Exploring places in Ahmedabad, India using Foursquare API and Zomato API

Maulik Doshi

1.Introduction

1.1 Background

Whenever someone searches for a restaurant to visit in a new city, they're highly interested in the places that offer good food and the best services. He will be interested in knowing how good a given restaurant is based on various factors such as rating and pricing. This information will help him to select a restaurant to visit which is also according to the budget. Combining the location of these venues in the city with their price and rating information can help him to make data-driven decisions and he can get the best out of his available choices.

Ahmedabad is the largest city in the Indian state of Gujarat. It is spread across approximately 460 square kilometers. It is situated on the banks of the Sabarmati river, mostly known for its tourist attractions. This project will help tourist to choose restaurant as per their choice of price and ratings.

1.2 Audience

The primary audience for this project is people who are visiting Ahmedabad or want to explore various cuisines in Ahmedabad. They can explore various restaurants and select the one which is best suited for their requirements.

This project may also help people who want to start any restaurant in Ahmedabad. They can explore various restaurants and their categories to survey the area and select the best option available to start their business.

2.Data

2.1 Data Sources

To get the solution to the above-discussed problem, we will require data or restaurants in Ahmedabad. Various fields required are the name of the restaurant, location details containing longitude and latitude information, restaurant's ratings, pricing, and other useful information. To get this infomation, following two APIs are mainly going to be userd:

- Foursquare API
- Zomato API

Using the Foursquare's explore API places within a range of 5 kilometers from the center of Ahmedabad will be fetched and various information like restaurant name, category and location details containing latitude and longitude will be collected.

Using Zomato Search API, restaurant will be searched by restaurant's name collected above by Foursquare API. Data about Name, address,longitude, lattitude,rating and Pricing details will be collected.

The information collected from the above two APIs will be analyzed and restaurants will be classified into a set of clusters. Users can use information obtained by this analysis to take the call on which are the restaurants he would like to visit.

