

# **Exploring venues in Hi-tech city Hyderabad, India using Foursquare and Zomato API**

## **1. Introduction**

### **1.1 Background**

Any person who visits or planning to travel to new city will tend to visit good cuisine which is available best in the city based on certain parameters like varieties in cuisine, budget friendly and types of restaurants available around the city. In order to give better insights combining the location venues of the city with the price and rating information would surely help visitors in a new city.

Hi-Tech city is the prime location in Hyderabad, it has many venues specially restaurants, coffee shops, hotels etc. Which can be explored, this project explores various venues in Hi-tech city and attributes the data based on the positions of both the Foursquare API and Zomato API to fetch complete information of various venues (name, add, category, rating and price). Further, a map of the venues with specific colour attributes will be plotted to highlight their position and information about the venues. This enables any visitor to take a quick glance and make the decision.

### **1.2 Interested audience**

The target audience are who are visiting Hi-Tech Hyderabad 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 Hi-tech city Hyderabad, used two APIs and combined the data from both together.

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

Using the name, latitude and longitude values, used 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.

From Foursquare API (<https://developers.zomato.com/api>), 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. Later converted 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.

## 1.1 2.2 Data Cleaning

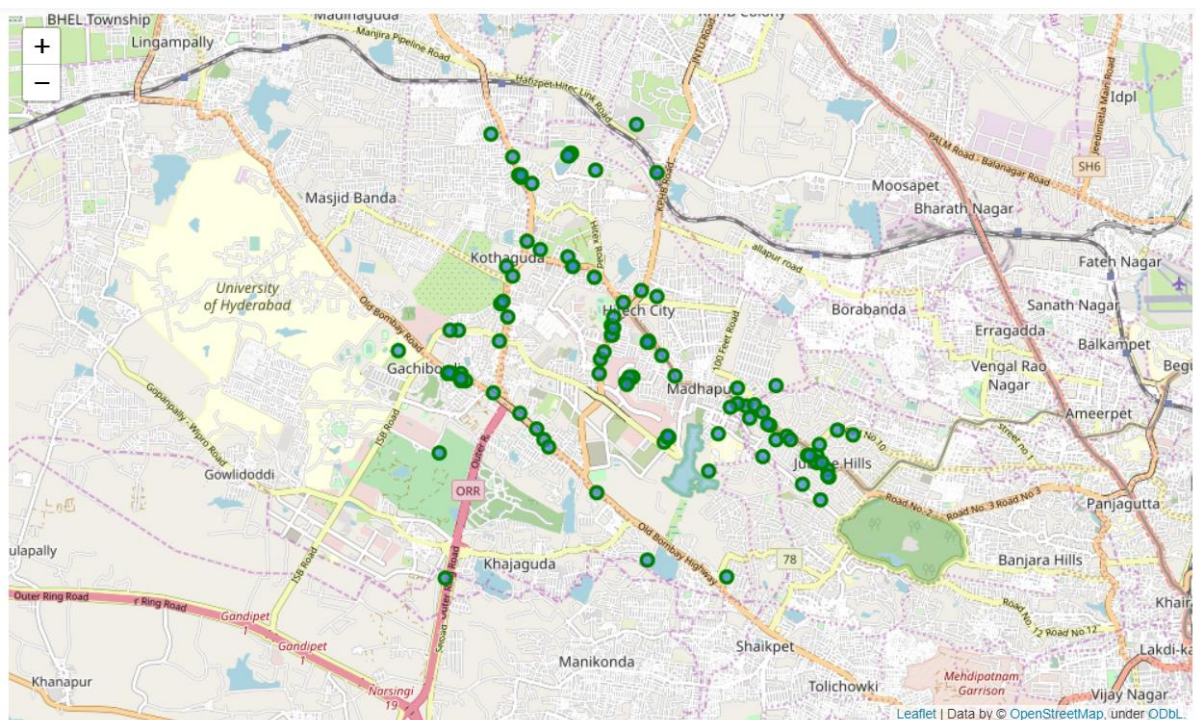
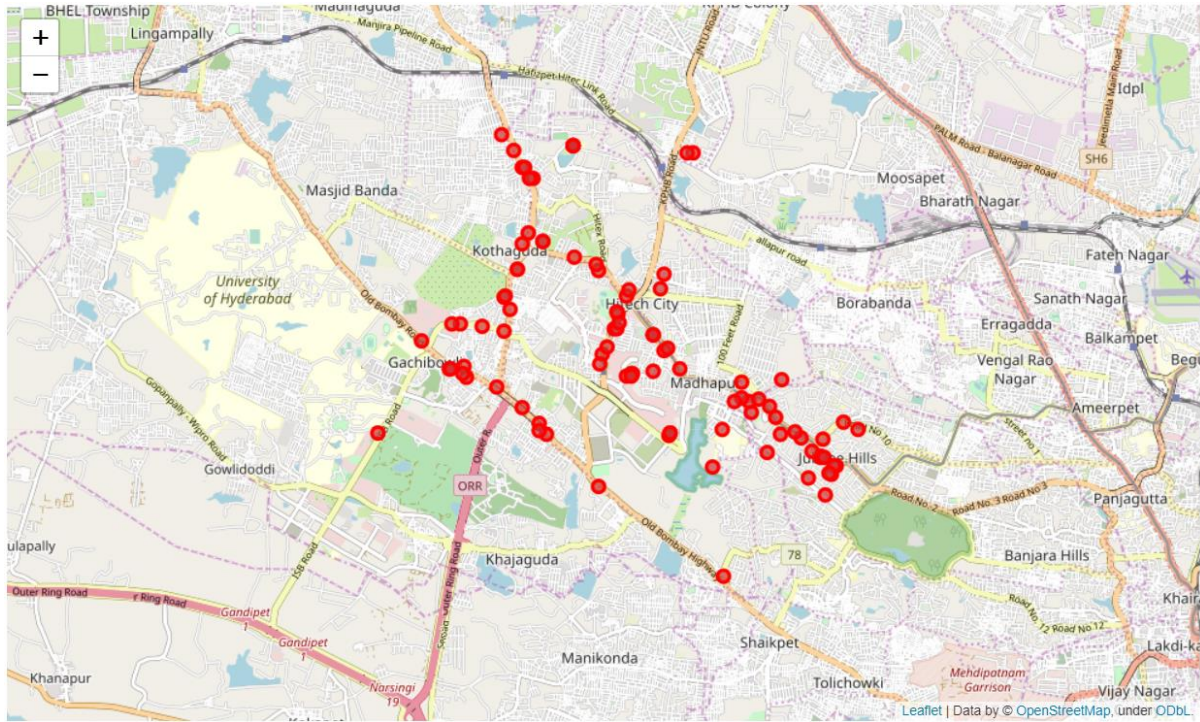


Figure 1: Venues retrieved from Foursquare API



*Figure 2: Venues retrieved from Zomato API*

From figure 1 and figure 2, it is observed that some venues from the two APIs do not align with each other. Combined both the datasets using their latitude and longitude values.

To combine the two datasets, checked that the latitude and longitude values of each corresponding venue match. After analyzing dropped all corresponding venues from the two datasets that had their latitude and longitude values different by more than 0.0004 from one another. rounded both the latitude and longitude values up to 4 decimal places. Then, calculated the difference between the corresponding latitude and longitude values and observed 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, It is observed that there were still some venues which were not correctly aligned.

They can be categorized 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 49 venues with which we can work.

As a final dataset, we're left with 59 venues with 8 columns as described in figure 3.

	categories	venue	latitude	longitude	price_range	rating	address	average_price
0	Coffee Shop	Starbucks Coffee	17.4480	78.3788	2.0	4.2	Upper Ground Floor, Phoenix Tower A, Opposite ...	350.0
1	Restaurant	Firewater Kitchen & Bar	17.4481	78.3790	3.0	4.2	5th Floor, Phoenix Tower A, Opposite Trident H...	650.0
2	Bakery	Karachi Bakery	17.4451	78.3857	1.0	4.4	Shop 05, SBR CV Towers, Plot 2, HUDA Techno En...	150.0
3	Indian Restaurant	Ohri's Nautanki Gali	17.4447	78.3770	3.0	4.5	Plot 9, Survey 64, Cyber Pearl, Madhapur, Hite...	550.0
4	Hotel	Pi - Hotel Avasa	17.4472	78.3842	4.0	4.1	Hotel Avasa, Survey 64, Sector 1, HUDA Techno ...	1000.0

