Analyzing the Neighborhoods in Mumbai for Starting a Restaurant

Applied Data Science Capstone Project

By: Dashavidh Dixit

Table of Contents

Introduction

Mumbai is the financial capital of India and is one of the most densely populated cities in the world. It lies on the west coast of India and attracts heavy tourism from all over the globe every year. Personally, I have been brought up in Mumbai and have loved the city from the bottom of my heart. It is one of the major hubs of the world and is extremely diverse with people from various ethnicities residing here. The multi-cultural nature of the city of Mumbai has brought along with it numerous cuisines from all over the world. The people of India generally love food and I personally love to try different cuisines and experience different flavors. Thus, the aim of this project is to study the neighborhoods in Mumbai to determine possible locations for starting a restaurant. This project can be useful for business owners and entrepreneurs who are looking to invest and open a restaurant in Mumbai. The main objective of this project is to carefully analyze appropriate data and find recommendations for the stakeholders.

Data Collection

The following data is required for the project:

1) Neighborhood data of Mumbai
2) Geographical coordinates of Mumbai and all neighborhoods in Mumbai
3) Venue data for neighborhoods in Mumbai

Neighborhoods Data

The data of the neighborhoods in Mumbai was scraped from https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Mumbai. The data is read into a pandas data frame using the read_html() method. The main reason for doing so is that the Wikipedia page provides a comprehensive and detailed table of the data which can easily be scraped using the read_html() method of pandas. The top 10 rows of the dataframe are shown in Figure 1.

	Neighborhood	Location	Latitude	Longitude
0	Amboli	Andheri,Western Suburbs	19.129300	72.843400
1	Chakala, Andheri	Western Suburbs	19.111388	72.860833
2	D.N. Nagar	Andheri,Western Suburbs	19.124085	72.831373
3	Four Bungalows	Andheri,Western Suburbs	19.124714	72.827210
4	Lokhandwala	Andheri,Western Suburbs	19.130815	72.829270
5	Marol	Andheri,Western Suburbs	19.119219	72.882743
6	Sahar	Andheri,Western Suburbs	19.098889	72.867222
7	Seven Bungalows	Andheri,Western Suburbs	19.129052	72.817018
8	Versova	Andheri,Western Suburbs	19.120000	72.820000
9	Mira Road	Mira-Bhayandar,Western Suburbs	19.284167	72.871111

Figure 1: Top 10 rows of Mumbai neighborhoods data scraped from Wikipedia.

Geographical Coordinates

The geographical coordinates for Mumbai has been obtained from the GeoPy library in python. This data is relevant for plotting the map of Mumbai using the Folium library in python. The code for getting the geographical coordinates of Mumbai is shown in Figure 2.

```
address = 'Mumbai, IN'
geolocator = Nominatim()
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geograpical coordinates of Mumbai are {}, {}.'.format(latitude, longitude))
```

```
The geograpical coordinates of Mumbai are 19.0759899, 72.8773928.
```

Figure 2: Obtaining geographical coordinates of Mumbai.

Venue Data

The venue data has been extracted using the Foursquare API. This data contains venue recommendations for all neighborhoods in Mumbai and is used to study the popular venues of different neighborhoods as well as build the unsupervised learning model to cluster neighborhoods. The venue recommendations of all neighborhoods were obtained with a limit of 200, that is, maximum of 200 venue recommendations per neighborhood and a radius of 1 km around the neighborhood's geographical coordinates. Figure 3 shows the top 10 rows depicting the results obtained after cleaning the data from Foursquare API.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Amboli	19.1293	72.84644	Cafe Arfa	19.128930	72.847140	Indian Restaurant
1	Amboli	19.1293	72.84644	5 Spice , Bandra	19.130421	72.847206	Chinese Restaurant
2	Amboli	19.1293	72.84644	Shawarma Factory	19.124591	72.840398	Falafel Restaurant
3	Amboli	19.1293	72.84644	Jaffer Bhai's Delhi Darbar	19.137714	72.845909	Mughlai Restaurant
4	Amboli	19.1293	72.84644	Narayan Sandwich	19.121398	72.850270	Sandwich Place
5	Amboli	19.1293	72.84644	Persia Darbar	19.136952	72.846822	Indian Restaurant
6	Amboli	19.1293	72.84644	Domino's Pizza	19.131000	72.848000	Pizza Place
7	Amboli	19.1293	72.84644	Garden Court	19.127188	72.837478	Indian Restaurant
8	Amboli	19.1293	72.84644	Subway	19.127860	72.844461	Sandwich Place
9	Amboli	19.1293	72.84644	Sarvodaya Veg. Restaurant	19.123760	72.850893	Indian Restaurant

Figure 3: Data obtained from Foursquare API after cleaning.

Methodology

This section provides details for the methodology used in the project.

Data Visualization

In order to understand the data obtained for Mumbai neighborhoods, basic visualization was carried out. Figure 4 shows a bar plot depicting the number of neighborhoods in each location in Mumbai.

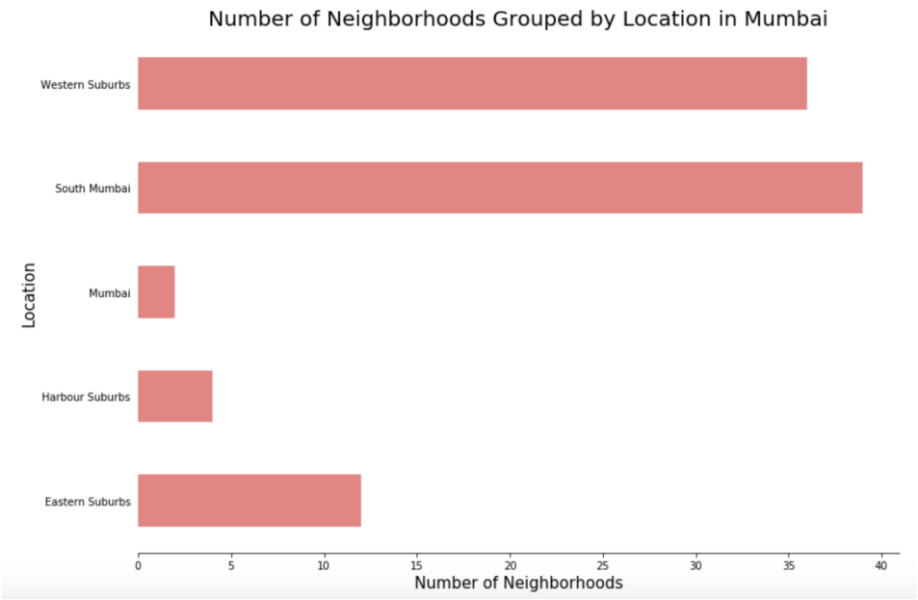

Figure 4: Number of neighborhoods grouped by location.

It is evident from Figure 4 that South Mumbai and Western Suburbs have the most number of neighborhoods. Notice how we see one of the locations as Mumbai itself? This is because the neighborhoods contained in this location are located at the outskirts of the city and thus have been termed as just Mumbai.

Using folium, a map was plotted to show how the different neighborhoods are spread all across Mumbai. This is shown in Figure 5.

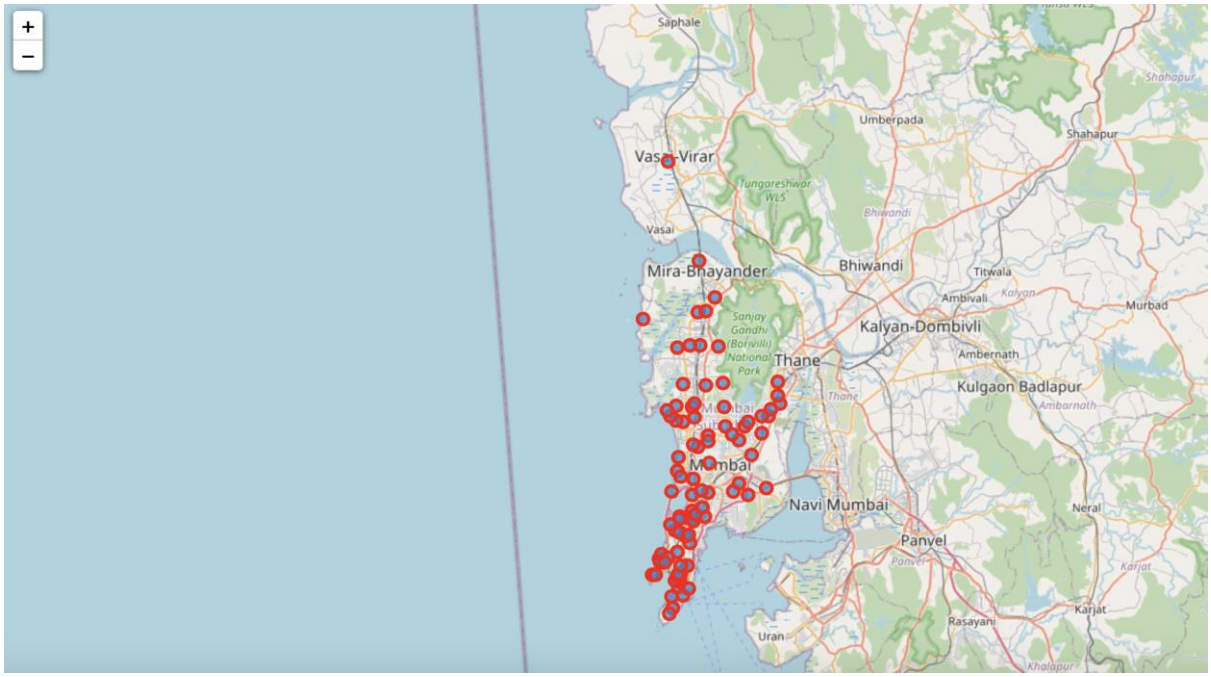

Figure 5: Depicting the neighborhood spread across Mumbai.

Feature Extraction

Feature extraction was carried out to obtain features from the Foursquare API data (as shown in Figure 3) which was used for building the unsupervised learning model. In order to achieve this, the "Venue Category" column had to be converted to some form of numeric value to be used for building the model. This was achieved by the One-hot Encoding method which takes all the unique categories and creates a column for each category. Then, if a neighborhood venue belongs to that category, it would get a value of 1 for that row in that specific category column and if a neighborhood venue does not belong to the particular category, the value would be 0. This process was repeated for all venues in all neighborhoods and the result was a sparse matrix containing the neighborhood name and all unique category columns with either 1 or 0 based on whether the

neighborhood venue belonged to that category or not. This dataframe was then grouped by the neighborhood name and the average value was taken for all categories. The result is shown in Figure 6 which shows only the top 10 rows.

	Neighborhood	ATM	Accessories Store	Airport Terminal	American Restaurant	Antique Shop	Aquarium	Arcade	Art Gallery	Arts & Crafts Store	...	Trail	Train	Train Station	Vegetarian / Vegan Restaurant	Whisky Bar	Wine Bar	Wine Shop	Women's Store	Yoga Studio	Zoo
0	Amboli	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.000000	0.0	0.0	0.000	0.000000	0.0	0.0
1	Chakala, Andheri	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.047619	0.0	0.0	0.000	0.000000	0.0	0.0
2	D.N. Nagar	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.043478	0.0	0.0	0.000	0.021739	0.0	0.0
3	Four Bungalows	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.030303	0.0	0.0	0.000	0.015152	0.0	0.0
4	Lokhandwala	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.010753	0.0	0.0	0.000	0.010753	0.0	0.0
5	Marol	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.000000	0.0	0.0	0.000	0.000000	0.0	0.0
6	Sahar	0.0	0.0	0.033333	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.000000	0.0	0.0	0.000	0.000000	0.0	0.0
7	Seven Bungalows	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.014925	...	0.0	0.0	0.0	0.029851	0.0	0.0	0.000	0.000000	0.0	0.0
8	Versova	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.025000	...	0.0	0.0	0.0	0.000000	0.0	0.0	0.025	0.000000	0.0	0.0
9	Mira Road	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.000000	0.0	0.0	0.000	0.066667	0.0	0.0

10 rows × 221 columns

Figure 6: One-hot Encoding resulting dataframe.

Notice that most of the values are 0 since there were a large number of unique categories and not all neighborhoods had venues belonging to each category. This data was used for the unsupervised learning model with the neighborhood name dropped. The unsupervised learning model is explained in the next section.

A dataframe was also created which contained the top 10 most common venues of all neighborhoods. Though this is not a part of Feature Extraction, it is important to provide a glimpse into what this dataframe looks like as it will be used later to combine the results from the unsupervised learning model. The top 10 rows of this dataframe are shown in Figure 7.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Amboli	Indian Restaurant	Coffee Shop	Bakery	Bar	Asian Restaurant	Pizza Place	Sandwich Place	Bowling Alley	Bus Station	Bike Rental / Bike Share
1	Chakala, Andheri	Hotel	Indian Restaurant	Café	Fast Food Restaurant	Pizza Place	Asian Restaurant	Hotel Bar	Vegetarian / Vegan Restaurant	Restaurant	Gym
2	D.N. Nagar	Bar	Indian Restaurant	Pub	Gym / Fitness Center	Pizza Place	Lounge	Coffee Shop	Vegetarian / Vegan Restaurant	Snack Place	Gym
3	Four Bungalows	Pub	Café	Indian Restaurant	Gym / Fitness Center	Chinese Restaurant	Bar	Seafood Restaurant	Lounge	Vegetarian / Vegan Restaurant	Coffee Shop
4	Lokhandwala	Indian Restaurant	Chinese Restaurant	Café	Pub	Bakery	Bar	Italian Restaurant	Gym / Fitness Center	Coffee Shop	Asian Restaurant
5	Marol	Indian Restaurant	Hotel	Diner	Bakery	Dance Studio	Ice Cream Shop	Chinese Restaurant	Fast Food Restaurant	Restaurant	Lounge
6	Sahar	Hotel	Café	Indian Restaurant	Lounge	Gym	Asian Restaurant	Pizza Place	Seafood Restaurant	Restaurant	Falafel Restaurant
7	Seven Bungalows	Café	Pub	Seafood Restaurant	Chinese Restaurant	Pizza Place	Coffee Shop	Bar	Ice Cream Shop	Asian Restaurant	Bistro
8	Versova	Café	Ice Cream Shop	Beach	Pizza Place	Coffee Shop	Chinese Restaurant	Salon / Barbershop	Frozen Yogurt Shop	Bistro	Sandwich Place
9	Mira Road	Indian Restaurant	Convenience Store	Coffee Shop	Mexican Restaurant	Fast Food Restaurant	Food Truck	Motorcycle Shop	Movie Theater	Basketball Court	Bar

Figure 7: Top 10 most common venues for neighborhoods.

Unsupervised Learning

K-means unsupervised learning technique was used to cluster the neighborhoods based on the category of venues near the neighborhoods. One important aspect of the k-means model is to determine the number of clusters to use in model development. This was determined by the Silhouette score which was calculated for a range of clusters from 2 to 15. The resulting number of clusters and their respective Silhouette scores are shown in Figure 8.

