CAPSTONE PROJECT

Explore and Compare Hotspots in Bengaluru

ABSTRACT

Analysing 3 main hotspots in Bengaluru in terms of distribution of venues such as restaurants, cafes etc. based on rating and average cost to help understand and compare the trends and in each of these places.

Vismay Radhakrishnan

Electrical Engineer, VJIT, Bengaluru, India

Date: 15th June 2020

CONTENTS

1. INTRODUCTION

1.1 BACKGROUND

The project is aimed at exploring 3 key venues in *Bengaluru, Karnataka, India* based on their rating and prices. In this project, we will be exploring various venues in 3 main hotspots in Bengaluru namely ***MG Road, Indiranagar*** *and* ***Koramangala***. We will be making use of **Foursquare API and Zomato API** to get the coordinates and details of the interested places. These 3 locations have some of the best restaurant, cafes, and pubs in various categories and also belong to the top revenue generating areas in Bengaluru.

The main aim of this project is to help the newbies, foodies, weekend wanderlusts, working professionals explore some of the most happening places in Bengaluru. Be it weekend or weekday, food is something which keeps our foodgasmic mind in place to explore. Also, the clustering details obtained will help the upcoming businesses to analyse the categories of venues in each location and setup business accordingly based on missing categories.

We explore the trends in each of these 3 places by analysing the distribution of each type of venues, the ratings of each place as well as the class of the places based on the average price for two.

These details will be helpful for individuals as well as business to **explore the trends and see which the right place is to invest or go out**. Various opportunities can be explored based on the **lack of any venue in any said hotspot locations**. These locations are frequented by each individual living here or new to the place.

1.2 A BRIEF INTRO TO THE SILICON VALLEY OF INDIA – BENGALURU

Bengaluru (Previously known as Bangalore) was initially the **capital of Mysore State** until Karnataka was formed in **1956**. Then Bangalore was declared as **Karnataka's capital**.

The city of Bangalore has a **population of over 10 million**, thus making it the third most populous city in India and a **megacity**. It is the second fastest-growing major city in India, owing to its diverse demographics.

The social and economic diversity in the city is clearly reflected in Bangalore food. This diversity goes in parallel with the cuisines served here too. You can find all sorts of cuisines here. Thanks to the great variety, it will not be wrong to call Bengaluru a foodie's paradise.

The city has various options for food win all categories of price ranges. We will explore the **Food Culture** in **Bengaluru**.

2. TARGET AUDIENCE

The target audience for this project is twofold. Firstly, any person who is visiting Bengaluru or living here can use the plots and maps from this project to quickly select places that suit their budget and rating preferences. Secondly, a company can utilize 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. Any individual or businesses will be benefited by the comparison data and geospatial plots to explore venues in quick and effortless manner.

3. DATA

3.1 DATA SOURCE

The information pertaining to the venues explored were acquired using 2 API's (Application Programming Interface).

1. **Foursquare API:** (https://developer.foursquare.com/)
   This API was used to collect information regarding the following details:

   a) The coordinates (Latitudes and Longitudes),
   b) Category of Venue
   c) Names of the venues.

   The data collected will be in the radius of 3Km from the 3 hotspot locations. Details of 100 venues will be collected. The final dataset will consist of details regarding 300 venues.

2. **Zomato API:** (https://developers.zomato.com/api)

   The data collected using Foursquare API was fed to Zomato API to get detailed information such as:

   a) The coordinates (Latitudes and Longitudes),
   b) Names of the venues,
   c) Address,
   d) Price for two,
   e) Price range,
   f) Rating.

The datasets collected from both the API's was analysed and compared to remove the outliers and make accurate dataset. These datasets had mismatching coordinate values for a venue which resulted in different venues appearing for same coordinates.

Also, there were data where the new venues replaced the old ones. They had to be filtered to check the accuracy of the data obtained.

Hence, cleaning of the data was an essential activity which had to be taken care of. So, the next activity was Data Cleaning.

Both the datasets were combined to analyse effectively which resulted int a complete dataset of 300 venues.

3.2 DATA CLEANING

The datasets obtained from the Foursquare API and Zomato API was analysed based on the coordinates obtained. The differences were compared, and matching venues obtained by comparison was used as the final dataset. The datasets obtained from both the API's were plotted using Folium to see the differences graphically.

