Capstone Project: The Battle of Neighborhoods

September 2, 2020

1. Introduction

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,
Over last decades it is continuously grow because of the city's important role in government and
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, and French etc.

So as part of this project, we will list and visualize 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 Restaurant Market? - Which all areas have less number of restaurants? - Which is the best place to stay if I prefer Chinese Cuisine? - What places are having the best restaurant in New Delhi?

2. Data

For this project we need the following data: New Delhi Restaurants data

that contains list Locality, Restaurants name, rating along with their

Latitude and longitude.

Data source: <a href="https://www.kaggle.com/brianmathew/zomato-restaurants-data">

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.

*Data source: Fousquare API *
Description: By using this api we will get all the venues in each neighborhood.

3 Methodology

• 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 restaurant to find the best places. • Visualize the Ranking of neighborhoods using folium library(python)

3.1 Read the Zomato restaurant data from csv file

```
In [18]: df = pd.read_csv('zomato.csv',encoding='ISO-8859-1') df_india = df[df['Country Code'] == 1] ## New Delhi df_NDLS = df_india[df_india['City'] == 'New Delhi']

df_NDLS.reset_index(drop=True, inplace=True) df_NDLS.head()
```

3.2 Data Cleaning remove the unwanted columns and rows from dataset

```
In [32]: df_Res= df_NDLS[df_NDLS.Longitude !=0.000000][['Restaurant Name','Locality','Longitude

In [344]: df_Res = df_Res[df_Res['Aggregate rating'] !=0.0]
```

```
In [358]: df_Res.head()
```

3.2.1 created map to show the restaurant cluters

```
In [346]: 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['Longitu label = folium.Popup(Locality,
parse_html=True) folium.CircleMarker(

[latitude,
longitude], radius=5, popup=label, color='black', fill=True, fill_color=colors[c
luster], fill_opacity=0.7).add_to(New_Delhi_Rest)

        New_Delhi_Rest
/home/zettadevs/anaconda3/lib/python3.7/site-
packages/ipykernel_launcher.py:11: SettingWithCopy
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: http://pandas.pydata.org/pandas-
docs/stable/indexing.html
  # This is added back by InteractiveShellApp.init_path()
```

3.3 What places are have best restaurant in New Delhi?

```
In [575]: 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()
```

3.4 what places are have worst restaurants in New Delhi?

```
In [576]: 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()
```

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

```
In [577]: 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()
```

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

```
In [579]: import matplotlib.pyplot as plt 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()
```

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

```
In [580]: 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')['Rest
aurant
        plt.xlabel('Resturant Locality in New Delhi')
```

3.8 Which places are the best chinese resturants in New Delhi?

```
In [584]: import matplotlib.pyplot as plt plt.figure(figsize=(9,5), dpi = 100)
```

```
# title

plt.title('The best places for Chinese restaurant in New Delhi city') #On x-axis

        #giving a bar plot

df_Res[df_Res['Cuisines'].str.startswith('Chinese')].groupby('Locality')['Aggr
egate r
plt.xlabel('Resturant Locality in New Delhi') #On y-axis plt.ylabel('Rating of
resturants')

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

3.8.1 Data transformation

Based on Locality grouping the data

```
In [467]: 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
```

```
In [468]: df_final = df_final[df_final['Aggregate rating'] != 0.000000]
        df_final.columns =['Locality','Lat','Lng', 'No_of_Restaurant','Cusines',
'Agg_Rating'
        df_final.head()
```

3.9 Define Foursquare Credentials and Version

In [593]: *## Define Foursquare Credentials and Version* CLIENT_ID = 'ClientId' *# Foursquare ID*

CLIENT_SECRET = 'SecretID' *# Foursquare Secret* VERSION = '20200902' *# Foursquare API version*

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

Your credentails:
CLIENT_ID: XXXXXXXXXXXXXXXXXXXXXXX
CLIENT_SECRET:XXXXXXXXXXXXXXXXXXXX

3.10 create a function to repeat the same process to all the Locality in New Delhi

In [484]: *## create a function to repeat the same process to all the Locality in New Delhi*

3.100000
3.292308
3.275000
3.200000
3.033333

```
def getNearbyVenues(names, latitudes, longitudes, radius=500,LIMIT = 100):

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_secr
CLIENT_ID,

                CLIENT_SECRET,
                VERSION,
```

```
            lat,
            lng,
            radius,
            LIMIT)
        # make the GET request
        results = requests.get(url).json()["response"]['groups'][0]['items']
        # 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 for item
in venu nearby_venues.columns = ['Locality',

                  'Locality Latitude',
                  'Locality Longitude',
                  'Venue',
                  'Venue Latitude',
                  'Venue Longitude',
                  'Venue Category']
return(nearby_venues)
```

3.11 find the venues in all New Delhi Locality

```
In [485]: # find the venues in all New Delhi Locality new_Delhi_venues =
getNearbyVenues(names=df_final['Locality'],

                        latitudes=df_final['Lat'],
                        longitudes=df_final['Lng']
                       )
```

3.11.1 Print frequency of venues in each locality:

```
In [491]: New_Delhi_grouped.shape Out[491]: (239, 216) In [492]: ## print
each Locality along with the top 5 most common venues

    num_top_venues = 5
```

```
for hood in New_Delhi_grouped['Locality']: print("----"+hood+"----") temp =
New_Delhi_grouped[New_Delhi_grouped['Locality'] ==
hood].T.reset_index() 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=True).head(num_t print('\n')
```

4 Results

- **List of venues per locality**

```
Sheraton New Delhi Hotel, Saket
South Extension 1
South Extension 2
Southern Park Mall, Saket
Spark Mall, Kamla Nagar
Star City Mall, Mayur Vihar Phase 1
Subhash Nagar
Sunder Nagar
T3 Domestic Arrival, Aerocity
TDI Mall, Rajouri Garden
Tagore Garden
Taj Vivanta, Khan Market
The Ashok, Chanakyapuri
```

- **Number of venues per locality**

Bhikaji Cama Place	10
Chanakyapuri	5
Chander Nagar	1
Chandni Chowk	15
Chawri Bazar	6
Chhatarpur	5
Chittaranjan Park	4
City Centre Mall, Rohini	7
City Square Mall, Rajouri Garden	29

- **Types of Venues in each locality**

	Locality	ATM	Accessories Store	Afghan Restaurant
0	ARSS Mall, Paschim Vihar	0	0	0
1	ARSS Mall, Paschim Vihar	0	0	0
2	ARSS Mall, Paschim Vihar	0	0	0
3	ARSS Mall, Paschim Vihar	0	0	0
4	ARSS Mall, Paschim Vihar	0	0	0

	African Restaurant	Airport	Airport Food Court	Airport Lounge
0	0	0	0	0
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0

	Airport Service	Airport Terminal	American Restaurant	Arcade
0	0	0	0	0
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0

	Art Gallery	Art Museum	Arts & Crafts Store	Asian Restaurant
0	0	0	0	0
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0

- **Mean weight score for the venues and localities**

Locality	ATM
ARSS Mall, Paschim Vihar	0.111111
Adchini	0.000000
Aditya Mega Mall, Karkardooma	0.000000
Aerocity	0.000000
Aggarwal City Mall, Pitampura	0.000000
Aggarwal City Plaza, Rohini	0.000000
Alaknanda	0.000000
Ambience Mall, Vasant Kunj	0.000000
Anand Lok	0.000000
Anand Vihar	0.000000
Andaz Delhi, Aerocity	0.000000
Ansal Plaza Mall, Khel Gaon Marg	0.000000

- **Types of venues and their frequency**

```
----Rohini----
                venue  freq
0   Indian Restaurant  0.25
1        Shopping Mall  0.25
2           Multiplex  0.25
3                 Gym  0.25
4                 ATM  0.00

----Roseate House, Aerocity----
            venue  freq
0           Hotel  0.37
1             Spa  0.07
2   Shopping Mall  0.04
3   Train Station  0.04
4             Gym  0.04

----SDA----
                venue  freq
0                Café  0.23
1   Indian Restaurant  0.15
2         Coffee Shop  0.08
3               Hotel  0.08
4   Chinese Restaurant  0.08
```

5 Discussion

- Chanakyapuri, Pitampura, Safdarjung are some of the best neighborhoods for Chinese cui- sine.

- Pancsheel park, Nehru place have the best Chinese Resturant.

- Cannaught place, Rajouri garden, Malviya nagar are the best places for edible person.

- Greater kailash, Feroze shah road, Saket have best resturants in New Delhi. #### Cluster 1: It is most recommended for Indian Restaurants. #### Cluster 2: It is most recommended for Hotels and nightclub. #### Cluster 3 and Cluster 5: It is most recommended for Fast food. #### Cluster 4: It is most recommended for the cafe and pizza.

6 Conclusion

From the analysis that I made I have come to realize that New Delhi is a city with a large number of options of cuisines. Due to the diverse socio-economic stand here multiple cuisines will thirve.
This analysis will hopefully make it easier for people in the restaurant business in New Delhi.