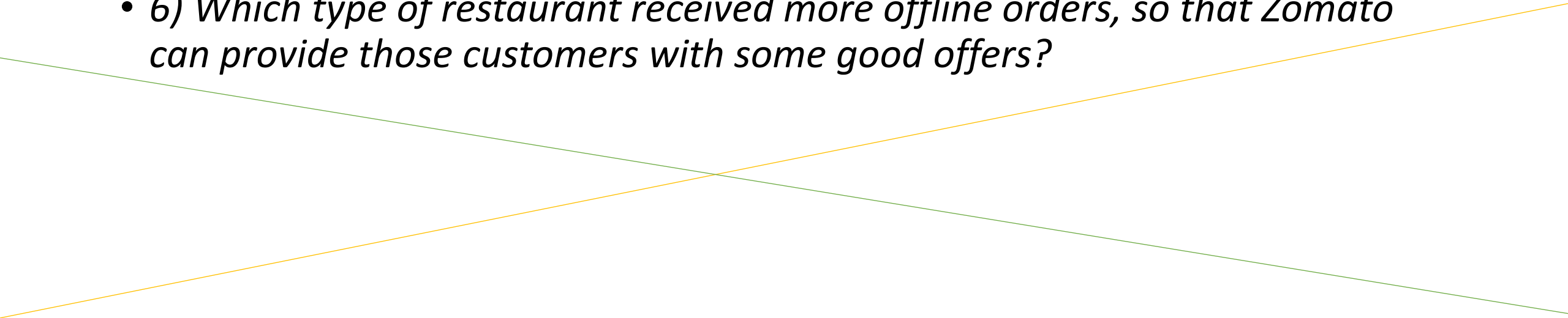


Data Analysis Project Using Python

Zomato has an average of 17.5 million monthly transacting customers for its food delivery business. The average monthly active food delivery restaurant partners on Zomato's platform have also increased by 8.7% year-on-year, from 208,000 to 226,000. You are working in a data-driven role at Zomato. You have a dataset of customers. As a data professional, you need to analyze the data, perform EDA (Exploratory Data Analysis) and visualization, and answer the following questions

Data Analysis Project Using Python

- *1) What type of restaurant do the majority of customers order from?*
 - *2) How many votes has each type of restaurant received from customers?*
 - *3) What are the ratings that the majority of restaurants have received?*
 - *4) Zomato has observed that most couples order most of their food online. What is their average spending on each order?*
 - *5) Which mode (online or offline) has received the maximum rating?*
 - *6) Which type of restaurant received more offline orders, so that Zomato can provide those customers with some good offers?*
- 
- Two decorative lines, one green and one yellow, cross each other diagonally across the bottom half of the slide.

Data Analysis Project Using Python

- Import Necessary Python Libraries

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

- Create the data frame

```
[3]: data = pd.read_csv("Zomato data .csv")
#print(data)
df = pd.DataFrame(data)
print(df)
```

	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			

[148 rows x 7 columns]

Data Analysis Project Using Python

DATA CLEANING

- Convert the data type of the “rate ”column to float and remove the denominator.

DATA CLEANING

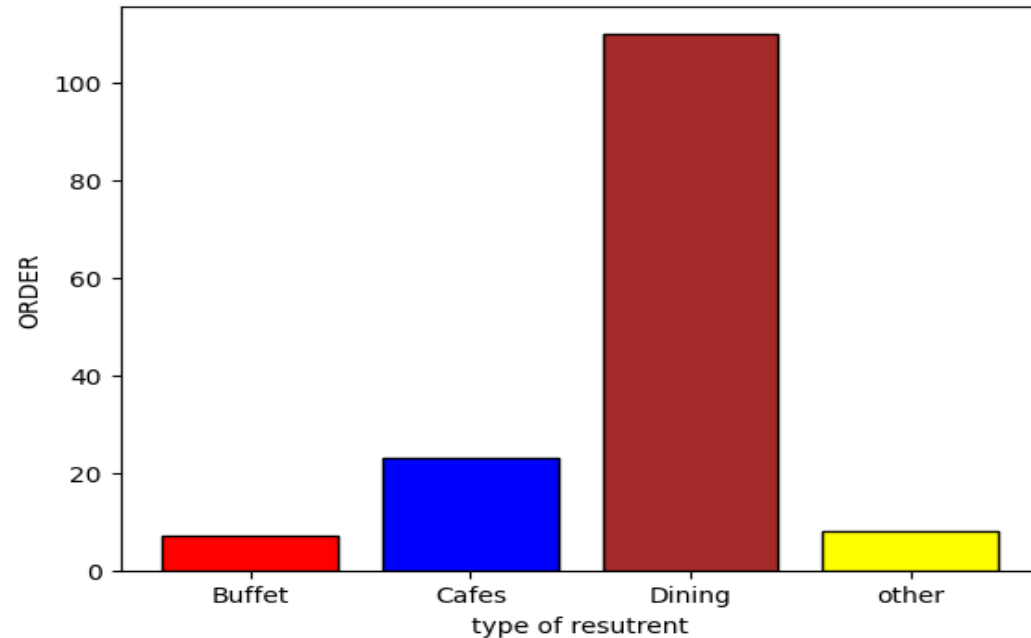
```
[5]: def handleRate(value):  
      value =str(value).split("/")  
      value =value[0];  
      return float(value)  
      df["rate"]=df["rate"].apply(handleRate)  
      print(df)
```

	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	
..	
143	Melting Melodies	No	No	3.3	0	
144	New Indraprasta	No	No	3.3	0	
145	Anna Kuteera	Yes	No	4.0	771	
146	Darbar	No	No	3.0	98	
147	Vijayalakshmi	Yes	No	3.9	47	

1) What type of restaurant do the majority of customers order from?

```
[5]: gp1= df.groupby("listed_in(type)")["listed_in(type)"].count()
print(gp1)
plt.bar(gp1.index, gp1.values, color = ["red", "blue", "brown", "yellow"], edgecolor="black")
x=gp1.index
y= gp1.values
plt.xlabel("type of resutrent")
plt.ylabel("ORDER")
plt.show()
```

```
listed_in(type)
Buffet      7
Cafes      23
Dining     110
other       8
Name: listed_in(type), dtype: int64
```



CONCLUSION: Majority of resturant fall into Dining category.

2)How many votes has each type of restaurant received from customers?

Problem (2). How many votes has each type of restaurant received from customers?

```
[11]: gp2=df.groupby("listed_in(type)")["votes"].sum()
      print(gp2)
      c = ["red","blue","brown","green"]
      plt.plot(gp2.index,gp2.values,color="green",marker="D")
      plt.xlabel("Type of Resturant",color="red",size=20)
      plt.ylabel("Votes",color="red",size=20)
      plt.show()
```

```
listed_in(type)
Buffet      3028
Cafes       6434
Dining     20363
other       9367
Name: votes, dtype: int64
```

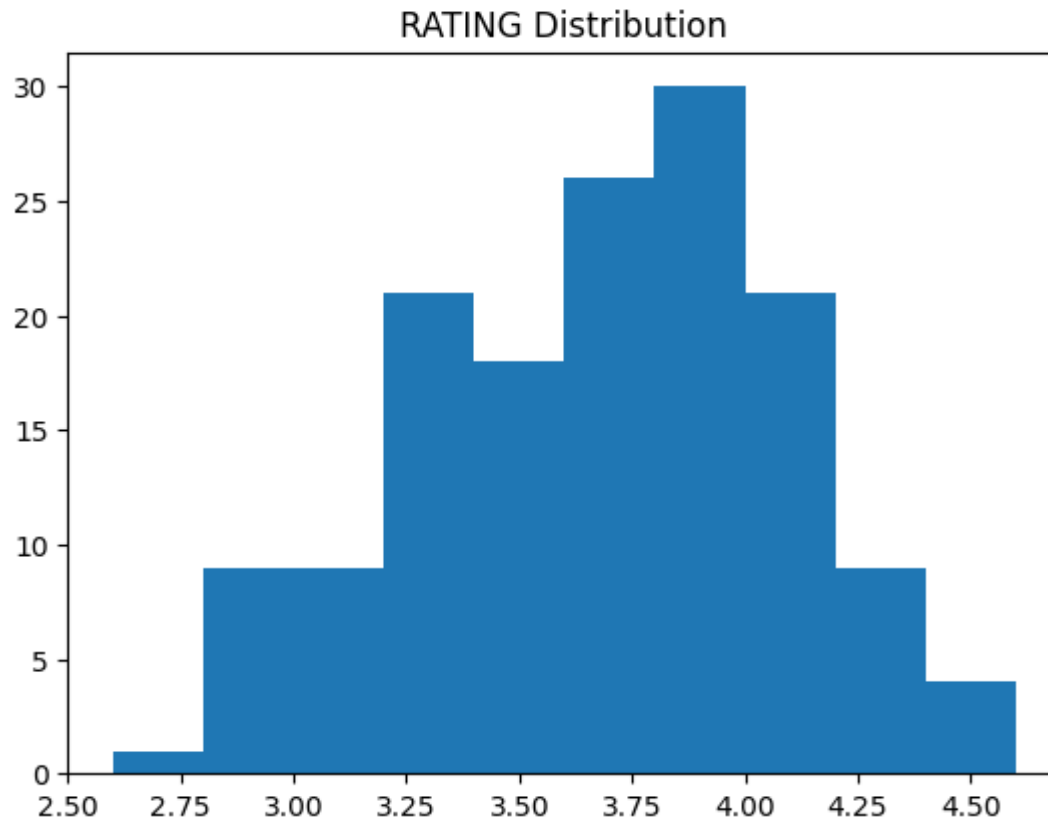


CONCLUSION - dinning restaurant has recieved maximum votes.

CONCLUSION: Dinning restaurant has recieved maximum votes.

3) What are the ratings that the majority of restaurants have received?

```
[7]: gp3=df.groupby(["listed_in(type)","rate"]).agg({"rate":"count"})  
plt.hist(df["rate"],bins=10)  
plt.title("RATING Distribution")  
plt.show()
```



CONCLUSION - The majority resturants received rating from 3.5 to 4

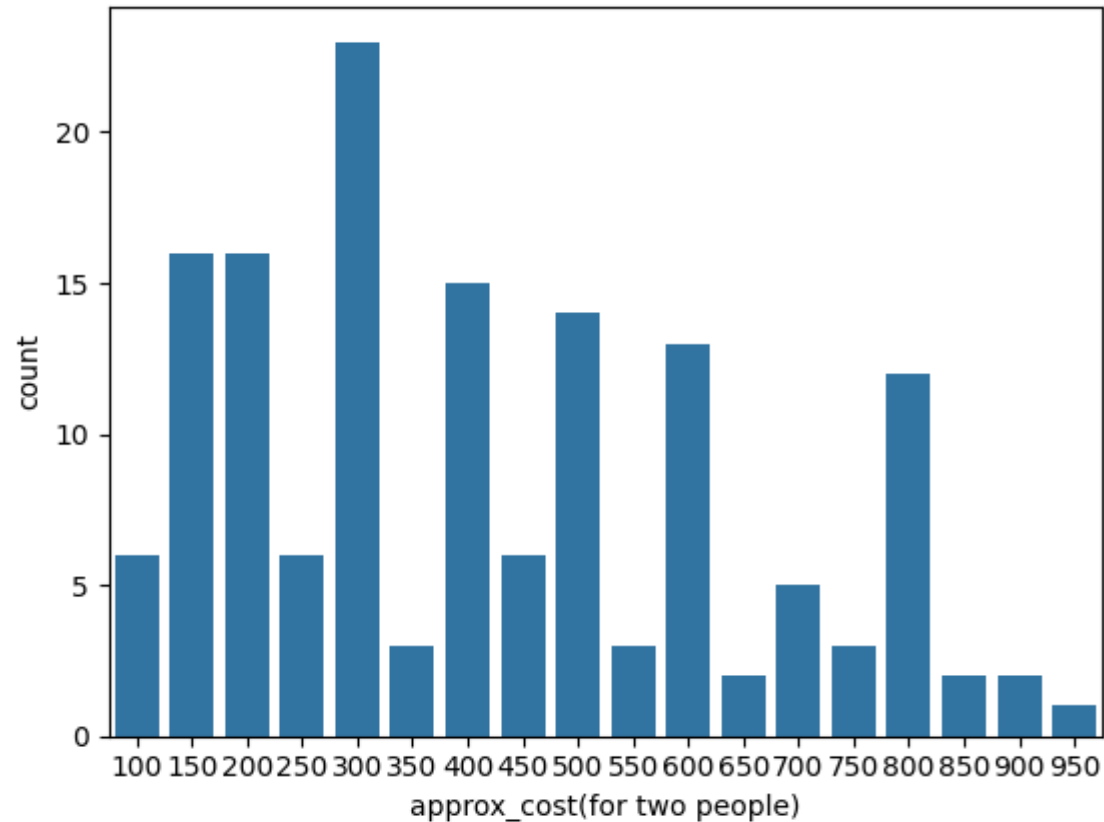


CONCLUSION: The majority resturants received rating from 3.5 to 4.

4) Zomato has observed that most couples order most of their food online. What is their average spending on each order?

```
[17]: couple_data = df["approx_cost(for two people)"]  
sns.countplot(x=couple_data)
```

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



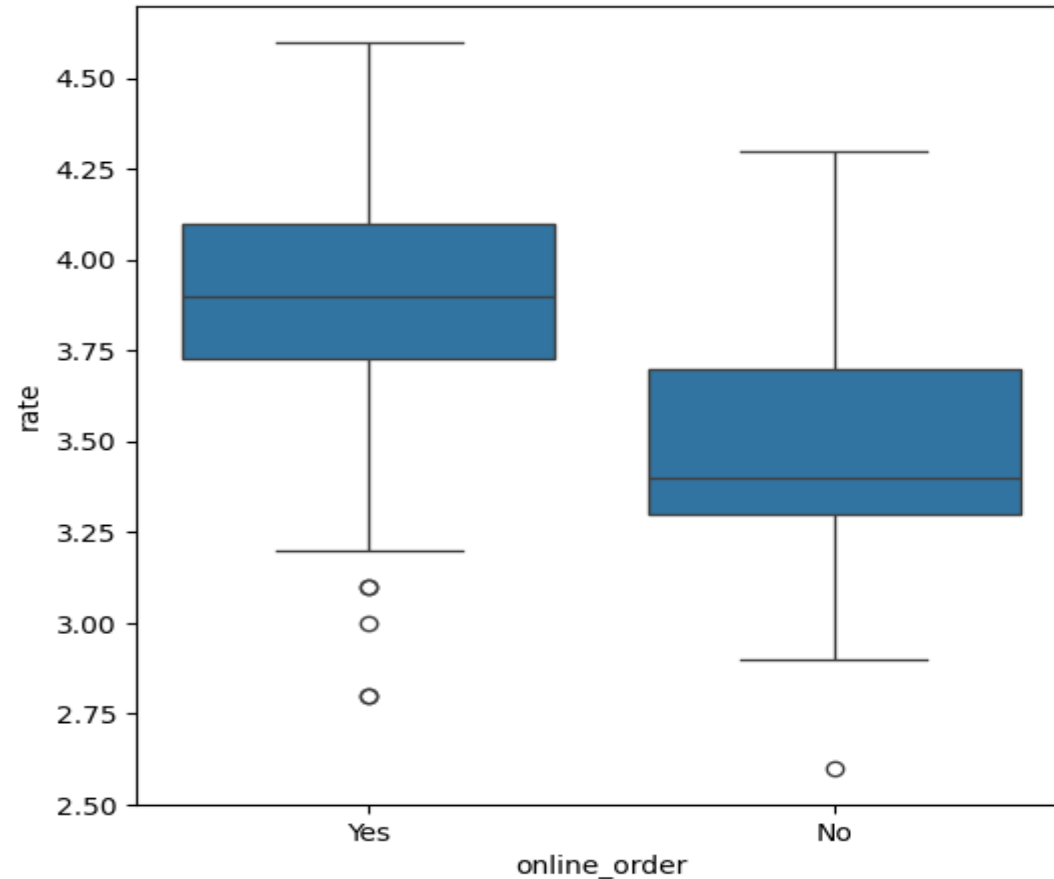
Conclusion - The majority of couple prefer restaurants with an approximate cost of 300 rupees.

CONCLUSION: The majority of couple prefer restaurants with an approximate cost of 300 rupees.

5) Which mode (online or offline) has received the maximum rating?

```
[18]: plt.figure(figsize = (6,6))  
      sns.boxplot(x = "online_order", y = "rate" , data =df)
```

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

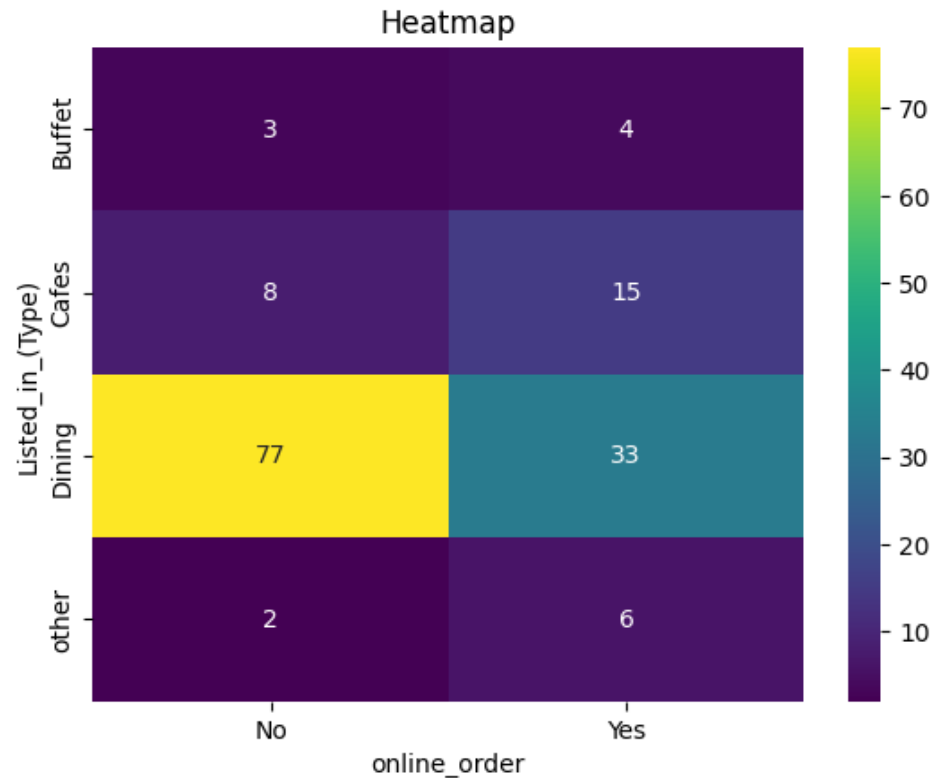


CONCLUSION - Offline orders recieved lower rating in comparision to online orders, which obtain excellent ratings.

CONCLUSION: Offline orders recieved lower rating in comparision to online orders, which obtain excellent ratings.

6) Which type of restaurant received more offline orders, so that Zomato can provide those customers with some good offers?

```
[31]: pivot_table = df.pivot_table(index = "listed_in(type)", columns = "online_order", aggfunc = "size" , fill_value = 0)
sns.heatmap(pivot_table, annot = True , cmap = "viridis",fmt = 'd')
plt.title("Heatmap")
plt.xlabel("online_order")
plt.ylabel("Listed_in_(Type)")
plt.show()
```



CONCLUSION - Dining restaurant primarily accept offline orders. This suggest that client prefer to place order in person at restaurant ,but prefer online ordering at cafes.

CONCLUSION: Dining restaurant primarily accept offline orders. This suggest that client prefer to place order in person at restaurant ,but prefer online ordering at cafes.