Analysing Food Quality and Culture in Indian Cities

Sandeep A. Thorat

15th June 2021

1. Introduction

India is the second most populated country in the world with nearly a fifth of the world's population. India is one of biggest country in the world with total geo-graphical area of 3,287,240 sq.km. The country has very diverse culture which changes across the geogrphical areas.

Quality and variety of food being served in the Indian restaurants is my area of interest. Many people travel across the country for various reasons like tourism, jobs, business, family etc. After visiting a new place foody people always want to taste the best cuisines of every part of the city which lies in their budget. Analysing best restaurants of different cities is goal of this project.

Zomato is an Indian multinational restaurant aggregator and food delivery company founded by Pankaj Chaddah and Deepinder Goyal in 2008. Zomato provides information, menus and user-reviews of restaurants as well as food delivery options from partner restaurants in many cities. As of 2019, the service is available in 24 countries and in more than 10,000 cities. [Source : Wikipedia].

Zomato restaurant data across various cities from the globe is available on Kaggle.

2. Problems of Interest

The objectives behind this project activity are to answer following questions:

1. Find cumulative number of restaurants in each City of India.
2. Focus on the City having maximum Restaurants (in this case it is New Delhi)
3. Show restaurant clusters from Delhi using folium maps. Use k-means for clustering.
4. Find out 5 places from New Delhi which have best restaurant.
5. Find out New Delhi's 5 worst places for eating
6. Find out top 5 localities of New Delhi having maximum and minimum number of Restaurants
7. Find best places for Chinese restaurant in New Delhi city
8. Group and summarize the data based on localities in New Delhi

3. Data Acquisition and Pre-processing

3.1 Data Sources

For this project we are going to use following datasets :

- Zomato kaggle dataset that contains list of Indian cities, Localities, Restaurant names, Ratings.
    - Data source : Zomato kaggel dataset
    - Description : This data set contains the required information. And we will use this data set to explore various cities of India.
- Foursquare Places API.
    - Data source : Fousquare API
    - Description : The Foursquare Places API provides location based experiences with diverse information about venues, users, photos, and check-ins. The API supports real time access to places, Snap-to-Place that assigns users to specific locations, and Geo-tag.

3.2 Data Pre-processing

Following steps are performed as part of data pre-processing.

a. Remove the unwanted columns and rows from dataset
   Following attributes are considered as important and all remaining attributes are discarded : 'Restaurant Name', 'City', 'Locality', 'Longitude', 'Latitude', 'Cuisines', 'Aggregate rating', 'Rating text', 'Votes'
b. Remove all records which have incorrect Longitude value (i.e. 0.0)
c. Remove Restaurants having aggregate rating 0.0

4. Exploratory Data Analysis

Following methodology is used for performing exploratory data analysis of Kaggle Zomato dataset.

a. Collect the Indian restaurant's data from Zomato kaggel dataset
b. Perform data pre-processing to remove unnecessary fields from the dataset
c. Use FourSquare API to find all venues for each city.
d. Answer Questions of interest using data analytics appraoch.
e. Visualize the results using different libraries in Python

Following are some important findings from this analysis.

There are total 9551 records on Kaggle dataset; out of which 8652 records are from India.

After pre-processing 6301 records related with Indian restaurants is available for analysis.

Following are top 5 cities from India having maximum records in Kaggle dataset.

```
New Delhi       3975
Gurgaon          874
Noida            676
```

```
Faridabad          147
Ghaziabad           23
```

Since Restaurants from Delhi are maximum, we will focus on Delhi city only. The methodology can be applied on all other cities.

Lets find out restaurant clusters from New Delhi using folium maps. We used k-means for clustering with value of k=7.The clusters are shown in following figure 1.

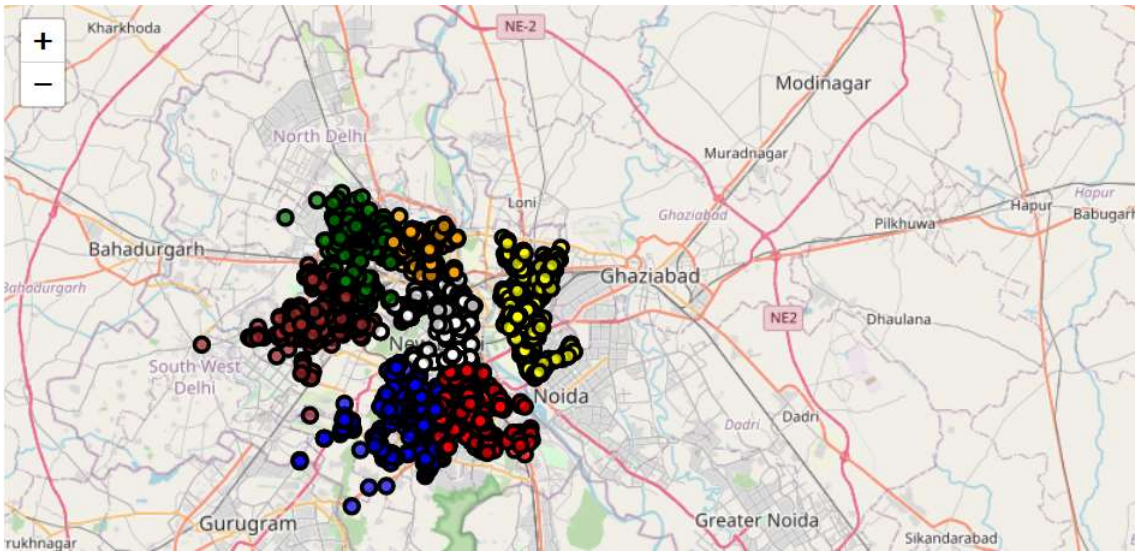
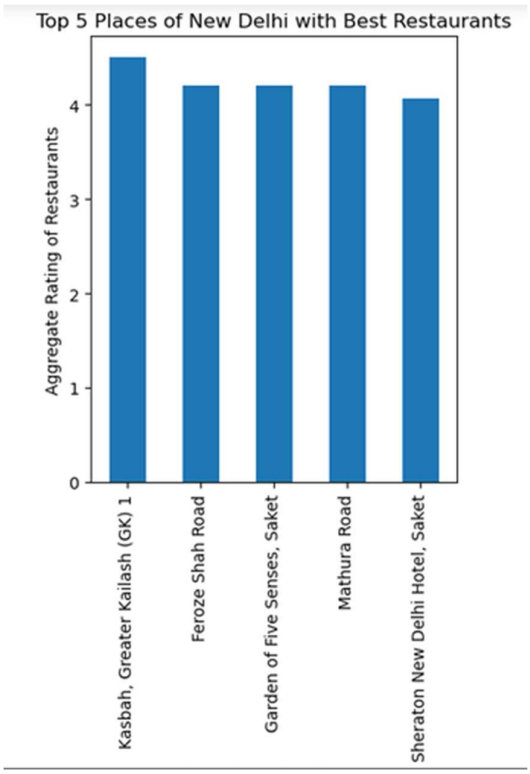
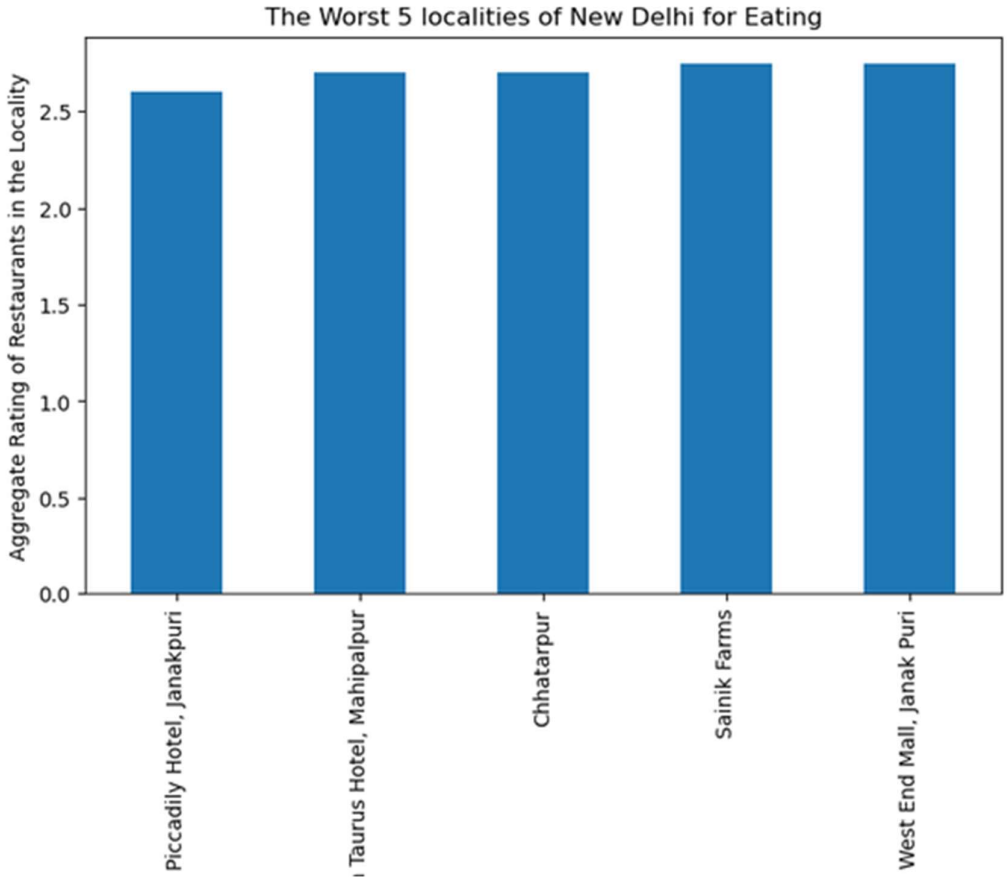


Figure 1: Locality-wise Clusters of Restaurants in New Delhi

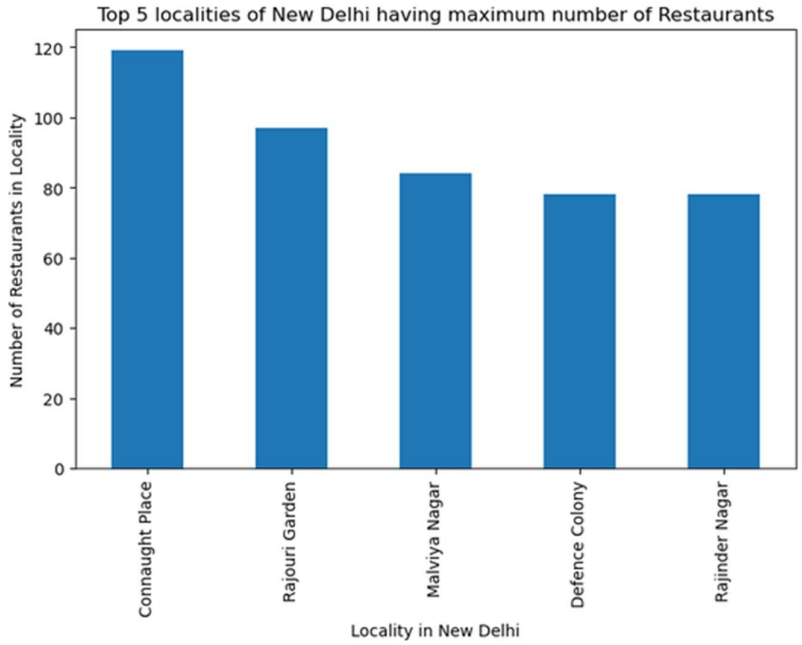
Following graph shows top 5 places of New Delhi which have best restaurants from the city.



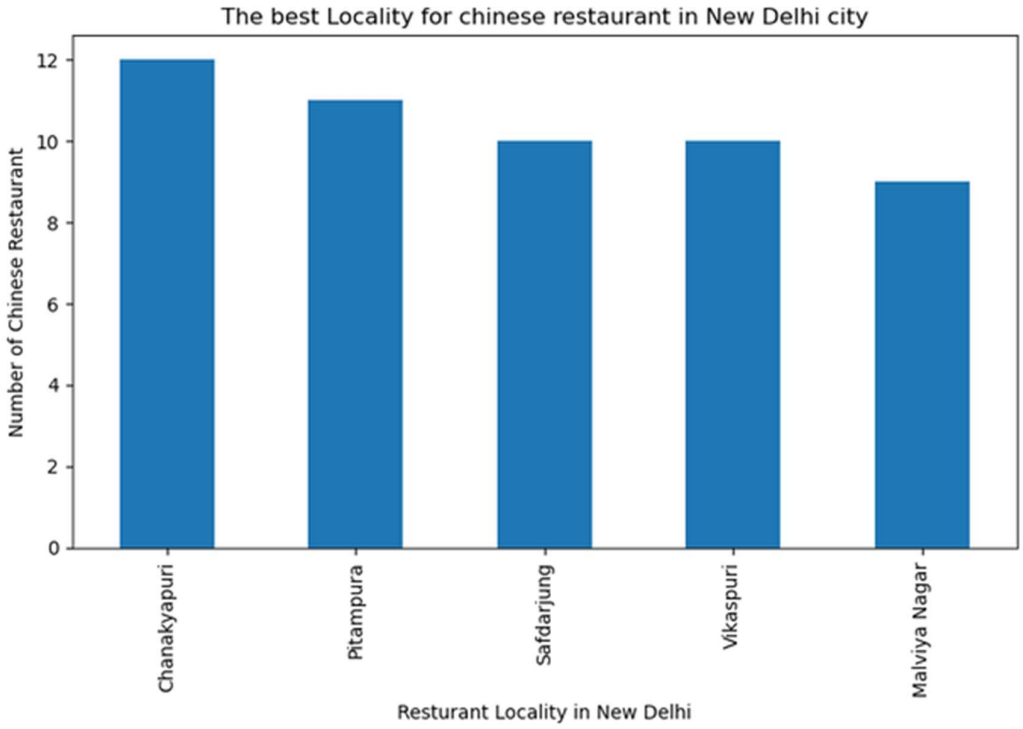
Following graph shows places from New Delhi which have restaurants having lowest average ratings.



The top localities with maximum number of restaurants is shown below.



The best localities for chinses cushion from New Delhi are as follows :



FourSquare API used to correlate location data and Zomato restaurant data. After analysing both datasets together following table was observed. The table shows food speciality of each venue and best place to visit at that place for eating the food.

	Locality	Locality Latitude	Locality Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	ARSS Mall, Paschim Vihar	28.668945	77.101544	Subway	28.669999	77.102546	Sandwich Place
1	ARSS Mall, Paschim Vihar	28.668945	77.101544	Pizza Hut	28.670321	77.103853	Pizza Place
2	ARSS Mall, Paschim Vihar	28.668945	77.101544	Café Coffee Day	28.670009	77.102480	Coffee Shop
3	ARSS Mall, Paschim Vihar	28.668945	77.101544	Baljeet's Amritsari Koolcha	28.665768	77.100481	Indian Restaurant
4	ARSS Mall, Paschim Vihar	28.668945	77.101544	Little Chef	28.670000	77.101459	Chinese Restaurant

5. Conclusions

Following are conclusions of this capstone project activity :

1) Kasbha Greater Kailas is the best place for eating in New Delhi
2) Picadially Hotel, Janakpuri is the worst place for eating in New Delhi.
3) The Connaghut have maximum number of restaurants.
4) Chanakyapuri is the best place for Chinese cushions.
5) In ARSS Mall, Pashim Vihar the Subway venue is best for eating sandwich.