```
In [2]: # Data handling
import pandas as pd
import numpy as np

# Visualization
import matplotlib.pyplot as plt
import seaborn as sns
```

```
In [25]: # Read dataset into pandas dataframe.
data = pd.read_csv("D:\Python Files Excel\Zomato-data-.csv")
```

```
In [26]: df = pd.DataFrame(data)
```

```
In [8]: print(data)
```

```
                      name online_order book_table   rate  votes  \
0                    Jalsa          Yes        Yes  4.1/5    775
1           Spice Elephant          Yes         No  4.1/5    787
2          San Churro Cafe          Yes         No  3.8/5    918
3    Addhuri Udupi Bhojana           No         No  3.7/5     88
4            Grand Village           No         No  3.8/5    166
..                     ...          ...        ...    ...    ...
143       Melting Melodies           No         No  3.3/5      0
144        New Indraprasta           No         No  3.3/5      0
145           Anna Kuteera          Yes         No  4.0/5    771
146                 Darbar           No         No  3.0/5     98
147          Vijayalakshmi          Yes         No  3.9/5     47

     approx_cost(for two people) listed_in(type)
0                            800          Buffet
1                            800          Buffet
2                            800          Buffet
3                            300          Buffet
4                            600          Buffet
..                           ...             ...
143                          100          Dining
144                          150          Dining
145                          450          Dining
146                          800          Dining
147                          200          Dining

[148 rows x 7 columns]
```

```
In [27]: # Lists all column names.
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 148 entries, 0 to 147
Data columns (total 7 columns):
 #   Column                       Non-Null Count  Dtype 
---  ------                       --------------  ----- 
 0   name                         148 non-null    object
 1   online_order                 148 non-null    object
 2   book_table                   148 non-null    object
 3   rate                         148 non-null    object
 4   votes                        148 non-null    int64 
 5   approx_cost(for two people)  148 non-null    int64 
 6   listed_in(type)              148 non-null    object
dtypes: int64(2), object(5)
memory usage: 8.2+ KB
```

In [28]:
```
# Statistical summary
df.describe()
```

Out[28]:

	votes	approx_cost(for two people)
count	148.000000	148.000000
mean	264.810811	418.243243
std	653.676951	223.085098
min	0.000000	100.000000
25%	6.750000	200.000000
50%	43.500000	400.000000
75%	221.750000	600.000000
max	4884.000000	950.000000

In [15]:
```
# Data Cleaning
```

In [29]:
```
# Check Missing Values
df.isnull().sum()
```

Out[29]:
```
name                           0
online_order                   0
book_table                     0
rate                           0
votes                          0
approx_cost(for two people)    0
listed_in(type)                0
dtype: int64
```

In [30]:
```
# Remove Duplicates
df.drop_duplicates(inplace=True)
```

In [32]:
```
# Clean Cost Column
df['approx_cost(for two people)'] = df['approx_cost(for two people)']\
```

```
    .astype(str)\
    .str.replace(',', '')\
    .astype(float)
```

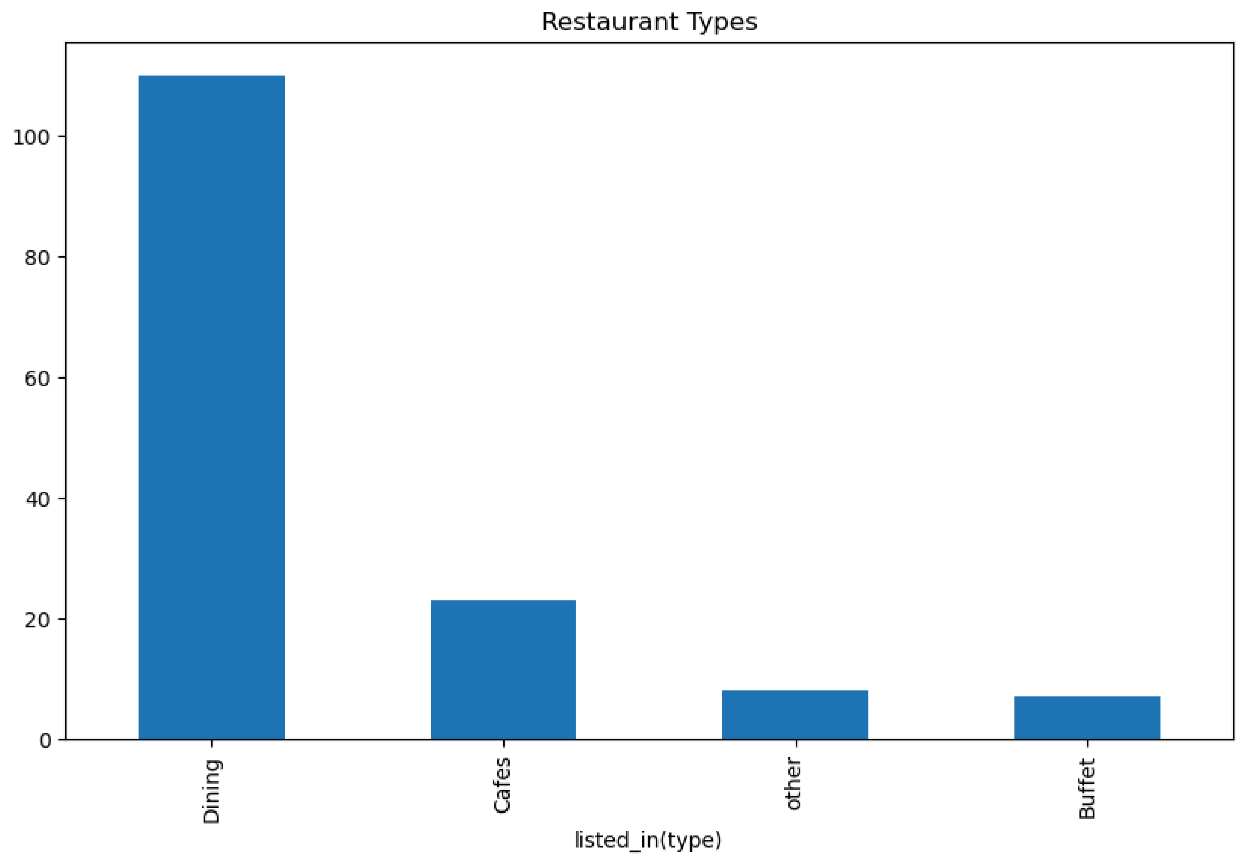
In [33]:
```
print(df)
```

```
                     name online_order book_table   rate  votes  \
0                   Jalsa          Yes        Yes  4.1/5    775   
1          Spice Elephant          Yes         No  4.1/5    787   
2         San Churro Cafe          Yes         No  3.8/5    918   
3   Addhuri Udupi Bhojana           No         No  3.7/5     88   
4           Grand Village           No         No  3.8/5    166   
..                    ...          ...        ...    ...    ...   
143      Melting Melodies           No         No  3.3/5      0   
144       New Indraprasta           No         No  3.3/5      0   
145          Anna Kuteera          Yes         No  4.0/5    771   
146                Darbar           No         No  3.0/5     98   
147         Vijayalakshmi          Yes         No  3.9/5     47   

     approx_cost(for two people) listed_in(type)  
0                          800.0          Buffet  
1                          800.0          Buffet  
2                          800.0          Buffet  
3                          300.0          Buffet  
4                          600.0          Buffet  
..                           ...             ...  
143                        100.0          Dining  
144                        150.0          Dining  
145                        450.0          Dining  
146                        800.0          Dining  
147                        200.0          Dining  

[148 rows x 7 columns]
```

In [43]:
```
# Types of Restaurant
df['listed_in(type)'].value_counts().head(10)
```

Out[43]:
```
listed_in(type)
Dining    110
Cafes      23
other       8
Buffet      7
Name: count, dtype: int64
```

In [34]:
```
# Types of Restaurant
plt.figure(figsize=(10,6))
df['listed_in(type)'].value_counts().head(10).plot(kind='bar')
plt.title("Restaurant Types")
plt.show()
```



The majority of restaurants fall under the dining category, indicating that customers prefer casual dine-in experiences. Cafes form the second largest category, while buffet and other restaurant types are comparatively fewer, suggesting niche demand.

In [46]:
```
# Online Order Availability
df['online_order'].value_counts()
```

Out[46]:
```
online_order
No     90
Yes    58
Name: count, dtype: int64
```

In [35]:
```
# Online Order Availability
sns.countplot(x='online_order', data=df)
plt.title("Online Order Availability")
plt.show()
```

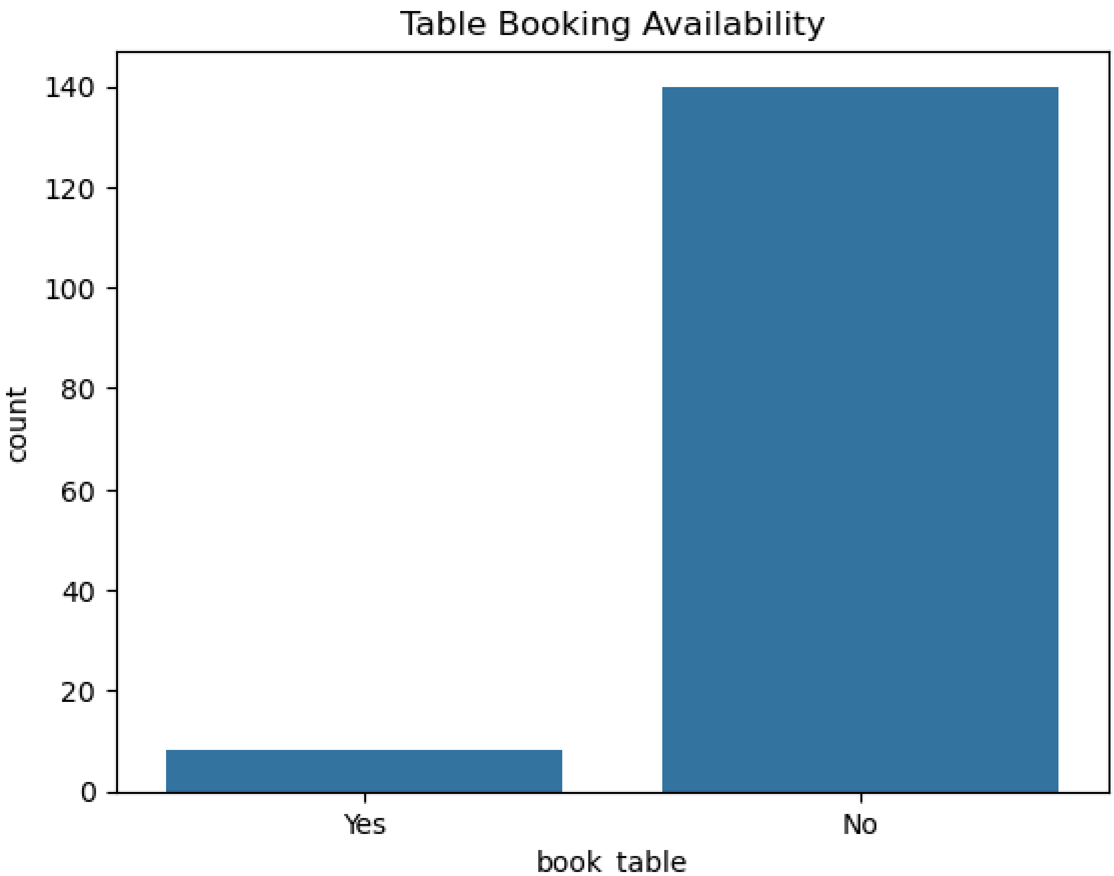
Most restaurants in the dataset do not offer online ordering, suggesting that traditional dine-in services still dominate. However, a significant portion of restaurants providing online ordering indicates growing adoption of digital food delivery services.

In [47]:
```
# Table Booking Trend
df['book_table'].value_counts()
```

Out[47]:
```
book_table
No     140
Yes      8
Name: count, dtype: int64
```

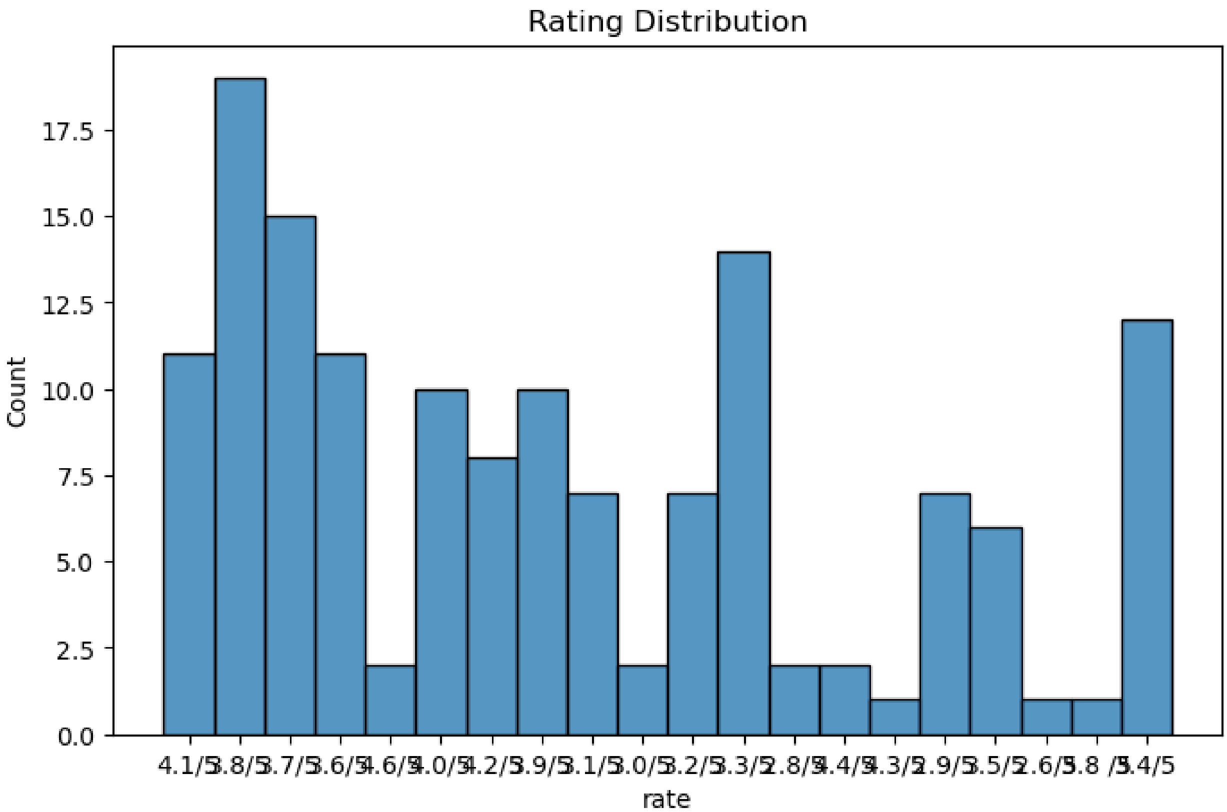
In [36]:
```
# Table Booking Availability

sns.countplot(x='book_table', data=df)
plt.title("Table Booking Availability")
```

```
plt.show()
```



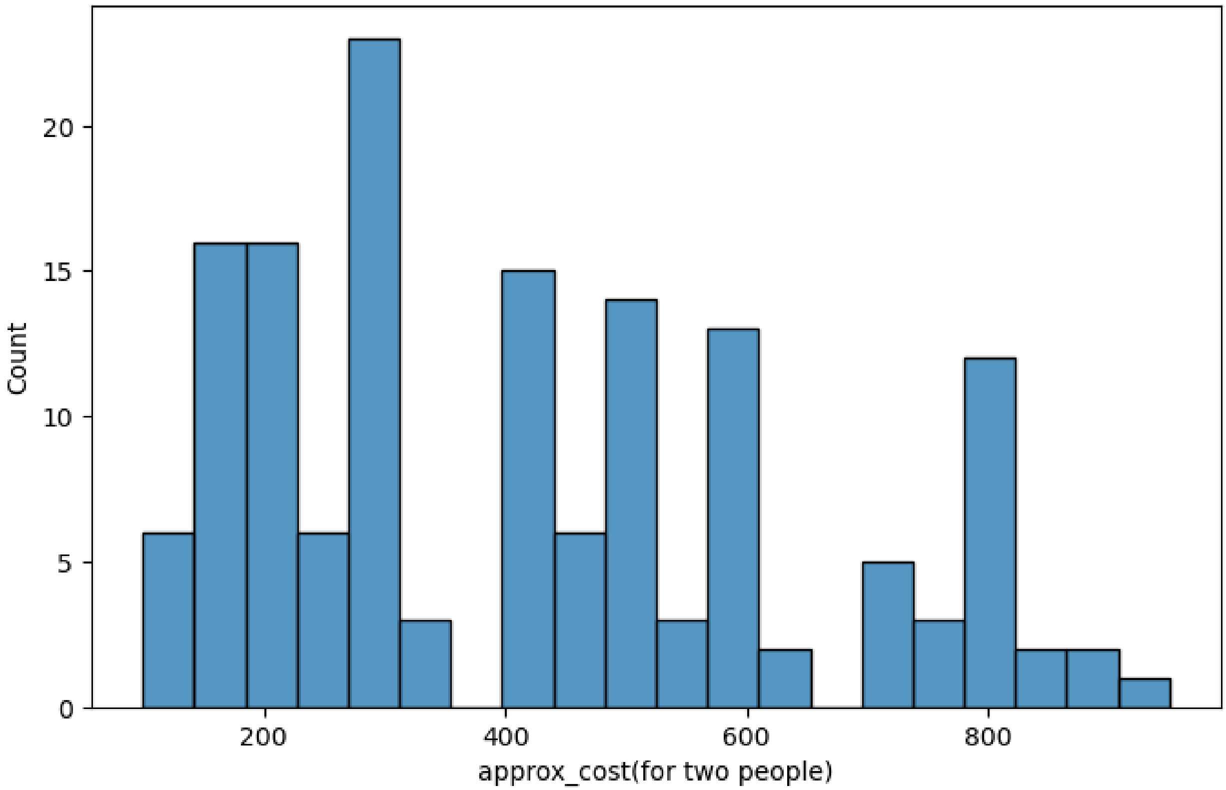
Only a small number of restaurants provide table booking facilities. This suggests that most restaurants operate on walk-in customers rather than reservation-based dining, which is typical for casual or budget dining segments.

In [37]:
```
# Rating
plt.figure(figsize=(8,5))
sns.histplot(df['rate'], bins=20)
plt.title("Rating Distribution")
plt.show()
```



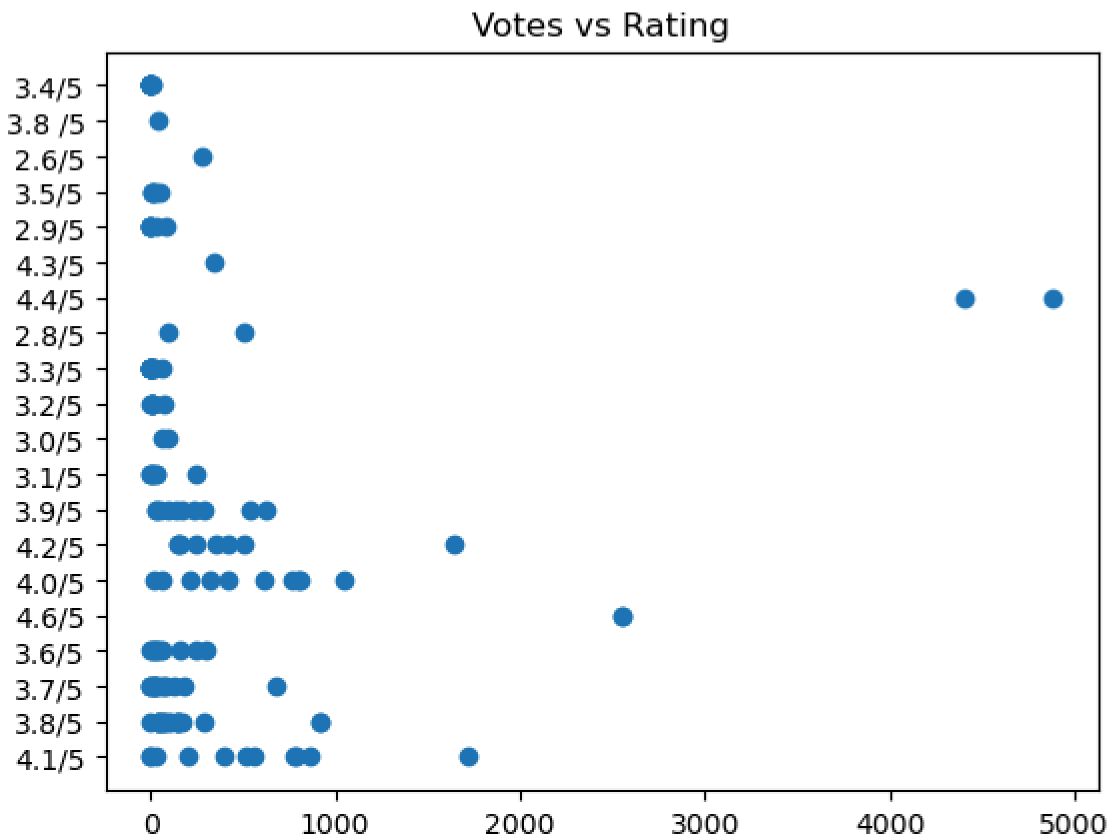
Restaurant ratings are mostly concentrated between 3.5 and 4.0, indicating generally satisfactory customer experiences. Very low or very high ratings are less common, suggesting consistent service quality across most restaurants.

In [44]:
```
# Cost Distribution
plt.figure(figsize=(8,5))
sns.histplot(df['approx_cost(for two people)'], bins=20)
plt.show()
```



The cost distribution shows that most restaurants fall within the affordable to mid-range pricing category. High-cost restaurants are relatively fewer, indicating that budget-friendly dining options dominate the market.

In [51]:
```
# Votes vs Rating
plt.figure()
plt.scatter(df['votes'], df['rate'])
plt.title("Votes vs Rating")
plt.show()
```



Votes vs Rating
3.4/5
3.8 /5
2.6/5
3.5/5
2.9/5
4.3/5
4.4/5
2.8/5
3.3/5
3.2/5
3.0/5
3.1/5
3.9/5
4.2/5
4.0/5
4.6/5
3.6/5
3.7/5
3.8/5
4.1/5
0
1000
2000
3000
4000
5000

The scatter plot indicates a mild positive relationship between customer votes and ratings. Restaurants with higher ratings often receive more votes, although some highly rated restaurants have fewer votes, suggesting that popularity depends on factors beyond rating such as location, marketing, and brand visibility.

CONCLUSION

The analysis indicates that most restaurants operate in the affordable dining segment with moderate customer ratings. Online ordering adoption is growing but traditional dining remains dominant. Customer engagement varies significantly, with only a few restaurants receiving very high votes, highlighting competitive market dynamics.