Exploring venues in Chandigarh, India using Foursquare and Zomato API

June 12, 2020

1. Introduction

1.1 Background

Whenever a person visits a city for the first time, he/she searches for a venue in a city, where they can stay and dine. While doing so he/she wants to evaluate the venues available in the city from various perspectives like quality, reachability, value for money and affordability. The person might want to know how good a given restaurant is or the price range it falls under. This extra information would help him/her decide which venue to choose amongst the many venues available 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.

Chandigarh is composed of a number of sectors spread across a total area of 114 sq Km. There are many venues (especially restaurants, hotels and cafes) which can be explored. This project explores various venues in Chandigarh city and analyze them based on attribute data like user ratings, average price and location. To explore this information, this project involves the merging of information extracted from two sources, the Foursquare API and the Zomato API. The complete information of various venues includes 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 embed meaningful 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 multi-fold. First, any person who is visiting Chandigarh, India can use these plots and maps from this project to quickly select places that suit their budget and rating preferences. Second, a tourist company can use this information to create a website or a mobile application, which is updated on a regular basis, to provide value added services to its customers. Third the person or company who is planning to setup a new venue in Chandigarh can use these maps to identify and finalize the suitable location for their startup.

2. Data

2.1 Data Sources

To get location and other information about various venues in Chandigarh, 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 5 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://api.foursquare.com/v2/venues/explore), 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

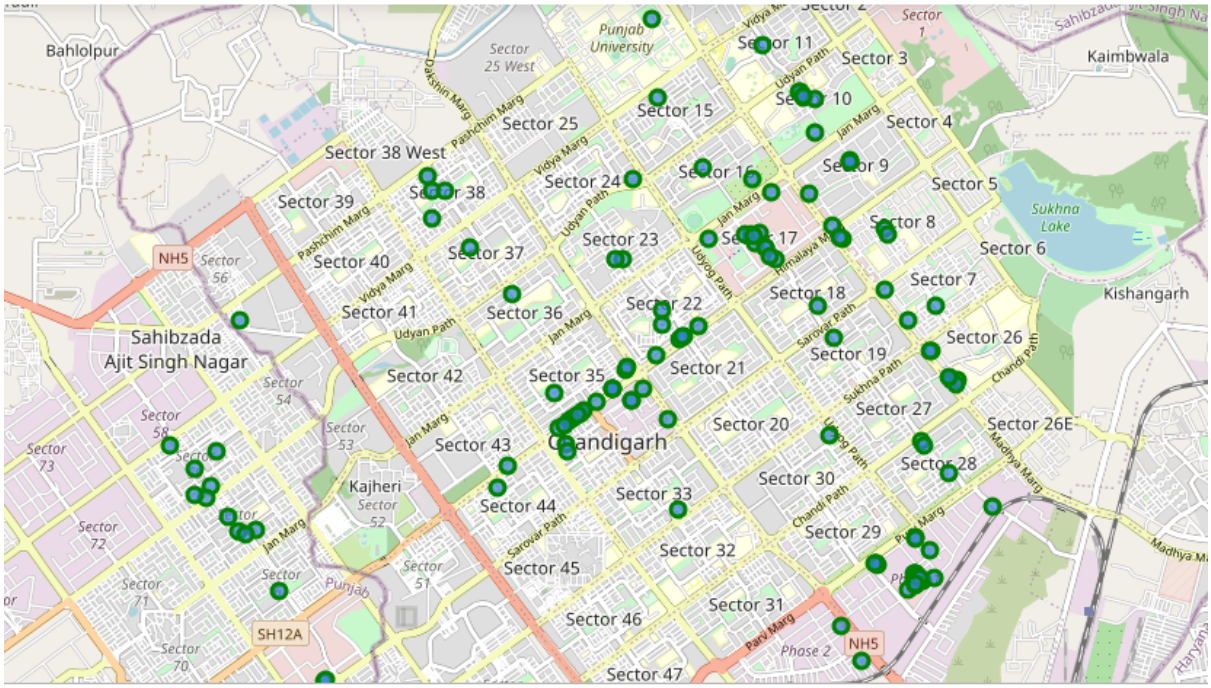

Figure 1: Venues retrieved from Foursquare API

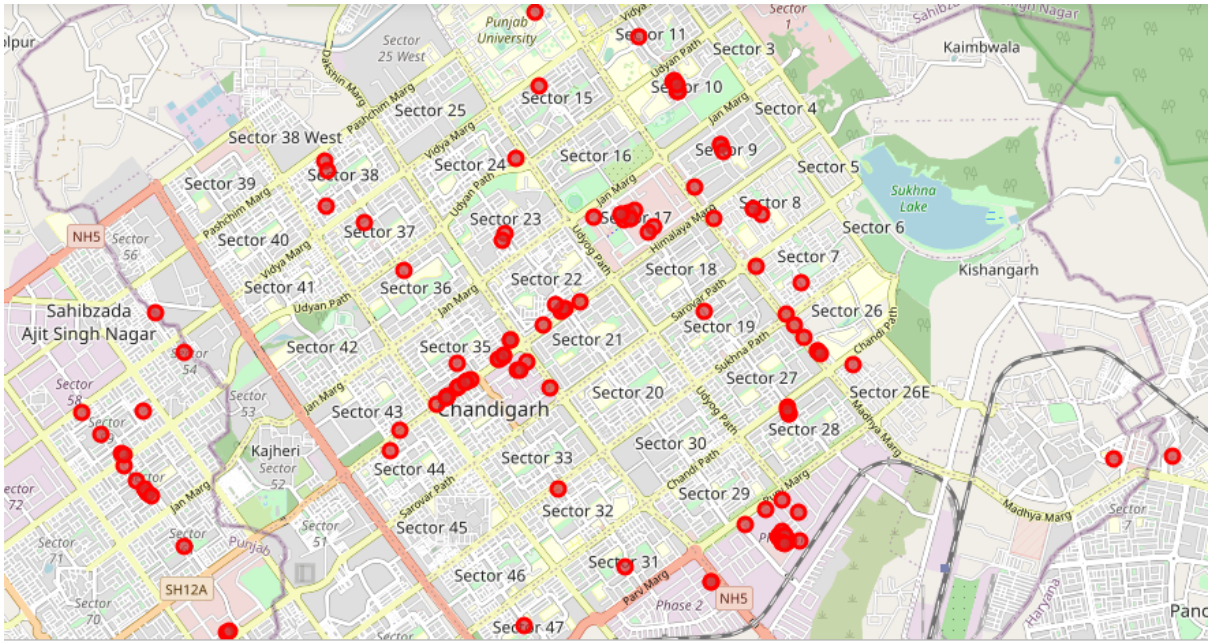

Figure 2: Venues retrieved from Zomato API

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 those in another. To proceed with, I first 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 assumed that if the difference in latitude and longitude both is less than 0.0004 then logically both the locations are the same. This removed many outliers from the two datasets. Once this was done, I observed that there were some venues for which the rating value was zero, which means these venues were never rated by any user. Hence, considering these venues in our analysis is virtually of no use. Hence, I decided to drop these venues as well.

After careful inspection and cleaning of data, the final dataset that was left for analysis had a total of 63 venues with which we can work.

As a final dataset, we're left with 63 venues with 8 columns as described in figure 3.

	categories	venue	latitude	longitude	price_range	rating	address	average_price
0	Bakery	Nik Baker's	30.7216	76.7601	2.0	4.5	SCO 441 - 442, Sector 35 C, Sector 35, Chandigarh	450.0
1	Coffee Shop	Starbucks Coffee	30.7299	76.7732	2.0	4.0	Sector 22 C, Near Sector 22, Chandigarh	250.0
2	American Restaurant	OvenFresh	30.7217	76.7604	2.0	4.7	SCO 437 & 438, Sector 35 C, Sector 35, Chandigarh	425.0
3	Bakery	Classic 44	30.7148	76.7517	1.0	3.3	SCO 413, Sector 44 D, Near Sector 44, Chandigarh	100.0
4	Hotel Bar	McDonald's	30.7302	76.7737	2.0	4.3	Himalaya Marg, Sector 22 C, Near Sector 22, Ch...	250.0

Figure 3: Final data aggregated from both APIs

3. Methodology and Exploratory Data Analysis

As a first step, I retrieve the venues in Chandigarh from Foursquare and Zomato APIs. I extract the location data from the Foursquare API for all venues up to a distance of 5 kilometers from the center of Chandigarh. Using this, I fetch the venue information including price and rating data from Zomato API.

Using data cleaning, the dataset from the two APIs were 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 began with the analyzing of the top venue types that exist in Chandigarh. I then explored the venues on maps. This allowed us to study the location of various venues and the places where many venues co-exist and place worth visiting. I also explored the venues based on the ratings and price range of various venues. The venues were plot on the map 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 also clustered the venues and see if we can draw meaningful information out of what kind of venues exist in Chandigarh.

As a final step, I analyzed these plots and try to draw conclusions on what places can be recommended to visitors. I have discussed my findings and what inferences I can draw.

3.1 Categories

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

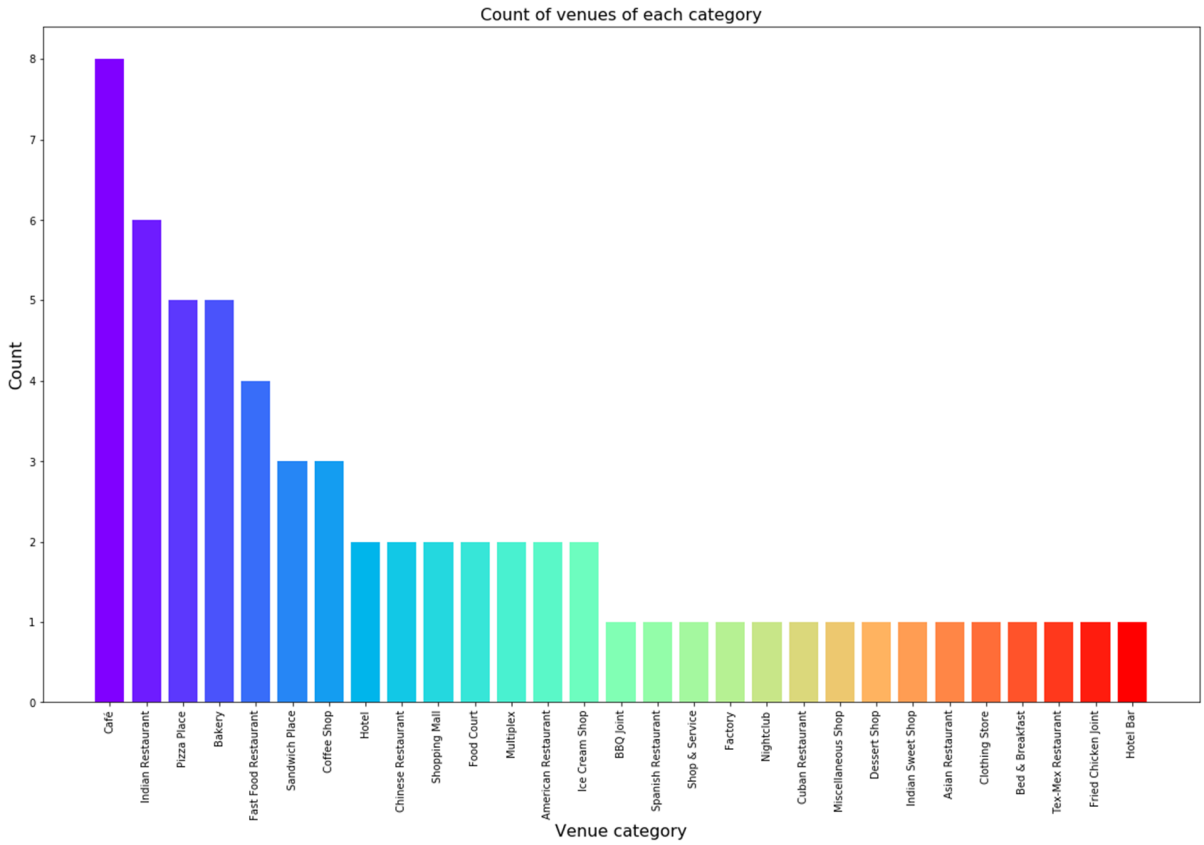

Figure 4: Count of various types of venues in Chandigarh

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

3.2 Rating

Next, I explored the ratings of various venues in Chandigarh. 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 Chandigarh. This can be seen in figure 5.

While the range available to rate the venues was from 1 to 5, but the average rating for the venues 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 Chandigarh. 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 blue. 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 Sector 35, and Sector 17. Whereas, Elante Mall has venues with wide range of rating from low to high. Also, the venues from Sector 7 and Sector 26 have high rated venues.

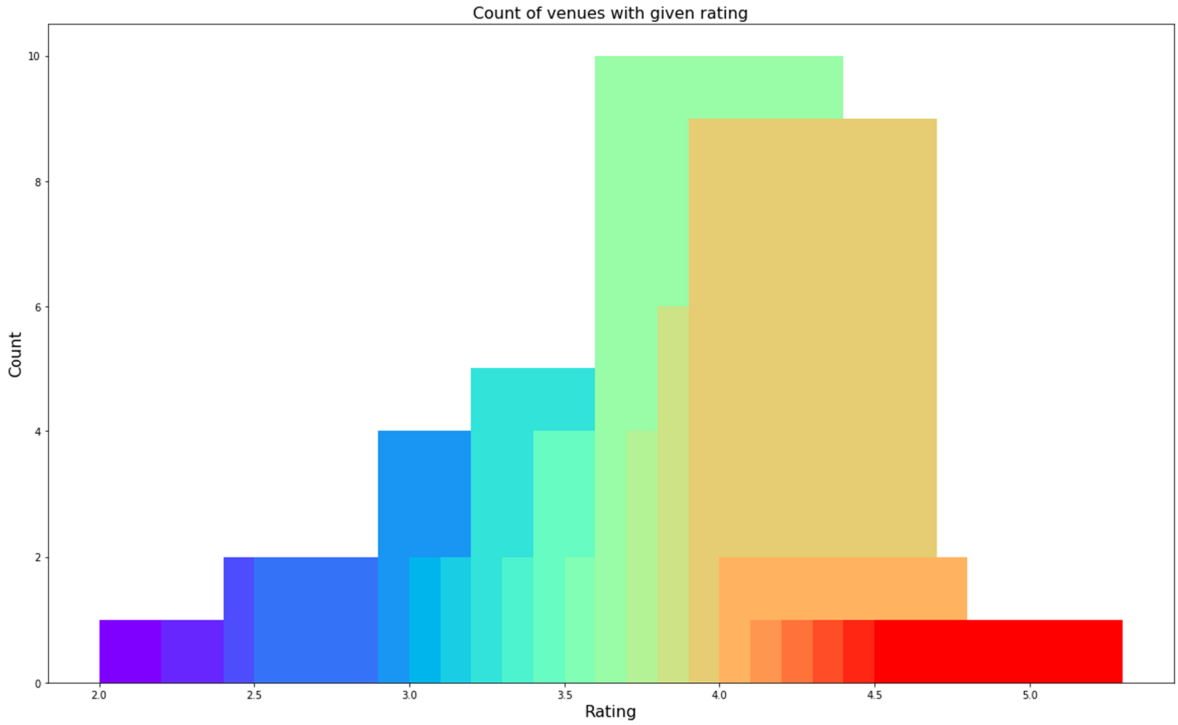

Figure 5: Rating and count of venues with that rating

Figure 6: Plot of venues with different ratings

The venues in sectors that have few venues have rating more than 3. Overall, venues in Chandigarh have on an average a good rating for its venues.

3.3 Price

Next, I explored 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 1000+ for one person.

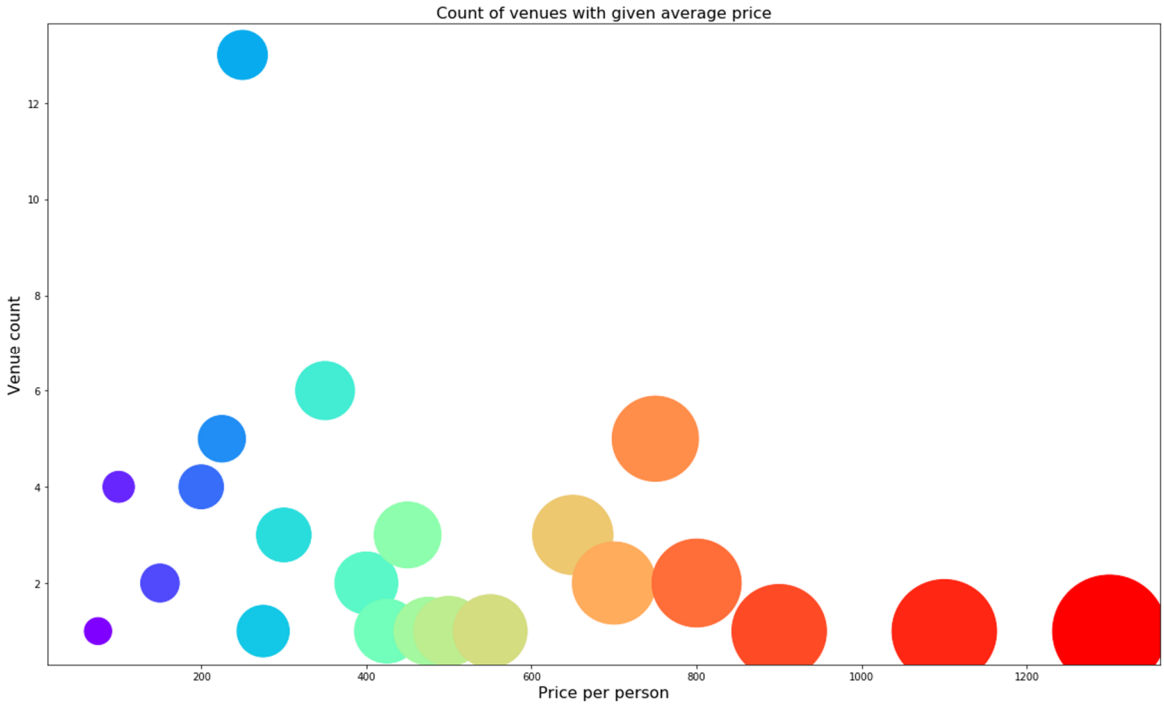

Figure 7: Price per person with count of venues with that price

I also plot the venues based on their price range.

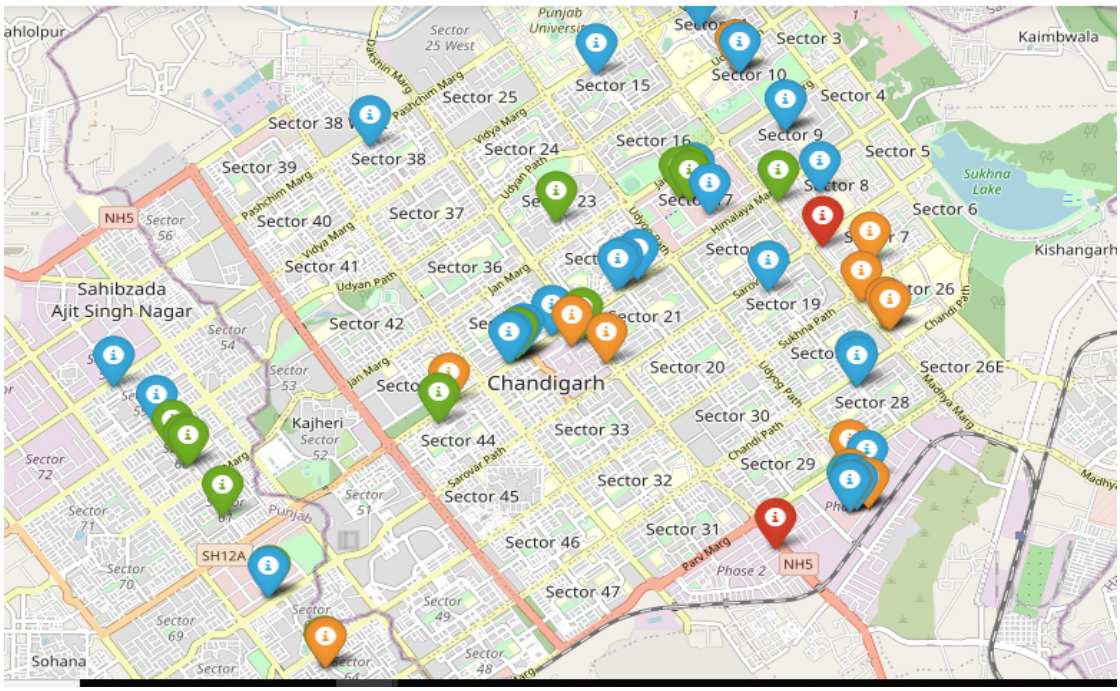

Figure 8: Plot of venues with different prices

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 blue. From the plot, we observe that venues near Sector 35 and Sector 17 are primarily lower priced. The venues near Sector 7 and Sector 26 have steep prices. Elante Mall seems to have a mix of both high priced and low priced venues.

3.4 Clustering

Finally, I clusterd 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.

