level-1-task-2-city-analysis

September 6, 2024

1 Level 1 Task 2. City Analysis

1.1 Import necessary libraries

[1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

1.2 read csv file

[2]:
```
df = pd.read_csv("D:\Data Analytics\Internships\Cognifyz\Dataset .csv")
df.head(3)
```

[2]:
```
   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

                                           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…

                                  Locality  \
0   Century City Mall, Poblacion, Makati City
1  Little Tokyo, Legaspi Village, Makati City
2  Edsa Shangri-La, 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

                          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
```

```
  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

   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 rows x 21 columns]
```

[3]:
```
df.shape
```

[3]:
```
(9551, 21)
```

[4]:
```
df.columns
```

[4]:
```
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')
```

1.3 1. Identify the city with the highest number of restaurants in the dataset.

[5]:
```
df['City'].value_counts().head(1)
```

[5]:
```
New Delhi    5473
Name: City, dtype: int64
```

1.4 2. Calculate the average rating for restaurants in each city.

[11]:
```
avg_rating = df.groupby('City')['Aggregate rating'].mean()
avg_rating
```

[11]:
```
City
Abu Dhabi          4.300000
Agra               3.965000
Ahmedabad          4.161905
Albany             3.555000
Allahabad          3.395000
                     …
```

```
Weirton             3.900000
Wellington City     4.250000
Winchester Bay      3.200000
Yorkton             3.300000
 stanbul           4.292857
Name: Aggregate rating, Length: 141, dtype: float64
```

1.5 3. Determine the city with the highest average rating.

[12]:
```
avg_rating.sort_values(ascending=False).head(1)
```

[12]:
```
City
Inner City    4.9
Name: Aggregate rating, dtype: float64
```