IBM DATA SCIENCE

Applied Data Science Capstone

The Battle of Neighbourhoods (Part-1)

Finding Best Venues, Finding Best Neighbourhoods, Segmenting and Clustering Neighbourhoods in Mumbai City

FINAL REPORT

Introduction

Mumbai is the capital city of the Indian state of Maharashtra. According to UN, as of 2018, Mumbai was the second most populous city in India after Delhi and the seventh most populous city in the world with a population of almost 20 million. In 2008, Mumbai was named an alpha world city. It has the highest number of millionaires and billionaires among all cities in India. Mumbai is home to three UNESCO World Heritage Sites: the Elephanta Caves, Chhatrapati Shivaji Maharaj Terminus, and the city's distinctive ensemble of Victorian and Art Deco buildings.

Mumbai is the financial, commercial and entertainment capital of India. It is also one of the world's top ten centres of commerce in terms of global financial flow, generating 6.16% of India's GDP and accounting for 25% of industrial output, 70% of maritime trade in India (Mumbai Port Trust and JNPT), and 70% of capital transactions to India's economy. Mumbai's billionaires had the highest average wealth of any city in the world in 2008. The city houses important financial institutions and the corporate headquarters of numerous Indian companies and multinational corporations. It is also home to some of India's premier scientific and nuclear institutes. The city also houses India's Hindi (Bollywood) and Marathi cinema industries. Mumbai's business opportunities attract migrants from all over India. Tourists from various countries around the world visit Mumbai.

With such a huge population and many venues for various purposes its often hard to choose the best one. Therefore, this project will help user decide best neighborhoods, best venues, etc.

Problem:

Note: Best is selected for the category of venue given by user as input.

To find the answers to the following questions:

- What are the places of Mumbai City that have top rated venues of the given category?
- What are the top-rated venues of Mumbai City of the given category?
- Which are the best places to stay if you prefer particular type of venues such as Gyms, Pubs, Italian Restaurant, Chinese Restaurant, etc. ?
- How are the neighborhoods around the Mumbai clustered?

Data Section

Mumbai is the financial, commercial and entertainment capital of India. It is also one of the world's top ten centres of commerce in terms of global financial flow, generating 6.16% of India's GDP. Therefore, there are venues of various categories and given below are the sources which will be used to collect data.

Data needed for this project is as follows:

- Mumbai City data that contains list Boroughs, Neighbourhoods along with their Latitude and Longitude.
  - Data source : https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Mumbai
  - Description : This website contains the required information. After cleaning this data, it will be used to explore various neighbourhoods of Mumbai City.

- Venues of certain category given by user, details of venues of certain category and top venues in each neighbourhood of Mumbai City.
  - Data source : Foursquare API
  - Description : By using this API required information can be fetched and after converting it to data frames it can be cleaned and used to get required data.

- Locating the coordinates of addresses and cities.
  - Data source : geopy-geocoders (Third party geocoders and other data sources)
  - Description : This will be used to fetch co-ordinates of locations.

Methodology:

1. Importing necessary libraries and defining functions to get venues of certain type, get details of venue and get venues around neighborhood.

2. Download dataset containing of Postal Code, Boroughs and Neighborhoods from link.

3. Cleaning of dataset obtained above.

	Neighborhood	Borough	Latitude	Longitude
0	Amboli	Western Suburbs	19.129300	72.843400
1	Chakala\tAndheri,	Western Suburbs	19.111388	72.860833
2	D.N. Nagar	Western Suburbs	19.124085	72.831373
3	Four Bungalows	Western Suburbs	19.124714	72.827210
4	Lokhandwala	Western Suburbs	19.130815	72.829270

	Neighborhood	Borough	Latitude	Longitude
88	Parel	South Mumbai	18.990000	72.840000
89	Gowalia Tank	South Mumbai	18.962450	72.809703
90	Dava Bazaar	South Mumbai	18.946882	72.831362
91	Dharavi	Mumbai	19.040208	72.850850
92	Thane	Thane	19.200000	72.970000

4. Fetching co-ordinates of Mumbai using geopy.

5. Visualizing Map of Mumbai City with Neighborhoods superimposed on top.
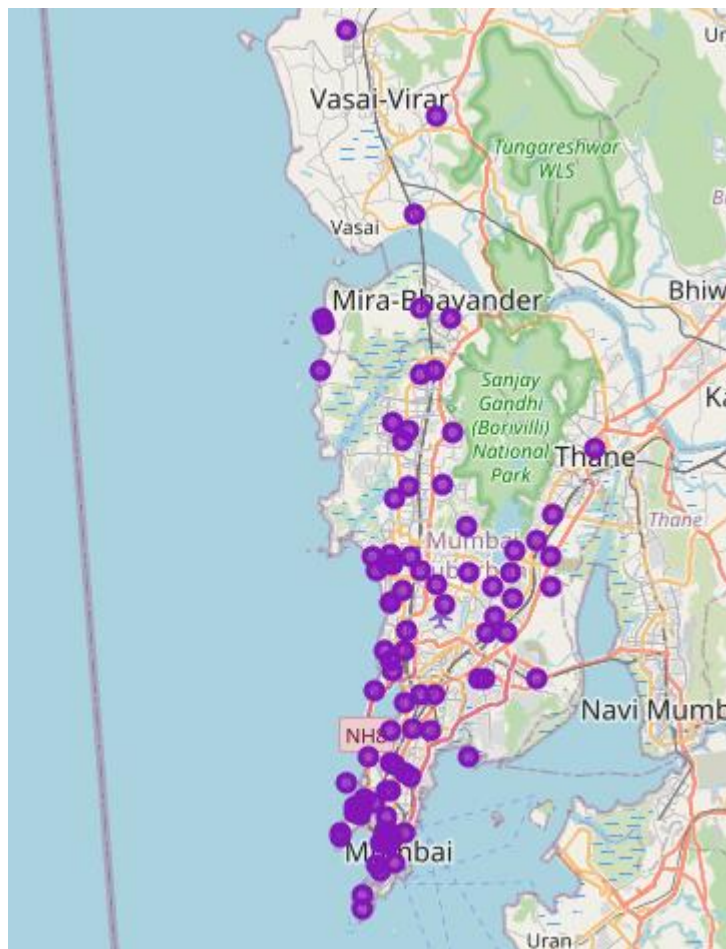
6. Bar plot of Number of Neighborhoods in each borough of Mumbai.

Number of Neighborhood for each Borough in Mumbai City
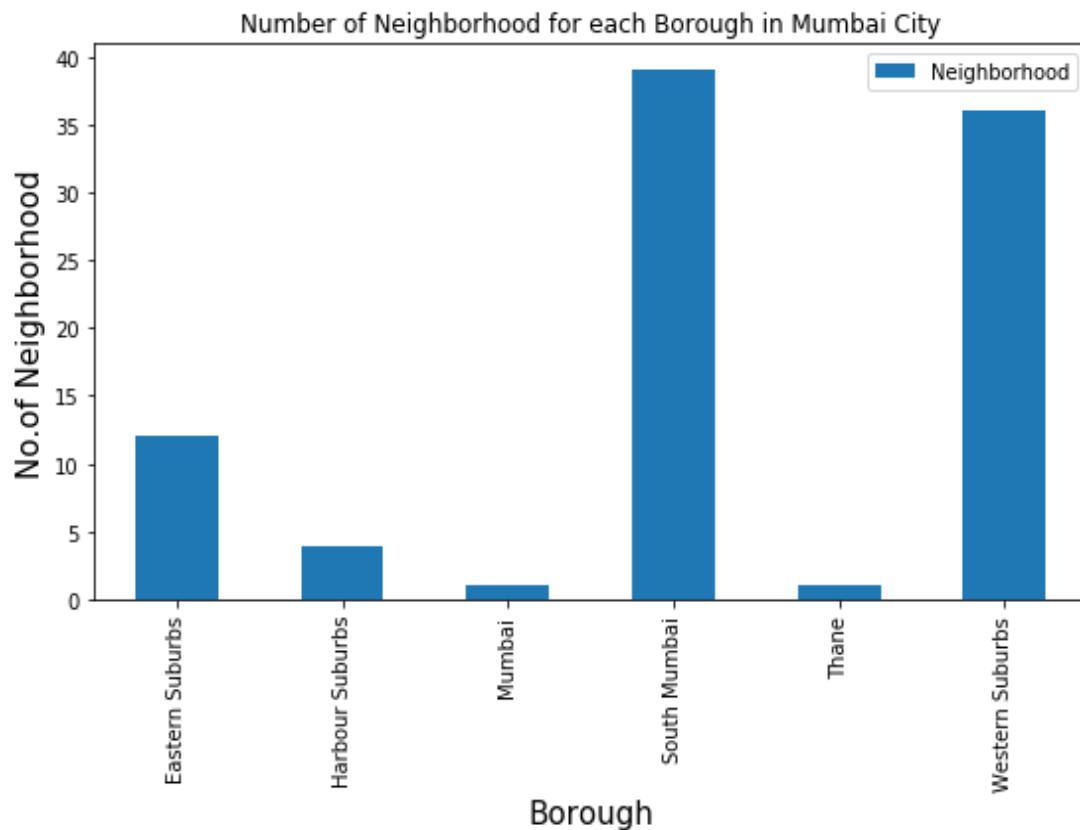
7. Get input to find venues of certain category.

```
category = input('\nEnter Category ')
```

Enter Category Italian

8. Using Foursquare API to download data of venues containing of category given as input.

9. Converting data of venues to data frame and cleaning it.

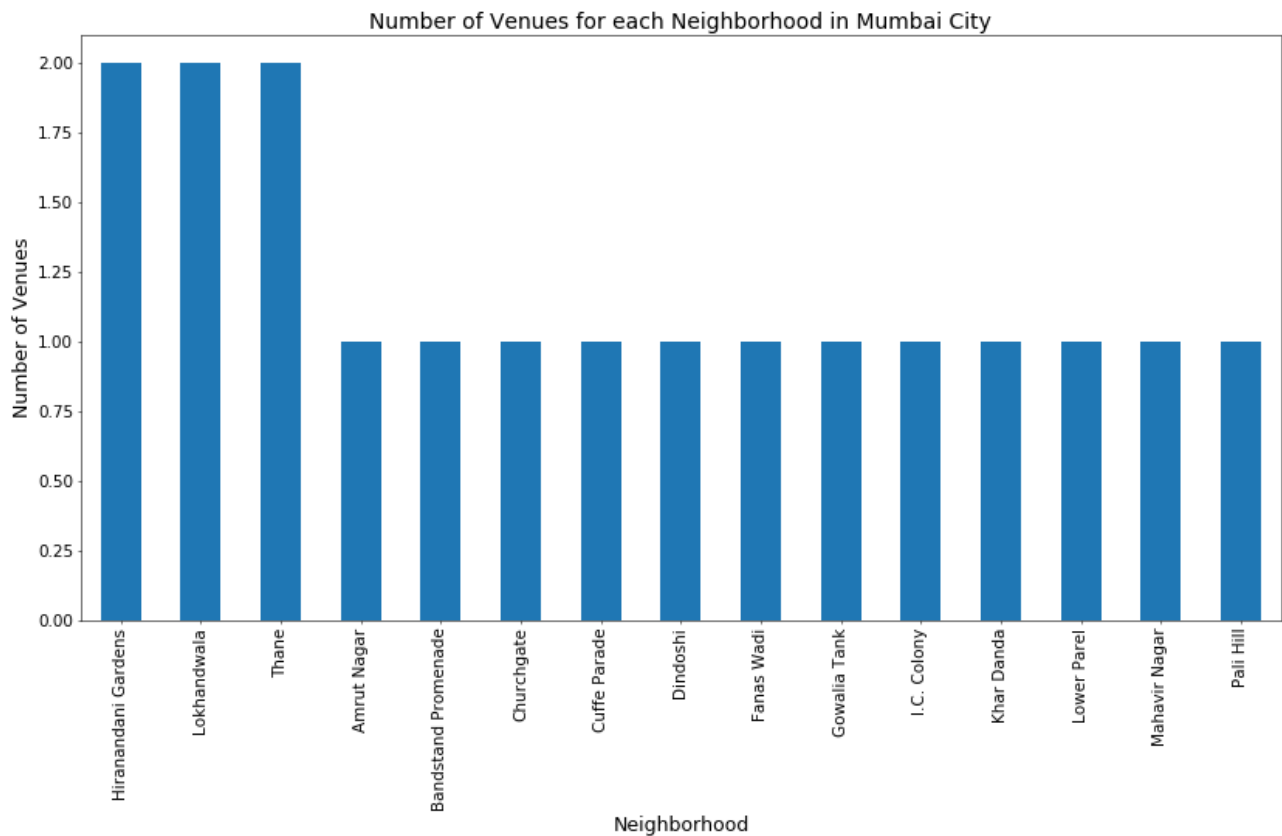|	Borough	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue ID	Venue Latitude	Venue Longitude	Venue Category
0	Western Suburbs	Lokhandwala	19.130815	72.829270	Indigo Deli	5128e04ce4b07aedc43f68bc	19.136430	72.827553	Italian Restaurant
2	Western Suburbs	Lokhandwala	19.130815	72.829270	Apicius	4f72ab6c0039782e9bff2d1d	19.136980	72.833115	Italian Restaurant
3	Western Suburbs	Sahar	19.098889	72.867222	Stax	50782e31e4b07204c0420bb5	19.103659	72.871614	Italian Restaurant
5	Western Suburbs	Bandstand Promenade	19.042718	72.819132	Maritime By San Lorenzo	5230378dbce6d10f7410f8a6	19.043696	72.819263	Italian Restaurant
6	Western Suburbs	Pali Hill	19.068000	72.826000	Out of the Blue	4b0587ccf964a52035a222e3	19.071756	72.825535	Italian Restaurant

10. Visualizing Map of Mumbai City with Venues superimposed on top.
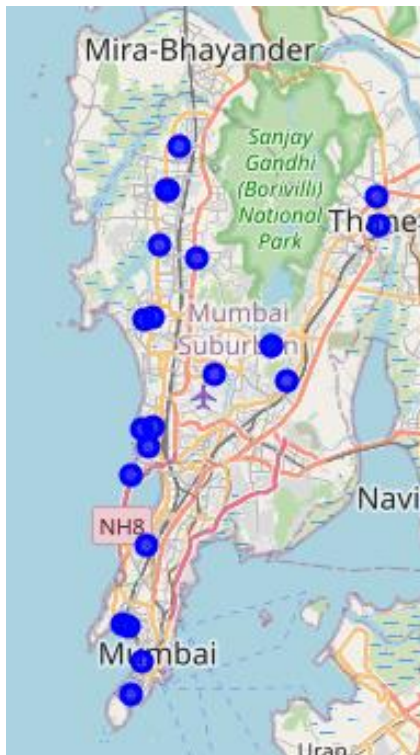
11. Bar plot of number of venues in each Borough of Mumbai
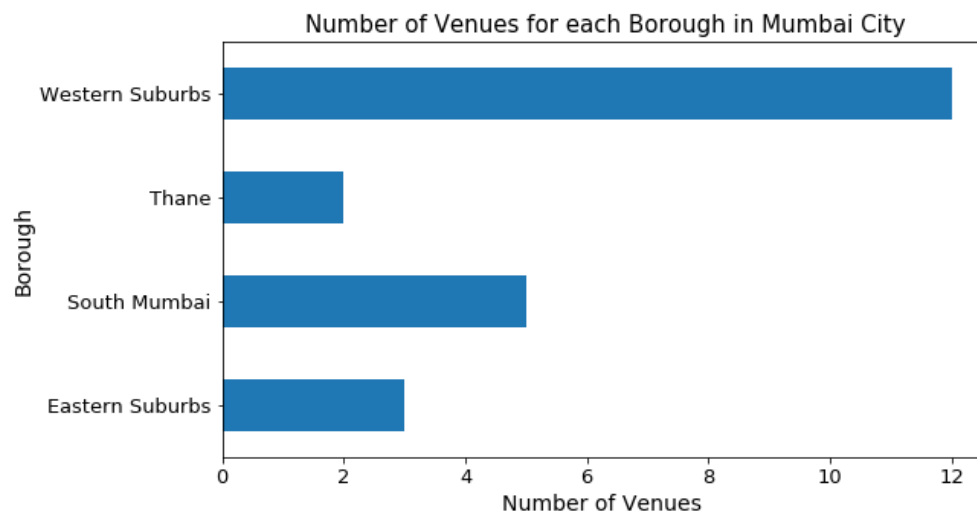
12. Bar plot of number of venues in top 15 neighborhoods of Mumbai

Number of Venues for each Neighborhood in Mumbai City

13. Using Foursquare API to download data of details of venues containing of category given as input

	Borough	Neighborhood	Venue	Venue ID	Likes	Rating	Tips
0	Western Suburbs	Lokhandwala	Indigo Deli	5128e04ce4b07aedc43f68bc	7	7.5	1
1	Western Suburbs	Lokhandwala	Apicius	4f72ab6c0039782e9bff2d1d	13	7.2	4
2	Western Suburbs	Sahar	Stax	50782e31e4b07204c0420bb5	17	7.6	3
3	Western Suburbs	Bandstand Promenade	Maritime By San Lorenzo	5230378dbce6d10f7410f8a6	8	6.9	3
4	Western Suburbs	Pali Hill	Out of the Blue	4b0587ccf964a52035a222e3	219	7.8	94

14. Creating dataframe containing neighborhood and average rating

	Neighborhood	Avg Rating
11	Lower Parel	8.1
7	Hiranandani Gardens	8.1
13	Pali Hill	7.8
2	Churchgate	7.7
15	Sahar	7.6

15. Creating dataframe containing borough and average rating

|	Borough	Average Rating
0	Eastern Suburbs	7.40
1	South Mumbai	7.28
2	Thane	7.30
3	Western Suburbs	7.10

16. Creating dataframe best neighborhood to stay

	Neighborhood	Avg Rating	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Rating	Likes
0	Lower Parel	8.1	18.995278	72.830000	Pronto	19.002965	72.828912	8.1	13.0
1	Hiranandani Gardens	8.1	19.118986	72.911767	Cafe Mangii	19.120260	72.907052	8.2	141.0
2	Hiranandani Gardens	8.1	19.118986	72.911767	Mia Cucina	19.120170	72.907217	8.0	90.0
3	Pali Hill	7.8	19.068000	72.826000	Out of the Blue	19.071756	72.825535	7.8	219.0
4	Churchgate	7.7	18.930000	72.820000	Salt Water Cafe	18.933500	72.826300	7.7	87.0
5	Sahar	7.6	19.098889	72.867222	Stax	19.103659	72.871614	7.6	17.0

17. Visualizing Map of Mumbai City with Best Neighborhood to stay superimposed on top.
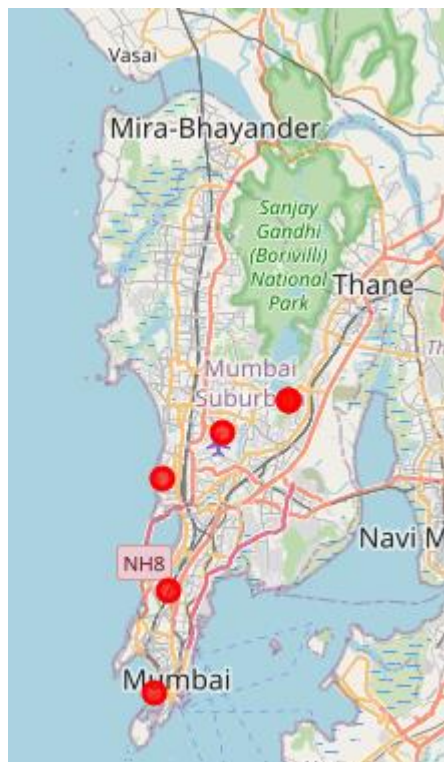
18. Visualizing Map of Mumbai City with Top Rated Venues superimposed on top.
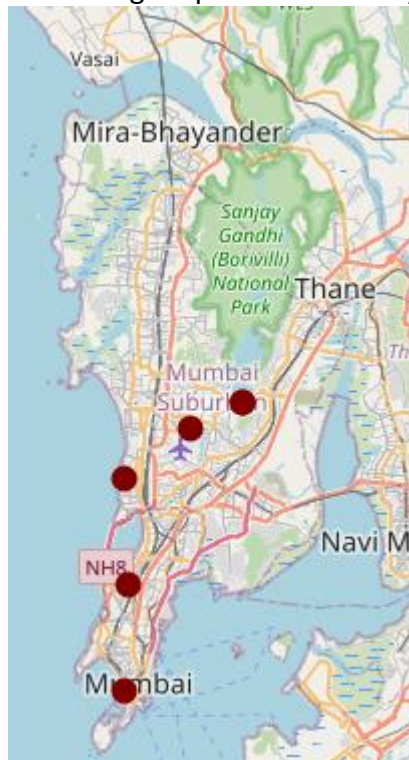
19. Using Foursquare API to download data of top venues containing of in neighborhoods and converting it to dataframe.

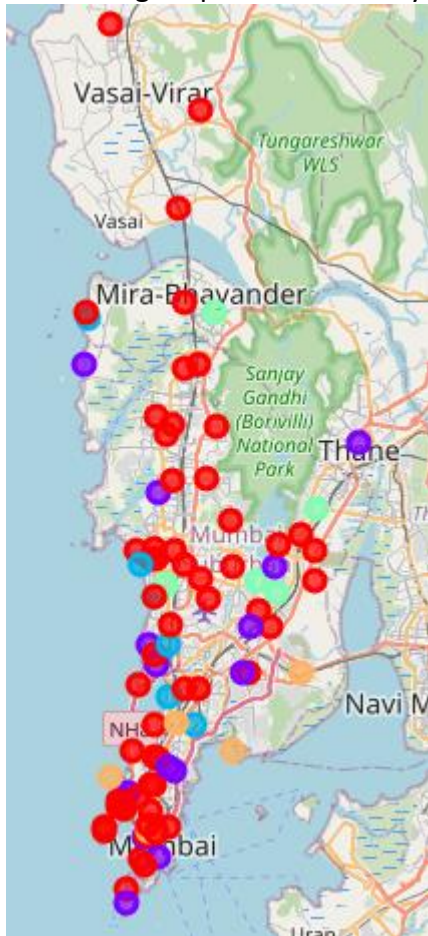|	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Amboli	19.1293	72.8434	Shawarma Factory	19.124591	72.840398	Falafel Restaurant
1	Amboli	19.1293	72.8434	Merwans Cake shop	19.119300	72.845418	Bakery
2	Amboli	19.1293	72.8434	Jaffer Bhai's Delhi Darbar	19.137714	72.845909	Mughlai Restaurant
3	Amboli	19.1293	72.8434	Hard Rock Cafe Andheri	19.135995	72.835335	American Restaurant
4	Amboli	19.1293	72.8434	Pizza Express	19.131893	72.834668	Pizza Place

20. Finding number of unique categories and one-hot encoding according to those categories

21. Creating Dataframe consisting of Top 10 venue categories in each Neighborhood

	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	Aarey Milk Colony	BBQ Joint	Zoo	Dim Sum Restaurant	Flea Market	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store
1	Agripada	Seafood Restaurant	Zoo	Dhaba	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Diner
2	Altamount Road	Indian Restaurant	Zoo	Dhaba	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop
3	Amboli	Falafel Restaurant	Zoo	Dim Sum Restaurant	Flea Market	Fast Food Restaurant	Farmers Market	Farm	Event Space	Electronics Store	Donut Shop
4	Amrut Nagar	Pub	Zoo	Dessert Shop	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Diner

22. Applying KMeans on top 10 venues and adding cluster labels to top 10 venues dataframe.

23. Visualizing Map of Mumbai City with Clusters superimposed on top.

24. Displaying Neighborhoods in each cluster.

25. Cluster 1:

	Borough	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	Western Suburbs	Falafel Restaurant	Zoo	Dim Sum Restaurant	Flea Market	Fast Food Restaurant	Farmers Market	Farm	Event Space	Electronics Store	Donut Shop
1	Western Suburbs	Bakery	Zoo	Food	Flea Market	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store
2	Western Suburbs	Mughlai Restaurant	Zoo	Dim Sum Restaurant	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop
3	Western Suburbs	American Restaurant	Zoo	Dim Sum Restaurant	Flea Market	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store
4	Western Suburbs	Pizza Place	Zoo	Dessert Shop	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Diner

26. Cluster 2:

	Borough	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
13	Western Suburbs	Indian Restaurant	Zoo	Dhaba	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop
14	Western Suburbs	Indian Restaurant	Zoo	Dhaba	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop
16	Western Suburbs	Indian Restaurant	Zoo	Dhaba	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop
19	Western Suburbs	Indian Restaurant	Zoo	Dhaba	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop
44	Eastern Suburbs	Indian Restaurant	Zoo	Dhaba	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop

27. Cluster 3:

	Borough	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
8	Western Suburbs	Chinese Restaurant	Dhaba	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Diner
11	Western Suburbs	Chinese Restaurant	Dhaba	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Diner
21	Western Suburbs	Chinese Restaurant	Dhaba	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Diner
27	Western Suburbs	Chinese Restaurant	Dhaba	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Diner
41	Eastern Suburbs	Chinese Restaurant	Dhaba	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Diner

28. Cluster 4:

	Borough	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
9	Western Suburbs	Pub	Zoo	Dessert Shop	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Diner
34	Western Suburbs	Pub	Zoo	Dessert Shop	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Diner
37	Eastern Suburbs	Pub	Zoo	Dessert Shop	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Diner
42	Eastern Suburbs	Pub	Zoo	Dessert Shop	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Diner
43	Eastern Suburbs	Pub	Zoo	Dessert Shop	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop	Diner

29. Cluster 5:

	Borough	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
50	Harbour Suburbs	Hotel	Zoo	Dhaba	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop
51	Harbour Suburbs	Hotel	Zoo	Dhaba	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop
57	South Mumbai	Hotel	Zoo	Dhaba	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop
63	South Mumbai	Hotel	Zoo	Dhaba	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop
68	South Mumbai	Hotel	Zoo	Dhaba	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Event Space	Electronics Store	Donut Shop

Result

Solutions to the problems given above in the Problems section:

1. The neighborhoods of Mumbai City that have top rated venues of the given category.
   a. Lower Parel
   b. Hiranandani Gardens
   c. Pali Hill
   d. Churchgate
   e. Sahar

	Neighborhood	Avg Rating	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Rating	Likes
0	Lower Parel	8.1	18.995278	72.830000	Pronto	19.002965	72.828912	8.1	13.0
1	Hiranandani Gardens	8.1	19.118986	72.911767	Cafe Mangii	19.120260	72.907052	8.2	141.0
2	Hiranandani Gardens	8.1	19.118986	72.911767	Mia Cucina	19.120170	72.907217	8.0	90.0
3	Pali Hill	7.8	19.068000	72.826000	Out of the Blue	19.071756	72.825535	7.8	219.0
4	Churchgate	7.7	18.930000	72.820000	Salt Water Cafe	18.933500	72.826300	7.7	87.0
5	Sahar	7.6	19.098889	72.867222	Stax	19.103659	72.871614	7.6	17.0

2. The top-rated venues of Mumbai City of the given category are:
   a. Pronto
   b. Café Mangli
   c. Mia Cucina
   d. Out of the Blue
   e. Salt Water Café
   f. Stax

3. Best places to stay if you prefer particular type of venues:
   a. Lower Parel
   b. Hiranandani Gardens
   c. Pali Hill
   d. Churchgate
   e. Sahar

4. Neighborhoods around Mumbai can be clustered into 5 types with top visited venues:
   Cluster 1: Mixed top visited venues
   Cluster 2: Indian Restaurant, Zoo, Dhaba, Fast Food Restaurant, etc.
   Cluster 3: Chinese Restaurant, Dhaba, Fast Food Restaurant, Farmers Market, etc.
   Cluster 4: Pub, Zoo, Dessert Shop, Farmers Market, etc.
   Cluster 5: Hotel, Zoo, Dhaba, Fast Food Restaurant, etc.

Conclusion

Best venues and neighborhoods for the category of venue given as Input were found and successfully visualized on maps and also using plots. KMeans Clustering model was built and used to successfully cluster neighborhoods around the Mumbai City.

There is always room for improvement and hence the solution provided above can also be improved to get best results depending upon the data.