PREDICTING THE BEST VENUES TO VISIT IN BENGALURU TO EXPERIENCE THE NIGHTLIFE

Finding the **Best Venue for outings** is always a complex issue…….

Bengaluru city is well known for its nightlife. Thousands of Bengalurians look for escape venues on Friday nights and weekends. However, finding the best venues for hangouts is a conundrum, given the plethora of choices available.

This Data Analytics project intends to help make better choices by clustering similar venues based on three parameters: Rating, Price, and the number of likes.

Data Acquisition:

Foursquare DataBase was used to collect necessary data for the analysis. Two end points were used to get the data:

1. **Venue Categories:**

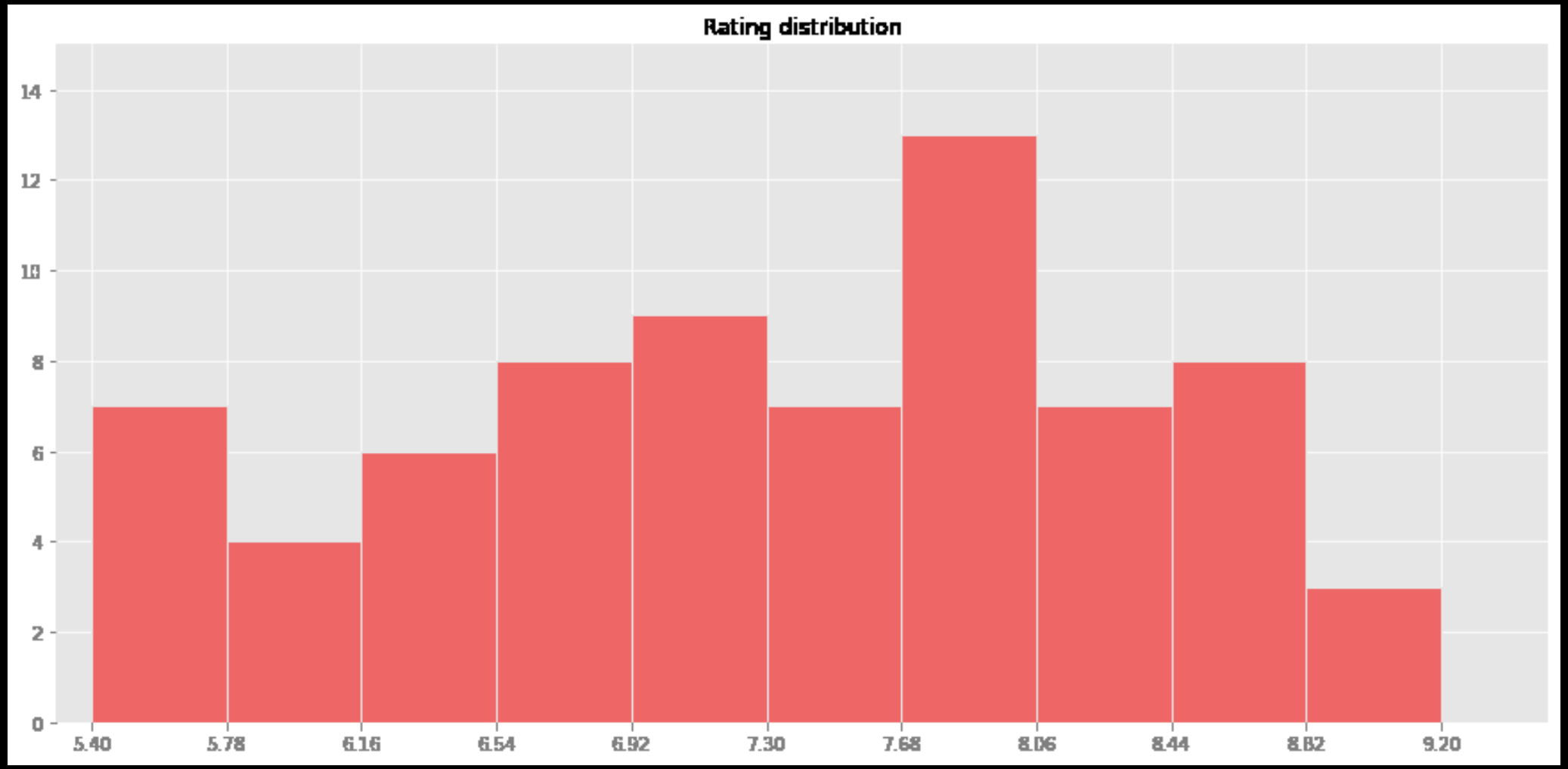
Description: Returns a hierarchical list of categories applied to venues.

2. **Venue Details:**

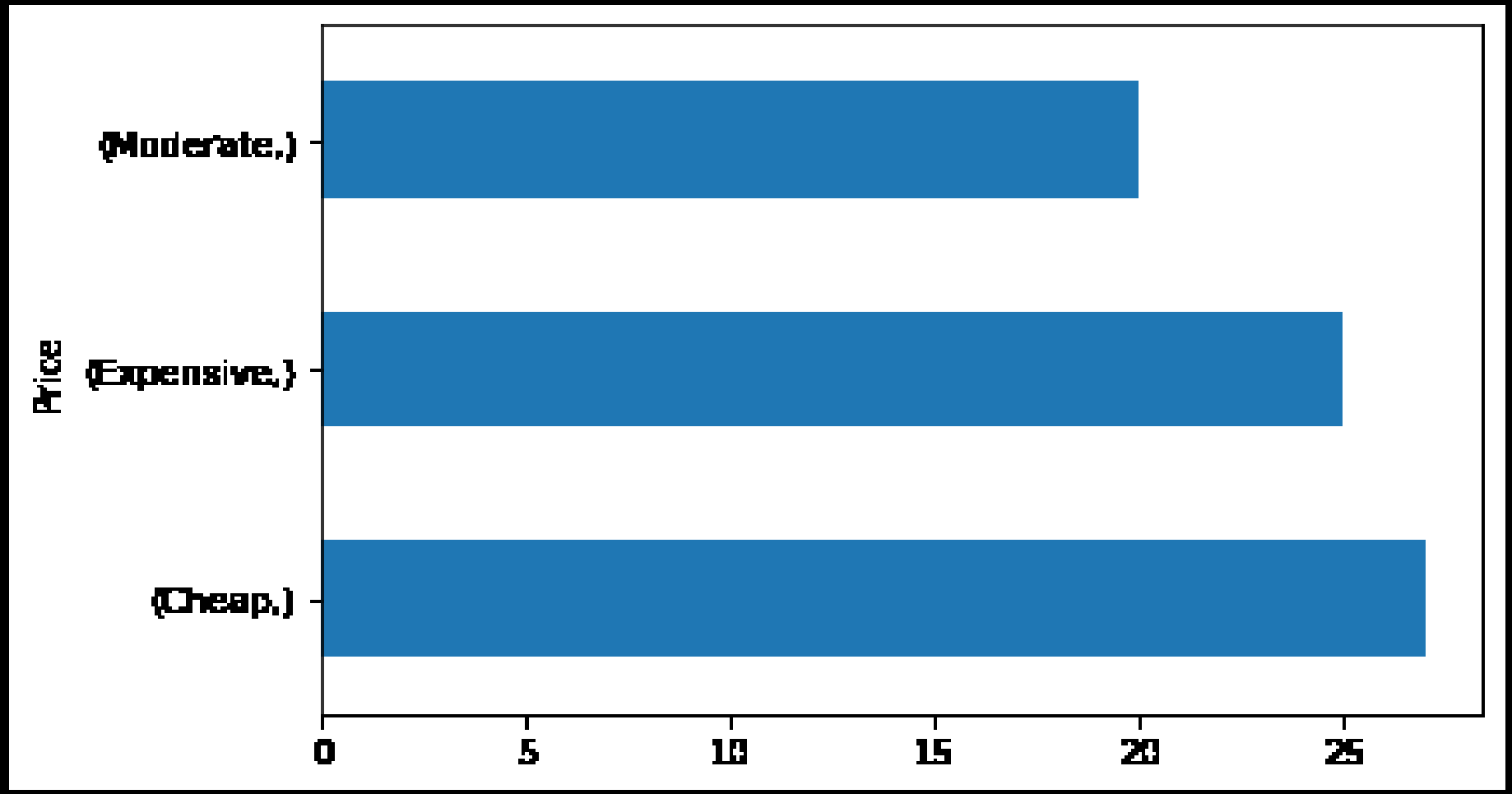
Description: Gives the full details about a venue including location, tips, and categories.

