

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

```
In [2]: dataframe = pd.read_csv("Zomato data .csv")
dataframe.head()
```

```
Out[2]:
```

	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 [3]: # Let's convert the data type of the "rate" column to float and remove the denominator
def handleRate(value):
    value=str(value).split('/')
    value=value[0];
    return float(value)

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

```
Out[3]:
```

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

```
In [4]: dataframe.info()
```

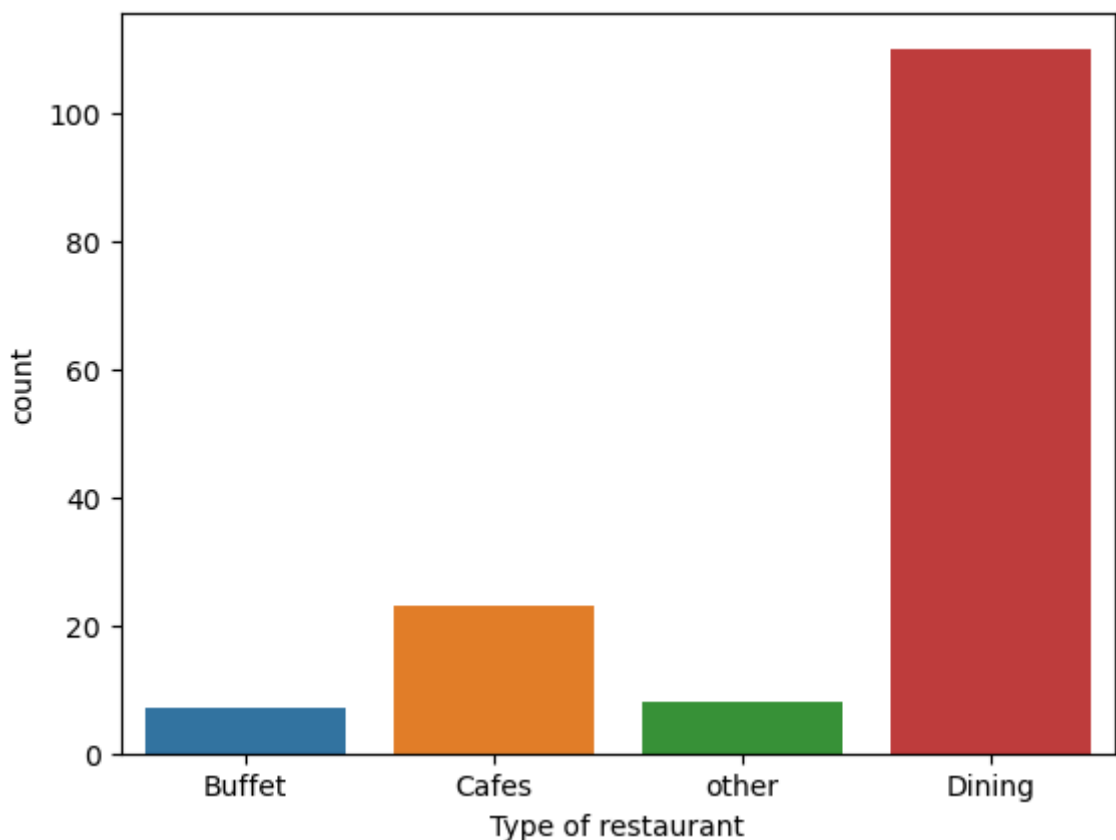
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 148 entries, 0 to 147
Data columns (total 7 columns):
#   Column                                Non-Null Count  Dtype
---  ---                                -
0   name                                148 non-null    object
1   online_order                        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 [5]: # Checking for null values
dataframe.isnull().sum()
```

```
Out[5]: 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 [6]: sns.countplot(x=dataframe['listed_in(type)'])
plt.xlabel("Type of restaurant")
```

```
Out[6]: Text(0.5, 0, 'Type of restaurant')
```

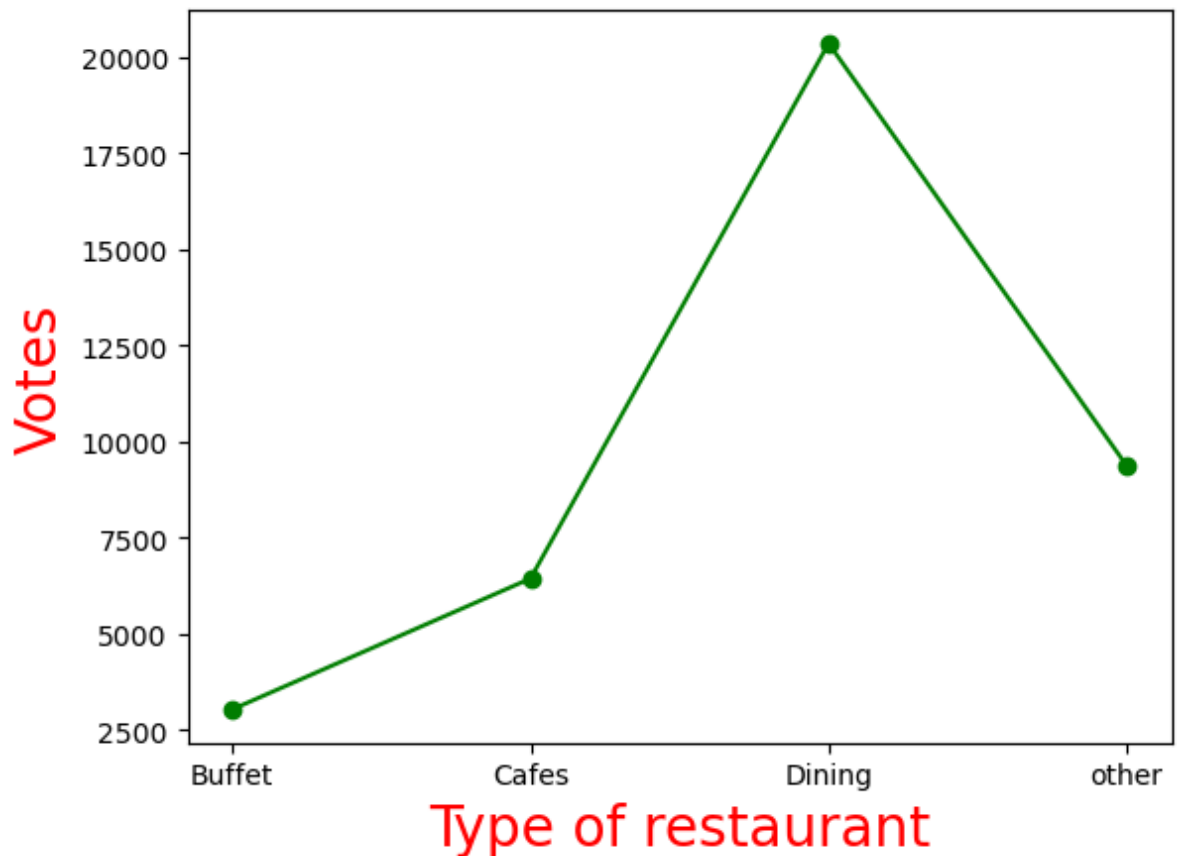


The majority of the restaurants fall into the dining category.

```
In [7]: grouped_data = dataframe.groupby('listed_in(type)')['votes'].sum()
result = pd.DataFrame({'votes': grouped_data})
plt.plot(result, c="green", marker="o")
```

```
plt.xlabel("Type of restaurant", c="red", size=20)
plt.ylabel("Votes", c="red", size=20)
```

Out[7]: Text(0, 0.5, 'Votes')



Dining restaurants are preferred by a larger number of individuals.

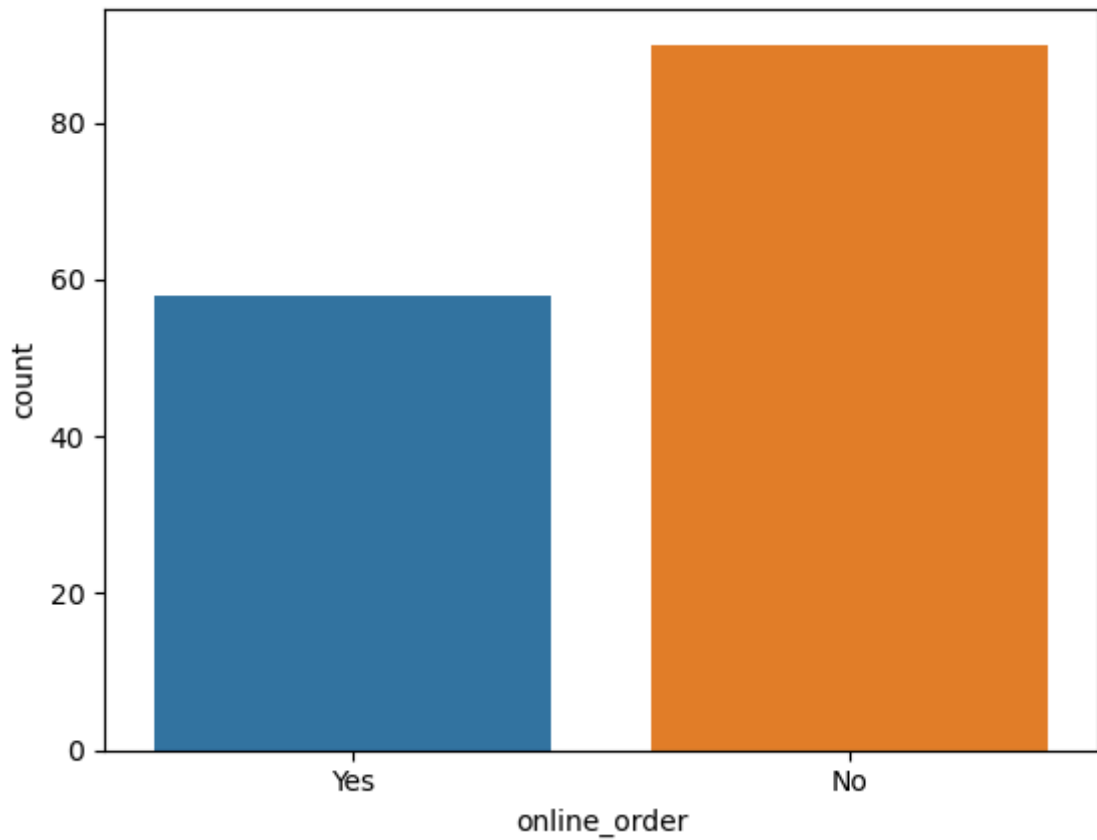
```
In [8]: # Checking for restaurant that received the maximum votes
max_votes = dataframe['votes'].max()
restaurant_with_max_votes = dataframe.loc[dataframe['votes'] == max_votes, 'name']

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
```

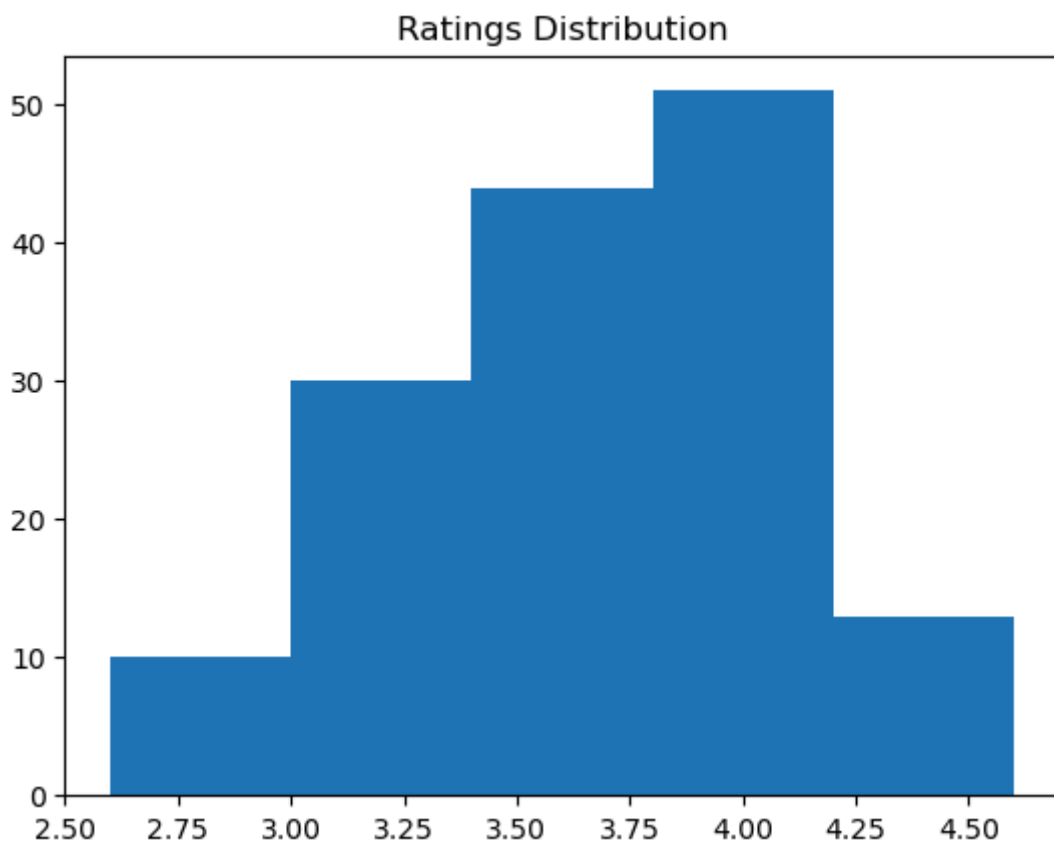
```
In [9]: sns.countplot(x=dataframe['online_order'])
```

Out[9]: <Axes: xlabel='online_order', ylabel='count'>



This suggests that a majority of the restaurants do not accept online orders.

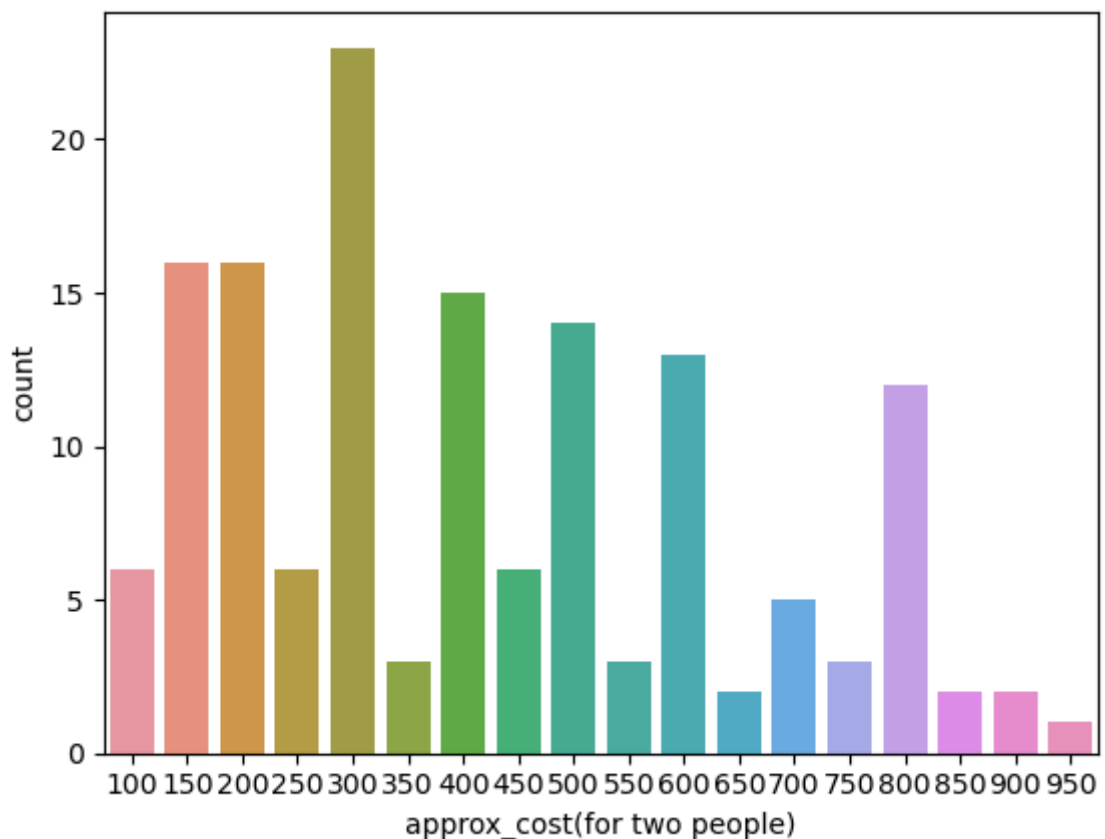
```
In [10]: plt.hist(dataframe['rate'],bins=5)
plt.title("Ratings Distribution")
plt.show()
```



The majority of restaurants received ratings ranging from 3.5 to 4.

```
In [11]: couple_data=dataframe['approx_cost(for two people)']  
sns.countplot(x=couple_data)
```

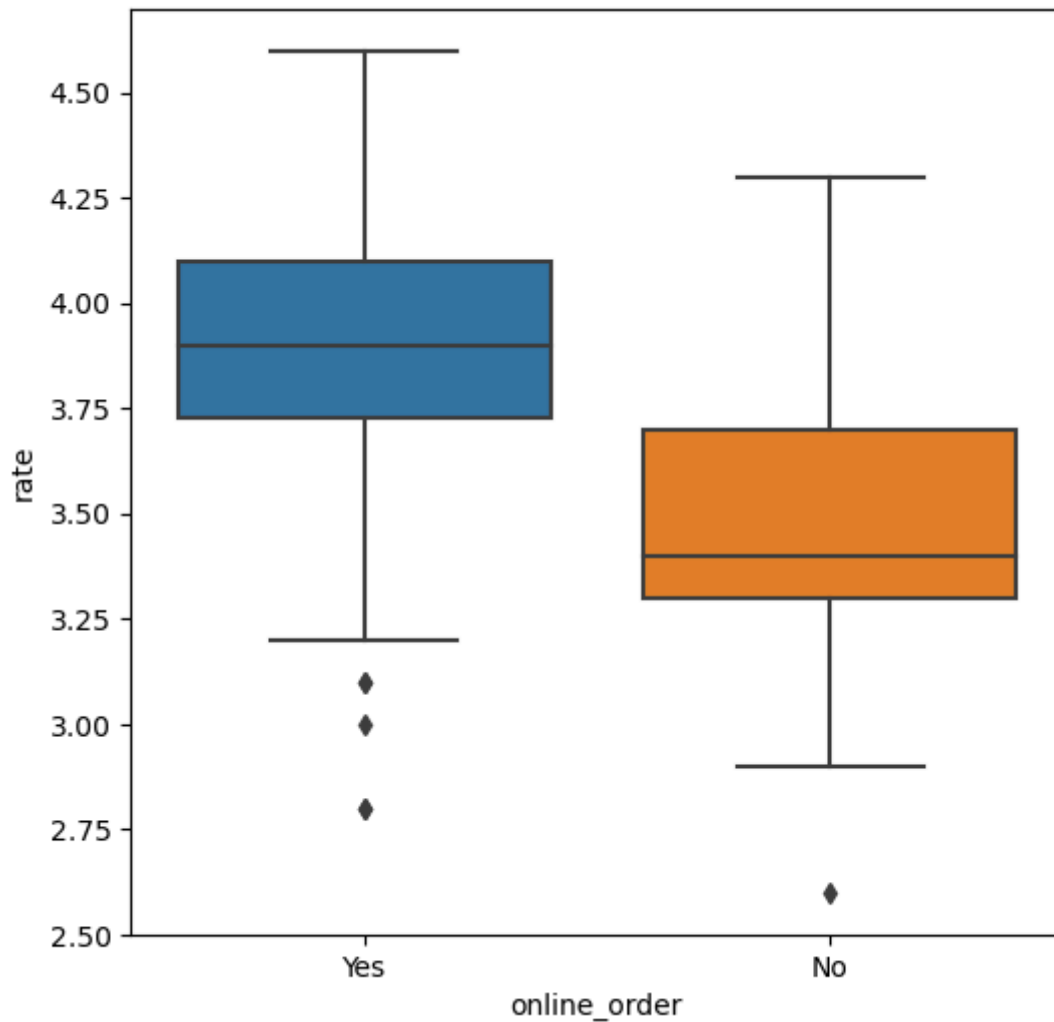
```
Out[11]: <Axes: xlabel='approx_cost(for two people)', ylabel='count'>
```



The majority of couples prefer restaurants with an approximate cost of 300 rupees.

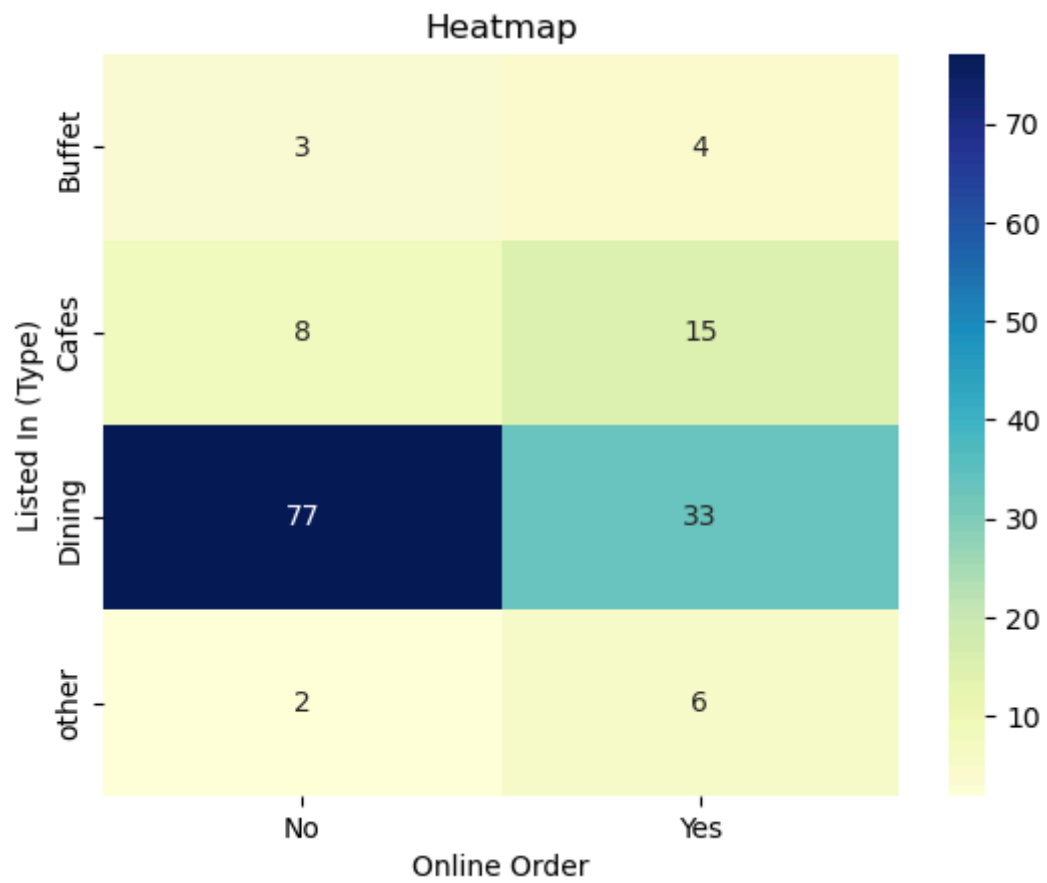
```
In [12]: # Checking if whether online orders receive higher ratings tha offline orders  
plt.figure(figsize = (6,6))  
sns.boxplot(x = 'online_order', y = 'rate', data = dataframe)
```

```
Out[12]: <Axes: xlabel='online_order', ylabel='rate'>
```



Offline orders received lower ratings in comparison to online orders, which obtained excellent ratings.

```
In [13]: pivot_table = dataframe.pivot_table(index='listed_in(type)', columns='online_order',
sns.heatmap(pivot_table, annot=True, cmap="YlGnBu", fmt='d')
plt.title("Heatmap")
plt.xlabel("Online Order")
plt.ylabel("Listed In (Type)")
plt.show()
```



Dining restaurants primarily accept offline orders, whereas cafes primarily receive online orders. This suggests that clients prefer to place orders in person at restaurants, but prefer online ordering at cafes.