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

In [2]:
```
data = pd.read_csv(r"C:\Users\santo\OneDrive\Desktop\Data science\Dec2025\14th dec EDA\Zomato-data-.csv")
```

In [5]:
```
(data.head())
```

Out[5]:

	name	online_order	book_table	rate	votes	approx_cost(for two people)	listed_in(type)
0	Jalsa	Yes	Yes	4.1/5	775	800	Buffet
1	Spice Elephant	Yes	No	4.1/5	787	800	Buffet
2	San Churro Cafe	Yes	No	3.8/5	918	800	Buffet
3	Addhuri Udupi Bhojana	No	No	3.7/5	88	300	Buffet
4	Grand Village	No	No	3.8/5	166	600	Buffet

In [7]:
```
def handleRate(value):
    value = str(value).split("/")
    value = value[0];
    return float(value)

data['rate'] = data['rate'].apply(handleRate)
(data.head())
(data.tail())
```

Out[7]:

	name	online_order	book_table	rate	votes	approx_cost(for two people)	listed_in(type)
143	Melting Melodies	No	No	3.3	0	100	Dining
144	New Indraprasta	No	No	3.3	0	150	Dining
145	Anna Kuteera	Yes	No	4.0	771	450	Dining
146	Darbar	No	No	3.0	98	800	Dining
147	Vijayalakshmi	Yes	No	3.9	47	200	Dining

In [8]:
```
data.info()
```

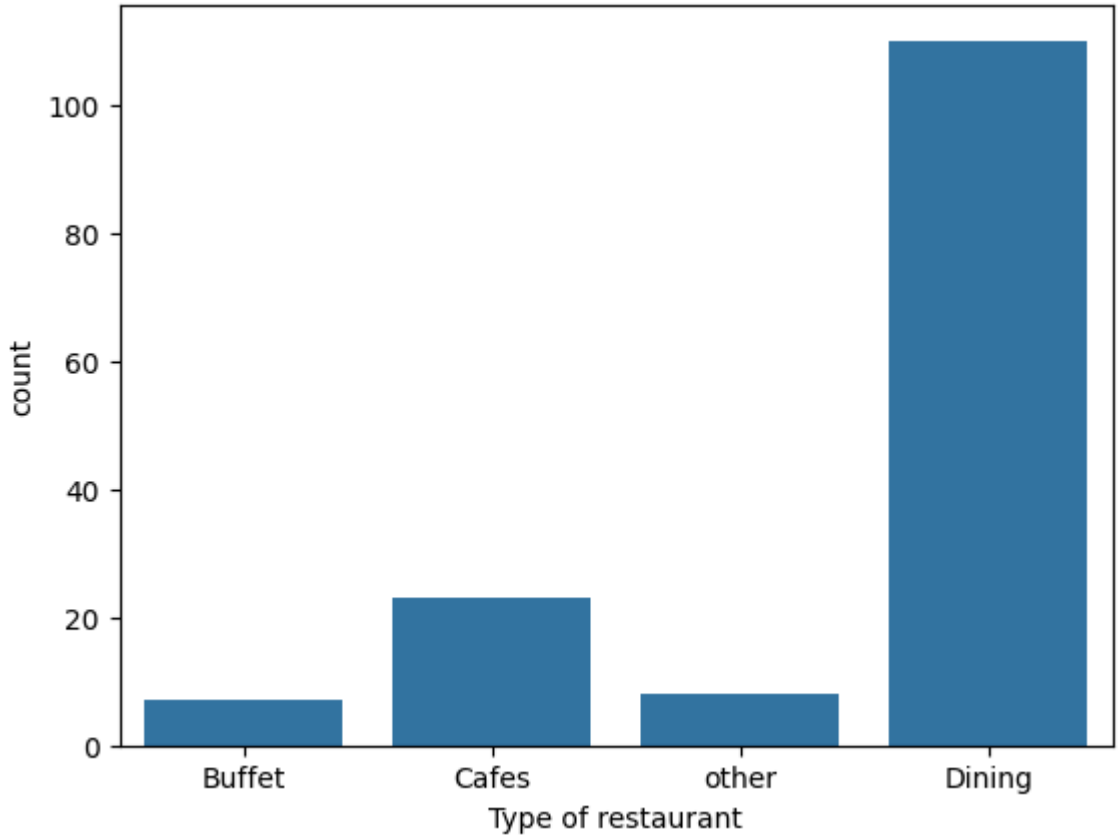
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 148 entries, 0 to 147
Data columns (total 7 columns):
 #   Column                       Non-Null Count  Dtype  
---  ------                       --------------  -----  
 0   name                         148 non-null    object 
 1   online_order                 148 non-null    object 
 2   book_table                   148 non-null    object 
 3   rate                         148 non-null    float64
 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: float64(1), int64(2), object(4)
memory usage: 8.2+ KB
```

In [10]:
```
print(data.isnull().sum())
```

```
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 [11]:
```
sns.countplot(x=data['listed_in(type)'])
plt.xlabel("Type of restaurant")
```

Out[11]:
```
Text(0.5, 0, 'Type of restaurant')
```



In [12]:
```
grouped_data = data.groupby('listed_in(type)')['votes'].sum()
result = pd.DataFrame({'votes':grouped_data})
plt.plot(result, c='green', marker='o')
plt.xlabel('Type of restaurant')
plt.ylabel('votes')
```

Out[12]:
```
Text(0, 0.5, 'votes')
```

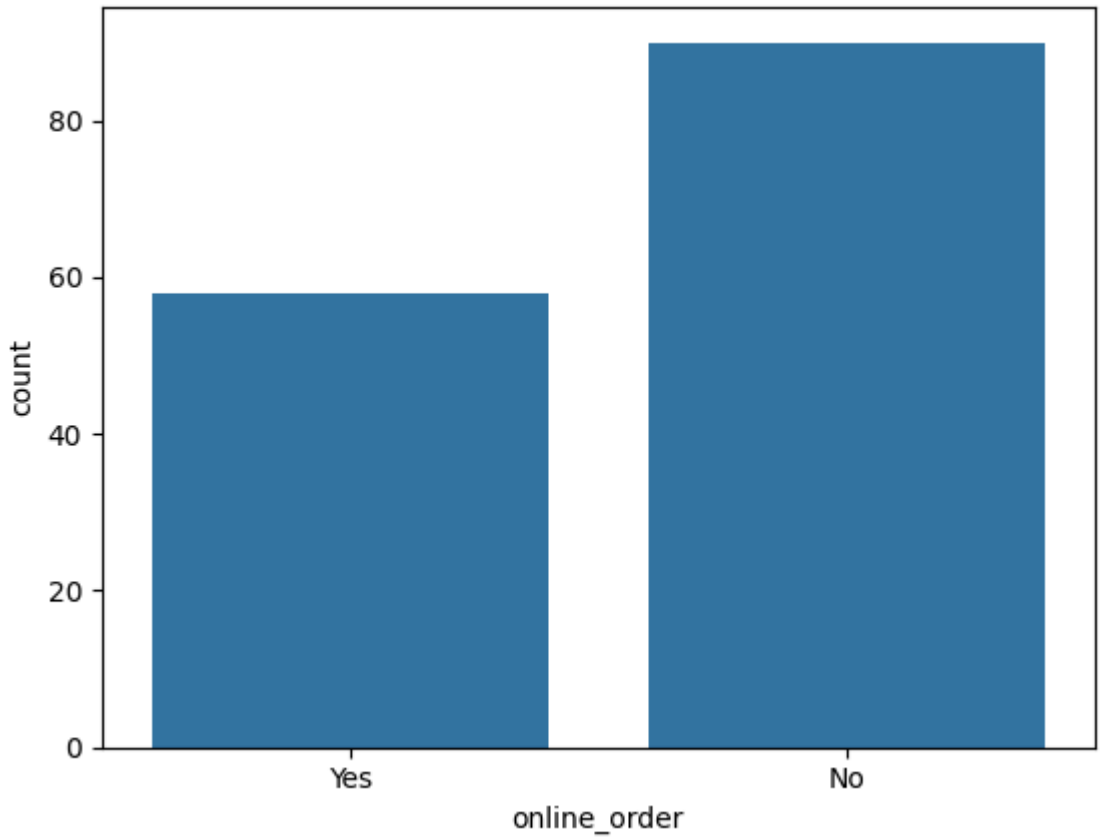
In [15]:
```
max_votes = data['votes'].max()
restaurant_with_max_votes = data.loc[data['votes'] == max_votes, 'name']

print('Restaurant(s) with the maximum votes:')
print(restaurant_with_max_votes)
```

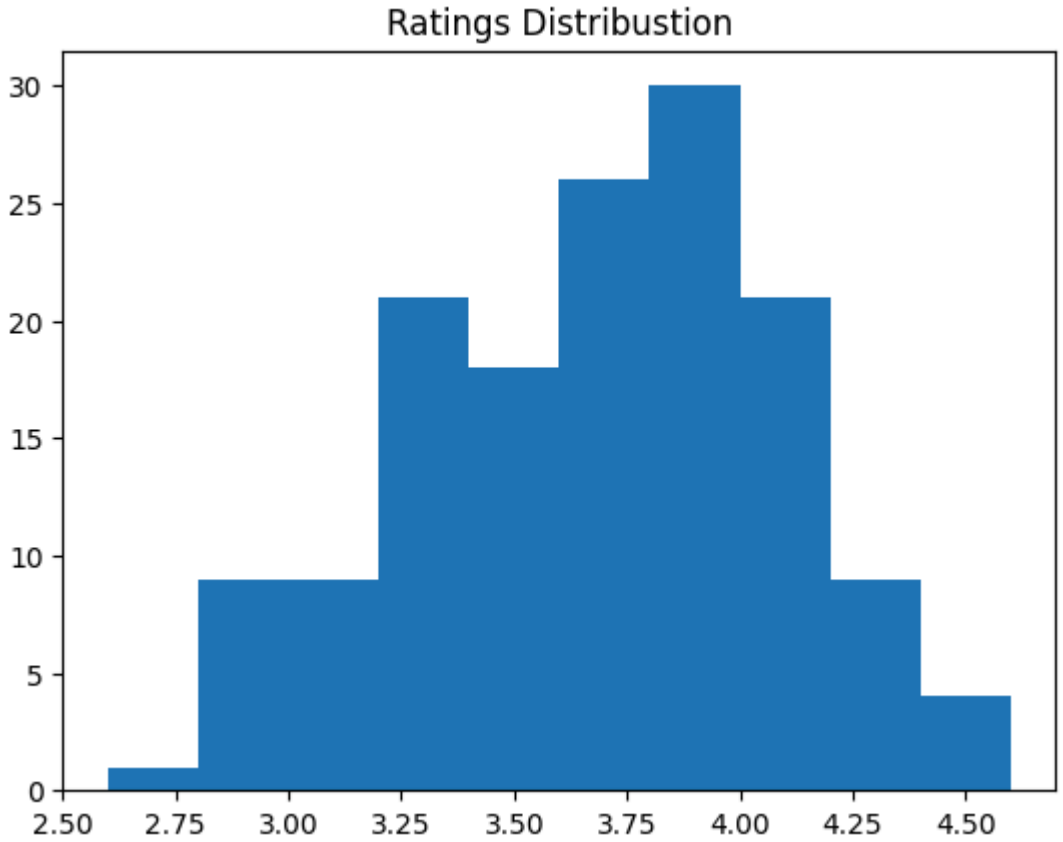
```
Restaurant(s) with the maximum votes:
38    Empire Restaurant
Name: name, dtype: object
```

In [16]:
```
sns.countplot(x=data['online_order'])
```

Out[16]: <Axes: xlabel='online_order', ylabel='count'>

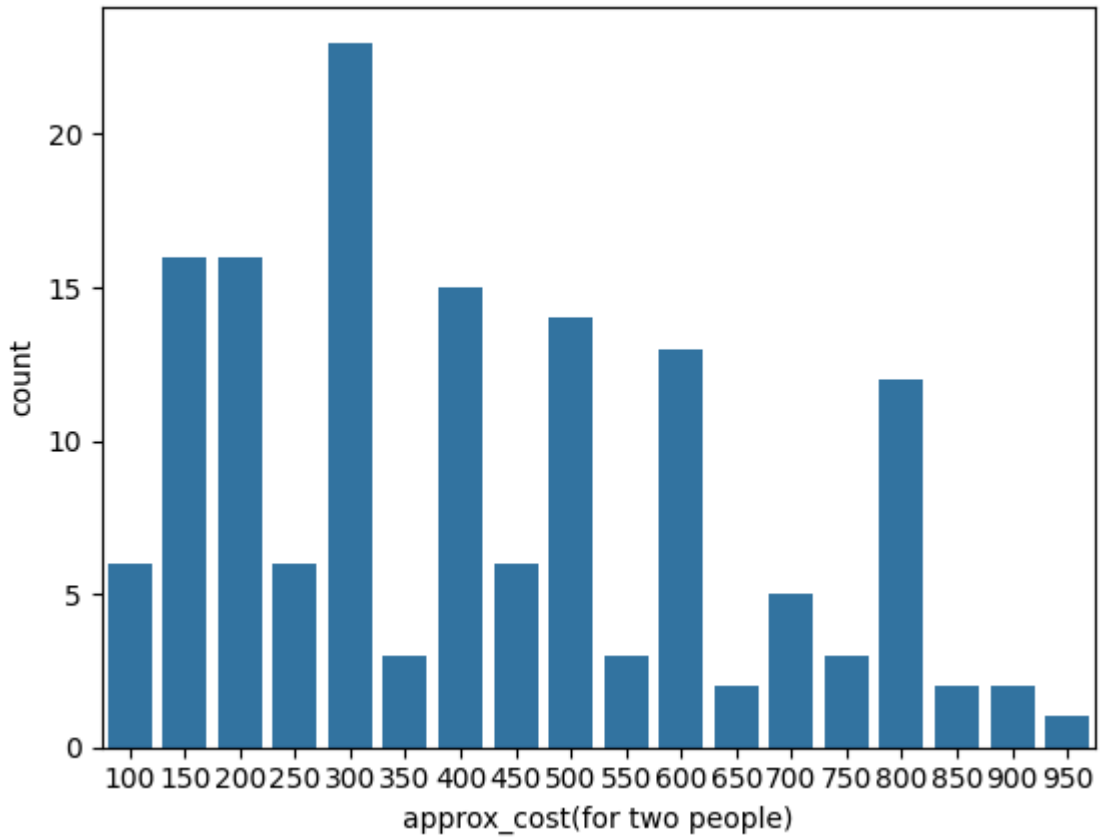


In [18]:
```
plt.hist(data['rate'],bins = 10)
plt.title("Ratings Distribustion")
plt.show()
```



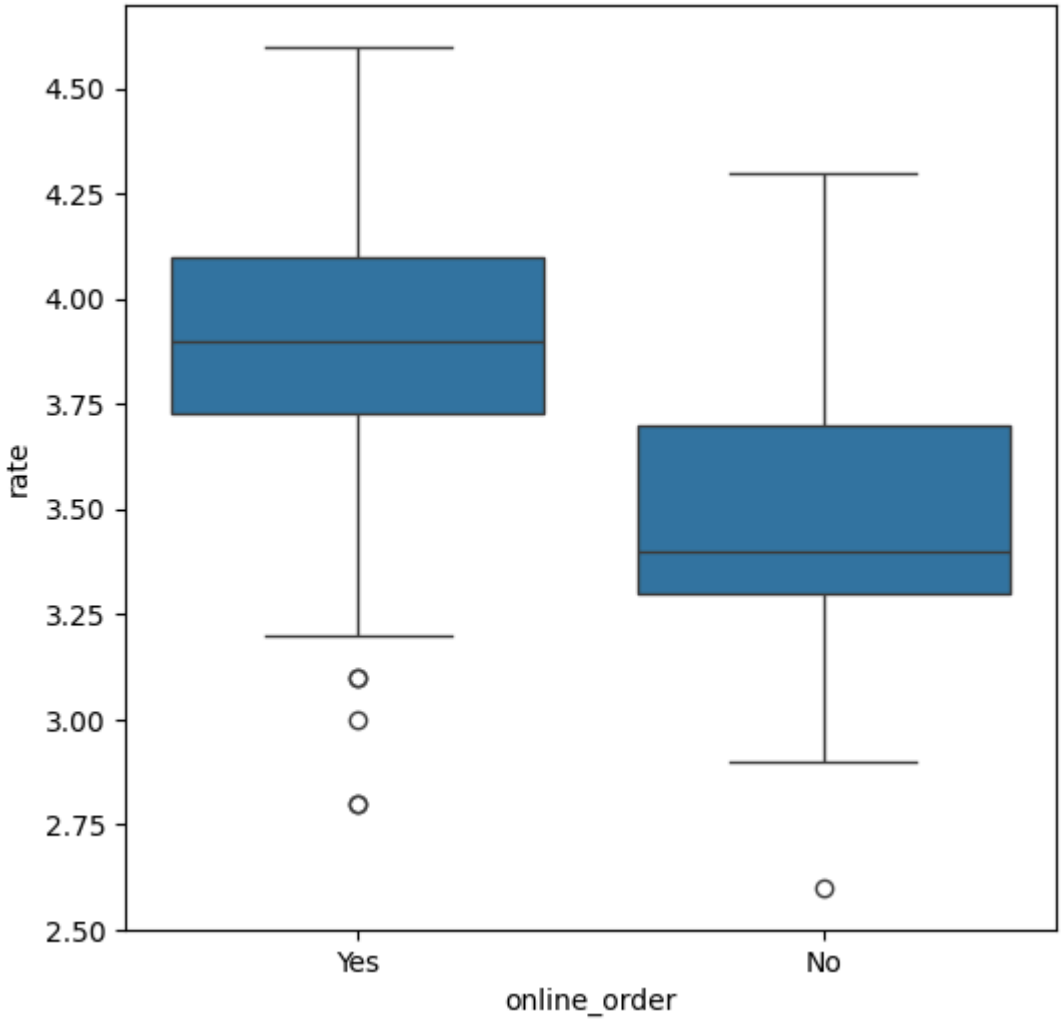
In [20]:
```
couple_data=data['approx_cost(for two people)']
sns.countplot(x=couple_data)
```

Out[20]:
```
<Axes: xlabel='approx_cost(for two people)', ylabel='count'>
```



In [21]:
```
plt.figure(figsize = (6,6))
sns.boxplot(x = 'online_order', y = 'rate', data = data)
```

Out[21]:
```
<Axes: xlabel='online_order', ylabel='rate'>
```



In [ ]: