Exploring venues in Hyderabad, India using Foursquare and Zomato API

Laxman Kishore K

Feb 24th 2020

1. Introduction

1.1 Background

Whenever a person searches for a venue in a new city, they're highly interested in the best places that the city has to offer. The person might want to know how good a given restaurant is or the price range it falls under. This extra information would help decide which venue to choose amongst the many venues in the city. Combining the location of the venues in the city with their price and rating information would surely help visitors in a city make better informed decisions about the places they should visit.

Hyderabad is the capital of southern India's Telangana state. A major center for the technology industry, it's home to many upscale restaurants which can be explored. This project explores various venues in Hyderabad and attributes the data based on user ratings and average price. To explore this information, this project involves the usage of both the Foursquare API and the Zomato API to fetch complete information of various venues (including name, address, category, rating, and price). Further, a map of the venues with specific color attributes will be plotted to highlight their position, and information about these venues. Such plots imbibe bountiful information in the form of their colored representations and location on the map. This enables any visitor to take a quick glance and decide what place to visit.

1.2 Interested audience

The target audience for such a project is twofold. Firstly, any person who is visiting Hyderabad, India can use the plots and maps from this project to quickly select places that suit their budget and rating preferences. Secondly, a company can use this information to create a website or a mobile application, which is updated on a regular basis, to allow individuals to the city or even expand same functionality to other places.

2. Data

2.1 Data Sources

To get location and other information about various venues in Ameerpet, I used two APIs and decided to combine the data from both of them together.

Using the Foursquare's explore API (which gives venues recommendations), I fetched venues up to a range of 4 kilometers from the center of Chandigarh and collected their names, categories and locations (latitude and longitude).

Using the name, latitude and longitude values, I used the Zomato search API to fetch venues from its database. This API allows to find venues based on search criteria (usually the name), latitude and longitude values and more. Given that the data from the two APIs did not align completely, I had to use data cleaning to combine the two datasets properly.

From Foursquare API (https://developers.zomato.com/api), I retrieved the following for each venue:

- **Name:** The name of the venue.
- **Category:** The category type as defined by the API.
- **Latitude:** The latitude value of the venue.
- **Longitude:** The longitude value of the venue.

From Zomato API (https://developers.zomato.com/api), I retrieved the following for each venue:

- **Name:** The name of the venue.
- **Address:** The complete address of the venue.
- **Rating:** The ratings as provided by many users.
- **Price range:** The price range the venue belongs to as defined by Zomato.
- **Price for two:** The average cost for two people dining at the place. I later convert the same to average price per person by dividing by 2.
- **Latitude:** The latitude value of the venue.
- **Longitude:** The longitude value of the venue.

2.2 Data Cleaning

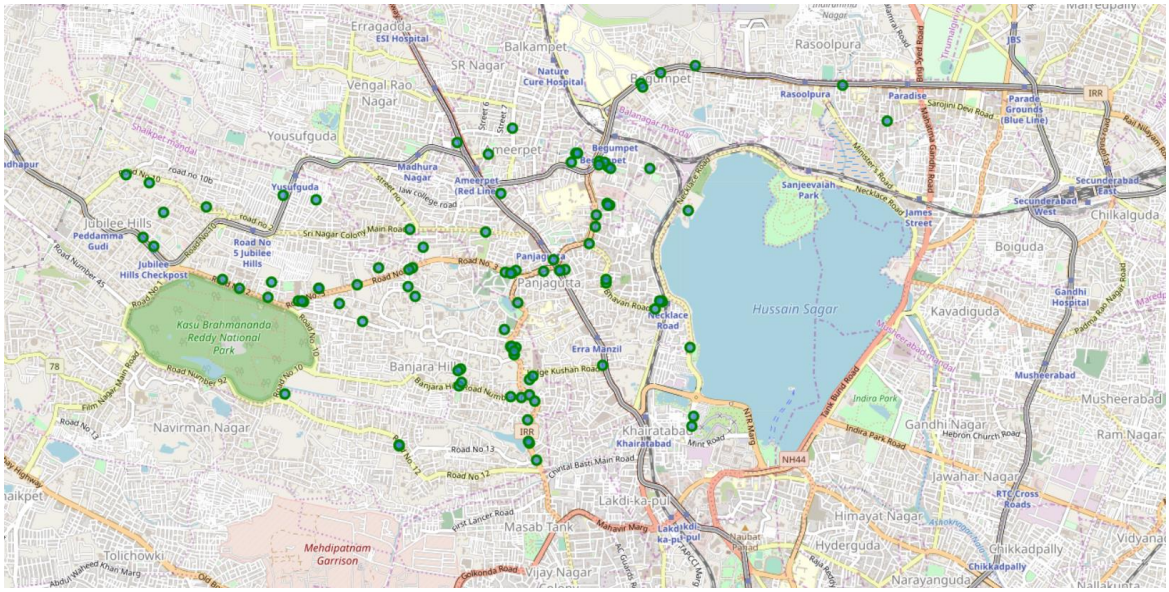

Figure 1: Venues retrieved from Foursquare API

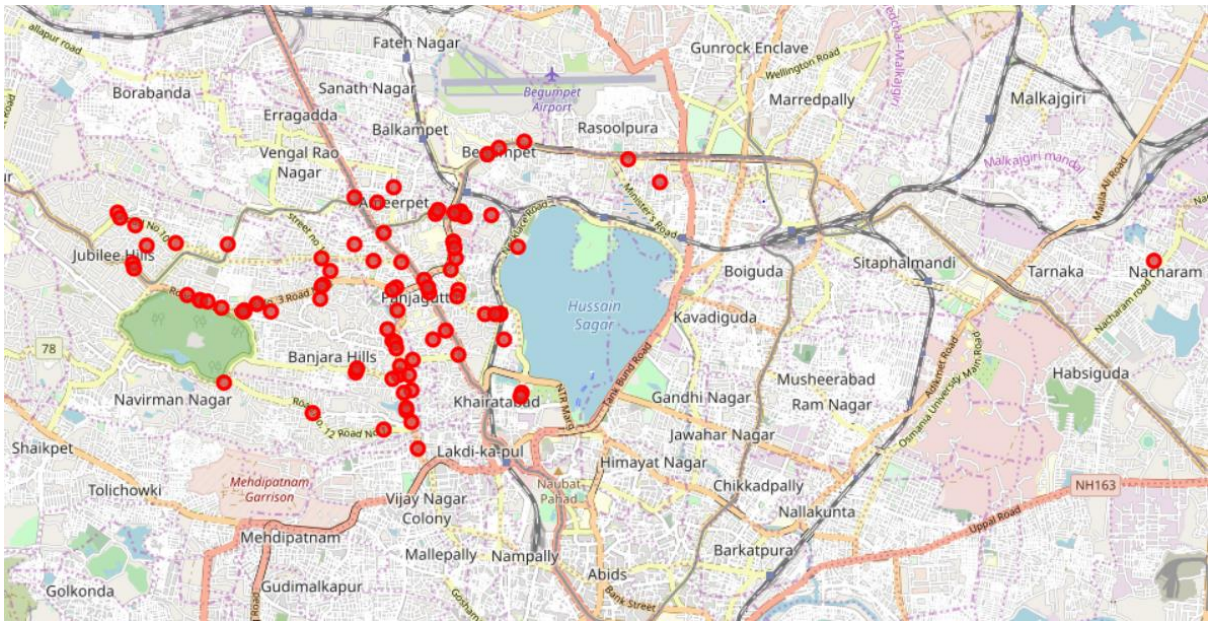

Figure 2: Venues retrieved from Zomato API

From figure 1 and figure 2, we can clearly see that some venues from the two APIs do not align with each other. Thus, I decided to combine them using their latitude and longitude values.

To combine the two datasets, I had to check that the latitude and longitude values of each corresponding venue match. After careful analysis, I decided to drop all corresponding venues from the two datasets that had their latitude and longitude values different by more than 0.0004 from one another. Thus, 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 0.0004 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 still some venues which were not correctly aligned.

They can be categorised as follows:

1. There are venues that have specific restaurants/cafes inside them as provided by Zomato API .
2. Two locations are so close that they have practically same latitude and longitude values.
3. Some venues have been replaced with new venues.

Venues belonging to category 1 and 3 are perfect to keep. However, the venues that belong to category 2 should be dropped. After careful inspection and removal, the final dataset had a total of 58 venues with which we can work.

As a final dataset, we're left with 58 venues with 8 columns as described in figure 3.

	categories	venue	latitude	longitude	price_range	rating	address	average_price
0	Diner	Hyderabad Street Kitchen	17.4335	78.4469	2.0	3.2	6-3-793, Officers Colony, Punjagutta, Kakatiya...	350.0
1	Pub	10 Downing Street	17.4359	78.4574	3.0	4.1	10, Ground Floor, My Home Tycoon, Begumpet, Hy...	900.0
2	Ice Cream Shop	Natural Ice Cream	17.4286	78.4396	1.0	4.6	G 3, Raghavendra Grand Residency, Opposite Rat...	100.0
3	Indian Restaurant	Minerva Coffee Shop	17.4288	78.4560	2.0	3.9	Ground Floor, Amrutha Mall, Opposite GRT Jewel...	400.0
4	Hotel	The Grill - Vivanta By Taj	17.4437	78.4609	4.0	3.9	Vivanta By Taj, 1-10-147 & 148, Mayuri Marg, B...	1100.0

Figure 3: Final data aggregated from both APIs

3. Methodology and Exploratory Data Analysis

As a first step, I retrieve the venues in Ameerpet, Hyderabad from Foursquare and Zomato APIs. I extract the location data from the Foursquare API for all venues up to a distance of 4 kilometers from the center of Ameerpet. Using this, I fetch the venue information including price and rating data from Zomato API.

Using data cleaning, the dataset from the two APIs will be combined based on the venue names, latitude, and longitude values. One to one matching and careful data inspection would be used to remove any remaining outliers such as multiple venues at the same location from the two datasets. The final data will include the venue name, category, address, latitude, longitude, rating, price range, and average cost per person.

Using this dataset, I begin by analyzing the top venue types that exist in Ameerpet. I will then explore the venues on maps. This will allow us to better understand the location of various venues and the places where many venues co-exist and create place worth visiting. I'll also explore the venues based on the ratings and price range of various venues. The venues will be plot using proper color coding such that a simple glance at the map would reveal the location of the venues as well as give information about them. I aim to identify places which can be recommended to visitors based on their price and rating preferences. I'll also cluster the venues and see if we can draw meaningful information out of what kind of venues exist in Ameerpet.

As a final step, I will analyse these plots and try to draw conclusions on what places can be recommended to visitors. I'll discuss my findings and any inferences I can draw.

3.1 Categories

I begin my analysis by taking a look at the various categories of venues that exist in Ameerpet. As there are many restaurants, I believe that the majority venues shall include restaurants.

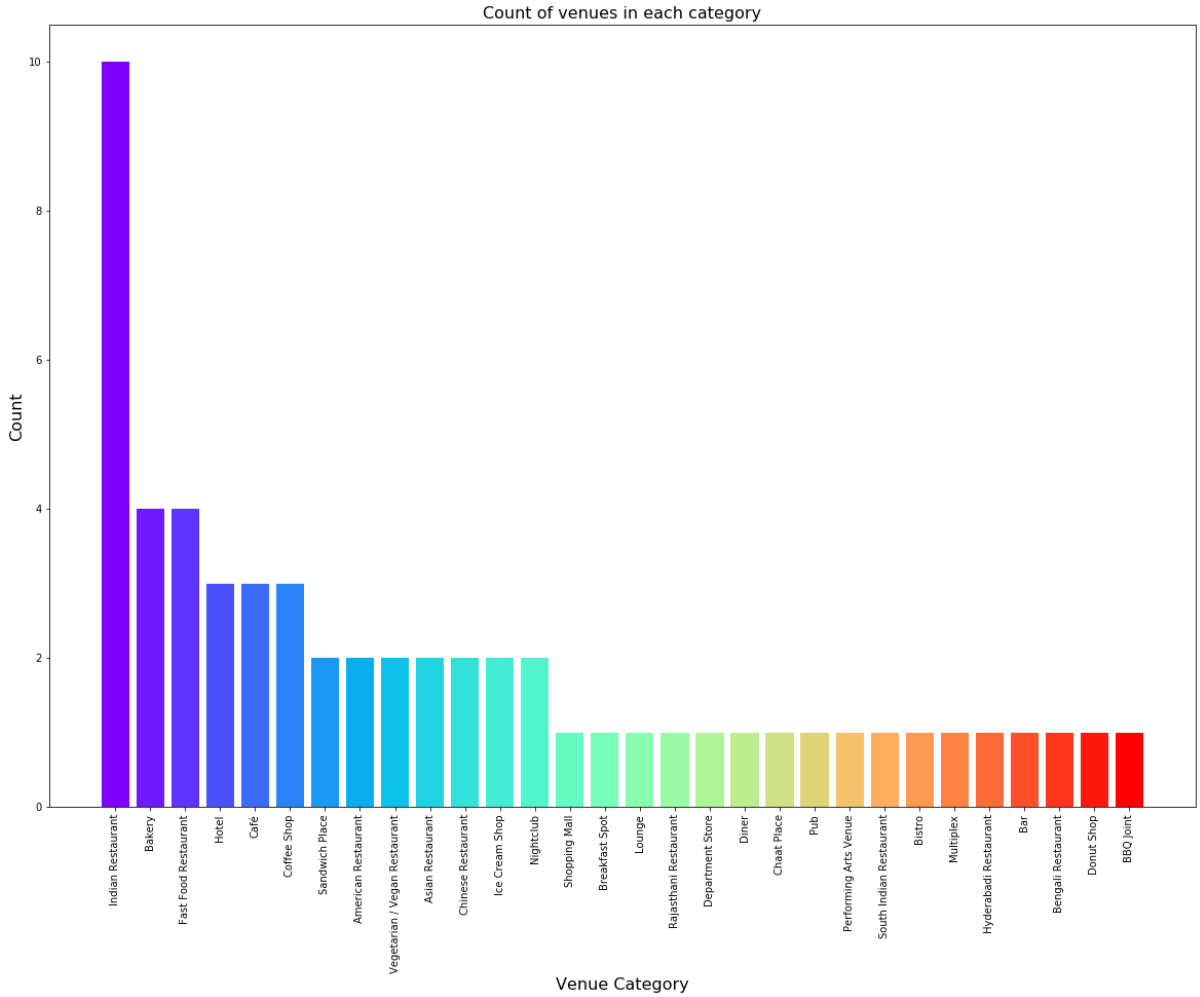

Figure 4: Count of various types of venues in Ameerpet

From figure 4, we see that the majority venues are actually Indian Restaurants. This is closely followed by Bakery and Fast food Restaurants. For someone who is visiting Ameerpet and loves either Fast foods or Indian Restaurants, they'd surely love their stay.

3.2 Rating

Next, I'll explore the ratings of various venues in Ameerpet. I decided to plot a bar chart with x-axis as the rating from 1 to 5 and the y-axis as the count of venues with that rating. I decided to plot the bar chart to see what average rating venues get in Hyderabad. This can be seen in figure 5.

While the whole range of rating of venues might stretch from 1 to 5, the average rating is spread across 4 with maximum number of venues scoring between 3 and 5.

I followed this information by plotting the venues on the map of Ameerpet. The venues that were rated below 3 were marked by red and orange while the venues that were rated more than or equal to 3 were plot as green and dark green. Taking a look at figure 6 reveals the same results as the bar plot. However, it is interesting to note that many high rated venues are located

near Road no.3 and Road no.1(Punjagutta). Punjagutta has venues with rating in the complete range from low to high. Also, the belt of venues around punjagutta have high rated venues.

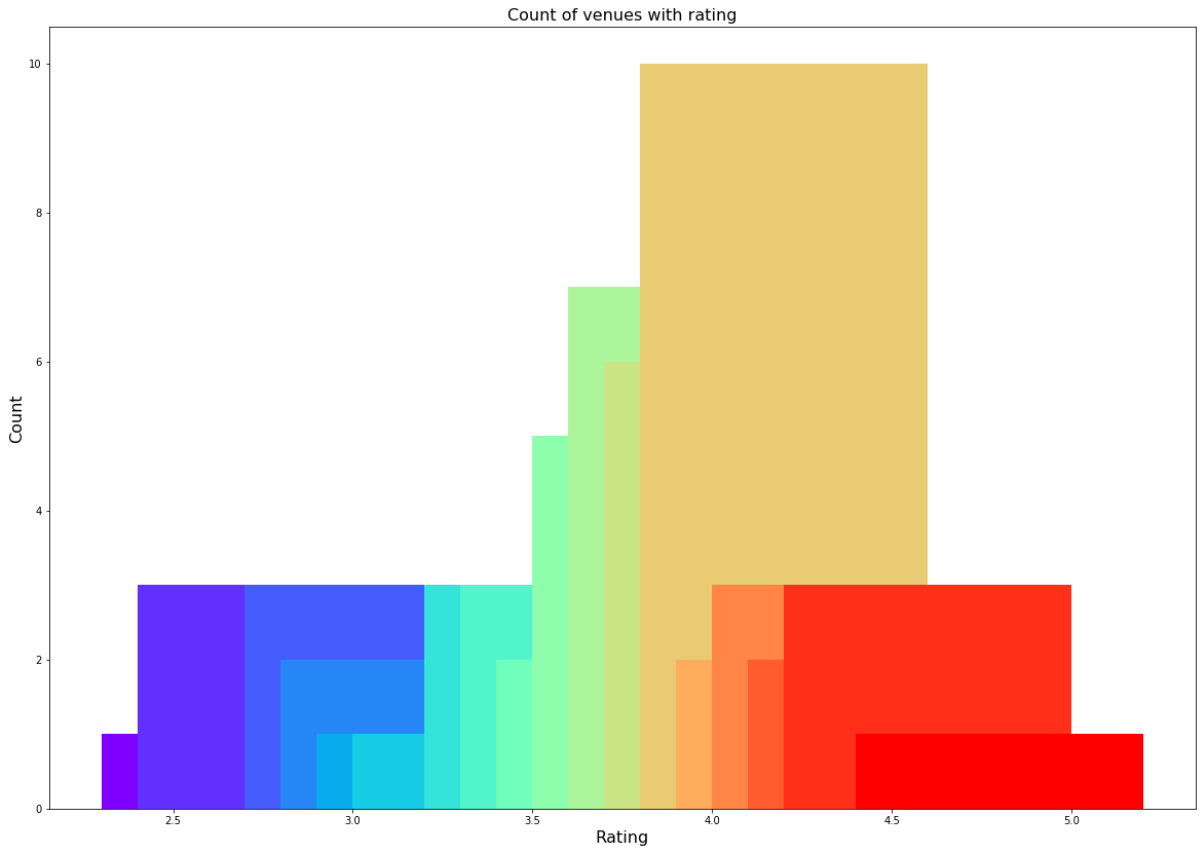

Figure 5: Rating and count of venues with that rating

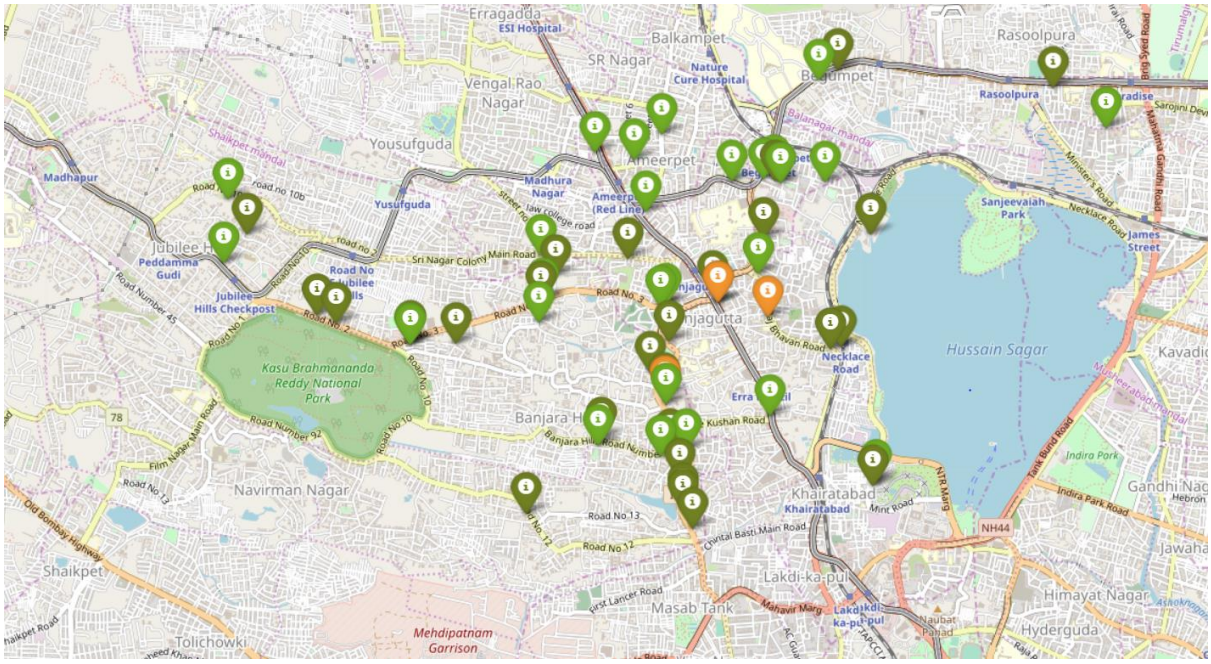

Figure 6: Plot of venues with different ratings

The venues in sectors that do not have many venues have rating more than 3. Overall, Ameerpet on an average has good rating for its venues.

3.3 Price

Next, I explore the average prices of all venues for one person using a scatter plot along with the count of venues with that average price per person. Taking a look at figure 7, reveals that the majority venues have an average cost of Rs 200 to Rs 400 for one person. Even though the maximum venues lie in that range, the actual range of prices is very different. There are places with average price even as high as Rs 1400+ for one person.

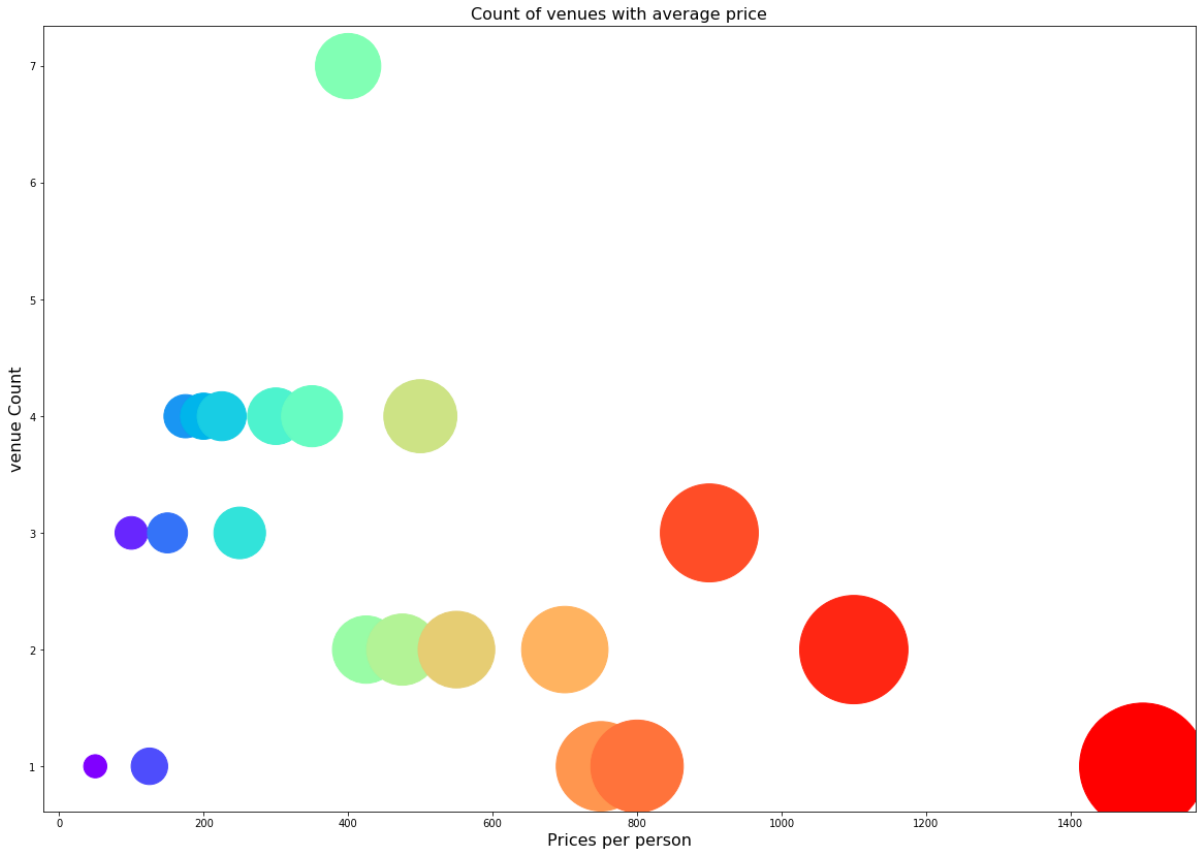

Figure 7: Price per person with count of venues with that price

I also plot the venues based on their price range.

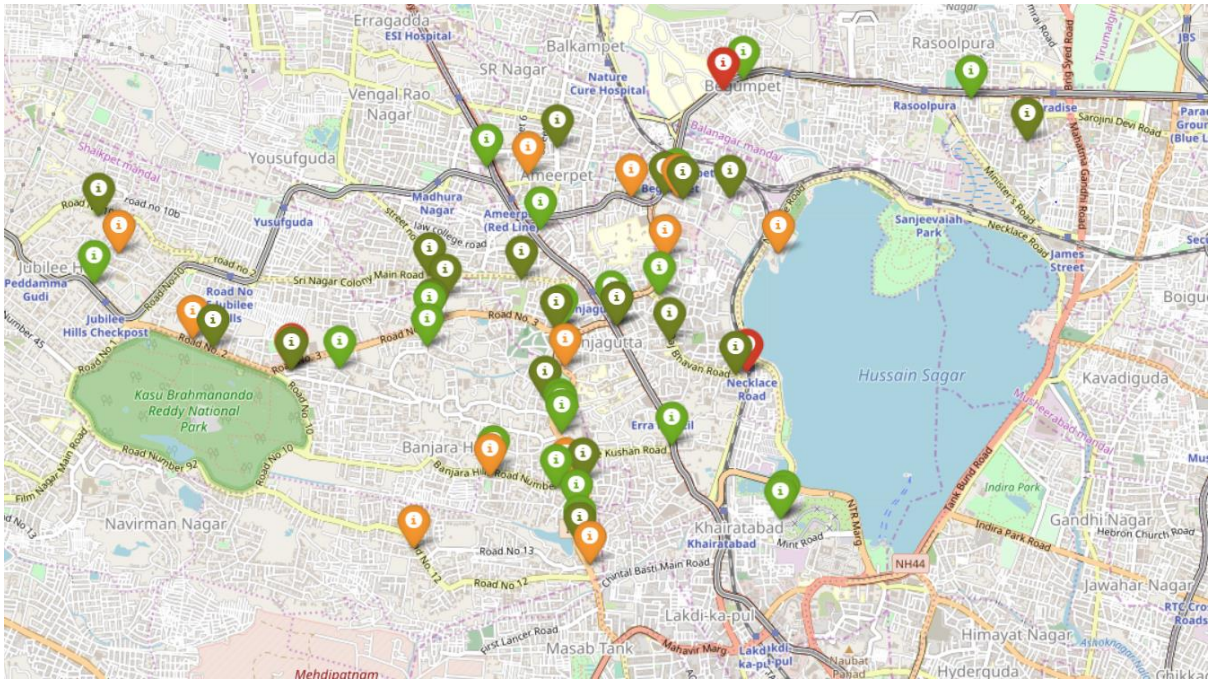

Figure 8: Plot of venues with different prices

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. From the plot, we observe that venues near Punjagutta and Srinagar colony road are primarily lower priced. The venues near Road no.3 Necklace road and Begumpet have steep prices.

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.

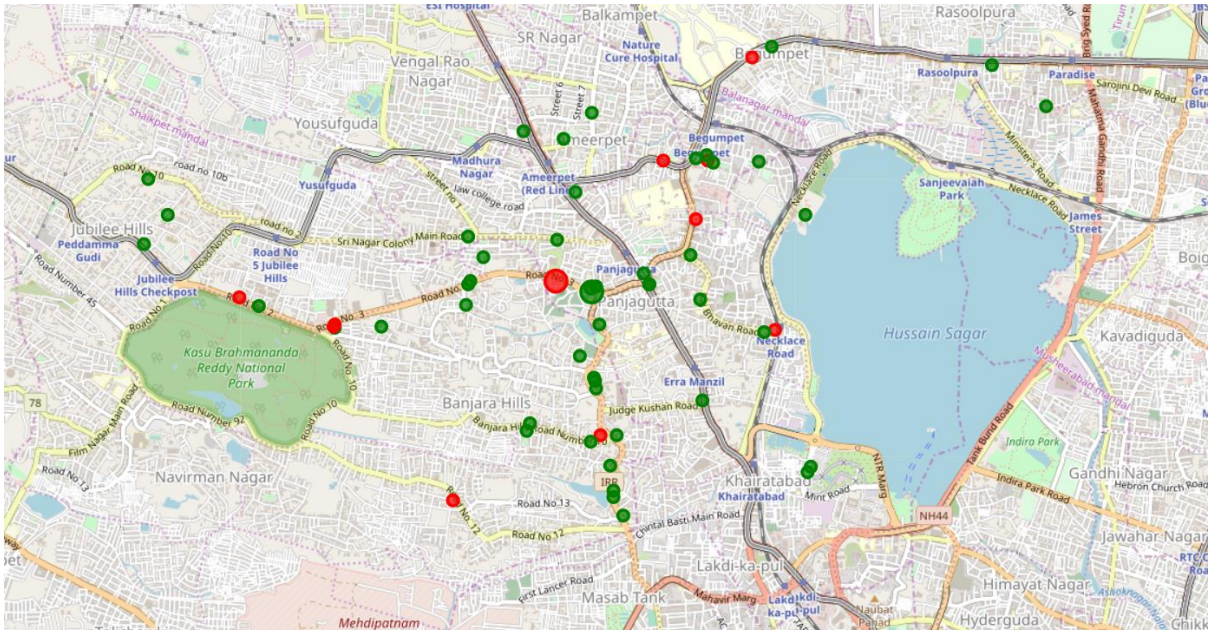

Figure 9: Clusters of venues

In figure 9, we see the two clusters:

1. The first cluster (green) is spread across the whole city and includes the majority venues. These venues have mean price range of 1.71 and rating spread around 3.85.
2. The second cluster (red) is very sparsely spread and has very limited venues. These venues have mean price range of 3.3 and rating spread around 4.15.

4. Results and Discussion

Based on our analysis above we can get some conclusions which can be useful for any visitor visiting Hyderabad (Ameerpet) India.

From combining four_square and Zomato APIs we got a list of around 150 different venues, and not all the venues are identical. By inspecting the latitude and longitude values and removing outliers we came with 58 selected venues.

By analysing all the venues we observed that **majority of venues are Indian Restaurants Fast food centers and Bakery** A visitor who is searching for items fried fast food items like fried rice or noodles can get food easily and also we have many Indian Restaurants here.

The **majority of venues are with rating of 4** which means that majority of restaurants are providing good quality of food. We also observed that there are cluster of venues in Road no.3 banjara hills, Begumpet, and Road no.1 Banjara hills near punjagutta with ratings which are Good.

Based on the price range there are many restaurants with average price range of 200-400 in the city. Also, the price range varies from 100 - 3000. on plotting the map of this venues we observed costlier venues in Begumpet, Necklace Road and Near Lv Prasad eye hospital Road no 3. Near punjagutta we can observe many restaurants with low prices.

Finally, from the clusters we **identified venues** which are with **relatively low prices and average rating of 3.85**. Also, few venues which are **too expensive but have a good rating of 4.15**

5. Conclusion

The purpose of this project is to explore restaurants near Ameerpet for the people who are visiting Hyderabad. Also, to taste good food. The venues have been identified using Foursquare and Zomato API and have been plotted on the map. The map reveals that there are three major areas a person can visit: **`Road no.3 banjara hills, Begumpet, and Road no.1 Banjara hills near punjagutta`.**