Zomato Data Analysis Project

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

In [3]:
```
dataframe =pd.read_csv('Zomato data .csv')
```

In [4]:
```
dataframe
```

Out[4]:

	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

In [7]:
```
def handdlerate(value):
    value=str(value).split('/')
    value=value[0];
    return float(value)

dataframe['rate']=dataframe['rate'].apply(handdlerate)
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 [8]:
```
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
```

Type of Restraunt

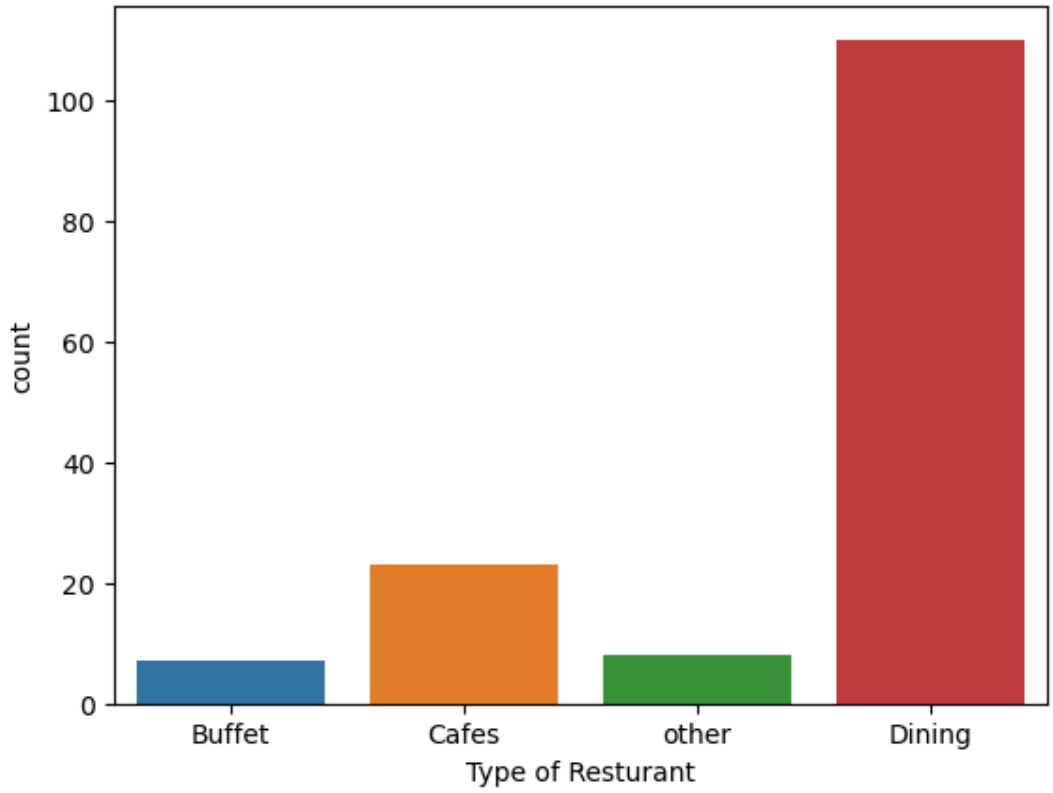
In [9]:
```
dataframe.head()
```

Out[9]:

	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 [11]:
```
sns.countplot(x=dataframe['listed_in(type)'])
plt.xlabel('Type of Resturant')
```

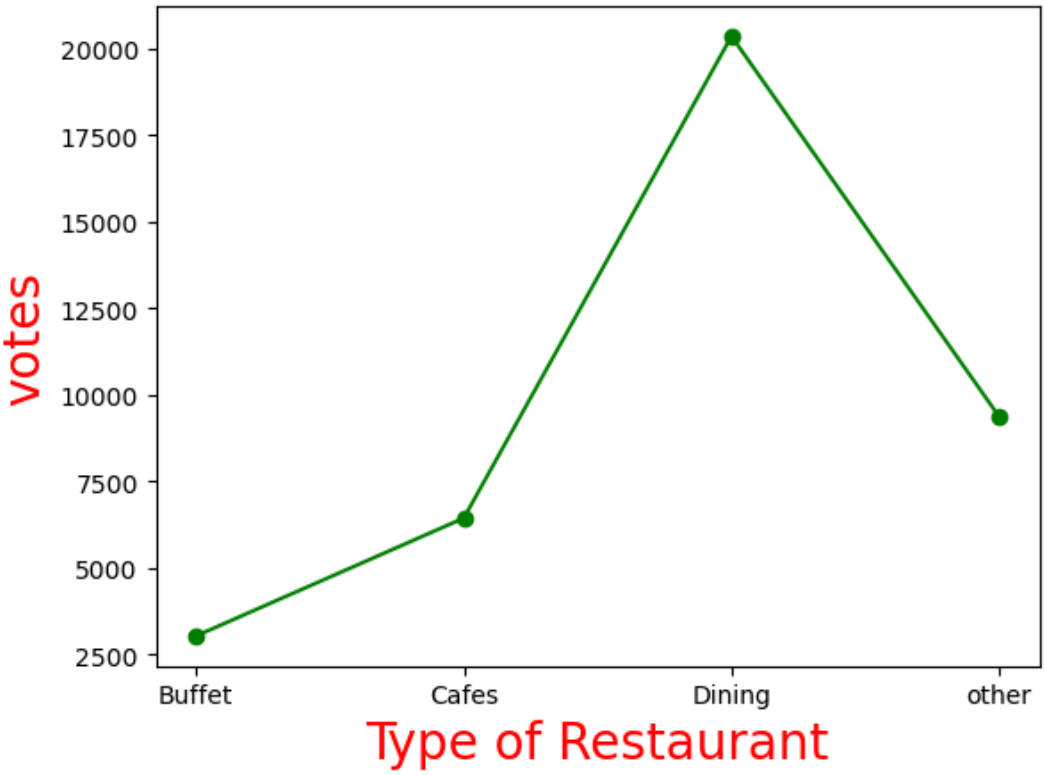
Out[11]:
```
Text(0.5, 0, 'Type of Resturant')
```



Conclusion : Majority of the resturant falls in Dinning category followed by Cafes

Total number of votes by Resturants

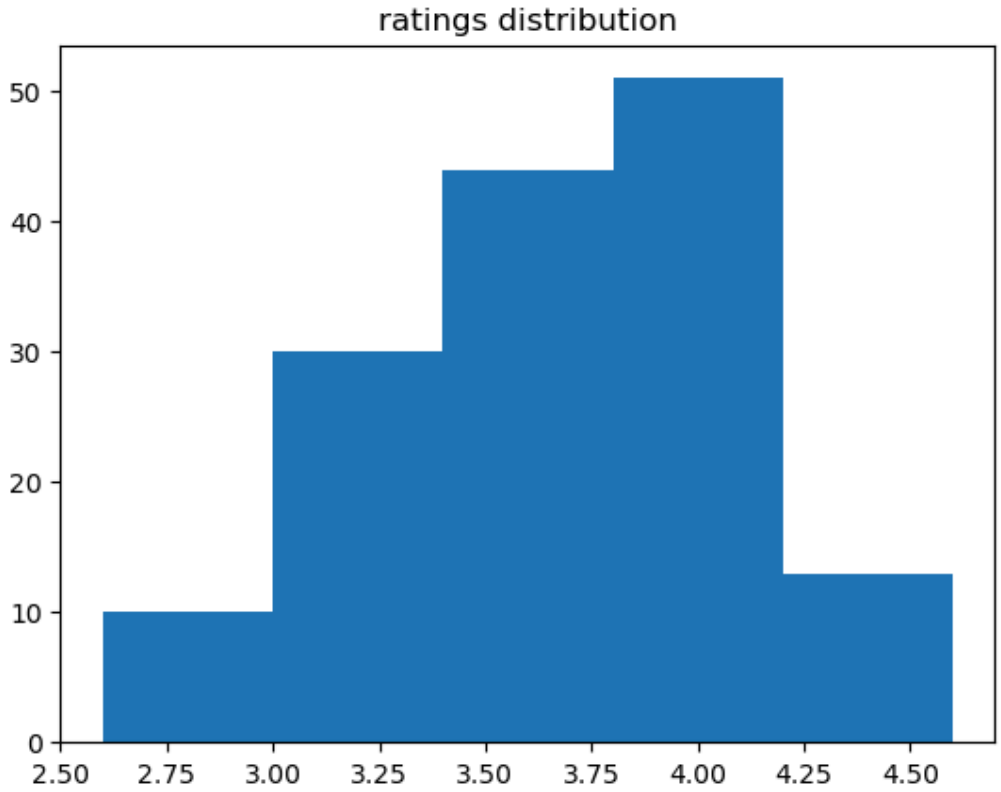
In [24]:
```
grouped_data = dataframe.groupby('listed_in(type)')['votes'].sum()
result = pd.DataFrame({'votes': grouped_data })
plt.plot(result , c='green' , marker='o')
plt.xlabel('Type of Restaurant', c='red', size = 20)
plt.ylabel('votes' ,c='red' , size = 20)
plt.show()
```



Conclusion : Dinning Resturants has recieved maximum votes

Frequency of the Ratings by the Resturants

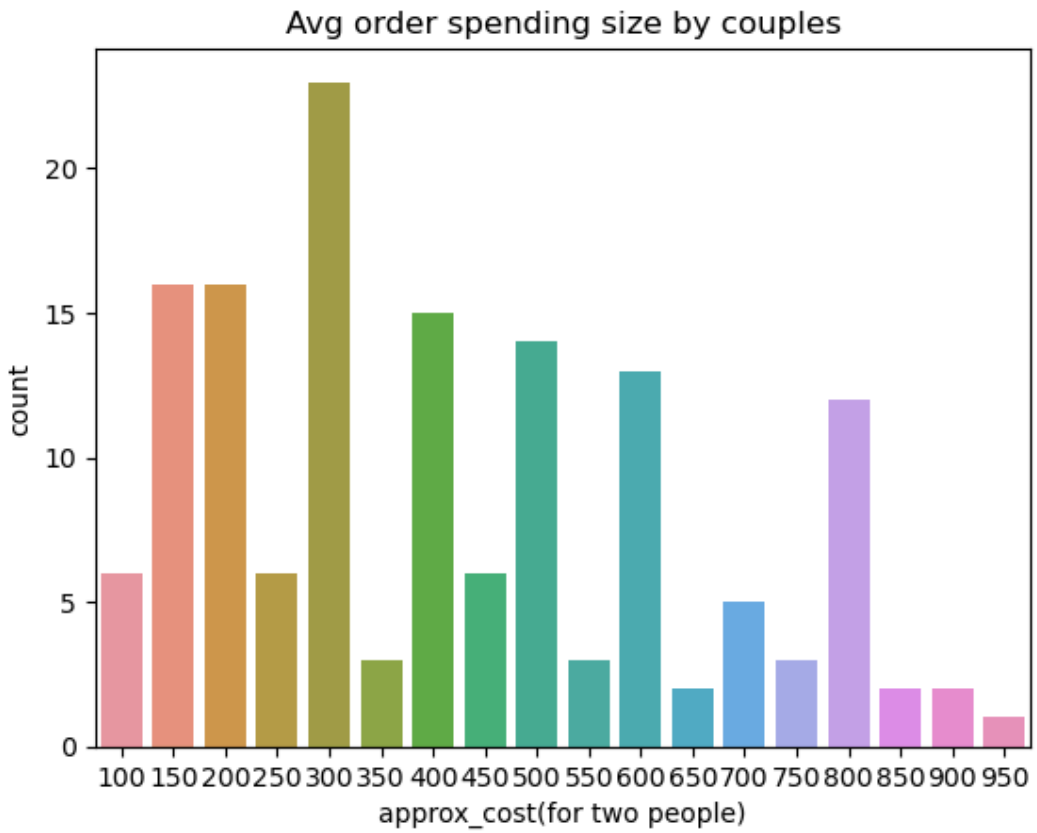
In [19]:
```
plt.hist(dataframe['rate'], bins =5)
plt.title('ratings distribution')
plt.show()
```



Conclusion : The majority of the resturants received ratings from 3.5 to 4

Average order spending size by couples

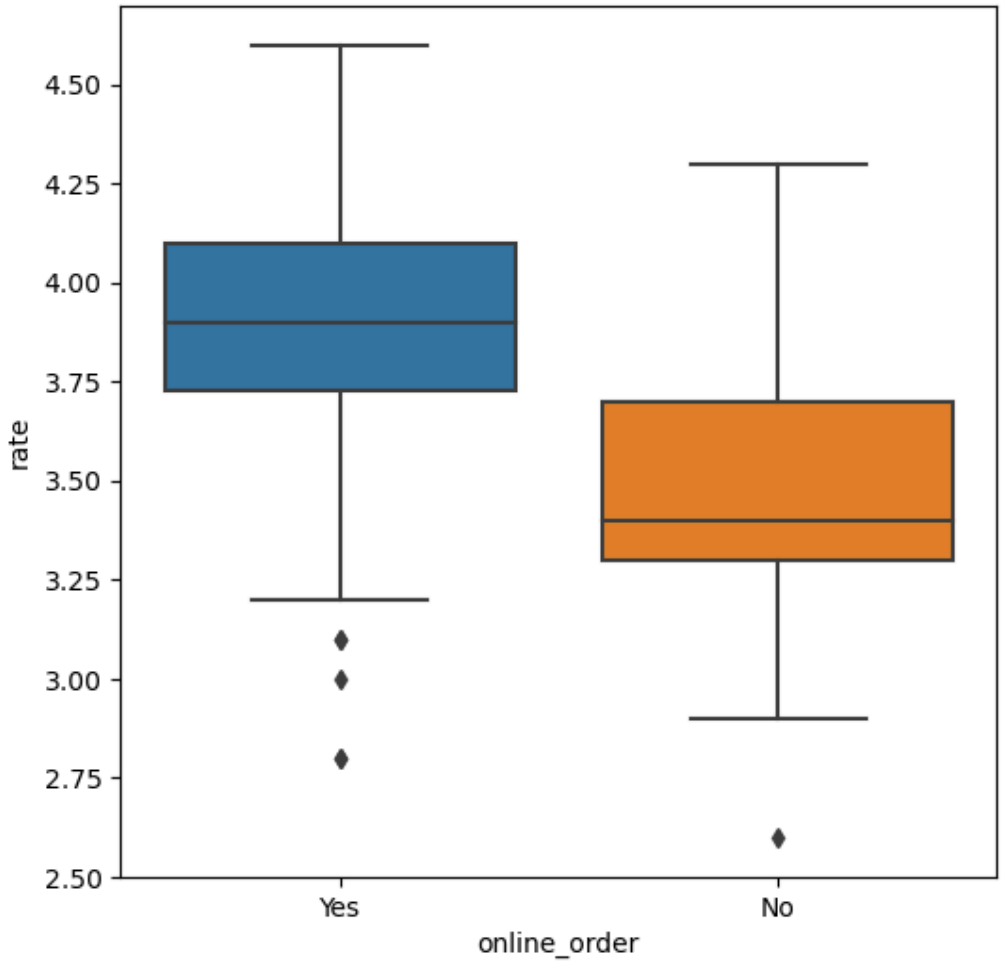
In [23]:
```
couple_data = dataframe['approx_cost(for two people)']
sns.countplot(x=couple_data)
plt.title('Avg order spending size by couples')
plt.show()
```



Conclusion : The majority of the couples likely spend an average amount of 300 rs

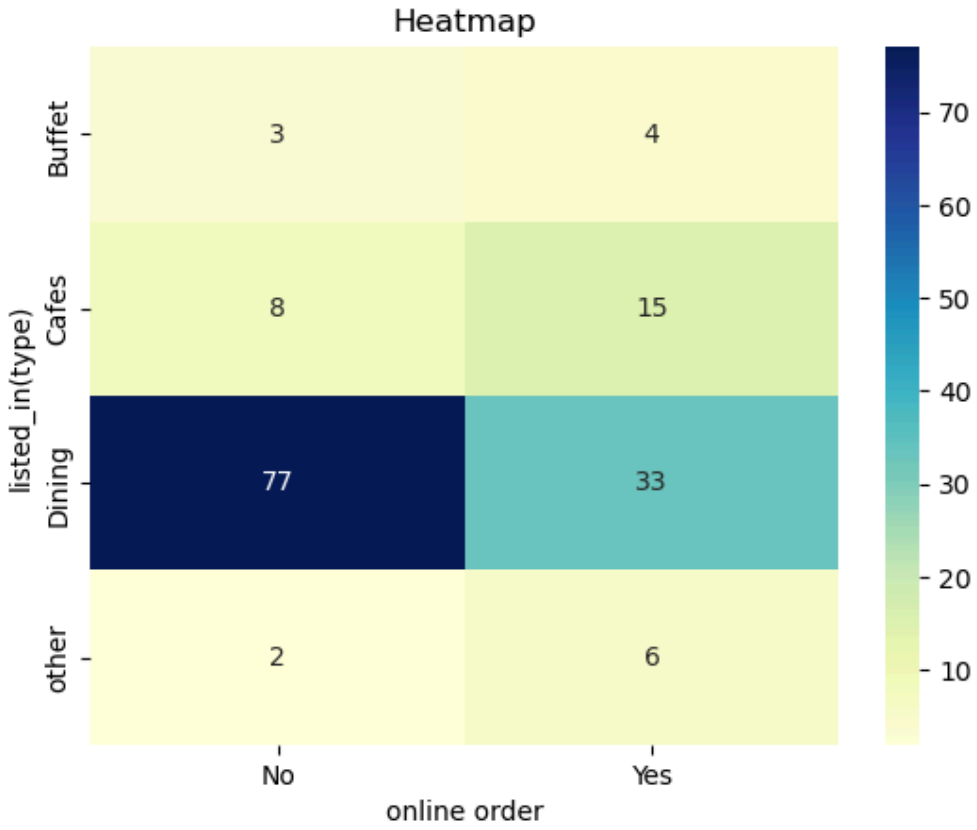
Which Mode receives maximum ratings

In [29]:
```
plt.figure(figsize =(6,6))
sns.boxplot (x= 'online_order' , y ='rate', data = dataframe)
plt.show()
```



Conclusion : offline order receives lower ratings incomparison to online order

In [35]:
```
pivot_table = dataframe.pivot_table(index='listed_in(type)' , columns='onli
sns.heatmap(pivot_table , annot = True , cmap='YlGnBu' , fmt='d')
plt.title('Heatmap')
plt.xlabel('online order')
plt.ylabel('listed_in(type)')
plt.show()
```



Conslusion : Dining Restaurants primarily accept offine orders , whereas cafes primarily receive online orders . This suggests that customers prefers to order in person at restaurants , but prefer online ordering at cafes .

In [ ]:
```

```