2.2 Data Cleaning

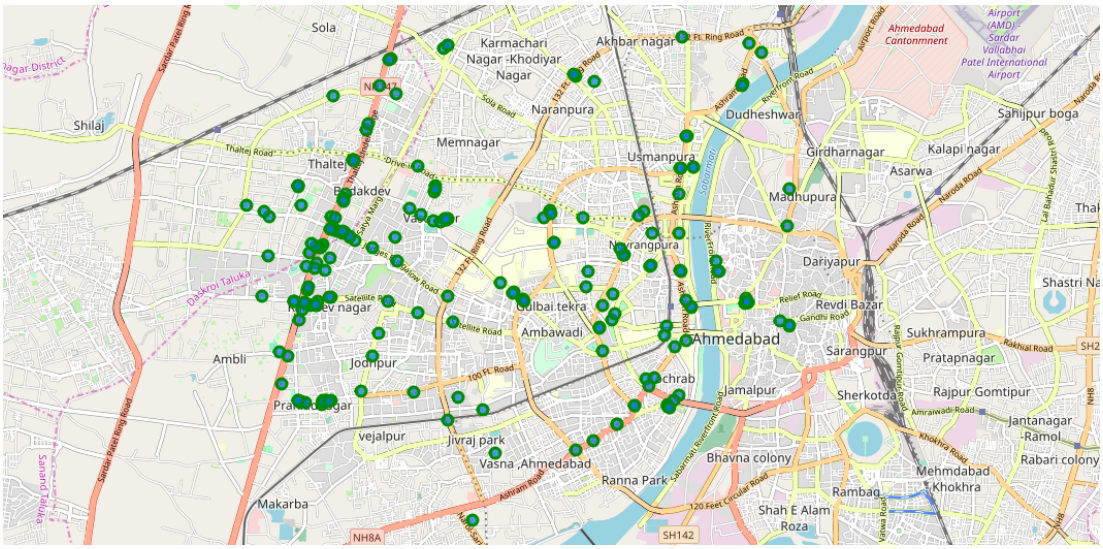

Figure 1: Venues retrieved from Foursquare API

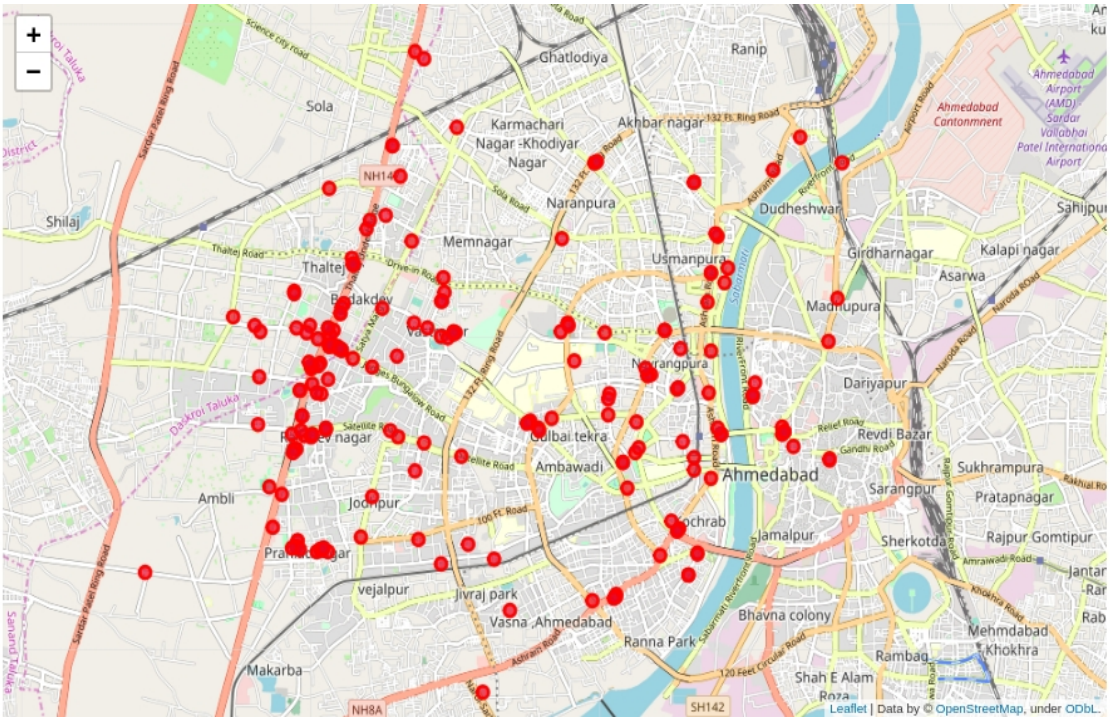

Figure 2: Venues retrieved from Zomato API

From figure 1 and figure 2, we can clearly observer that few venues from the two APIs do not align with each other. So we are going to combine these venues using their latitude and longitude values.

In order to combine the two datasets obtained from Zomato API and Foursquare API, I first check that the latitude and longitude values of each corresponding venue match. After this analysis,all the venues that had their latitude and longitude values different by more than 0.0006 from one another are dropped. To do this, I rounded both the latitude and longitude values up to 4 decimal places. Then, I calculated the difference between the corresponding latitude and longitude values and saw if the difference was less than 0.0006 which should ideally mean that the two locations are the same. This removed many outliers from the two datasets. Once this was done, I observed that there were some venues which were not rated and also some places' categories were not for food like 'Farm','Men's Store','Zoo','Clothing Store','Historic Site','Park','Farmers Market','Multiplex','Arcade','Bookstore','Shopping Mall'. I have dropped this places.

As a final dataset, we're left with 129 venues with 8 columns as described in figure 3.

	categories	venue	latitude	longitude	price_range	rating	address	rating_bin	average_price
0	Café	Zen Cafe	23.0363	72.5497	3.0	4.2	Kasturbhai Lalbhai Campus, University Area, Na...	Very good	400.0
1	Diner	Cafe Upper Crust	23.0411	72.5487	3.0	4.3	Aarohi Complex, Vijay Cross Road, Navrangpura,...	Very good	500.0
2	Bakery	7Violettes	23.0401	72.5477	2.0	4.3	A 2, Vishal Apartments, Off University Road, ...	Very good	250.0
4	Café	Mocha	23.0294	72.5549	3.0	4.6	10, Vasant Baug Society, Near IDBI Bank, CA Ci...	Very good	600.0
5	Café	Ristretto - Behind the Rods	23.0409	72.5488	3.0	4.3	A-1, Ground Floor, Maharaja Palace, University...	Very good	500.0

Figure 3: First 5 records from final data aggregated from both APIs

3. Methodology

Aim of this project is to find best places in Ahmedabad based on their rating and average costs. It will help any visitor or tourist to identify places to visit according to their preferences. To do this we retrieved data from two APIs (Foursquare and Zomato). Places within a range of 5 Km from Gujarat University, Ahmedabad were extracted using Foursquare API. The latitude and longitude values are then used to fetch venue ratings and price from Zomato.

After, we then explored the data retrieved from the two APIs on the map. The data from the two sources is carefully combined based on the name, latitude and longitude values from the two sources. The final dataset would include the rating and price values for each place.

We will analyze this dataset based on the price, ratings and category and location. We will plot this on a map using folium to analyze it further. After that, we will cluster the venues based on the available information of each venue. This will allow us to identify the places which can be recommended.

3.1 Analyzing places by categories

To analyse various places by categories, we have plot number of categories vs categories chart as shown in Figure 4 below. In this figure we can see that the majority places are Indian Restaurant, Cafe, Hotel, Coffee Shop, Restaurant and Pizza Place.

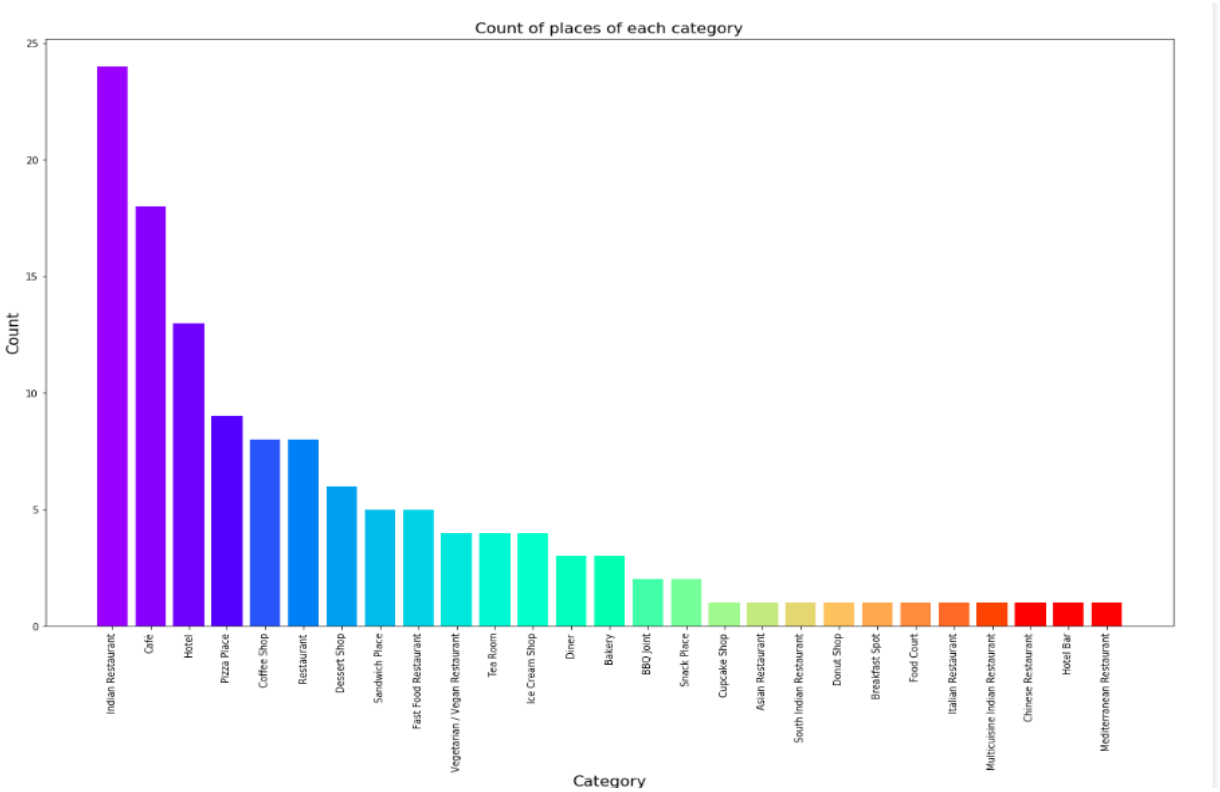

Figure 4: Count of various types of venues in Ahmedabad

3.2 Analyzing places by rating

To analyse places by rating, a bar chart is plotted with x-axis as the rating from 1 to 5 and the y-axis as the count of venues with that rating.

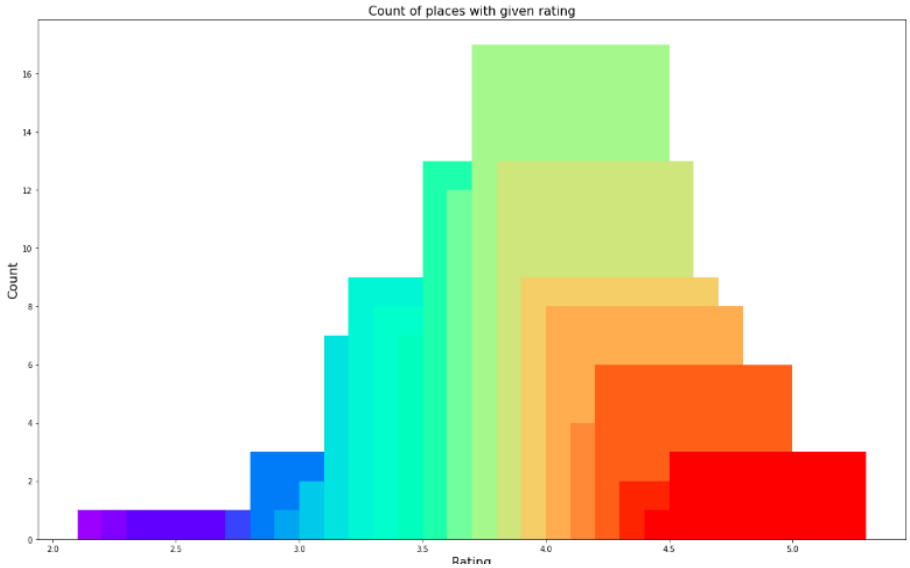

Figure 5: Rating and count of venues with that rating

It can be said as shown in Figure 5 that most of the places have rating around 4 with maximum number of venues scoring between 3 and 5. To see distribution of these locations I have plotted this place on maps as shown in Figure 6. Places with rating 1-2 are mapped as red,2-3 are mapped as orange,3-4 are mapped as green and 4-5 are mapped as dark green in map.

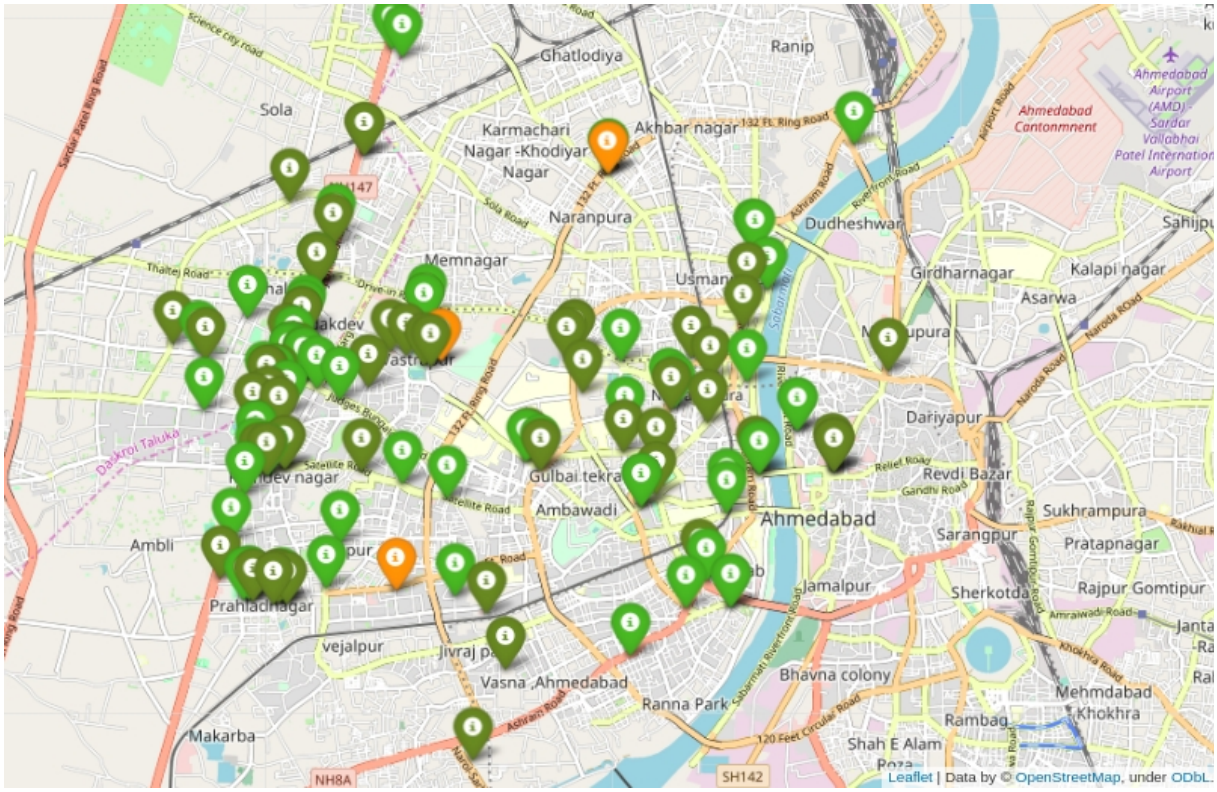

Figure 6: Plot of venues with different ratings

In Figure 6, we can see the most of the places are with very good ratings. Places on SG highway and Ashram road are dark green. So it is very likely that people visiting Ahmedabad will love to visit this venues.

3.3 Analyzing places by price

To explore the average prices of all venues for one person ,let us usea scatter plot along with the count of venues with that average price per person. After plotting as shown in figure 7, it can be said that the majority venues have an average cost of Rs 200 to Rs 600 for one person. Even though the maximum venues lie in that range, the actual range of prices is very different. There are few places with average price even as high as Rs 1000+ for one person.

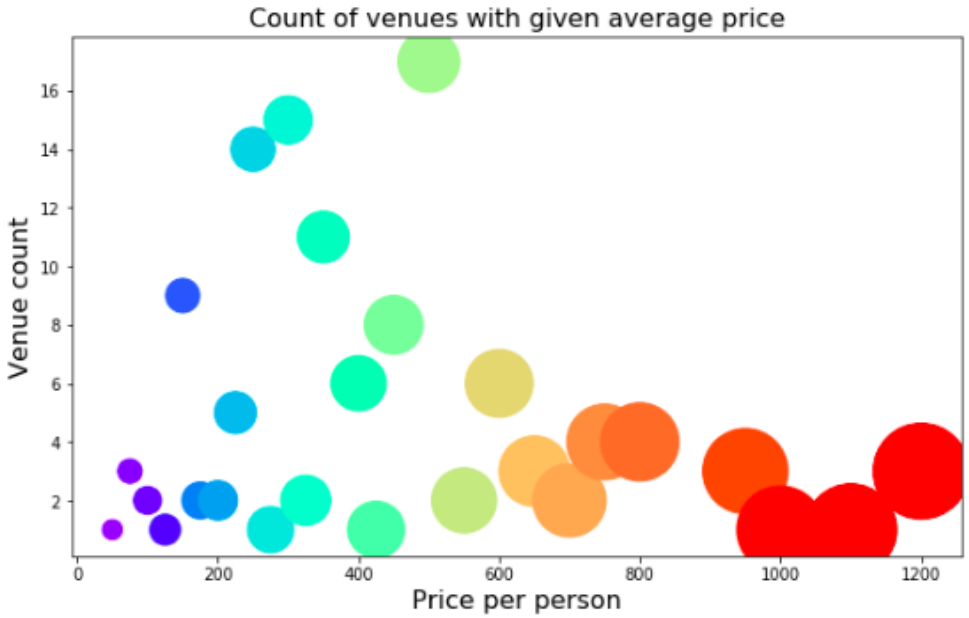

Figure 7: Price per person with count of venues with that price

I have also plotted this places on maps and given color coding as per price of the venues. Figure 8 includes all the venues where high priced venues are marked by orange and red while the low priced venues are marked with green and dark green.

In figure 8 we can see that venues of all price range are distributed across the area selected. It makes visitor easier to find restaurant of his price preferences in the any area.

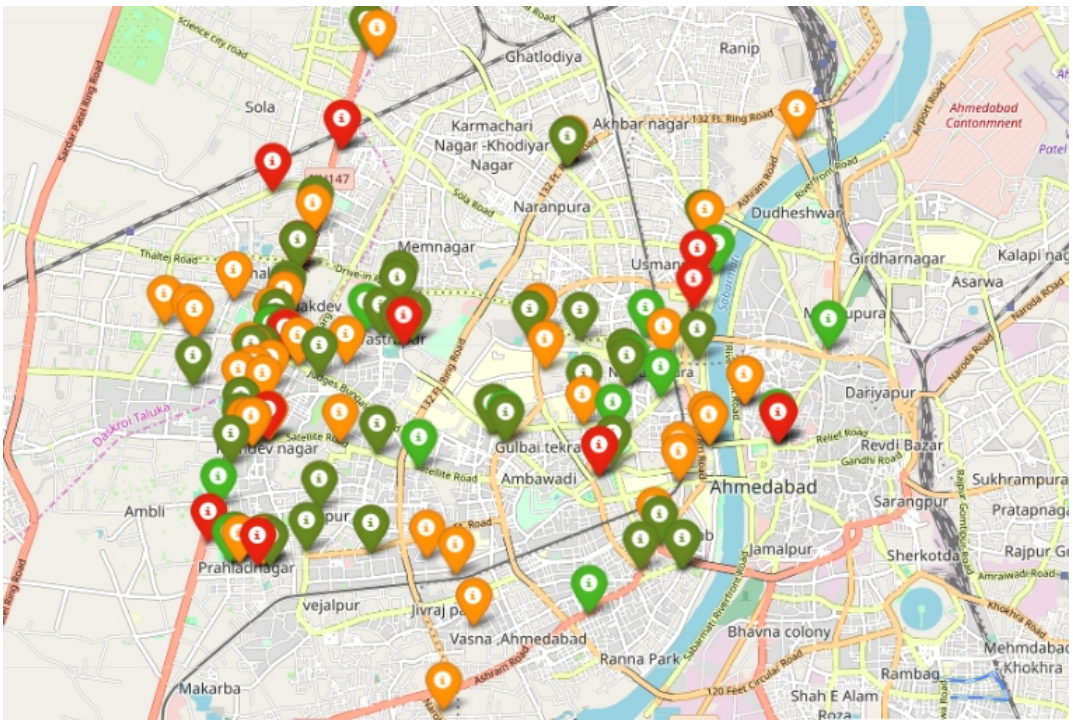

Figure 8: Plot of venues with different prices

3.4 Clustering

Finally, I cluster all the venues based on their price range, location and more to identify similar venues and the relationship amongst them. I used KMeans clustering and decided to cluster the venues into two separate groups.

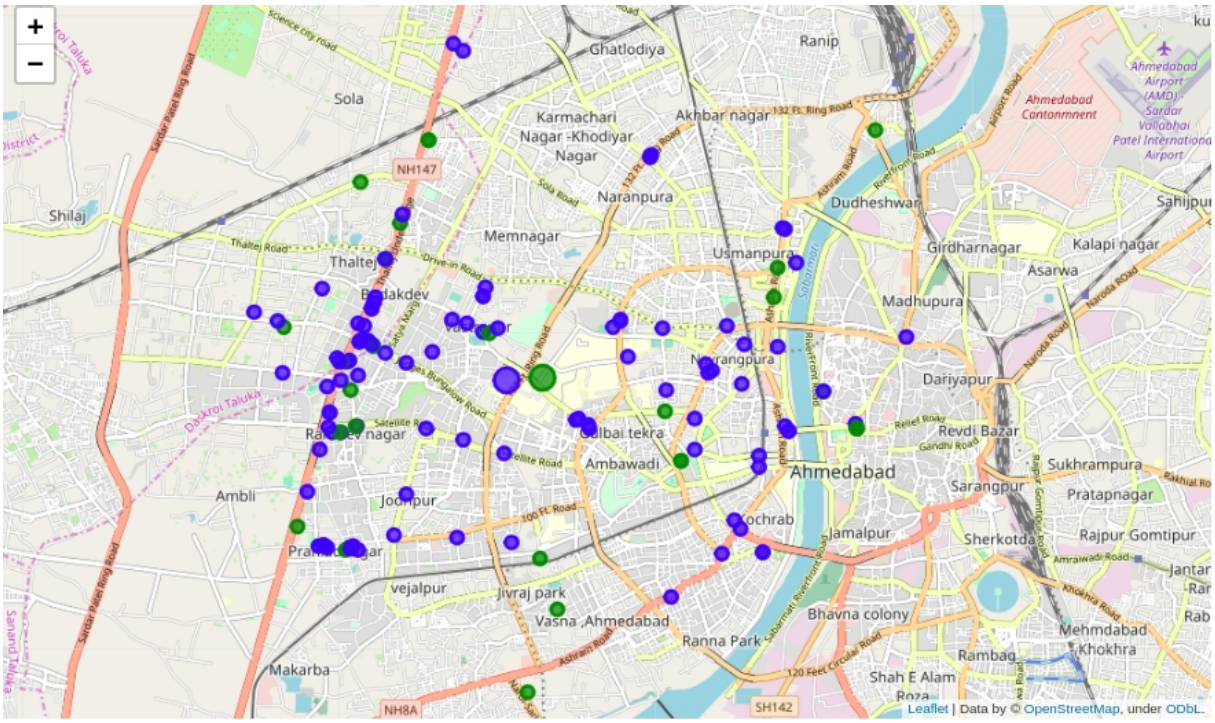

Figure 9: Clusters of venues

In figure 9, we see the two clusters:

1. The first cluster (blue) is spread across the whole area and includes the majority venues. These venues have mean price range of 2.18 and rating spread around 3.90.
2. The second cluster (green) is very sparsely spread and has very limited venues. These venues have mean price range of 3.59 and rating spread around 4.33.

4. Results and Discussion

With the help of above analysis, we can draw conclusions which will help people visiting Ahmedabad to locate places according to their preferences.

Initially we got list of 203 venues from Foursqare and Zomato APIs. After that we dropped places with unwanted categories and places without rating available.We also combined data of places obtained from Zomato and Foursquare APIs. Finally there were 129 places to analyze.

We find that he majority places are Indian Restaurant, Cafe, Hotel, Coffee Shop, Restaurant and Pizza Place. People who like places of these categories will love to visit Ahmedabad frequently.

While the complete range of ratings range from 1 to 5, the majority venues have ratings close to 4. This indicates the people who have previously visited this places have liked the quality

of food and pricing. When we plot these venues on the map, we discover that there are clusters of venues around SG Highway and Ashram road.

While analyzing price range, majority of places falls in price range of 200 to 600. However, the variation in prices is very large, given the complete range starts from Rs 100 and goes uptil Rs 1200. On plotting these places on maps, we can see that venues of both high and low price range are available at both SG Highway and Ashram road area.

Finally, through clusters we identified that there are many venues which are relatively lower priced but have an average rating of 3.90. On the other hand, there are few venues which are high priced and have average rating of 4.33.

5. Conclusion

The goal of this project was to explore various places that a person can visit while he is in Ahmedabad,Gujarat.The venues have been identified using Foursquare and Zomato API and have been plotted on the map. The map reveals that there are two major following areas near Gujarat University, Ahmedabad: S G Highway and Ashram Road. A person can easily identify places of their choice based on their priorities from the places plotted on maps.