1. **Foursquare API:**

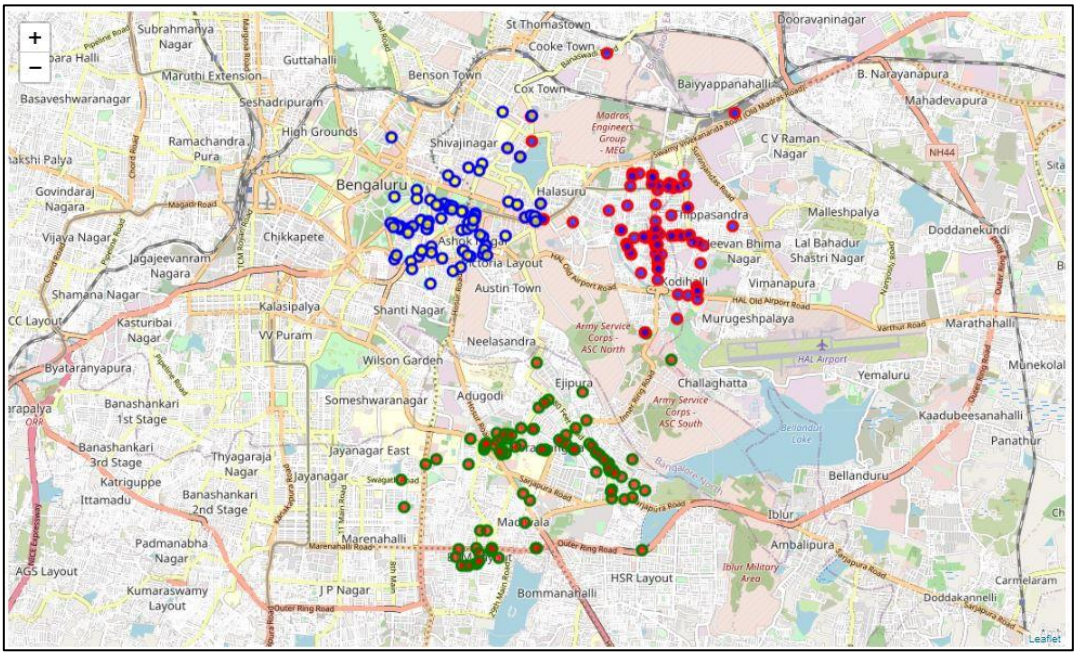

Figure 1: Venues retrieved from Foursquare API

Green Markers: Represents venues in Koramangala, Bengaluru.
Red Markers: Represents venues in Indiranagar, Bengaluru.
Blue Markers: Represents venues in MG Road, Bengaluru.

2. **Zomato API:**

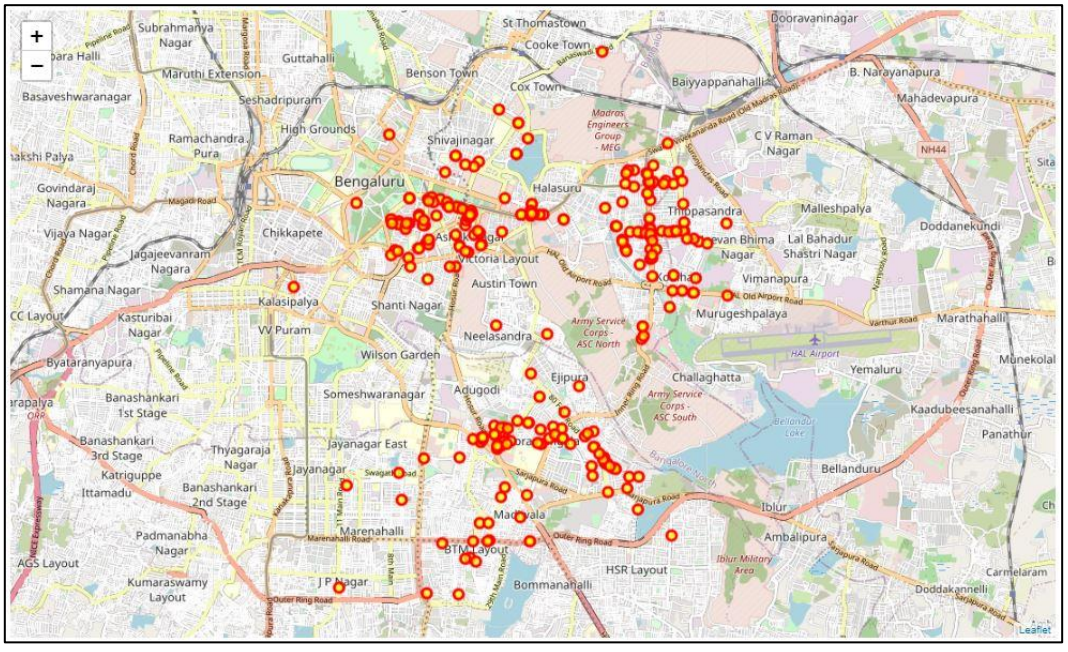

Figure 2: Venues retrieved from Zomato API

From figure 1 and figure 2, we can clearly see that some venues from the two APIs do not align with each other. So, the datasets were combined and filtered using their latitude and longitude values.

To combine the two datasets, the latitude and longitude values were analyzed, and it was made sure the difference between the values obtained from both the API's were not less than **0.0004**. The values of the coordinates were rounded off to 4 decimal places. This helped in filtering the data to a very accurate level.

Also, there were some venues which were still a mismatch. There were removed to get accurate dataset.

After filtering and cleaning, the final dataset had information for **176 venues.**

The final venues in the dataset obtained was plotted to show the distribution.

	Name	Categories	Lat	Long	Neighborhood	Code	Venue	Latitude	Longitude	Price_For_Two	Price_Range	Rating	Address	Latitude_diff	Longitude_diff
0	Chianti	Italian Restaurant	12.9335	77.6219	Koramangala	0	Chianti	12.9335	77.6219	2000.0	4.0	4.4	12, 5th A Block, Koramangala 5th Block, Bangalore	0.0000	0.0000
1	Sree Krishna Kafe	Breakfast Spot	12.9371	77.6197	Koramangala	0	Sree Krishna Kafe	12.9371	77.6196	200.0	1.0	4.3	5th Main, Near Muni Reddy Kalyana Mantap, Kora...	0.0000	-0.0001

Figure 3: Final dataset

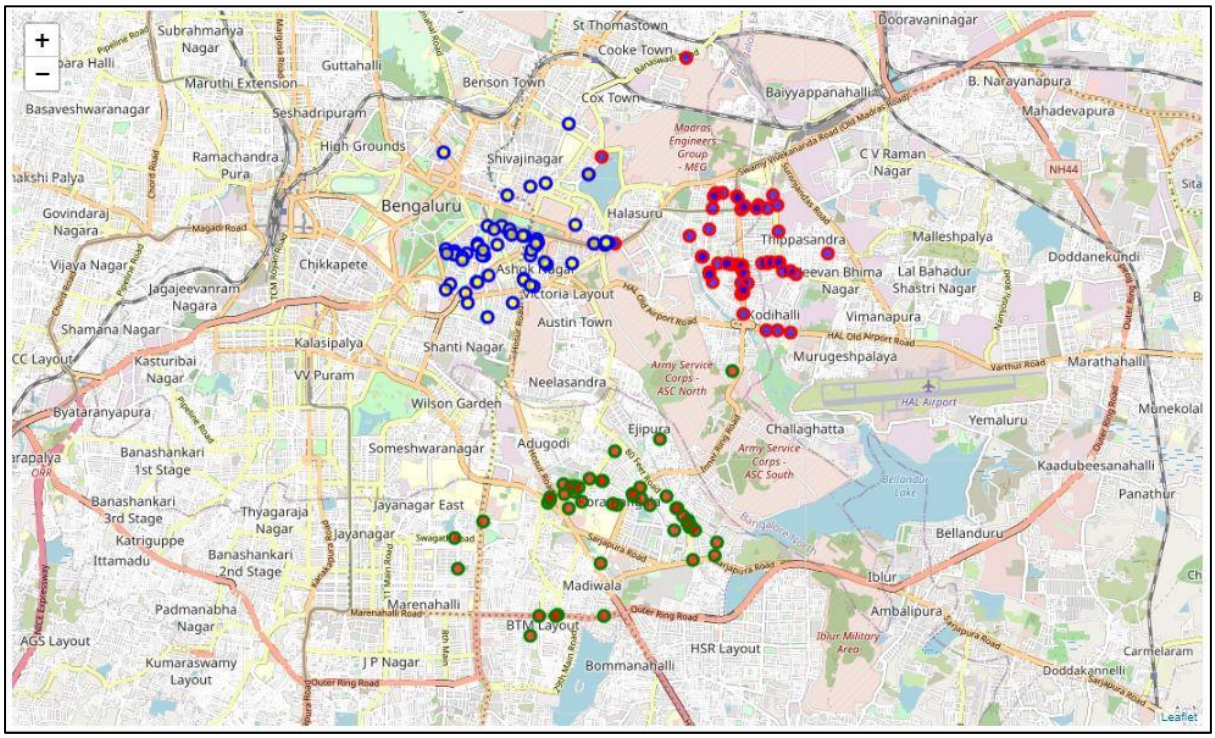

Figure 4: Final list of venues plotted in map

4. METHODOLOGY AND EXPLORATORY DATA ANALYSIS

Using the final data obtained, various relations between the data were analyzed. Following relations were explored:

1. Category count in each Location,
2. Count of venue categories in each location,
3. Average rating of venues,
4. The price range of the venues.
5. K-Means Clustering.