Exploratory Data Analysis:

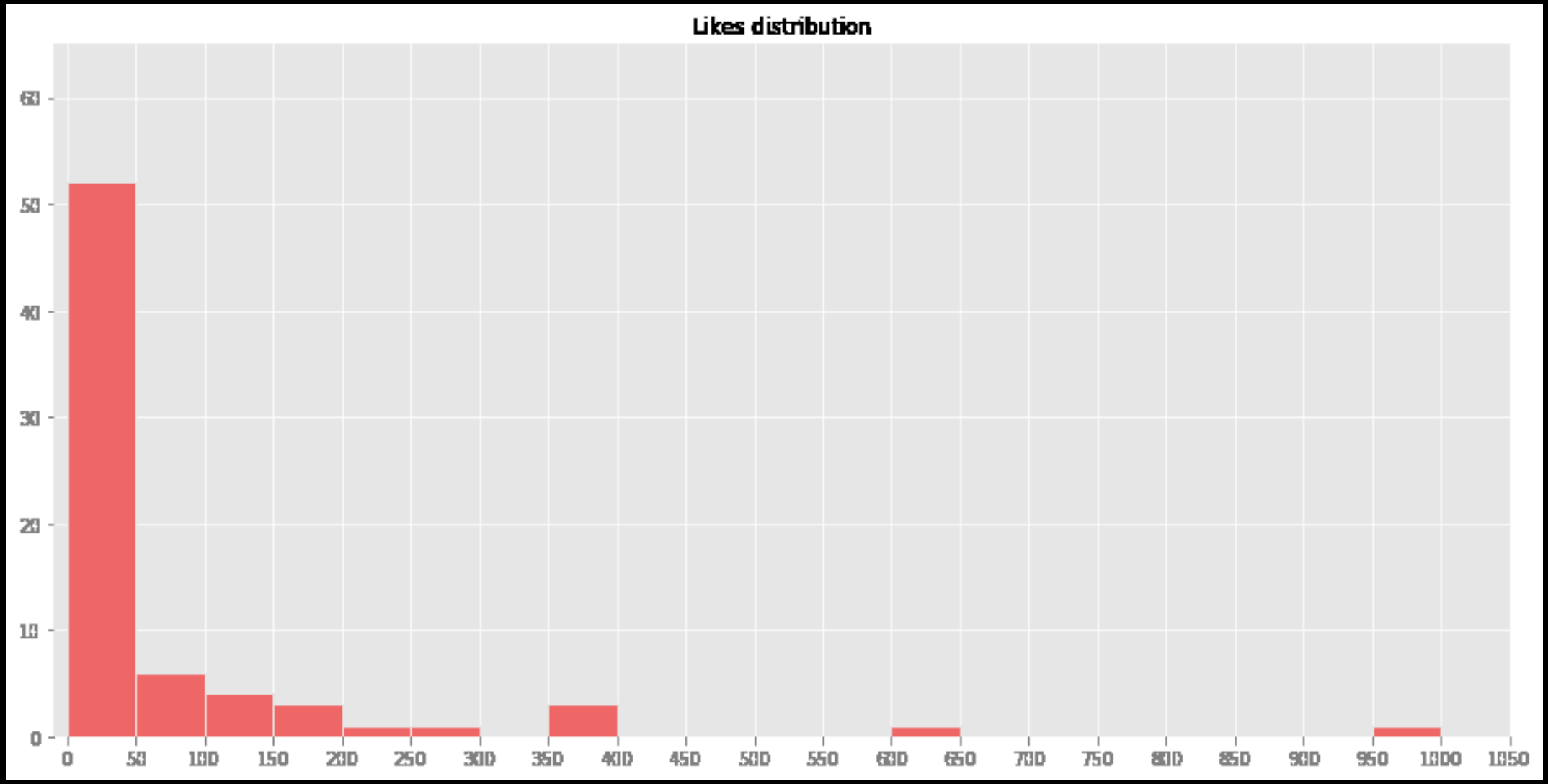
1. Rating Distribution:



