**Analyzing different Neighbourhoods of New Delhi and finding the best place to start an Italian Restaurant**

1. Introduction
   1. Background

New Delhi is the capital city of India. The city itself has a population of 257,803. However, the much larger metro area has a population that exceeds 26 million. With its diverse culture, comes diverse food items. There are many restaurants in New Delhi, each belonging to different categories like Chinese, Italian, French, Mughlai, North Indian etc.

* 1. The Idea

So in this project, I would need to list and visualize all major parts of the city and find out which part is the go-to place for Italian food so that a future restaurateur who wants to open a Italian cuisine restaurant can find the best part of the city in which the demand for Italian cuisine is very high.

I would need to use Foursquare API to get the location data on nearby venues and their longitudes and other information to make a good analysis of the data that I have and create an efficient cluster model.

**Questions a future restaurateur might ask if he/she were to open an Italian cuisine restaurant in New Delhi**

* Which neighbourhoods have greater number of restaurants?
* Which neighbourhoods have fewer number of restaurants?
* What neighbourhoods has greater demand for Italian Cuisine?

1. Data acquisition and cleaning
   1. Data sources

### For this project we need the following data:

### Delhi Neighbourhood Dataset that contains a list of sort Borough, Neighbourhood, latitude and longitude.

### Data source: Delhi Neighbourhood Dataset

### Dataset: [https://www.kaggle.com/shaswatd673/delhi-neighbourhood-data](https://www.kaggle.com/shaswatd673/delhi-neighborhood-data)

### This data set contains the required information. And we will use this data set to explore various neighbourhoods of New Delhi city.

### Nearby places in each locality of New Delhi city.

### Data source: Foursquare API: <https://developer.foursquare.com/>

### By using this API, we can get all the venues info including longitudes and latitudes, reviews and ratings of each venues in the neighbourhood.

### Approach

•Collect the necessary the New Delhi city data from Delhi Neighbourhood Dataset that was downloaded from Kaggle

•Using Foursquare API, we will find all venues for each neighbourhood.

•Filter out all venues that are nearby by the neighbourhood.

•Cluster the neighbourhoods and find the best neighbourhoods in the cluster for opening an Italian Restaurant

•Visualize the neighbourhoods clusters using folium library on a map

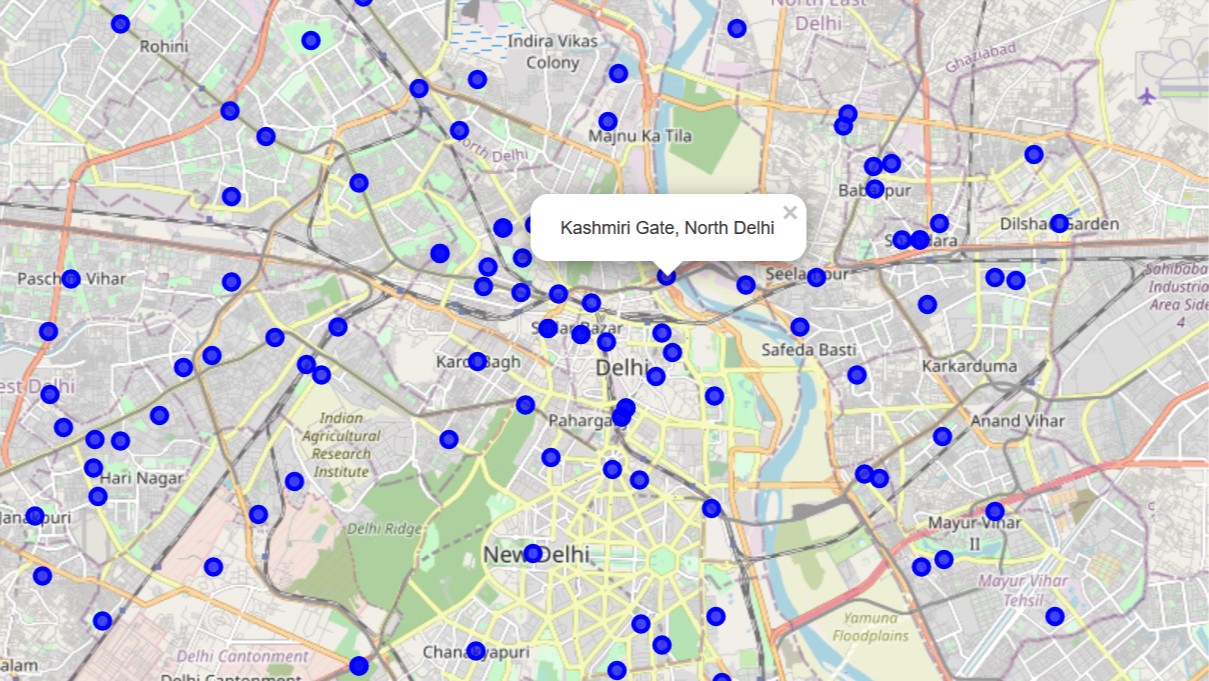
* 1. Data cleaning

After reading the CSV file into a data frame which contains the neighbourhood data of New Delhi, I cleaned the data frame by dropping unnecessary columns and removed Null records or records which had missing values.

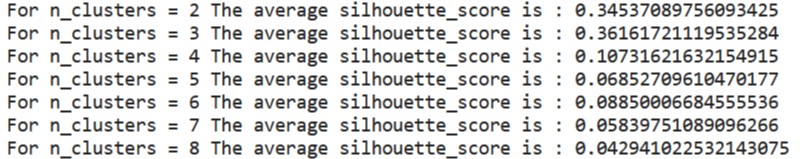
The dataset had had 184 records of neighbourhood with longitudes and latitudes. Some of the coordinates were either missing or wrong, so I fixed them by adding correct coordinates.

3. Exploratory Data Analysis

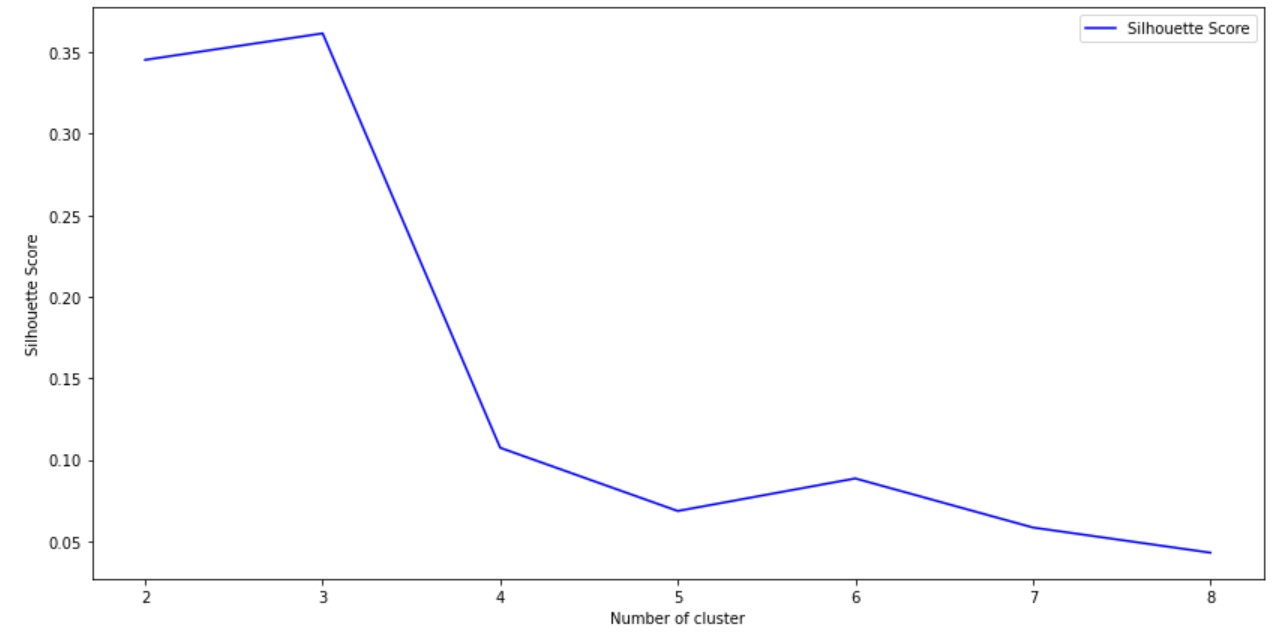
From the map of the New Delhi that I visualized using the folium library, I could observe the different neighbourhoods of New Delhi.



I used the Silhouette Coefficient Method to find the optimal numbers of clusters.



Then I visualized it using matplotlib-

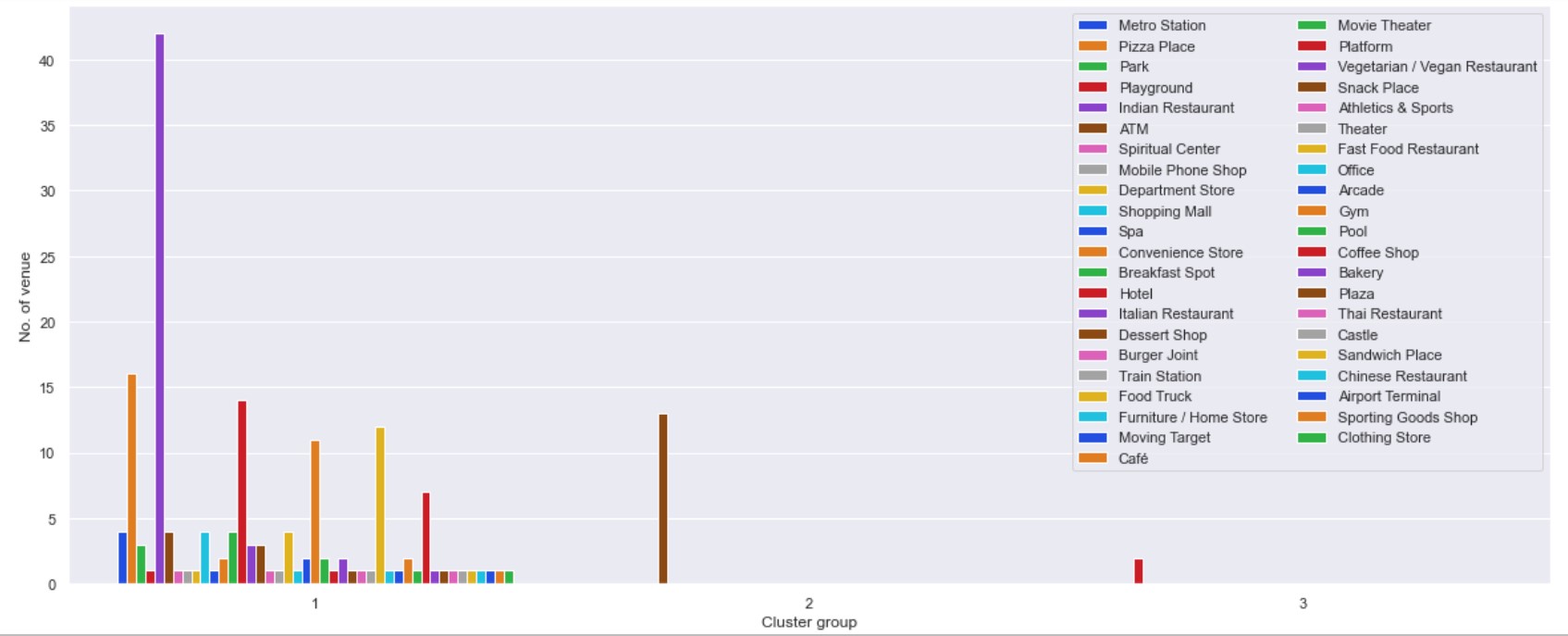


From the above plot, we observe that 3 is the optimal number of clusters.

So, I applied K-Means Cluster to divide the neighbourhoods into three clusters.



I visualized the neighbourhoods clusters with countplot using Seaborn library to check which Cluster has the most numbers of Italian Restaurants.



So, in Cluster 1, we have the highest number of Italian Restaurants and in Cluster 2, we mostly have ATMs and in Cluster 3, we have Playgrounds.

I found the best neighbourhoods for opening an Italian Restaurant by comparing the Most Common Venues between the neighbourhoods which had an Italian Restaurant in Cluster 1. I created three data frames where each of them had “Italian Restaurant” as 1st, 2nd or 3rd Most Common Venues.

Fig 1: Neighbourhoods with “Italian Restaurant” as 1st Most Common Venue



Fig 2: Neighbourhoods with “Italian Restaurant” as 2nd Most Common Venue



Fig 3: Neighbourhoods with “Italian Restaurant” as 3rd Most Common Venue

The data frames with the Clusters are:

Fig 4: Data frame containing neighbourhoods of Cluster 1



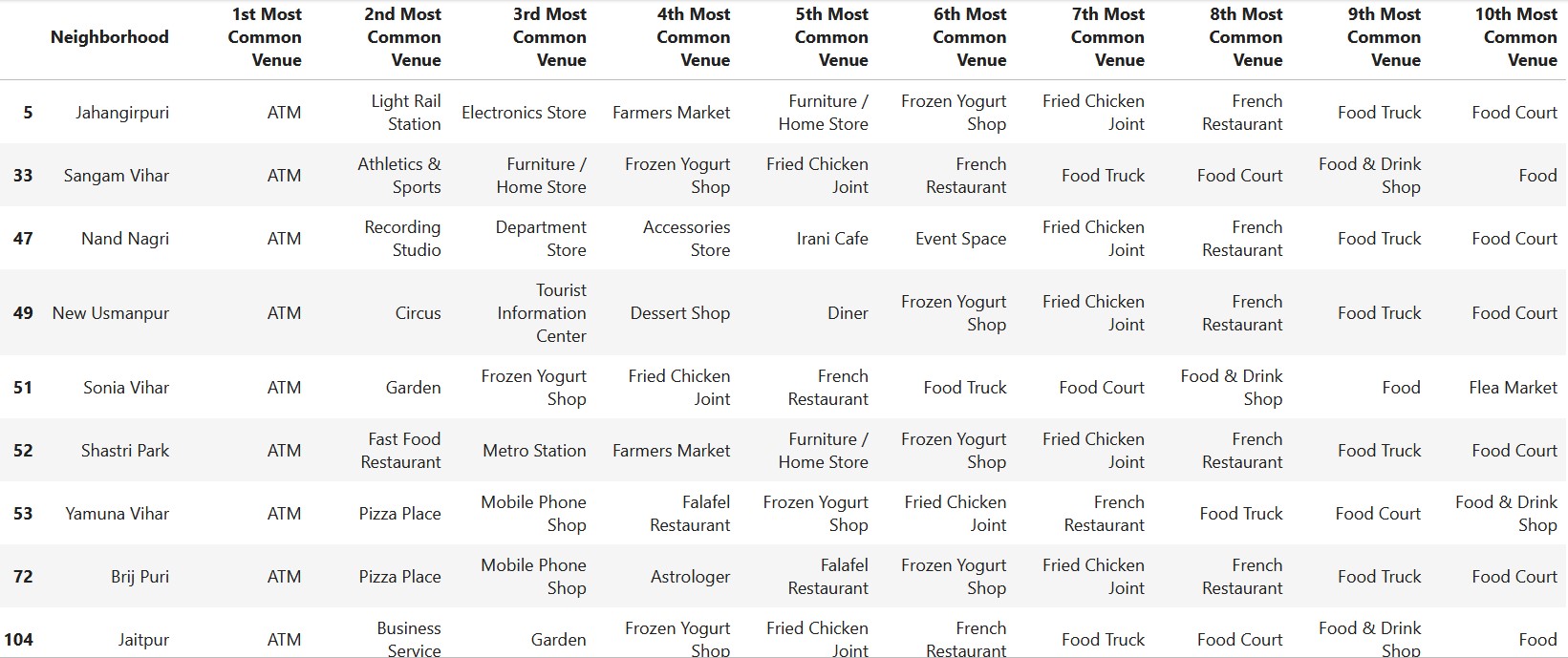
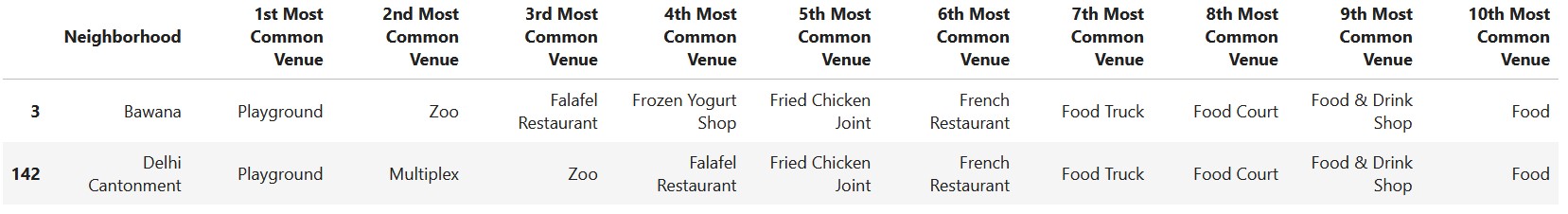
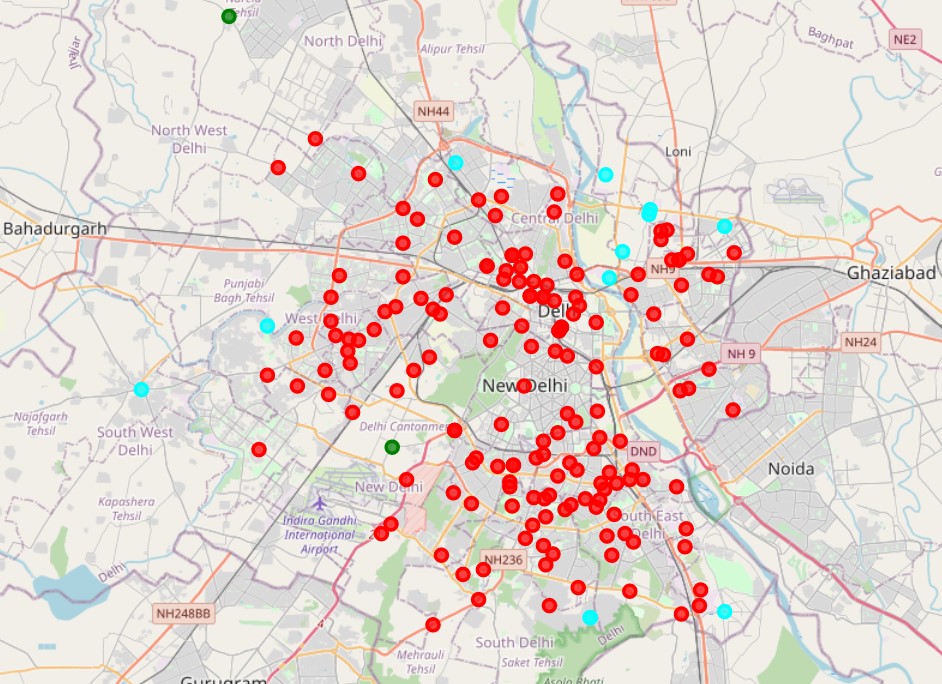
Fig 5: Data frame containing neighbourhoods of Cluster 2

Fig 6: Data frame containing neighbourhoods of Cluster 3



And then I visualized the clusters into the map of New Delhi using Folium library to find the similarities between neighbourhoods.



4. Conclusion

* In Cluster 1:
  + Chhatarpur, Civil Lines and Kashmiri Gate are the best place to start a new Italian cuisine restaurant.
  + Defence Colony and Panjabi Bagh, Krishna Nagar, Kotla Mubarakpur and Sundar Nagar are good alternatives.
* Cluster 2 and Cluster 3 mostly has Pizza places and Fast Food Restaurants. So, if an in-depth survey is done on the neighbourhoods of Cluster 2, we can find out if there are any interest in Italian Restaurants there.