*Figure 3: Final data aggregated from both APIs*

### 3.Methodology and Exploratory Data Analysis

As a first step, retrieved venues in Hi-tech city Hyderabad from Foursquare and Zomato APIs. Then extracted location data from the Foursquare API for all venues up to a distance of 4 kilometers from the center of Hi-tech city. Using this, retrieved 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 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, started analyzing the top venue types that exist in Hi-tech city. Then explored 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. And explored 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. Aim is to identify places which can be recommended to visitors based on their price and rating preferences. Also clustered the venues and see if we can draw meaningful information out of what kind of venues exist in Hi-tech city.

As a final step, analyzed these plots and try to draw conclusions on what places can be recommended to visitors.

#### 3.1 Categories

Started analysis by looking at various categories of venues that exist in Hi-tech city. As there are many restaurants.



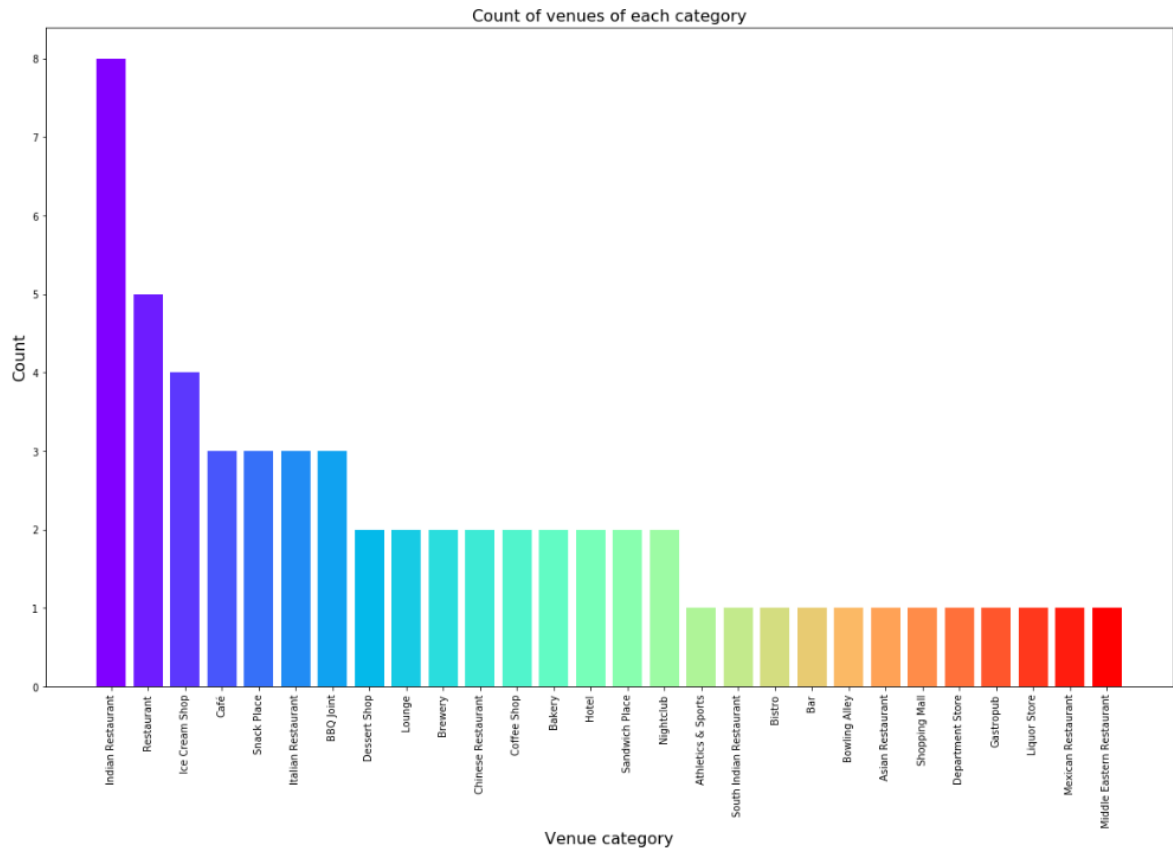


Figure 4: Count of various types of venues in Hi-tech city

### 3.2 Rating

Next, explored the ratings of various venues in Hi-tech city. 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. To see what the average rating of venues in Hi-Tech city are. 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.

