Zomatoproject

February 27, 2025

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

[2]:
```
dataframe = pd.read_csv('Zomato data .csv')
```

[3]:
```
dataframe.head()
```

[3]:
```
                  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
```

1 Convert the data type of column - Rate

[4]:
```
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
```

[5]:
```
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
```

Q1) #what type of resturant do the majority of customer order from?

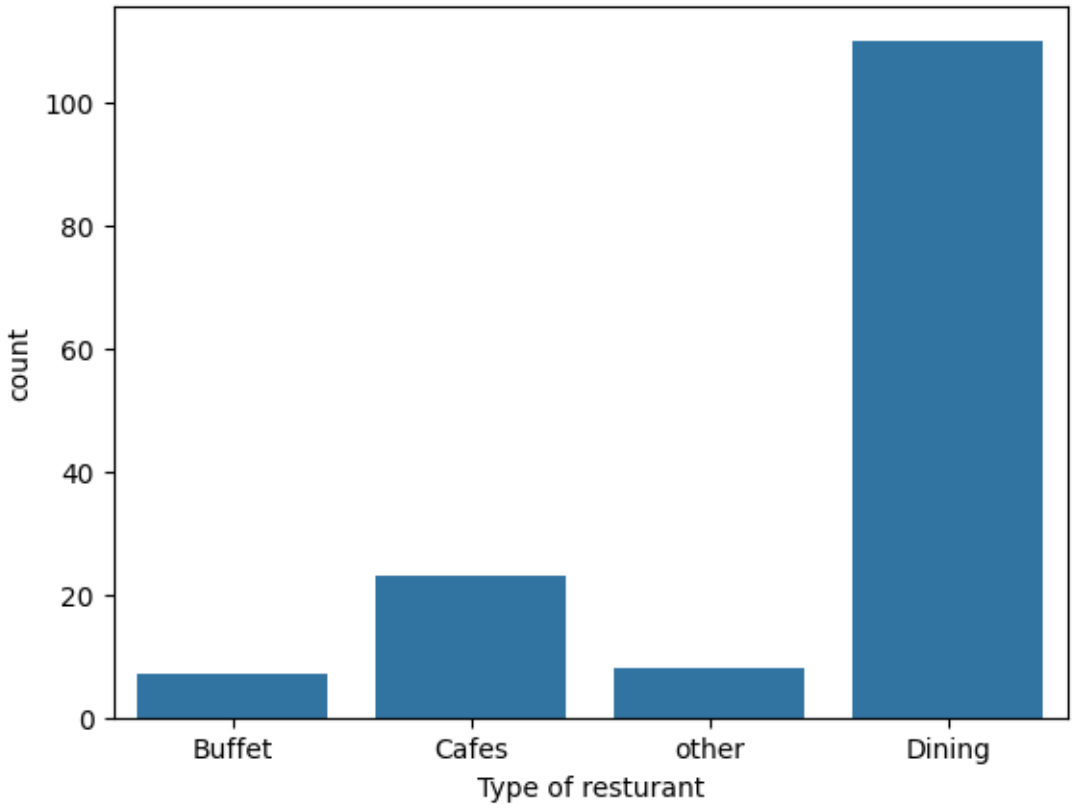
[6]:
```
dataframe.head()
```

[6]:
```
                  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
```

[7]:
```
sns.countplot(x = dataframe['listed_in(type)'])
plt.xlabel("Type of resturant")
```

