The Battle of Neighborhoods (Week 2)

Finding optimal Location for opening a Restaurant

1. Introduction

New Delhi, national capital of India. It is situated in the north-central part of the country on the west bank of the Yamuna River, adjacent to and just south of Delhi city (Old Delhi) and within the Delhi national capital territory. A symbol of the country's rich past and thriving present, Delhi is a city where ancient and modern blend seamlessly together. It is a place that not only touches your pulse but even fastens it to a frenetic speed. Home to millions of dreams, the city takes on unprecedented responsibilities of realizing dreams bringing people closer and inspiring their thoughts. No visit to Delhi is complete without an experience of its famed specialties such as its delicious curries, barbecued tikkas and kebabs. The city is famous for a variety of foods, dishes, and restaurants. You can't miss trying the famous Indian foods when you're visiting.

1.1 Problem

New Delhi is rich in its Indian cuisine. It has huge variety of Indian foods from different parts of India. New Delhi seems to be a better place for setting up an Indian restaurant. Since there are lots of restaurants in New Delhi we will try to detect locations that are not already crowded with restaurants. We are also particularly interested in areas with no Indian restaurants in vicinity.

1.2 Target Audience

Companies or Individuals looking into opening a restaurant would be interested in prediction of optimal location. This project will provide an analysis whether the venture is feasible or not.

2.Data acquisition and cleaning

2.1 Data Sources

Delhi has a total of 9 boroughs and 163 neighborhoods. In order to analyze the neighborhoods and explore them, we will essentially need a dataset that contains the 9 boroughs and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood.

Luckily, this dataset exists for free on the Kaggle's website, here is the link to the dataset: https://www.kaggle.com/shaswatd673/delhi-neighborhood-data#delhi_dataSet.csv

Based on definition of our problem, factors that will influence our decision is the number of existing restaurants in the neighborhood (any type of restaurant). Number of restaurants and their type and location in every neighborhood will be obtained using Foursquare API .

2.2 Data Cleaning

Data downloaded from Kaggle's website contains Borough, their Neighborhoods and latitude and longitude coordinates of each neighborhood.
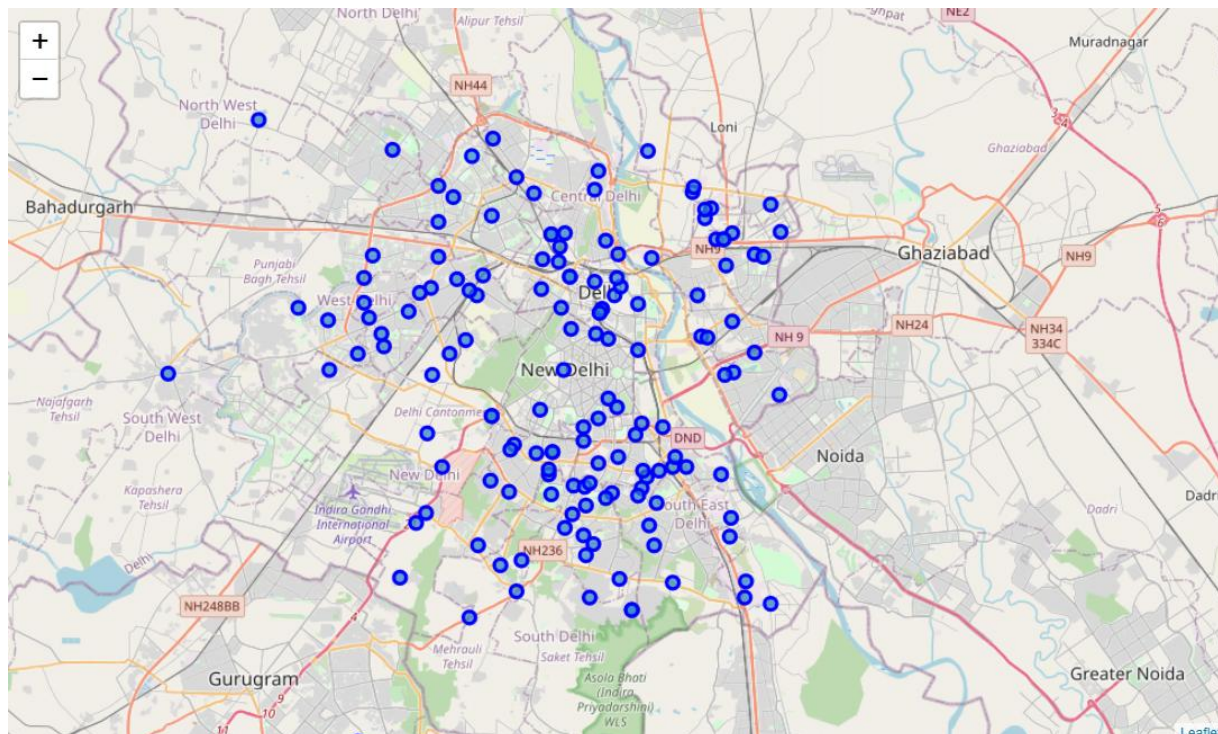
	Unnamed: 0	Borough	Neighborhood	latitude	longitude
0	0	North West Delhi	Adarsh Nagar	28.614192	77.071541
1	1	North West Delhi	Ashok Vihar	28.699453	77.184826
2	2	North West Delhi	Azadpur	28.707657	77.175547
3	3	North West Delhi	Bawana	28.799660	77.032885
4	4	North West Delhi	Begum Pur	NaN	NaN

