Capstone Project Report

Exploring venues of a location and using Foursquare

<u>Introduction</u>

- 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.Chandigarh is composed of a number of sectors spread across a total area of 114 sq Km.This project explores various venues in Chandigarh and attributes the data based on user ratings and average price. To explore this information, this project involves the juxtaposition 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.

Data Source And Data Cleaning

- To get location and other information about various venues in Chandigarh, two APIs were used and decided to combine the data from both of them together. Using the Foursquare's explore API (which gives venues recommendations), venues were 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, Zomato search API is used 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.

- **Name:** The name of the venue.

- **Category:** The category type as defined by the API.

- **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.

Data Analysis

- As a first step,retrieved the venues in Chandigarh from Foursquare and Zomato APIs.Extracted the location data from the Foursquare API for all venues up to a distance of 4 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 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,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, also the venues based on the ratings and price range of various venues is explored. 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. Venues analysis is done based on rating, price.
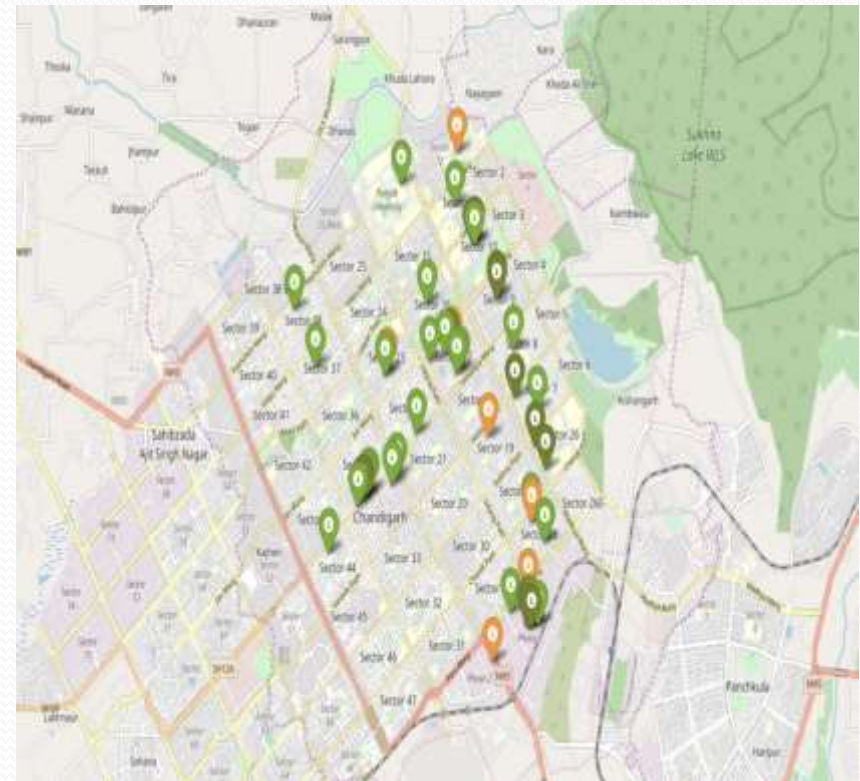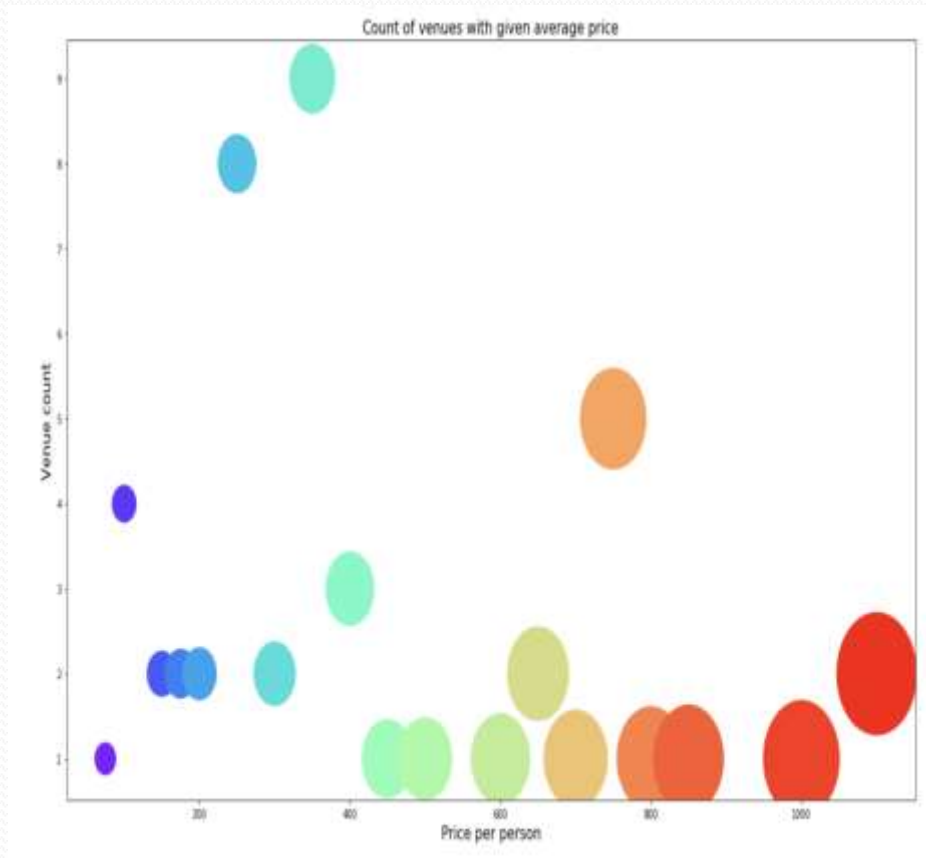
Categories

Count of venues of each category
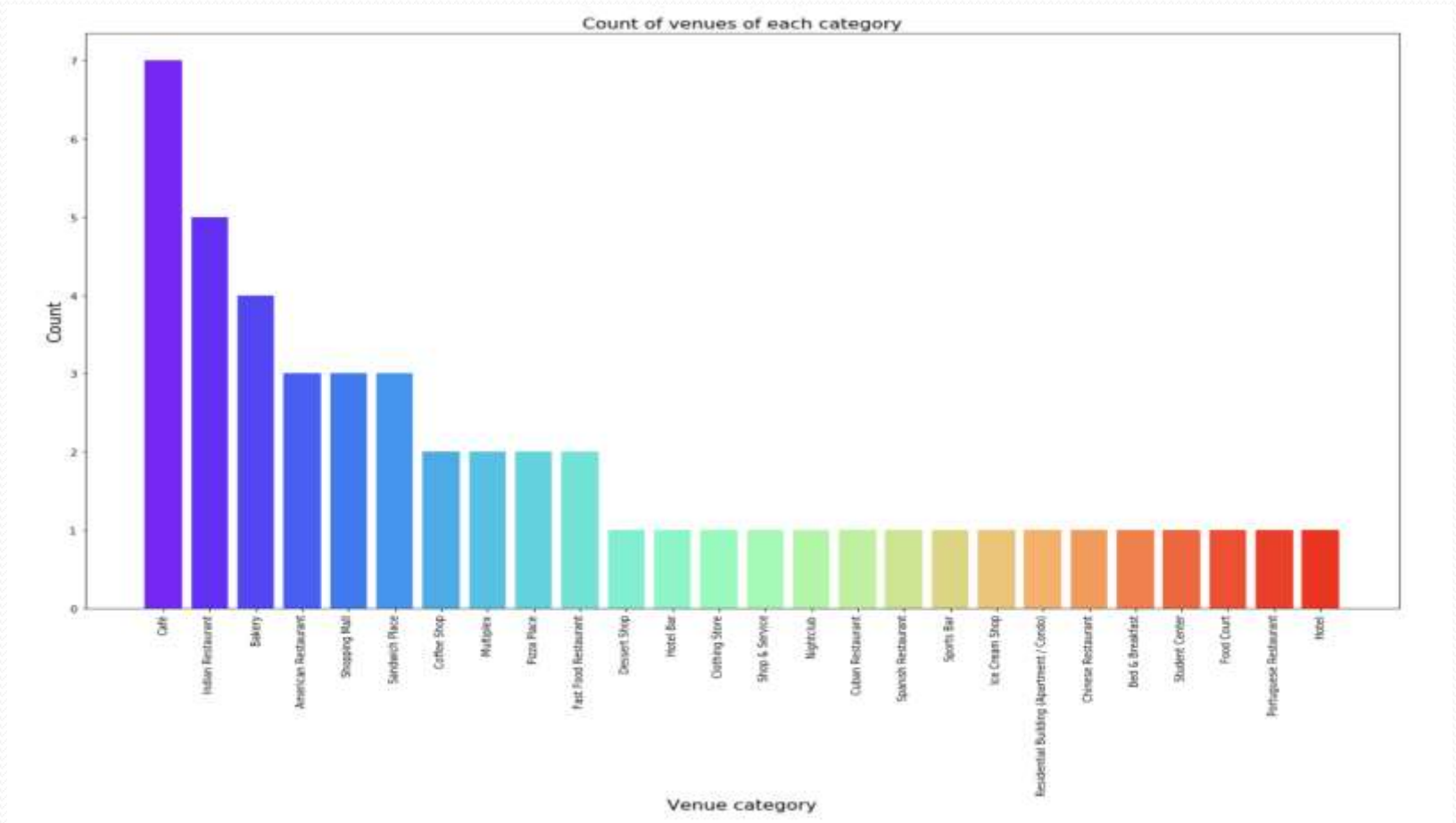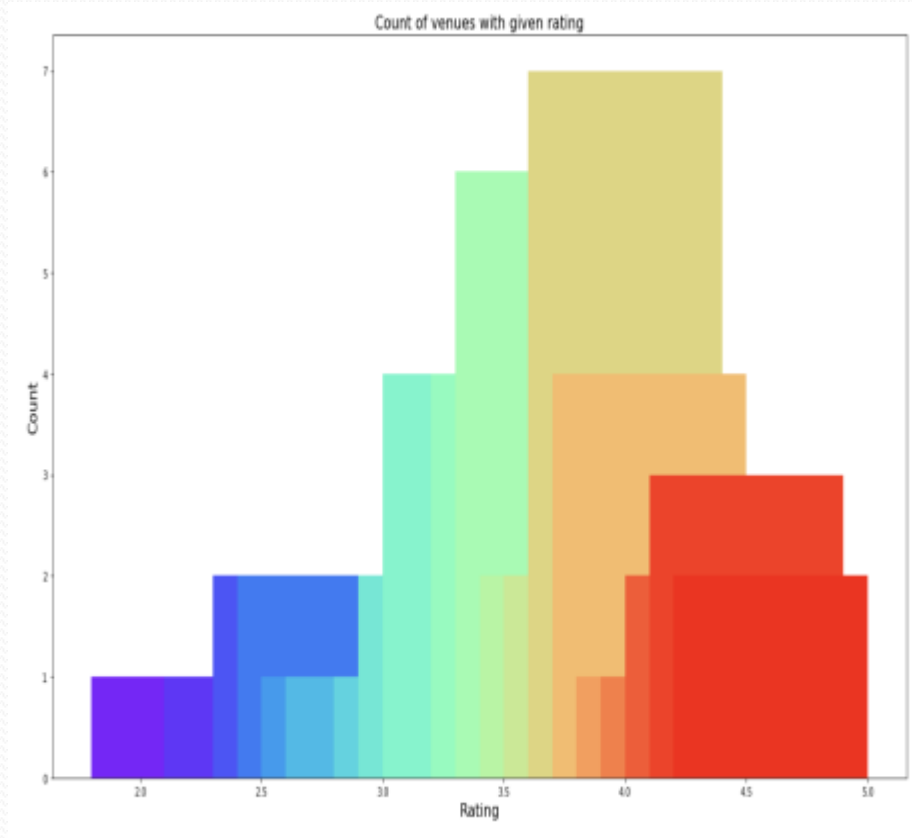
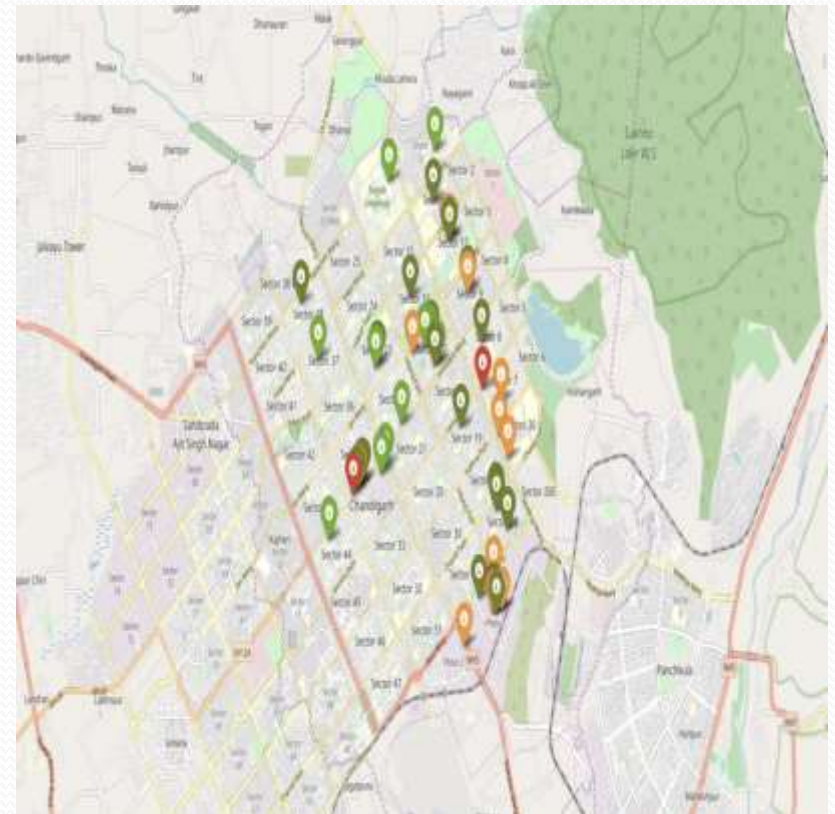# **Rating**

Rating and count of venues with that rating

Plot of venues with different ratings
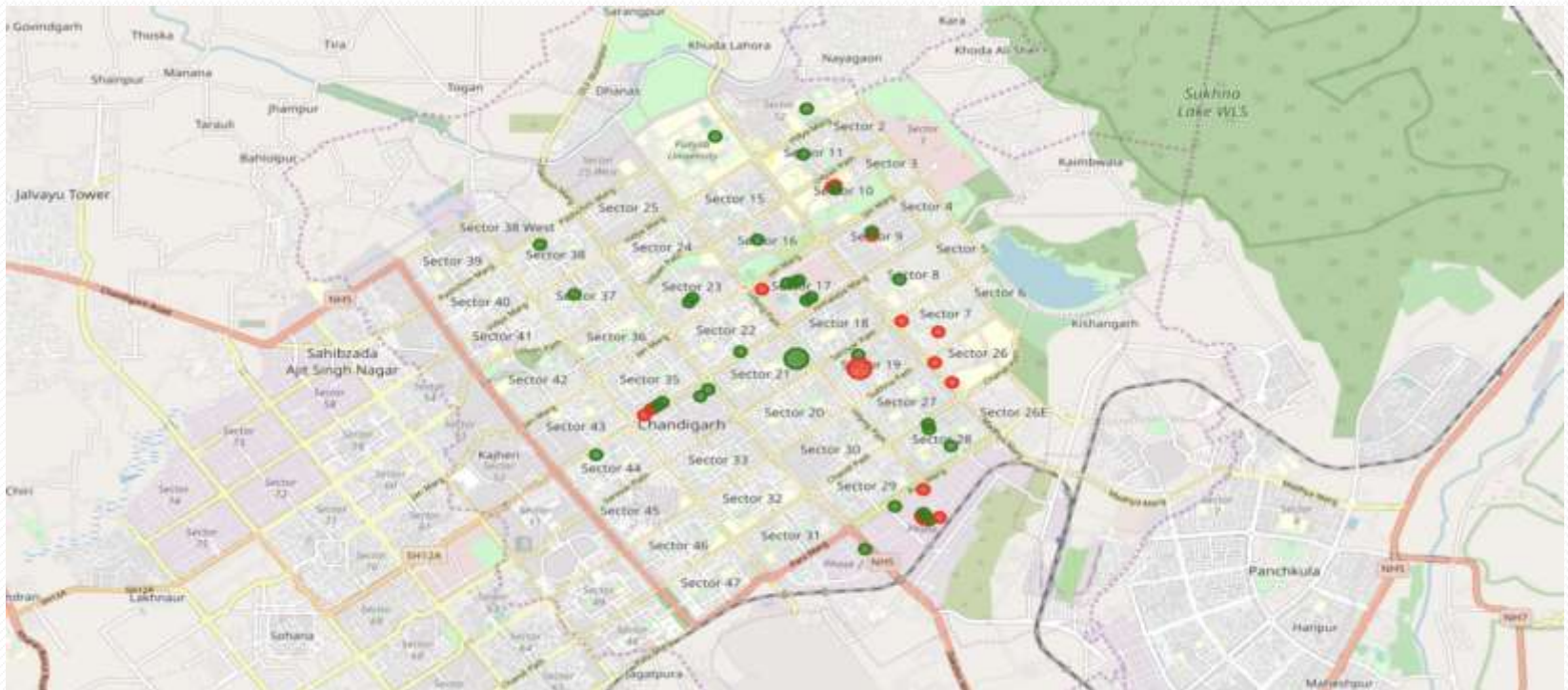
Price

Price per person with count of venues with that price

Plot of venues with different prices

Clustering

- Finally clustering 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.

Clusters of venues

Conclusion

- The purpose of this project was to explore the places that a person visiting Chandigarh could explore. 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: Sector 35, Sector 7 & 26 and Elante Mall. Based on the visitor's venue rating and price preferences, he/she can choose amongst the three places.