



Capstone Project - The Battle of Neighborhoods

Vinodhini V.

Introduction

- ❖ Business Objective :
 - ❖ Identify the best location to setup a coffee shop in Bangalore based on location proximity and ratings of existing shops in the neighborhood.
 - ❖ For a new business to start-up a coffee-shop in Bangalore, sufficient research must be done to identify the best location. The current analytics project will aid to narrow down the locations based on the coffee-shops in the neighborhood and their ratings.

Data Requirements

- ❖ Data required for this analysis :
 1. List of Neighborhoods in Bangalore -
https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Bangalore
 2. Latitude and Longitude details of the Neighborhoods
 3. List of coffeeshops in each neighborhood (Extracted from Foursquare)
 4. Ratings of the coffee shops (Extracted from Foursquare)
- ❖ Neighborhood details and their latitude and longitude information helps to fetch the details of the coffee shops in each locality.
- ❖ Ratings of each coffee shop helps in the analysis to answer the objective.

Methodology

- ❖ Data Extraction :
 - ❖ Neighborhood Data is extracted from Wikipedia using web-scraping techniques (beautifulsoup).
 - ❖ Latitude and Longitude information is prepared with data available on internet as a CSV.
 - ❖ Coffee shop information and ratings are extracted using Foursquare API.

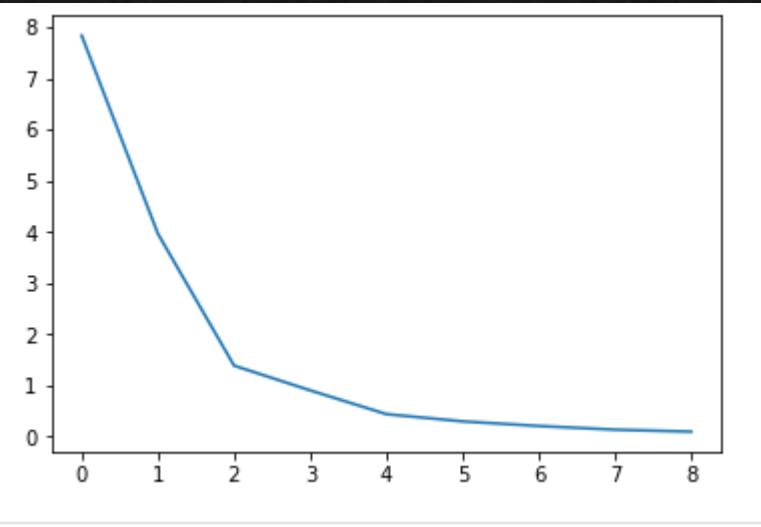
- ❖ Exploratory Analysis :
 - ❖ There are about 64 prominent neighborhoods in the city of Bangalore, spread on the north, south, east and west of Bangalore.
 - ❖ In these 64 locations, 128 coffee shops are located as found on Foursquare.
 - ❖ Only 37 out of 128 coffee shops have a rating listed on Foursquare.
 - ❖ The mean of the rating is 6.3 with maximum being 7.7 and minimum as 5.1.

Methodology

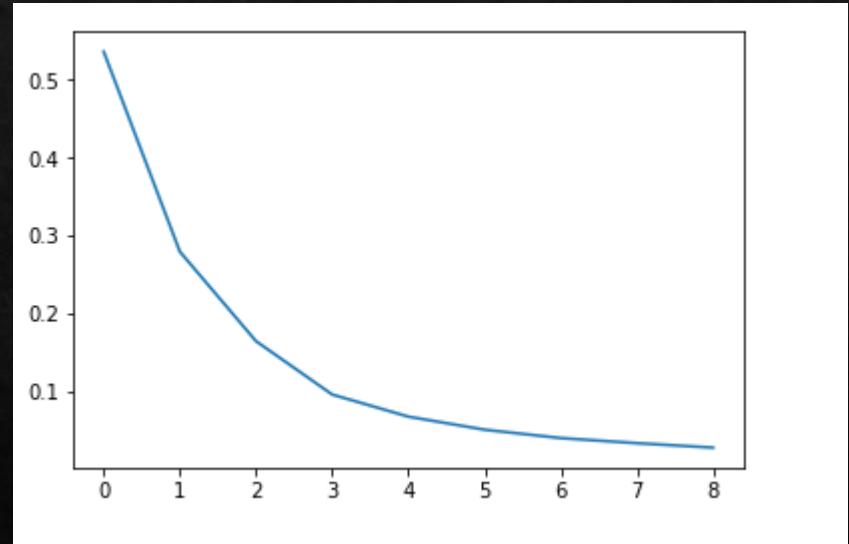
- ❖ Data Clean-up/Preparation:
 - ❖ Any null values are treated by dropping the records.
 - ❖ Coffee shops with no rating are imputed suitably.
- ❖ Machine Learning Model :
 - ❖ Unsupervised Clustering using KMeans is preferred approach.
 - ❖ Clustering will be applied to understand the proximity and distribution of coffee-shops in the neighborhood.
 - ❖ Clustering on the ratings will be done to understand how the coffee shops are distributed based on rating and locality.
 - ❖ **Elbow method** will be applied to identify the suitable number of clusters.
 - ❖ Validations done by plotting the output of cluster over maps using Folium and reviewed manually.

Results

Optimum number of clusters for neighborhood proximity and rating of coffee shops was 2.



Optimum number of clusters for neighborhood proximity of coffee shops was 3.

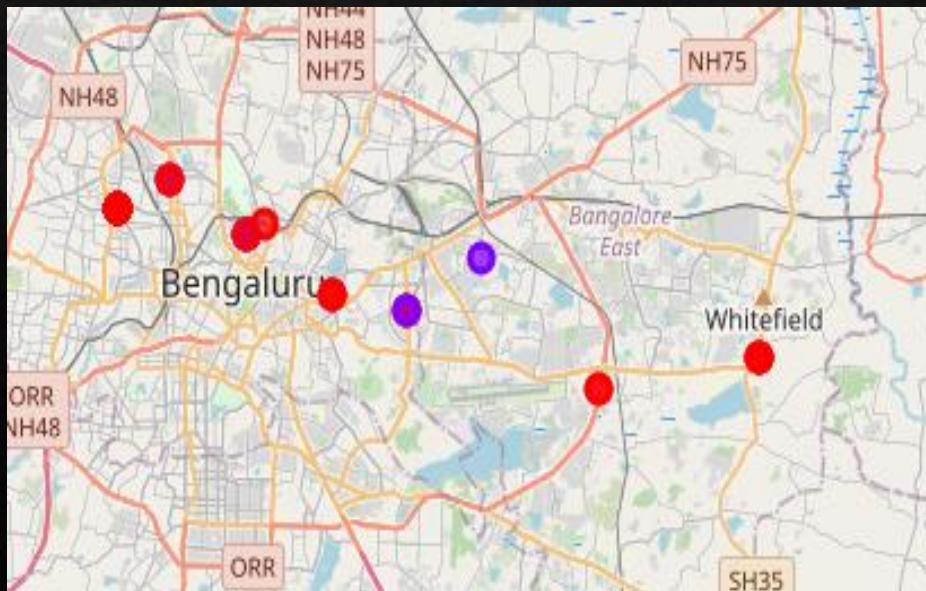


Results

Cluster distribution – Coffeeshops with ratings

Cluster 0 – 32 coffee shops [rating avg – 6.47]

Cluster 1 – 5 coffee shops [rating avg – 5.52]

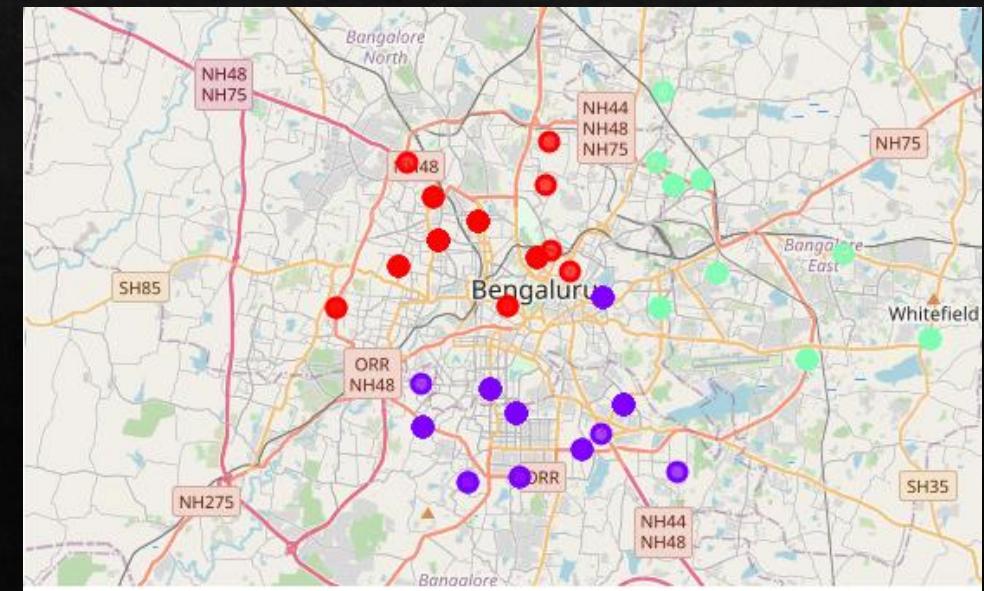


Cluster distribution – Coffeeshops in Neighborhood

Cluster 0 – 52 neighborhoods.

Cluster 1 – 51 neighborhoods

Cluster 2 – 33 neighborhoods



Discussion and Recommendation

- ❖ Bangalore East has coffee shops sparsely distributed compared to other regions.
- ❖ Out of the many Coffee shops listed, only 37 have been rated. Most of the ratings hover around 5-7. Ratings of 6-7 are mostly seen around Bangalore North-West.
- ❖ The outcome of this project along with other Market Research related to
 - i. population and surroundings
 - ii ratings on other platforms for Coffee shops
 - iii Coffee variants and other beverages offeredcan help finalize the optimal location for setting up a Coffee shop.
- ❖ This standalone analysis suggests potential opportunities in East Bangalore as there is lack of good rated coffee shops.

Conclusion

- ❖ Bangalore East seems to be a suitable location for new coffee shop if the business objective is to look for a less competitive environment.
- ❖ If the business strategy is to go for more crowded place, then Bangalore Central seems to be a suitable option.