The Battle of Neighbourhoods- Noida
Alvin philip Cheriyan
07/11/2020

1 Introduction

Noida, short for **New Okhla Industrial Development Authority**, is a planned city located in Gautam Buddh Nagar district of the Indian state of Uttar Pradesh. It is a satellite city of Delhi and is a part of the National Capital Region of India. As per provisional reports of Census of India, the population of Noida in 2011 was 642,381. The official language of Noida and the one that is most widely spoken is Hindi. With its diverse culture, comes diverse food items. There are many restaurants in Noida, 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 Noida.

Questions that can be asked using the above-mentioned datasets

- Which places In Noida have some of the finest Restaurants?
- Which Places in Noida have Restaurants with Lowest rating?
- Which place in Noida is suitable for a Foodie?
- Which place in Noida is not suitable for a Foodie?
- What are the best places in Noida for Chinese Restaurant?
- Which places in Noida have highest rated Chinese Restaurants?

2 Data

For this project we need the following data:

- Noida Restaurants data that contains list Locality, Restaurant name, Rating along with their latitude and longitude.

    - Data Source: zomatoKaggle
    - Description: This data set contains the required information which we will use to explore various localities of Noida city.
- Nearby places in each locality of Noida city.
    - Data Source: FourSquareAPI
    - Description: By using this API we will get all the venues in each neighbourhood.

3 Approach

- Collect the Noida city data from zomatoKaggle
- Using Foursquare API we will find all venues for each neighbourhoods.
- Filter out all venues that are nearby by locality.
- Using aggregative rating for each restaurant to find the best places.
- Visualize the Ranking of neighbourhoods using folium library(python)

In [54]:
```
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 p
andas 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\alvin\anaconda3\lib
```

```
Requirement already satisfied: geocoder in c:\users\alvin\anaconda3\lib
\site-packages (1.38.1)
Requirement already satisfied: future in c:\users\alvin\anaconda3\lib\s
ite-packages (from geocoder) (0.18.2)
Requirement already satisfied: requests in c:\users\alvin\anaconda3\lib
\site-packages (from geocoder) (2.24.0)
Requirement already satisfied: ratelim in c:\users\alvin\anaconda3\lib
\site-packages (from geocoder) (0.1.6)
Requirement already satisfied: click in c:\users\alvin\anaconda3\lib\si
te-packages (from geocoder) (7.1.2)
Requirement already satisfied: six in c:\users\alvin\anaconda3\lib\site
-packages (from geocoder) (1.15.0)
Requirement already satisfied: idna<3,>=2.5 in c:\users\alvin\anaconda3
\lib\site-packages (from requests->geocoder) (2.10)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\alvin\ana
conda3\lib\site-packages (from requests->geocoder) (2020.6.20)
Requirement already satisfied: chardet<4,>=3.0.2 in c:\users\alvin\anac
onda3\lib\site-packages (from requests->geocoder) (3.0.4)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1
in c:\users\alvin\anaconda3\lib\site-packages (from requests->geocoder)
(1.25.9)
Requirement already satisfied: decorator in c:\users\alvin\anaconda3\li
b\site-packages (from ratelim->geocoder) (4.4.2)
```

3.1 Read the zomato resturant data from csv file

In [57]:
```
df = pd.read_csv('https://raw.githubusercontent.com/Alvy001/PeerGradedA
ssignment-Battle-of-Neighbourhood/main/zomato.csv',encoding='ISO-8859-
1')
df.head()
```

Out[57]:

Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitud

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitud
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...	121.02753
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...	121.01410
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...	121.05683
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...	121.05647
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...	121.05750

5 rows × 21 columns

In [58]:
```
df_india = df[df['Country Code'] == 1]
df_Nda = df_india[df_india['City'] == 'Noida']
df_Nda.reset_index(drop=True, inplace=True)
df_Nda.head()
```

Out[58]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude
0	310063	Ali Baba Caves	1	Noida	1st Floor, Ansal Plaza Mall, Pari Chowk, Great...	Ansal Plaza Mall, Greater Noida	Ansal Plaza Mall, Greater Noida, Noida	77.528129	28.458107
1	312214	Cafí© Doo Ghoont	1	Noida	AG-13, Atrium Floor, Ansal Plaza Mall, Greater...	Ansal Plaza Mall, Greater Noida	Ansal Plaza Mall, Greater Noida, Noida	77.507456	28.463957
2	309641	Knights Chaska	1	Noida	201 to 207, 2nd Floor, Ansal Plaza Mall, Great...	Ansal Plaza Mall, Greater Noida	Ansal Plaza Mall, Greater Noida, Noida	77.507834	28.464434
3	306688	Thirsty Scholar Cafe	1	Noida	SF-256, 2nd Floor, Ansal Plaza Mall, Near Pari...	Ansal Plaza Mall, Greater Noida	Ansal Plaza Mall, Greater Noida, Noida	77.507701	28.464201

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude	
4	18440427	Savoury Street	1	Noida	GF-37, 1st Floor, Ansal Plaza Mall, Greater No...	Ansal Plaza Mall, Greater Noida	Ansal Plaza Mall, Greater Noida, Noida	77.507701	28.464201	Co

5 rows × 21 columns

3.1.1 Data Cleaning

Remove the unwanted columns and rows from dataset

In [59]:
```
df_Res= df_Nda[df_Nda.Longitude !=0.000000][['Restaurant Name','Locality','Longitude','Latitude','Cuisines','Aggregate rating','Rating text','Votes']]
```

In [60]:
```
df_Res = df_Res[df_Res['Aggregate rating'] !=0.0]
```

In [61]:
```
df_Res.head()
```

Out[61]:

	Restaurant Name	Locality	Longitude	Latitude	Cuisines	Aggregate rating	Rating text	Votes
0	Ali Baba Caves	Ansal Plaza Mall, Greater Noida	77.528129	28.458107	Cafe, Chinese	3.1	Average	15
1	Cafí© Doo Ghoont	Ansal Plaza Mall, Greater Noida	77.507456	28.463957	Cafe	3.4	Average	50

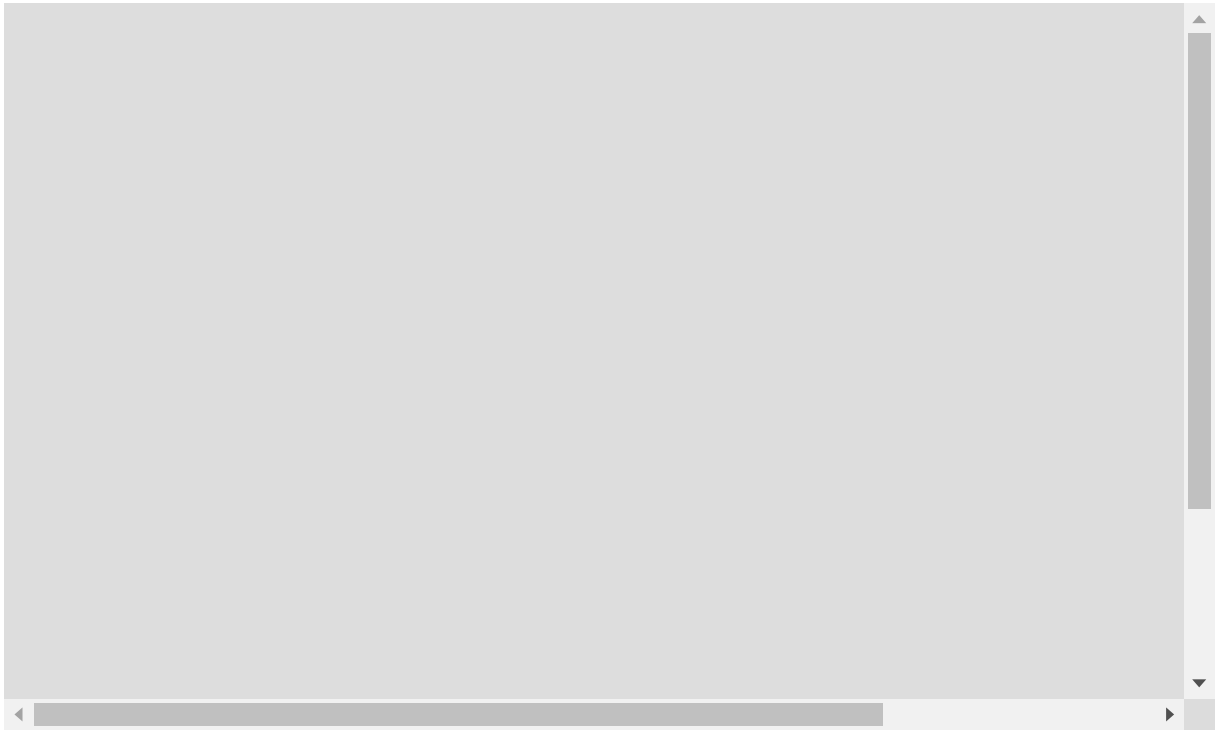
	Restaurant Name	Locality	Longitude	Latitude	Cuisines	Aggregate rating	Rating text	Votes
2	Knights Chaska	Ansal Plaza Mall, Greater Noida	77.507834	28.464434	North Indian, Chinese	2.9	Average	6
3	Thirsty Scholar Cafe	Ansal Plaza Mall, Greater Noida	77.507701	28.464201	Cafe	3.2	Average	45
5	JSB Evergreen Snack & Sweets	Brahmaputra Shopping Complex	77.332577	28.569899	North Indian, Chinese, Fast Food, Street Food	3.1	Average	47

3.1.2 Creating Map to show the Restaurant clusters

In [62]:
```
Noida_Rest = folium.Map(location=[28.60, 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'], d
f_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(Noida_Rest)

Noida_Rest
```

Out[62]:



In [63]:
```
df_Res.head()
```

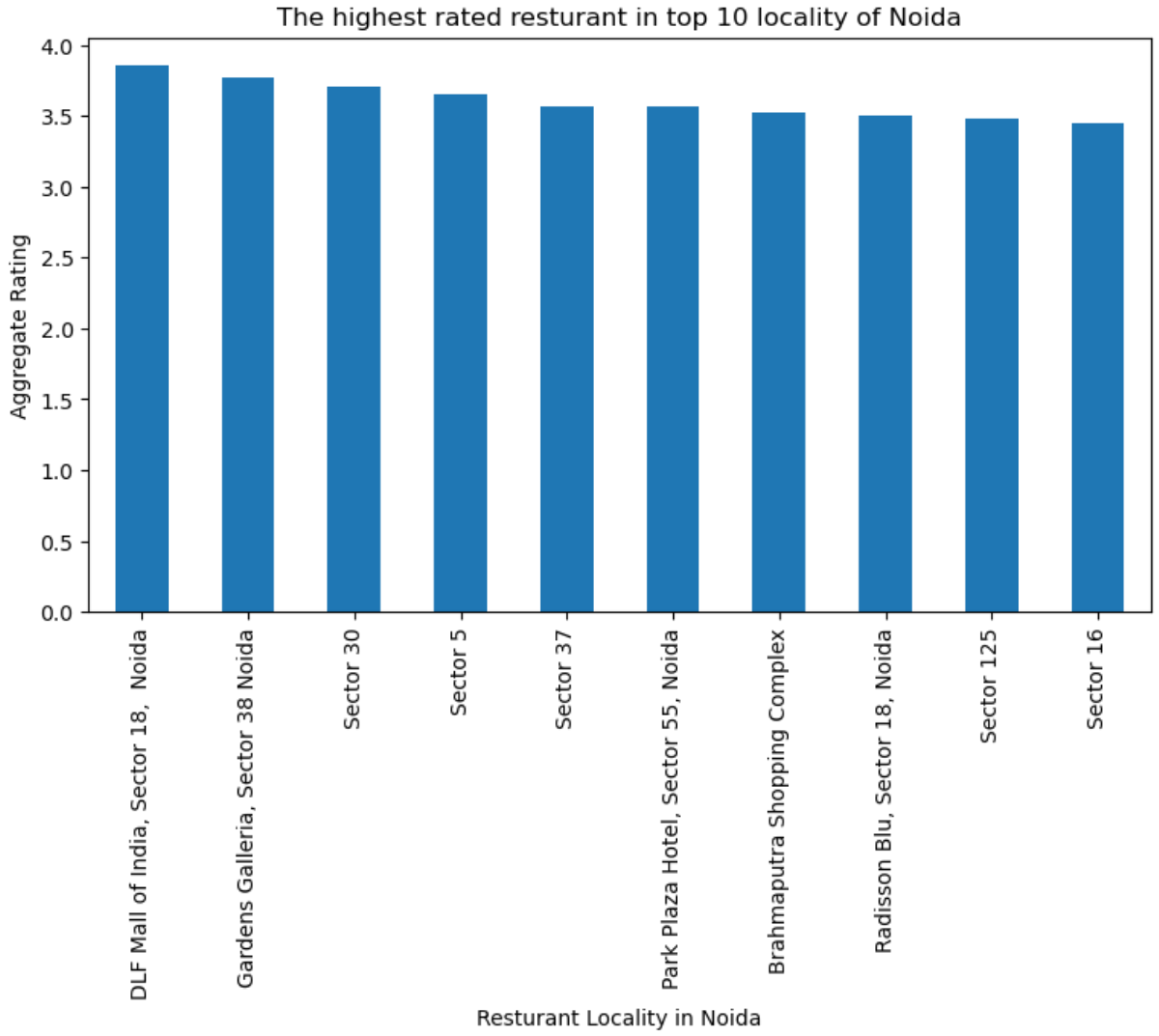
Out[63]:

	Restaurant Name	Locality	Longitude	Latitude	Cuisines	Aggregate rating	Rating text	Votes	Cluster
0	Ali Baba Caves	Ansal Plaza Mall, Greater Noida	77.528129	28.458107	Cafe, Chinese	3.1	Average	15	1
1	Cafí© Doo Ghoont	Ansal Plaza Mall, Greater Noida	77.507456	28.463957	Cafe	3.4	Average	50	1

	Restaurant Name	Locality	Longitude	Latitude	Cuisines	Aggregate rating	Rating text	Votes	Cluster
2	Knights Chaska	Ansal Plaza Mall, Greater Noida	77.507834	28.464434	North Indian, Chinese	2.9	Average	6	1
3	Thirsty Scholar Cafe	Ansal Plaza Mall, Greater Noida	77.507701	28.464201	Cafe	3.2	Average	45	1
5	JSB Evergreen Snack & Sweets	Brahmaputra Shopping Complex	77.332577	28.569899	North Indian, Chinese, Fast Food, Street Food	3.1	Average	47	0

3.1.3 Which places In Noida have some of the finest Restaurants?

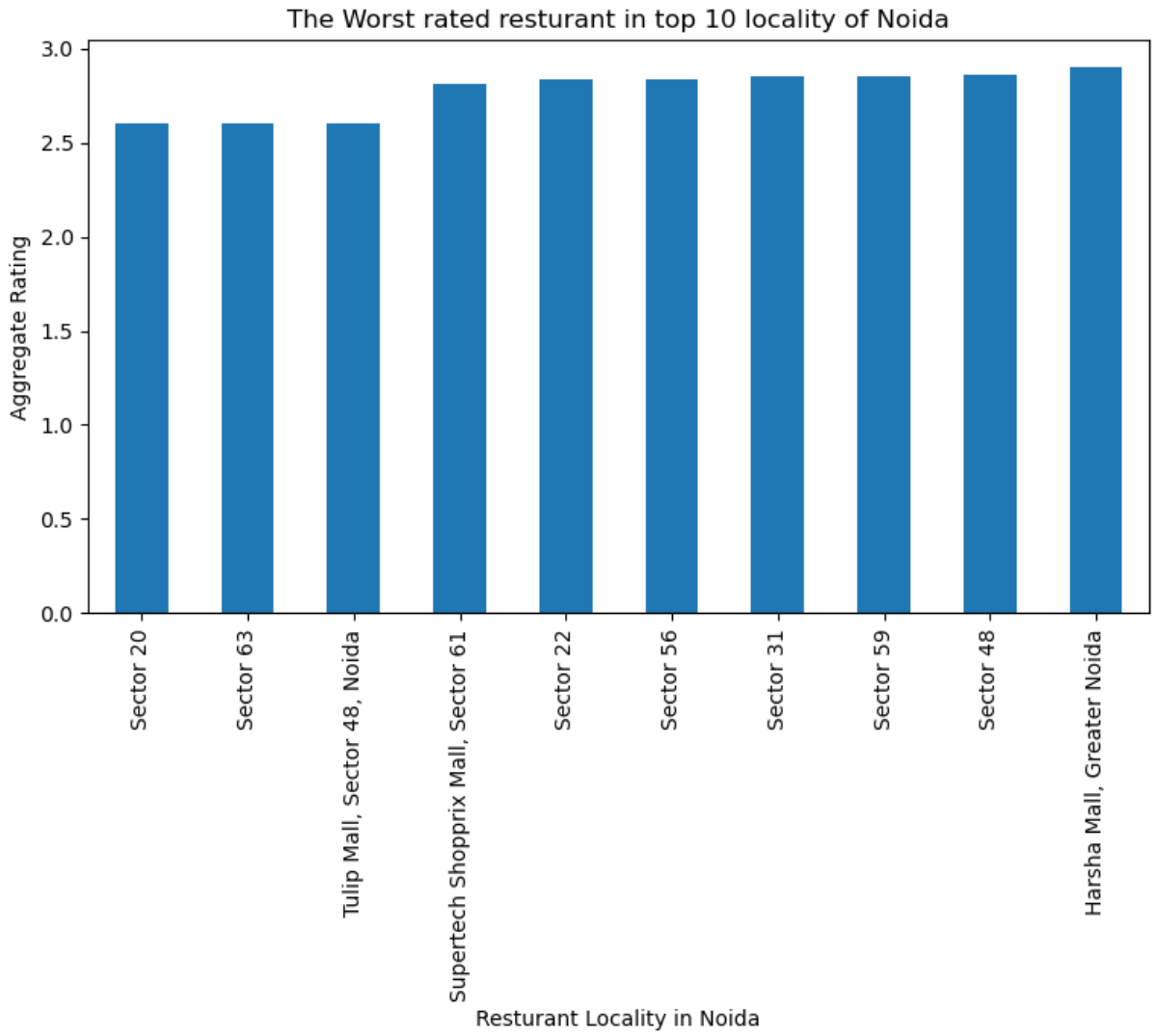
In [64]:
```
import matplotlib.pyplot as plt
plt.figure(figsize=(9,5), dpi = 100)
# title
plt.title('The highest rated resturant in top 10 locality of Noida')
#On x-axis

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

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



The best restarants are available in DLF Mall of India, Sector 18, Noida.

3.1.4 Which Places in Noida have Restaurants with Lowest rating?

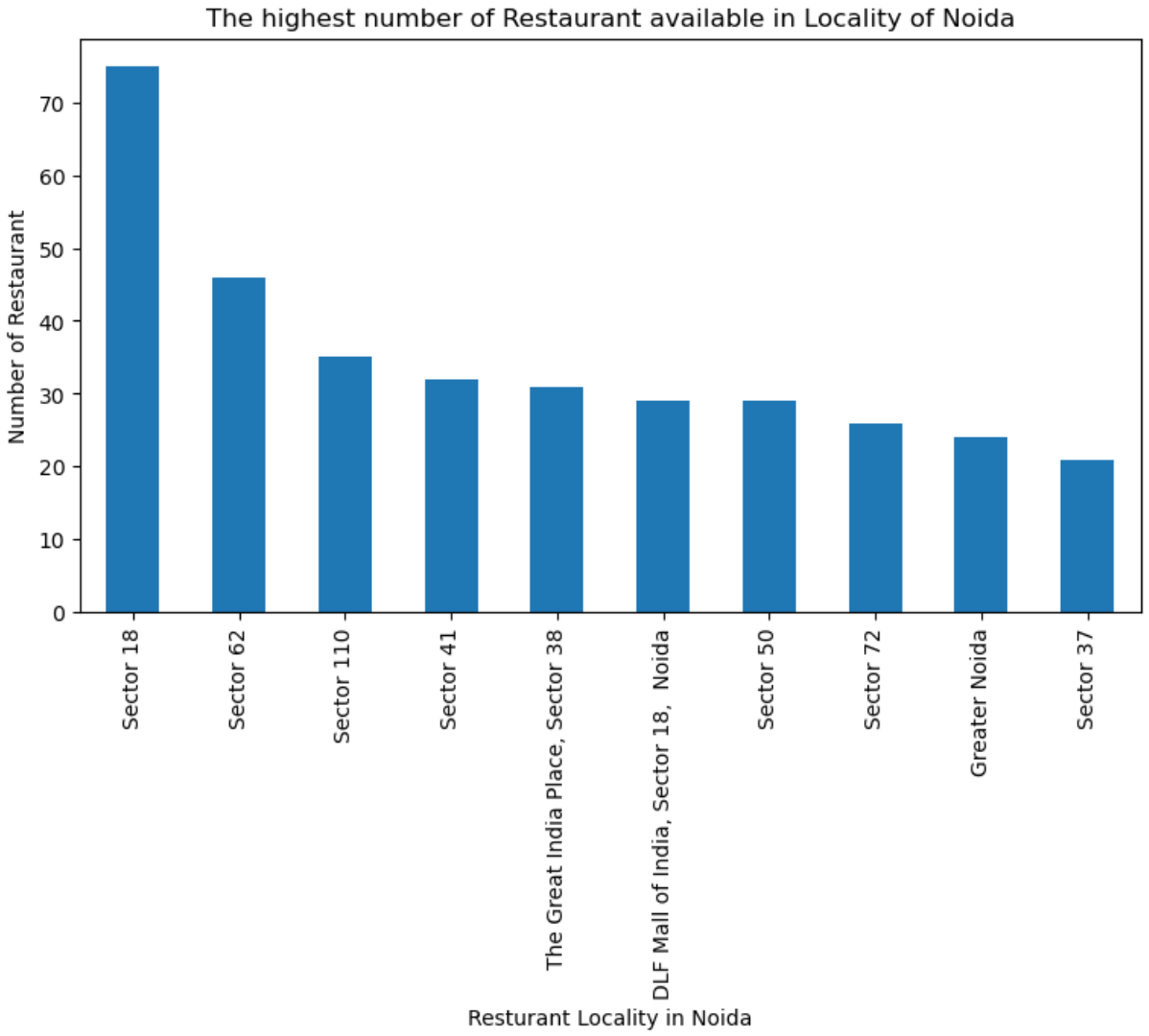
In [17]:
```
import matplotlib.pyplot as plt
plt.figure(figsize=(9,5), dpi = 100)
# title
plt.title('The Worst rated resturant in top 10 locality of Noida')
#On x-axis

#giving a bar plot

df_Res.groupby('Locality')['Aggregate rating'].mean().nsmallest(10).plo
t(kind='bar')

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

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



The worst rated Restaurants is in Sector 20.

3.1.5 Which place in Noida is suitable for a Foodie?

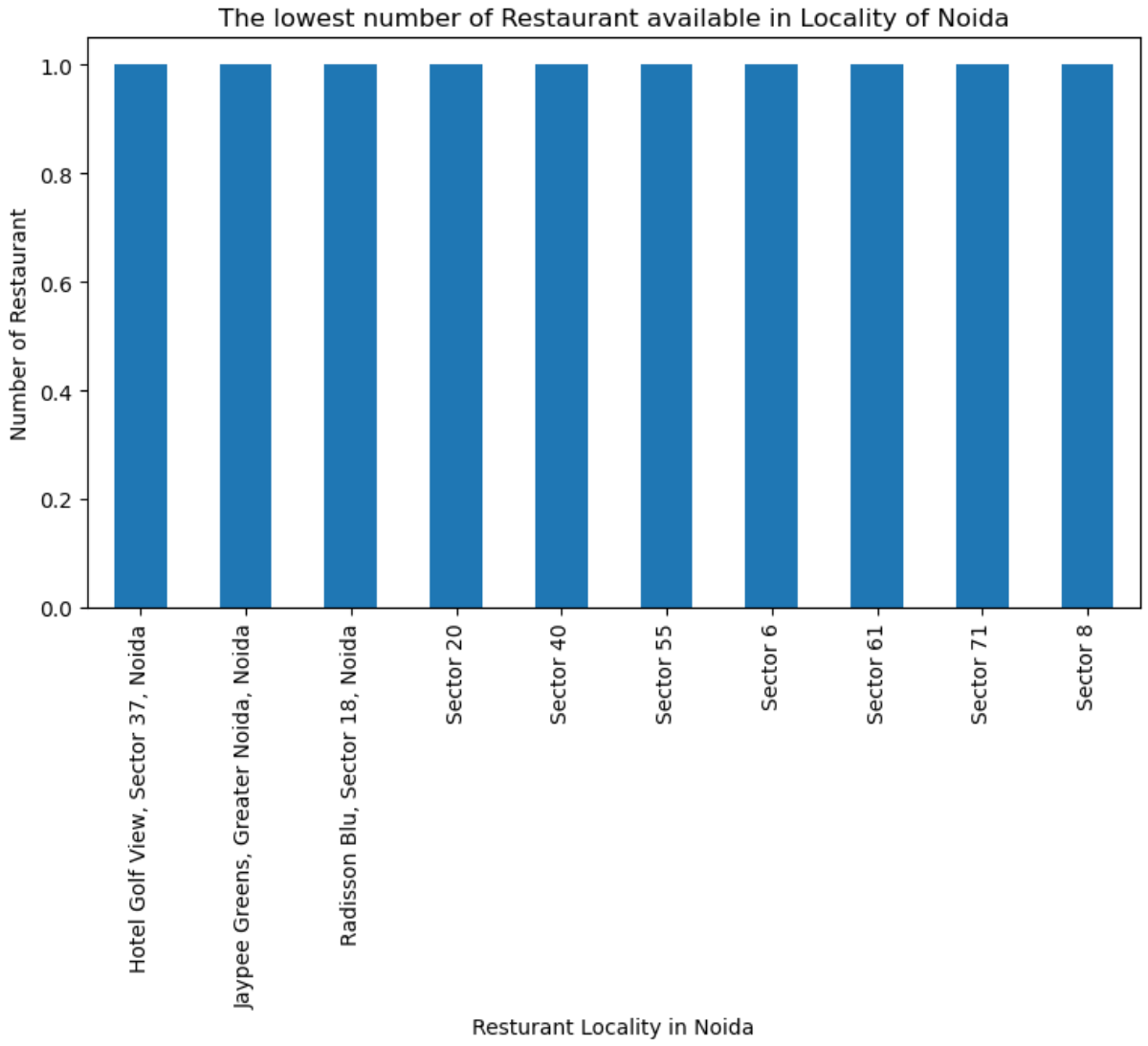
In [18]:
```
import matplotlib.pyplot as plt
plt.figure(figsize=(9,5), dpi = 100)
# title
plt.title('The highest number of Restaurant available in Locality of No
ida')
#On x-axis

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

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

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



Sector 18 is the best place for Foodie.

3.1.6 Which place in Noida is not suitable for a Foodie?

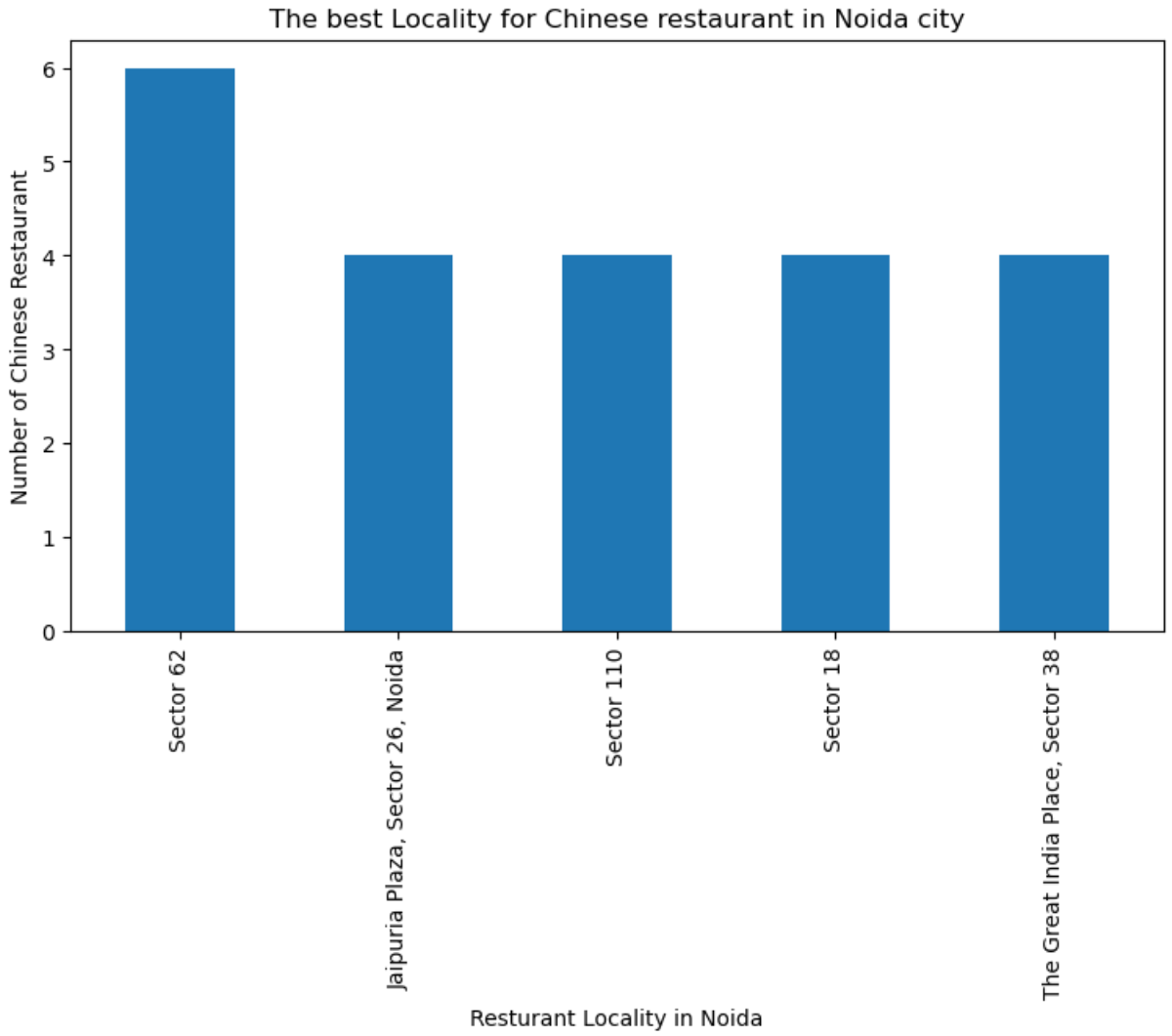
In [19]:
```
import matplotlib.pyplot as plt
plt.figure(figsize=(9,5), dpi = 100)
# title
plt.title('The lowest number of Restaurant available in Locality of Noi
da')
#On x-axis

#giving a bar plot
df_Res.groupby('Locality')['Restaurant Name'].count().nsmallest(10).plo
t(kind='bar')

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

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



No Sector is bad enough for the Foodie as per the above Statistics.

3.1.7 What are the best places in Noida for Chinese Restaurant?

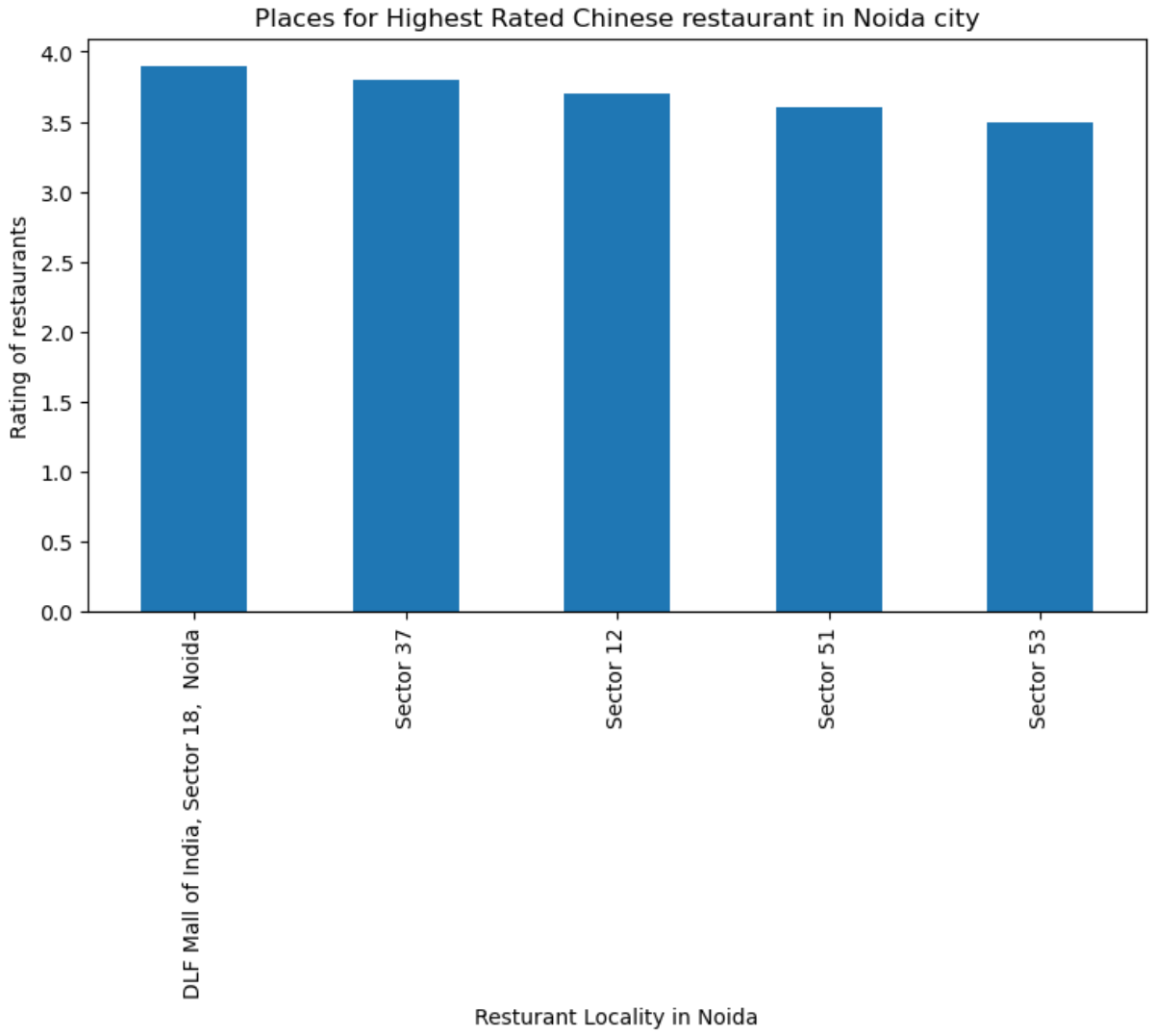
In [65]:
```
import matplotlib.pyplot as plt
plt.figure(figsize=(9,5), dpi = 100)
# title
plt.title('The best Locality for Chinese restaurant in Noida city')
#On x-axis

#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 Noida')
#On y-axis
plt.ylabel('Number of Chinese Restaurant')

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



Sector 62 is the best place for Chinese Restaurant.

3.1.8 Which places in Noida have highest rated Chinese Restaurants?

In [66]:
```
import matplotlib.pyplot as plt
plt.figure(figsize=(9,5), dpi = 100)
# title
plt.title('Places for Highest Rated Chinese restaurant in Noida city')
#On x-axis

#giving a bar plot
df_Res[df_Res['Cuisines'].str.startswith('Chinese')].groupby('Locality'
)['Aggregate rating'].mean().nlargest(5).plot(kind='bar')

plt.xlabel('Resturant Locality in Noida')
#On y-axis
plt.ylabel('Rating of restaurants')

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



DLF Mall of India, Sector 18, Noida has best Chinese Resturants.

3.2 Data transformation

Based on Locality grouping the data

In [67]:
```
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).re
set_index()
d_R = df_Res.groupby(['Locality'])['Rating text'].unique().agg(', '.joi
n).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='Lo
cality').merge(d_Cuisines, on='Locality').merge(df_Res_rating,on ='Loca
lity').merge(d_R, on ='Locality').merge(d_V, on ='Locality')
```

In [68]:
```
df_final = df_final[df_final['Aggregate rating'] != 0.000000]
df_final.columns =['Locality','Lat','Lng', 'No_of_Restaurant','Cusines'
, 'Agg_Rating','Comments' ,'No_of_Votes']
df_final.head()
```

Out[68]:

	Locality	Lat	Lng	No_of_Restaurant	Cusines	Agg_Rating	Comments	No
0	Ansal Plaza Mall, Greater Noida	28.462675	77.512780	4	Cafe, Chinese, Cafe, North Indian, Chinese, Cafe	3.150000	Average	
1	Brahmaputra Shopping Complex	28.570019	77.332701	4	North Indian, Chinese, Fast Food, Street Food,...	3.525000	Average, Good	

	Locality	Lat	Lng	No_of_Restaurant	Cusines	Agg_Rating	Comments	No
2	Centre Stage Mall, Sector 18	28.568185	77.323030	4	Asian, Continental, Italian, North Indian, Nor...	3.100000	Average, Good, Poor	
3	DLF Mall of India, Sector 18, Noida	28.567267	77.320868	29	North Indian, Mughlai, North Indian, European,...	3.858621	Average, Excellent, Good, Very Good	
4	Fortune Inn Grazia, Sector 27, Noida	28.577435	77.328359	2	North Indian, Mughlai, Chinese	3.300000	Average	

In [69]:
```
df_final.shape
```

Out[69]:
```
(77, 8)
```

3.2.1 Defining Foursquare Credentials and Version

In [70]:
```
## Define Foursquare Credentials and Version
CLIENT_ID = 'ZNUQA0Y0JUXEKGWLC2HQAS1HVNDJ4UBDTZIKKUZ40DZH1EL1' # Foursq
uare ID
CLIENT_SECRET = 'CVQOTVVVTKJACYGLJUY1W2GVZ0UC21YS5S5W0LKOLI1UODKB' # Fo
ursquare Secret
VERSION = '20180604' # Foursquare API version

print('Your credentails:')
print('CLIENT_ID: ' + CLIENT_ID)
print('CLIENT_SECRET:' + CLIENT_SECRET)
```

```
Your credentails:
CLIENT_ID: ZNUQA0Y0JUXEKGWLC2HQAS1HVNDJ4UBDTZIKKUZ40DZH1EL1
```

```
CLIENT_SECRET:CVQOTVVVTKJACYGLJUY1W2GVZ0UC21YS5S5W0LKOLI1UODKB
```

