Neighborhood Analysis of Bangalore, India

Applied Data Science Capstone by IBM / Coursera

By

Manish Kumar Lal

10th April 2019

1. Introduction

1.1. Background

Bangalore, a city in southern India, is called "**The garden city of India**". It has moderate climate all year round and is considered as a perfect retirement place. However, in the past few decades, Bangalore has risen to be the "**IT capital of India**" and is also referred to as the "**Silicon Valley of India**". It has become a popular destination for many companies from all around the world to open their offices. Likewise, the city has seen tremendous migration of people from all over India. As a result, there is a plethora of new businesses opening in the city.

1.2. Problem statement

In this project we will focus on finding an optimal location for opening an **Italian restaurant** in the city. For this we will focus on the following three criteria:

- find locations that have fewer restaurants
- particularly detect areas with no Italian restaurant in the vicinity
- give preference to locations closer to the city center if the first two conditions are satisfied.

1.3. Potential interest

This study is to aid stakeholders in determining the optimal location / neighborhood to open a new Italian Restaurant in Bangalore. It may also be of interest to people moving into Bangalore and for entrepreneurs to choose residential and / or office locations.

2. Data acquisition and cleaning

I utilized the tools discussed in the IBM Data Science Professional Certificate course on Coursera to segment and analyze the neighborhoods in Bangalore based upon the above criteria. I then used the analysis to identify the top locations and discussed these to give the stakeholders enough information to decide on the location for opening an Italian restaurant.

Based upon the project scope, I considered the following factors in the analysis:

- number of restaurants in the neighborhood
- number of and distance to Italian restaurants in the neighborhood, if any

2.1. Data sources

I collected the data for my analysis from the following sources:

- Used two web sites, Wikipedia and Popular Neighborhoods in Bangalore, to high-grade neighborhoods in Bangalore
- Used Search Nearby option on Google Maps to compile the coordinates for each neighborhood
- Queried Foursquare to get the top 100 venues within a radius of 500 meters of each neighborhood

2.2. Data cleaning

I combined the Bangalore neighborhoods data downloaded or scraped from multiple sources into one table that served as the data for my analysis.

2.3. Feature selection

After querying and cleaning the data, there were 15 neighborhoods **(Fig. 1)** and 206 features **(Fig. 2).**

```
Indiranagar
Marathali
Malleswaram
Koramangala
Hebbal
Jayanagar
Ulsoor
Whitefield
J. P. Nagar
Rajajinagar
HSR Layout
K R Puram
C V Raman Nagar
Electronic City
Mahadevapura
```

Fig. 1 Bangalore Neighborhoods

	Neighborhood	Neighborhood_Latitude	Neighborhood_Longitude	Venue	Venue_Latitude	Venue_Longitude	Venue_Category
0	Indiranagar	12.9719	77.6412	Vero Moda	12.972808	77.641225	Boutique
1	Indiranagar	12.9719	77.6412	Krispy Kreme Doughnuts	12.970094	77.640671	Bakery
2	Indiranagar	12.9719	77.6412	The Black Rabbit	12.969891	77.641251	Pub
3	Indiranagar	12.9719	77.6412	Chakum Chukum	12.972516	77.639152	Snack Place
4	Indiranagar	12.9719	77.6412	Red Fork	12.970314	77.642789	Café

Fig. 2 Neighborhood Venues

I then filtered the top 100 venues data queried from Foursquare to include only the restaurant data.

3. Exploratory data analysis

I plotted the neighborhoods on a map to visualize the locations of the neighborhoods across the city. These 15 neighborhoods were evenly distributed across the city center **(Fig. 3)**.

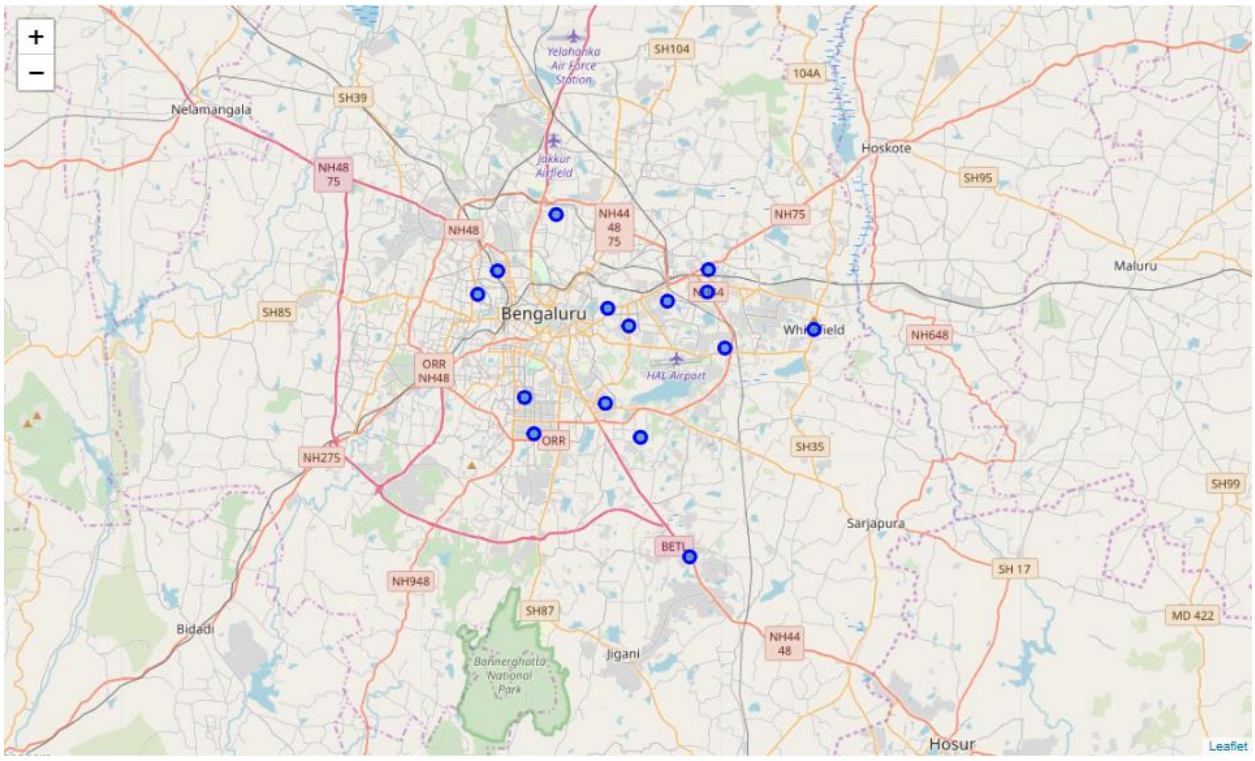

Fig. 3 Neighborhood Map of Bangalore

I selected the "Restaurant" category **(Fig. 4)** for each neighborhood.

	Neighborhood	Neighborhood_Latitude	Neighborhood_Longitude	Venue	Venue_Latitude	Venue_Longitude	Venue_Category
0	Indiranagar	12.9719	77.6412	Mamagoto	12.970113	77.639173	Asian Restaurant
1	Indiranagar	12.9719	77.6412	Bricklane Grill	12.969920	77.641093	Restaurant
2	Indiranagar	12.9719	77.6412	Sri Udupi Park	12.973617	77.641104	Udupi Restaurant
3	Indiranagar	12.9719	77.6412	Chianti Ristorante & Wine Bar	12.970167	77.640346	Italian Restaurant
4	Indiranagar	12.9719	77.6412	Mother Cluckers	12.970176	77.640243	Restaurant

Fig. 4 Neighborhood Restaurants

3.1. Calculation of target variable

There were some restaurants for which the restaurant type was not defined. I grouped these under "Others" category **(Fig. 5)**.

	Neighborhood	Neighborhood_Latitude	Neighborhood_Longitude	Restaurant	Restaurant_Latitude	Restaurant_Longitude	Restaurant_Type
0	Indiranagar	12.9719	77.6412	Mamagoto	12.970113	77.639173	Asian Restaurant
1	Indiranagar	12.9719	77.6412	Bricklane Grill	12.969920	77.641093	Others
2	Indiranagar	12.9719	77.6412	Sri Udupi Park	12.973617	77.641104	Udupi Restaurant
3	Indiranagar	12.9719	77.6412	Chianti Ristorante & Wine Bar	12.970167	77.640346	Italian Restaurant
4	Indiranagar	12.9719	77.6412	Mother Cluckers	12.970176	77.640243	Others

Fig. 5 Restaurants with No Defined Type Grouped as "Others"

3.2. Restaurant density by area

The top 3 neighborhoods with the highest number of restaurants were J. P. Nagar, Indiranagar, and Jayanagar **(Fig. 6)**.

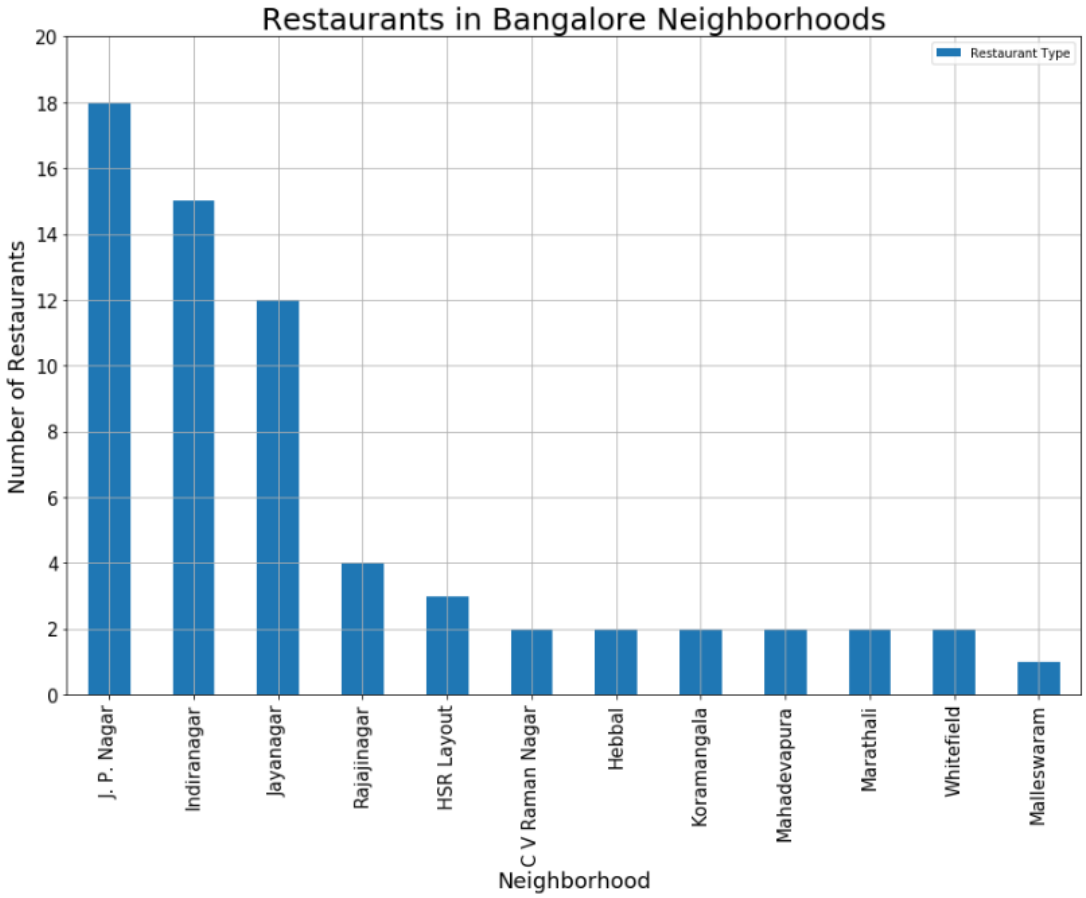

Fig. 6 Restaurants in Bangalore Neighborhoods

3.3. Restaurant density by type

After grouping the restaurants by type, I found that the highest number of restaurants was "Indian" type followed by "Others" **(Fig. 7)**.

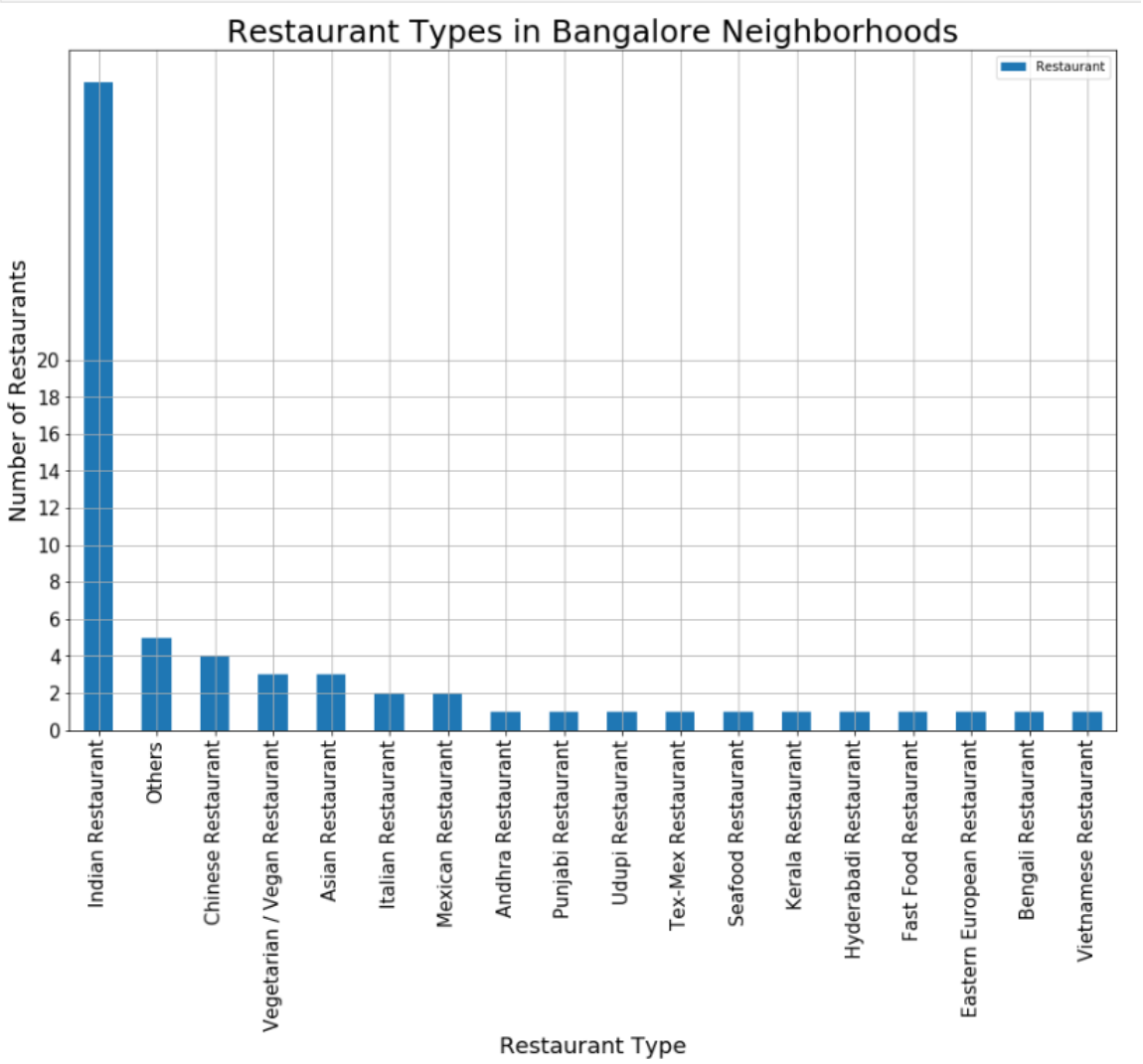

Fig. 7 Restaurant Types in Bangalore Neighborhoods

4. Predictive analysis

Since the Indian Restaurants were the highest amongst restaurant types, I reviewed the results after excluding the Indian restaurants **(Fig. 8)**. The restaurants classified as “Others’ were highest in number followed by Chinese, Vegetarian / Vegan, and Asian restaurants.

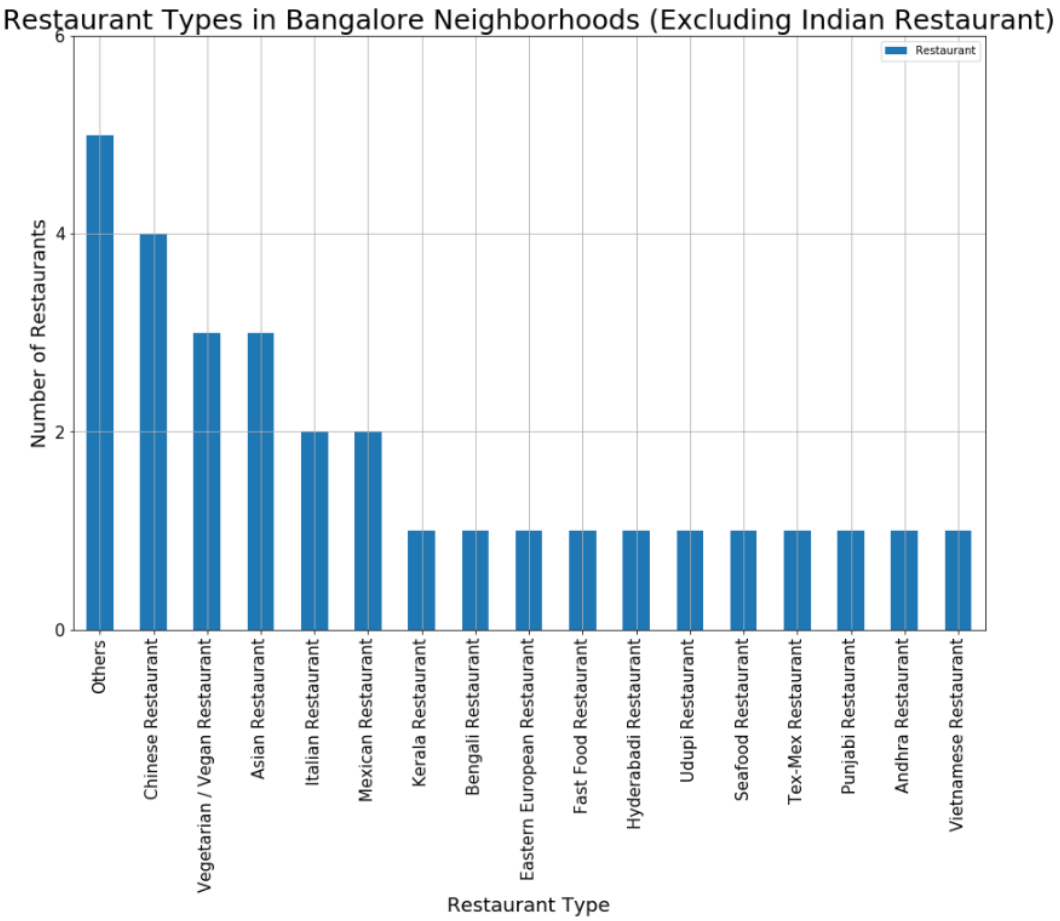

Fig. 7 Restaurant Types Excluding Indian Restaurants

4.1. Italian restaurant density

I filtered the dataset and found that there were only 2 Italian restaurants in the neighborhood **(Fig. 8)**. Both these restaurants were in Indiranagar. I used Folium map to display the location of these 2 Italian restaurants **(Fig. 9)**.

	Neighborhood	Neighborhood_Latitude	Neighborhood_Longitude	Restaurant	Restaurant_Latitude	Restaurant_Longitude	Restaurant_Type	Italian Restaurant
3	Indiranagar	12.9719	77.6412	Chianti Ristorante & Wine Bar	12.970167	77.640346	Italian Restaurant	1
13	Indiranagar	12.9719	77.6412	Terrazzo	12.969337	77.641404	Italian Restaurant	1

Fig. 8 Filtered dataset for Italian Restaurants

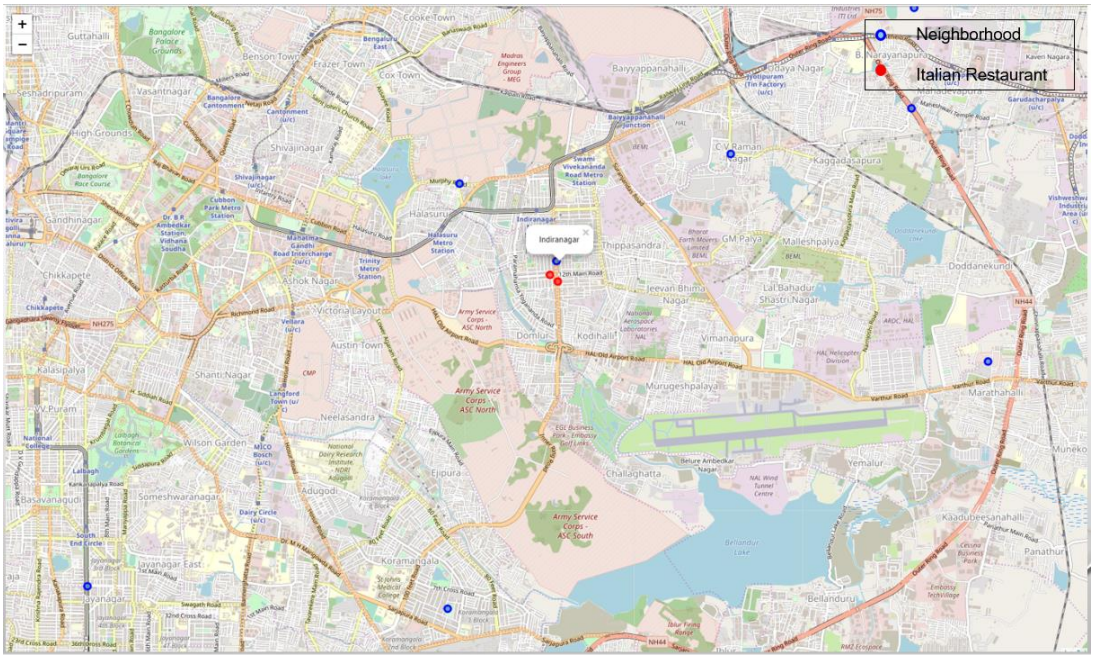

Fig. 9 Location of Existing Italian Restaurants

4.2. Potential locations for new Italian Restaurant

Since there are only 2 Italian restaurants in the neighborhoods, there is scope to open a new one in any of the neighborhoods. Based upon this study, we could possibly exclude J. P. Nagar, Indiranagar, and Jayanagar that have the highest number of restaurants to minimize competition. Also, we could particularly exclude Indiranagar that already has 2 Italian restaurants. All other neighborhoods appear as potential locations for the new Italian Restaurant.

5. Results and discussion

My analysis shows that there is a fairly large number of restaurants in the selected neighborhoods in Bangalore City (~65 restaurants in the 15 popular neighborhoods). The top 3 neighborhoods with the highest number of restaurants were J. P. Nagar, Indiranagar, and Jayanagar.

A closer look at the restaurant types showed that there is a predominance of Indian Restaurants in the neighborhoods. Of the 65 total restaurants in the 15 queried neighborhoods, 35 were Indian Restaurants. I narrowed the search to look for Italian Restaurants in the neighborhoods and found that there were only 2 Italian Restaurant, and both were in the Indiranagar.

The result of my analysis is that all neighborhoods, other than Indiranagar, are probable candidates for opening an Italian Restaurant. Further, since the neighborhoods of J. P. Nagar and Jayanagar have 18 and 12 restaurants (any kind) respectively, choosing a neighborhood other that these 3 neighborhoods (Indiranagar,

J. P. Nagar, and Jayanagar) could serve the dual purpose of lesser competition from any restaurant type and no competition from Italian Restaurant type.

This does not imply that Indiranagar, J. P. Nagar, and Jayanagar and not good locations for Italian Restaurant. The purpose of this analysis was to only provide information about Bangalore neighborhoods that are not crowded with restaurants, particularly Italian. It is possible that, for any reason, an Italian Restaurant in a neighborhood with smaller number of restaurants and no Italian Restaurant could be unsuitable for a new restaurant regardless of the lack of competition in the area. Recommended neighborhoods should be considered as a starting point for a detailed analysis which considers other factors in addition to the number of restaurants and competition.

6. Conclusion

The goal of this project was to analyze the neighborhoods in Bangalore, India to identify neighborhoods with lower number of restaurants, particularly Italian Restaurants. This would aid the stakeholders in narrowing down the search for an optimal location for an Italian Restaurant in Bangalore. By getting the restaurant data from Foursquare, calculating the restaurant density distribution, and visualizing the results in the form of bar charts and maps, it was apparent that the neighborhoods of Indiranagar, J. P. Nagar, and Jayanagar had the highest restaurant density. There were only 2 Italian Restaurants, and both were in Indiranagar. This suggested that the first step is to select any neighborhood other than Indiranagar, J. P. Nagar, and Jayanagar to avoid competition from both the existing restaurants, particularly Italian Restaurants. However, for reasons not explored in this study, these 3 locations could also be good candidates for a new restaurant despite competition from the existing restaurants.

7. Recommendation

Final decision on optimal restaurant location should be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended neighborhood, taking into consideration additional factors like attractiveness of each location (number of offices and malls), proximity to major roads and local train station locations, potential future growth of residential areas and new offices, social and economic dynamics of each neighborhood etc.