[2]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

[4]:
```
#Step-2) Creating the data frame
```

[6]:
```
dataframe = pd.read_csv("Zomato data .csv")
print(dataframe)
```

```
                    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
```

[7]:
```
dataframe
```

[7]:

	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
...	...	...	...	...	...	...	...
143	Melting Melodies	No	No	3.3/5	0	100	Dining
144	New Indraprasta	No	No	3.3/5	0	150	Dining
145	Anna Kuteera	Yes	No	4.0/5	771	450	Dining
146	Darbar	No	No	3.0/5	98	800	Dining
147	Vijayalakshmi	Yes	No	3.9/5	47	200	Dining

148 rows × 7 columns

[8]:
```
#Step-3) Data cleaning & data transformation
```

[10]:
```
def handleRate(value):
    value = str(value).split('/')
    value = value[0]
    return float(value)

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

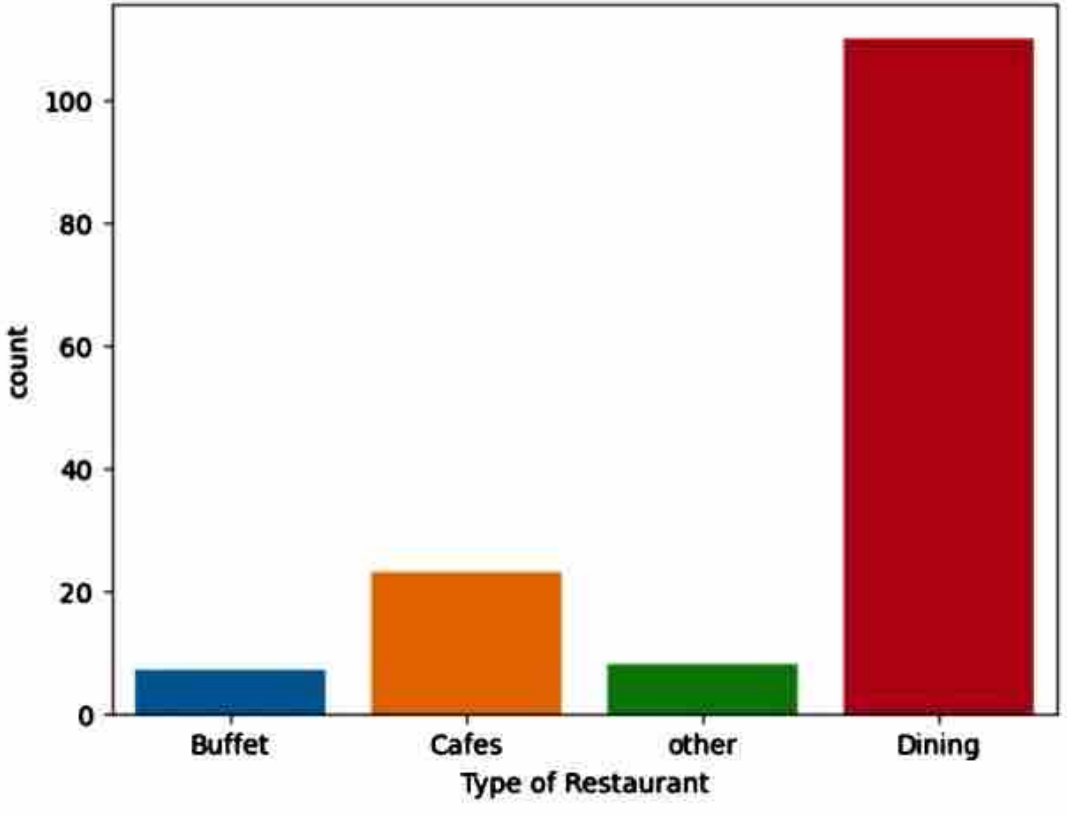
```
                  name online_order book_table  rate  votes  \
0                Jalsa          Yes        Yes   4.1    775
1        Spice Elephant          Yes         No   4.1    787
2       San Churro Cafe          Yes         No   3.8    918
3  Addhuri Udupi Bhojana           No         No   3.7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

[11]:
```
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
```

[14]:
```
sns.countplot(x=dataframe['listed_in(type)'])
plt.xlabel("Type of Restaurant")
```

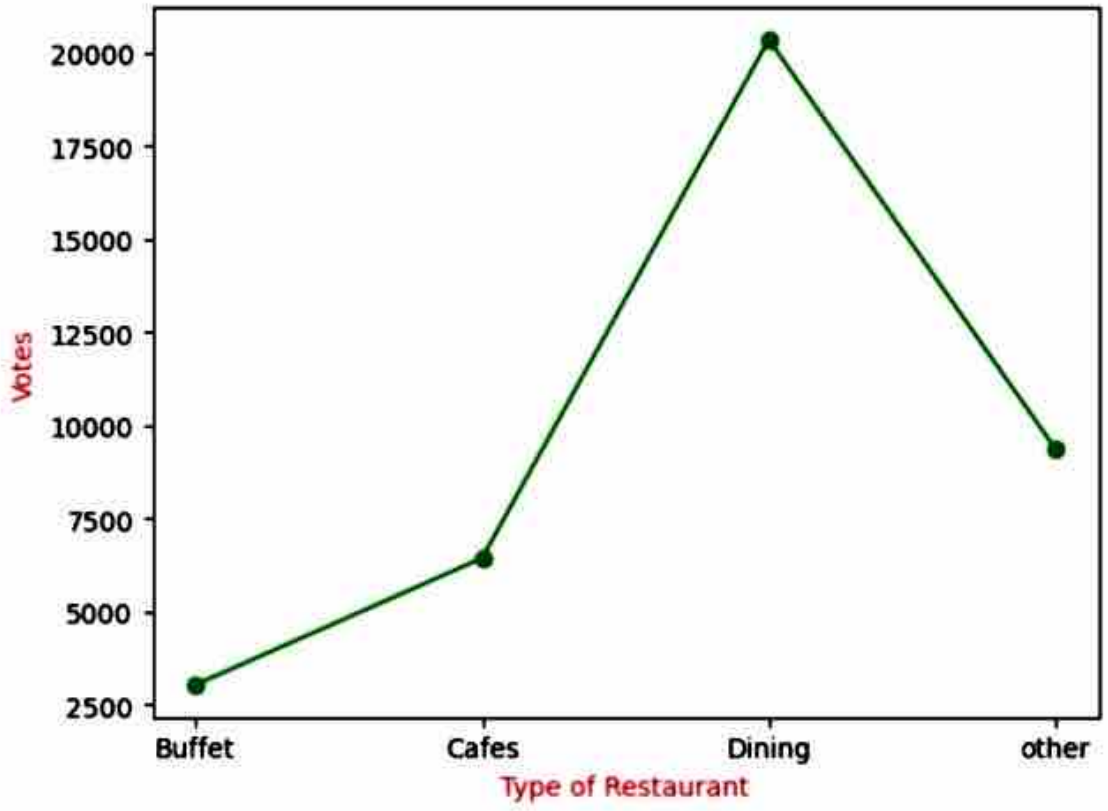
[14]:
```
Text(0.5, 0, 'Type of Restaurant')
```



[15]:
```
#Conclusion: Majority of the restaurant falls in dinning category
```

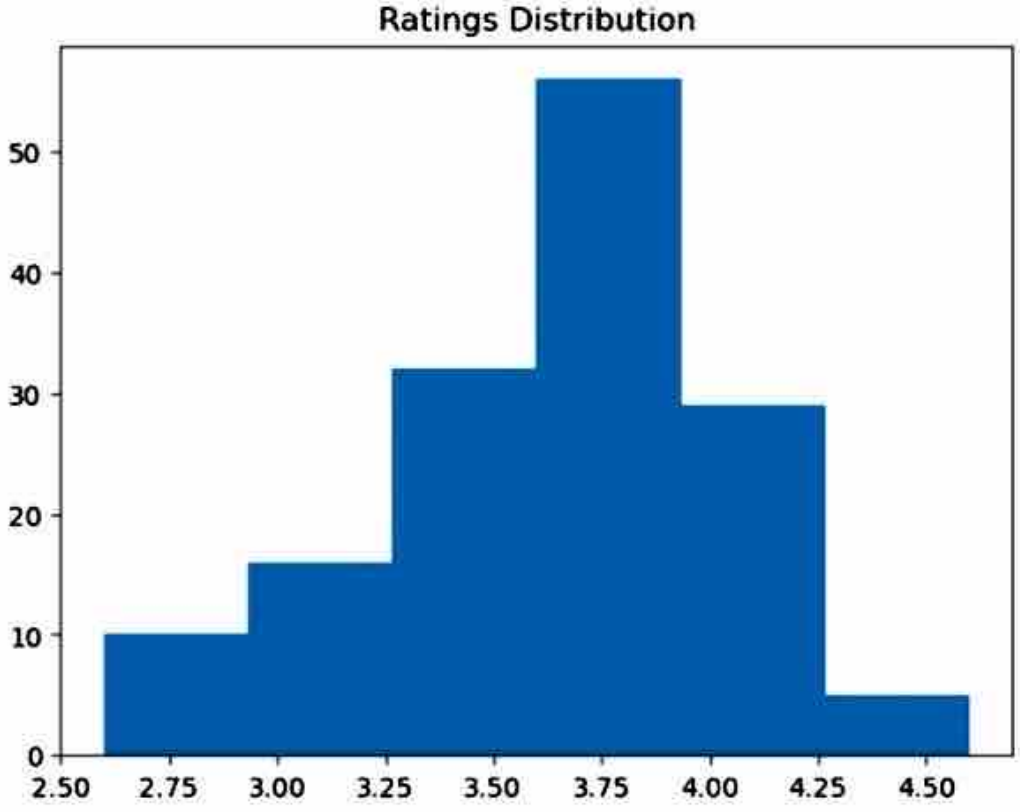
[23]:
```
grouped_data = dataframe.groupby('listed_in(type)')['votes'].sum()
result = pd.DataFrame({'votes':grouped_data})
plt.plot(result, c="green", marker="o")
plt.xlabel("Type of Restaurant", c="red", size=10)
plt.ylabel("Votes", c="red", size=10)
```

[23]:
```
Text(0, 0.5, 'Votes')
```



[24]:
```
#Conclusion: Dining restaurants has recieved maximum votes
```

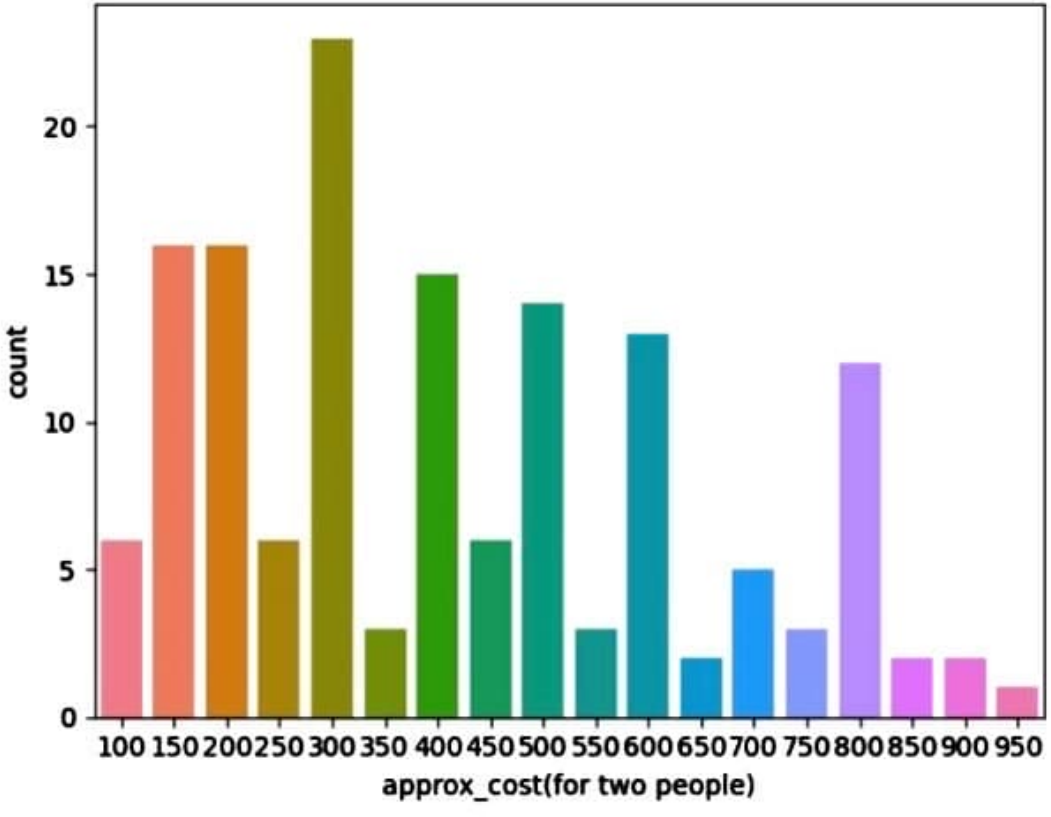
[28]:
```
plt.hist(dataframe['rate'], bins= 6)
plt.title("Ratings Distribution")
plt.show()
```



[29]:
```
#Conclusion: The majority restaurants recieved ratings from 3.5 to 4
```

[31]:
```
couple_data=dataframe['approx_cost(for two people)']
sns.countplot(x=couple_data)
```

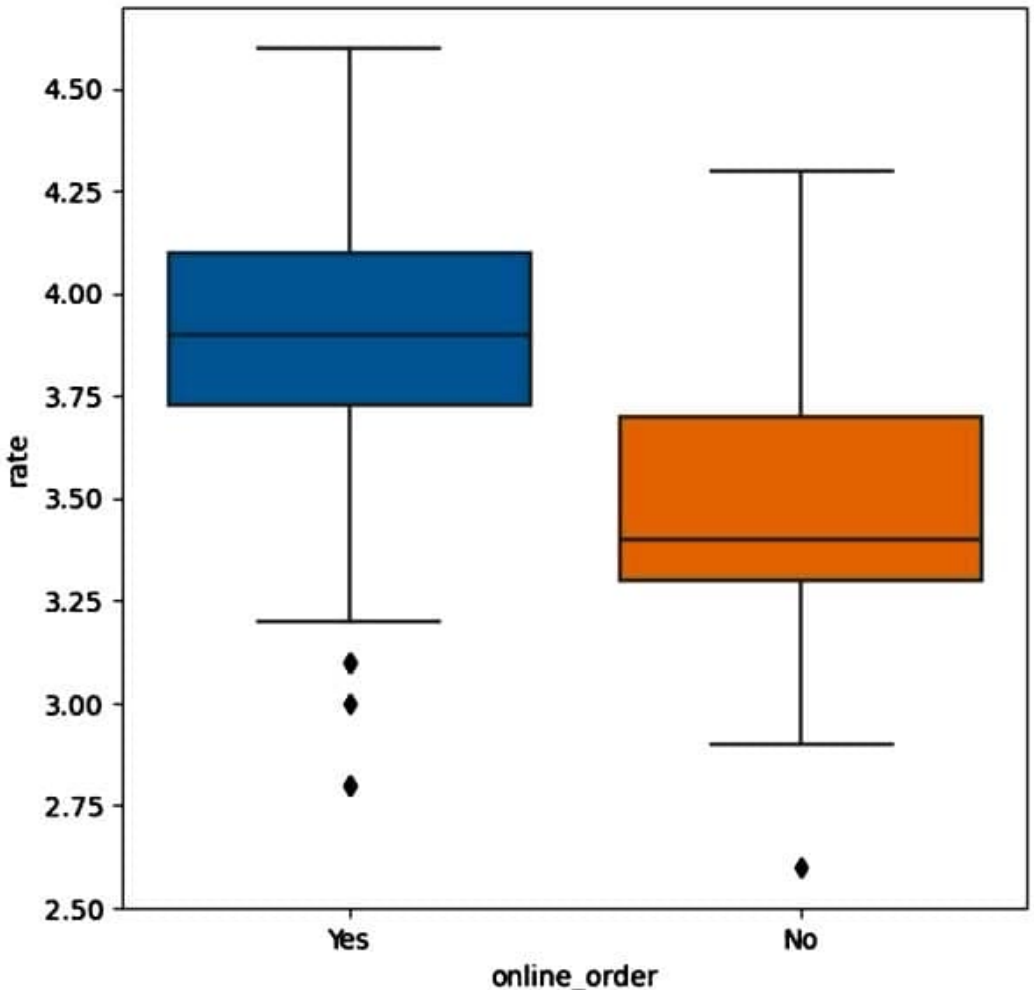
[31]:
```
<Axes: xlabel='approx_cost(for two people)', ylabel='count'>
```



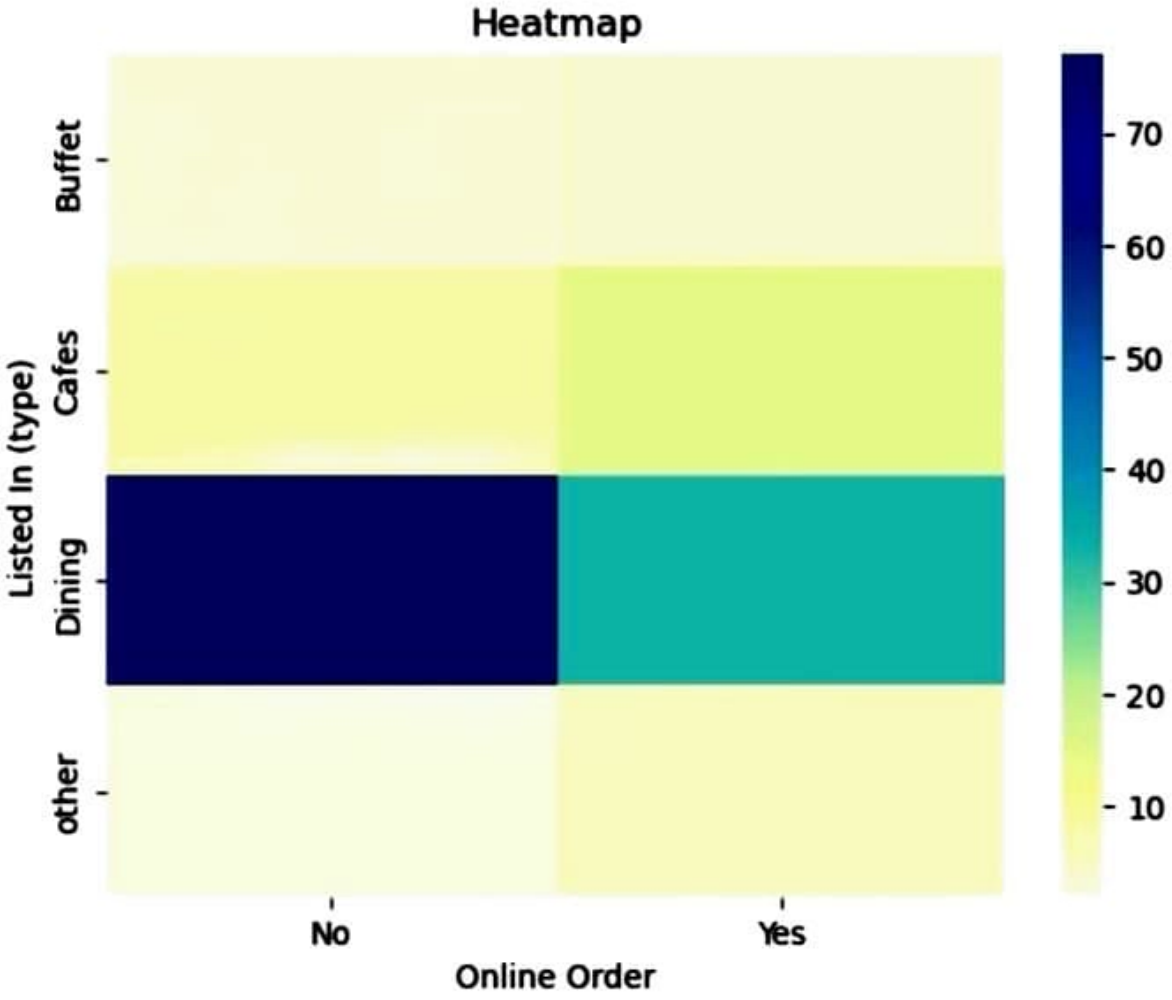
[32]:
```
#Conclusion: The majority of couples prefer restaurants with an approximate cost of 300 rupees
```

[34]:
```
plt.figure(figsize = (6,6))
sns.boxplot(x = 'online_order', y = 'rate', data = dataframe)
```

[34]:
```
<Axes: xlabel='online_order', ylabel='rate'>
```



[46]:
```
pivot_table = dataframe.pivot_table(index='listed_in(type)', columns='online_order', aggfunc='size', fill_value=0)
sns.heatmap(pivot_table, cmap="YlGnBu")
plt.title("Heatmap")
plt.xlabel("Online Order")
plt.ylabel("Listed In (type)")
plt.show()
```



[41]:
```
# Conclusion: Dinning restaurants primarily accept offline orders, whereas cafe primarily recieve online orders.
```