Task 2: Descriptive Analysis

- Calculate basic statistical measures (mean,median, standard deviation, etc.) for numerical columns.
- Explore the distribution of categoricalvariables like "Country Code," "City," and"Cuisines."
- Identify the top cuisines and cities with the highest number of restaurants.

In [2]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [3]:
```
rdata_df=pd.read_csv('Dataset.csv')
rdata_df
```

Out[3]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality
0	6317637	Le Petit Souffle	162	Makati City	Third Floor, Century City Mall, Kalayaan Avenu...	Century City Mall, Poblacion, Makati City
1	6304287	Izakaya Kikufuji	162	Makati City	Little Tokyo, 2277 Chino Roces Avenue, Legaspi...	Little Tokyo, Legaspi Village, Makati City
2	6300002	Heat - Edsa Shangri-La	162	Mandaluyong City	Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal...	Edsa Shangri-La, Ortigas, Mandaluyong City
3	6318506	Ooma	162	Mandaluyong City	Third Floor, Mega Fashion Hall, SM Megamall, O...	SM Megamall, Ortigas, Mandaluyong City
4	6314302	Sambo Kojin	162	Mandaluyong City	Third Floor, Mega Atrium, SM Megamall, Ortigas...	SM Megamall, Ortigas, Mandaluyong City
...	...	...	...	...	...	...
9546	5915730	Naml\ Gurme	208	��stanbul	Kemanke�� Karamustafa Pa��a Mahallesi, Rlhtlm ...	Karak�_y
9547	5908749	Ceviz A��acl	208	��stanbul	Ko��uyolu Mahallesi, Muhittin ��st�_nda�� Cadd...	Ko��uyolu
9548	5915807	Huqqa	208	��stanbul	Kuru�_e��me Mahallesi, Muallim Naci Caddesi, N...	Kuru�_e��me
9549	5916112	A���k Kahve	208	��stanbul	Kuru�_e��me Mahallesi, Muallim Naci Caddesi, N...	Kuru�_e��me
9550	5927402	Walter's Coffee Roastery	208	��stanbul	Cafea��a Mahallesi, Bademaltl Sokak, No 21/B, ...	Moda

9551 rows × 21 columns

basic statistical measures

In [4]:
```
rdata_df.describe()
```

Out[4]:

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

Exploring the categoricalvariables

In [5]:
```
categorical_varibles=rdata_df.select_dtypes(include='object').columns
```

In [6]:
```
categorical_varibles
```

Out[6]:
```
Index(['Restaurant Name', 'City', 'Address', 'Locality', 'Locality Verbose',
       'Cuisines', 'Currency', 'Has Table booking', 'Has Online delivery',
       'Is delivering now', 'Switch to order menu', 'Rating color',
       'Rating text'],
      dtype='object')
```

In [7]:
```
numerical_varibles=rdata_df.select_dtypes(exclude='object').columns
numerical_varibles
```

Out[7]:
```
Index(['Restaurant ID', 'Country Code', 'Longitude', 'Latitude',
       'Average Cost for two', 'Price range', 'Aggregate rating', 'Votes'],
      dtype='object')
```

Identify the top cuisines and cities with the highest number of restaurants.

In [11]:
```
rdata_df.head(), rdata_df.columns
# here iam taking the all the columns top 10 rows data
```

Out[11]: (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

                                              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...

                                     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

                                    Locality Verbose   Longitude   Latitude \
0 Century City Mall, Poblacion, Makati City, Mak... 121.027535 14.565443
1 Little Tokyo, Legaspi Village, Makati City, Ma... 121.014101 14.553708
2 Edsa Shangri-La, Ortigas, Mandaluyong City, Ma... 121.056831 14.581404
3 SM Megamall, Ortigas, Mandaluyong City, Mandal... 121.056475 14.585318
4 SM Megamall, Ortigas, Mandaluyong City, Mandal... 121.057508 14.584450

                           Cuisines ...   Currency Has Table booking
\
0   French, Japanese, Desserts ... Botswana Pula(P)   Yes
1   Japanese ... Botswana Pula(P)   Yes
2 Seafood, Asian, Filipino, Indian ... Botswana Pula(P)   Yes
3   Japanese, Sushi ... Botswana Pula(P)   No
4   Japanese, Korean ... Botswana Pula(P)   Yes

  Has Online delivery Is delivering now Switch to order menu Price range \
0   No   No   No   3
1   No   No   No   3
2   No   No   No   4
3   No   No   No   4
4   No   No   No   4

   Aggregate rating Rating color Rating text Votes
0   4.8   Dark Green   Excellent   314
1   4.5   Dark Green   Excellent   591
2   4.4   Green   Very Good   270
3   4.9   Dark Green   Excellent   365
4   4.8   Dark Green   Excellent   229

[5 rows x 21 columns],
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]:
```
# Count the number of restaurants per cuisine type in the give data
cuisine_counts=rdata_df['Cuisines'].str.split(', ').explode().value_counts().hea
# Count the number of restaurants per city
city_counts=rdata_df['City'].value_counts().head(10)
cuisine_counts,city_counts
```

Out[17]:
```
(Cuisines
 North Indian    3960
 Chinese         2735
 Fast Food       1986
 Mughlai          995
 Italian          764
 Bakery           745
 Continental      736
 Cafe             703
 Desserts         653
 South Indian     636
 Name: count, dtype: int64,
 City
 New Delhi       5473
 Gurgaon         1118
 Noida           1080
 Faridabad        251
 Ghaziabad         25
 Bhubaneshwar      21
 Amritsar          21
 Ahmedabad         21
 Lucknow           21
 Guwahati          21
 Name: count, dtype: int64)
```

In [ ]: