

IBM Applied Data Science Capstone Project on The Battle of Neighborhoods

June, 2020

INTRODUCTION

Being the greatest HubSpot of some of the most exquisite cuisines and lip-smacking local food from every corner of the world, **New Delhi**, the national capital of India, is proud to have some of the most unique and best restaurants within its lap. Any cuisine, Indian or International, name it and they have it. In this project, we will explore the neighbourhood of New Delhi city and find out the best restaurants for certain cuisines.

The **foodies** being the target audience of this project idea, we will seek answers to some of the following questions:

- What places have best restaurants in New Delhi?
- Which areas have lesser number of restaurants in the city?
- Which is the best location for Chinese cuisine in New Delhi?
- Which is the best location to stay if one prefers Chinese cuisine?
- Where can we find large number of Chinese restaurants in the city?

DATA

The following data is used to solve the above-mentioned problem:

Source: <https://www.kaggle.com/shrutimehta/zomato-restaurants-data>

The dataset contains locality, restaurants, cuisines, ratings of the restaurants around the world along with their latitude and longitude. We extract the data of New Delhi using the latitude and longitudes provided.

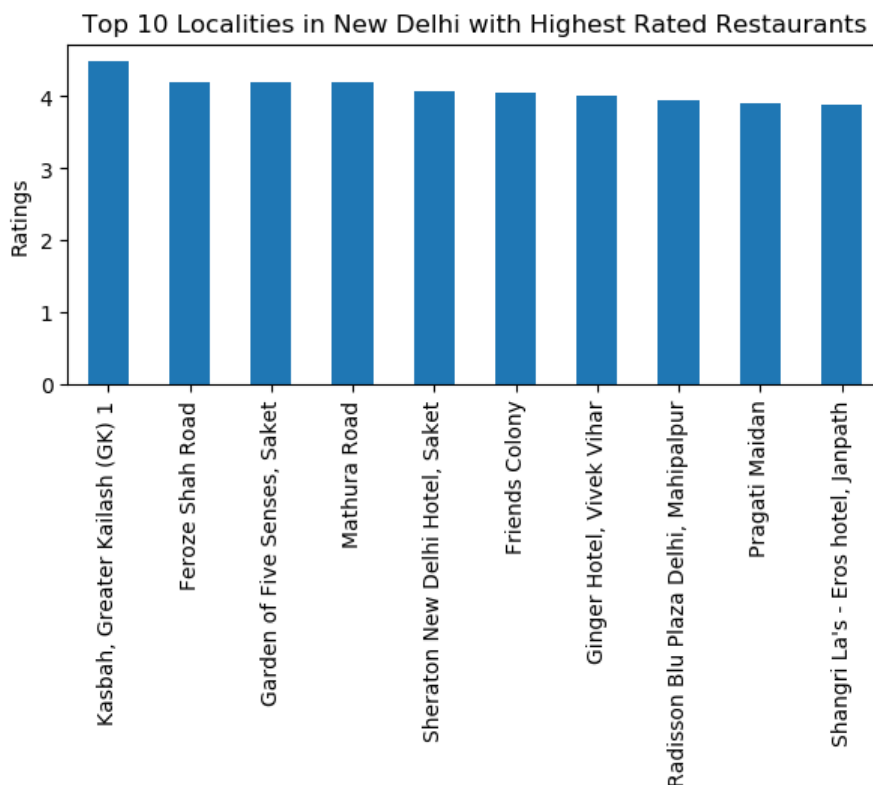
Source: <https://developer.foursquare.com/>

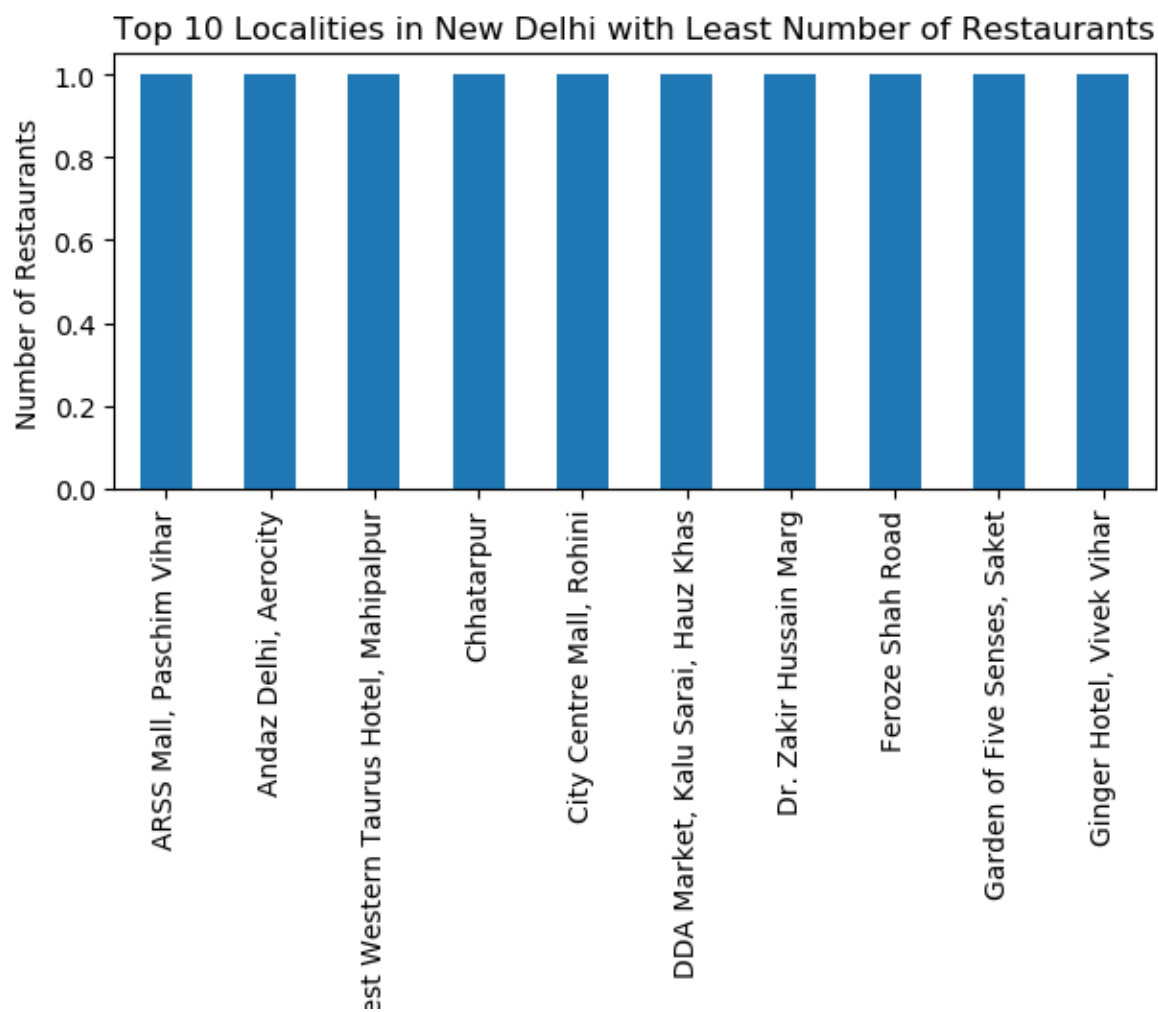
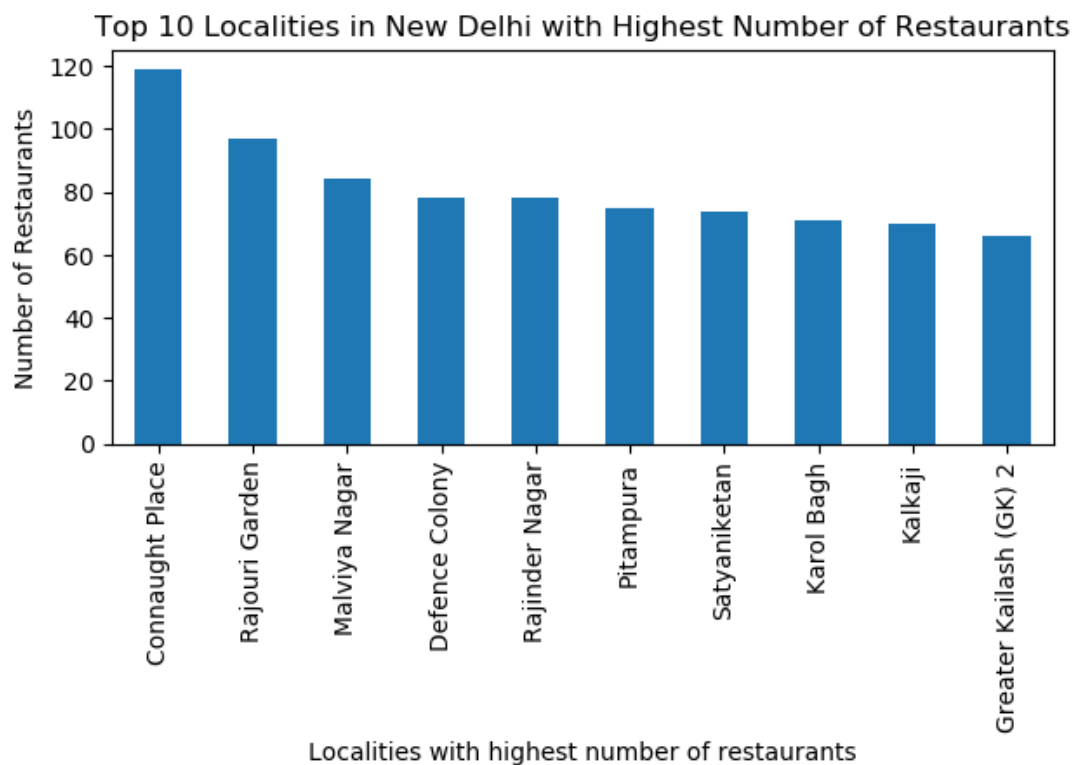
We use the Foursquare API to explore the neighbourhood of New Delhi and find the answers to some obvious question for the audience.

METHODOLOGY

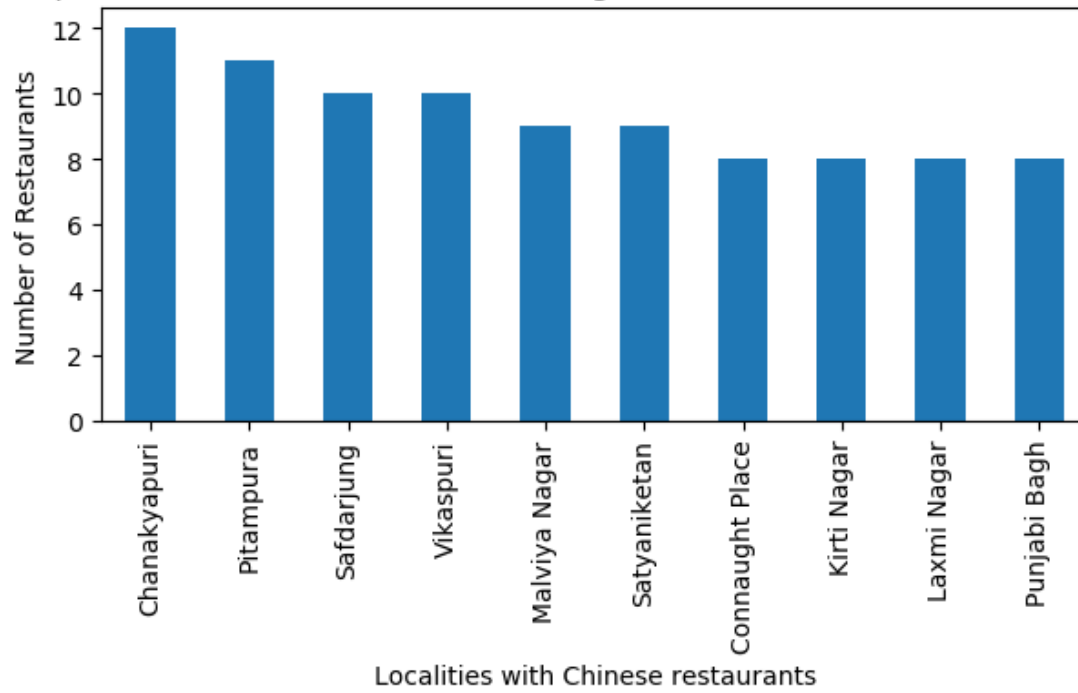
- After collecting the dataset, we extract the data of New Delhi from the Kaggle dataset. We transform the data and clean the given data set by removing the columns which seem to be of less importance.
- When data cleaning is completed, we use the Foursquare API to explore the nearby venues for each neighbourhood in New Delhi.
- We visualize the nearby venues of New Delhi using folium.
- We then filter the venues to find out the best restaurants around the city and the locality with highest as well as lowest number of restaurants.
- We also find the places with highest number of Chinese restaurants and the best places for Chinese cuisine lovers to eat.
- We transform the data by grouping them based on their locality and analyse them in order to cluster the data.
- We then cluster the most common venues in the neighbourhood into five clusters using K-Means Clustering machine learning algorithm and view the clusters formed. These clusters help the foodies to easily locate the places based on common venues and find a suitable place to stay to have delicious food within their reach.

RESULTS

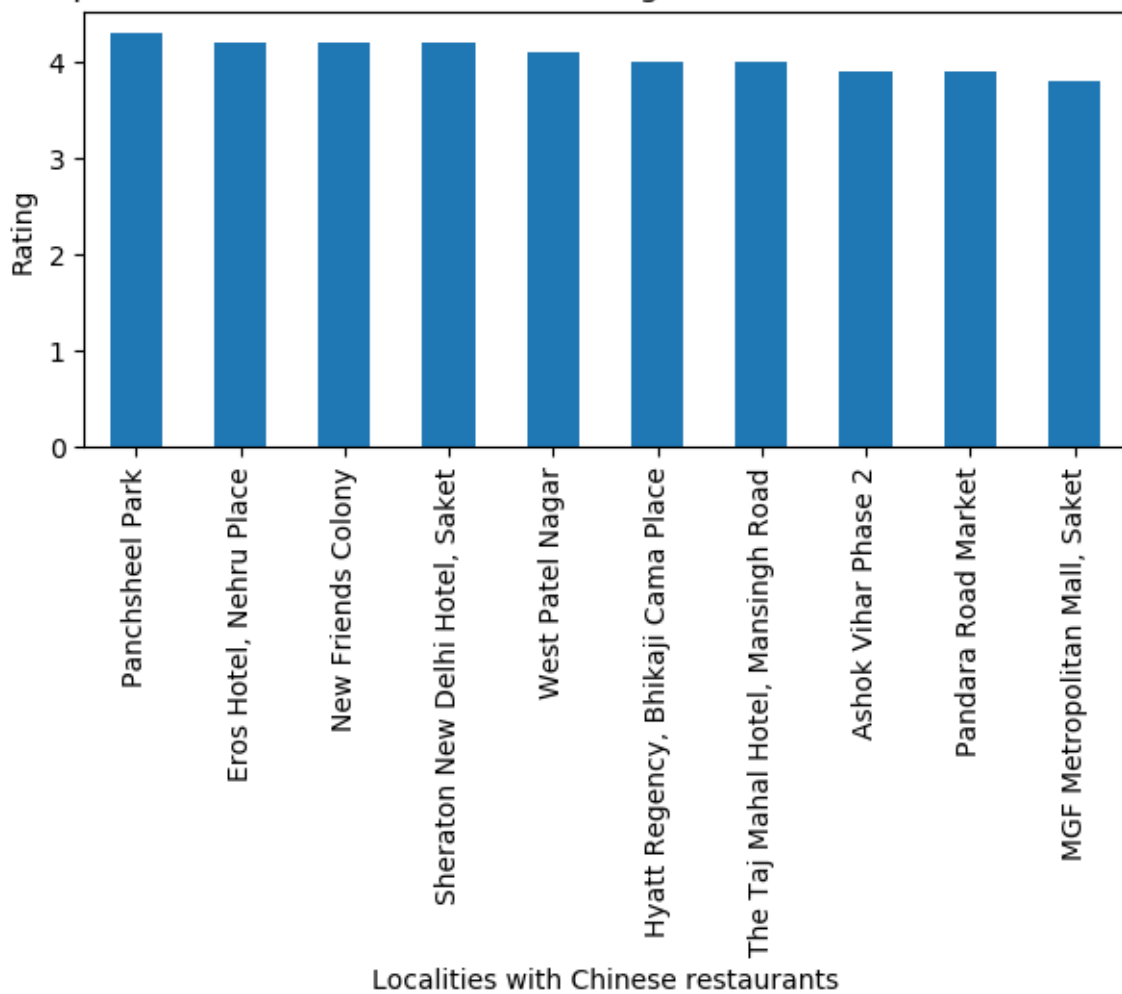


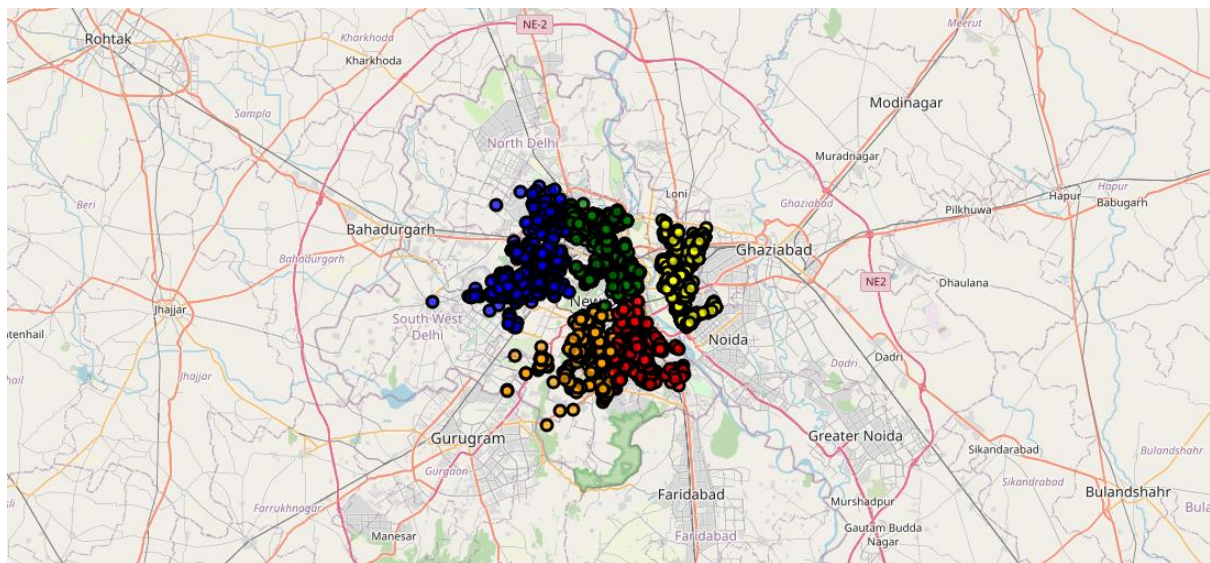


Top 10 Localities in New Delhi with Highest Number of Chinese Restaurants



Top 10 Localities in New Delhi with Highest Rated Chinese Restaurants





DISCUSSION

Based on the results obtained by clustering the data, we recommend:

- Cluster 1 for Indian restaurants
- Cluster 2 for Pizza Place
- Cluster 3 for Coffee Shops and Cafes
- Cluster 4 for Hotels
- Cluster 5 for Cafes

Cluster 1

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

	Lat	Agg_Rating	Comments	No_of_Votes	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	28.668945	3.100000	Average	117	0	Indian Restaurant	ATM	Pizza Place	Multicuisine Indian Restaurant	Coffee Shop
2	28.656131	3.275000	Average, Good	434	0	Indian Restaurant	Pizza Place	Multiplex	Shopping Mall	Hotel
4	28.690020	3.033333	Average	126	0	Department Store	Indian Restaurant	Pizza Place	Chinese Restaurant	Business Service
13	28.690182	3.090000	Average, Good, Poor	971	0	Snack Place	Dessert Shop	BBQ Joint	Market	Smoke Shop
16	28.630440	3.208333	Average, Good, Poor	1779	0	Indian Restaurant	Bakery	Hotel	Theater	Café

Cluster 2

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

]:

	Lat	Agg_Rating	Comments	No_of_Votes	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
5	28.700516	3.040000	Average, Good, Poor	594	1	Gym / Fitness Center	Pizza Place	Multiplex	Sandwich Place	Bus Station
15	28.691136	3.390909	Average, Good, Very Good	485	1	Gym / Fitness Center	Pizza Place	South Indian Restaurant	Fast Food Restaurant	Women's Store
22	28.654105	3.085714	Average	80	1	Pizza Place	Gym	Fast Food Restaurant	Frozen Yogurt Shop	French Restaurant
57	28.699837	3.376471	Average, Good, Very Good	2460	1	Pizza Place	Food Truck	Bike Rental / Bike Share	Women's Store	Falafel Restaurant
94	28.538153	3.117143	Average, Good, Poor	3642	1	Market	Fast Food Restaurant	Indian Restaurant	Pizza Place	Dessert Shop
98	28.651527	3.490909	Average, Good, Very Good	3222	1	Pizza Place	Food Truck	Train Station	Women's Store	Event Space

Cluster 3

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

]:

	Lat	Agg_Rating	Comments	No_of_Votes	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
6	28.527088	3.117391	Average, Good, Poor	1012	2	New American Restaurant	Coffee Shop	Steakhouse	Market	Worship Center
7	28.541298	3.425000	Average, Good, Very Good	2460	2	Coffee Shop	Shopping Mall	Clothing Store	Italian Restaurant	Fast Restaurant
8	28.555599	3.800000	Average, Very Good	2411	2	Café	Stadium	Dessert Shop	Hotel	Coffee Shop
11	28.562580	3.750000	Average, Good, Very Good	875	2	Performing Arts Venue	Department Store	Karaoke Bar	Fast Food Restaurant	Worship Center
14	28.693969	3.378571	Average, Good	2343	2	Indian Restaurant	Department Store	Pizza Place	South Indian Restaurant	Coffee Shop

Cluster 4

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

]:

	Lat	Agg_Rating	Comments	No_of_Votes	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
3	28.553077	3.200000	Average	59	3	Hotel	Airport Lounge	Coffee Shop	Airport Terminal	Fast Food Restaurant
10	28.554807	3.000000	Average	4	3	Hotel	Spa	Coffee Shop	Restaurant	Hotel Pool
12	28.641698	3.050000	Average	17	3	Hostel	Movie Theater	Indian Restaurant	Hotel	Asian Restaurant
19	28.548827	2.700000	Average	9	3	Hotel	Indian Restaurant	Accessories Store	Italian Restaurant	Shoe Store
20	28.568193	2.755556	Average, Poor	339	3	Hotel	Lounge	Asian Restaurant	Bakery	Café
25	28.495624	2.700000	Average	23	3	Indian Restaurant	Hotel	Donut Shop	Luggage Store	Japanese Restaurant
35	28.590297	3.350000	Average, Good	127	3	Hotel	Indian Restaurant	American Restaurant	Pizza Place	Restaurant
56	28.570246	4.050000	Average, Very Good	1040	3	Hotel	Café	Sculpture	Neighborhood	Gym

Cluster 5

```
ND_merged.loc[ND_merged['Cluster Labels'] == 4, ND_merged.columns[[1] + list(range(5, ND_merged.shape[1]))]]
]:
```

	Lat	Agg_Rating	Comments	No_of_Votes	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6
1	28.537063	3.292308	Average, Good, Poor, Very Good	1560	4	Indian Restaurant	Café	Pub	Restaurant	Dim Sum Restaurant	
9	28.659870	3.281481	Average, Good	1309	4	Café	Burrito Place	Burger Joint	Convenience Store	Diner	D
33	28.657388	3.320000	Average, Good	636	4	Café	Indian Restaurant	Pizza Place	Hotel	Burger Joint	
82	28.548371	3.137500	Average, Good, Poor	609	4	Indian Restaurant	Shopping Mall	Scenic Lookout	Café	Falafel Restaurant	
86	28.624632	3.315217	Average, Excellent, Good, Poor, Very Good	4011	4	Café	Park	Bakery	Gym	Falafel Restaurant	Res
89	28.539003	3.150000	Average, Good, Poor	893	4	Café	Pizza Place	Breakfast Spot	Soccer Stadium	Dessert Shop	Fa Res
100	28.556869	4.500000	Excellent	524	4	Café	Indian	Furniture / Home	Lounge	Pool	C

CONCLUSION

Based on the results, we conclude that best restaurants can be found in Kasbah, Greater Kailash. Connaught Place has highest number of restaurants while ARSS Mall, Pashchim Vihar has the least. Chanakyapuri has a lot of Chinese restaurants but Panchsheel park has the best Chinese restaurants. Upon clustering using K-Means, we found the most common recommended venues for each cluster which will help the foodies to choose the best spot to reside in based on their interests.