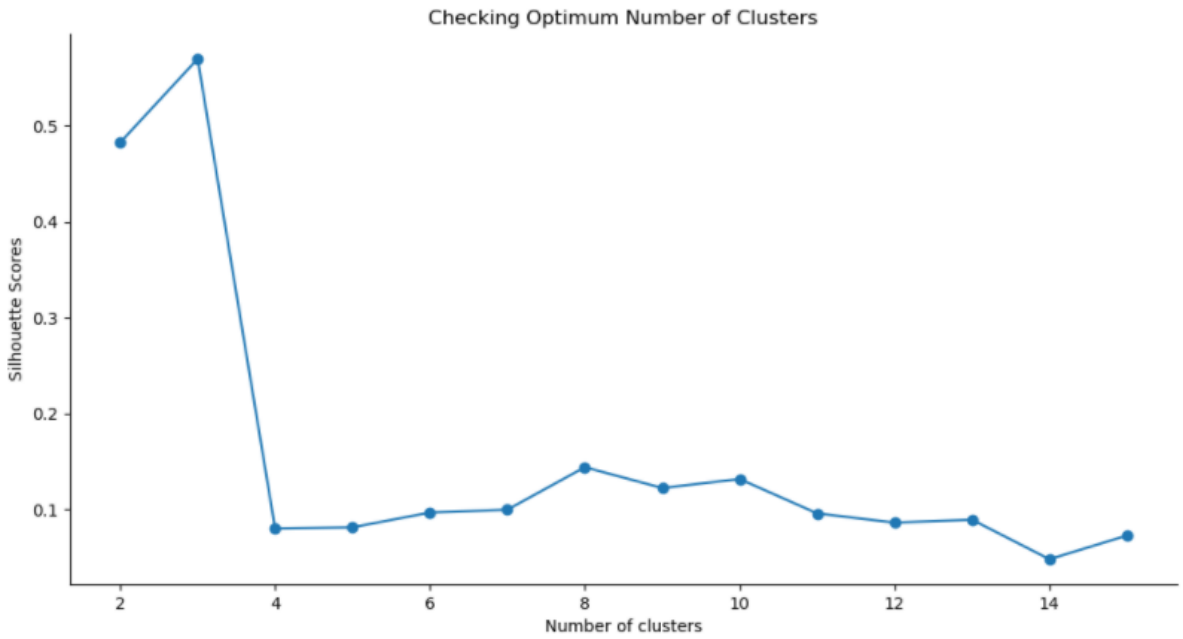

Figure 8: Silhouette scores for different number of clusters.

It is evident that the Silhouette scores are not very high even as the number of clusters increases. This means that the inter-cluster distance is not very high over the range of k-values. Despite this, the data will be clustered to the best possible extent. For this, 5 clusters will be used for the k-means clustering model since it provides the highest silhouette score as seen in Figure 8.

Results

The clustering model then clusters the neighborhoods in Mumbai and provides a label for each neighborhood which is representative of the cluster it belongs to. The cluster labels were then added to the dataframe in Figure 7 along with the

Location, Latitude, and Longitude columns to provide a complete summary of the clustering. The top 10 rows are shown in Figure 9.

	Neighborhood	Location	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Amboli	Western Suburbs	19.1293	72.8464	1	Indian Restaurant	Coffee Shop	Bakery	Bar	Asian Restaurant	Pizza Place	Sandwich Place	Bowling Alley	Bus Station	Bike Rental / Bike Share
1	Chakala, Andheri	Western Suburbs	19.1084	72.8623	1	Hotel	Indian Restaurant	Café	Fast Food Restaurant	Pizza Place	Asian Restaurant	Hotel Bar	Vegetarian / Vegan Restaurant	Restaurant	Gym
2	D.N. Nagar	Western Suburbs	19.1241	72.8325	0	Bar	Indian Restaurant	Pub	Gym / Fitness Center	Pizza Place	Lounge	Coffee Shop	Vegetarian / Vegan Restaurant	Snack Place	Gym
3	Four Bungalows	Western Suburbs	19.1263	72.8243	0	Pub	Café	Indian Restaurant	Gym / Fitness Center	Chinese Restaurant	Bar	Seafood Restaurant	Lounge	Vegetarian / Vegan Restaurant	Coffee Shop
4	Lokhandwala	Western Suburbs	19.1432	72.8249	0	Indian Restaurant	Chinese Restaurant	Café	Pub	Bakery	Bar	Italian Restaurant	Gym / Fitness Center	Coffee Shop	Asian Restaurant
5	Marol	Western Suburbs	19.1192	72.8827	1	Indian Restaurant	Hotel	Diner	Bakery	Dance Studio	Ice Cream Shop	Chinese Restaurant	Fast Food Restaurant	Restaurant	Lounge
6	Sahar	Western Suburbs	19.1027	72.8626	0	Hotel	Café	Indian Restaurant	Lounge	Gym	Asian Restaurant	Pizza Place	Seafood Restaurant	Restaurant	Falafel Restaurant
7	Seven Bungalows	Western Suburbs	19.1315	72.817	0	Café	Pub	Seafood Restaurant	Chinese Restaurant	Pizza Place	Coffee Shop	Bar	Ice Cream Shop	Asian Restaurant	Bistro
8	Versova	Western Suburbs	19.1377	72.8135	0	Café	Ice Cream Shop	Beach	Pizza Place	Coffee Shop	Chinese Restaurant	Salon / Barbershop	Frozen Yogurt Shop	Bistro	Sandwich Place
9	Mira Road	Western Suburbs	19.2657	72.8711	1	Indian Restaurant	Convenience Store	Coffee Shop	Mexican Restaurant	Fast Food Restaurant	Food Truck	Motorcycle Shop	Movie Theater	Basketball Court	Bar

Figure 9: Clustering neighborhoods in Mumbai.

