# Battle of Neighborhoods

# Comparing Neighborhoods of Bengaluru and Delhi

## Introduction

Bengaluru and Delhi are two of the leading IT hubs of India and is among the fastest growing metropolises of the country. Delhi is the fifth most populous city in the world and the largest city in India area-wise. Delhi has an estimated 2016 population of 18.6 million. Bangalore has an estimated population of 12.34 million in its urban area in 2017, up from 8.5 million in 2011 [1]. Both the cities welcome a wide range of residents, from students to families planning to settle down. There are over 65 neighborhoods in the Bengaluru metropolitan area and over 175 in Delhi. This project aims to compare these neighborhoods by grouping them using K-Means clustering algorithm based on the services available within a certain distance around the neighborhood. This will especially be a guide to people moving from Bengaluru to Delhi or vice-versa as they can choose a neighborhood with similar services they are used to.

## Data Acquisition

List of neighborhoods in Delhi and Bengaluru are available open-source in the Internet. BeautifulSoup library package in Python has been used to extract data from the tables of region wise neighborhoods in Bengaluru. The list of neighborhoods of Delhi was available in CSV format. MapQuest API has been used to obtain the coordinates of neighborhoods. This step was done externally, and the CSV file was then imported into the notebook. The explore function of Foursquare API has been used to get the list of services available around the neighborhood along with their coordinates.

## Methodology

After obtaining the required data, it was sorted into two tables containing neighborhoods of Delhi and Bengaluru with the region in which the neighborhoods were located along with their coordinates respectively.



Table 1 - Bengaluru Data



Table 2 Delhi Data

Folium library of Python was used to create a map of both the cities. Markers were added with pop-up labels for each neighborhood to visualize the spread of neighborhoods over the cities.

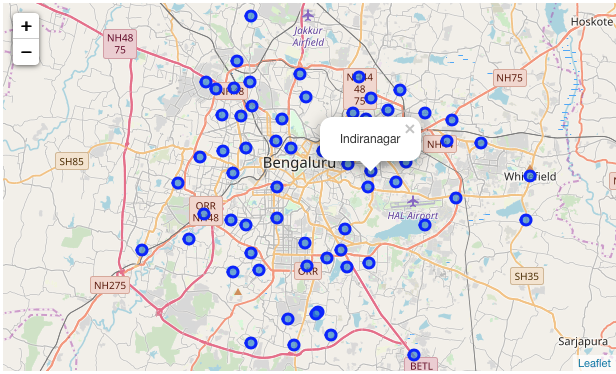


Figure 1 Map of Bengaluru

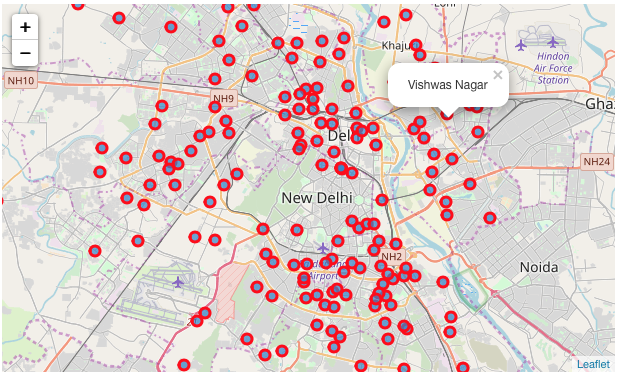


Figure 2 Map of Delhi

Foursquare API was used to explore each neighborhood. A list of all Venues within 1km around the neighborhoods was obtained for Bengaluru and Delhi. It was noted that Delhi neighborhoods has in total **217 unique categories** of services available whereas Bengaluru has **179**. **Indiranagar** hasthe greatest number of unique categories in Bengaluru (51 Categories) and **Lodi Colony** has the greatest number of unique categories in Delhi (54 Categories) **.**

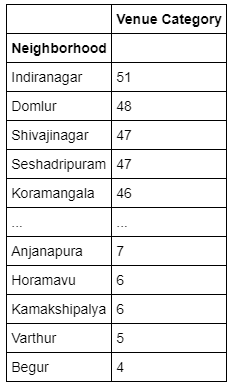


Table 3 Count of Unique Venue Categories in each Neighborhood – Bengaluru

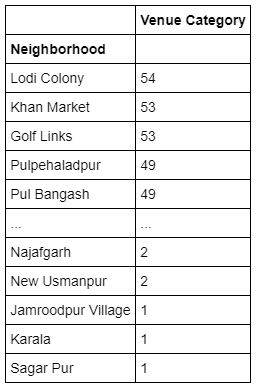


Table 4 Count of Unique Venue Categories in each Neighborhood – Delhi

A data frame was created to better understand the different categories of services and the number of them located around each neighborhood



Table 5 Count of Venue Category per Neighborhood - Bengaluru

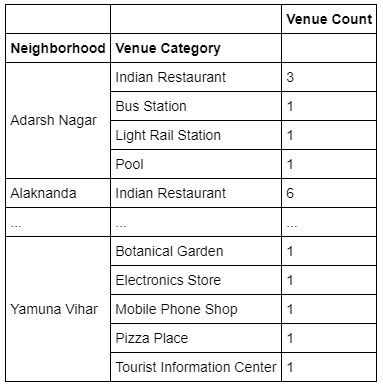
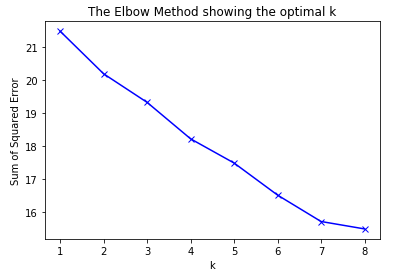


Table 6 Count of Venue Category per Neighborhood - Delhi

A new data frame was constructed by merging data of Bengaluru and Delhi. Before clustering the data using K-Means algorithm, one hot encoding was performed on the data and then grouped by neighborhood. In order to find the optimum ‘k’ in K-Means clustering, **elbow method** was initially tried.



As the graph didn’t show a clear elbow, the **Silhouette Coefficient** for each k was found and it was determined that the optimum ‘k’ was **7**

K-Means clustering was performed, and the neighborhoods were grouped into 7.



Table 7 Top 10 Venue Categories with Cluster Labels

## Results

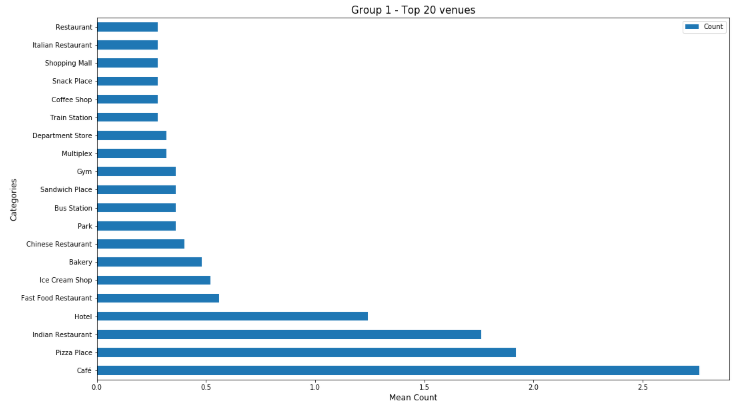
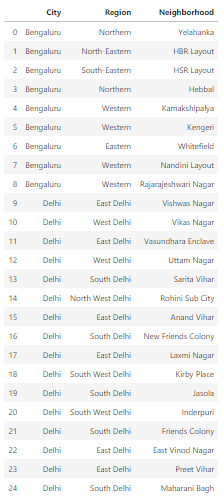
As a result of K-Means clustering, 7 Groups of neighborhoods in Delhi and Bengaluru was obtained. Data frames of each groups with the City, Region and Neighborhood was constructed. It was noted that 3 Groups had only Delhi Neighborhoods. Majority of Bengaluru Neighborhoods were clustered together in Group 2 and Group 7. Group 4 and Group 6 has only one Delhi Neighborhood each. This may account to lack of nearby venue information. Graphs for each group showing the mean count of top 20 categories were created.