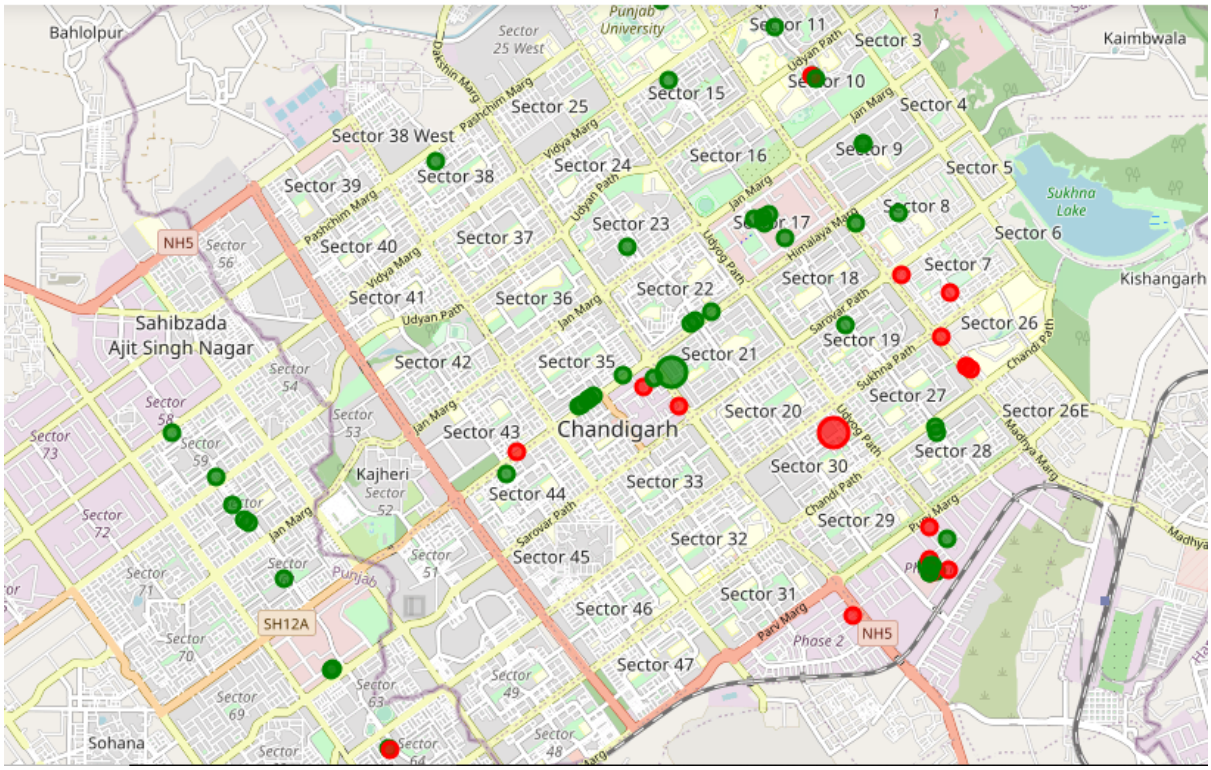

Figure 9: Clusters of venues

In figure 9, we see the two clusters:

1. The first cluster (green) is spread across the whole city and includes the majority of venues. These venues have mean price range of 1.68 and rating spread around 3.82.
2. The second cluster (red) is very sparsely spread and has very limited venues. These venues have mean price range of 3.12 and rating spread around 4.04.

4. Results and Discussion

After collecting data from the Foursquare and Zomato APIs, we got a list of 130 different venues. However, not all venues from the two APIs were identical. Hence, we had 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 63.

We identified that from the total set of venues, majority of them were Cafes and Indian Restaurants. A visitor who loves Cafes/Indian Restaurants would surely benefit from coming to Chandigarh.

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. When we plot these venues on the map, we discovered that there are clusters of venues around Sector 17, Sector 35 and Elante Mall. These clusters also have very high ratings (more than 3).

When we take a look at the price values of each venue, we observed that 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 uptil Rs 1200. On plotting the venues based on their price range on the map, we discovered that venues located near Sector 17 and Sector 35 are relatively priced lower than venues in Sector 7 and Sector 26. A mix of low price and high price exists in Elante Mall.

Finally, through clustering the venues, we identified that there are many venues which are relatively lower priced but have an average rating of 3.82. On the other hand, there are few venues which are high priced and have average rating of 4.04.

1. If you're looking for cheap places with relatively high rating, you should check Sector 35.
2. If you're looking for the best places, with the highest rating but might also carry a high price tag, you should visit Sector 7 and Sector 26.
3. If you're looking to explore the city and have no specific criteria to decide upon the places you want to visit, you should try Elante Mall.

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

- Customer Perspective
  - We can conclude that Sector 35, 17 and Elante Mall have venues with good quality of food and service and yet they have wide price range. Hence there are venues available for different category (Middle to High class) of consumers.
  - Venues in Sector 7 and 26 can only be afforded by upper-middle and high-class customers.
- Business Perspective
  - In Sector 35, 17 and Elante Mall venues are densely available, hence there will be high competition if anyone plans to setup new venue.
  - However, in spite of high competition the probability of success is very high due to heavy foot fall of customers in these sectors. Hence the risk is relatively low.
  - We can also explore other sectors in the proximity of sector 35, 17 and Elante Mall for setup of new venue. As these sectors have potential to get developed in future