

1) Introduction/Business Problem

New Delhi is the capital city of India. It is a part of the city of Delhi's 11 districts. The city itself has a population of 257,803. However, the much larger metro area has a population that exceeds 26 million.

New Delhi are used interchangeably to refer to the National Capital Territory of Delhi (NCT), these are two distinct entities, with New Delhi forming a small part of Delhi. The National Capital Region is a much larger entity comprising the entire NCT along with adjoining districts in neighboring states.

The official language of New Delhi and the one that is most widely spoken is Hindi. However, English is also spoken as a formal language within businesses and government agencies. Over last decades it is continuously grow because of the city's important role in government and commercial business.

With it's diverse culture , comes diverse food items. There are many restaurants in New Delhi City, each belonging to different categories like Chinese , Italian , French etc. So as part of this project , we will list and visualise all major parts of New Delhi City . Questions that can be asked using the above mentioned datasets What is best location in New Delhi City for Chinese Cuisine ?\ Which areas have large number of Chinese Resturant Market ?\ Which all areas have less number of resturant ?

2) Data

For this project we need the following data :

New Delhi Resturants data that contains list Locality, Resturant name,Rating along with their latitude and longitude. Data source : Zomato kaggel dataset Description : This data set contains the required information. And we will use this data set to explore various locality of new delhi city. Nearby places in each locality of new delhi city.

3) Approach

Collect the new delhi city data from Zomato kaggel dataset Using FourSquare API we will find all venues for each neighborhood. Filter out all venues that are nearby by locality. Using aggregative rating for each resturant to find the best places. Visualize the Ranking of neighborhoods using folium library(python)

```
In [7]: import pandas as pd
import numpy as np
import requests # library to handle requests
from pandas.io.json import json_normalize # tranform JSON file into a pandas dataframe
# Matplotlib and associated plotting modules
import matplotlib.cm as cm
import matplotlib.colors as colors
# import k-means from clustering stage
from sklearn.cluster import KMeans

!conda install -c conda-forge folium=0.5.0 --yes
import folium # map rendering library
! pip install geocoder
import geocoder
```

Collecting package metadata (current_repodata.json): ...working... done
Solving environment: ...working... done

All requested packages already installed.

Requirement already satisfied: geocoder in c:\users\jinda\anaconda3\lib\site-packages (1.38.1)
Requirement already satisfied: requests in c:\users\jinda\anaconda3\lib\site-packages (from geocoder) (2.22.0)
Requirement already satisfied: six in c:\users\jinda\anaconda3\lib\site-packages (from geocoder) (1.14.0)
Requirement already satisfied: future in c:\users\jinda\anaconda3\lib\site-packages (from geocoder) (0.18.2)
Requirement already satisfied: click in c:\users\jinda\anaconda3\lib\site-packages (from geocoder) (7.0)
Requirement already satisfied: ratelim in c:\users\jinda\anaconda3\lib\site-packages (from geocoder) (0.1.6)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in c:\users\jinda\anaconda3\lib\site-packages (from requests->geocoder) (1.25.8)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\jinda\anaconda3\lib\site-packages (from requests->geocoder) (2019.11.28)
Requirement already satisfied: chardet<3.1.0,>=3.0.2 in c:\users\jinda\anaconda3\lib\site-packages (from requests->geocoder) (3.0.4)
Requirement already satisfied: idna<2.9,>=2.5 in c:\users\jinda\anaconda3\lib\site-packages (from requests->geocoder) (2.8)
Requirement already satisfied: decorator in c:\users\jinda\anaconda3\lib\site-packages (from ratelim->geocoder) (4.4.1)

```
In [27]: import pandas as pd
df = pd.read_csv('https://raw.githubusercontent.com/ashishjindal94/Battle-of-neighbourhoods/master/zomato.csv', encoding='ISO-8859-1')
df.head()
```

Out[27]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude
0	6317637	Le Petit Souffle	162	Makati City	Third Floor, Century City Mall, Kalayaan Avenu...	Century City Mall, Poblacion, Makati City	Century City Mall, Poblacion, Makati City, Mak...	121.027535	14.56544
1	6304287	Izakaya Kikufuji	162	Makati City	Little Tokyo, 2277 Chino Roces Avenue, Legaspi...	Little Tokyo, Legaspi Village, Makati City	Little Tokyo, Legaspi Village, Makati City, Ma...	121.014101	14.55370
2	6300002	Heat - Edsa Shangri-La	162	Mandaluyong City	Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal...	Edsa Shangri-La, Ortigas, Mandaluyong City	Edsa Shangri-La, Ortigas, Mandaluyong City, Ma...	121.056831	14.58140
3	6318506	Ooma	162	Mandaluyong City	Third Floor, Mega Fashion Hall, SM Megamall, O...	SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City, Mandal...	121.056475	14.58531
4	6314302	Sambo Kojin	162	Mandaluyong City	Third Floor, Mega Atrium, SM Megamall, Ortigas...	SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City, Mandal...	121.057508	14.58445

5 rows × 21 columns

```
In [28]: df_india = df[df['Country Code'] == 1]
df_NDLS = df_india[df_india['City'] == 'New Delhi']
df_NDLS.reset_index(drop=True, inplace=True)
df_NDLS.head()
```

Out[28]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude	Cuisines	
0	18287358	Food Cloud	1	New Delhi	Aaya Nagar, New Delhi	Aaya Nagar	Aaya Nagar, New Delhi	0.000000	0.000000	Cuisine Varies	
1	18216944	Burger.in	1	New Delhi	84, Near Honda Showroom, Adchini, New Delhi	Adchini	Adchini, New Delhi	77.196923	28.535382	Fast Food	
2	313333	Days of the Raj	1	New Delhi	81/3, 1st Floor, Qutub Residency, Adchini, New...	Adchini	Adchini, New Delhi	77.197475	28.535493	North Indian, Seafood, Continental	
3	18384127	Dilli Ka Dhaba	1	New Delhi	66 A, Ground Floor, Sri Aurobindo Marg, Adchin...	Adchini	Adchini, New Delhi	77.198033	28.537547	South Indian, North Indian	
4	582	Govardhan	1	New Delhi	84, Adjacent Hero Motor Bike Showroom, Main Me...	Adchini	Adchini, New Delhi	77.196924	28.535523	South Indian, North Indian, Chinese	

5 rows × 21 columns

Data Cleaning\ remove the unwanted columns and rows from dataset

```
In [29]: df_Res= df_NDLS[df_NDLS.Longitude !=0.000000][['Restaurant Name','Locality','Longitude',
'Latitude','Cuisines','Aggregate rating','Rating text','Votes']]
```

```
In [30]: df_Res = df_Res[df_Res['Aggregate rating'] !=0.0]

df_Res.head()
```

Out[30]:

	Restaurant Name	Locality	Longitude	Latitude	Cuisines	Aggregate rating	Rating text	Votes
1	Burger.in	Adchini	77.196923	28.535382	Fast Food	3.2	Average	46
2	Days of the Raj	Adchini	77.197475	28.535493	North Indian, Seafood, Continental	3.4	Average	45
3	Dilli Ka Dhaba	Adchini	77.198033	28.537547	South Indian, North Indian	2.6	Average	11
4	Govardhan	Adchini	77.196924	28.535523	South Indian, North Indian, Chinese	3.4	Average	238
5	Mezbaan Grills	Adchini	77.198122	28.538134	Mughlai	3.1	Average	8

created map to show the restaurant clusters

```
In [32]: New_Delhi_Rest = folium.Map(location=[28.52, 77.25], zoom_start=12)
X = df_Res['Latitude']
Y = df_Res['Longitude']
Z = np.stack((X, Y), axis=1)

kmeans = KMeans(n_clusters=5, random_state=0).fit(Z)

clusters = kmeans.labels_
colors = ['red', 'green', 'blue', 'yellow', 'orange']
df_Res['Cluster'] = clusters

for latitude, longitude, Locality, cluster in zip(df_Res['Latitude'], df_Res['Longitude'], df_Res['Locality'], df_Res['Cluster']):
    label = folium.Popup(Locality, parse_html=True)
    folium.CircleMarker(
        [latitude, longitude],
        radius=5,
        popup=label,
        color='black',
        fill=True,
        fill_color=colors[cluster],
        fill_opacity=0.7).add_to(New_Delhi_Rest)

New_Delhi_Rest
```

Out[32]: Make this Notebook Trusted to load map: File -> Trust Notebook

In [33]: df_Res.head()

Out[33]:

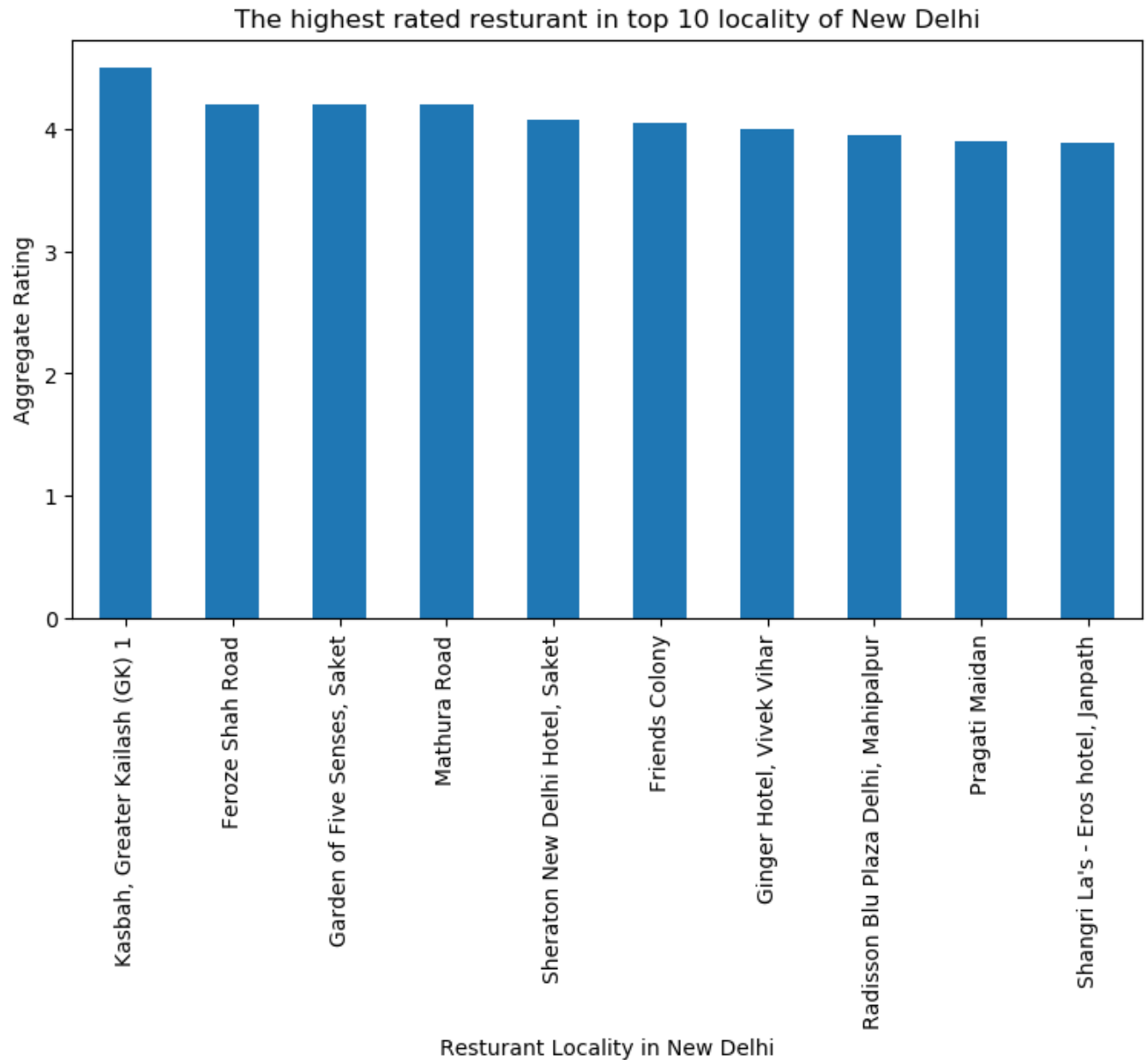
	Restaurant Name	Locality	Longitude	Latitude	Cuisines	Aggregate rating	Rating text	Votes	Cluster
1	Burger.in	Adchini	77.196923	28.535382	Fast Food	3.2	Average	46	1
2	Days of the Raj	Adchini	77.197475	28.535493	North Indian, Seafood, Continental	3.4	Average	45	1
3	Dilli Ka Dhaba	Adchini	77.198033	28.537547	South Indian, North Indian	2.6	Average	11	1
4	Govardhan	Adchini	77.196924	28.535523	South Indian, North Indian, Chinese	3.4	Average	238	1
5	Mezbaan Grills	Adchini	77.198122	28.538134	Mughlai	3.1	Average	8	1

What places have best restaurant in New Delhi?

```
In [34]: import matplotlib.pyplot as plt
plt.figure(figsize=(9,5), dpi = 100)
# title
plt.title('The highest rated resturant in top 10 locality of New Delhi')
#On x-axis

#giving a bar plot
df_Res.groupby('Locality')['Aggregate rating'].mean().nlargest(10).plot(kind='bar')

plt.xlabel('Resturant Locality in New Delhi')
#On y-axis
plt.ylabel('Aggregate Rating')
#displays the plot
plt.show()
```

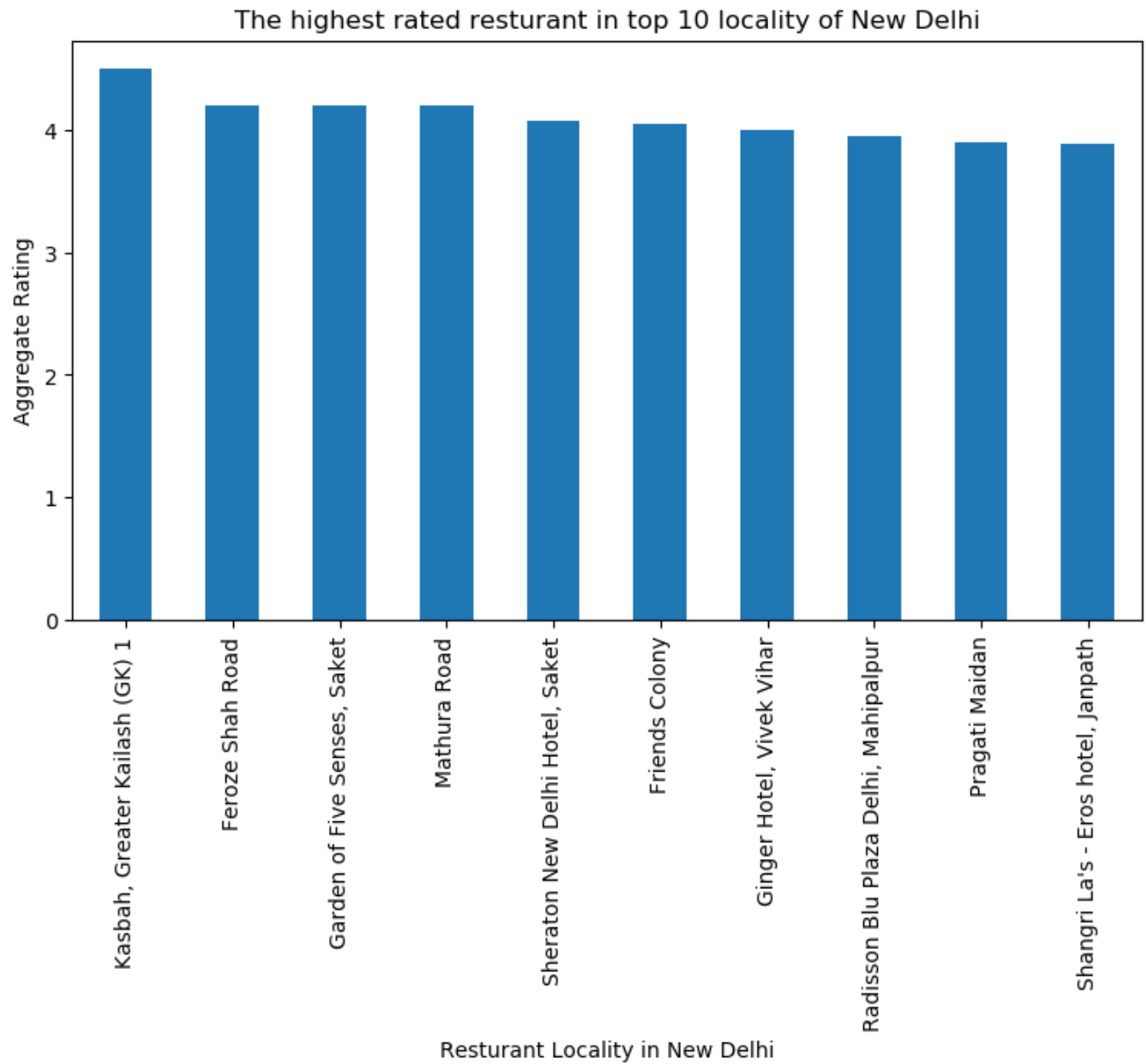


what places have worst restaurants in New Delhi?


```
In [35]: import matplotlib.pyplot as plt
plt.figure(figsize=(9,5), dpi = 100)
# title
plt.title('The highest rated resturant in top 10 locality of New Delhi')
#On x-axis

#giving a bar plot
df_Res.groupby('Locality')['Aggregate rating'].mean().nlargest(10).plot(kind='bar')

plt.xlabel('Resturant Locality in New Delhi')
#On y-axis
plt.ylabel('Aggregate Rating')
#displays the plot
plt.show()
```



what places have worst restaurants in New Delhi?

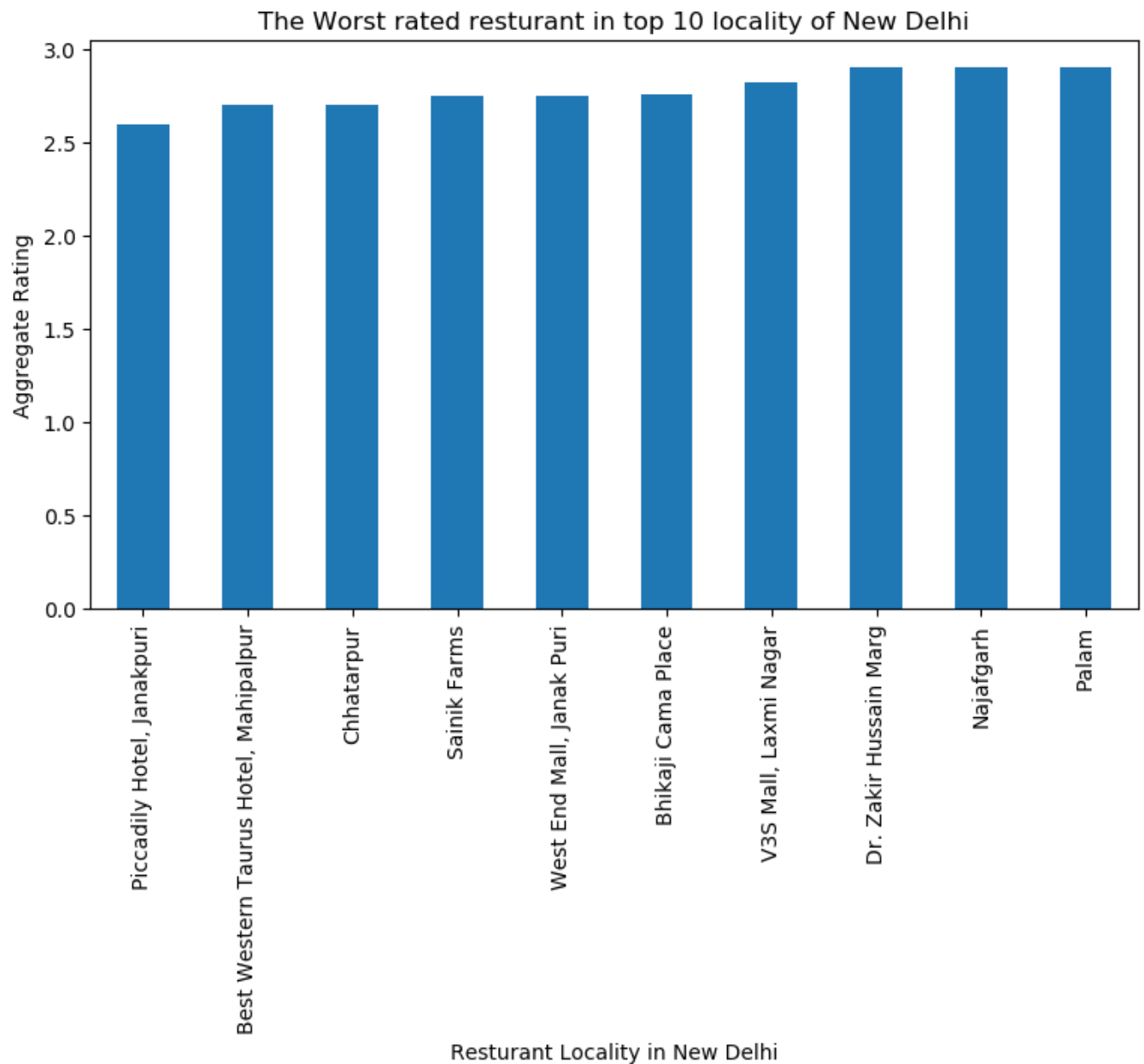
```
In [36]: import matplotlib.pyplot as plt
plt.figure(figsize=(9,5), dpi = 100)
# title
plt.title('The Worst rated resturant in top 10 locality of New Delhi')
#On x-axis

#giving a bar plot

df_Res.groupby('Locality')['Aggregate rating'].mean().nsmallest(10).plot(kind='bar')

plt.xlabel('Resturant Locality in New Delhi')
#On y-axis
plt.ylabel('Aggregate Rating')

#displays the plot
plt.show()
```



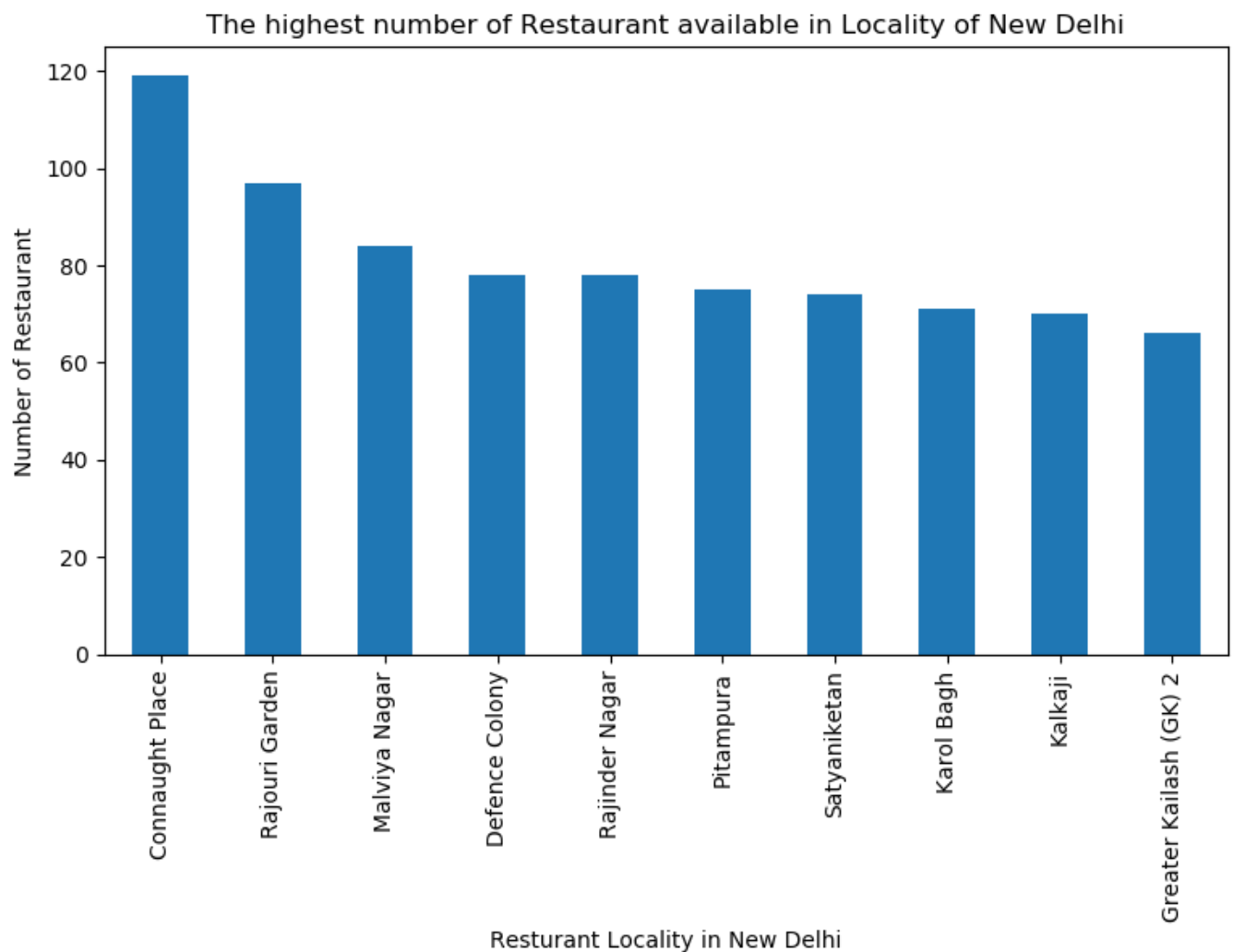
Which place are suitable for edible person in New Delhi city?

```
In [37]: import matplotlib.pyplot as plt
plt.figure(figsize=(9,5), dpi = 100)
# title
plt.title('The highest number of Restaurant available in Locality of New Delhi')
#On x-axis

#giving a bar plot
df_Res.groupby('Locality')['Restaurant Name'].count().nlargest(10).plot(kind='bar')

plt.xlabel('Resturant Locality in New Delhi')
#On y-axis
plt.ylabel('Number of Restaurant')

#displays the plot
plt.show()
```



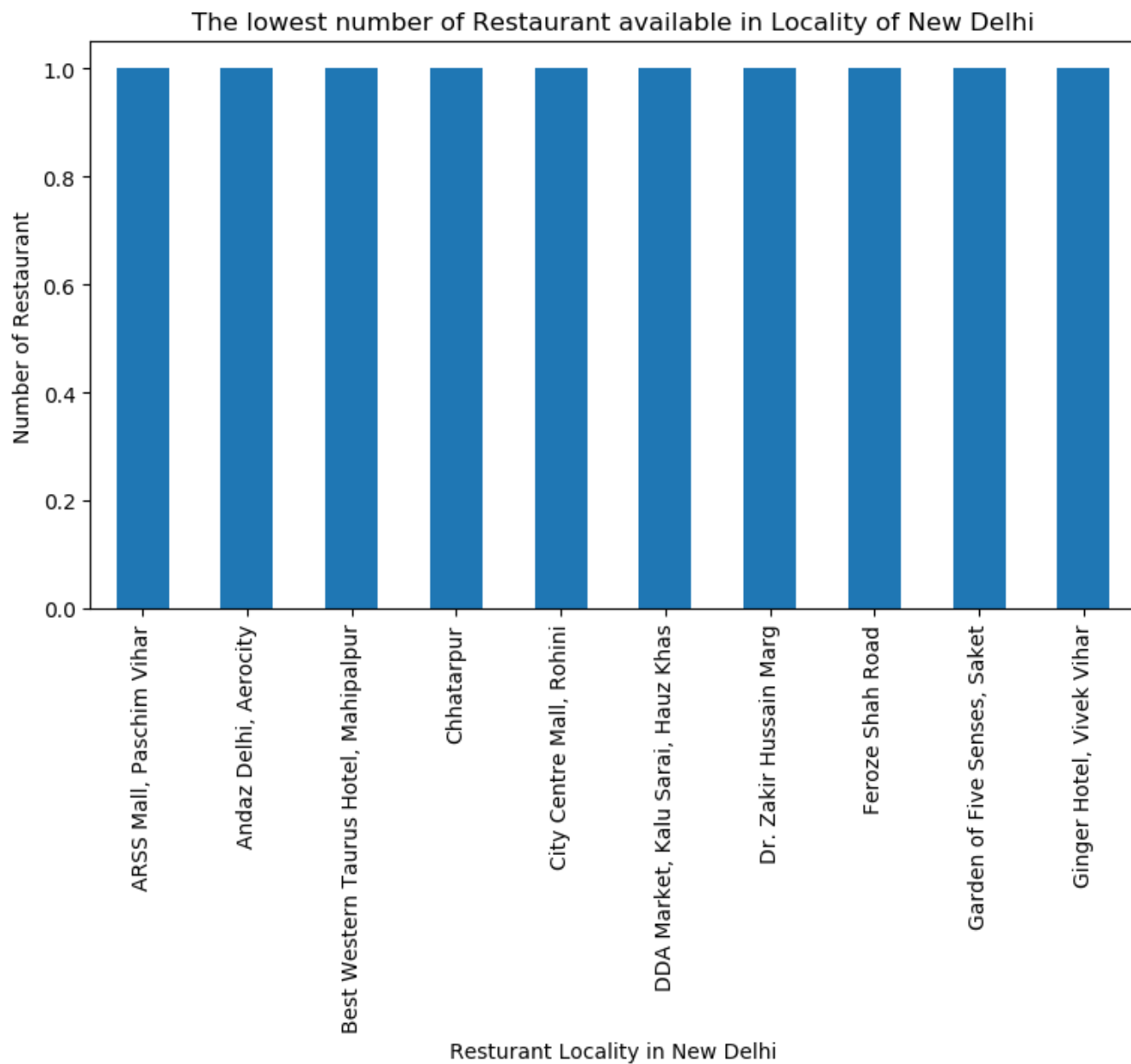
Which place are not suitable for edible person in New Delhi city?

```
In [40]: plt.figure(figsize=(9,5), dpi = 100)
# title
plt.title('The lowest number of Restaurant available in Locality of New Delhi')
#On x-axis

#giving a bar plot
df_Res.groupby('Locality')['Restaurant Name'].count().nsmallest(10).plot(kind='bar')

plt.xlabel('Resturant Locality in New Delhi')
#On y-axis
plt.ylabel('Number of Restaurant')

#displays the plot
plt.show()
```



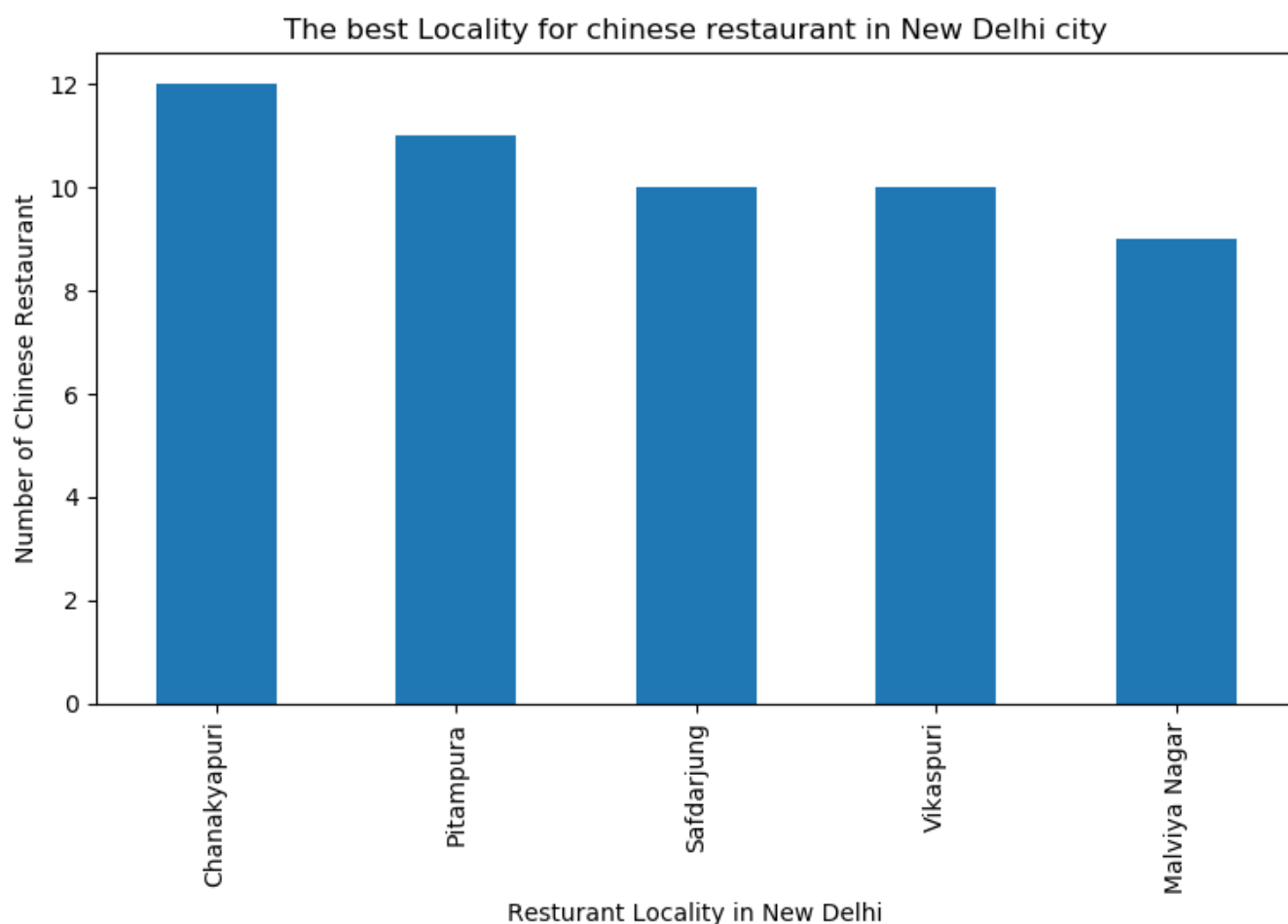
What are the best places for chinese restaurant in New Delhi city?

```
In [41]: import matplotlib.pyplot as plt
plt.figure(figsize=(9,5), dpi = 100)
# title
plt.title('The best Locality for chinese restaurant in New Delhi city')
#On x-axis

#giving a bar plot
df_Res[df_Res['Cuisines'].str.startswith('Chinese')].groupby('Locality')['Restaurant Name'].count().nlargest(5).plot(kind='bar')

plt.xlabel('Resturant Locality in New Delhi')
#On y-axis
plt.ylabel('Number of Chinese Restaurant')

#displays the plot
plt.show()
```