First we remove the unnamed column from the dataset, and the drop all the columns that contains missing (NaN) values.

	Borough	Neighborhood	latitude	longitude
0	North West Delhi	Adarsh Nagar	28.614192	77.071541
1	North West Delhi	Ashok Vihar	28.699453	77.184826
2	North West Delhi	Azadpur	28.707657	77.175547
3	North West Delhi	Bawana	28.799660	77.032885
5	North West Delhi	Dhaka	39.031714	-90.261223
6	North West Delhi	Jahangirpuri	28.725972	77.162658

2.3 Feature Selection

After cleaning we have 9 boroughs and 163 neighborhoods in dataset. Naw, we get coordinates of delhi and plot a map of Delhi,IN showing all the neighborhood on it.

3. Methodology

In this section, we will conduct exploratory data analysis. First, we slice the original dataframe and create a new dataframe of the New Delhi data.

	Borough	Neighborhood	latitude	longitude
0	New Delhi	Barakhamba Road	28.629142	77.226149
1	New Delhi	Chanakyapuri	28.594678	77.188521
2	New Delhi	Connaught Place	28.631383	77.219792
3	New Delhi	Gautampuri	28.511570	77.302623
4	New Delhi	Gole Market	28.633719	77.205627

Then we get location of New Delhi and create a map of New Delhi Using latitude and longitude values.
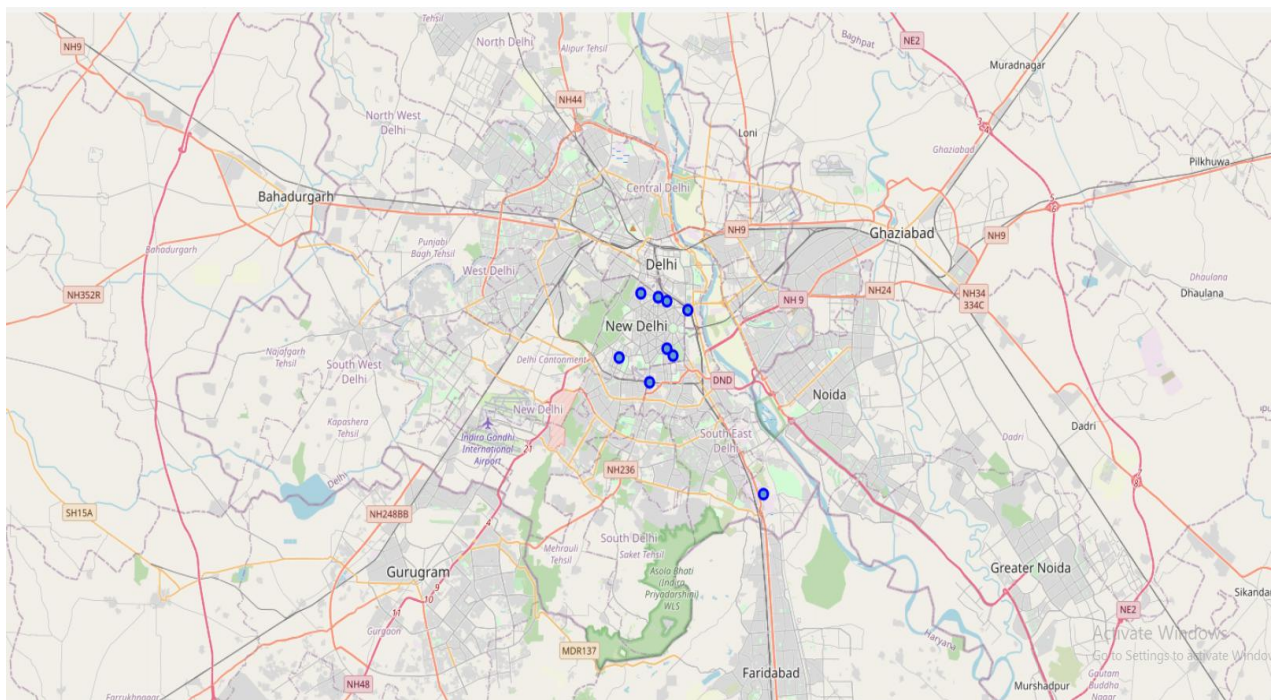
