

```
In [1]: #Zomato Data Analysis Using Python
```

```
In [1]: import pandas as pd
```

```
In [2]: import numpy as np
```

```
In [3]: import matplotlib.pyplot as plt
import seaborn as sns
```

```
In [14]: dataframe = pd.read_csv("G:\Data Science\DA_Projects\Zomato-data.csv")
print(dataframe.head())
```

	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	

	approx_cost(for two people)	listed_in(type)
0	800	Buffet
1	800	Buffet
2	800	Buffet
3	300	Buffet
4	600	Buffet

```
In [ ]: #Before proceeding, let's convert the data type of the "rate" column to float
```

```
In [15]: def handleRate(value):
    value=str(value).split('/')
    value=value[0];
    return float(value)

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

	name	online_order	book_table	rate	votes	\
0	Jalsa	Yes	Yes	4.1	775	
1	Spice Elephant	Yes	No	4.1	787	
2	San Churro Cafe	Yes	No	3.8	918	
3	Addhuri Udupi Bhojana	No	No	3.7	88	
4	Grand Village	No	No	3.8	166	

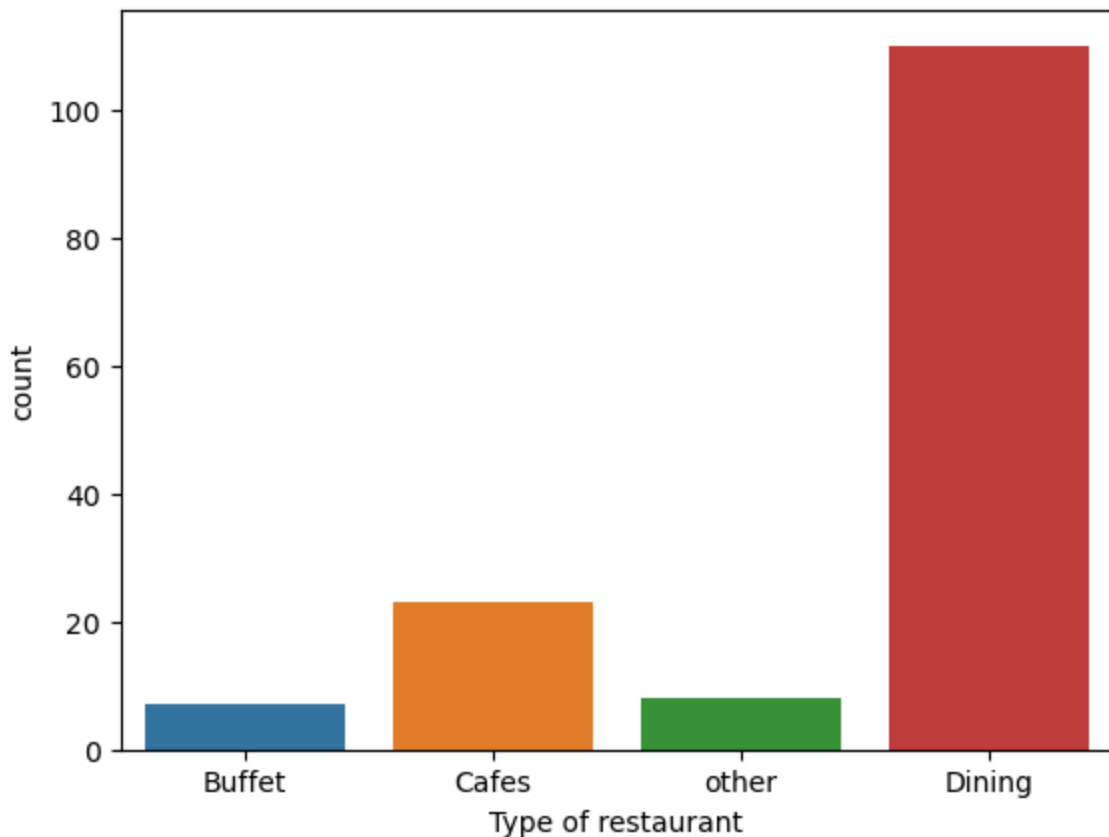
	approx_cost(for two people)	listed_in(type)
0	800	Buffet
1	800	Buffet
2	800	Buffet
3	300	Buffet
4	600	Buffet

```
In [16]: 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 [17]: sns.countplot(x=dataframe['listed_in(type)'])
plt.xlabel("Type of restaurant")
```

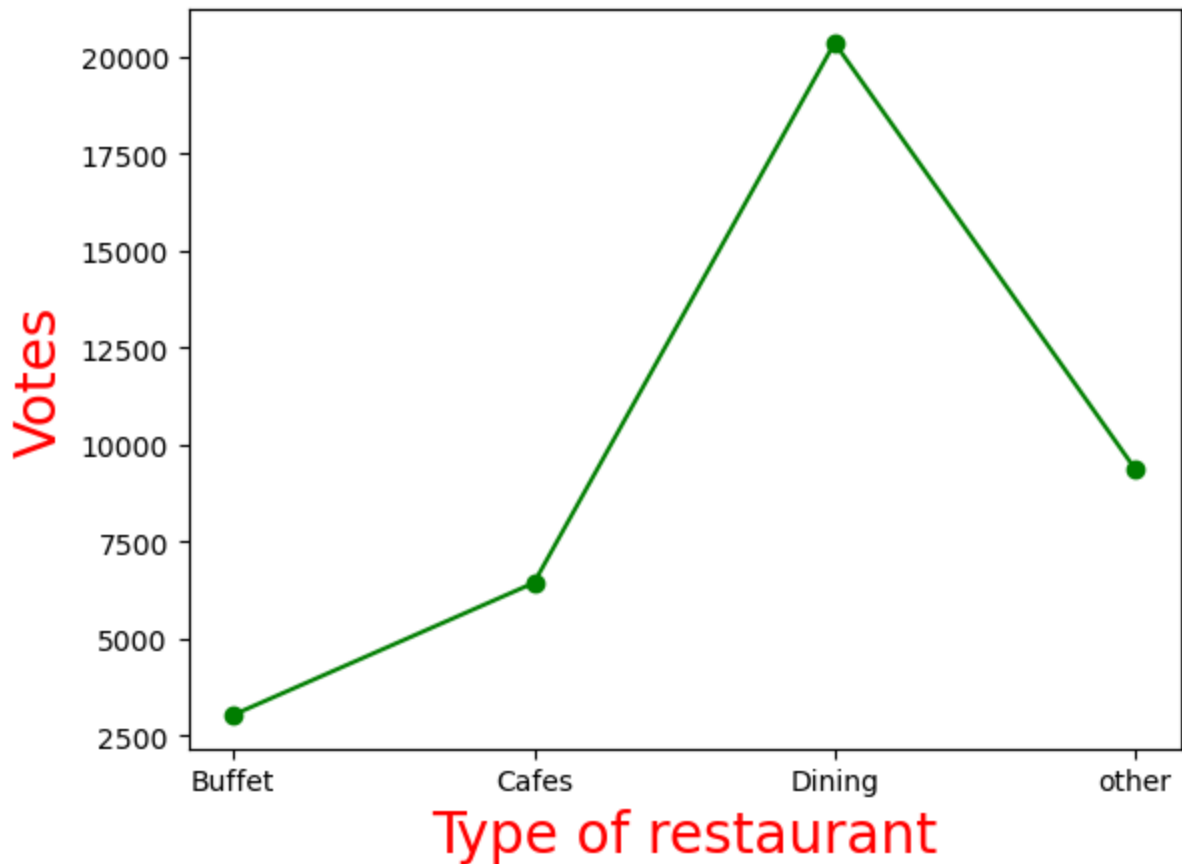
```
Out[17]: Text(0.5, 0, 'Type of restaurant')
```



```
In [ ]: #Conclusion: The majority of the restaurants fall into the dining category.
```

```
In [18]: 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[18]: Text(0, 0.5, 'Votes')
```



In []: *#Conclusion: Dining restaurants are preferred by a larger number of individuals*

In []: *#Now we will determine the restaurant's name that received the maximum votes*

```
In [19]: 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:

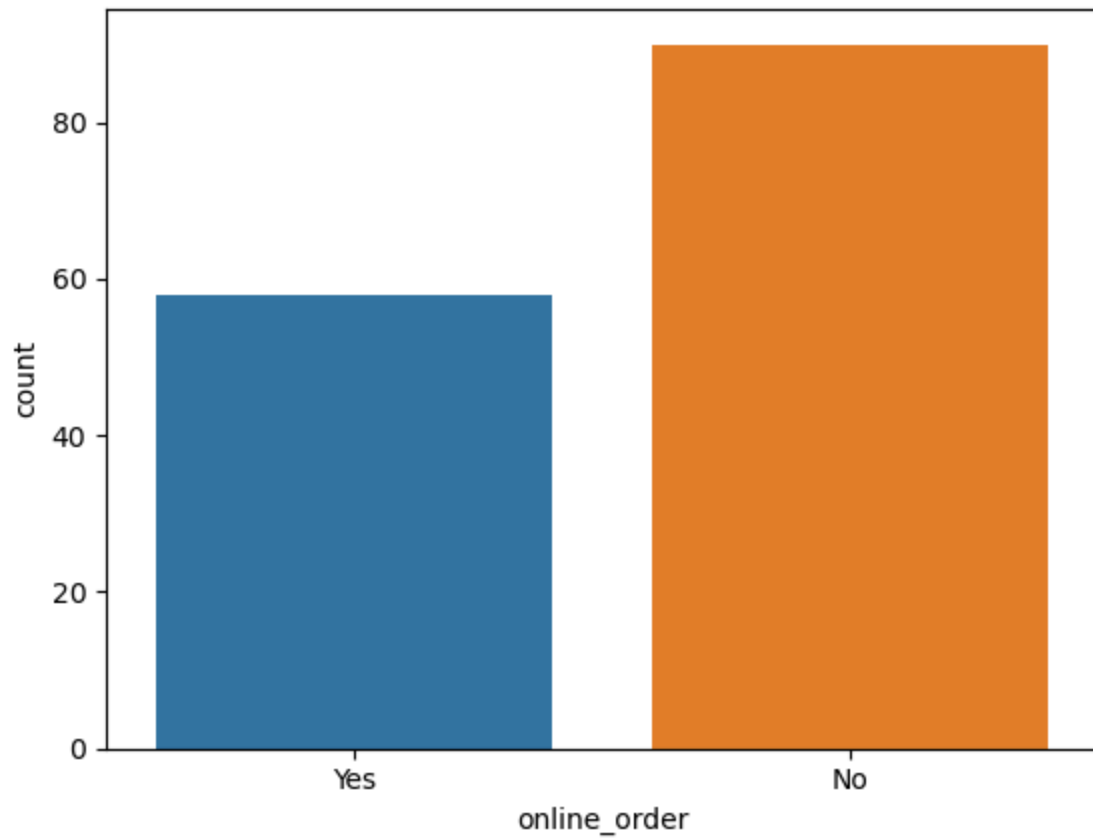
38 Empire Restaurant

Name: name, dtype: object

In []: *#Let's explore the online_order column.*

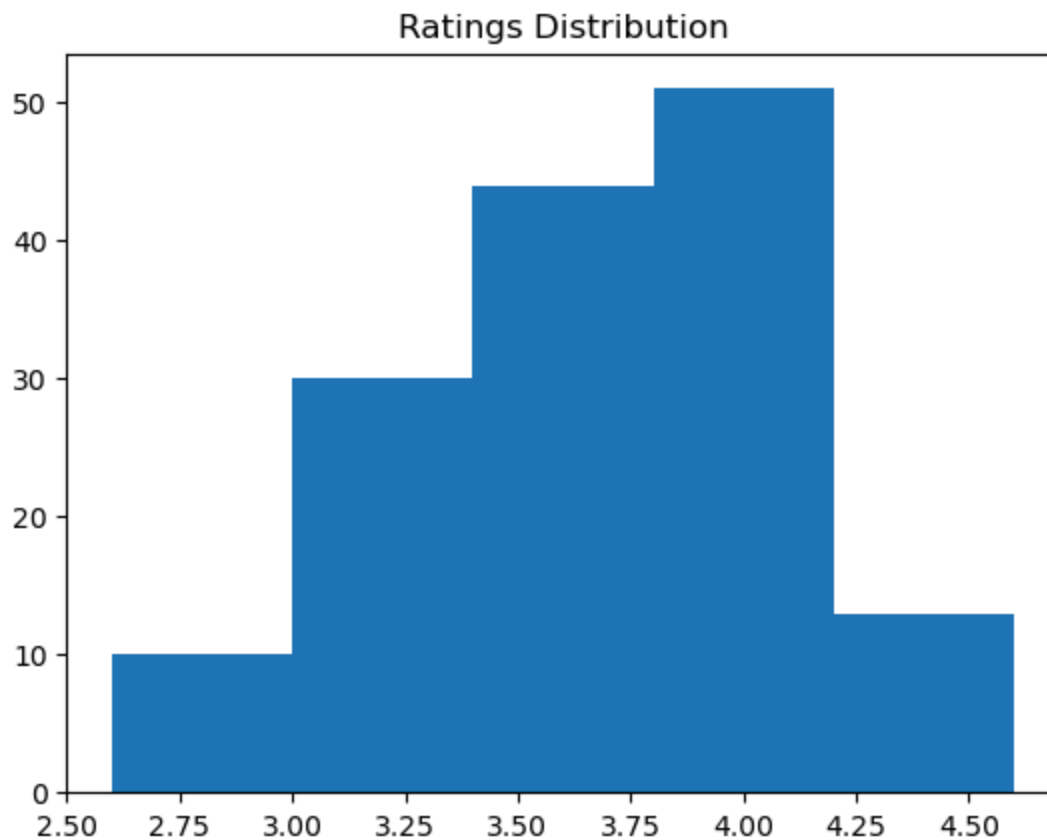
```
In [20]: sns.countplot(x=dataframe['online_order'])
```

Out[20]: <Axes: xlabel='online_order', ylabel='count'>



```
In [ ]: #Conclusion: This suggests that a majority of the restaurants do not accept
```

```
In [21]: plt.hist(dataframe['rate'],bins=5)
plt.title('Ratings Distribution')
plt.show()
```

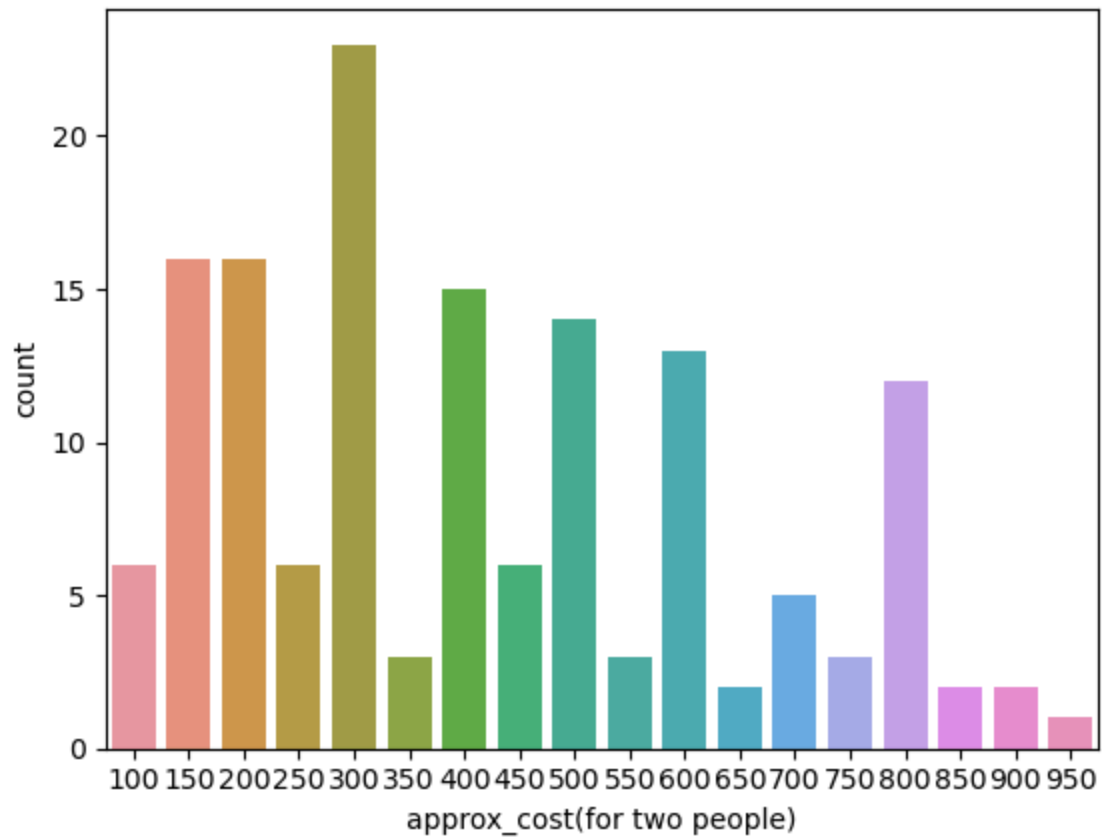


```
In [ ]: #Conclusion: The majority of restaurants received ratings ranging from 3.5 t
```

```
In [ ]: #Let's explore the approx_cost(for two people) column.
```

```
In [22]: couple_data=dataframe['approx_cost(for two people)']  
sns.countplot(x=couple_data)
```

```
Out[22]: <Axes: xlabel='approx_cost(for two people)', ylabel='count'>
```

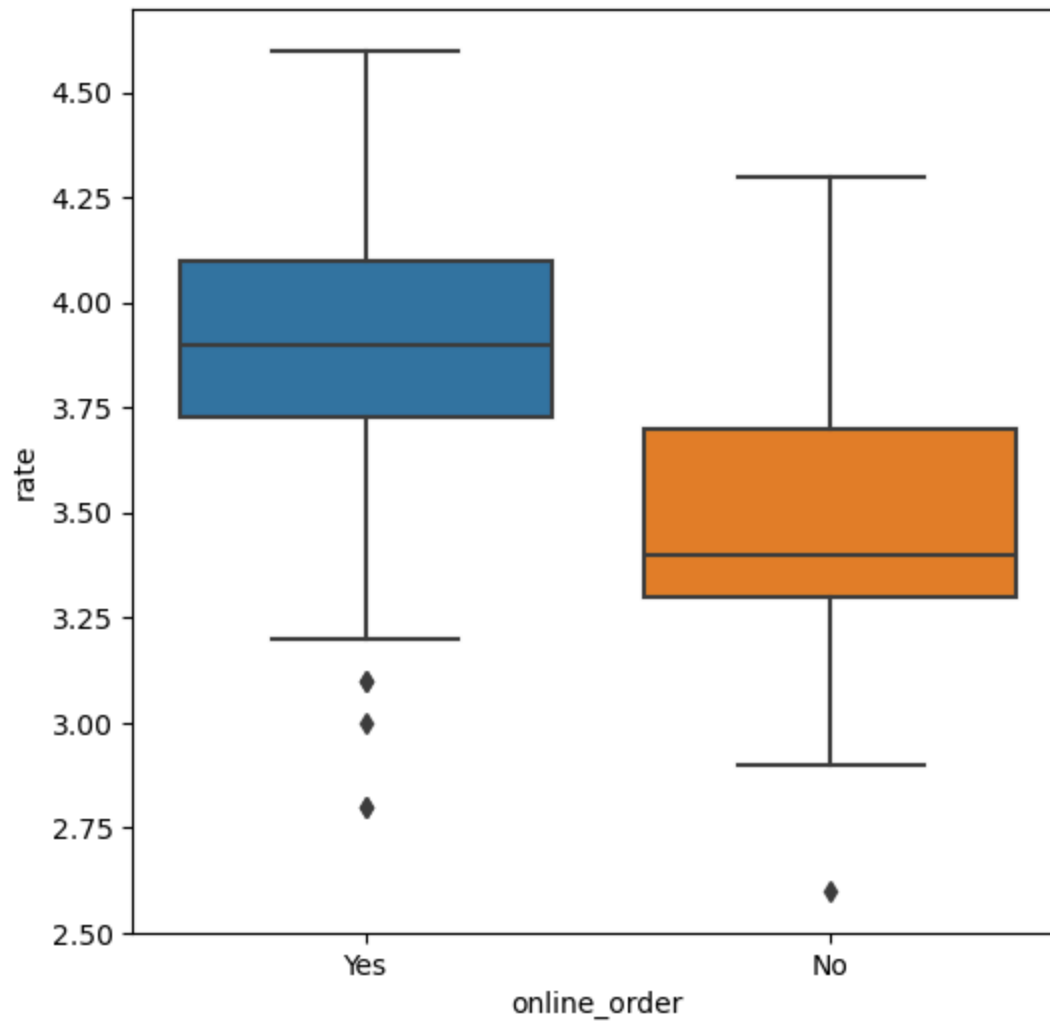


```
In [ ]: #Let's explore the approx_cost(for two people) column.
```

```
In [ ]: #Now we will examine whether online orders receive higher ratings than offli
```

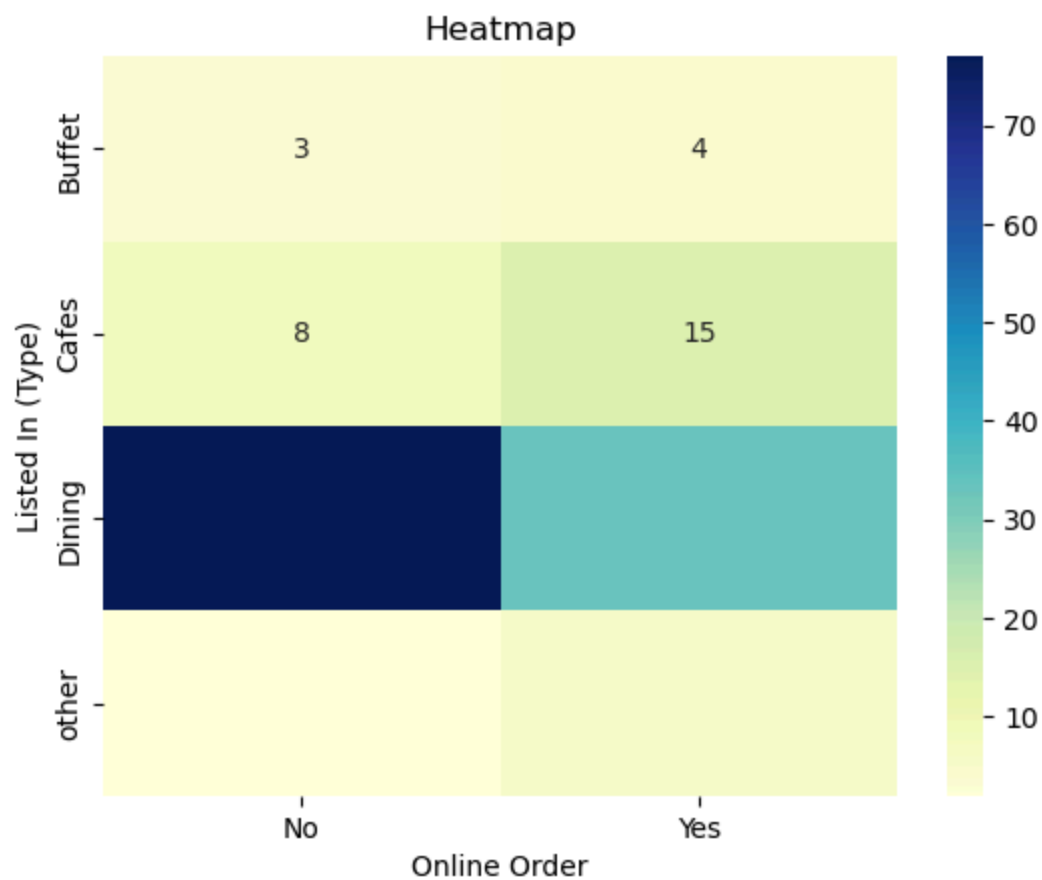
```
In [23]: plt.figure(figsize = (6,6))
sns.boxplot(x = 'online_order', y = 'rate', data = dataframe)
```

```
Out[23]: <Axes: xlabel='online_order', ylabel='rate'>
```



In []: *#CONCLUSION: Offline orders received lower ratings in comparison to online c*

```
In [24]: 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()
```



In []: *#CONCLUSION: Dining restaurants primarily accept offline orders, whereas cafes accept both online and offline orders.
#This suggests that clients prefer to place orders in person at restaurants,*