Plotting these venues on the map. 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. Take a look at figure 6 reveals the same results as the bar plot.

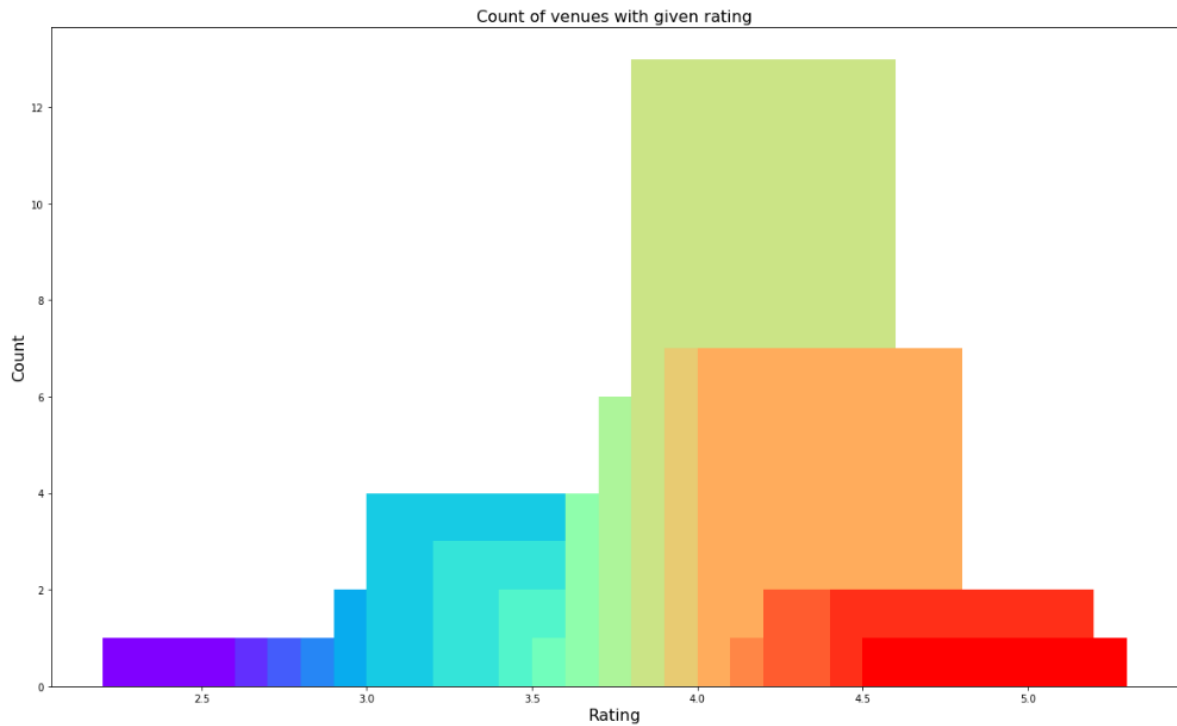


Figure 5: Rating and count of venues with that rating

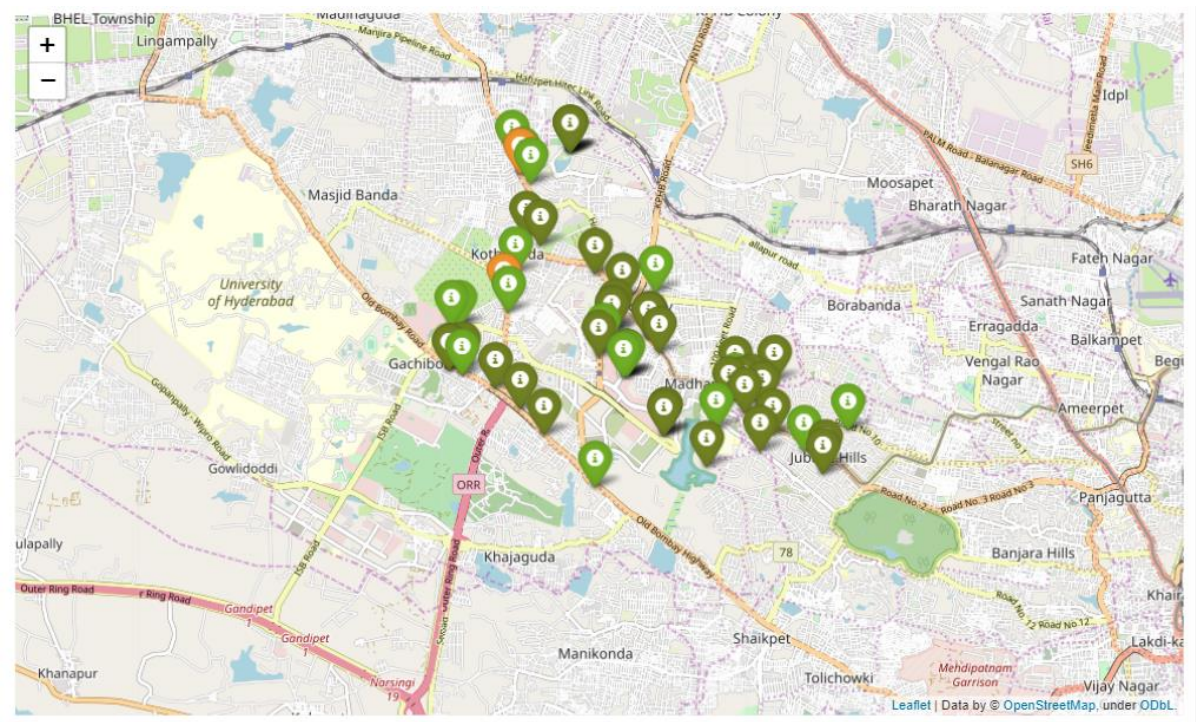


Figure 6: Plot of venues with different ratings

### 3.3 Price

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. Looking 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 1000+ for one person.

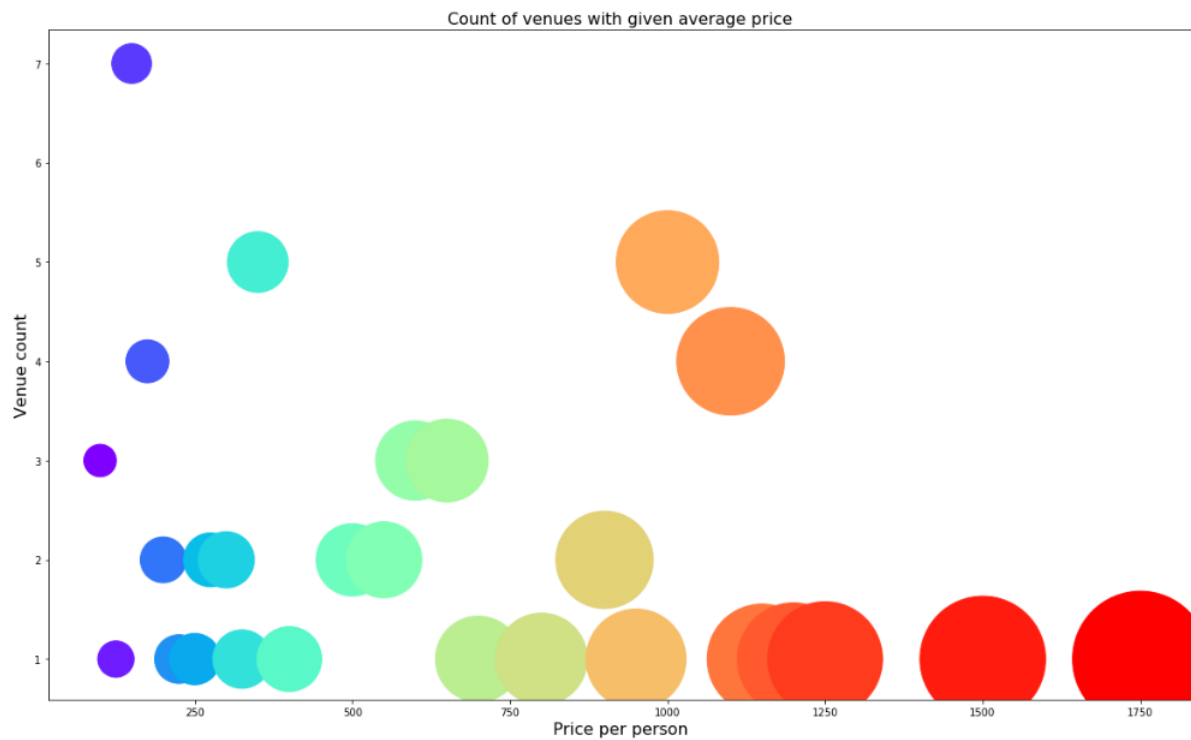


Figure 7: Price per person with count of venues with that price

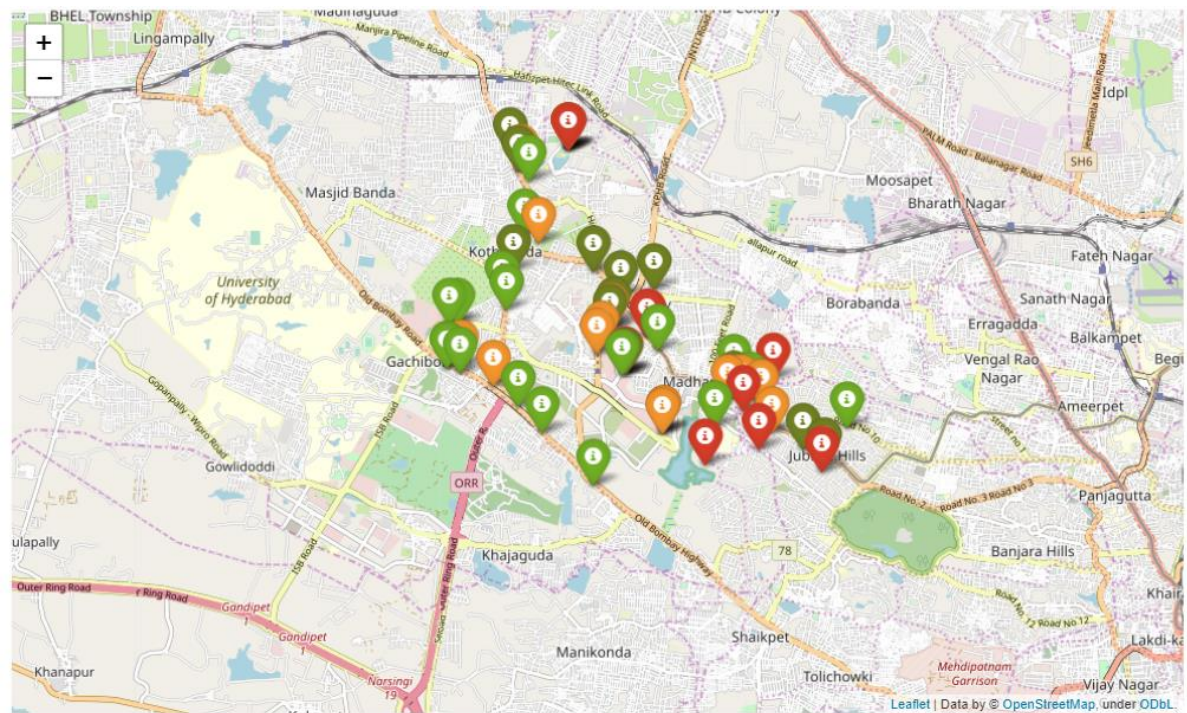


Figure 8: Plot of venues with different prices



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.

### 3.4 Clustering

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

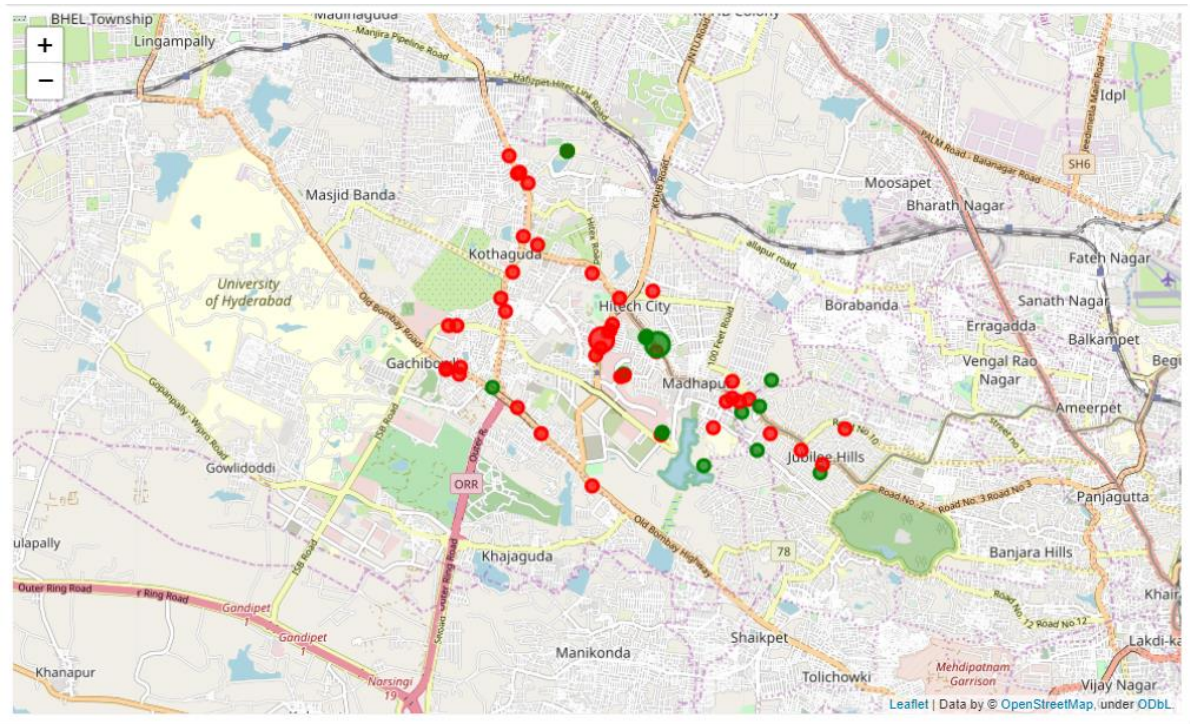


Figure 9: Clusters of venues

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 3.57.
2. The second cluster (red) is very sparsely spread and has very limited venues. These venues have mean price range of 3.21 and rating spread around 4.03.

## 4. Results and Discussion

After collecting data from the Foursquare and Zomato APIs, we got a list of 100+ different venues. However, not all venues from the two APIs were identical. Analyze to inspect their latitude and longitude values as well as their names to combine them and remove all the outliers. This resulted in a total venue count of 59.



Identified from the total set of venues, majority of them were Restaurants and Ice cream shops. A visitor who loves Indian Restaurants/ Ice cream would surely benefit from coming to Hi-Tech city.

While the ratings range from 1 to 5, majority venues have ratings close to 4. This means that most restaurants provide good quality food which is liked by the people of the city, thus indicating the high rating.

looking at the price values of each venue, many venues have prices which are in the range of Rs 200 to Rs 400 for one person. However, the variation in prices is very large, given the complete range starts from Rs 100 and goes up to Rs 1200+ On plotting the venues based on their price range on the map.

A company can use this information to build an online website/mobile application, to provide users with up to date information about various venues in the city based on the search criteria (name, rating and price).

## **5. Conclusion**

The purpose of this project was to explore the places that a person visiting Hi-Tech city could explore. The venues have been identified using Foursquare and Zomato API and have been plotted on the map.