

```
In [4]: import pandas as pd
import seaborn as sns
import numpy as np
import matplotlib.pyplot as plt

In [6]: df=pd.read_csv("Zomato data .csv")
df
```

Out[6]:

	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 x 7 columns

```
In [7]: df.shape
Out[7]: (148, 7)
```

In [9]: df.head()

Out[9]:

	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 [8]: df.info
Out[8]: <bound method DataFrame.info of
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

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

[148 rows x 7 columns]>
```

```
In [11]: def rate_handle(value):
value=str(value).split("/")
value=value[0]
return float(value)
df["rate"]=df["rate"].apply(rate_handle)
print(df.head())

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 [12]: df.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
```

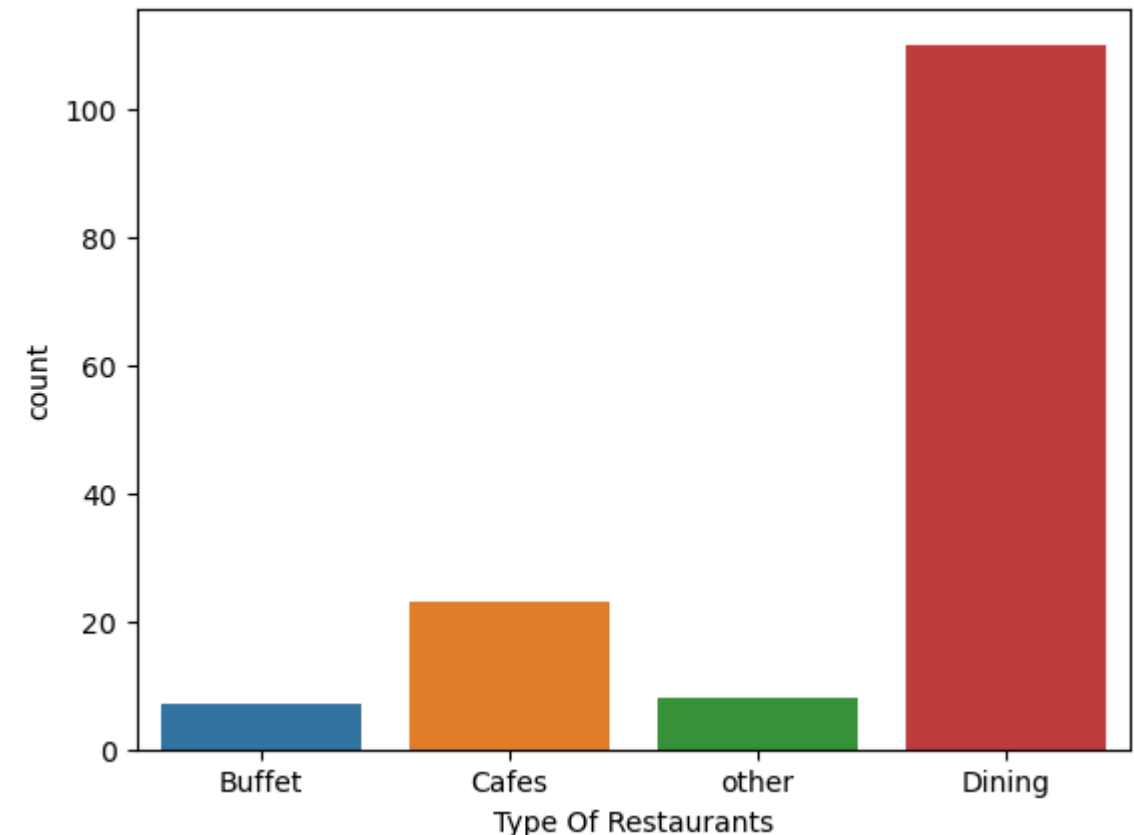
```
In [13]: df.isnull().sum()
Out[13]: name                0
online_order              0
book_table                0
rate                     0
votes                    0
approx_cost(for two people) 0
listed_in(type)           0
dtype: int64
```

```
In [19]: df.describe()
Out[19]:
```

	rate	votes	approx_cost(for two people)
count	148.000000	148.000000	148.000000
mean	3.633108	264.810811	418.243243
std	0.402271	653.676951	223.085098
min	2.600000	0.000000	100.000000
25%	3.300000	6.750000	200.000000
50%	3.700000	43.500000	400.000000
75%	3.900000	221.750000	600.000000
max	4.600000	4884.000000	950.000000

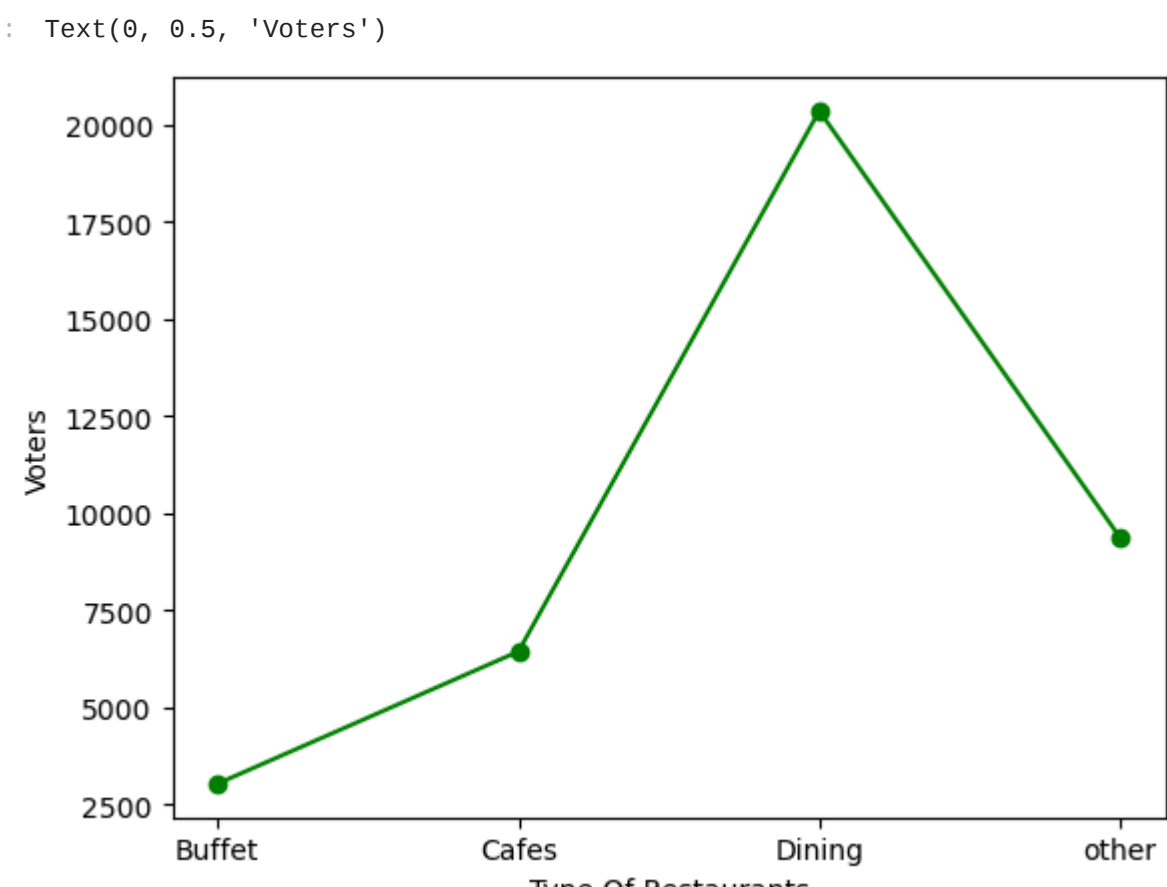
```
In [20]: df.dtypes
Out[20]: name                object
online_order             object
book_table               object
rate                    float64
votes                   int64
approx_cost(for two people) int64
listed_in(type)          object
dtype: object
```

```
In [90]: sns.countplot(x="listed_in(type)",data=df)
plt.xlabel("Type Of Restaurants")
plt.show()
```



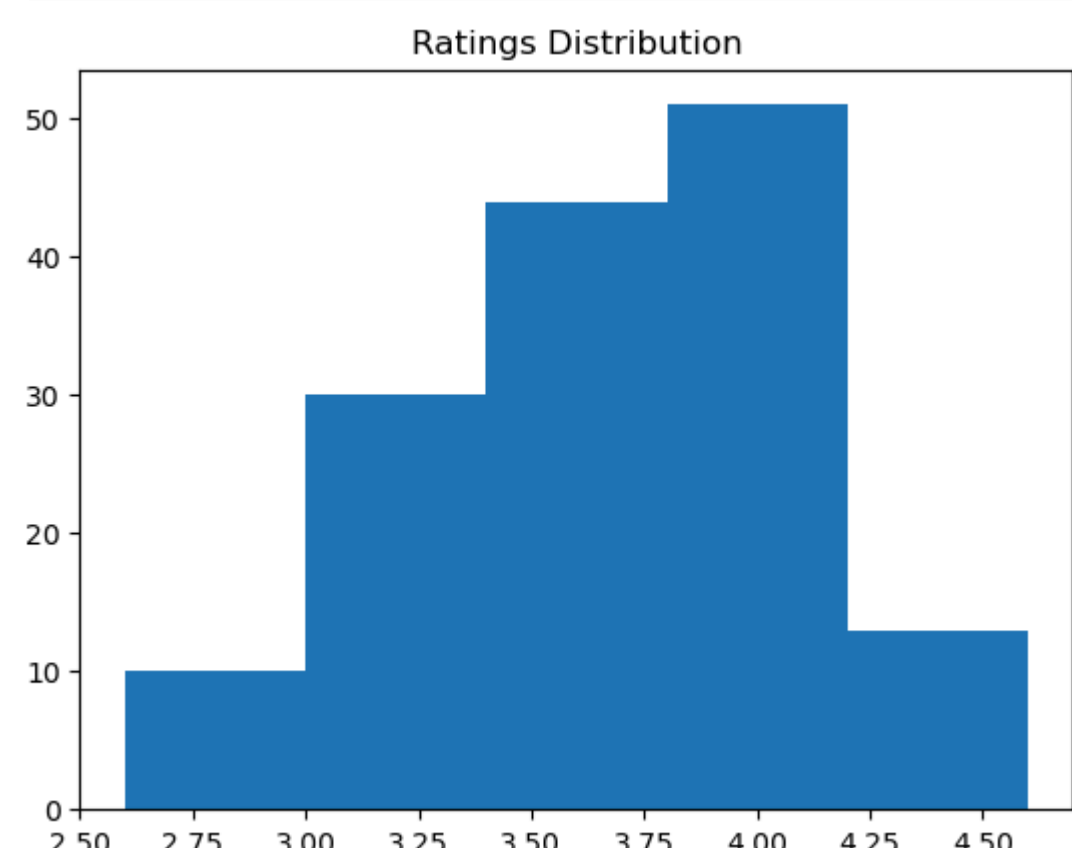
Conclusion: The majority of the restaurants fall into the dining category. Dining restaurants are preferred by a larger number of individuals.

```
In [36]: groupby_data=df.groupby("listed_in(type)")[ "votes"].sum()
result=pd.DataFrame({"votes":groupby_data})
plt.plot(result,c="green",marker="o")
plt.xlabel("Type Of Restaurants")
plt.ylabel("Voters")
```



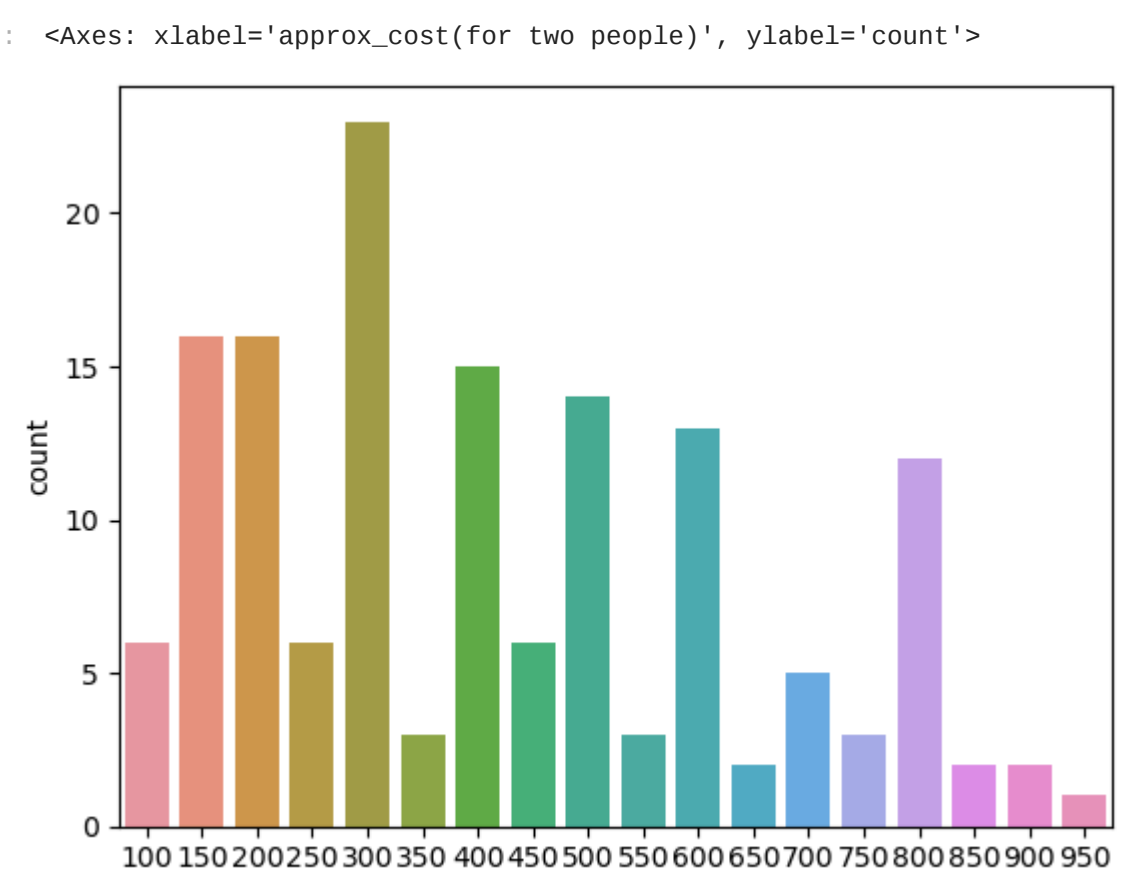
The majority of restaurants received ratings dining category.

```
In [45]: plt.hist(df["rate"],bins=5)
plt.title("Ratings Distribution")
plt.show()
```



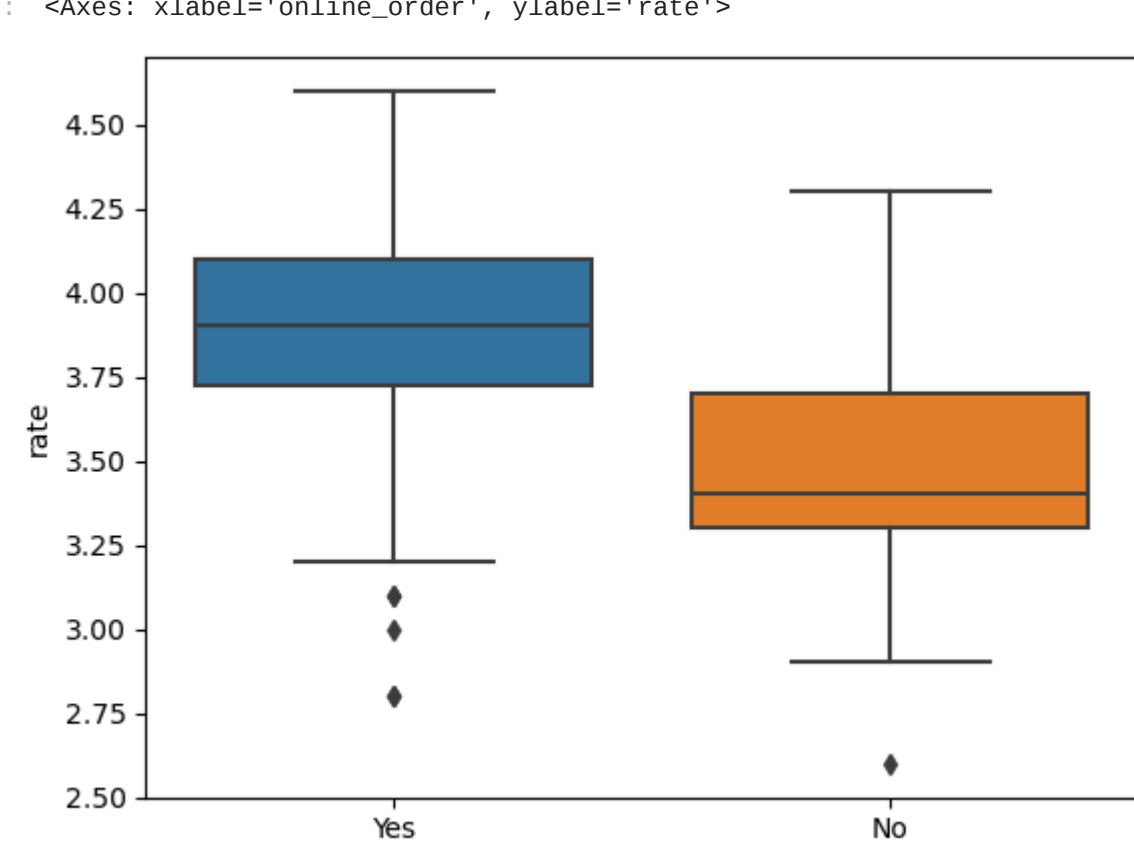
Conclusion: The majority of restaurants received ratings ranging from 3.5 to 4.

```
In [52]: data=df[["approx_cost(for two people)"]]
sns.countplot(x=data)
```



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

```
In [55]: sns.boxplot(x="online_order",y="rate",data=df)
Out[55]: <Axes: xlabel='online_order', ylabel='rate'>
```



CONCLUSION: Offline orders received lower ratings in comparison to online orders.

```
In [86]: data=df.pivot_table(index="listed_in(type)", columns="online_order",aggfunc="size",fill_value=0)
sns.heatmap(data,annot=True)
plt.title("Heatmap")
plt.xlabel("Online Order")
plt.ylabel("Listed in Type")
plt.show()
```