[7]:
```
Text(0.5, 0, 'Type of resturant')
```



2 conclusion- Majority of the customer order in dinning resturant

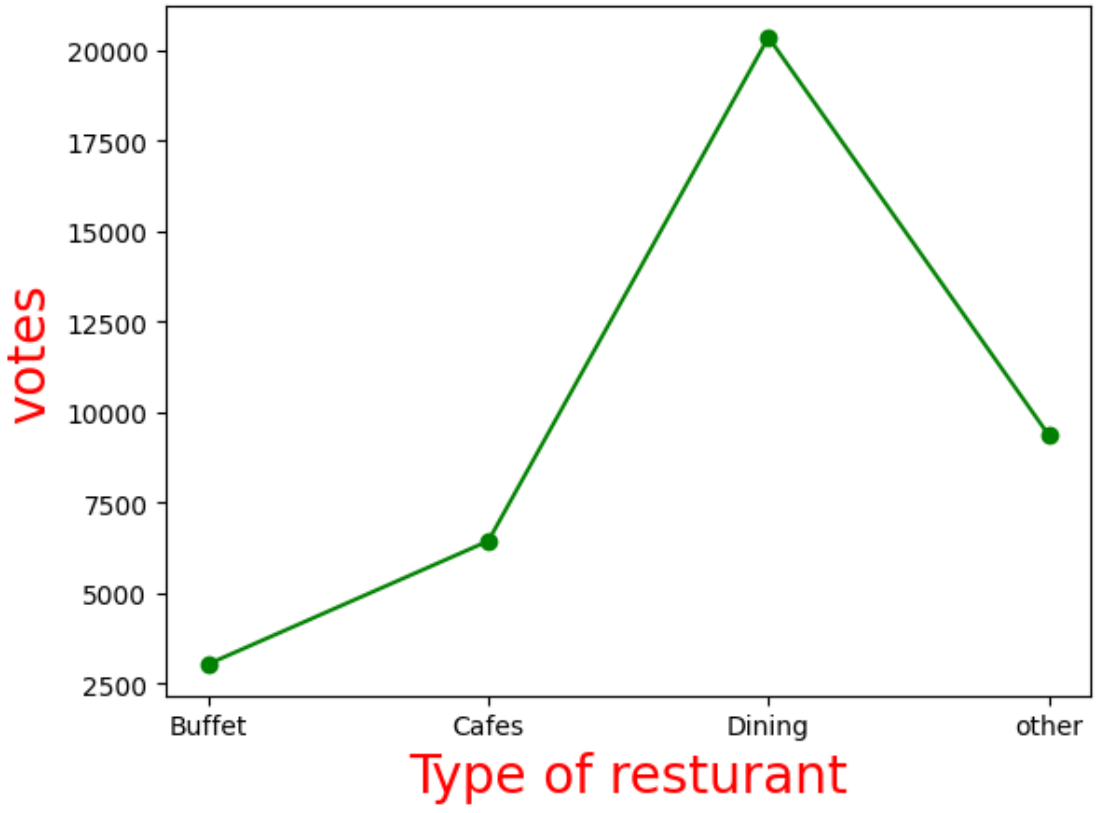
Q2) # How many votes has each type of resturant recived from customer?

[8]:
```
dataframe.head()
```

[8]:
```
                   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
```

[9]:
```
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 resturant", c="red", size=20)
plt.ylabel("votes", c="red", size=20)
plt.show()
```



3 Conclusion - Dinning resturant has recieved maxium votes

Q3) #What are the rating that the majority of resturamt have recieved?

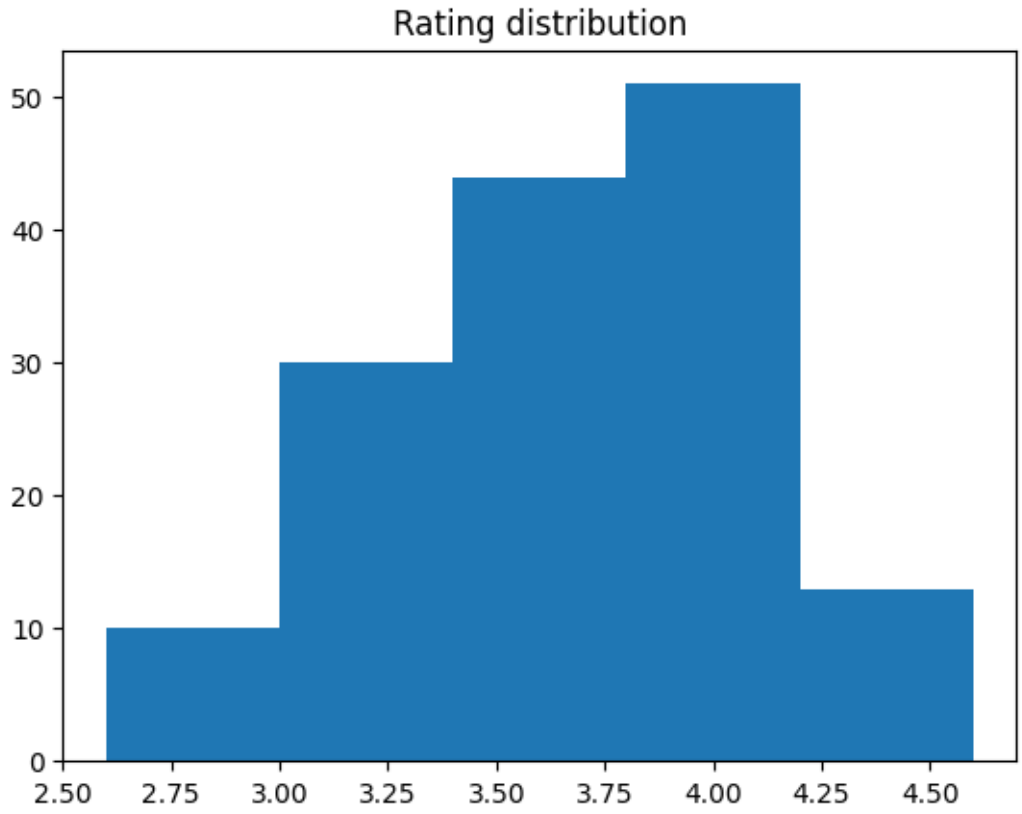
[10]:
```
dataframe.head()
```

[10]:

	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]:
```
plt.hist(dataframe['rate'],bins=5)
plt.title("Rating distribution")
plt.show()
```



4 Conclusion - The majority of resturant recieved rating from 3.5 to 4

Q4) # Zomato has observed that most couples prder most of their food online. What is their average spending on each oreder?

5 Average order spending by couples

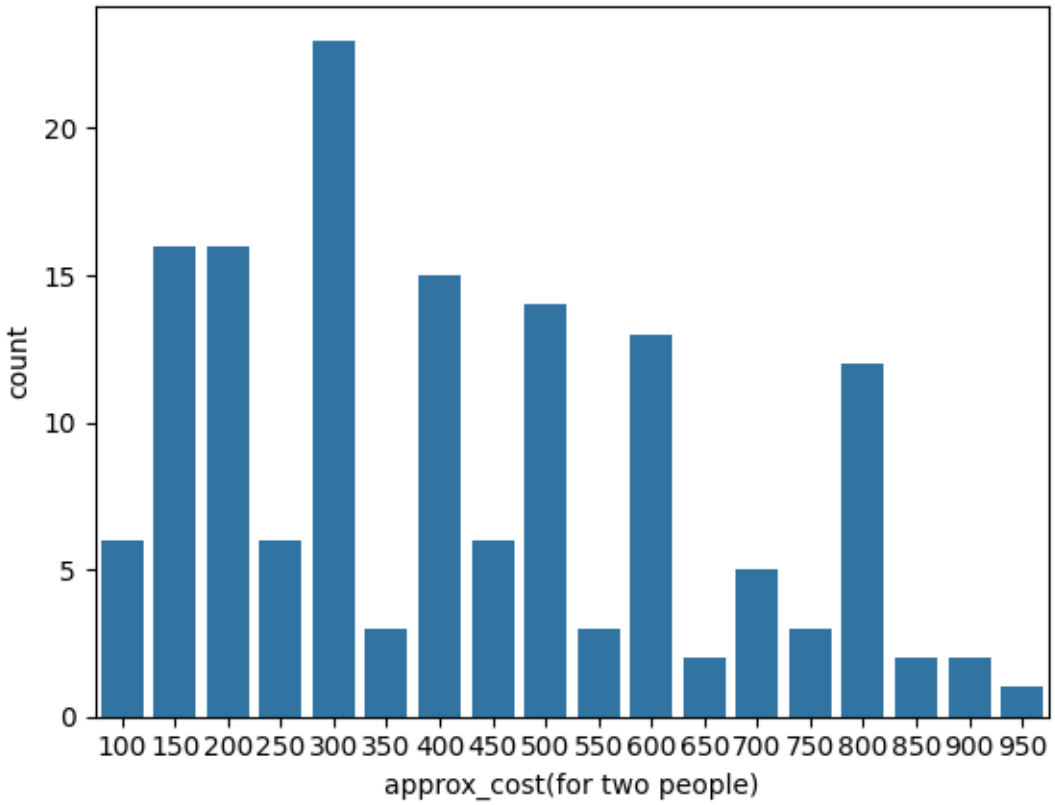
[12]:
```
dataframe.head()
```

[12]:
```
                   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
```

[13]:
```
couple_data=dataframe['approx_cost(for two people)']
sns.countplot(x=couple_data)
```

[13]:
```
<Axes: xlabel='approx_cost(for two people)', ylabel='count'>
```



6 Conclusion- the majority of couple prefer resturants with an approximate cost of 300 rupees

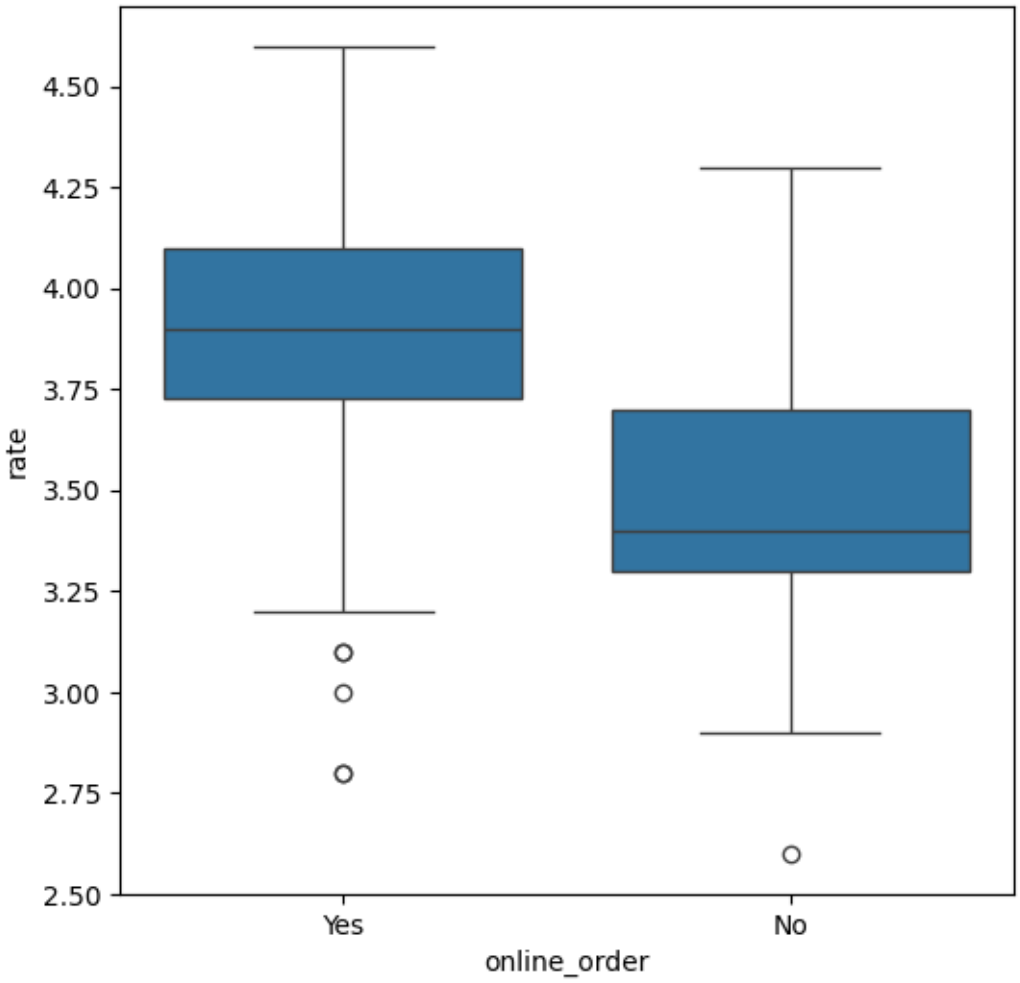
Q5) # Which mode recives maximum rating?

[14]:
```
dataframe.head()
```

[14]:

	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

[15]:
```
plt.figure(figsize = (6,6))
sns.boxplot(x = 'online_order', y = 'rate', data = dataframe)
plt.show()
```



7 Conclusion - offline order received lower rating in comparison to online order

[ ]:

[ ]:

[ ]:

[ ]:

[ ]: