Import libraries

In [3]:

```python
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
import seaborn as sns
from scipy import stats
import plotly.express as px
import plotly.graph_objects as go
import os,sys
from scipy.stats import norm, skew
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error
from sklearn.linear_model import LinearRegression
import warnings
warnings.filterwarnings('ignore')
```

import data

In [18]:

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

Out[18]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose
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...
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...
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...
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...
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...
...	...	...	...	...	...	...	...
9546	5915730	NamlÛ± Gurme	208	ÛÁstanbul	Kemanke ô Karamustafa Pa ôa Mahallesi, RÛ±htÛ±...	Karakí_y	Karakí_y, ÛÁstanbul
9547	5908749	Ceviz AÛôacÛ±	208	ÛÁstanbul	Ko ôuyolu Mahallesi, Muhittin îîstí_ndaÛô Cadd...	Ko ôuyolu	Ko ôuyolu, ÛÁstanbul
9548	5915807	Huqqa	208	ÛÁstanbul	Kuruí_e ôme Mahallesi, Muallim Naci Caddesi, N...	Kuruí_e ôme	Kuruí_e ôme, ÛÁstanbul
9549	5916112	A ô ôk Kahve	208	ÛÁstanbul	Kuruí_e ôme Mahallesi, Muallim Naci Caddesi, N...	Kuruí_e ôme	Kuruí_e ôme, ÛÁstanbul

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose
9550	5927402	Walter's Coffee Roastery	208	ÛÁstanbul	CafeaÛôa Mahallesi, BademaltÛ± Sokak, No 21/B,...	Moda	Moda, ÛÁstanbul

9551 rows × 21 columns

Data Understanding

In [19]:

```
df.head()
```

Out[19]:

	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 [20]:

```
df.shape
```

Out[20]:

```
(9551, 21)
```

In [21]:

```python
df.isna().sum()
```

Out[21]:

```
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 [22]:

```python
nan_values = df.isna()
nan_columns = nan_values.any()

columns_with_nan = df.columns[nan_columns].tolist()
print(columns_with_nan)
```

```
['Cuisines']
```

In [25]:

```
!pip install openpyxl
df1 = pd.read_excel('Country-Code.xlsx')
df1.head()
```

Defaulting to user installation because normal site-packages is not writeabl
e
Requirement already satisfied: openpyxl in c:\programdata\anaconda3\lib\site
-packages (3.0.9)
Requirement already satisfied: et-xmlfile in c:\programdata\anaconda3\lib\si
te-packages (from openpyxl) (1.1.0)

Out[25]:

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

In [26]:

```
df2 = pd.merge(df,df1,on='Country Code',how='left')
df2.head(2)
```

Out[26]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitud
0	6317637	Le Petit Souffle	162	Makati City	Third Floor, Century City Mall, Kalayaan Avenu...	Century City Mall, Poblacion, Makati City	Century City Mall, Poblacion, Makati City, Mak...	121.027535	14.56544
1	6304287	Izakaya Kikufuji	162	Makati City	Little Tokyo, 2277 Chino Roces Avenue, Legaspi...	Little Tokyo, Legaspi Village, Makati City	Little Tokyo, Legaspi Village, Makati City, Ma...	121.014101	14.55370

2 rows × 22 columns
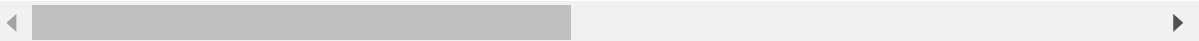
In [27]:

```python
print('List of counteris the survey is spread accross - ')
for x in pd.unique(df2.Country): print(x)
print()
print('Total number to country', len(pd.unique(df2.Country)))
```

```
List of counteris the survey is spread accross -
Phillipines
Brazil
United States
Australia
Canada
Singapore
UAE
India
Indonesia
New Zealand
United Kingdom
Qatar
South Africa
Sri Lanka
Turkey

Total number to country 15
```

In [29]:

```python
from plotly.offline import init_notebook_mode, plot, iplot

labels = list(df2.Country.value_counts().index)
values = list(df2.Country.value_counts().values)

fig = {
    "data":[
        {
            "labels" : labels,
            "values" : values,
            "hoverinfo" : 'label+percent',
            "domain": {"x": [0, .9]},
            "hole" : 0.6,
            "type" : "pie",
            "rotation":120,
        },
    ],
    "layout": {
        "title" : "Zomato's Presence around the World",
        "annotations": [
            {

                "font": {"size":20},
                "showarrow": True,
                "text": "Countries",
                "x":0.2,
                "y":0.9,
            },
        ]
    }
}

iplot(fig)
```

In [30]:

```
df3 = df2.groupby(['Aggregate rating','Rating color', 'Rating text']).size().reset_index().
df3
```

Out[30]:

	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 [31]:

```python
fig = px.bar(df3,
             x='Rating color',
             #marginal='box',
             color='Rating text',
         y='Rating Count',
             color_discrete_sequence=['Grey', 'Red','Orange','Yellow','Green',"Darkgr
             title='')
fig.update_layout(bargap=0.1)
fig.show()
```

In [32]:

```python
No_rating = df2[df2['Rating color']=='White'].groupby('Country').size().reset_index().renam
No_rating
```

Out[32]:

	Country	Rating Count
0	Brazil	5
1	India	2139
2	United Kingdom	1
3	United States	3

In [33]:

```
country_currency = df2[['Country','Currency']].groupby(['Country','Currency']).size().reset
country_currency.sort_values('Currency').reset_index(drop=True)
```

Out[33]:

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

In [34]:

```python
fig = px.pie(df2, names="Has Online delivery", title='Pie chart of online delivery')
fig.show()
```

In [35]:

```python
from plotly.offline import init_notebook_mode, plot, iplot
import plotly.graph_objs as go
plt.figure(figsize=(12,6))
# import plotly.plotly as py

labels = list(df2.City.value_counts().head(20).index)
values = list(df2.City.value_counts().head(20).values)

fig = {
    "data":[
        {
            "labels" : labels,
            "values" : values,
            "hoverinfo" : 'label+percent',
            "domain": {"x": [0, .9]},
            "hole" : 0.6,
            "type" : "pie",
            "rotation":120,
        },
    ],
    "layout": {
        "title" : "Zomato's Presence Citywise",
        "annotations": [
            {
                "font": {"size":20},
                "showarrow": True,
                "text": "Cities",
                "x":0.2,
                "y":0.9,
            },
        ]
    }
}
iplot(fig);
```

```
<Figure size 864x432 with 0 Axes>
```

In [36]:

```python
k = df2[(df2.Country == 'India')]
k.head(15)
```

Out[36]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	L
624	3400025	Jahanpanah	1	Agra	E 23, Shopping Arcade, Sadar Bazaar, Agra Cant...	Agra Cantt	Agra Cantt, Agra	78.011544	27.
625	3400341	Rangrezz Restaurant	1	Agra	E-20, Shopping Arcade, Sadar Bazaar, Agra Cant...	Agra Cantt	Agra Cantt, Agra	0.000000	0.
626	3400005	Time2Eat - Mama Chicken	1	Agra	Main Market, Sadar Bazaar, Agra Cantt, Agra	Agra Cantt	Agra Cantt, Agra	78.011608	27.
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.
628	3400017	Pinch Of Spice	1	Agra	23/453, Opposite Sanjay Cinema, Wazipura Road,...	Civil Lines	Civil Lines, Agra	78.007553	27.
629	3400325	MoMo Cafe	1	Agra	Courtyard by Marriott Agra, Phase 2, Fatehabad...	Courtyard by Marriott Agra, Tajganj	Courtyard by Marriott Agra, Tajganj, Agra	0.000000	0.
630	3400059	Peshawri - ITC Mughal	1	Agra	ITC Mughal, Fatehabad Road, Tajganj, Agra	ITC Mughal, Tajganj	ITC Mughal, Tajganj, Agra	78.044095	27.
631	3400060	Taj Bano - ITC Mughal	1	Agra	ITC Mughal, Fatehabad Road, Tajganj, Agra	ITC Mughal, Tajganj	ITC Mughal, Tajganj, Agra	78.044095	27.
632	3400348	G Thal	1	Agra	3/20, KPS Tower, Near Tulsi Talkies, Bypass Ro...	Khandari	Khandari, Agra	0.000000	0.
633	3400072	Dawat-e-Nawab - Radisson Blu	1	Agra	Radisson Blu, Taj East Gate Road, Tajganj, Agra	Radisson Blu, Tajganj	Radisson Blu, Tajganj, Agra	78.057044	27.
634	3400073	The Latitude - Radisson Blu	1	Agra	Radisson Blu, Taj East Gate Road, Tajganj, Agra	Radisson Blu, Tajganj	Radisson Blu, Tajganj, Agra	78.057044	27.

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	L
635	3400019	Dasaprakash Restaurant	1	Agra	Meher Cinema Complex, Gwalior Road, Rakabganj,...	Rakabganj	Rakabganj, Agra	78.015053	27.
636	3400033	The Charcoal Chimney	1	Agra	Hotel Samovar, Fatehabad Road, Tajganj, Agra	Tajganj	Tajganj, Agra	78.042990	27.
637	3400346	Sheroes Hangout	1	Agra	Opposite The Gateway Hotel, Fatehabad Road, Ta...	Tajganj	Tajganj, Agra	78.040165	27.
638	3400350	Bon Barbecue	1	Agra	Parador Hotel, 3A-3B, Phase 1, Fatehabad Road,...	Tajganj	Tajganj, Agra	0.000000	0.

15 rows × 22 columns
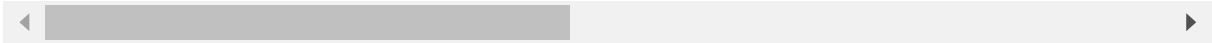
In [37]:

```python
fig = px.bar(k,
                x=k.Locality,
                #marginal='box',
                color=k['Has Table booking'],
            #y=k.Locality.value_counts().head(10).index,
                #color_discrete_sequence=['Grey',  'Red','Orange','Yellow','Green',"Darkg
                title='')
fig.update_layout(bargap=0.1)
fig.show()
```

In [38]:

```
ConnaughtPlace = k[(k.Locality.isin(['Park Street Area'])) & (k['Rating text'].isin(['Excel

ConnaughtPlace = ConnaughtPlace.Cuisines.value_counts().reset_index()

## Extracing all the cuisens in a single list
cuisien = []
for x in ConnaughtPlace['index']:
  cuisien.append(x)

# cuisien = '[%s]'%', '.join(map(str, cuisien))
cuisien
```

Out[38]:

```
['North Indian, Chinese', 'Chinese, North Indian', 'Continental, North India
n']
```

In [39]:

```
k2=k.nlargest(25, 'Average Cost for two')
k2
```

Out[39]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitu
7543	2701	Orient Express - Taj Palace Hotel	1	New Delhi	Taj Palace Hotel, Diplomatic Enclave, Chanakya...	The Taj Palace Hotel, Chanakyapuri	The Taj Palace Hotel, Chanakyapuri, New Delhi	77.170087	28.5950
4088	309548	Tian - Asian Cuisine Studio - ITC Maurya	1	New Delhi	ITC Maurya, Diplomatic Enclave, Chanakyapuri, ...	ITC Maurya, Chanakyapuri	ITC Maurya, Chanakyapuri, New Delhi	77.173455	28.5973
4087	2742	Bukhara - ITC	1	New Delhi	ITC Maurya, Chanakyapuri,	ITC Maurya, Chanakyapuri	ITC Maurya, Chanakyapuri,	77.173724	28.5974

In [40]:

```python
import plotly.express as px


fig = px.scatter_mapbox(k2, lat="Latitude", lon="Longitude", hover_name="City", hover_data=


                        color_discrete_sequence=["fuchsia"], zoom=10, height=300)
fig.update_layout(mapbox_style="open-street-map")
fig.update_layout(margin={"r":0,"t":0,"l":0,"b":0})
fig.update_layout(title='Top 25 costly Resturants Location',
                  autosize=True,
                  #hovermode='closest',
                  showlegend=False)
fig.update_layout(
    autosize=False,
    width=800,
    height=500,)
```

In [48]:

```python
ax = px.bar(k2, x='Restaurant Name',
           y='Average Cost for two',
           hover_data=['Has Online delivery'] ,
           title = 'Top 25 costly restaurant')
ax.show()
```

In [49]:

```python
fig=px.scatter(k,
         x="Average Cost for two",
         y="Aggregate rating",
         color="Price range",
         #size='petal_length',
         #hover_data=['petal_width']
         )
fig.show()
```

In [ ]:

```python
fig=px.scatter(k,
         x="Average Cost for two",
         y="Aggregate rating",
         color="Price range",
         #size='petal_length',
         #hover_data=['petal_width']
```