

A guide to Mumbai's restaurants

1. Introduction

1.1 Background

Nowadays, there are restaurants in every nook and corner of the city. Therefore, as a tourist, it becomes quite confusing for him/her to look at the huge number of options available, especially in metropolitan cities like Mumbai. It would come in handy to get extra information on it, like the price of food at the restaurant and the user ratings of the restaurant based on the nearby location. This would help the tourist to get a clear idea on which place to visit and what kind of food to expect.

Mumbai is a city that attracts a large number of tourists from around the world. It has restaurants, hotels, cafes, etc. ranging from various price points. This project explores various venues in Mumbai and provides efficient data based on the average price, price per person and user ratings of a restaurant. This project uses Foursquare API along with Zomato API to do so. It provides crucial information such as the name of the restaurant, address, category, average price and user ratings. To get information on the nearby location, we plot the places on the map with specific color attributes. These color coded plots will help the user decide which place to visit and which to skip.

1.2 Interested audience:

The target audience for this project are the tourists who visit Mumbai. Mumbai, being the financial, commercial and entertainment of India, it attracts a lot of tourists, performers, entrepreneurs and celebrities from all over the world.

They can use these color-coded plots to get info on places that suits their needs. Also, companies can use this detailed information to create a website or a web application similar to this one. This project can also be extended to different places.

2. Data:

2.1 Data Source

In order to get rid of any inconsistencies regarding the venue location, I have used two APIs, Foursquare API and Zomato API, and combined them together. The foursquare API was used to fetch venues in the range of 4 km radius from the coordinates of Mumbai which points out near Santacruz. This helped me fetch the restaurant name, its category and its latitude and longitude.

I used this name, latitude and longitude information on the Zomato API to fetch venues from its database. The location inconsistencies were removed by cleaning the data and getting it in the appropriate form that was needed for the analysis of the data.

The following things were retrieved from the Foursquare API:

1. Name: The name of the restaurant

2. Category: The category of the restaurant
3. Latitude: The latitude value of the restaurant
4. Longitude: The longitude value of the restaurant

The following things were retrieved from the Zomato API:

1. Name: The name of the restaurant
2. Address: The full address of the restaurant
3. Rating: Ratings by users
4. Price Range: The price range of the restaurant defined by Zomato
5. Price for two: The price for two defined by Zomato
5. Latitude: The latitude value of the restaurant
6. Longitude: The longitude value of the restaurant

2.2 Data Cleaning

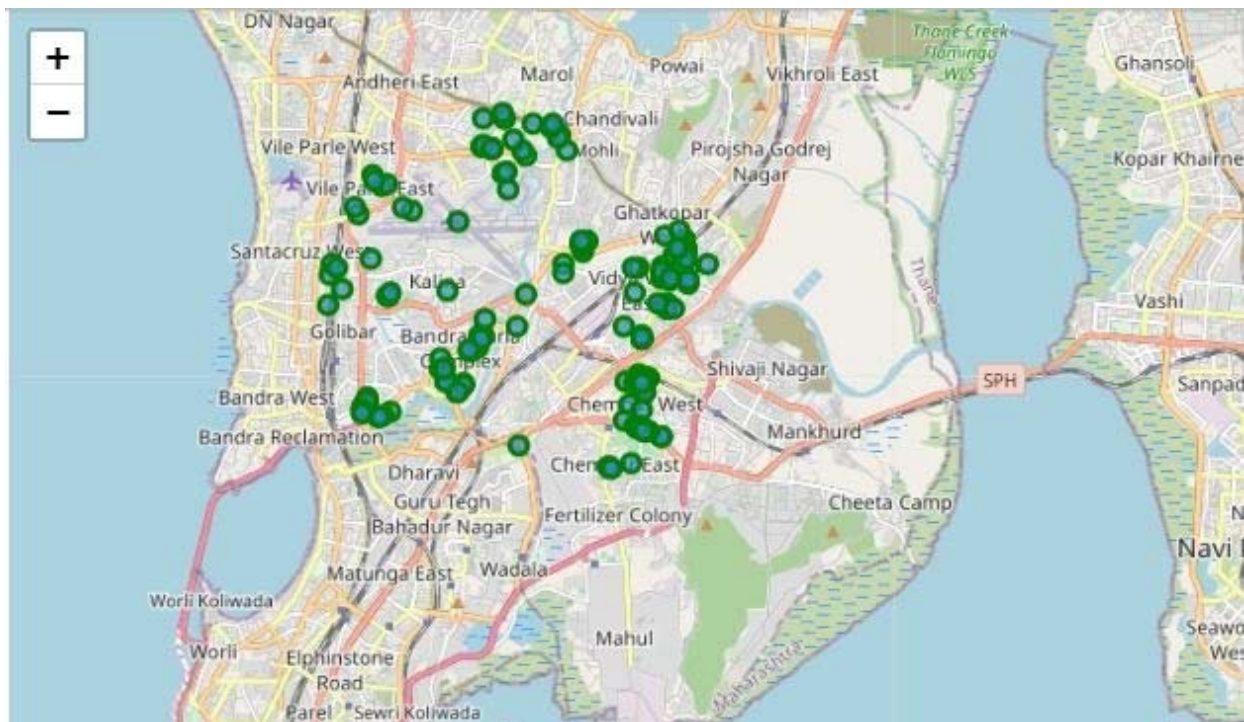


Figure 1: Venues retrieved from Foursquare API

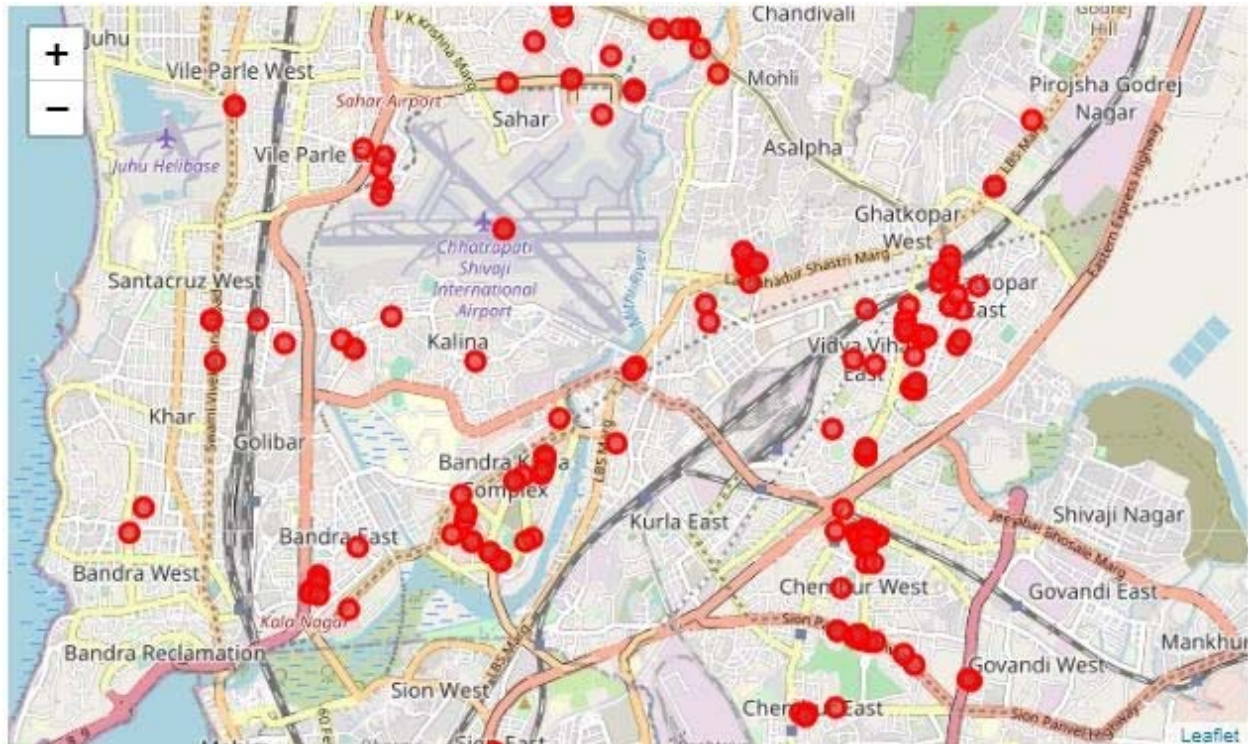


Figure 2: Venues retrieved from Zomato

It is evident from the two figures that not all venues match their locations and there is a bit of variation in the location of venues according to different APIs. Hence, I required the latitudes and longitudes of these venues in order to get their correct location on the map without any error.

If two venues had pretty similar coordinates, it was possible that a restaurant or a cafe was a part of the mall and hence such locations were accepted. However, there were some venue names which were not related to each other and still had pretty similar latitude and longitudes. The data of such venues had to be dropped in order to increase the correctness and accuracy of the project.

After dropping those venues, I was left with 60 rows of data. The first five rows of the updated dataset are as follows:

	categories	venue	latitude	longitude	price_range	rating	address	average_price
0	Hotel	Tuskers - Sofitel	19.0673	72.8692	4.0	4.0	Sofitel Hotel, C 57, Bandra Kuria Complex, Mumbai	1650.0
1	Hotel	O22 - Trident	19.0672	72.8675	4.0	4.0	Trident Hotel, C 56, G Block, Bandra Kuria Com...	2000.0
2	Indian Restaurant	Masala Library	19.0690	72.8696	4.0	4.4	Ground Floor, First International Financial Ce...	2500.0
3	Ice Cream Shop	Natural Ice Cream	19.0776	72.8628	1.0	4.1	3, Gokul Harmony, Kalina Market, Sunder Nagar,...	150.0
4	Deli / Bodega	Smoke House Deli	19.0688	72.8695	4.0	4.3	3A, Ground Floor, 1st International Financial ...	1250.0

Figure 3: First five rows of updated dataset

3. Methodology

A guide to Mumbai's restaurants' aims at identifying the restaurants in Mumbai based on their user ratings and average costs.

I first retrieved the necessary data from two popular APIs namely Foursquare and Zomato. Restaurants' information can be extracted from these APIs and can be put to use. The latitude and longitude values can fetch the venue rating and price information.

Then the explored data is retrieved from the two APIs and plotted on the map and the top category types are identified. The data from the two sources is precisely combined based on the name and 'latitude and longitude' values. The final dataset includes the rating and price values for each venue.

Next, I'll analyze the data that has been created on the basis of average price and user ratings of each restaurant. Then I'll cluster the venues based on the available information of each venue. This will allow us to clearly identify which restaurants can be recommended to the interested audience.

The discussions and conclusions will be seen at the end of this project.

3.1 Analysis

3.1.1 Categories

I tried to distinguish all the venues by the services they offered (e.g. Cafe, Indian Restaurants, Hotels, Bakery, etc.). A bar graph was created by me which indicated the clear majority of 'Indian Restaurants' in the area taken for consideration. To give a proper view of how the diversity is, here is a graph:

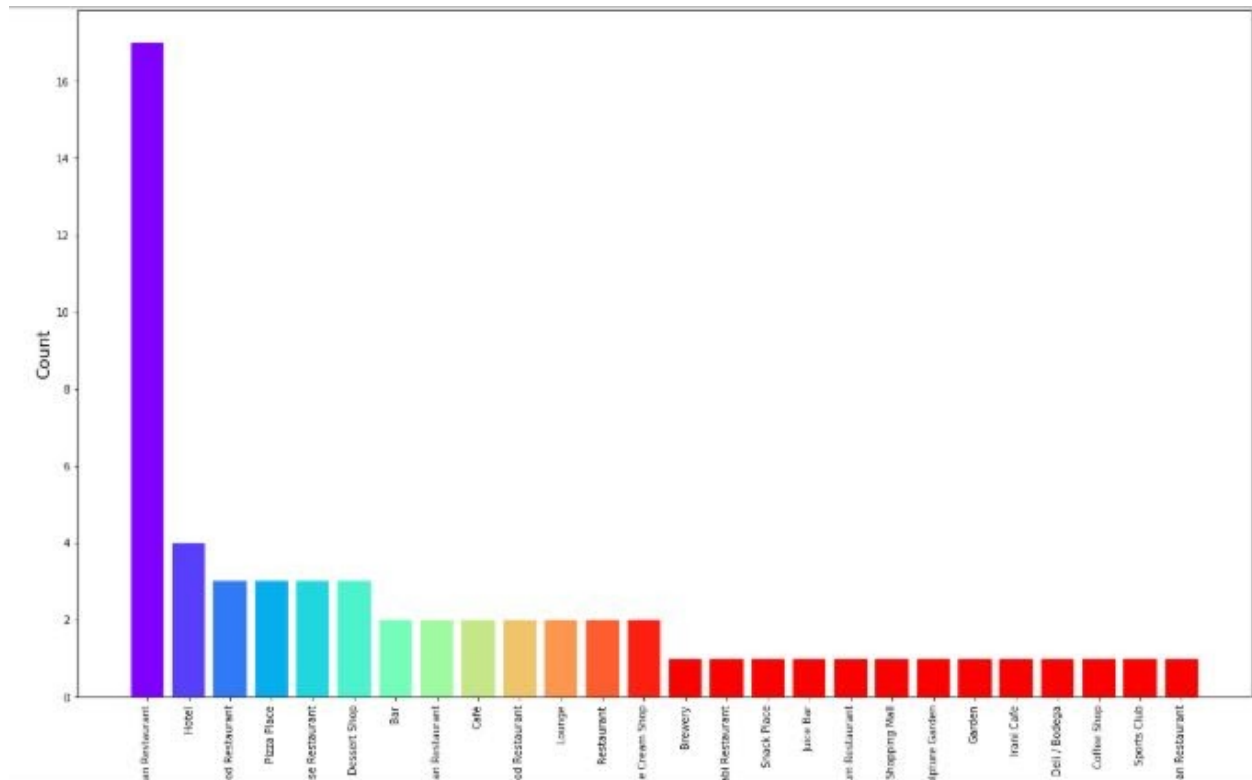


Figure 4: Bar graph shows the density of 'Indian restaurants'

The next-most densely populated venues after Indian restaurants are 'Hotels' and are actually very few in number compared to the former.

3.1.2 Rating

The users of the Zomato application were considered to be the prime factor for the use of this data in order to serve other users who do not have access to the app. The users have the option to rate a venue from 1 to 5.

I collected the information of the ratings received by the various restaurants and tried to find something out of it. It was understood that a majority of those venues had received a rating of 3.4 from the users. To prove my point, here is a graph I plotted based on the ratings:

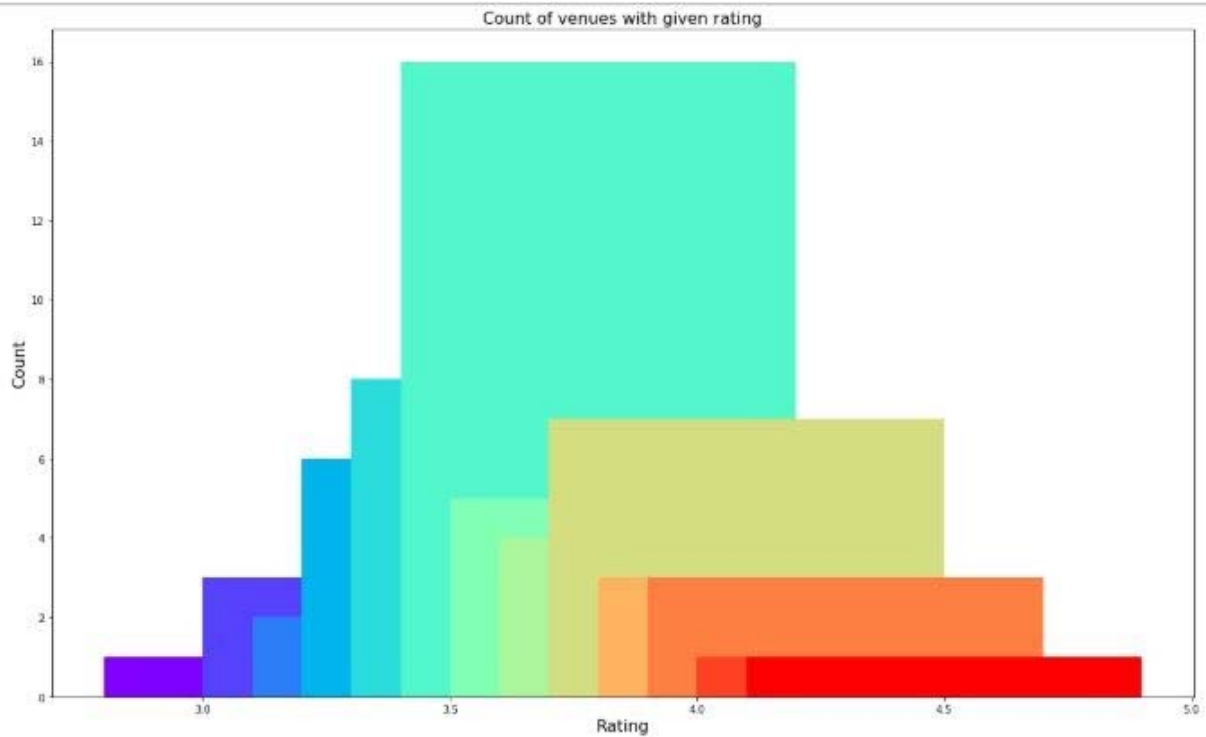


Figure 5: Majority of the venues have a rating of 3.4

I decided to go a step further and plot these points on the map of Mumbai and explore where these places are situated according to their ratings. Here is the plot:

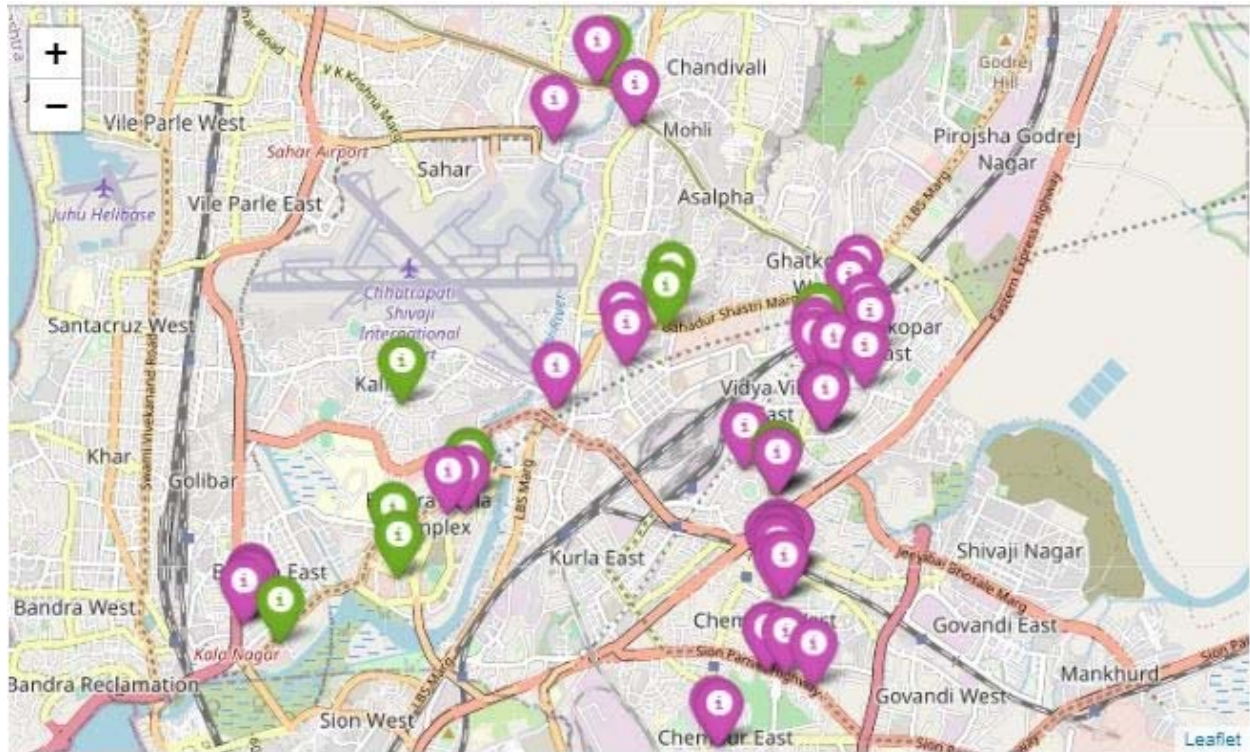


Figure 6: Plot based on ratings of venues

A lot of 'Good' quality venues were found near Chembur West and Ghatkopar areas.

3.1.3 Price

The other area of classification was the pricing of these venues and whether they fit a suitable price range was the question to be solved. I went ahead with a 'price vs venue count' plot and found out that a major chunk of the venues were in the price range of Rs. 0 to Rs. 500. While the venues were evenly distributed in the other price ranges with a maximum being above Rs. 2500, this price range had a significant number of them.

Here is the plot:

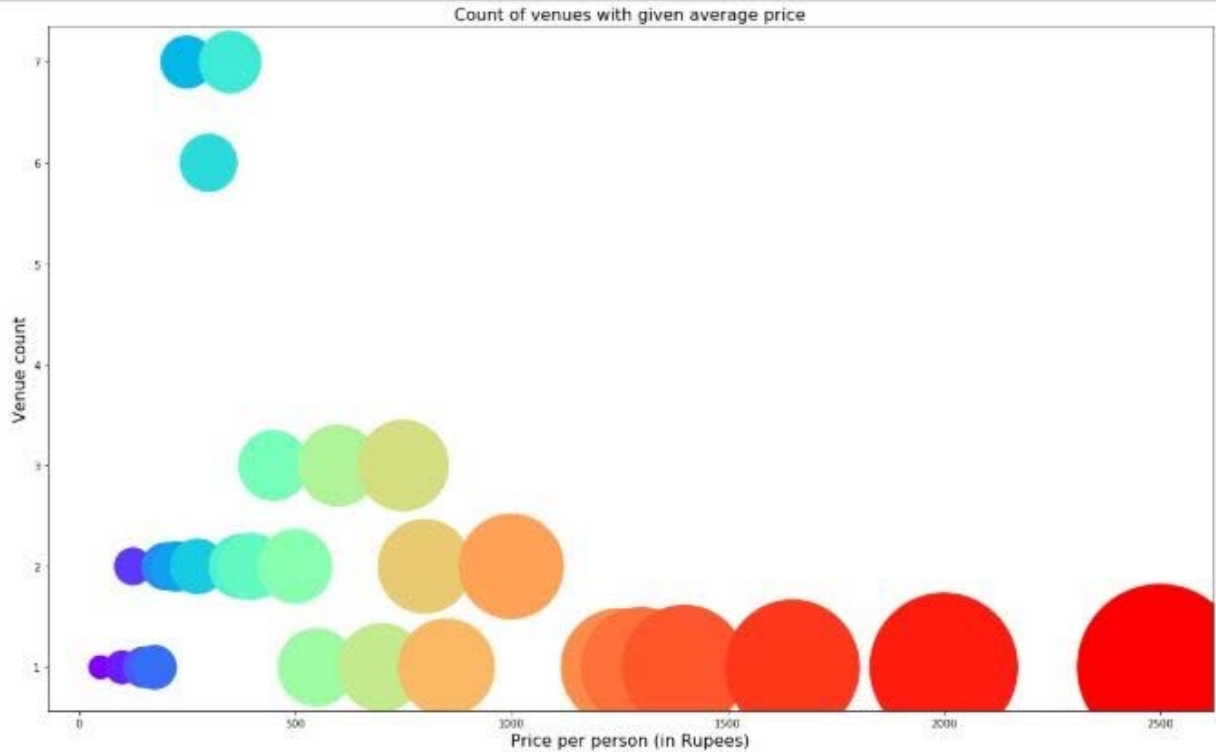


Figure 7: Price range with venue count

The plot of these venues on the map of Mumbai told a story that a lot of cheaper-priced ones were situated where also a lot of 'Good' quality restaurants were found i.e. Chembur and Ghatkopar. However, the area with the costliest venues was Bandra Kurla Complex (BKC).

The point I am trying to prove is visible here:

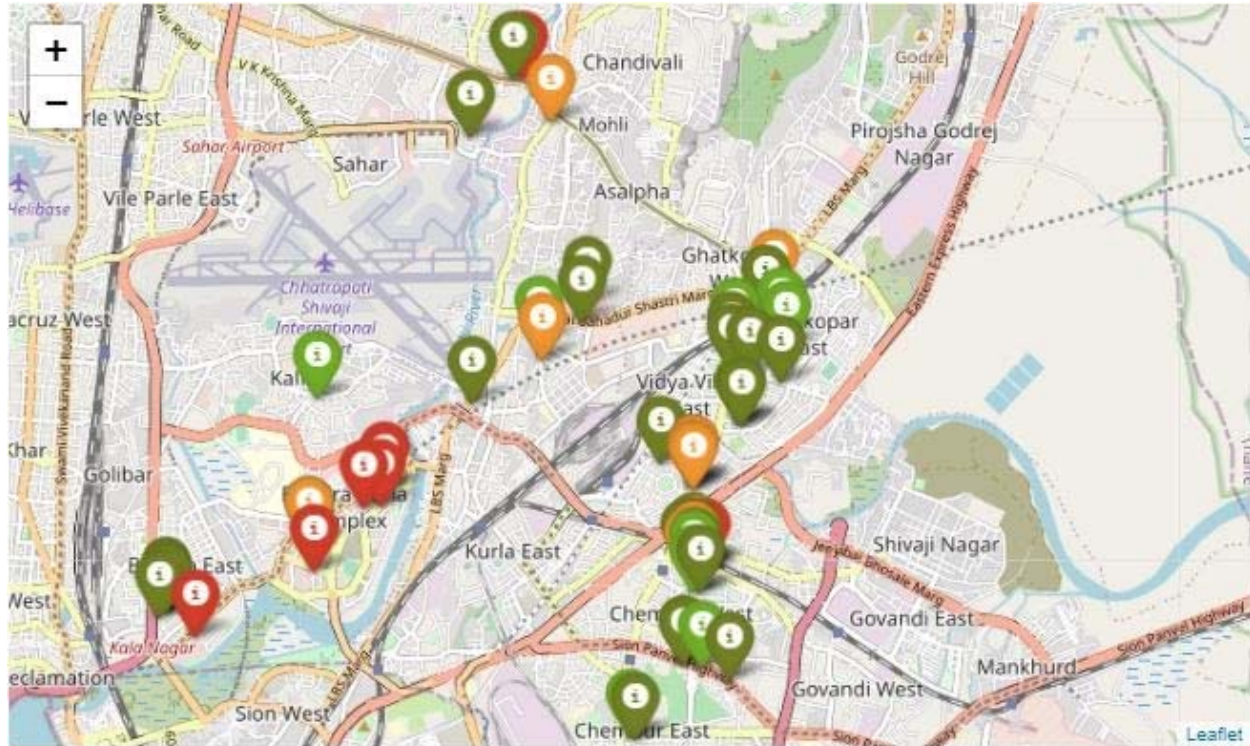


Figure 8: Locations of price-based venues

3.1.4 Clustering

The K-means clustering algorithm was used to separate these venues in two groups. The two clusters are as follows:

1. The green marked cluster captures a wider area and is spread across the entire region. The mean price range of these venues was computed to be 2.06 and rating spread around 3.80.
2. The red marked cluster encompasses a smaller and limited area. The mean price range of these venues was computed to be 4 and rating spread around 4.17.

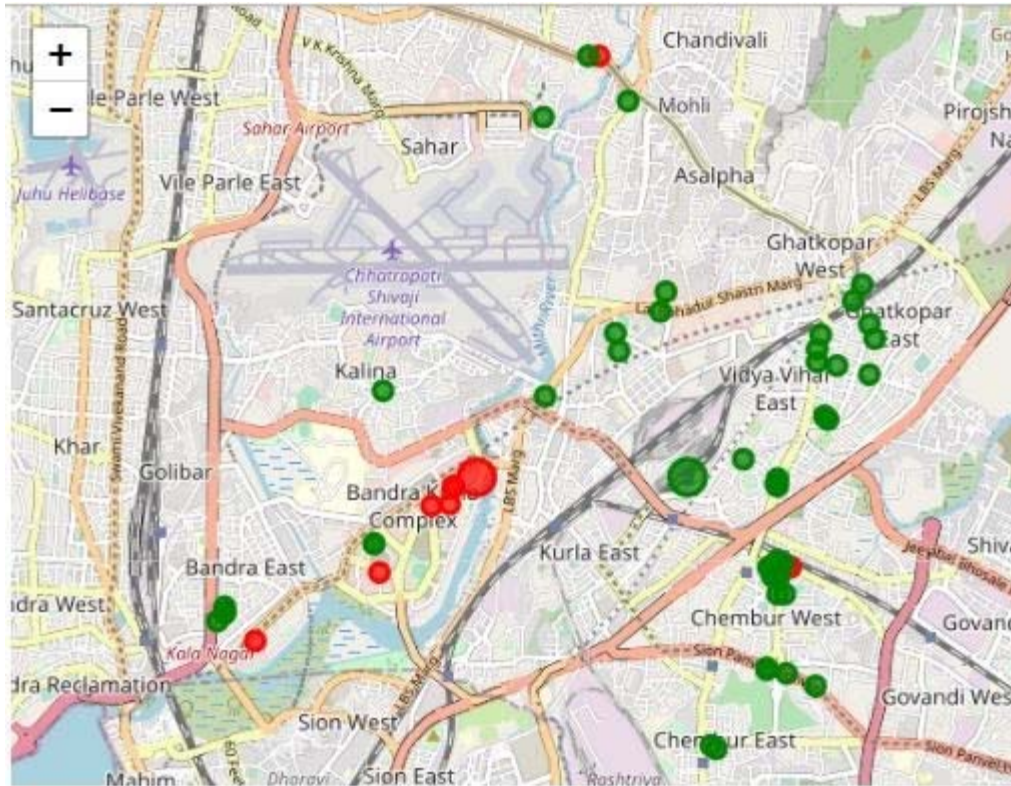


Figure 9: Clusters

4. Results and Discussion

The Foursquare and Zomato APIs were indeed helpful in finding out some really good restaurants in Mumbai. By good, I can clearly say that those restaurants are classified with respect to the budget and the rating received.

I got a list of 183 restaurants to research on and I had to keep out some of those because of the inefficiency of the connection of the two APIs and the location variances. To be precise, 17 of the 183 venues were discarded and the rest were acted upon.

From the initial plot, it was evident that the people of Mumbai were attracted to 'Indian Restaurants' more than anything else and therefore there were a lot of those present in the area selected for analysis. Someone who has visited the country for a short period would really like to try out Indian food and he/she would be spoiled for choices with the opportunity in Mumbai.

Just one disappointment to the visitor will be that a maximum of restaurants in Mumbai are rated close to 3.4 out of 5 and it is not really impressive. However, it is not degrading either as it is considered to be in the 'Good' quality of ratings or the person can have a good time dining out. When I plotted these places, a lot of worthy restaurants were found to be situated at Chembur and Ghatkopar.

Talking about the prices of food, I got many restaurants with average price in the range of Rs. 0 to Rs. 500. There was a variation in the price ranges and the maximum was above Rs. 2500. But it can be said that except the range with most restaurants, the other ranges were quite balanced

with the number of restaurants. The lower priced restaurants were more towards Chembur East and Ghatkopar.

5. Conclusion

The area around Chembur and Ghatkopar is perfect for tourists wanting Indian Restaurants finding 'Good' quality food at a reasonable budget and these areas also have a wide variety of venues if someone not of the aforementioned preferability visits Mumbai.