**Group 1**

Number of Neighborhoods –

Delhi - 16

Bengaluru - 9

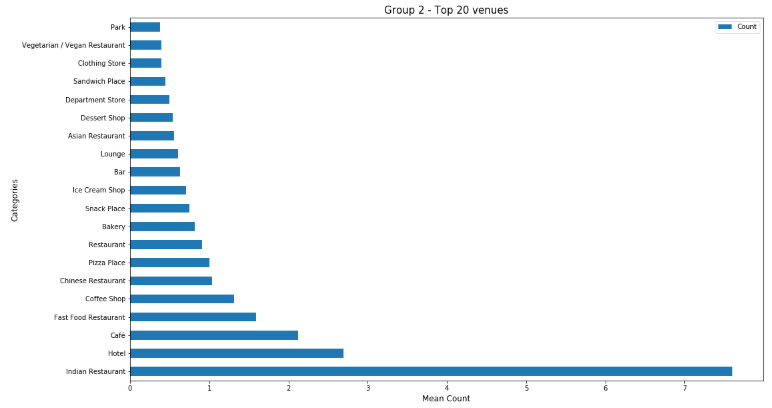
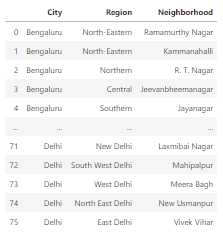


**Group 2**

Number of Neighborhoods –

Delhi - 43

Bengaluru - 33

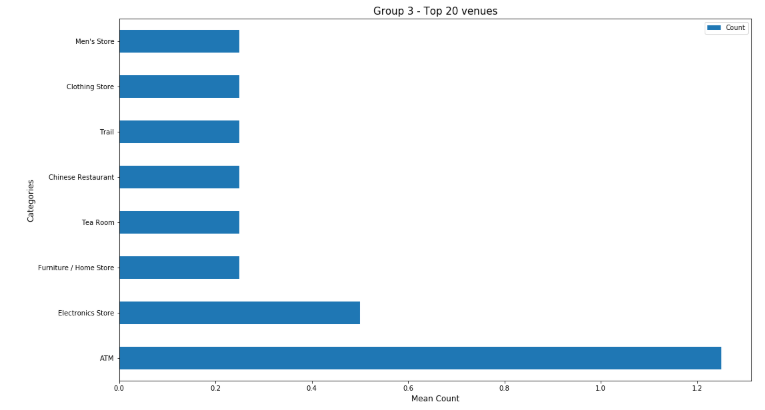
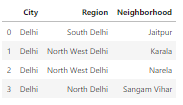


**Group 3**

Number of Neighborhoods –

Delhi - 4

Bengaluru - 0

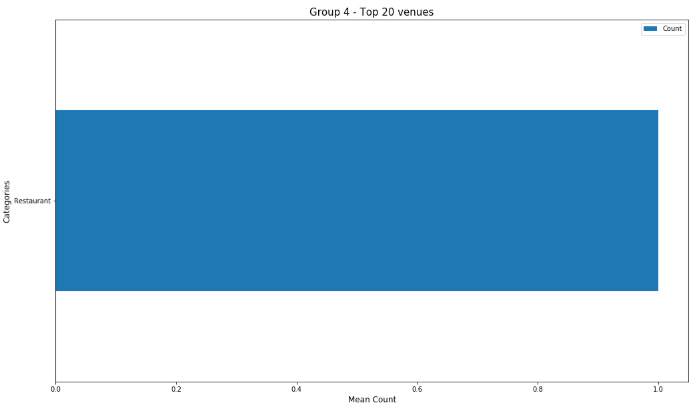


**Group 4**

Number of Neighborhoods –

Delhi - 4

Bengaluru – 0

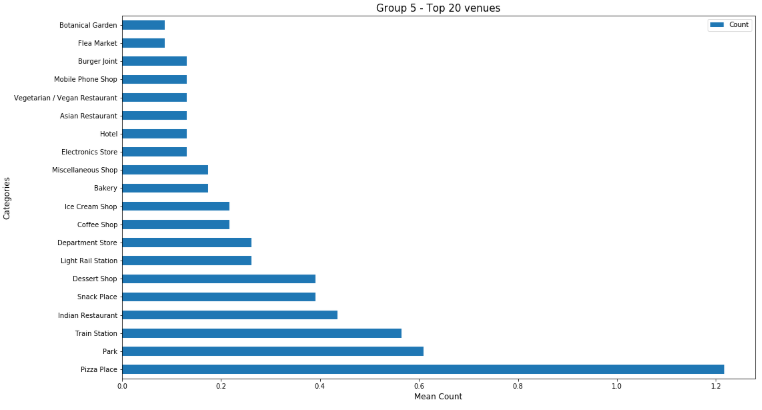
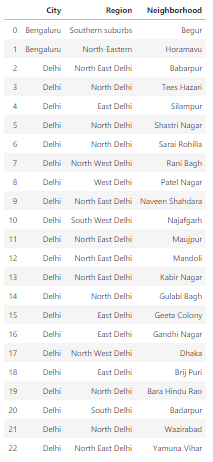


**Group 5**

Number of Neighborhoods –

Delhi - 21

Bengaluru - 2

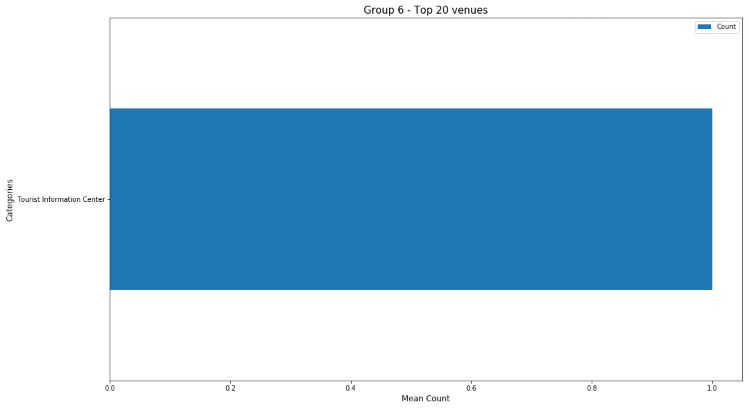


**Group 6**

Number of Neighborhoods –

Delhi - 1

Bengaluru - 0

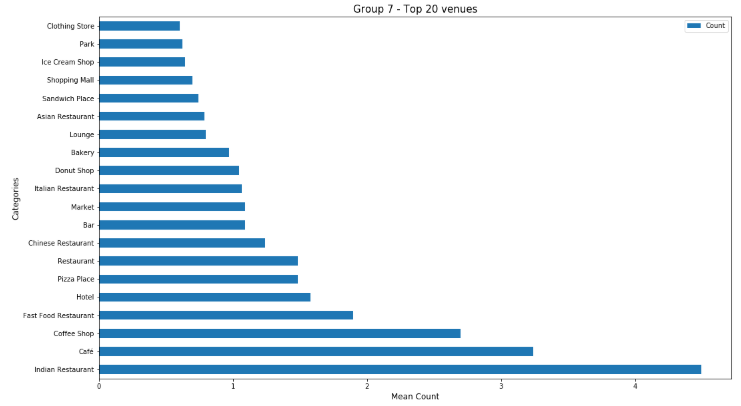
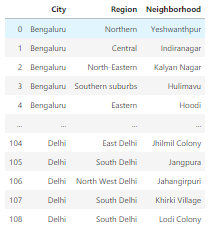


**Group 7**

Number of Neighborhoods –

Delhi - 88

Bengaluru - 21



## Discussion

Delhi, having greater number of neighborhoods, made 3 groups of itself. Indiranagar was found to be the neighborhood with most number of unique venue categories while Lodi Colony has the most in Delhi. Both the cities were found to be comparable between their neighborhoods and the services they offer. People shifting between these cities can easily find a neighborhood they are similar to.

## Conclusion

The project turned out to offer an effective guide to compare the neighborhoods of two most developing metropolitan cities in India. The analysis was performed by only comparing the services each neighborhood has to offer. Though this is an important criteria, other factors such as traffic data, housing prices etc. can be used in the future studies for a better result.