4.1 CATEGORY COUNT IN EACH LOCATION

The total count of each venue category was summed up from all the three locations and the same were plotted as bar chart.

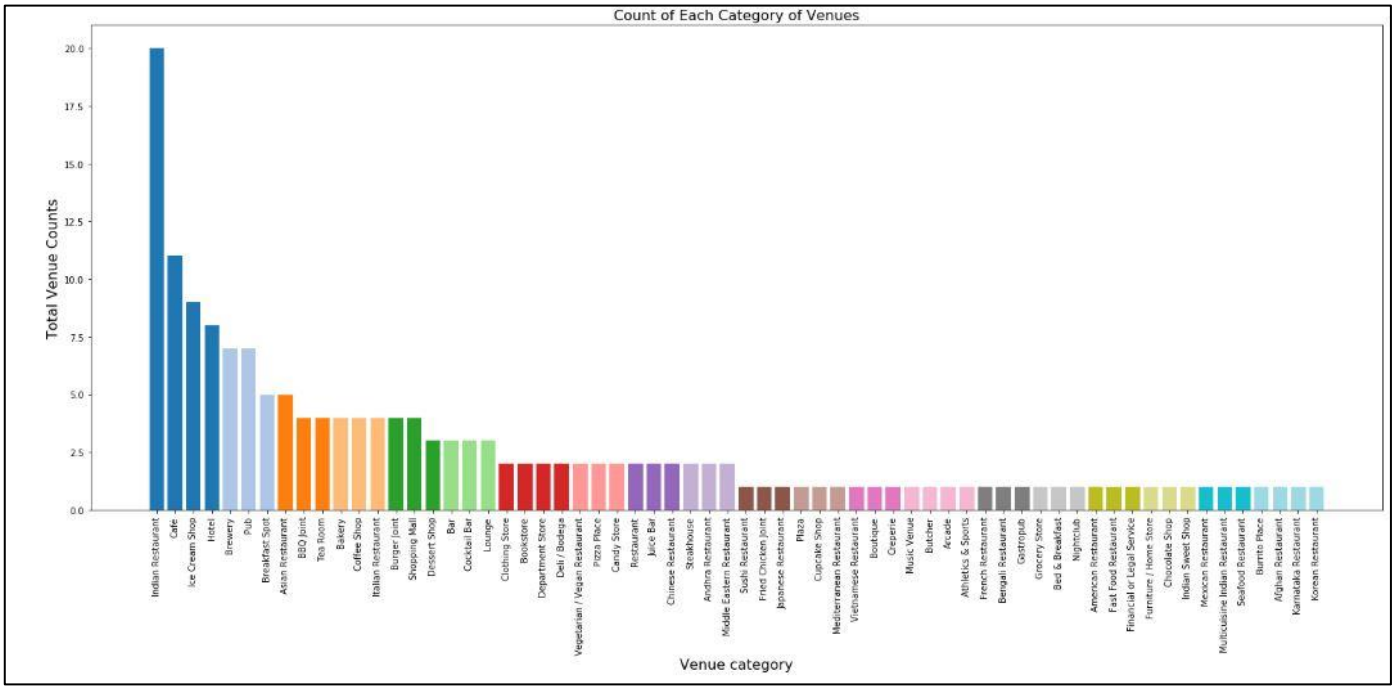

Figure 5: Count of each venue category across 3 hotspots combined

The above plot shows that, the count of ***Indian restaurants, Cafes, Ice cream shops, Hotel, Brewery and Pubs*** are the most from all the 3 locations.

Indian restaurants topping the list was obvious but looking at the other venues clearly shows their widespread influence of the choice of people. All the 3 places namely **MG Road**, **Koramangala** and **Indiranagar** are known for lively weekends and after office hangouts. They have some of the best places to go to when you are hungry or want to explore any new experience in terms of cuisines.

4.2 COUNT OF VENUE CATEGORIES IN EACH LOCATION

The total count of venues obtained above was split according the 3 locations: **MG Road**, **Koramangala** and **Indiranagar** as mentioned in the label '*Neighborhood*'.

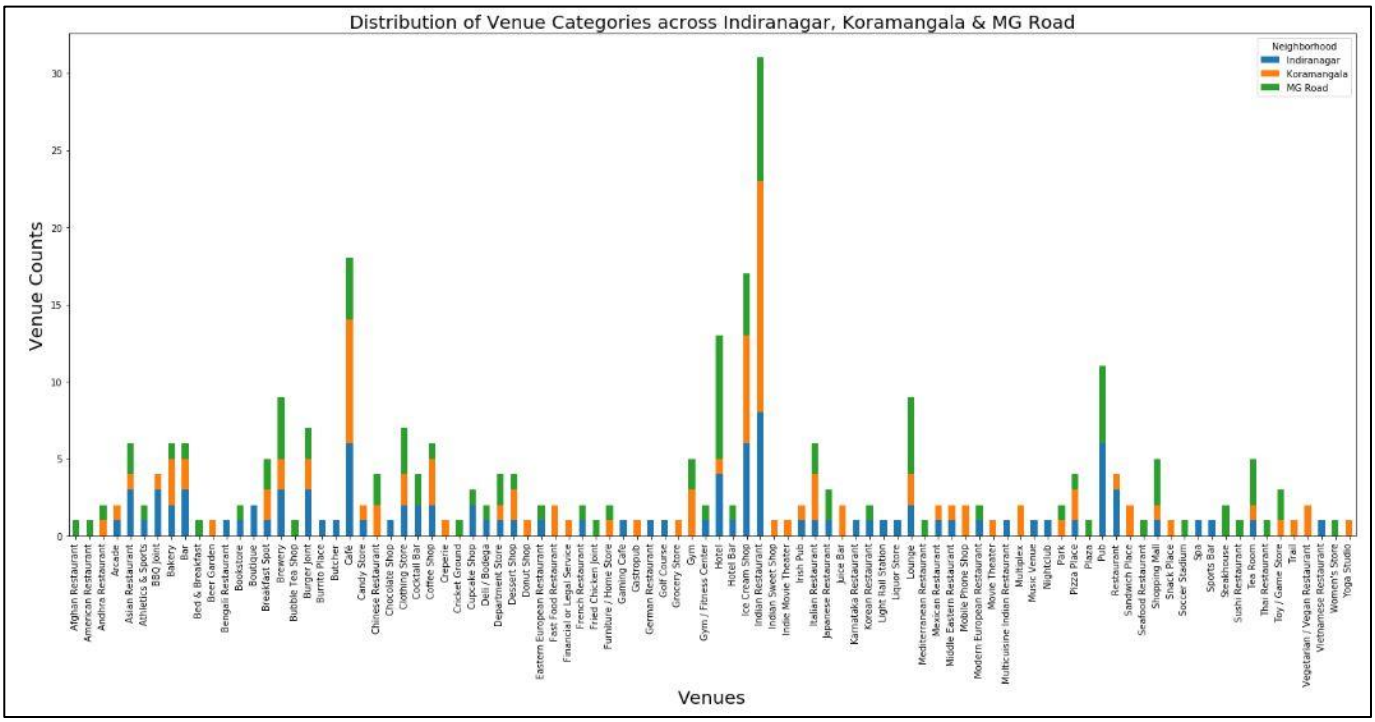

Figure 6: Venue counts location wise

4.3 PRICE RANGE OF VENUES

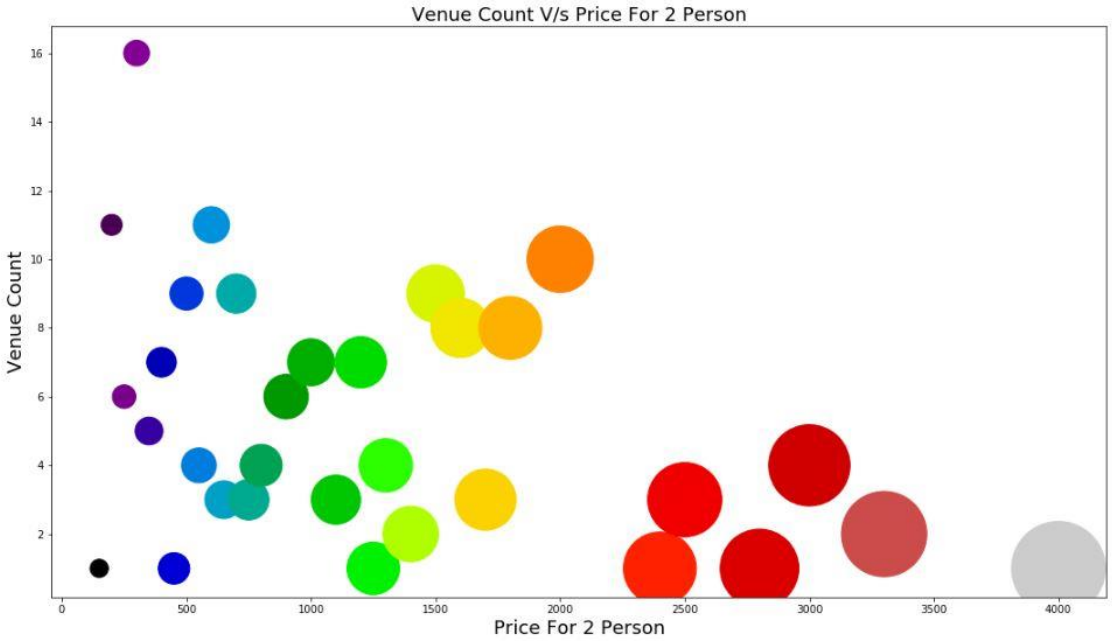

Figure 7: Price range of venues

From the above scatter plot, we can see the distribution of the venues based on the price range. The distribution of the venues in each location was plotted in the map indicating the price for 2 people.

