```
In [9]:    1 import numpy as np
           2 import pandas as pd
           3
           4 import matplotlib.pyplot as plt
           5 import seaborn as sns
```

```
In [10]:   1 data=pd.read_csv(r"C:\Users\abhis\Downloads\Zomato data .csv")
```

```
In [11]:   1 data.head()
```

Out[11]:

	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 [12]:   1 data.describe()
```

Out[12]:

	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 [14]:   1  data.dtypes
```

```
Out[14]:  name                          object
          online_order                  object
          book_table                    object
          rate                          object
          votes                          int64
          approx_cost(for two people)    int64
          listed_in(type)               object
          dtype: object
```

```
In [ ]:    1  # First converting datatype of rate from object to  flot value and removing denominator
```

```
In [18]:   1  def HandleRate(value):
           2      value=str(value).split('/')
           3      value=value[0]
           4      return float(value)
           5
           6  data['rate']=data['rate'].apply(HandleRate)
           7
           8
```

```
In [19]:   1  data.head()
```

Out[19]:

	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 [20]:   1  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 [ ]:   1  # no null value is there
```

```
In [ ]:   1  #Type of restaurant
```

In [23]:
```
1 sns.countplot(x=data['listed_in(type)'])
2 plt.xlabel("Type of Restaurant")
```

Out[23]: Text(0.5, 0, 'Type of Restaurant')
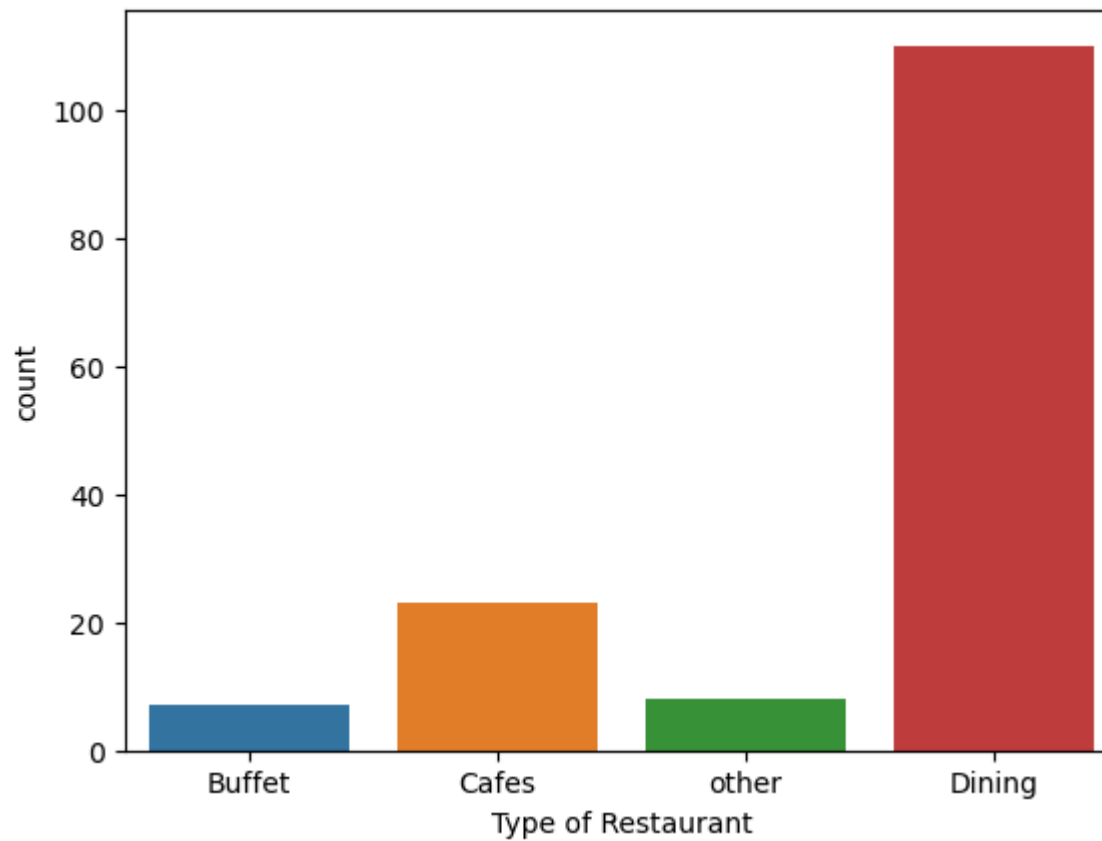
In [ ]:
```
1 # the majority of the taurants fallinto dining category
```

In [ ]:
```
1 #Dining restaurants preferred by a larger number of people
```

In [26]:
```python
1  grouped_data=data.groupby('listed_in(type)')['votes'].sum()
2  result=pd.DataFrame({'votes': grouped_data})
3  plt.plot(result,c="green",marker="o")
4  plt.xlabel('Type of Restaurant',c="red",size=20)
5  plt.ylabel("votes",c="red",size=20)
```

Out[26]: Text(0, 0.5, 'votes')
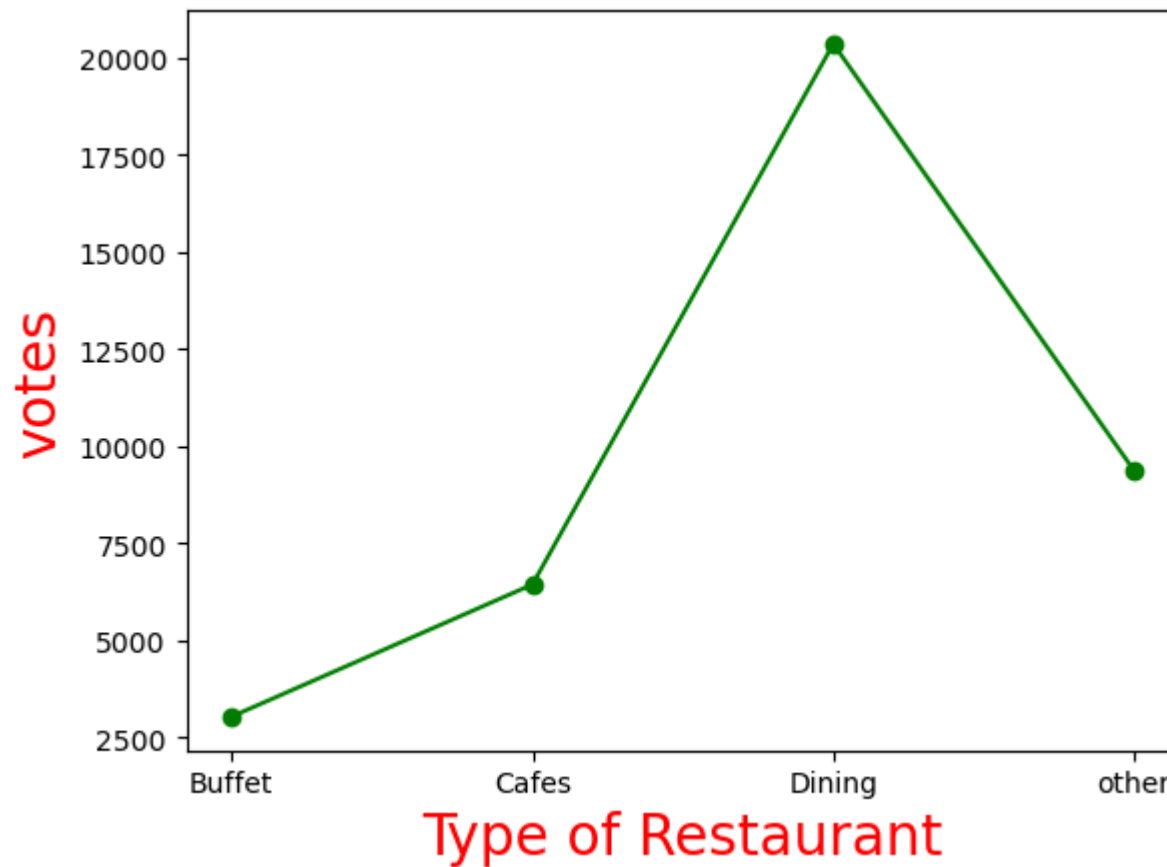
In [ ]:
```python
1  #The majority of restaurants received ratings
```

```
In [27]:    1  plt.hist(data['rate'],bins=5)
            2  plt.title('Ratings Distribution')
            3  plt.show()
```
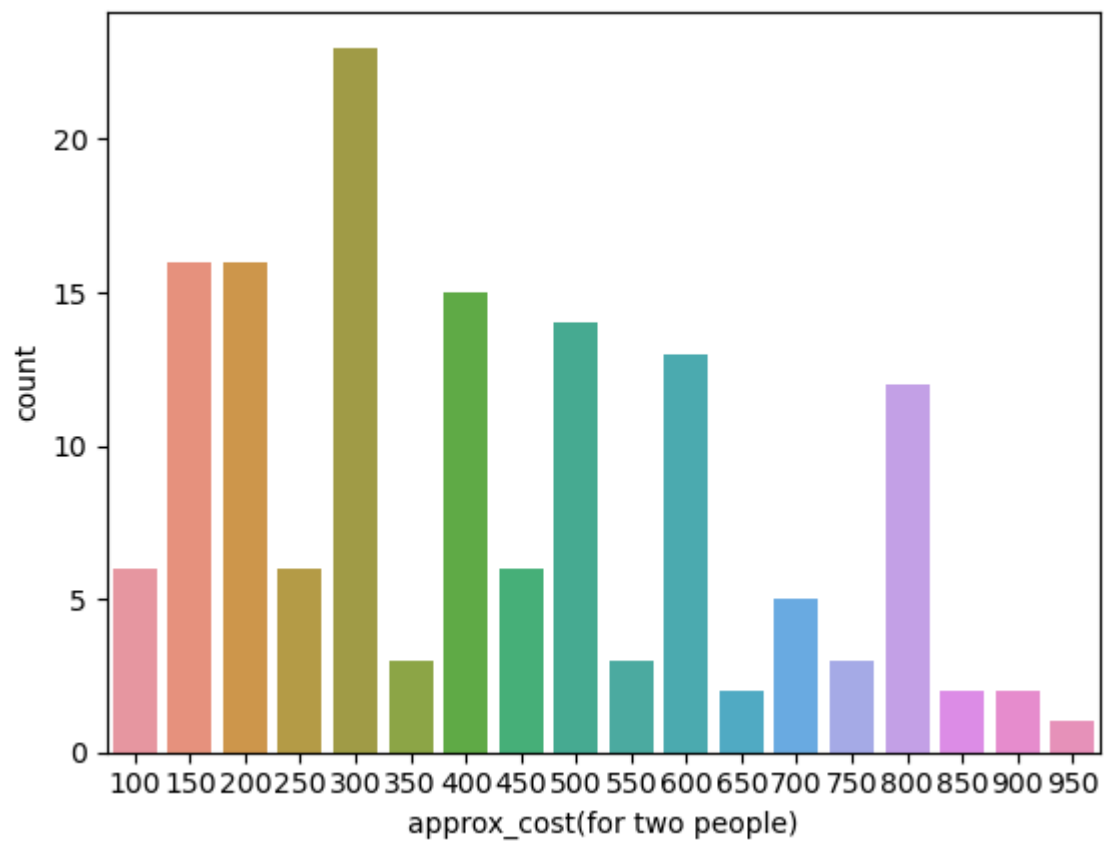

Ratings Distribution
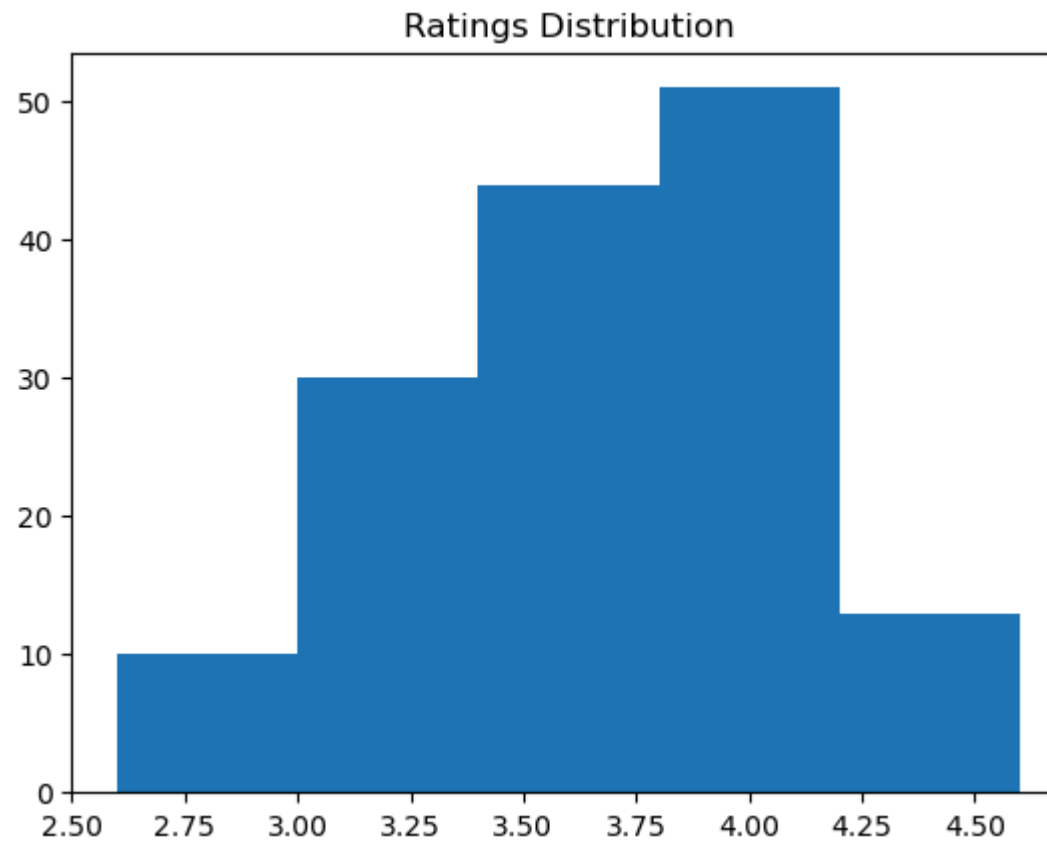
```
In [29]:    1  #conclusion: The majority of restaurants received ratings ranging from 3.5 to 4.
```

```
In [30]:    1  #The majority of couples prefer restaurants with an approximate cost of 300 rupees.
```

```
In [31]:  1  couple_data=data['approx_cost(for two people)']
          2  sns.countplot(x=couple_data)
```
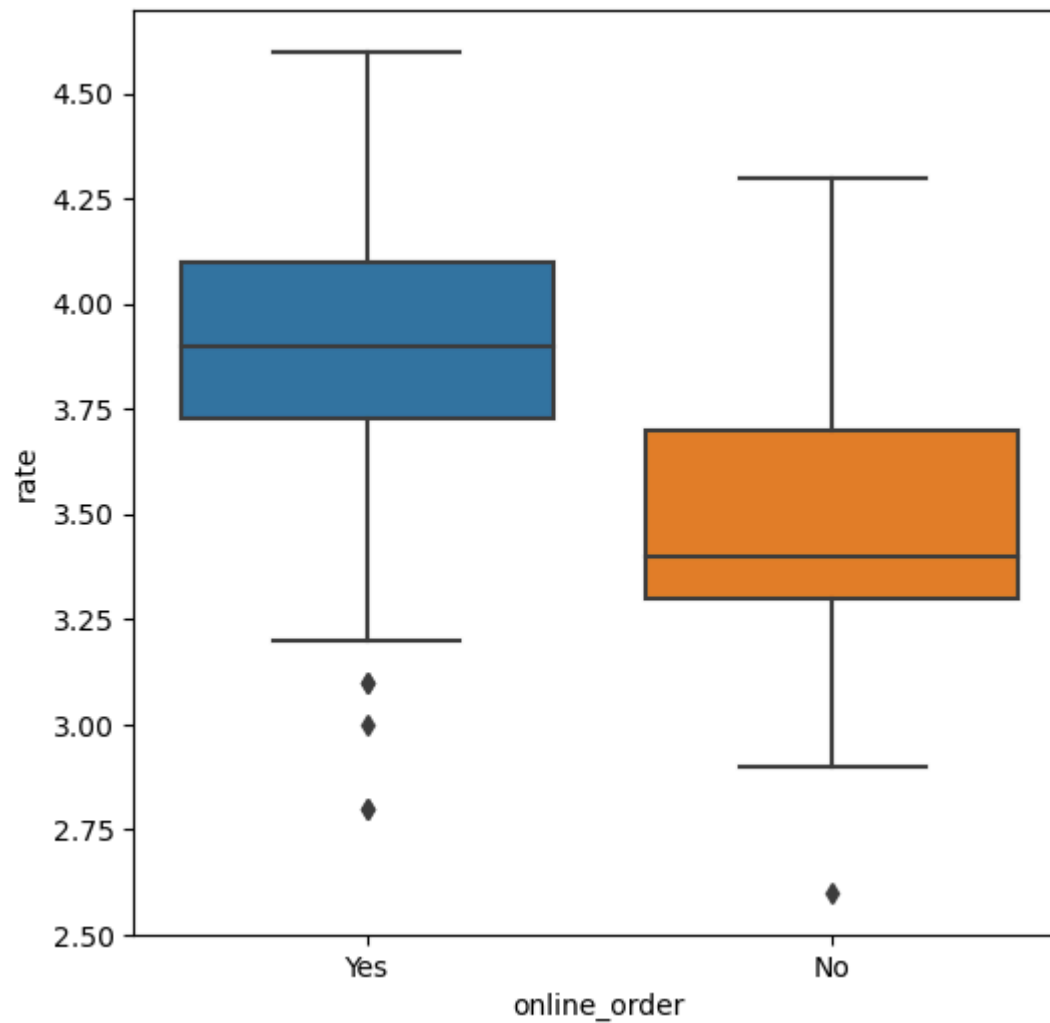
Out[31]: <Axes: xlabel='approx_cost(for two people)', ylabel='count'>


```
In [ ]:  1  #whether online orders receive higher ratings than offline orders
```

```
1 plt.figure(figsize=(6,6))
2 sns.boxplot(x='online_order', y='rate',data=data)
```

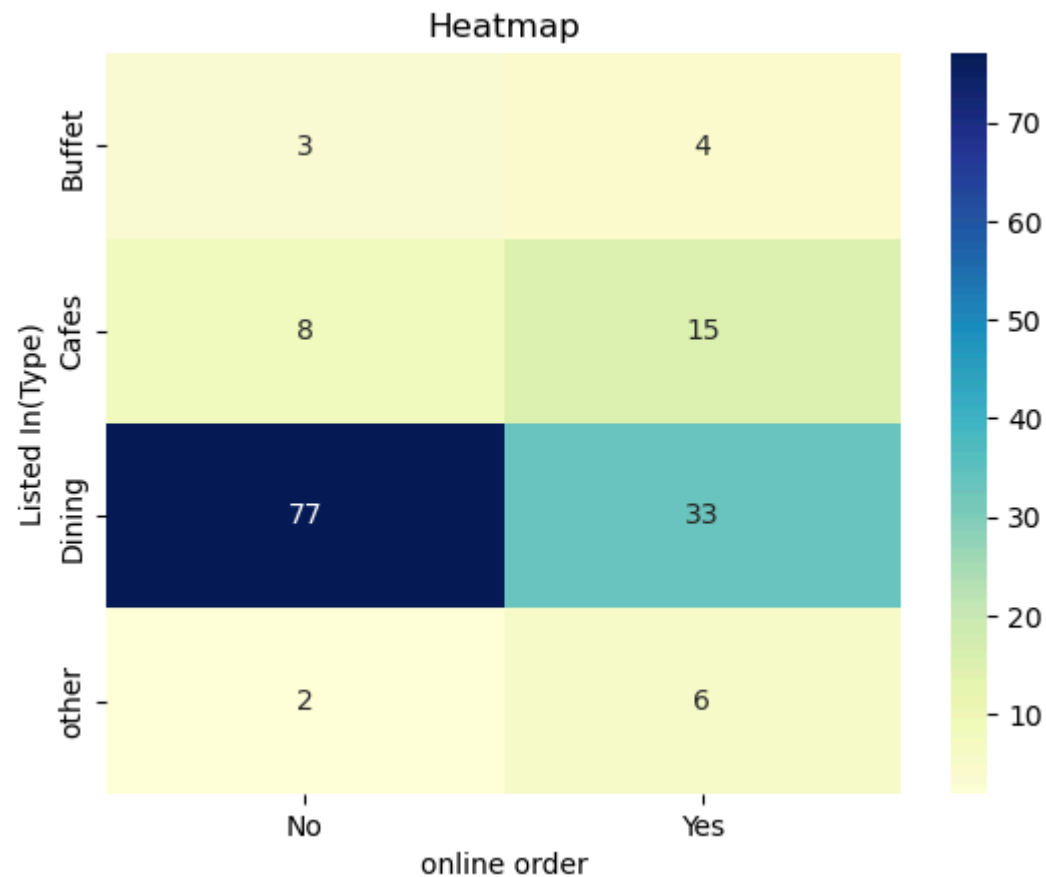Out[32]: <Axes: xlabel='online_order', ylabel='rate'>

In [ ]:

```
1 #conclusion: offline orders received lower ratings in comparison to online orders,which obtained excellent rating
```

```
In [33]:   1
           2  pivot_table=data.pivot_table(index='listed_in(type)',columns='online_order',aggfunc='size',fill_value=0)
           3  sns.heatmap(pivot_table,annot=True,cmap="YlGnBu",fmt='d')
           4
           5  plt.title('Heatmap')
           6  plt.xlabel("online order")
           7  plt.ylabel("Listed In(Type)")
           8  plt.show()
```


In [ ]:   1