better scopes for opening restaurant in Lower Parel and Hiranandani Gardens; Dhobitlao and Churchgate are also good location. This analysis can be helpful for the individuals looking for opening a restaurant or expanding business. This analysis shows the feasible venture and competition landscape of the area.
- Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location, levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.