Red Label: High Price Range

Orange Label: Above Average Price Range

Blue Label: Medium Price Range

Green Label: Low Price Range

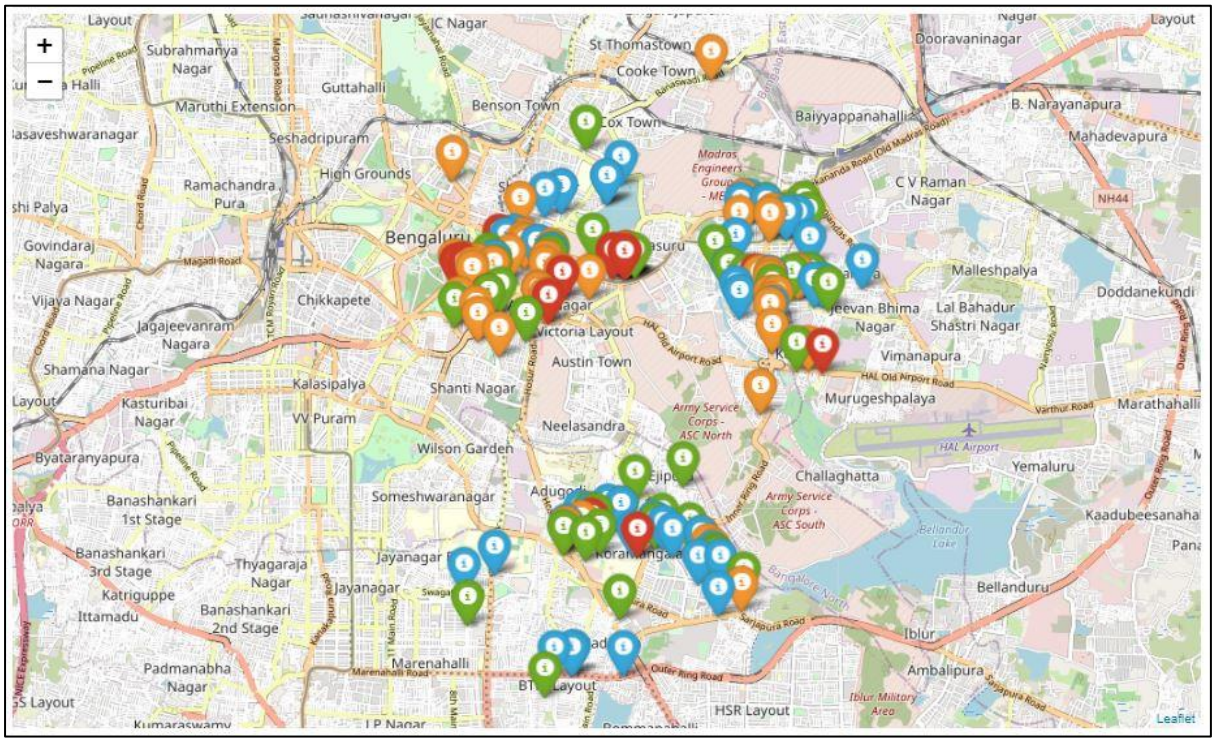

Figure 8: Venues segregated by Price Range

From the above map, we can observe some trends with distribution of venues based on price.

Most of the Red labelled venues are Fine Dine Restaurant's and most of them are clustered around **MG Road**. So, you know here to head when you are up for a fine dine. **MG Road** also has the most Orange labelled venues.

Koramangala has most affordable places shown by Green and Blue label.

As for **Indiranagar**, we can see a **mix of the High and Low-price range** venues. It has a *perfect mix of both the worlds*.

4.4 RATING OF VENUES ACROSS 3 LOCATIONS

The venues were labelled with the ratings and the same were plotted in the Map to understand the average distribution of venues based on Rating.

Green Label: Best Rating

Blue Label: Good Rating

Orange Label: Satisfactory Rating

Red Label: Low Rating

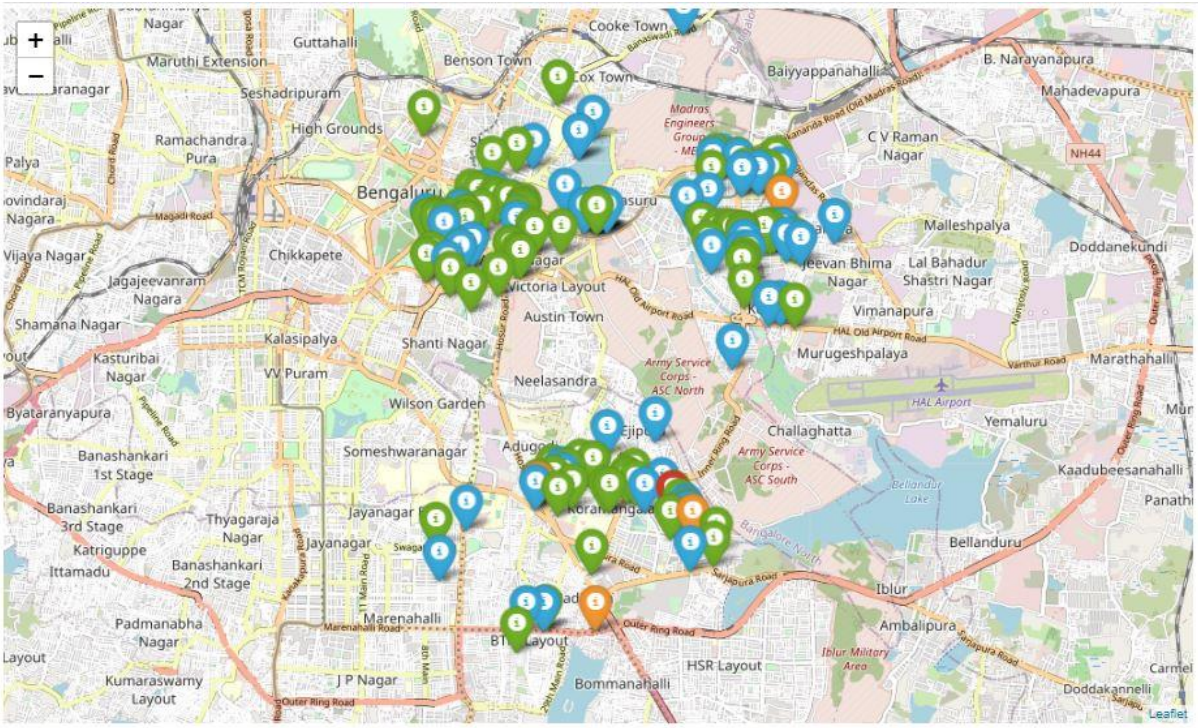

Figure 9: Map of Venues based on Rating

We see that most of the venues are having good rating across all the 3 locations. Only a selected few venue has satisfactory and low rating.

4.5 K- MEANS CLUSTERING

All the venues were clustered based on their price range, location and rating to identify similar venues and the relationship amongst them. K-Means clustering was used to decide the cluster. The venues were distributed into 4 clusters. Each cluster is indicated by a different color in the map shown below.

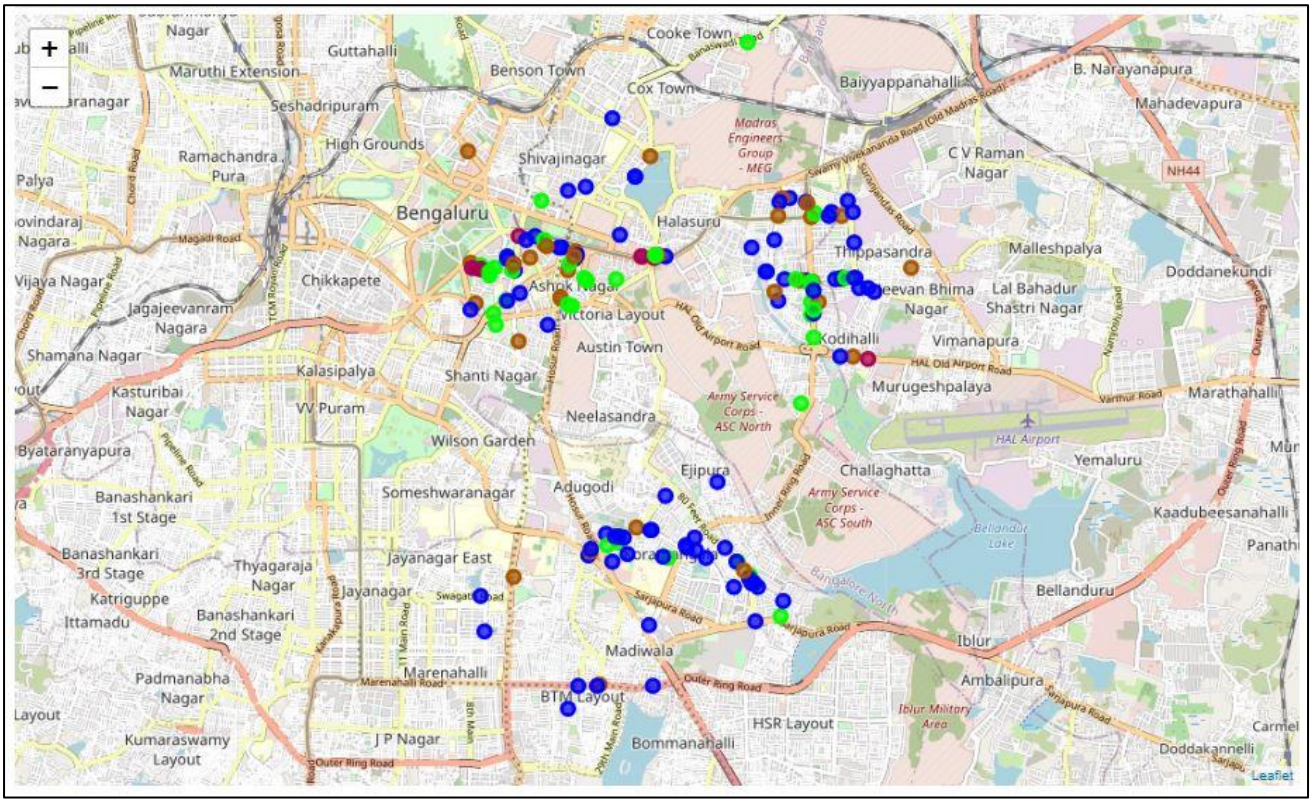

Figure 10: K-Means Clustering

5. RESULTS AND DISCUSSIONS:

From the above analysis, we get a clear understanding on the distribution of venue categories across the 3 main hotspot locations in Bengaluru namely MG Road, Koramangala and Indiranagar based on rating and average price for two people.

MG Road being one of the *prime and oldest hotspots*, it has some of the most ***premium Hotels and Restaurants*** which adds to its supremacy as one of the most happening places in Bengaluru. It has quite a few ***Fine Dine restaurants*** which can be explored for a unique experience. **Brigade Road** adds additional oomph to the popularity of MG Road which is frequented by shoppers.

Koramangala and **Indiranagar** were once peaceful residential areas which became quite happening on the onset of commercialization. But in competition they are keeping up with MG Road.

Indiranagar has proximity with **MG Road** which helps it to get more attention from people preferring to visit MG Road. Indiranagar has a **mix of high and affordable venues** making it quite popular as compared to Koramangala.

As for **Koramangala**, it has some of the **best affordable places** which are pocket friendly compared to the venues in MG road and Indiranagar. It is far away from other 2 hotspots, which gives it an added advantage of being the go-to place for people in areas such as **Electronic City, HSR, BTM and Madiwala**.

We can see that MG road is already well established as prime location to go out to. But Koramangala and Indiranagar has lot of opportunities for upcoming businesses as the competition is tough. There are many categories of venues which needs to be explore based on the options available in other locations.

This study amplifies the existing trends followed by the businesses already established giving a fair idea on the competitiveness across the 3 locations in Bengaluru.

As for the foodies, wanderlusts and explorers this study gives a basic understanding on where they should be headed to for a certain experience.

5.1 FURTHER SCOPE FOR STUDY

While doing the analysis, I noticed that some of the most common places were missing in the dataset obtained. Also, while cleaning the data, a few of the best venues had to be discarded because of mismatch of data from the coordinates. This can be worked upon with more dataset to get a wider scope with more accuracy.

One more detail that could be included in future studies is Cuisine. It can show the popularity of each cuisine in each location. This will give a fair idea on the preferences of the people based on the clustering. Also, this can give businesses a fair idea on what cuisine is missing at a location depending on which they can lookup for setting up those businesses for which they will have edge over others.

6. CONCLUSION

The purpose of this project was to explore 3 main places in Bengaluru, which has a lot of potential for businesses as a potential market and for individuals to explore new culinary experience. The venues were identified using Foursquare API and Zomato API and the same were plotted in maps to understand the trends followed. Also, we were able to cluster all the venues into 4 clusters using M-Means based on location, rating and price range. This opens a lot of opportunities for businesses to look for options for investment. Additionally, for anyone exploring Bengaluru these are the 3 best places to start exploring based on the cuisines and categories of venues available.