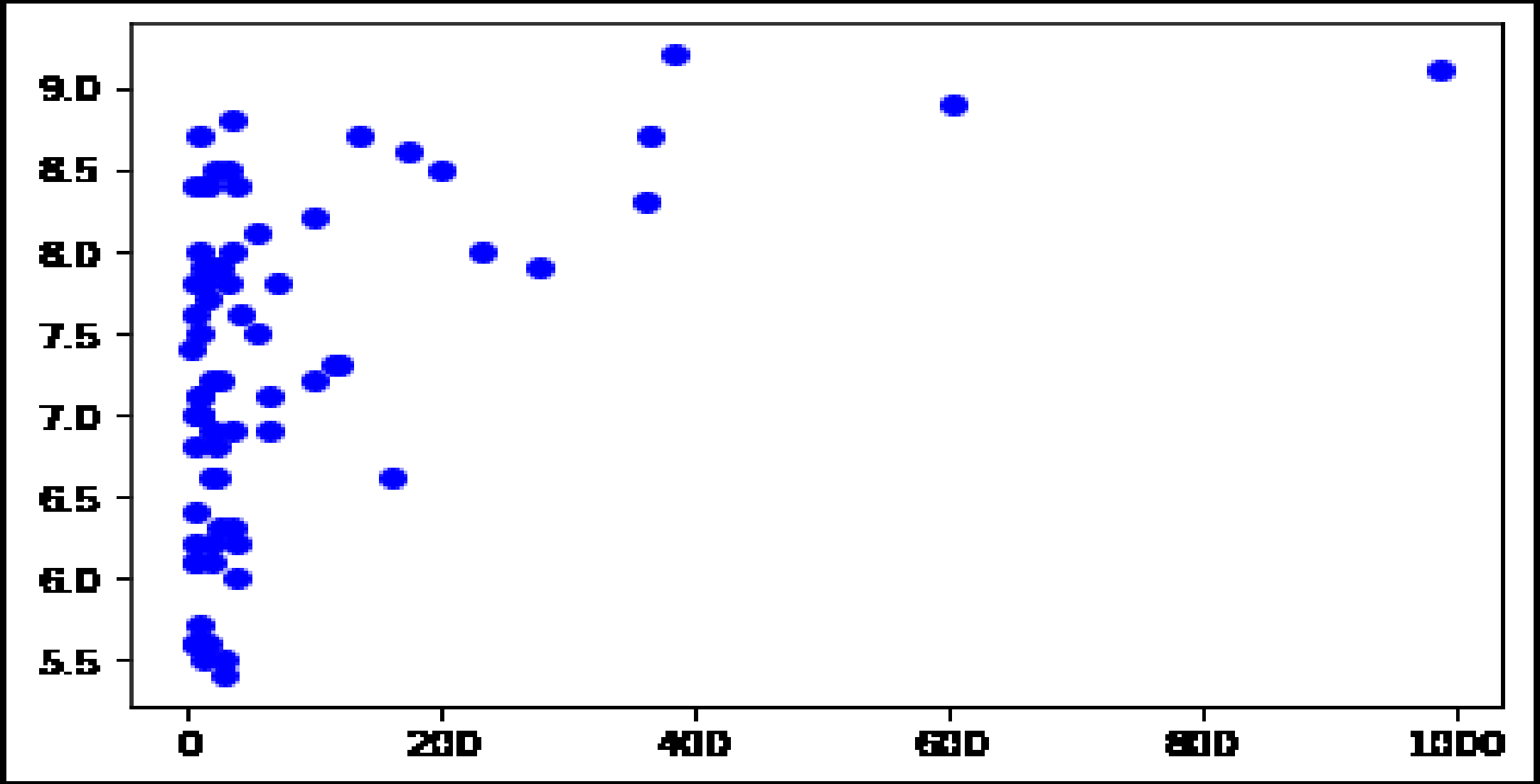
2. Price Distribution :



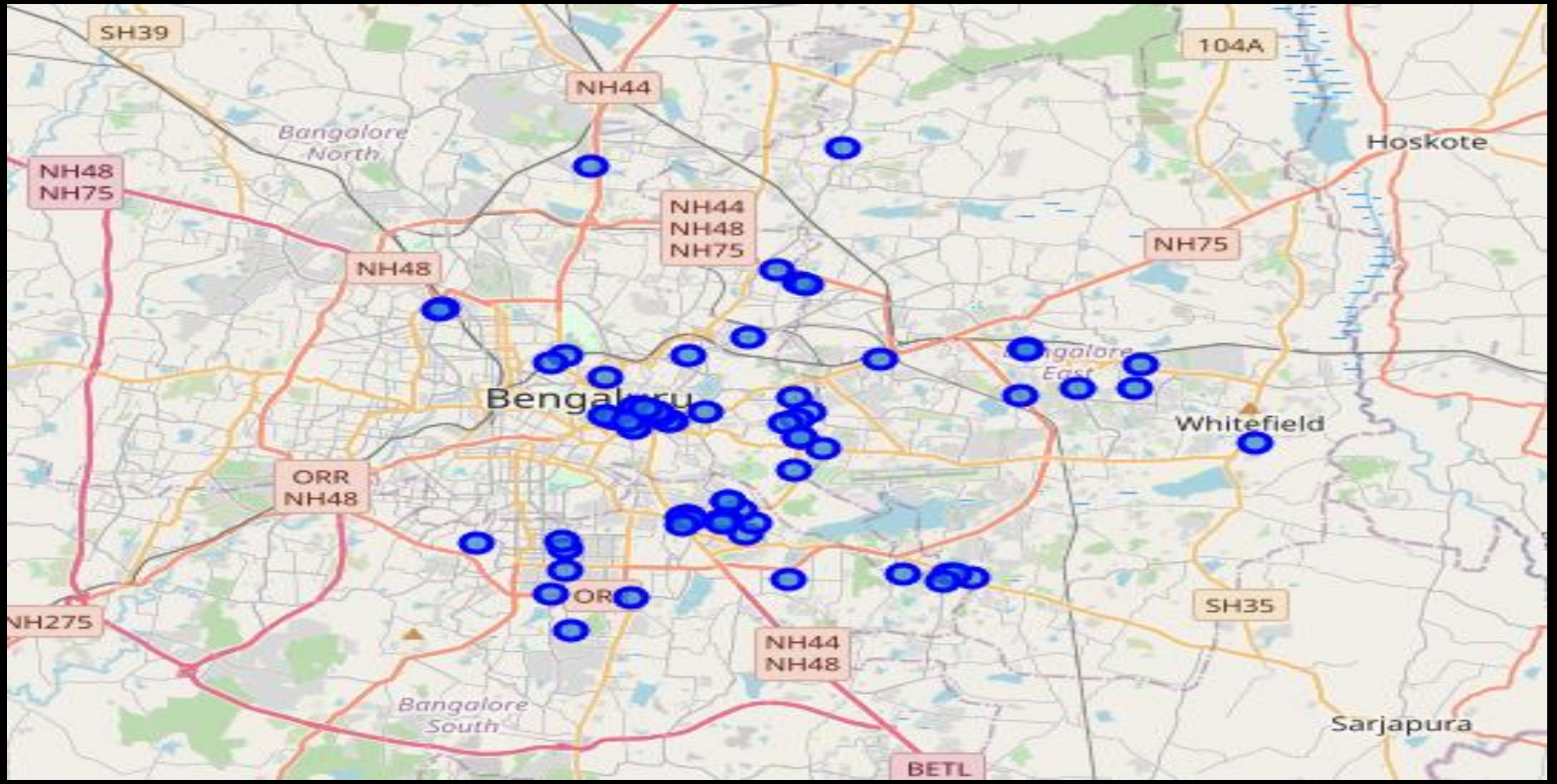
3. Like counts Distribution :



4. Likes counts Vs. Rating



Venues on Map :



Clustering Model:

The Kmeans clustering algorithm from the sci-kit learn library was used to cluster the venues into different groups.

The number of clusters based on the following assumption:

1. Under the pricing column, the venues fall under Cheap, Moderate, and Expensive.
2. Under the Rating column, the venues fall under: Bad, Decent, and Good.
3. From high-level permutation and combinations, we can have 3*3 = 9 possibilities.

Results:

On analyzing the type of venues in each cluster based on the three metrics, I labeled each cluster as follows:

1. **Cluster 0: 'Moderate, good but not well known'**
2. **Cluster 1: 'Cheap and Decent'**
3. **Cluster 2: 'Expensive and Decent'**
4. **Cluster 3:'Expensive, Excellent, and Very well known'**
5. **Cluster 4: 'Moderate, Good, and well known'**
6. **Cluster 5: 'Moderate and Bad'**
7. **Cluster 6: 'Expensive, good, and well known'**
8. **Cluster 7: 'Cheap and bad'**
9. **Cluster 8:'Expensive, bad, and not well known'**

Conclusion :

From the above analysis, I was able to sort venues based on these three metrics: Price, Rating, and likes. This analysis could be used to select the most appropriate venue for an outing based on preferences. It also forms a basis for the recommendation system.