

# **Capstone Project – The Battle of Neighborhoods | Finding a Better Place in Mumbai, India**

**Shubhangi E. Ingale**

**23-06-2021**

## **1. Introduction**

The purpose of this Capstone Project is to assist people in discovering superior amenities around their neighborhood. It will help people making smart and efficient decision on selecting great neighborhood out of numbers of other neighborhoods in Mumbai, India.

### **1.1 Background**

Mumbai, formerly Bombay city, is the capital of Maharashtra state, southwestern India. It is the country's financial and commercial hub and its principal port on the Arabian Sea. Mumbai may be called the City of Dreams, but it is also the Financial capital of India because of the stock market centres. It is the Commercial capital of India because of its vast global reach, and it is the Entertainment capital of India because of Bollywood - the Hindi cinema industry. Mumbai is a Global city or an Alpha World city that acts as a primary connection for the global economic network.

Mumbai city that acts as a primary connection for the global economic network. Lots of people are migrating to various areas of Mumbai and needed lots of research for good housing prices and reputed schools for their children.

### **1.2 Business Problem**

This Capstone Project aim to provide an analysis of features for people migrating to Mumbai to find the best. The features include median housing price and better school according to ratings, crime rates of that particular area, road connectivity, weather conditions, good management for emergency, water resources both fresh and waste water and excrement conveyed in sewers and recreational facilities.

It will help people to get awareness of the area and neighborhood before moving to a new city, state, country or place for their work or to start a new fresh life.

### **1.3 Who would be interested**

This project is for those people who are looking for better neighborhoods. For ease of accessing to Cafe, School, Super market, medical shops, grocery shops, mall, theatre, hospital, like minded people, etc.

## **2. Data**

The data needed for this project, is required to make analysis of housing prices and various venues in neighborhoods of Mumbai. It will help to take decision to select better place to migrate in Mumbai.

## 2.1 Data Description

User is interested to select a place in neighborhoods of Mumbai, where amenities are available at nearby locations and housing prices should be affordable to him. Hence, first we should find out a list of neighborhoods of Mumbai. Then, data of housing prices and various venues located in nearby locations in neighborhoods of Mumbai is required.

## 2.2 Data Sources

### 2.2.1 List of neighborhoods in Mumbai

**Data link:** [https://en.wikipedia.org/wiki/List\\_of\\_neighbourhoods\\_in\\_Mumbai](https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Mumbai) This web page contains a list of neighborhoods in Mumbai along with latitude and longitude in tabular format. This table is extracted by using BeautifulSoup and pandas.

	Neighborhood	Location	Latitude	Longitude
1	Amboli	Andheri, Western Suburbs	19.1293	72.8434
2	Chakala, Andheri	Western Suburbs	19.111388	72.860833
3	D.N. Nagar	Andheri, Western Suburbs	19.124085	72.831373
4	Four Bungalows	Andheri, Western Suburbs	19.124714	72.82721
5	Lokhandwala	Andheri, Western Suburbs	19.130815	72.82927

### 2.2.2 Housing Prices in neighborhoods in Mumbai

**Data Link:** <https://www.magicbricks.com/Property-Rates-Trends/ALL-RESIDENTIAL-rates-in-Mumbai> This website contains property rates in various regions of Mumbai city. It helps us to analyze housing prices in neighborhoods in Mumbai.

	Locality Name	Avg price per sqft
1	Thane West	101187.33
2	Mira Road East	92444.61
3	Kharghar	56700.82
4	Chembur	39984.16
5	Kandivali East	68358.17

### 2.2.3 Foursquare API Data:

As user wants to select a better place to migrate in Mumbai, where various facilities are serviced. We need to retrieve data about various venues in each neighborhood of Mumbai.

In order to retrieve that information we will use “Foursquare” locational information. Foursquare is a location data provider with information about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos. As such, the foursquare location platform will be used as the sole data source since all the stated required information can be obtained through the API.

After finding the list of neighborhoods, we then connect to the Foursquare API to gather information about venues inside each and every neighborhood. For each neighborhood, we have chosen the radius to be 100 meter.

The data retrieved from Foursquare contained information of venues within a specified distance of the longitude and latitude of the postcodes. The information obtained per venue as follows:

1. Neighborhood
2. Neighborhood Latitude
3. Neighborhood Longitude
4. Venue
5. Name of the venue e.g. the name of a store or restaurant
6. Venue Latitude
7. Venue Longitude
8. Venue Category

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Amboli	19.1293	72.8434	Cafe Arfa	19.128930	72.847140	Indian Restaurant
1	Amboli	19.1293	72.8434	5 Spice , Bandra	19.130421	72.847206	Chinese Restaurant
2	Amboli	19.1293	72.8434	Subway	19.127860	72.844461	Sandwich Place
3	Amboli	19.1293	72.8434	Cafe Coffee Day	19.127748	72.844663	Coffee Shop
4	Amboli	19.1293	72.8434	Apple Service Centre	19.128036	72.842718	IT Services

### 3. Methodology

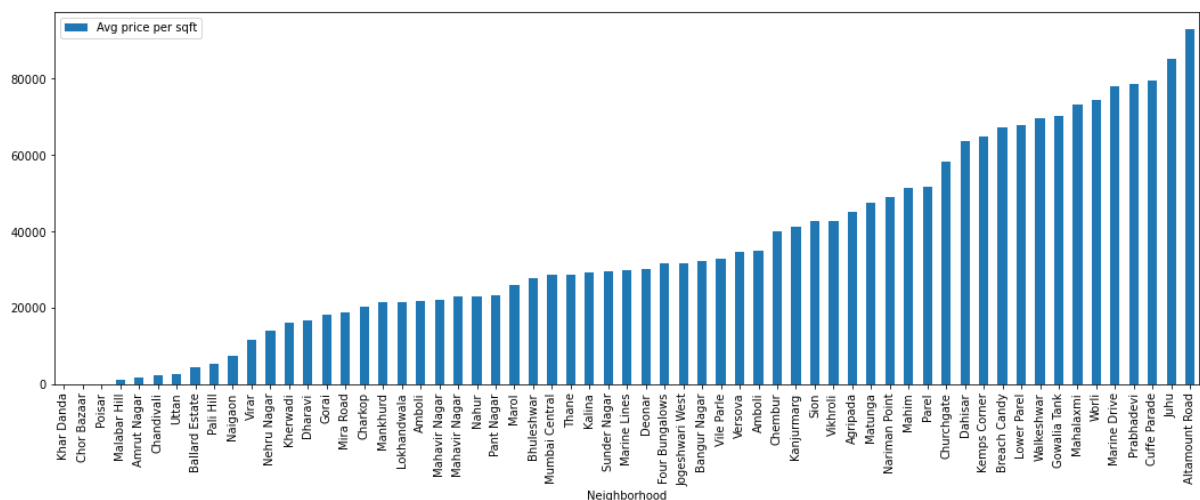
Now, we have the neighborhoods data of Mumbai (93 neighborhoods). We also have the most popular venues in each neighborhood obtained using Foursquare API. A total of 1204 venues have been obtained in the whole city and 169 unique categories. But as seen we have multiple neighborhoods with less than 10 venues returned. In order to create a good analysis let's consider only the neighborhoods with more than 10 venues.

We can perform one hot encoding on the obtained data set and use it find the 10 most common venue category in each neighborhood. Then clustering can be performed on the dataset. Here K - Nearest Neighbor clustering technique have been used. To find the optimal number of clusters silhouette score metric technique is used.

The clusters obtained can be analysed to find the major type of venue categories in each cluster. This data can be used to suggest better place to migrate, suitable locations based on the category.

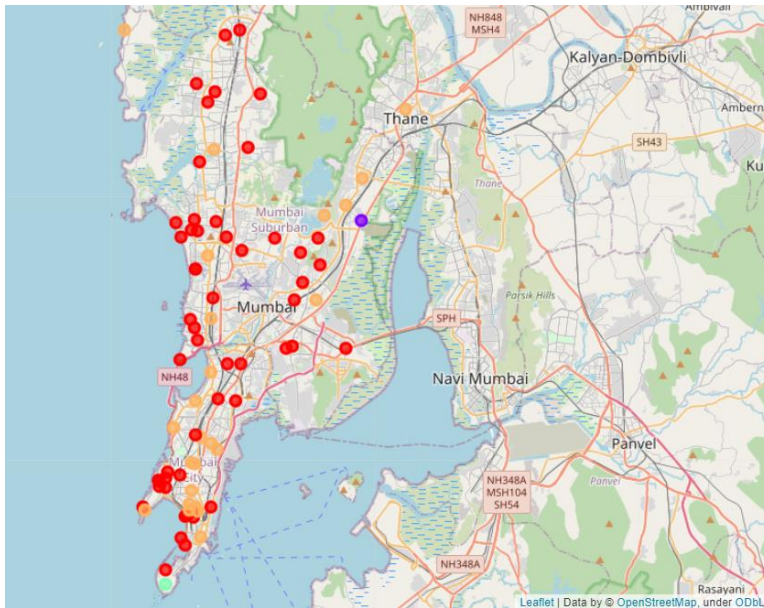
### 4. Analysis

Here we analyzed housing prices in neighborhoods of Mumbai. Following plot shows bar plot of housing prices in neighborhoods of Mumbai.



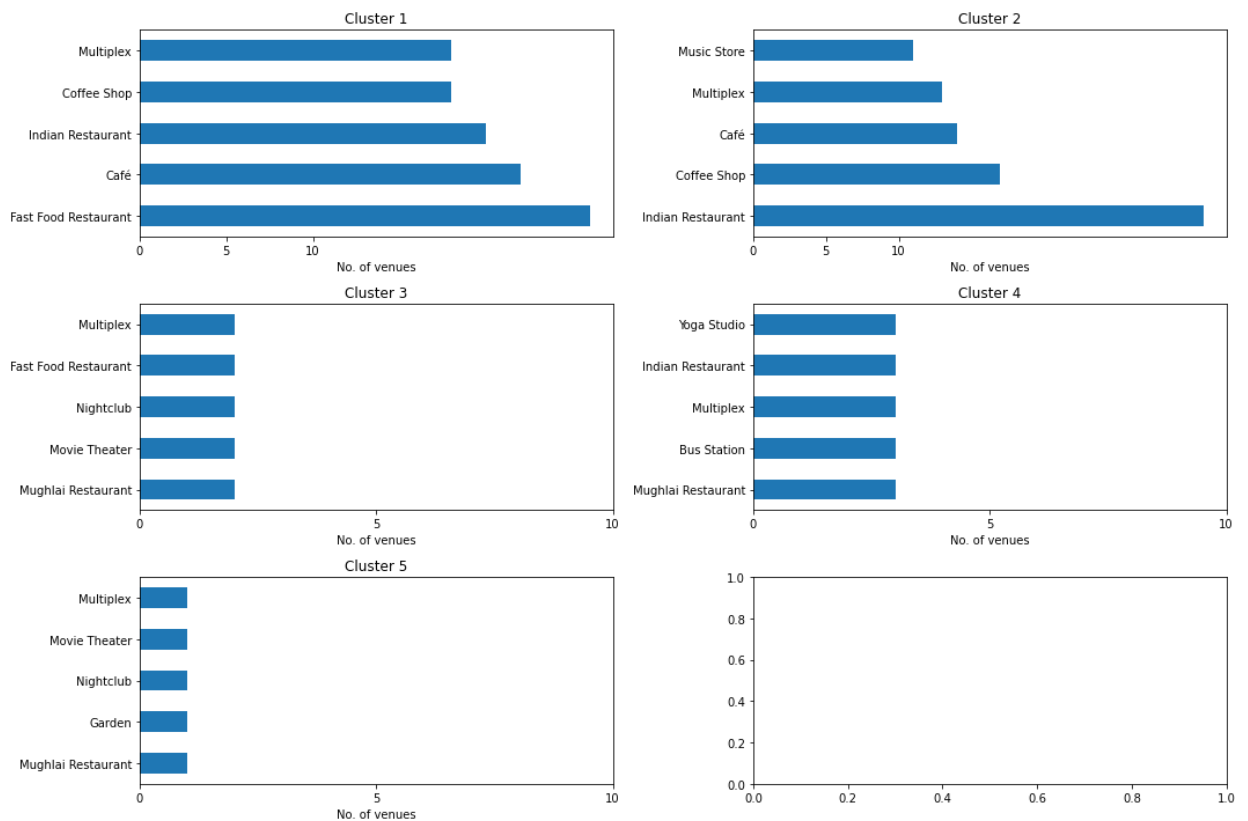
## Cluster Analysis:

To make clusters, we implemented K-means algorithm. As a result, 5 clusters are formed. The following is a map of Mumbai with the neighborhood clusters superimposed on top of it.



## 5. Results

Using the clusters and the top venue categories let's visualize the top 5 venue category in each Cluster for comparison.



## 6. Discussion

From housing prices analysis plot, we can summarize housing rates in Altamount Road neighborhood are costly.

From cluster analysis plots, we can summarize that

1. Cluster 1 contains more number of multiplex, coffee shops, indian resturants, Cafe and Food Resturants.
2. Cluster 2 has more number of multiplex, fast food restaurant, nightclub, Movie theater and Mughalai Restaurants.
3. Cluster 3 has more number of music store, multiplex, cafe, Coffee Shop and Indian Restaurants.
4. Cluster 4 has more number of Yoga Studio, Indian restaurant, multiplex, Bus Station and Mughalai Restaurants.
5. Cluster 5 has more number of multiplex, Movie theater, nightclub, garden and Mughalai Restaurants.

## 7. Conclusion

1. If person is interested to have multiplex, coffee shops, indian resturants, Cafe and Food Resturants nearer to his house, then he can choose place from cluster 1 like Uttan, Gorai, etc. But at the same time he has to see the budget to purchase a house.
2. If person is interested to have multiplex, fast food restaurant, nightclub, Movie theater and Mughalai Restaurants nearer to his house, then he can choose place from cluster 2 like Uttan, Gorai, etc.
3. If person is interested to have music store, multiplex, cafe, Coffee Shop and Indian Restaurants nearer to his house, then he can choose place from cluster 3 like Uttan, Gorai, etc.
4. If person is interested to have Yoga Studio, Indian restaurant, multiplex, Bus Station and Mughalai Restaurants nearer to his house, then he can choose place from cluster 4 like Uttan, Gorai, etc.
5. If person is interested to have multiplex, Movie theater, nightclub, garden and Mughalai Restaurants nearer to his house, then he can choose place from cluster 5 like Uttan, Gorai, etc.