We will start with importing libraries which we will use for EDA

In [6]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

Data is encoded in latin-1 so to read data we need to apply encodng of latin-1 to it

In [5]:
```
df=pd.read_csv('zomato.csv',encoding="latin-1")
df.head()
```

Out[5]:

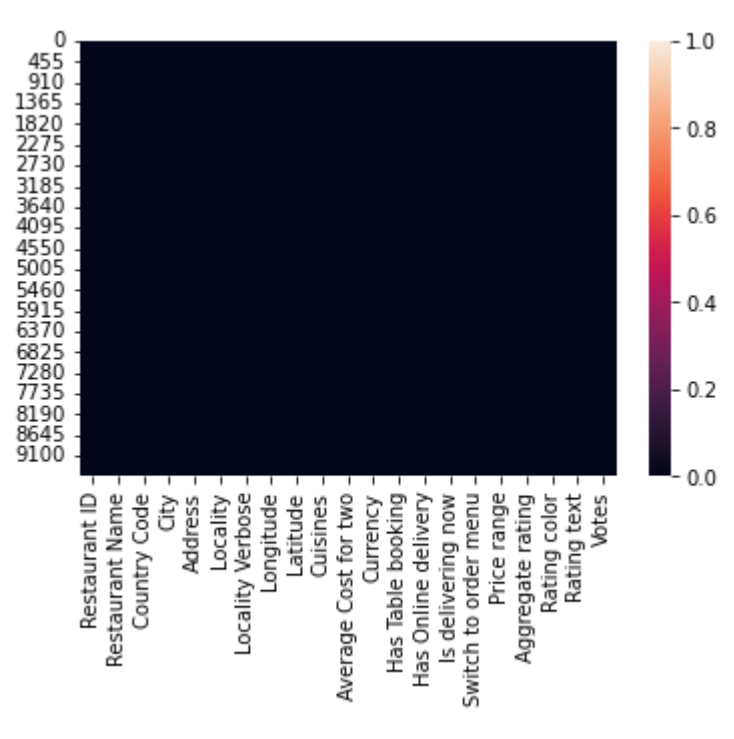
	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude	Cuisines	...	Currency	Has Table booking	Has Online delivery	Is delivering now	Switch to order menu	Price range	Aggregate rating	Ratin colo
0	6317637	Le Petit Souffle	162	Makati City	Third Floor, Century City Mall, Kalayaan Avenu...	Century City Mall, Poblacion, Makati City	Century City Mall, Poblacion, Makati City, Mak...	121.027535	14.565443	French, Japanese, Desserts	...	Botswana Pula(P)	Yes	No	No	No	3	4.8	Dar Gree
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	14.553708	Japanese	...	Botswana Pula(P)	Yes	No	No	No	3	4.5	Dar Gree
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	14.581404	Seafood, Asian, Filipino, Indian	...	Botswana Pula(P)	Yes	No	No	No	4	4.4	Gree
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	14.585318	Japanese, Sushi	...	Botswana Pula(P)	No	No	No	No	4	4.9	Dar Gree
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	14.584450	Japanese, Korean	...	Botswana Pula(P)	Yes	No	No	No	4	4.8	Dar Gree

5 rows × 21 columns

we will find out null values and if it can effect our analysis then we will fill or drop those values

In [9]:
```
sns.heatmap(df.isnull())
```

Out[9]: <AxesSubplot:>



we are not able to find null values from here we check again with pandas

In [8]:
```
df.isnull().sum()
```

Out[8]:
```
Restaurant ID           0
Restaurant Name         0
Country Code            0
City                    0
Address                 0
Locality                0
Locality Verbose        0
Longitude               0
Latitude                0
Cuisines                9
Average Cost for two    0
Currency                0
Has Table booking       0
Has Online delivery     0
Is delivering now       0
Switch to order menu    0
Price range             0
Aggregate rating        0
Rating color            0
Rating text             0
Votes                   0
dtype: int64
```

yes cuisines column have 9 null values , it will not effect our analysis

In [18]:
```
df1=pd.read_excel('Country-Code.xlsx')
df1.head()
```

Out[18]:

	Country Code	Country
0	1	India
1	14	Australia
2	30	Brazil
3	37	Canada
4	94	Indonesia

As we have data of 15 countries with their country codes ,to ease the process of EDA, we will merge both the dataset.

In [19]:
```
zomato=pd.merge(df,df1,on="Country Code",how="left")
zomato.head(2)
```

Out[19]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude	Cuisines	...	Has Table booking	Has Online delivery	Is delivering now	Switch to order menu	Price range	Aggregate rating	Rating color	Rating text	Votes	C
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	14.565443	French, Japanese, Desserts	...	Yes	No	No	No	3	4.8	Dark Green	Excellent	314	Ph
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	14.553708	Japanese	...	Yes	No	No	No	3	4.5	Dark Green	Excellent	591	Ph

2 rows × 22 columns

In [20]:
```
zomato.Country.value_counts()
```

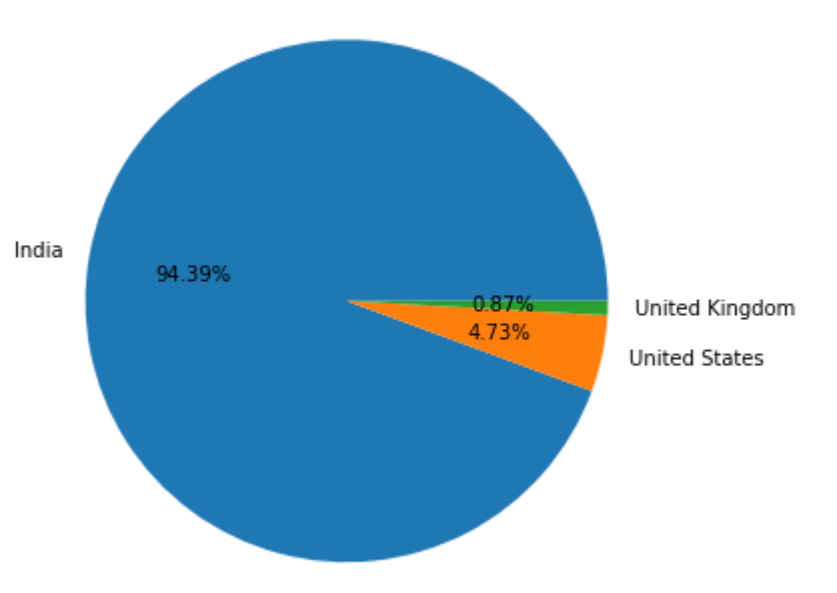
Out[20]:
```
India             8652
United States      434
United Kingdom      80
Brazil              60
UAE                 60
South Africa        60
New Zealand         40
Turkey              34
Australia           24
Phillipines         22
Indonesia           21
Singapore           20
Qatar               20
Sri Lanka           20
Canada               4
Name: Country, dtype: int64
```

In [32]:
```
country_names=zomato.Country.value_counts().index
```

In [34]:
```
country_val=zomato.Country.value_counts().values
```

In [77]:
```
## Pie Chart- Top 3 countries that uses zomato

plt.pie(country_val[:3],labels=country_names[:3],autopct='%1.2f%%')
plt.show()
plt.figure(figsize=(12,6))
```