Data transformation\ Based on Locality grouping the data

```
In [49]: df_Res_Loc = df_Res.groupby('Locality').count()['Restaurant Name'].to_frame()

df_Res_Loc.head()
```

Out[49]:

Locality	Restaurant Name
ARSS Mall, Paschim Vihar	1
Adchini	13
Aditya Mega Mall, Karkardooma	4
Aerocity	2
Aggarwal City Mall, Pitampura	3

```
In [51]: d_Lat = df_Res.groupby('Locality').mean()['Latitude'].to_frame()

d_Lat.head()
```

Out[51]:

Locality	Latitude
ARSS Mall, Paschim Vihar	28.668945
Adchini	28.537063
Aditya Mega Mall, Karkardooma	28.656131
Aerocity	28.553077
Aggarwal City Mall, Pitampura	28.690020

```
In [52]: df_Res_Loc = df_Res.groupby('Locality').count()['Restaurant Name'].to_frame()
df_Res_rating= df_Res.groupby('Locality')['Aggregate rating'].mean().to_frame()
d_Cuisines = df_Res.groupby(['Locality'])['Cuisines'].agg(', '.join).reset_index()
d_R = df_Res.groupby(['Locality'])['Rating text'].unique().agg(', '.join).reset_index()
d_V = df_Res.groupby(['Locality'])['Votes'].sum().to_frame()
d_Lat = df_Res.groupby('Locality').mean()['Latitude'].to_frame()
d_Lng = df_Res.groupby('Locality').mean()['Longitude'].to_frame()
df_final = pd.merge(d_Lat,d_Lng,on='Locality').merge(df_Res_Loc, on='Locality').merge(d_Cuisines, on='Locality').merge(df_Res_rating,on ='Locality').merge(d_R, on ='Locality').merge(d_V, on ='Locality')
```

```
In [53]: df_final = df_final[df_final['AggregateRating'] != 0.000000]
df_final.columns = ['Locality', 'Lat', 'Lng', 'No_of_Restaurant', 'Cusines', 'Agg_Rating', 'Comments', 'No_of_Votes']
df_final.head()
```

Out[53]:

	Locality	Lat	Lng	No_of_Restaurant	Cusines	Agg_Rating	Comments	No_of_Votes
0	ARSS Mall, Paschim Vihar	28.668945	77.101544	1	North Indian, South Indian, Chinese, Mithai, F...	3.100000	Average	117
1	Adchini	28.537063	77.197808	13	Fast Food, North Indian, Seafood, Continental,...	3.292308	Average, Good, Poor, Very Good	1560
2	Aditya Mega Mall, Karkardooma	28.656131	77.301266	4	Finger Food, North Indian, Mughlai, Pizza, Fas...	3.275000	Average, Good	434
3	Aerocity	28.553077	77.104270	2	Fast Food, Italian, Pizza, North Indian, Conti...	3.200000	Average	59
4	Aggarwal City Mall, Pitampura	28.690020	77.134650	3	North Indian, Chinese, Street Food, Mithai, No...	3.033333	Average	126

```
In [79]: df_final.shape
```

Out[79]: (240, 8)

Conclusion\ Chanakyapuri, Pitampura, Safdarjung are some of the best neighborhoods for Chinese cuisine.\ Pancsheel park, Nehru place have the best Chinese Resturant.\ Cannaught place, Rajouri garden, Malviya nagar are the best places for edible person.

```
In [ ]:
```