3.2.2 Function to analyse all localities in Noida in similar way

In [71]:

```
## create Function to analyse all localities in Noida in similar way

def getNearbyVenues(names, latitudes, longitudes, radius=500,LIMIT = 50
):

    venues_list=[]
    for name, lat, lng in zip(names, latitudes, longitudes):
        print(name)

        # create the API request URL
        url = 'https://api.foursquare.com/v2/venues/explore?&client_id=
{}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
            CLIENT_ID,
            CLIENT_SECRET,
            VERSION,
            lat,
            lng,
            radius,
            LIMIT)

        # make the GET request
        results = requests.get(url).json()["response"]['groups'][0]['it
ems']

        # return only relevant information for each nearby venue
        venues_list.append([(
            name,
            lat,
            lng,
            v['venue']['name'],
            v['venue']['location']['lat'],
            v['venue']['location']['lng'],
```

```
            v['venue']['categories'][0]['name']) for v in results])

    nearby_venues = pd.DataFrame([item for venue_list in venues_list fo
r item in venue_list])
    nearby_venues.columns = ['Locality',
                  'Locality Latitude',
                  'Locality Longitude',
                  'Venue',
                  'Venue Latitude',
                  'Venue Longitude',
                  'Venue Category']

    return(nearby_venues)
```

3.2.3 Retrieve venues for Noida Localities

In [72]:
```
# Retrieve venues for Noida Localities
Noida_venues = getNearbyVenues(names=df_final['Locality'],
                               latitudes=df_final['Lat'],
                               longitudes=df_final['Lng']
                              )
```

```
Ansal Plaza Mall, Greater Noida
Brahmaputra Shopping Complex
Centre Stage Mall, Sector 18
DLF Mall of India, Sector 18,  Noida
Fortune Inn Grazia, Sector 27, Noida
Ganga Shopping Complex, Sector 29
Gardens Galleria, Sector 38 Noida
Greater Noida
Harsha Mall, Greater Noida
Hotel Golf View, Sector 37, Noida
Jaipuria Plaza, Sector 26, Noida
Jaypee Greens Golf & Spa Resort, Surajpur
Jaypee Greens, Greater Noida, Noida
Logix City Centre, Sector 32, Noida
MSX Mall, Greater Noida
Mosaic Hotels, Sector 18, Noida
```

```
Park Plaza Hotel, Sector 55, Noida
Radisson Blu, Sector 18, Noida
Sector 10
Sector 11
Sector 110
Sector 12
Sector 125
Sector 132
Sector 15
Sector 16
Sector 18
Sector 19
Sector 2
Sector 20
Sector 21
Sector 22
Sector 25
Sector 26
Sector 27
Sector 28
Sector 29
Sector 3
Sector 30
Sector 31
Sector 33
Sector 34
Sector 37
Sector 38
Sector 39
Sector 40
Sector 41
Sector 44
Sector 45
Sector 47
Sector 48
Sector 49
Sector 5
Sector 50
Sector 51
```

```
Sector 52
Sector 53
Sector 55
Sector 56
Sector 58
Sector 59
Sector 6
Sector 61
Sector 62
Sector 63
Sector 65
Sector 7
Sector 71
Sector 72
Sector 8
Sector 83
Sector 93
Shopprix Mall, Sector 61, Noida
Spice World Mall, Sector 25
Supertech Shopprix Mall, Sector 61
The Great India Place, Sector 38
Tulip Mall, Sector 48, Noida
```

In [74]:
```
Noida_venues.head()
```

Out[74]:

	Locality	Locality Latitude	Locality Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Ansal Plaza Mall, Greater Noida	28.462675	77.512780	A Fresh	28.463620	77.511691	Supermarket
1	Ansal Plaza Mall, Greater Noida	28.462675	77.512780	BIKANERVALA	28.463515	77.510857	Indian Restaurant
2	Ansal Plaza Mall, Greater Noida	28.462675	77.512780	Savoy Suites	28.464374	77.509695	Hotel
3	Brahmaputra Shopping Complex	28.570019	77.332701	Brahmaputra Shopping Complex	28.570067	77.332539	Snack Place

	Locality	Locality Latitude	Locality Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
4	Brahmaputra Shopping Complex	28.570019	77.332701	Lakshmi Coffee House	28.570140	77.332734	South Indian Restaurant

In [75]:
```
Noida_venues.groupby('Locality').count()
```

Out[75]:

Locality	Locality Latitude	Locality Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Ansal Plaza Mall, Greater Noida	3	3	3	3	3	3
Brahmaputra Shopping Complex	4	4	4	4	4	4
Centre Stage Mall, Sector 18	49	49	49	49	49	49
DLF Mall of India, Sector 18, Noida	45	45	45	45	45	45
Fortune Inn Grazia, Sector 27, Noida	7	7	7	7	7	7
...	...	...	...	...	...	...
Shopprix Mall, Sector 61, Noida	9	9	9	9	9	9
Spice World Mall, Sector 25	9	9	9	9	9	9
Supertech Shopprix Mall, Sector 61	9	9	9	9	9	9
The Great India Place, Sector 38	41	41	41	41	41	41
Tulip Mall, Sector 48, Noida	3	3	3	3	3	3

76 rows × 6 columns

In [76]:
```
print('There are {} uniques categories.'.format(len(Noida_venues['Venue
 Category'].unique())))
```

```
There are 81 uniques categories.
```

In [77]:
```
## Analyze Each Locality

# one hot encoding
Noida_onehot = pd.get_dummies(Noida_venues[['Venue Category']], prefix=
"", prefix_sep="")

# add Locality column back to dataframe
Noida_onehot['Locality'] = Noida_venues['Locality']

# move Locality column to the first column
column_list = Noida_onehot.columns.tolist()
column_number = int(column_list.index('Locality'))
column_list = [column_list[column_number]] + column_list[:column_number
] + column_list[column_number+1:]
Noida_onehot = Noida_onehot[column_list]

Noida_onehot.head()
```

Out[77]:

	Locality	ATM	Arcade	Asian Restaurant	BBQ Joint	Bakery	Bar	Bed & Breakfast	Bistro	Bowling Alley	...	Re
0	Ansal Plaza Mall, Greater Noida	0	0	0	0	0	0	0	0	0	...	
1	Ansal Plaza Mall, Greater Noida	0	0	0	0	0	0	0	0	0	...	
2	Ansal Plaza Mall, Greater Noida	0	0	0	0	0	0	0	0	0	...	

	Locality	ATM	Arcade	Asian Restaurant	BBQ Joint	Bakery	Bar	Bed & Breakfast	Bistro	Bowling Alley	...	Re
3	Brahmaputra Shopping Complex	0	0	0	0	0	0	0	0	0	...	
4	Brahmaputra Shopping Complex	0	0	0	0	0	0	0	0	0	...	

5 rows × 82 columns

In [78]:
```
Noida_grouped = Noida_onehot.groupby('Locality').mean().reset_index()
Noida_grouped
```

Out[78]:

	Locality	ATM	Arcade	Asian Restaurant	BBQ Joint	Bakery	Bar	Bed & Breakfast	Bistro	Bowling Alley	.
0	Ansal Plaza Mall, Greater Noida	0.0	0.000000	0.000000	0.0	0.000000	0.0	0.0	0.0	0.000000	.
1	Brahmaputra Shopping Complex	0.0	0.000000	0.000000	0.0	0.000000	0.0	0.0	0.0	0.000000	.
2	Centre Stage Mall, Sector 18	0.0	0.020408	0.020408	0.0	0.020408	0.0	0.0	0.0	0.020408	.
3	DLF Mall of India, Sector 18, Noida	0.0	0.022222	0.000000	0.0	0.022222	0.0	0.0	0.0	0.022222	.
4	Fortune Inn Grazia, Sector 27, Noida	0.0	0.000000	0.000000	0.0	0.142857	0.0	0.0	0.0	0.000000	.
...	...	...	...	...	...	...	...	...	...	...	.

	Locality	ATM	Arcade	Asian Restaurant	BBQ Joint	Bakery	Bar	Bed & Breakfast	Bistro	Bowling Alley	.
71	Shopprix Mall, Sector 61, Noida	0.0	0.000000	0.000000	0.0	0.000000	0.0	0.0	0.0	0.000000	.
72	Spice World Mall, Sector 25	0.0	0.000000	0.000000	0.0	0.000000	0.0	0.0	0.0	0.000000	.
73	Supertech Shopprix Mall, Sector 61	0.0	0.000000	0.000000	0.0	0.000000	0.0	0.0	0.0	0.000000	.
74	The Great India Place, Sector 38	0.0	0.024390	0.024390	0.0	0.024390	0.0	0.0	0.0	0.024390	.
75	Tulip Mall, Sector 48, Noida	0.0	0.000000	0.000000	0.0	0.333333	0.0	0.0	0.0	0.000000	.

76 rows × 82 columns

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

Out[79]:
```
(76, 82)
```

In [80]:
```
## print each Locality along with the top 5 most common venues

num_top_venues = 5

for hood in Noida_grouped['Locality']:
    print("----"+hood+"----")
    temp = Noida_grouped[Noida_grouped['Locality'] == hood].T.reset_ind
ex()
    temp.columns = ['venue','freq']
    temp = temp.iloc[1:]
    temp['freq'] = temp['freq'].astype(float)
```

```
    temp = temp.round({'freq': 2})
    print(temp.sort_values('freq', ascending=False).reset_index(drop=Tr
ue).head(num_top_venues))
    print('\n')
```

```
----Ansal Plaza Mall, Greater Noida----
               venue  freq
0              Hotel  0.33
1  Indian Restaurant  0.33
2        Supermarket  0.33
3                ATM  0.00
4             Lounge  0.00


----Brahmaputra Shopping Complex----
                     venue  freq
0                Hotel Bar  0.25
1        Indian Restaurant  0.25
2              Snack Place  0.25
3  South Indian Restaurant  0.25
4        Indian Sweet Shop  0.00


----Centre Stage Mall, Sector 18----
                  venue  freq
0     Indian Restaurant  0.16
1           Coffee Shop  0.10
2    Chinese Restaurant  0.10
3           Pizza Place  0.06
4  Fast Food Restaurant  0.06


----DLF Mall of India, Sector 18,  Noida----
                  venue  freq
0                  Café  0.11
1     Indian Restaurant  0.11
2           Coffee Shop  0.09
3           Pizza Place  0.07
4  Fast Food Restaurant  0.07
```

```
----Fortune Inn Grazia, Sector 27, Noida----
               venue  freq
0         Food Truck  0.29
1        Pizza Place  0.14
2  Convenience Store  0.14
3            Bakery  0.14
4             Hotel  0.14


----Ganga Shopping Complex, Sector 29----
               venue  freq
0  Indian Restaurant   0.1
1        Bus Station   0.1
2        Pizza Place   0.1
3      Metro Station   0.1
4              Café   0.1


----Gardens Galleria, Sector 38 Noida----
               venue  freq
0         Coffee Shop  0.12
1   Indian Restaurant  0.09
2         Pizza Place  0.09
3  Chinese Restaurant  0.09
4       Shopping Mall  0.06


----Greater Noida----
                 venue  freq
0                  ATM  0.12
1                  Spa  0.12
2          Coffee Shop  0.12
3  Monument / Landmark  0.12
4                Café   0.12


----Harsha Mall, Greater Noida----
            venue  freq
0     Pizza Place   0.2
```

```
0      Pizza Place   0.2

1  Business Center   0.2
2      Coffee Shop   0.2
3             Café   0.2
4              Spa   0.2


----Hotel Golf View, Sector 37, Noida----
                       venue  freq
0          Indian Restaurant  0.11
1                  Tea Room  0.11
2  Middle Eastern Restaurant  0.11
3                       Café  0.11
4                     Market  0.11


----Jaipuria Plaza, Sector 26, Noida----
                venue  freq
0   Department Store   0.2
1                Bar   0.2
2                Gym   0.2
3       Burger Joint   0.2
4  Indian Restaurant   0.2


----Jaypee Greens Golf & Spa Resort, Surajpur----
          venue  freq
0  Pizza Place  0.17
1        Hotel  0.17
2  Coffee Shop  0.17
3          Spa  0.17
4         Café  0.17


----Jaypee Greens, Greater Noida, Noida----
          venue  freq
0  Golf Course  0.50
1        Diner  0.25
2        Hotel  0.25
3          ATM  0.00
```

```
3          ATM  0.00

4       Market  0.00


----Logix City Centre, Sector 32, Noida----
           venue  freq
0      Multiplex   0.4
1  Movie Theater   0.2
2  Metro Station   0.2
3     Food Truck   0.2
4            ATM   0.0


----MSX Mall, Greater Noida----
                       venue  freq
0               Shopping Mall   0.2
1  Indian Chinese Restaurant   0.2
2                Pizza Place   0.2
3          Indian Sweet Shop   0.2
4                Supermarket   0.2


----Mosaic Hotels, Sector 18, Noida ----
                   venue  freq
0     Indian Restaurant  0.24
1    Chinese Restaurant  0.15
2  Fast Food Restaurant  0.09
3                 Café  0.09
4        Clothing Store  0.06


----Park Plaza Hotel, Sector 55, Noida----
               venue  freq
0              Hotel   0.5
1  Indian Restaurant   0.5
2                ATM   0.0
3             Lounge   0.0
4          Multiplex   0.0
```

```
----Radisson Blu, Sector 18, Noida----
                venue  freq
0   Indian Restaurant  0.17
1         Coffee Shop  0.11
2  Chinese Restaurant  0.11
3               Café  0.06
4         Pizza Place  0.06


----Sector 10----
         venue  freq
0  IT Services   0.5
1   Food Truck   0.5
2       Market   0.0
3  Music Store   0.0
4    Multiplex   0.0


----Sector 11----
                  venue  freq
0     Electronics Store  0.25
1  Fast Food Restaurant  0.25
2           Golf Course  0.25
3             Gift Shop  0.25
4                   ATM  0.00


----Sector 110----
                    venue  freq
0                  Bakery   0.4
1  Health & Beauty Service   0.4
2                Mountain   0.2
3                     ATM   0.0
4                  Market   0.0


----Sector 12----
                venue  freq
0   Electronics Store   0.2
```

```
0   Electronics Store   0.2

1  Chinese Restaurant   0.2
2           Pool Hall   0.2
3         Golf Course   0.2
4           Gift Shop   0.2


----Sector 125----
              venue  freq
0  Asian Restaurant   0.2
1       Food Court   0.2
2       Food Truck   0.2
3             Café   0.2
4    Sandwich Place   0.2


----Sector 132----
           venue  freq
0      Cafeteria   0.5
1  Metro Station   0.5
2            ATM   0.0
3         Market   0.0
4    Music Store   0.0


----Sector 15----
                   venue  freq
0     Indian Restaurant  0.29
1  Fast Food Restaurant  0.29
2             BBQ Joint  0.14
3            Restaurant  0.14
4        Sandwich Place  0.14


----Sector 16----
                   venue  freq
0  Fast Food Restaurant  0.33
1         Train Station  0.17
2             BBQ Joint  0.17
3     Indian Restaurant  0.17
```

```
3     Indian Restaurant  0.17

4        Sandwich Place  0.17


----Sector 18----
                  venue  freq
0     Indian Restaurant  0.20
1    Chinese Restaurant  0.12
2                 Café  0.07
3  Fast Food Restaurant  0.07
4           Coffee Shop  0.07


----Sector 19----
            venue  freq
0     Pizza Place   0.2
1           Hotel   0.2
2          Bakery   0.2
3      Food Truck   0.2
4  Sandwich Place   0.2


----Sector 2----
                  venue  freq
0  Fast Food Restaurant  0.33
1     Indian Restaurant  0.17
2             BBQ Joint  0.17
3                 Hotel  0.17
4        Sandwich Place  0.17


----Sector 20----
               venue  freq
0             Market  0.25
1  Electronics Store  0.25
2     Sandwich Place  0.25
3         Food Truck  0.25
4      Metro Station  0.00
```

```
----Sector 21----
               venue  freq
0             Market  0.12
1            Stadium  0.12
2  Electronics Store  0.12
3        Coffee Shop  0.12
4          Multiplex  0.12


----Sector 22----
                venue  freq
0   Indian Restaurant  0.50
1  Chinese Restaurant  0.25
2    Lottery Retailer  0.25
3                 ATM  0.00
4       Metro Station  0.00


----Sector 25----
               venue  freq
0             Market  0.12
1            Stadium  0.12
2  Electronics Store  0.12
3        Coffee Shop  0.12
4          Multiplex  0.12


----Sector 26----
               venue  freq
0   Department Store   0.2
1                Bar   0.2
2                Gym   0.2
3       Burger Joint   0.2
4  Indian Restaurant   0.2


----Sector 27----
                venue  freq
0   Indian Restaurant   0.2
```

```
0    Indian Restaurant   0.2

1              Bakery   0.2
2       Sandwich Place   0.1
3                Hotel   0.1
4  Fried Chicken Joint   0.1


----Sector 28----
                     venue  freq
0                Hotel Bar  0.25
1       Indian Restaurant  0.25
2             Snack Place  0.25
3  South Indian Restaurant  0.25
4       Indian Sweet Shop  0.00


----Sector 29----
               venue  freq
0  Indian Restaurant  0.11
1        Bus Station  0.11
2        Pizza Place  0.11
3        Flea Market  0.11
4               Café  0.11


----Sector 3----
                  venue  freq
0     Indian Restaurant  0.43
1  Fast Food Restaurant  0.14
2        Sandwich Place  0.07
3         Train Station  0.07
4             BBQ Joint  0.07


----Sector 30----
               venue  freq
0  Convenience Store  0.25
1                Gym  0.25
2         Food Court  0.25
3     Sandwich Place  0.25
```

```
3     Sandwich Place  0.25

4                ATM  0.00


----Sector 31----
               venue  freq
0  Convenience Store   0.5
1          Multiplex   0.5
2           Pharmacy   0.0
3        Music Store   0.0
4      Movie Theater   0.0


----Sector 33----
           venue  freq
0      Multiplex   0.4
1  Movie Theater   0.2
2         Lawyer   0.2
3  Shopping Mall   0.2
4            ATM   0.0


----Sector 34----
                venue  freq
0         Bus Station  0.25
1  Chinese Restaurant  0.25
2                 Gym  0.25
3     Breakfast Spot  0.25
4              Market  0.00


----Sector 37----
                       venue  freq
0          Indian Restaurant  0.22
1                     Bistro  0.11
2             Clothing Store  0.11
3  Middle Eastern Restaurant  0.11
4                     Market  0.11
```

```
----Sector 38----
                venue  freq
0   Indian Restaurant  0.21
1  Chinese Restaurant  0.16
2         Coffee Shop  0.11
3      Clothing Store  0.11
4       Shopping Mall  0.05


----Sector 39----
                 venue  freq
0       Ice Cream Shop  0.25
1  Fried Chicken Joint  0.25
2                 Pool  0.25
3   Light Rail Station  0.25
4            Multiplex  0.00


----Sector 40----
                    venue  freq
0                    Park   0.2
1  Furniture / Home Store   0.2
2        Asian Restaurant   0.2
3       Electronics Store   0.2
4                   Café   0.2


----Sector 41----
          venue  freq
0  Pizza Place  0.25
1       Bakery  0.25
2       Lounge  0.25
3       Arcade  0.25
4   Theme Park  0.00


----Sector 44----
                venue  freq
0  Convenience Store   1.0
```

```
0  Convenience Store   1.0

1           Pharmacy   0.0
2       Music Store    0.0
3         Multiplex    0.0
4     Movie Theater    0.0


----Sector 45----
                  venue  freq
0                   ATM  0.33
1  Gym / Fitness Center  0.33
2            Restaurant  0.33
3                Lounge  0.00
4             Multiplex  0.00


----Sector 47----
                  venue  freq
0                 Hotel  0.33
1                Bakery  0.33
2  Gym / Fitness Center  0.33
3                   ATM  0.00
4                Lounge  0.00


----Sector 48----
                  venue  freq
0                   ATM   0.2
1                   Spa   0.2
2  Gym / Fitness Center   0.2
3           Pizza Place   0.2
4              Pharmacy   0.2


----Sector 49----
                  venue  freq
0                   ATM  0.25
1           Pizza Place  0.25
2  Gym / Fitness Center  0.25
3              Pharmacy  0.25
```

```
3              Pharmacy  0.25

4              Tea Room  0.00


----Sector 5----
         venue  freq
0          ATM  0.33
1   Hookah Bar  0.33
2          Bar  0.33
3       Market  0.00
4  Music Store  0.00


----Sector 50----
                venue  freq
0         Pizza Place  0.14
1              Bakery  0.14
2  Italian Restaurant  0.14
3      Sandwich Place  0.14
4                Park  0.14


----Sector 51----
                venue  freq
0  Convenience Store  0.33
1               Pool  0.33
2                Gym  0.33
3             Market  0.00
4        Music Store  0.00


----Sector 52----
           venue  freq
0   Dessert Shop  0.25
1    Pizza Place  0.25
2     Restaurant  0.25
3  Shopping Mall  0.25
4            ATM  0.00
```

```
----Sector 53----
            venue  freq
0     Snack Place  0.25
1  Clothing Store  0.25
2      Food Truck  0.25
3   Shopping Mall  0.25
4            Park  0.00


----Sector 55----
              venue  freq
0            Garden  0.25
1  Indian Restaurant  0.25
2            Hotel  0.25
3             Café  0.25
4            Market  0.00


----Sector 56----
              venue  freq
0  Department Store  0.25
1      Dessert Shop  0.25
2            Garden  0.25
3       IT Services  0.25
4          Tea Room  0.00


----Sector 58----
           venue  freq
0    Coffee Shop  0.33
1   Liquor Store  0.33
2           Food  0.33
3            ATM  0.00
4  Metro Station  0.00


----Sector 59----
                  venue  freq
0  Fast Food Restaurant  0.25
```

```
0  Fast Food Restaurant  0.25

1              Tea Room  0.25
2                  Food  0.25
3                 Café  0.25
4                   ATM  0.00


----Sector 61----
                  venue  freq
0     Indian Restaurant  0.18
1                   Gym  0.18
2         Shopping Mall  0.09
3  Fast Food Restaurant  0.09
4           Pizza Place  0.09


----Sector 62----
              venue  freq
0    Sandwich Place  0.25
1  Department Store  0.12
2         Multiplex  0.12
3       Coffee Shop  0.12
4        Food Truck  0.12


----Sector 63----
                  venue  freq
0  Fast Food Restaurant   1.0
1                   ATM   0.0
2                  Park   0.0
3            Multiplex   0.0
4         Movie Theater   0.0


----Sector 65----
           venue  freq
0          Hotel  0.75
1            ATM  0.25
2         Lounge  0.00
3     Multiplex  0.00
```

```
3      Multiplex  0.00

4  Movie Theater  0.00


----Sector 7----
              venue  freq
0              Gym   0.2
1             Café   0.2
2        Food Truck   0.2
3           Market   0.2
4  Business Service   0.2


----Sector 71----
         venue  freq
0         Park  0.33
1  Cheese Shop  0.33
2  Golf Course  0.33
3       Market  0.00
4  Music Store  0.00


----Sector 72----
                  venue  freq
0  Fast Food Restaurant  0.33
1                 Café  0.33
2              Pharmacy  0.33
3                   ATM  0.00
4         Metro Station  0.00


----Sector 8----
             venue  freq
0  Bed & Breakfast   0.5
1       Food Truck   0.5
2              ATM   0.0
3           Market   0.0
4     Music Store   0.0
```

```
----Sector 83----
                  venue  freq
0  Gym / Fitness Center   1.0
1                   ATM   0.0
2                Lounge   0.0
3             Multiplex   0.0
4         Movie Theater   0.0


----Sector 93----
             venue  freq
0  Ice Cream Shop   0.2
1           Plaza   0.2
2     Coffee Shop   0.2
3     Golf Course   0.2
4            Café   0.2


----Shopprix Mall, Sector 61, Noida----
                venue  freq
0         Pizza Place  0.22
1   Indian Restaurant  0.22
2                 Gym  0.22
3  Chinese Restaurant  0.11
4   Convenience Store  0.11


----Spice World Mall, Sector 25----
               venue  freq
0            Stadium  0.22
1             Market  0.11
2  Electronics Store  0.11
3        Coffee Shop  0.11
4          Multiplex  0.11


----Supertech Shopprix Mall, Sector 61----
                venue  freq
0         Pizza Place  0.22
```

```
0         Pizza Place  0.22

1   Indian Restaurant  0.22
2                 Gym  0.22
3  Chinese Restaurant  0.11
4   Convenience Store  0.11


----The Great India Place, Sector 38----
                  venue  freq
0     Indian Restaurant  0.12
1           Coffee Shop  0.10
2    Chinese Restaurant  0.10
3  Fast Food Restaurant  0.07
4        Clothing Store  0.07


----Tulip Mall, Sector 48, Noida----
                  venue  freq
0                 Hotel  0.33
1                Bakery  0.33
2  Gym / Fitness Center  0.33
3                   ATM  0.00
4                Lounge  0.00
```

In [81]:
```
## put that into a pandas dataframe
## First, write a function to sort the venues in descending order.

def return_most_common_venues(row, num_top_venues):
    row_categories = row.iloc[1:]
    row_categories_sorted = row_categories.sort_values(ascending=False)

    return row_categories_sorted.index.values[0:num_top_venues]
```

In [82]:
```
## create the new dataframe and display the top 10 venues for each Loca
lity.
```

```
num_top_venues = 10

indicators = ['st', 'nd', 'rd']

# create columns according to number of top venues
columns = ['Locality']
for ind in np.arange(num_top_venues):
    try:
        columns.append('{}{} Most Common Venue'.format(ind+1, indicator
s[ind]))
    except:
        columns.append('{}th Most Common Venue'.format(ind+1))

# create a new dataframe
Locality_venues_sorted = pd.DataFrame(columns=columns)
Locality_venues_sorted['Locality'] = Noida_grouped['Locality']

for ind in np.arange(Noida_grouped.shape[0]):
    Locality_venues_sorted.iloc[ind, 1:] = return_most_common_venues(No
ida_grouped.iloc[ind, :], num_top_venues)

Locality_venues_sorted
```

Out[82]:

	Locality	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 Mo Comm Ven
0	Ansal Plaza Mall, Greater Noida	Hotel	Supermarket	Indian Restaurant	Food Truck	Dessert Shop	Diner	Electron St
1	Brahmaputra Shopping Complex	Hotel Bar	Indian Restaurant	South Indian Restaurant	Snack Place	Food Court	Department Store	Dess Sh
2	Centre Stage Mall, Sector 18	Indian Restaurant	Chinese Restaurant	Coffee Shop	Café	Fast Food Restaurant	Pizza Place	Cloth St
3	DLF Mall of India, Sector 18, Noida	Café	Indian Restaurant	Coffee Shop	Chinese Restaurant	Fast Food Restaurant	Pizza Place	Cloth St

	Locality	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 Mo Comm Ven
4	Fortune Inn Grazia, Sector 27, Noida	Food Truck	Hotel	Pizza Place	Bakery	Flea Market	Convenience Store	Fo Co
...	...	...	...	...	...	...	...	
71	Shopprix Mall, Sector 61, Noida	Gym	Pizza Place	Indian Restaurant	Vegetarian / Vegan Restaurant	Chinese Restaurant	Convenience Store	Fo
72	Spice World Mall, Sector 25	Stadium	Music Store	Indian Restaurant	Coffee Shop	Market	Electronics Store	Restaur
73	Supertech Shopprix Mall, Sector 61	Gym	Pizza Place	Indian Restaurant	Vegetarian / Vegan Restaurant	Chinese Restaurant	Convenience Store	Fo
74	The Great India Place, Sector 38	Indian Restaurant	Coffee Shop	Chinese Restaurant	Clothing Store	Café	Fast Food Restaurant	Multip
75	Tulip Mall, Sector 48, Noida	Hotel	Gym / Fitness Center	Bakery	Food Court	Department Store	Dessert Shop	Di

76 rows × 11 columns

In [83]:
```
## Cluster Locality
## Run k-means to cluster the Locality into 5 clusters.

# set number of clusters
kclusters = 5

Noida_clustering = Noida_grouped.drop('Locality', 1)
```

```
# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(Noida_cluster
ing)

# check cluster labels generated for each row in the dataframe
kmeans.labels_[0:10]
kmeans.labels_.shape
```

Out[83]:
```
(76,)
```

In [84]:
```
# add clustering labels
Noida_merged = df_final.head(76)
Noida_merged['Cluster Labels'] = kmeans.labels_

# merge Noida_grouped with df_Chinese to add latitude/longitude for eac
h Locality
Noida_merged = Noida_merged.join(Locality_venues_sorted.set_index('Loca
lity'), on='Locality')

Noida_merged.head()
```

```
<ipython-input-84-55d391002207>:3: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  Noida_merged['Cluster Labels'] = kmeans.labels_
```

Out[84]:

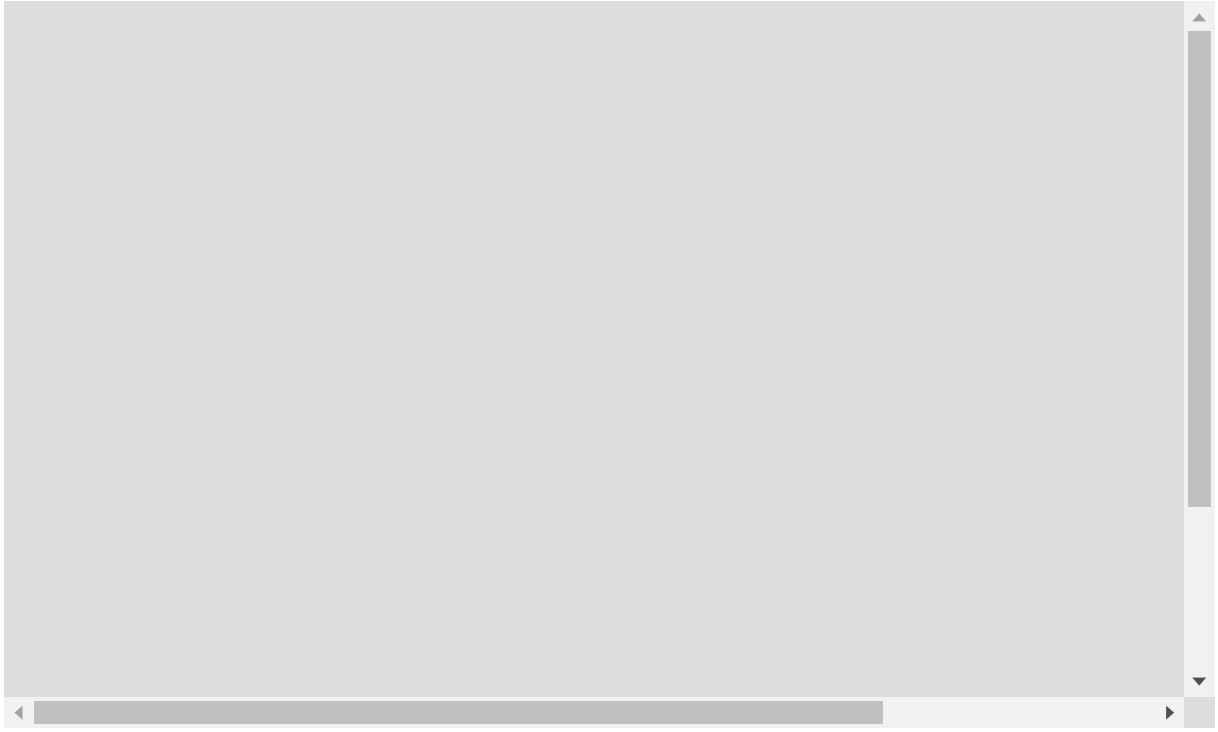
	Locality	Lat	Lng	No_of_Restaurant	Cusines	Agg_Rating	Comments	No
0	Ansal Plaza Mall, Greater Noida	28.462675	77.512780	4	Cafe, Chinese, Cafe, North Indian, Chinese, Cafe	3.150000	Average	

	Locality	Lat	Lng	No_of_Restaurant	Cusines	Agg_Rating	Comments	No
1	Brahmaputra Shopping Complex	28.570019	77.332701	4	North Indian, Chinese, Fast Food, Street Food,...	3.525000	Average, Good	
2	Centre Stage Mall, Sector 18	28.568185	77.323030	4	Asian, Continental, Italian, North Indian, Nor...	3.100000	Average, Good, Poor	
3	DLF Mall of India, Sector 18, Noida	28.567267	77.320868	29	North Indian, Mughlai, North Indian, European,...	3.858621	Average, Excellent, Good, Very Good	
4	Fortune Inn Grazia, Sector 27, Noida	28.577435	77.328359	2	North Indian, Mughlai, Chinese	3.300000	Average	

In [85]:
```
# create final map
map_clusters = folium.Map(location=[latitude, longitude], zoom_start=10
)

# set color scheme for the clusters
x = np.arange(kclusters)
ys = [i+x+(i*x)**2 for i in range(kclusters)]
#colors_array = cm.rainbow(np.linspace(0, 1, len(ys)))
#rainbow = [colors.rgb2hex(i) for i in colors_array]
colors = ['red', 'green', 'blue', 'yellow','orange']

# add markers to the map
markers_colors = []
```

```
for lat, lon, poi, cluster in zip(Noida_merged['Lat'], Noida_merged['Lng'], Noida_merged['Locality'], Noida_merged['Cluster Labels']):
    label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
    folium.CircleMarker(
        [lat, lon],
        radius=5,
        popup=label,
        color='black',
        fill=True,
        fill_color=colors[cluster],
        fill_opacity=0.7).add_to(map_clusters)

map_clusters
```

Out[85]:



In [86]:

```
## Examine Clusters
```

```
## Cluster 1
Noida_merged.loc[Noida_merged['Cluster Labels'] == 0, Noida_merged.colu
mns[[1] + list(range(5, Noida_merged.shape[1]))]]
```

Out[86]:

	Lat	Agg_Rating	Comments	No_of_Votes	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
1	28.570019	3.525000	Average, Good	1075	0	Hotel Bar	Indian Restaurant	South Indian Restaurant
2	28.568185	3.100000	Average, Good, Poor	1247	0	Indian Restaurant	Chinese Restaurant	Coffee Shop
3	28.567267	3.858621	Average, Excellent, Good, Very Good	8242	0	Café	Indian Restaurant	Coffee Shop
4	28.577435	3.300000	Average	94	0	Food Truck	Hotel	Pizza Place
5	28.567820	3.214286	Average, Good	747	0	Metro Station	Bakery	Café
...	...	...	...	...	...	...	...	...
70	28.528872	3.200000	Average	15	0	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Food Court
71	28.518748	3.125000	Average, Good, Poor	561	0	Coffee Shop	Ice Cream Shop	Golf Course
72	28.595794	3.100000	Average	8	0	Gym	Pizza Place	Indian Restaurant
73	28.585808	3.050000	Average, Good, Poor	1959	0	Stadium	Music Store	Indian Restaurant
74	28.597117	2.811765	Average, Good, Poor	1392	0	Gym	Pizza Place	Indian Restaurant

67 rows × 15 columns

In [87]:
```
## Examine Clusters

## Cluster 2
Noida_merged.loc[Noida_merged['Cluster Labels'] == 1, Noida_merged.colu
mns[[1] + list(range(5, Noida_merged.shape[1]))]]
```

Out[87]:

	Lat	Agg_Rating	Comments	No_of_Votes	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Mc Comm Ven
39	28.577396	2.850000	Average	61	1	Multiplex	Convenience Store	Departme Sto
47	28.558628	3.141667	Average, Good	663	1	Convenience Store	Hotel	Departme Sto

In [88]:
```
## Examine Clusters

## Cluster 3
Noida_merged.loc[Noida_merged['Cluster Labels'] == 2, Noida_merged.colu
mns[[1] + list(range(5, Noida_merged.shape[1]))]]
```

Out[88]:

	Lat	Agg_Rating	Comments	No_of_Votes	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	28.462675	3.150000	Average	116	2	Hotel	Supermarket	Indian Restaurant
16	28.603867	3.566667	Average, Very Good	285	2	Hotel	Indian Restaurant	Food Truck
64	28.621141	2.600000	Average	15	2	Fast Food Restaurant	Vegetarian / Vegan Restaurant	Hotel

In [89]:
```
## Examine Clusters

## Cluster 4
Noida_merged.loc[Noida_merged['Cluster Labels'] ==3 , Noida_merged.colu
mns[[1] + list(range(5, Noida_merged.shape[1]))]]
```

Out[89]:

	Lat	Agg_Rating	Comments	No_of_Votes	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4 C
63	28.616566	3.041304	Average, Good, Poor	2512	3	Sandwich Place	Multiplex	Coffee Shop	Dep

In [90]:
```
## Examine Clusters

## Cluster 5
Noida_merged.loc[Noida_merged['Cluster Labels'] == 4, Noida_merged.colu
mns[[1] + list(range(5, Noida_merged.shape[1]))]]
```

Out[90]:

	Lat	Agg_Rating	Comments	No_of_Votes	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4 C
49	28.553746	3.200000	Average	29	4	Hotel	Gym / Fitness Center	Bakery	
69	28.596372	3.200000	Average	9	4	Food Truck	Bed & Breakfast	Vegetarian / Vegan Restaurant	
75	28.567462	3.158065	Average, Good, Poor, Very Good	6377	4	Indian Restaurant	Coffee Shop	Chinese Restaurant	

4 Conclusion

- DLF Mall OF India are some of the best neighbourhoods for Chinese cuisine
- Sector 62 place have the highest number of Chinese Resturant.
- Sector 18 and Sector 32 are the best places for foodie.
- DLF Mall OF India, Gardens Galleria have best resturants in Noida. #### Cluster 1 and Cluster 5 : It is most recommended for family #### Cluster 2: It is most recommended for convenient Store and Entertainment #### Cluster 3: It is most recommended for restaurants #### Cluster 4: It is most recommended for the cafe and pizza

In [ ]: