exploratory data analysis(EDA)

In [1]:

```
import pandas as pd
import numpy as np
import seaborn as sns
from matplotlib import pyplot as plt
%matplotlib inline
```

In [2]:

```
#file_encoding= 'utf8'

#input_fd=open(input_file_and_path, encoding=file_encoding, errors='blackslashreplace')
#pd.read_csv(input_fd, "C:\Users\mujjj\AppData\Local\Temp\Rar$DIa10856.37308.csv" )
```

In [3]:

```
#df=pd.read_csv('C:\Users\mujjj\AppData\Local\Temp\Rar$DIa10856.37308.csv')
```

In [4]:

```
#import csv
#df=pd.read_csv('C:\Users\mujjj\AppData\Local\Temp\Rar$DIa10856.37308.csv')
```

In [5]:

```
df=pd.read_csv('zomato.CSV.csv', encoding='latin-1')
```

In [6]:

```
df.head()
```

Out[6]:

	Restaurant ID	**Restaurant Name**	**Country Code**	**City**	**Address**	**Locality**	**Locality Verbose**	**Longitude**	**Latitude**	**Cuisines**	**...**	**Currency**
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)
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)
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)
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)
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)

5 rows × 21 columns

In [7]:

```
df.columns
```

Out[7]:

```
Index(['Restaurant ID', 'Restaurant Name', 'Country Code', 'City', 'Address',
       'Locality', 'Locality Verbose', 'Longitude', 'Latitude', 'Cuisines',
       'Average Cost for two', 'Currency', 'Has Table booking',
       'Has Online delivery', 'Is delivering now', 'Switch to order menu',
       'Price range', 'Aggregate rating', 'Rating color', 'Rating text',
       'Votes'],
      dtype='object')
```

In [8]:

```
df.info
```

Out[8]:

```
<bound method DataFrame.info of       Restaurant ID          Restaurant Name  Country Code              City  
\
```

```
0           6317637          Le Petit Souffle           162       Makati City
1           6304287          Izakaya Kikufuji           162       Makati City
2           6300002    Heat - Edsa Shangri-La           162  Mandaluyong City
3           6318506                      Ooma           162  Mandaluyong City
4           6314302               Sambo Kojin           162  Mandaluyong City
...             ...                       ...           ...               ...
9546        5915730              NamlÛ± Gurme           208         ÛÁstanbul
9547        5908749             Ceviz AÛôacÛ±           208         ÛÁstanbul
9548        5915807                     Huqqa           208         ÛÁstanbul
9549        5916112             A^ô^ôk Kahve           208         ÛÁstanbul
9550        5927402  Walter's Coffee Roastery           208         ÛÁstanbul

                                                Address  \
0     Third Floor, Century City Mall, Kalayaan Avenu...   
1     Little Tokyo, 2277 Chino Roces Avenue, Legaspi...   
2     Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal...   
3     Third Floor, Mega Fashion Hall, SM Megamall, O...   
4     Third Floor, Mega Atrium, SM Megamall, Ortigas...   
...                                                 ...   
9546  Kemanke^ô Karamustafa Pa^ôa Mahallesi, RÛ±htÛ±...   
9547  Ko^ôuyolu Mahallesi, Muhittin íìstí_ndaÛô Cadd...   
9548  Kuruí_e^ôme Mahallesi, Muallim Naci Caddesi, N...   
9549  Kuruí_e^ôme Mahallesi, Muallim Naci Caddesi, N...   
9550  CafeaÛôa Mahallesi, BademaltÛ± Sokak, No 21/B,...   

                                        Locality  \
0      Century City Mall, Poblacion, Makati City   
1     Little Tokyo, Legaspi Village, Makati City   
2     Edsa Shangri-La, Ortigas, Mandaluyong City   
3         SM Megamall, Ortigas, Mandaluyong City   
4         SM Megamall, Ortigas, Mandaluyong City   
...                                          ...   
9546                                    Karakí_y   
9547                                   Ko^ôuyolu   
9548                                 Kuruí_e^ôme   
9549                                 Kuruí_e^ôme   
9550                                        Moda   

                                       Locality Verbose   Longitude  \
0     Century City Mall, Poblacion, Makati City, Mak...  121.027535   
1     Little Tokyo, Legaspi Village, Makati City, Ma...  121.014101   
2     Edsa Shangri-La, Ortigas, Mandaluyong City, Ma...  121.056831   
3     SM Megamall, Ortigas, Mandaluyong City, Mandal...  121.056475   
4     SM Megamall, Ortigas, Mandaluyong City, Mandal...  121.057508   
...                                                 ...         ...   
9546                                Karakí_y, ÛÁstanbul   28.977392   
9547                               Ko^ôuyolu, ÛÁstanbul   29.041297   
9548                             Kuruí_e^ôme, ÛÁstanbul   29.034640   
9549                             Kuruí_e^ôme, ÛÁstanbul   29.036019   
9550                                    Moda, ÛÁstanbul   29.026016   

       Latitude                          Cuisines  ...          Currency  \
0     14.565443        French, Japanese, Desserts  ...  Botswana Pula(P)   
1     14.553708                          Japanese  ...  Botswana Pula(P)   
2     14.581404  Seafood, Asian, Filipino, Indian  ...  Botswana Pula(P)   
3     14.585318                   Japanese, Sushi  ...  Botswana Pula(P)   
4     14.584450                  Japanese, Korean  ...  Botswana Pula(P)   
...         ...                               ...  ...               ...   
9546  41.022793                           Turkish  ...  Turkish Lira(TL)   
9547  41.009847   World Cuisine, Patisserie, Cafe  ...  Turkish Lira(TL)   
9548  41.055817            Italian, World Cuisine  ...  Turkish Lira(TL)   
9549  41.057979                   Restaurant Cafe  ...  Turkish Lira(TL)   
9550  40.984776                              Cafe  ...  Turkish Lira(TL)   

     Has Table booking Has Online delivery Is delivering now  \
0                  Yes                  No                No   
1                  Yes                  No                No   
2                  Yes                  No                No   
3                   No                  No                No   
4                  Yes                  No                No   
...                ...                 ...               ...   
9546                No                  No                No   
9547                No                  No                No   
9548                No                  No                No   
9549                No                  No                No   
9550                No                  No                No   

     Switch to order menu Price range  Aggregate rating  Rating color  \
0                      No           3               4.8    Dark Green   
1                      No           3               4.5    Dark Green   
2                      No           4               4.4         Green   
3                      No           4               4.9    Dark Green   
4                      No           4               4.8    Dark Green   
...                   ...         ...               ...           ...   
```

```
9546                  No           3              4.1        Green
9547                  No           3              4.2        Green
9548                  No           4              3.7       Yellow
9549                  No           4              4.0        Green
9550                  No           2              4.0        Green

     Rating text Votes
0      Excellent   314
1      Excellent   591
2      Very Good   270
3      Excellent   365
4      Excellent   229
...          ...   ...
9546   Very Good   788
9547   Very Good  1034
9548        Good   661
9549   Very Good   901
9550   Very Good   591

[9551 rows x 21 columns]>
```

we find in Data analysis

1. missing values
2. explore about the numerical variables
3. explore about categorical variables
4. finding relation between feature

In [9]:

```
df.isnull().sum()
```

Out[9]:

```
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
```

In [10]:

```
df.shape
```

Out[10]:

```
(9551, 21)
```

In [11]:

```
df.describe()
```

Out[11]:

	Restaurant ID	Country Code	Longitude	Latitude	Average Cost for two	Price range	Aggregate rating	Votes
count	9.551000e+03	9551.000000	9551.000000	9551.000000	9551.000000	9551.000000	9551.000000	9551.000000
mean	9.051128e+06	18.365616	64.126574	25.854381	1199.210763	1.804837	2.666370	156.909748
std	8.791521e+06	56.750546	41.467058	11.007935	16121.183073	0.905609	1.516378	430.169145
min	5.300000e+01	1.000000	-157.948486	-41.330428	0.000000	1.000000	0.000000	0.000000
25%	3.019625e+05	1.000000	77.081343	28.478713	250.000000	1.000000	2.500000	5.000000
50%	6.004089e+06	1.000000	77.191964	28.570469	400.000000	2.000000	3.200000	31.000000
75%	1.835229e+07	1.000000	77.282006	28.642758	700.000000	2.000000	3.700000	131.000000
max	1.850065e+07	216.000000	174.832089	55.976980	800000.000000	4.000000	4.900000	10934.000000

In [12]:

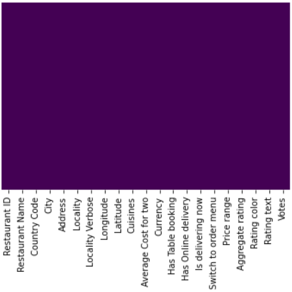
```
[features for features in df.columns if df[features].isnull().sum()>0]
```

Out[12]:

```
['Cuisines']
```

In [13]:
```
sns.heatmap(df.isnull(),yticklabels=False,cbar=False,cmap='viridis')
```

Out[13]:
```
<AxesSubplot:>
```



In [14]:
```
df_country=pd.read_excel('country-code.xlsx')
```

In [15]:
```
df_country.head()
```

Out[15]:

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

In [16]:
```
df.columns
```

Out[16]:
```
Index(['Restaurant ID', 'Restaurant Name', 'Country Code', 'City', 'Address',
       'Locality', 'Locality Verbose', 'Longitude', 'Latitude', 'Cuisines',
       'Average Cost for two', 'Currency', 'Has Table booking',
       'Has Online delivery', 'Is delivering now', 'Switch to order menu',
       'Price range', 'Aggregate rating', 'Rating color', 'Rating text',
       'Votes'],
      dtype='object')
```

In [17]:
```
final_df=pd.merge(df,df_country,on='Country Code', how='left')
```

In [18]:
```
final_df
```

Out[18]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude	Cuisines	...	H Tal booki
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	...	Y
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	...	Y
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	...	Y
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	...	I

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	...
...	...	...	...	...	...	...	...	...	...	...	...
9546	5915730	NamlÛ± Gurme	208	ÛÁstanbul	Kemankeˆô Karamustafa Paˆôa Mahallesi, RÛ±htÛ±...	Karakí_y	Karakí_y, ÛÁstanbul	28.977392	41.022793	Turkish	...
9547	5908749	Ceviz AÛôacÛ±	208	ÛÁstanbul	Koˆôuyolu Mahallesi, Muhittin íìstí_ndaÛô Cadd...	Koˆôuyolu	Koˆôuyolu, ÛÁstanbul	29.041297	41.009847	World Cuisine, Patisserie, Cafe	...
9548	5915807	Huqqa	208	ÛÁstanbul	Kuruí_eˆôme Mahallesi, Muallim Naci Caddesi, N...	Kuruí_eˆôme	Kuruí_eˆôme, ÛÁstanbul	29.034640	41.055817	Italian, World Cuisine	...
9549	5916112	Aˆôˆôk Kahve	208	ÛÁstanbul	Kuruí_eˆôme Mahallesi, Muallim Naci Caddesi, N...	Kuruí_eˆôme	Kuruí_eˆôme, ÛÁstanbul	29.036019	41.057979	Restaurant Cafe	...
9550	5927402	Walter's Coffee Roastery	208	ÛÁstanbul	CafeaÛôa Mahallesi, BademaltÛ± Sokak, No 21/B,...	Moda	Moda, ÛÁstanbul	29.026016	40.984776	Cafe	...

9551 rows × 22 columns

In [19]:

```
final_df.dtypes
```

Out[19]:

```
Restaurant ID             int64
Restaurant Name          object
Country Code              int64
City                     object
Address                  object
Locality                 object
Locality Verbose         object
Longitude               float64
Latitude                float64
Cuisines                 object
Average Cost for two      int64
Currency                 object
Has Table booking        object
Has Online delivery      object
Is delivering now        object
Switch to order menu     object
Price range               int64
Aggregate rating        float64
Rating color             object
Rating text              object
Votes                     int64
Country                  object
dtype: object
```

In [20]:

```
final_df.columns
```

Out[20]:

```
Index(['Restaurant ID', 'Restaurant Name', 'Country Code', 'City', 'Address',
       'Locality', 'Locality Verbose', 'Longitude', 'Latitude', 'Cuisines',
       'Average Cost for two', 'Currency', 'Has Table booking',
       'Has Online delivery', 'Is delivering now', 'Switch to order menu',
       'Price range', 'Aggregate rating', 'Rating color', 'Rating text',
       'Votes', 'Country'],
      dtype='object')
```

In [21]:

```
country_names=final_df.Country.value_counts().index
```

In [22]:

```
country_names
```

Out[22]:

```
Index(['India', 'United States', 'United Kingdom', 'Brazil', 'UAE',
       'South Africa', 'New Zealand', 'Turkey', 'Australia', 'Phillipines',
       'Indonesia', 'Singapore', 'Qatar', 'Sri Lanka', 'Canada'],
      dtype='object')
```

In [23]:

```
country_val=final_df.Country.value_counts().values
```

In [24]:

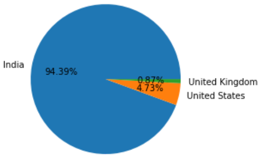
```
country_val
```

```
Out[24]: array([8652,  434,   80,   60,   60,   60,   40,   34,   24,   22,   21,
                  20,   20,   20,    4], dtype=int64)
```

```
In [25]: ## pie chart
```

```
In [26]: #from matplotlib import pyplt as plt
         #%matplotlib inline
         #for top three country
         plt.pie(country_val[:3],labels=country_names[:3],autopct="%1.2f%%")
```

```
Out[26]: ([<matplotlib.patches.Wedge at 0x24570850e80>,
           <matplotlib.patches.Wedge at 0x2456f874e50>,
           <matplotlib.patches.Wedge at 0x2456f82c6a0>],
          [Text(-1.0829742700952103, 0.19278674827836725, 'India'),
           Text(1.077281715838356, -0.22240527134123297, 'United States'),
           Text(1.0995865153823035, -0.03015783794312073, 'United Kingdom')],
          [Text(-0.590713238233751, 0.10515640815183668, '94.39%'),
           Text(0.5876082086391032, -0.12131196618612707, '4.73%'),
           Text(0.5997744629358018, -0.01644972978715676, '0.87%')])
```



```
In [27]: final_df.columns
```

```
Out[27]: Index(['Restaurant ID', 'Restaurant Name', 'Country Code', 'City', 'Address',
                'Locality', 'Locality Verbose', 'Longitude', 'Latitude', 'Cuisines',
                'Average Cost for two', 'Currency', 'Has Table booking',
                'Has Online delivery', 'Is delivering now', 'Switch to order menu',
                'Price range', 'Aggregate rating', 'Rating color', 'Rating text',
                'Votes', 'Country'],
               dtype='object')
```

```
In [28]: ratings=final_df.groupby(['Aggregate rating','Rating color','Rating text']).size().reset_index().rename(columr
```

```
In [29]: ratings
```

Out[29]:

	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
5	2.2	Red	Poor	27
6	2.3	Red	Poor	47
7	2.4	Red	Poor	87
8	2.5	Orange	Average	110
9	2.6	Orange	Average	191
10	2.7	Orange	Average	250
11	2.8	Orange	Average	315
12	2.9	Orange	Average	381
13	3.0	Orange	Average	468
14	3.1	Orange	Average	519
15	3.2	Orange	Average	522
16	3.3	Orange	Average	483
17	3.4	Orange	Average	498
18	3.5	Yellow	Good	480
19	3.6	Yellow	Good	458
20	3.7	Yellow	Good	427

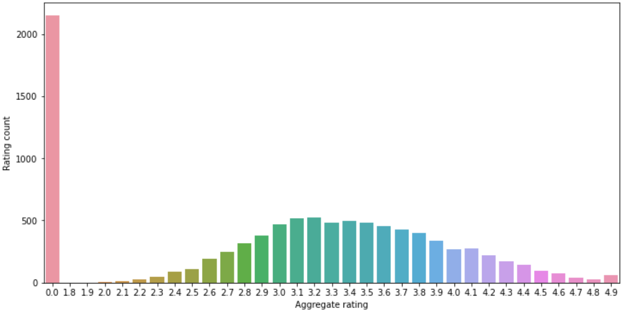
21	3.8	Yellow	Good	400
22	3.9	Yellow	Good	335
23	4.0	Green	Very Good	266
24	4.1	Green	Very Good	274
25	4.2	Green	Very Good	221
26	4.3	Green	Very Good	174
27	4.4	Green	Very Good	144
28	4.5	Dark Green	Excellent	95
29	4.6	Dark Green	Excellent	78
30	4.7	Dark Green	Excellent	42
31	4.8	Dark Green	Excellent	25
32	4.9	Dark Green	Excellent	61

In [30]:

```
import matplotlib
matplotlib.rcParams['figure.figsize']=(12,6)
sns.barplot(x="Aggregate rating",y="Rating count",data=ratings)
```

Out[30]:

```
<AxesSubplot:xlabel='Aggregate rating', ylabel='Rating count'>
```

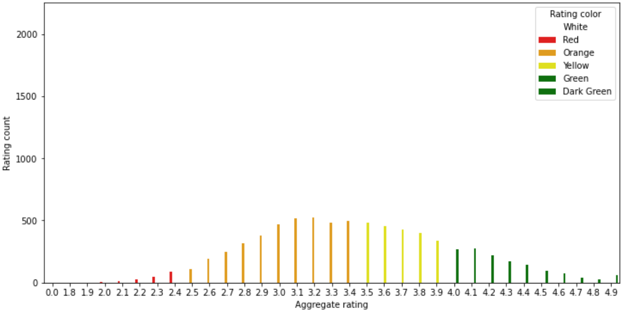


In [31]:

```
sns.barplot(x="Aggregate rating",y="Rating count",hue='Rating color',data=ratings,palette=['white','red','orar
```

Out[31]:

```
<AxesSubplot:xlabel='Aggregate rating', ylabel='Rating count'>
```



Observation

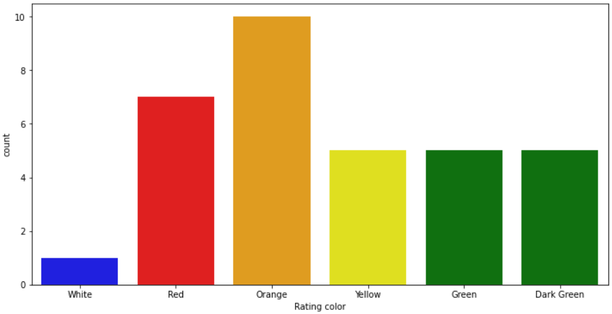
1. Not rated Is very high
2. 2.1 to 2.5 rating is maximum

In [35]:

```
sns.countplot(x="Rating color",data=ratings,palette=['blue','red','orange','yellow','green','green'])
```

Out[35]:

```
<AxesSubplot:xlabel='Rating color', ylabel='count'>
```



In [34]:

```
ratings
```

Out[34]:

	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
5	2.2	Red	Poor	27
6	2.3	Red	Poor	47
7	2.4	Red	Poor	87
8	2.5	Orange	Average	110
9	2.6	Orange	Average	191
10	2.7	Orange	Average	250
11	2.8	Orange	Average	315
12	2.9	Orange	Average	381
13	3.0	Orange	Average	468
14	3.1	Orange	Average	519
15	3.2	Orange	Average	522
16	3.3	Orange	Average	483
17	3.4	Orange	Average	498
18	3.5	Yellow	Good	480
19	3.6	Yellow	Good	458
20	3.7	Yellow	Good	427
21	3.8	Yellow	Good	400
22	3.9	Yellow	Good	335
23	4.0	Green	Very Good	266
24	4.1	Green	Very Good	274
25	4.2	Green	Very Good	221
26	4.3	Green	Very Good	174
27	4.4	Green	Very Good	144
28	4.5	Dark Green	Excellent	95
29	4.6	Dark Green	Excellent	78
30	4.7	Dark Green	Excellent	42
31	4.8	Dark Green	Excellent	25
32	4.9	Dark Green	Excellent	61

In [38]:

```python
final_df.columns
```

Out[38]:

```
Index(['Restaurant ID', 'Restaurant Name', 'Country Code', 'City', 'Address',
       'Locality', 'Locality Verbose', 'Longitude', 'Latitude', 'Cuisines',
       'Average Cost for two', 'Currency', 'Has Table booking',
       'Has Online delivery', 'Is delivering now', 'Switch to order menu',
       'Price range', 'Aggregate rating', 'Rating color', 'Rating text',
       'Votes', 'Country'],
      dtype='object')
```

In [43]:

```python
final_df.groupby(['Aggregate rating','Country']).size().reset_index().head(5)
```

Out[43]:

	Aggregate rating	Country	0
0	0.0	Brazil	5
1	0.0	India	2139
2	0.0	United Kingdom	1
3	0.0	United States	3
4	1.8	India	1

In [46]:

```python
#find which currancy is used in which country
final_df.groupby(['Country','Currency']).size().reset_index()
```

Out[46]:

	Country	Currency	0
0	Australia	Dollar($)	24
1	Brazil	Brazilian Real(R$)	60
2	Canada	Dollar($)	4
3	India	Indian Rupees(Rs.)	8652
4	Indonesia	Indonesian Rupiah(IDR)	21
5	New Zealand	NewZealand($)	40
6	Phillipines	Botswana Pula(P)	22
7	Qatar	Qatari Rial(QR)	20
8	Singapore	Dollar($)	20
9	South Africa	Rand(R)	60
10	Sri Lanka	Sri Lankan Rupee(LKR)	20
11	Turkey	Turkish Lira(TL)	34
12	UAE	Emirati Diram(AED)	60
13	United Kingdom	Pounds(£)	80
14	United States	Dollar($)	434

In [47]:

```python
#which country do have online deliveries
final_df.columns
```

Out[47]:

```
Index(['Restaurant ID', 'Restaurant Name', 'Country Code', 'City', 'Address',
       'Locality', 'Locality Verbose', 'Longitude', 'Latitude', 'Cuisines',
       'Average Cost for two', 'Currency', 'Has Table booking',
       'Has Online delivery', 'Is delivering now', 'Switch to order menu',
       'Price range', 'Aggregate rating', 'Rating color', 'Rating text',
       'Votes', 'Country'],
      dtype='object')
```

In [48]:

```python
final_df.groupby(['Country','Has Online delivery']).size().reset_index()
```

Out[48]:

	Country	Has Online delivery	0
0	Australia	No	24
1	Brazil	No	60
2	Canada	No	4
3	India	No	6229
4	India	Yes	2423
5	Indonesia	No	21
6	New Zealand	No	40
7	Phillipines	No	22
8	Qatar	No	20
9	Singapore	No	20

10	South Africa	No	60
11	Sri Lanka	No	20
12	Turkey	No	34
13	UAE	No	32
14	UAE	Yes	28
15	United Kingdom	No	80
16	United States	No	434

In [55]:

```
final_df[final_df['Has Online delivery']=='yes'].Country.value_counts()
```

Out[55]:

```
Series([], Name: Country, dtype: int64)
```

OBSERVATION

1. online delivery is in UAE and India

In [64]:

```
 #creating pie chart for cities distribution
```

In [67]:

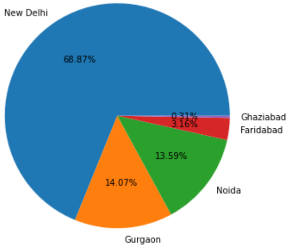
```
    city_values=final_df.City.value_counts().values
    city_labels=final_df.City.value_counts().index
```

In [69]:

```
plt.pie(city_values[:5],labels=city_labels[:5],autopct='%1.2f%%')
```

Out[69]:

```
([<matplotlib.patches.Wedge at 0x24577b020d0>,
  <matplotlib.patches.Wedge at 0x24577b027f0>,
  <matplotlib.patches.Wedge at 0x24577b02f10>,
  <matplotlib.patches.Wedge at 0x24577b0d670>,
  <matplotlib.patches.Wedge at 0x24577b0da60>],
 [Text(-0.6145352824185932, 0.9123301960708633, 'New Delhi'),
  Text(0.0623675251198054, -1.0982305276263407, 'Gurgaon'),
  Text(0.8789045225625368, -0.661458116753524б, 'Noida'),
  Text(1.0922218418223437, -0.13058119407559224, 'Faridabad'),
  Text(1.099946280005612, -0.010871113182029924, 'Ghaziabad')],
 [Text(-0.3352010631374145, 0.497634652402289, '68.87%'),
  Text(0.0340186500653484, -0.5990348332507311, '14.07%'),
  Text(0.47940246685229276, -0.36079533641101336, '13.59%'),
  Text(0.5957573682667329, -0.07122610585941394, '3.16%'),
  Text(0.5999706981848791, -0.005929698099289049, '0.31%')])
```



In [70]:

```
final_df.columns
```

Out[70]:

```
Index(['Restaurant ID', 'Restaurant Name', 'Country Code', 'City', 'Address',
       'Locality', 'Locality Verbose', 'Longitude', 'Latitude', 'Cuisines',
       'Average Cost for two', 'Currency', 'Has Table booking',
       'Has Online delivery', 'Is delivering now', 'Switch to order menu',
       'Price range', 'Aggregate rating', 'Rating color', 'Rating text',
       'Votes', 'Country'],
      dtype='object')
```

In [71]:

```
# find top 10 cuisines
```

In [79]:

```
final_df.Cuisines.value_counts().head(10)
```

Out[79]:

```
North Indian                      936
North Indian, Chinese             511
Chinese                           354
Fast Food                         354
North Indian, Mughlai             334
Cafe                              299
Bakery                            218
North Indian, Mughlai, Chinese    197
Bakery, Desserts                  170
Street Food                       149
Name: Cuisines, dtype: int64
```

In [ ]: