level-2-task-1-restaurant-ratings

September 8, 2024

1 Level 2 Task 1. Restaurant Ratings.ipynb

1.1 Import necessary libraries

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

1.2 read csv file

[46]:
```
df = pd.read_csv("D:\Data Analytics\Internships\Cognifyz\Dataset .csv")
df.head(3)
```

[46]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude	Cuisines	…	Currency	Has Table booking
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)	Yes
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)	Yes
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)	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]
```

1.3 1. Analyze the distribution of aggregate ratings and determine the most common rating range.

[47]:
```
agg_rating_distribution = df["Aggregate rating"].value_counts()
agg_rating_distribution
```

[47]:
```
0.0    2148
3.2     522
3.1     519
3.4     498
3.3     483
3.5     480
3.0     468
3.6     458
3.7     427
3.8     400
2.9     381
3.9     335
2.8     315
4.1     274
4.0     266
2.7     250
4.2     221
2.6     191
4.3     174
4.4     144
2.5     110
4.5      95
2.4      87
4.6      78
4.9      61
2.3      47
4.7      42
2.2      27
```

```
4.8      25
2.1      15
2.0       7
1.9       2
1.8       1
Name: Aggregate rating, dtype: int64
```

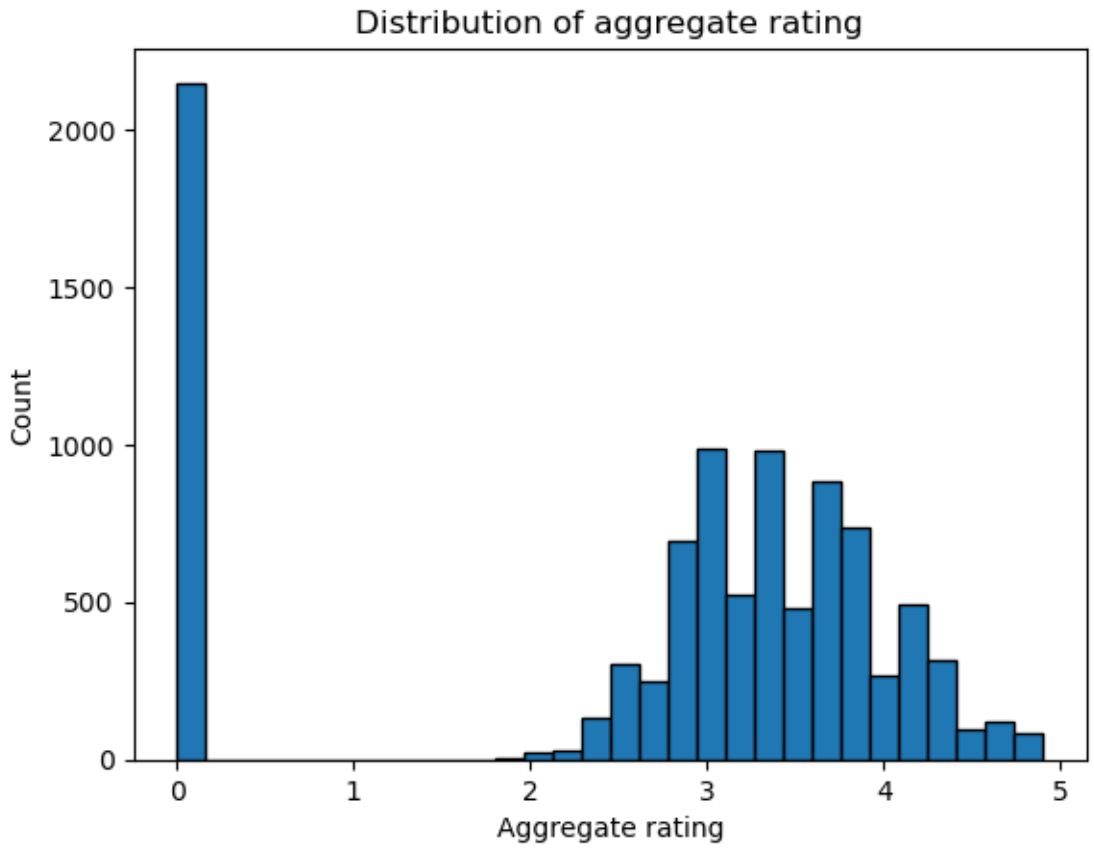
[48]:
```
agg_rating_distribution
```

[48]:
```
0.0    2148
3.2     522
3.1     519
3.4     498
3.3     483
3.5     480
3.0     468
3.6     458
3.7     427
3.8     400
2.9     381
3.9     335
2.8     315
4.1     274
4.0     266
2.7     250
4.2     221
2.6     191
4.3     174
4.4     144
2.5     110
4.5      95
2.4      87
4.6      78
4.9      61
2.3      47
4.7      42
2.2      27
4.8      25
2.1      15
2.0       7
1.9       2
1.8       1
Name: Aggregate rating, dtype: int64
```

[49]:
```
print("The most common rating is = ", agg_rating_distribution.idxmax())
```

```
The most common rating is =  0.0
```

[52]:
```
plt.hist(df["Aggregate rating"], bins=30, edgecolor='black')
plt.xlabel("Aggregate rating")
plt.ylabel("Count")
plt.title("Distribution of aggregate rating")
plt.show()
```



1.4 2. Calculate the average number of votes received by restaurants.

[54]:
```
avg_vote = round(df['Votes'].mean(), 3)
print("The average number of votes received by restaurants: ", avg_vote)
```

```
The average number of votes received by restaurants:  156.91
```