cognify-l3-t3

January 20, 2024

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

```python
[3]: df = pd.read_csv('./L1T2_Dataset.csv')
     df.head()
```

```
[3]:           Restaurant Name  Country Code              City  \
     0        Le Petit Souffle           162       Makati City
     1        Izakaya Kikufuji           162       Makati City
     2  Heat - Edsa Shangri-La           162  Mandaluyong City
     3                    Ooma           162  Mandaluyong City
     4             Sambo Kojin           162  Mandaluyong City

                                                   Address   Longitude   Latitude  \
     0  Third Floor, Century City Mall, Kalayaan Avenu…  121.027535  14.565443
     1  Little Tokyo, 2277 Chino Roces Avenue, Legaspi…  121.014101  14.553708
     2  Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal…  121.056831  14.581404
     3  Third Floor, Mega Fashion Hall, SM Megamall, O…  121.056475  14.585318
     4  Third Floor, Mega Atrium, SM Megamall, Ortigas…  121.057508  14.584450

                             Cuisines  Average Cost for two          Currency  \
     0      French, Japanese, Desserts                  1100  Botswana Pula(P)
     1                        Japanese                  1200  Botswana Pula(P)
     2  Seafood, Asian, Filipino, Indian                4000  Botswana Pula(P)
     3                 Japanese, Sushi                  1500  Botswana Pula(P)
     4                 Japanese, Korean                 1500  Botswana Pula(P)

        Has Table booking  Has Online delivery  Is delivering now  Price range  \
     0                  1                    0                  0            3
     1                  1                    0                  0            3
     2                  1                    0                  0            4
     3                  0                    0                  0            4
     4                  1                    0                  0            4

        Aggregate rating  Rating color  Rating text  Votes
```

```
        0                 4.8                 0                 1    314
        1                 4.5                 0                 1    591
        2                 4.4                 1                 5    270
        3                 4.9                 0                 1    365
        4                 4.8                 0                 1    229
```
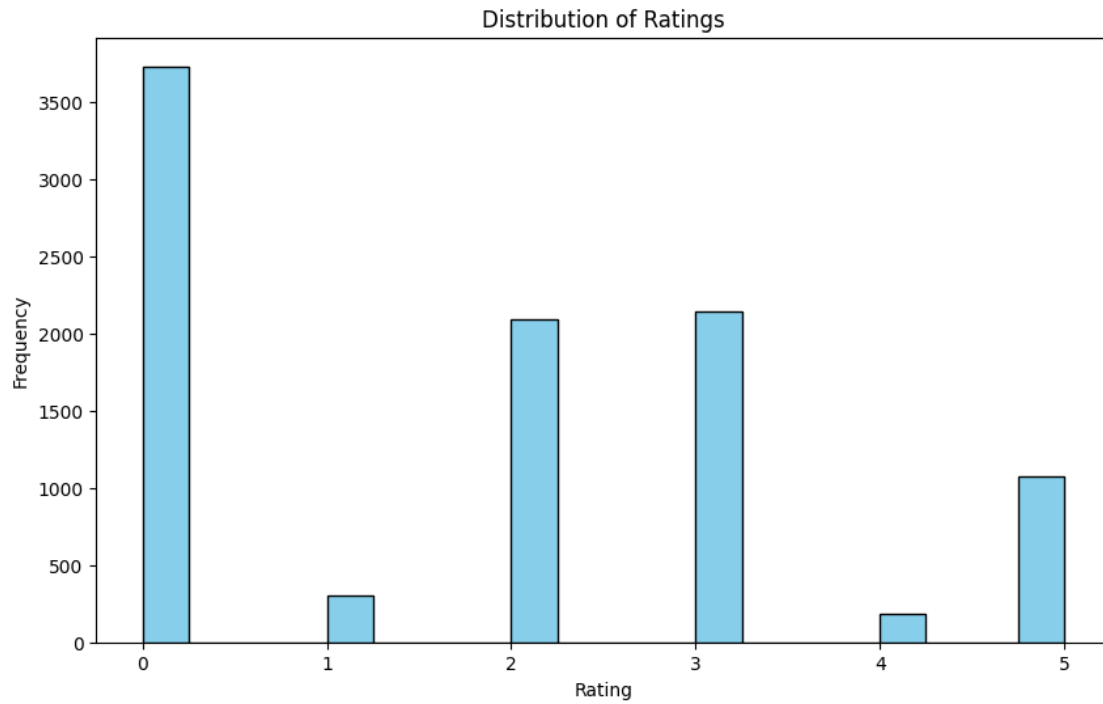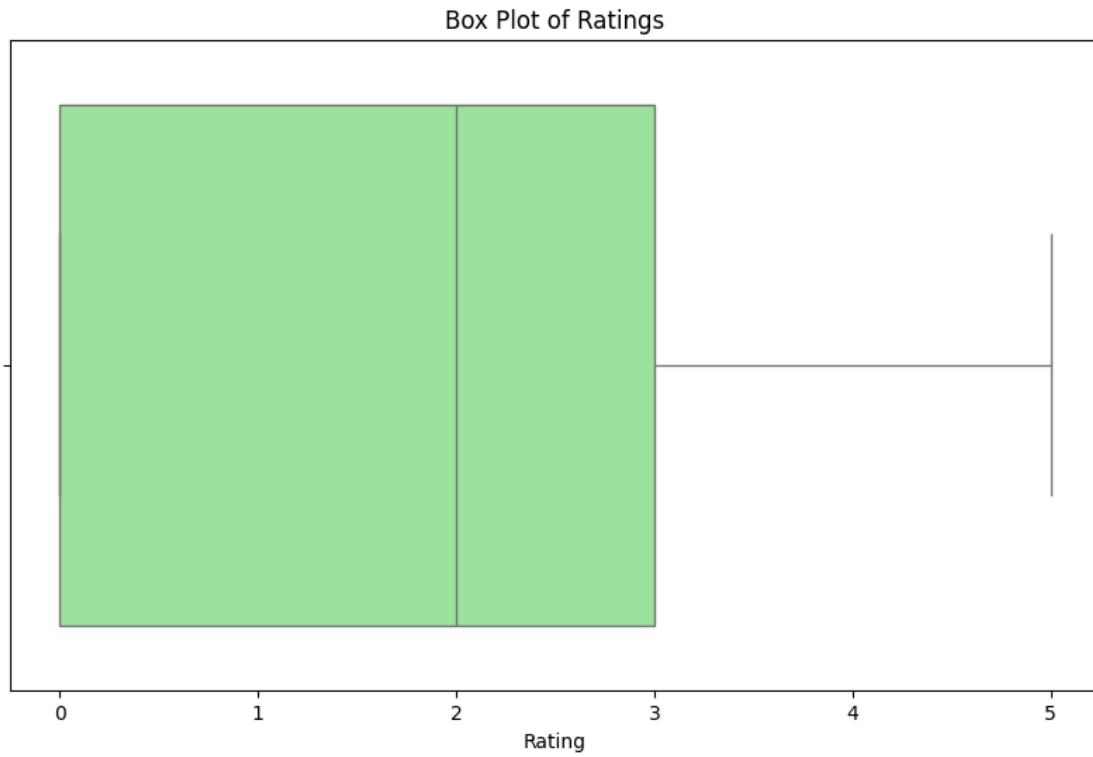
[4]: `df.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9542 entries, 0 to 9541
Data columns (total 17 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   Restaurant Name     9542 non-null   object
 1   Country Code        9542 non-null   int64
 2   City                9542 non-null   object
 3   Address             9542 non-null   object
 4   Longitude           9542 non-null   float64
 5   Latitude            9542 non-null   float64
 6   Cuisines            9542 non-null   object
 7   Average Cost for two 9542 non-null  int64
 8   Currency            9542 non-null   object
 9   Has Table booking   9542 non-null   int64
 10  Has Online delivery 9542 non-null   int64
 11  Is delivering now   9542 non-null   int64
 12  Price range         9542 non-null   int64
 13  Aggregate rating    9542 non-null   float64
 14  Rating color        9542 non-null   int64
 15  Rating text         9542 non-null   int64
 16  Votes               9542 non-null   int64
dtypes: float64(3), int64(9), object(5)
memory usage: 1.2+ MB
```
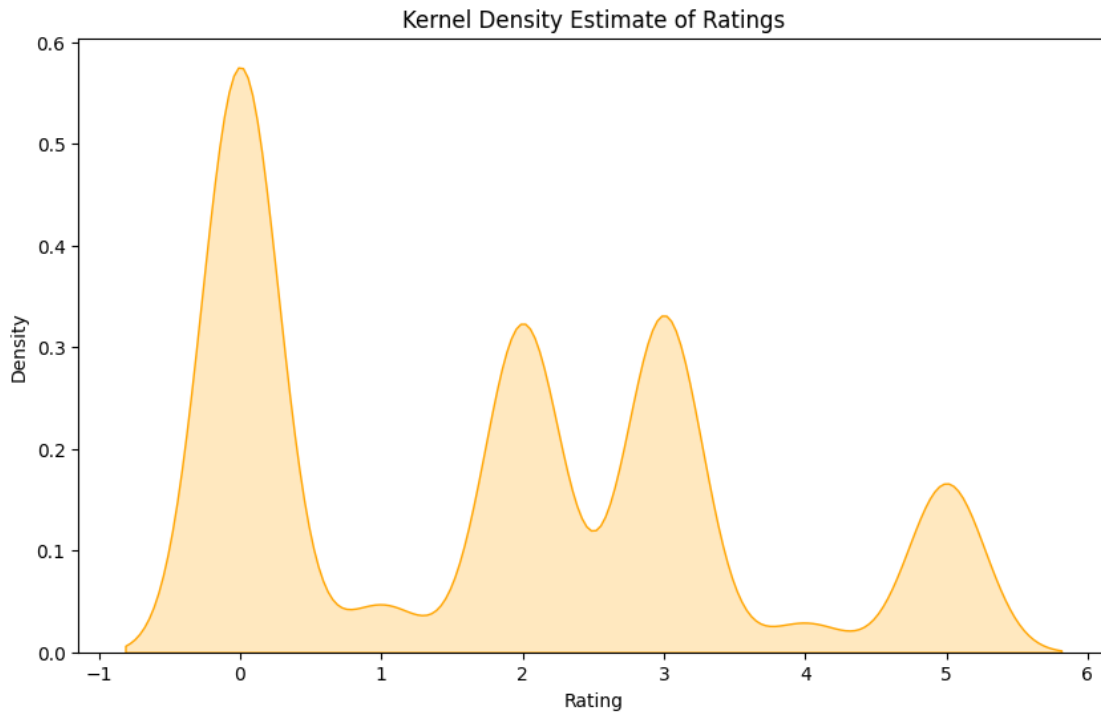
[6]:
```python
# Histogram
plt.figure(figsize=(10, 6))
plt.hist(df['Rating text'], bins=20, color='skyblue', edgecolor='black')
plt.title('Distribution of Ratings')
plt.xlabel('Rating')
plt.ylabel('Frequency')
plt.show()
```

Distribution of Ratings


```python
[7]:  # Box Plot
      plt.figure(figsize=(10, 6))
      sns.boxplot(x=df['Rating text'], color='lightgreen')
      plt.title('Box Plot of Ratings')
      plt.xlabel('Rating')
      plt.show()
```

Box Plot of Ratings

Rating

[8]:
```python
# KDE Plot (Kernel Density Estimate)
plt.figure(figsize=(10, 6))
sns.kdeplot(df['Rating text'], fill=True, color='orange')
plt.title('Kernel Density Estimate of Ratings')
plt.xlabel('Rating')
plt.show()
```

Kernel Density Estimate of Ratings


```
[17]: cuisine_ratings = df.groupby('Cuisines')['Rating text'].mean().reset_index()

      # Sorting by Ratings in descending order
      cuisine_ratings = cuisine_ratings.sort_values(by='Rating text', ascending=False)

      # Visualization: Bar plot
      plt.figure(figsize=(300, 24))
      sns.barplot(x='Cuisines', y='Rating text', data=cuisine_ratings)
      plt.xticks(rotation=90)
      plt.title('Average Ratings by Cuisine Type')
      plt.show()
```

Output hidden; open in https://colab.research.google.com to view.

```
[11]: # Group by City and calculating show the average rating for the city
      city_ratings = df.groupby('City')['Rating text'].mean().reset_index()

      # Sort by Ratings in descending order
      city_ratings = city_ratings.sort_values(by='Rating text', ascending=False)

      plt.figure(figsize=(40, 6))
      sns.barplot(x='City', y='Rating text', data=city_ratings)
      plt.xticks(rotation=90)
      plt.title('Average Ratings by City')
```

```
plt.show()
```


Average Ratings by City