I utilized the Foursquare API to explore the boroughs and segment them. I designed the limit as 30 venue and the radius 1000 meter for each borough from their given latitude and longitude informations. Here is a head of the list Venues name, category, latitude and longitude informations from Foursquare API.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Barakhamba Road	28.629142	77.226149	Bengali Market | बंगाली मार्केट | বাংলা বাজার	28.629498	77.232020	Indian Restaurant
1	Barakhamba Road	28.629142	77.226149	Naturals Ice Cream	28.634455	77.222139	Ice Cream Shop
2	Barakhamba Road	28.629142	77.226149	Rajdhani Thali	28.629999	77.220401	Indian Restaurant
3	Barakhamba Road	28.629142	77.226149	The Imperial	28.625548	77.218664	Hotel
4	Barakhamba Road	28.629142	77.226149	HOTEL SARAVANA BHAVAN	28.627041	77.219514	South Indian Restaurant
5	Barakhamba Road	28.629142	77.226149	Connaught Place | कनॉट प्लेस (Connaught Place)	28.632731	77.220018	Plaza
6	Barakhamba Road	28.629142	77.226149	Tamasha	28.629663	77.221835	Gastropub
7	Barakhamba Road	28.629142	77.226149	Farzi Cafe	28.632581	77.221125	Molecular Gastronomy Restaurant
8	Barakhamba Road	28.629142	77.226149	Johnny Rockets	28.630457	77.219594	Bistro
9	Barakhamba Road	28.629142	77.226149	Triveni Kala Sangam | त्रिवेणी कला संगम	28.627278	77.232137	Arcade
10	Barakhamba Road	28.629142	77.226149	Triveni Terrace Cafe	28.627295	77.232677	Café

There are 60 unique categories returned by Foursquare API.

Then I have made dataset of top 9 venues for each neighborhood returned by Foursquare API.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
0	Barakhamba Road	Indian Restaurant	Café	Lounge	Bakery	Plaza	Ice Cream Shop	Gastropub	Food Truck	Coffee Shop
1	Chanakyapuri	Indian Restaurant	Nightclub	Café	Chinese Restaurant	Park	Asian Restaurant	Coffee Shop	Hotel	Italian Restaurant
2	Connaught Place	Indian Restaurant	Lounge	Café	South Indian Restaurant	Bistro	Ice Cream Shop	Gastropub	Food Truck	Deli / Bodega
3	Gautampuri	Indian Restaurant	Train Station	Snack Place	Café	Hotel	Historic Site	Golf Course	Gastropub	Furniture / Home Store
4	Gole Market	Snack Place	Hotel	Indian Restaurant	Japanese Restaurant	Bar	Breakfast Spot	Art Museum	Coffee Shop	Antique Shop

Then, Find optimum k value for k-means clustering.
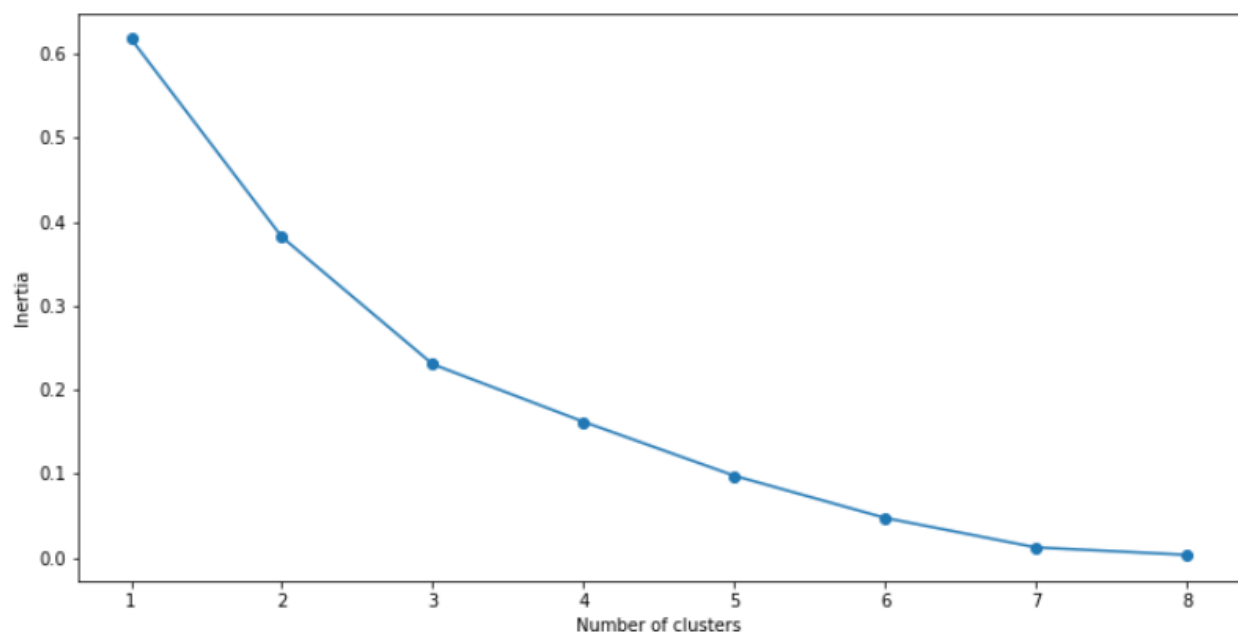
When we changed the cluster value from 1 to 3, the inertia value reduced very sharply. This decrease in the inertia value reduces and eventually becomes constant as we increase the number of clusters further. Here, we can choose any number of clusters between 4 and 8. We can have 5, 6, or even 8 clusters. We must also look at the computation cost while deciding the number of clusters. If we increase the number of clusters, the computation cost will also increase. So, We will now run K-means clustering for K=4.

I have used unsupervised learning K-means algorithm to cluster the borough. K-Means algorithm is one of the most common cluster method of unsupervised learning. First, created a new dataframe that includes the cluster as well as the top 9 venues for each neighborhood.

	Borough	Neighborhood	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
0	New Delhi	Barakhamba Road	28.629142	77.226149	0	Indian Restaurant	Café	Lounge	Bakery	Plaza	Ice Cream Shop	Gastropub	Food Truck	Coffee Shop
1	New Delhi	Chanakyapuri	28.594678	77.188521	0	Indian Restaurant	Nightclub	Café	Chinese Restaurant	Park	Asian Restaurant	Coffee Shop	Hotel	Italian Restaurant
2	New Delhi	Connaught Place	28.631383	77.219792	0	Indian Restaurant	Lounge	Café	South Indian Restaurant	Bistro	Ice Cream Shop	Gastropub	Food Truck	Deli / Bodega
3	New Delhi	Gautampuri	28.511570	77.302623	2	Indian Restaurant	Train Station	Snack Place	Café	Hotel	Historic Site	Golf Course	Gastropub	Furniture / Home Store
4	New Delhi	Gole Market	28.633719	77.205627	3	Snack Place	Hotel	Indian Restaurant	Japanese Restaurant	Bar	Breakfast Spot	Art Museum	Coffee Shop	Antique Shop

4. Result

Here I have examine all the 4 clusters.

Cluster 1

```
ND_merged.loc[ND_merged['Cluster Labels'] == 0, ND_merged.columns[[1] + list(range(5, ND_merged.shape[1]))]]
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
0	Barakhamba Road	Indian Restaurant	Café	Lounge	Bakery	Plaza	Ice Cream Shop	Gastropub	Food Truck	Coffee Shop
1	Chanakyapuri	Indian Restaurant	Nightclub	Café	Chinese Restaurant	Park	Asian Restaurant	Coffee Shop	Hotel	Italian Restaurant
2	Connaught Place	Indian Restaurant	Lounge	Café	South Indian Restaurant	Bistro	Ice Cream Shop	Gastropub	Food Truck	Deli / Bodega
5	Golf Links	Indian Restaurant	Restaurant	Café	Coffee Shop	Bookstore	Hotel	Asian Restaurant	Golf Course	Hotel Bar
7	Khan Market	Indian Restaurant	Café	Restaurant	Coffee Shop	Bookstore	Chinese Restaurant	Asian Restaurant	Mediterranean Restaurant	Hotel
8	Pragati Maidan	Theater	Art Gallery	Light Rail Station	Indian Restaurant	Plaza	Udupi Restaurant	Furniture / Home Store	Flea Market	Coffee Shop

Cluster 2

```
ND_merged.loc[ND_merged['Cluster Labels'] == 1, ND_merged.columns[[1] + list(range(5, ND_merged.shape[1]))]]
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
6	INA Colony	Indian Restaurant	Restaurant	Antique Shop	Athletics & Sports	Bakery	Coffee Shop	Market	Metro Station	Airport

Cluster 3

```
ND_merged.loc[ND_merged['Cluster Labels'] == 2, ND_merged.columns[[1] + list(range(5, ND_merged.shape[1]))]]
```

|	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
3	Gautampuri	Indian Restaurant	Train Station	Snack Place	Café	Hotel	Historic Site	Golf Course	Gastropub	Furniture / Home Store

Cluster 4

```
ND_merged.loc[ND_merged['Cluster Labels'] == 3, ND_merged.columns[[1] + list(range(5, ND_merged.shape[1]))]]
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
4	Gole Market	Snack Place	Hotel	Indian Restaurant	Japanese Restaurant	Bar	Breakfast Spot	Art Museum	Coffee Shop	Antique Shop

Here, We can clearly see that Cluster 1 is the cluster with the restaurants as the most common venue. Then visualize the resulting clusters, showing all the clusters in different colures.
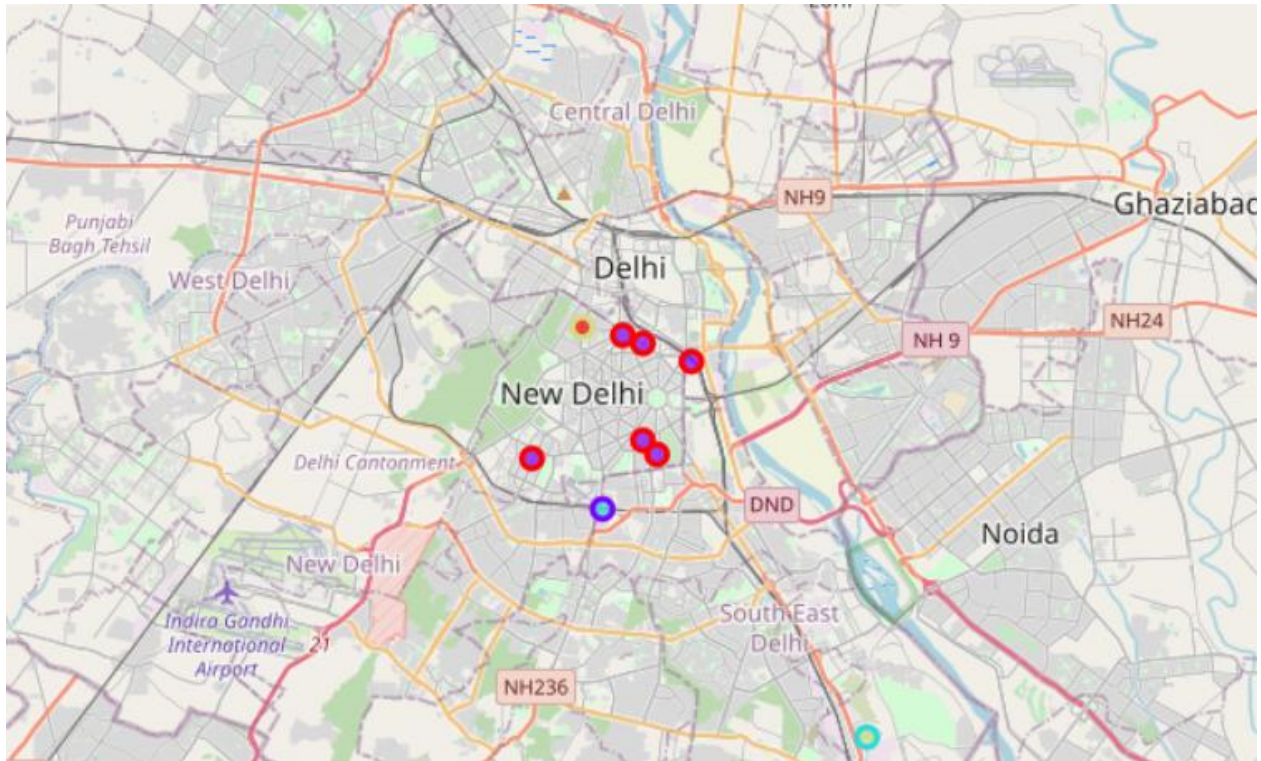
We will plot the mean value of Indian restaurants for each selected neighborhood in the cluster to study the presence of competition in each neighborhood and visualize them.

Density of Indian Restaurant in each Neighborhood
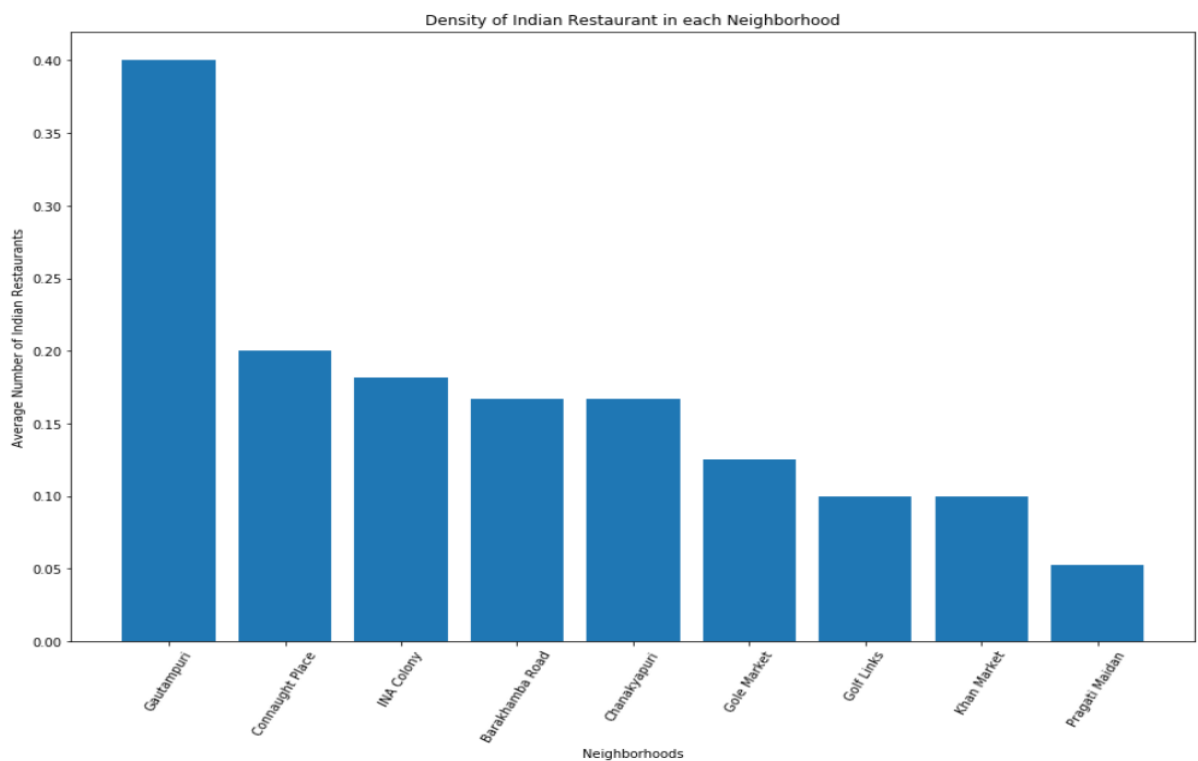
5. Discussion

From the graph, we can see that the first neighborhood, Gautampuri, has the most competition for Indian Restaurants in New Delhi. This indicating that it has the greatest obstacles in opening a new restaurant. Gautampuri has almost double the competition than any other neighborhoods. Connaught Place is the second neighborhood with the most Indian restaurants, followed by INA Colony. Barakhamba Road and Chanakyapuri has almost same but moderate competition. However, the following neighborhoods have moderate competition, which will enable a new business to establish easily:

> 1- Gole Market
>
> 2- Golf Links
>
> 3- Khan Market
>
> 4- Pragati Maidan

Note that these recommendations are based on some assumptions of the analysis,like:

> 1- Radius of the opportunity of each neighborhood was considered as 1000 meters from the location,
>
> 2- Recommendation opportunities are based on absence of a restaurant which is likely to be appreciated in the top 9 venues

6. Conclusion

This project recommends some of the ideal places to open Indian restaurant in New Delhi, India. The analysis shows there are better chances for opening restaurant in Gole Market, Golf Links, Khan Market, and Pragati Maidan. INA Colony, Barakhamba Road and Chanakyapuri are also good location. This analysis can be helpful for the individuals looking for opening a restaurant or expanding business. This analysis shows the feasible venture and competition landscape of the area.

Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location , levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.