Furthermore, neighborhoods in each individual cluster can be extracted using cluster labels and thus the details of specific clusters can be seen. This is done below for all clusters with only 10 rows for clusters that contain a high number of neighborhoods.

	Neighborhood	Location	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
32	Nalasopara	Western Suburbs	Bar	Bus Station	Zoo	Fast Food Restaurant	Farm	Falafel Restaurant	Factory	Event Space	Electronics Store	Duty-free Shop

Figure 10: Cluster 1.

	Neighborhood	Location	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	
0	Amboli	Western Suburbs	Indian Restaurant	Bar	Pizza Place	Coffee Shop	Asian Restaurant	Pub	Chinese Restaurant	Cocktail Bar	Mughlai Restaurant	Fa R
1	Chakala, Andheri	Western Suburbs	Indian Restaurant	Hotel	Café	Seafood Restaurant	Chinese Restaurant	Multiplex	Pizza Place	Restaurant	Vegetarian / Vegan Restaurant	Fa R
2	D.N. Nagar	Western Suburbs	Bar	Indian Restaurant	Pizza Place	Gym / Fitness Center	Pub	Vegetarian / Vegan Restaurant	Coffee Shop	Women's Store	Lounge	S Pl
3	Four Bungalows	Western Suburbs	Indian Restaurant	Pub	Bar	Coffee Shop	Café	Vegetarian / Vegan Restaurant	Pizza Place	Chinese Restaurant	Gym / Fitness Center	Lo
4	Lokhandwala	Western Suburbs	Bar	Indian Restaurant	Pub	Coffee Shop	Pizza Place	Lounge	Asian Restaurant	Italian Restaurant	Café	M
6	Sahar	Western Suburbs	Coffee Shop	Café	Fast Food Restaurant	Hotel	Airport Service	Indian Restaurant	Airport	Lounge	Airport Food Court	Ai Lo
7	Seven Bungalows	Western Suburbs	Café	Pub	Ice Cream Shop	Coffee Shop	Chinese Restaurant	Bar	Pizza Place	Indian Restaurant	Beach	S R
8	Versova	Western Suburbs	Indian Restaurant	Café	Coffee Shop	Multiplex	Chinese Restaurant	Pub	Donut Shop	Japanese Restaurant	Beach	Sp
9	Mira Road	Western Suburbs	Pizza Place	Indian Restaurant	Gym / Fitness Center	Café	Multiplex	Convenience Store	Asian Restaurant	Sculpture Garden	Sandwich Place	Fa R
12	Bandstand Promenade	Western Suburbs	Coffee Shop	Scenic Lookout	Fast Food Restaurant	Indian Restaurant	Boutique	Boat or Ferry	Lounge	Beach	Café	Ita R
13	Kherwadi	Western Suburbs	Café	Indian Restaurant	Bakery	Bar	Coffee Shop	Chinese Restaurant	Seafood Restaurant	Pizza Place	Snack Place	D B
14	Pali Hill	Western Suburbs	Indian Restaurant	Fast Food Restaurant	Bakery	Café	Dessert Shop	Bar	Asian Restaurant	Seafood Restaurant	Cupcake Shop	Ic Sl
15	I.C. Colony	Western Suburbs	Indian Restaurant	Chinese Restaurant	Bar	Dessert Shop	Fast Food Restaurant	Park	Ice Cream Shop	Italian Restaurant	Soccer Field	G
17	Dahisar	Western Suburbs	Indian Restaurant	Bar	Restaurant	Bakery	Café	Juice Bar	Chinese Restaurant	Diner	Department Store	Ic Sl
18	Aarey Milk	Western	Gym / Fitness	Hotel	Café	Resort	Indian	Restaurant	Farm	Golf Course	Dhaba	Fa

Figure 11: Cluster 2.

	Neighborhood	Location	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
68	Mahalaxmi	South Mumbai	Arcade	Zoo	Dim Sum Restaurant	Farm	Falafel Restaurant	Factory	Event Space	Electronics Store	Duty-free Shop	Dumpling Restaurant

Figure 12: Cluster 3.

	Neighborhood	Location	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
11	Uttan	Western Suburbs	Beach	Playground	Indian Restaurant	Resort	Bus Station	Zoo	Dim Sum Restaurant	Factory	Event Space	Electronics Store
47	Vikhroli	Eastern Suburbs	Comedy Club	Pool	Bakery	Scenic Lookout	Soccer Field	Dim Sum Restaurant	Factory	Event Space	Electronics Store	Duty-free Shop
60	Cuffe Parade	South Mumbai	Park	Asian Restaurant	Beach	Basketball Court	Department Store	Italian Restaurant	Dive Bar	Farm	Falafel Restaurant	Factory
64	Dongri	South Mumbai	Beach	Playground	Indian Restaurant	Resort	Bus Station	Zoo	Dim Sum Restaurant	Factory	Event Space	Electronics Store
81	Navy Nagar	South Mumbai	Golf Course	Asian Restaurant	Beach	General Entertainment	Dim Sum Restaurant	Farm	Falafel Restaurant	Factory	Event Space	Electronics Store

Figure 13: Cluster 4.

	Neighborhood	Location	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10tl Co
5	Marol	Western Suburbs	Indian Restaurant	Hotel	Chinese Restaurant	Fast Food Restaurant	Diner	Boat or Ferry	Ice Cream Shop	Restaurant	Coffee Shop	Cup Sho
10	Bhayandar	Western Suburbs	Fast Food Restaurant	Indian Restaurant	Multiplex	Burger Joint	Food Truck	Pizza Place	Coffee Shop	Mexican Restaurant	Restaurant	Ligh Stat
16	Gorai	Western Suburbs	Beach	Food	Seafood Restaurant	Indian Restaurant	Zoo	Diner	Falafel Restaurant	Factory	Event Space	Elec Stor
20	Jogeshwari West	Western Suburbs	Indian Restaurant	Vegetarian / Vegan Restaurant	Fast Food Restaurant	Electronics Store	Shop & Service	Restaurant	Light Rail Station	Breakfast Spot	Food Court	Café
24	Mahavir Nagar	Western Suburbs	Indian Restaurant	Fast Food Restaurant	Pizza Place	Food	Coffee Shop	Chinese Restaurant	Electronics Store	Park	Donut Shop	Spo Goo Sho
34	Irla	Western Suburbs	Indian Restaurant	Ice Cream Shop	Café	Snack Place	Seafood Restaurant	Sandwich Place	Park	Coffee Shop	Bar	Fast Rest
38	Asalfa	Eastern Suburbs	Indian Restaurant	Juice Bar	Coffee Shop	Donut Shop	Gym / Fitness Center	Men's Store	Bakery	Factory	Vegetarian / Vegan Restaurant	Ligh Stat
39	Pant Nagar	Eastern Suburbs	Indian Restaurant	Ice Cream Shop	Fast Food Restaurant	Gym / Fitness Center	Coffee Shop	Dessert Shop	Pizza Place	Vegetarian / Vegan Restaurant	Restaurant	Scul Gar
44	Hiranandani Gardens	Eastern Suburbs	Indian Restaurant	Bar	Ice Cream Shop	Dessert Shop	Restaurant	Lounge	Bakery	Chinese Restaurant	Italian Restaurant	Pizz Plac
45	Indian Institute of Technology Bombay campus	Eastern Suburbs	Indian Restaurant	Asian Restaurant	Dance Studio	Building	Café	Food Court	Department Store	Diner	Coffee Shop	Ever Spa
48	Chembur	Harbour Suburbs	Indian Restaurant	Pizza Place	Fast Food Restaurant	Bakery	Snack Place	Café	Ice Cream Shop	Bar	Coffee Shop	Spa
49	Deonar	Harbour Suburbs	Indian Restaurant	Gym	Snack Place	Fast Food Restaurant	Bar	Garden	Pizza Place	Coffee Shop	Pool	Dine
52	Agripada	South Mumbai	Indian Restaurant	Gym	Coffee Shop	Bakery	Asian Restaurant	History Museum	Fast Food Restaurant	Movie Theater	Nightclub	Cup Sho
54	Bhuleshwar	South Mumbai	Indian Restaurant	Chinese Restaurant	Jewelry Store	Café	Market	Bakery	Fast Food Restaurant	Soccer Field	Bar	Indi Mov The

Figure 14: Cluster 5.

Based on the clusters shown above, the neighborhoods can once again be plotted on a map of Mumbai, however, this time with different color markers to distinguish between different clusters. This is shown in Figure 17.

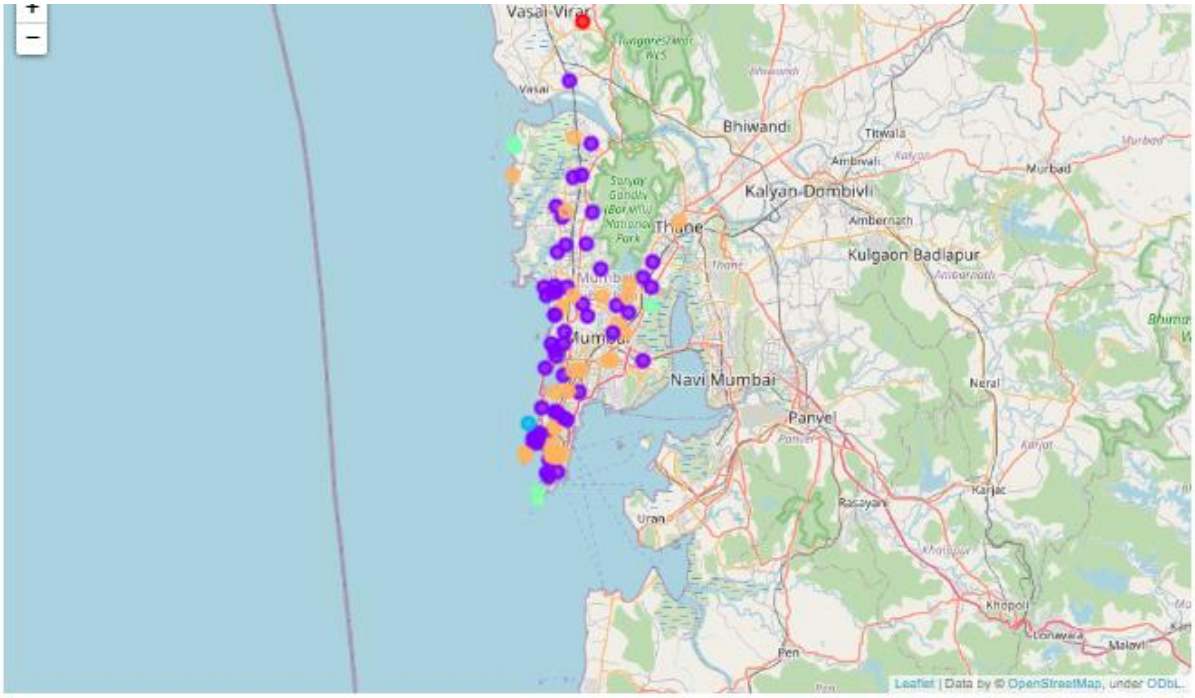

Figure 15: Visualizing the clustering of neighborhoods in Mumbai.

Discussion

By analyzing the five clusters obtained we can see that some of the clusters are more suited for restaurants and hotels, whereas, other clusters are less suited. Neighborhoods in clusters 1, 3, and 4 contain a small percentage of restaurants, hotels, cafe and pubs in their top 10 common venues. These clusters contain a higher degree of other venues like train station, bus station, fish market, gym, performing arts venue and smoke shop, to name a few. Thus, they are not well suited for opening a new restaurant. On the other hand, neighborhoods in clusters 1 and 2 contain a much higher degree of restaurants, hotels, multiplex, cafes, bars and other food joints. Thus, the neighborhoods in these clusters would be well suited for opening a new restaurant.

Conclusion

In this project, the neighborhoods in Mumbai, India have been successfully analyzed for determining which would be the best neighborhoods for opening a new restaurant. Based on the analysis carried out, neighborhoods in cluster 1 are recommended as locations for the new restaurant. This has also been plotted in the map in Figure 18. The stakeholders and investors can further tune this by considering various other factors like transport, legal requirements, and costs associated. These were out of the scope for this project and thus were not considered.