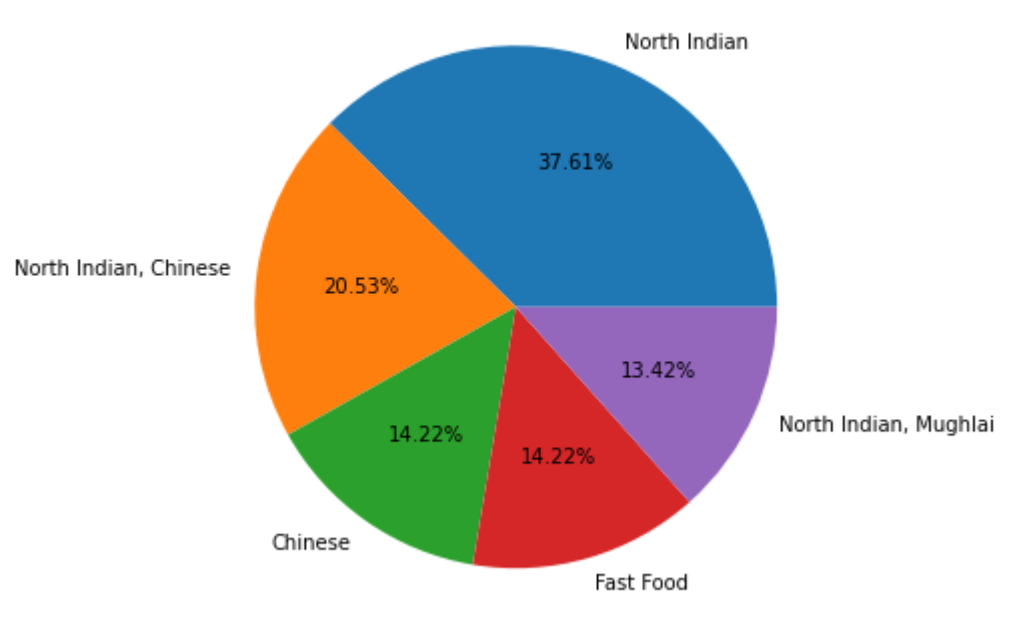
Out[77]: <Figure size 864x432 with 0 Axes>

<Figure size 864x432 with 0 Axes>

Observation: Zomato maximum transaction are from India followed by USA and then UK

In [39]:
```
Cuisines_values=zomato.Cuisines.value_counts().values
Cuisines_labels=zomato.Cuisines.value_counts().index
```

In [76]:
```
plt.pie(Cuisines_values[:5],labels=Cuisines_labels[:5],autopct='%1.2f%%' )
plt.show()
plt.figure(figsize=(12,6))
```



Out[76]: <Figure size 864x432 with 0 Axes>

<Figure size 864x432 with 0 Axes>

Observation: North Indian is the most demanded Cuisines.

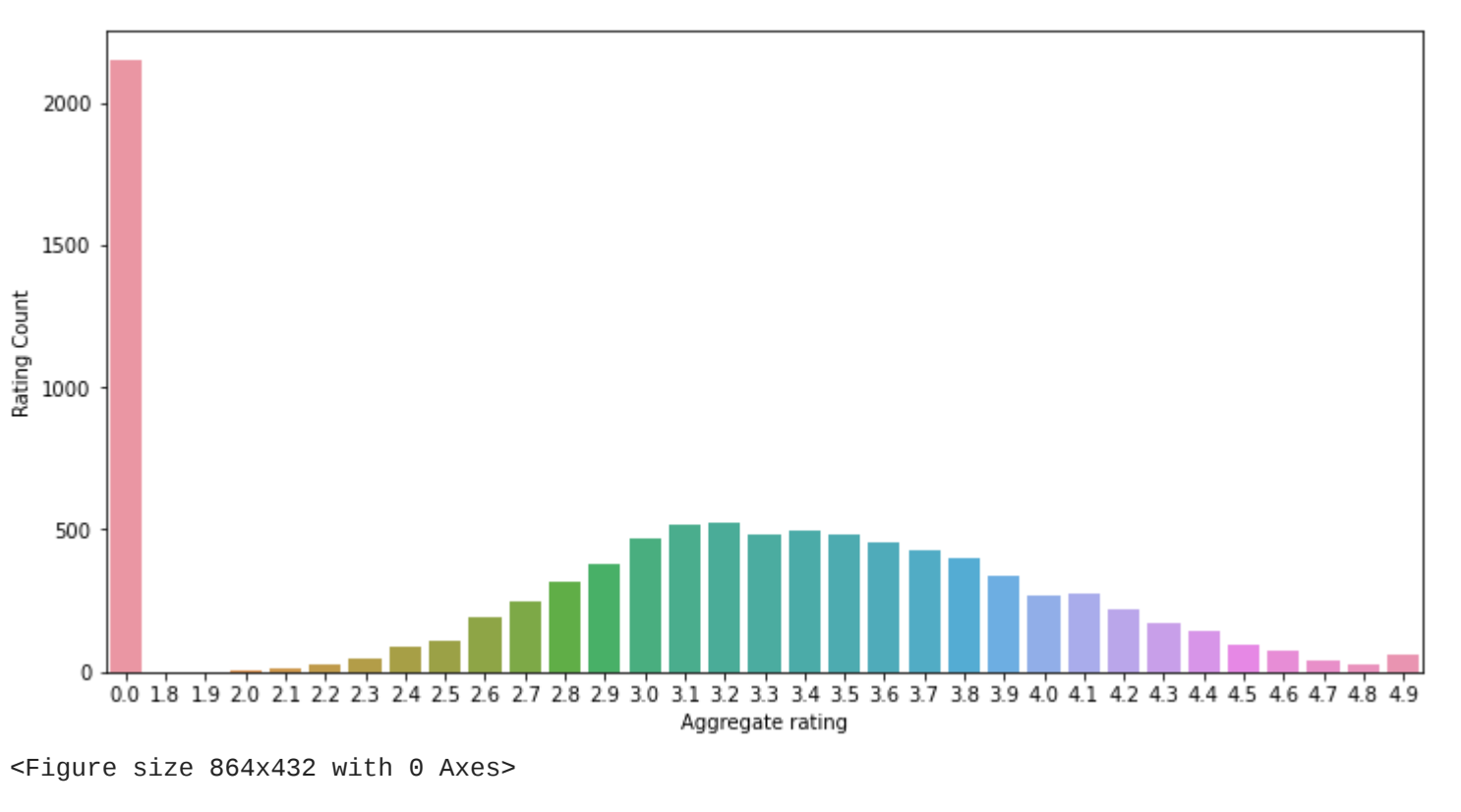
In [59]:
```
ratings = zomato.groupby(['Aggregate rating','Rating color','Rating text']).size().reset_index().rename(columns={0:'Rating Count'})
ratings.head()
```

Out[59]:

	Aggregate rating	Rating color	Rating text	Rating Count
0	0.0	White	Not rated	2148
1	1.8	Red	Poor	1
2	1.9	Red	Poor	2
3	2.0	Red	Poor	7
4	2.1	Red	Poor	15

In [75]:
```
sns.barplot(x="Aggregate rating",y="Rating Count",data=ratings)
plt.figure(figsize=(12,6))
```

Out[75]: <Figure size 864x432 with 0 Axes>



<Figure size 864x432 with 0 Axes>

Observation:

1) 0 denote Not rated , which is very high which is.

2)Maximum number of rating are between 2.8 to 3.8.

EDA on India's market

In [65]:
```
india=zomato[zomato['Country']=="India"]
india.head()
```

Out[65]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude	Cuisines	...	Has Table booking	Has Online delivery	Is delivering now	Switch to order menu	Price range	Aggregate rating	Rating color	Rating text	Votes	Country
624	3400025	Jahanpanah	1	Agra	E 23, Shopping Arcade, Sadar Bazaar, Agra Cant...	Agra Cantt	Agra Cantt, Agra	78.011544	27.161661	North Indian, Mughlai	...	No	No	No	No	3	3.9	Yellow	Good	140	India
625	3400341	Rangrezz Restaurant	1	Agra	E-20, Shopping Arcade, Sadar Bazaar, Agra Cant...	Agra Cantt	Agra Cantt, Agra	0.000000	0.000000	North Indian, Mughlai	...	No	No	No	No	2	3.5	Yellow	Good	71	India
626	3400005	Time2Eat - Mama Chicken	1	Agra	Main Market, Sadar Bazaar, Agra Cantt, Agra	Agra Cantt	Agra Cantt, Agra	78.011608	27.160832	North Indian	...	No	No	No	No	2	3.6	Yellow	Good	94	India
627	3400021	Chokho Jeeman Marwari Jain Bhojanalya	1	Agra	1/48, Delhi Gate, Station Road, Raja Mandi, Ci...	Civil Lines	Civil Lines, Agra	77.998092	27.195928	Rajasthani	...	No	No	No	No	2	4.0	Green	Very Good	87	India
628	3400017	Pinch Of Spice	1	Agra	23/453, Opposite Sanjay Cinema, Wazipura Road,...	Civil Lines	Civil Lines, Agra	78.007553	27.201725	North Indian, Chinese, Mughlai	...	No	No	No	No	3	4.2	Green	Very Good	177	India

5 rows × 22 columns

In [66]:
```
india.City.value_counts()
```

Out[66]:
```
New Delhi       5473
Gurgaon         1118
Noida           1080
Faridabad        251
Ghaziabad         25
Ahmedabad         21
Guwahati          21
Lucknow           21
Bhubaneshwar      21
Amritsar          21
Pune              20
Puducherry        20
Patna             20
Ludhiana          20
Ranchi            20
Surat             20
Vadodara          20
Nashik            20
Nagpur            20
Mysore            20
Mumbai            20
Varanasi          20
Mangalore         20
Agra              20
Kochi             20
Kolkata           20
Dehradun          20
Allahabad         20
Aurangabad        20
Bangalore         20
Bhopal            20
Chennai           20
Coimbatore        20
Goa               20
Indore            20
Jaipur            20
Kanpur            20
Vizag             20
Chandigarh        18
Hyderabad         18
Secunderabad       2
Panchkula          1
Mohali             1
Name: City, dtype: int64
```

Delhi has most number of resturants

In [70]:
```
delhi=zomato[(zomato.City=="New Delhi")]
delhi.head(2)
```

Out[70]:

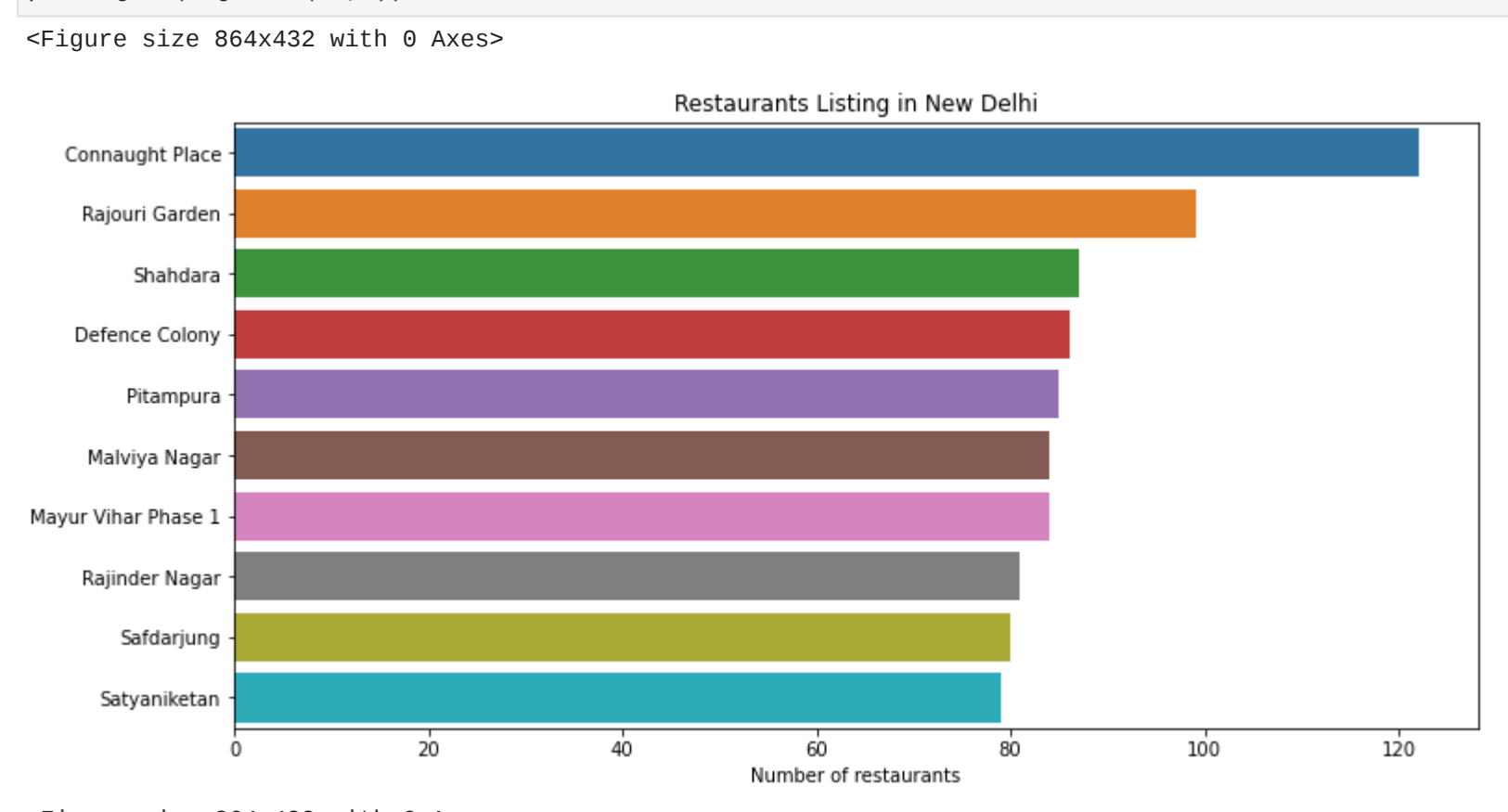
	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude	Cuisines	...	Has Table booking	Has Online delivery	Is delivering now	Switch to order menu	Price range	Aggregate rating	Rating color	Rating text	Votes	Cour
2560	18287358	Food Cloud	1	New Delhi	Aaya Nagar, New Delhi	Aaya Nagar	Aaya Nagar, New Delhi	0.000000	0.000000	Cuisine Varies	...	No	No	No	No	2	0.0	White	Not rated	2	In
2561	18216944	Burger.in	1	New Delhi	84, Near Honda Showroom, Adchini, New Delhi	Adchini	Adchini, New Delhi	77.196923	28.535382	Fast Food	...	No	Yes	No	No	1	3.2	Orange	Average	46	In

2 rows × 22 columns

In [72]:
```
sns.barplot(x=delhi.Locality.value_counts().head(10),y=delhi.Locality.value_counts().head(10).index)

plt.ylabel(None)
plt.xlabel("Number of restaurants")
plt.title("Restaurants Listing in New Delhi")
plt.figure(figsize=(12,6))
```

Out[72]: <Figure size 864x432 with 0 Axes>



<Figure size 864x432 with 0 Axes>

observation

Since max RESTAURANTS are in connaught place so we need more food delivery employee there.

Popular resturant of connaught place

In [86]:
```
cp=delhi[(delhi.Locality.isin(['Connaught Place'])) & (delhi["Rating text"].isin(['Excellent']))]
```

In [90]:
```
Best_rest=cp['Restaurant Name']
Best_rest
```

Out[90]:
```
3012               MOB Brewpub
3013          Naturals Ice Cream
3014    Zabardast Indian Kitchen
Name: Restaurant Name, dtype: object
```

These 3 are popular restaurant of CP with Excellent ratings.