ZOMATO SALES ANALYSIS

IMPORTING LIBRARIES

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

data = pd.read_csv('Zomato.csv')

data.head(5)
```

```
                     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
```

```
data.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
```

DATA CLEANING

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

data['rate'] = data['rate'].apply(rating)
print(df.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
```

```
data.head(3)
```

```
              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

   approx_cost(for two people) listed_in(type)
0                           800          Buffet
1                           800          Buffet
2                           800          Buffet
```

```
data.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
```

```
print(data.isnull().sum())
```

```
name                           0
online_order                   0
book_table                     0
rate                           0
votes                          0
approx_cost(for two people)    0
listed_in(type)                0
dtype: int64
```

Type of Restraunt

```
data.head()
```

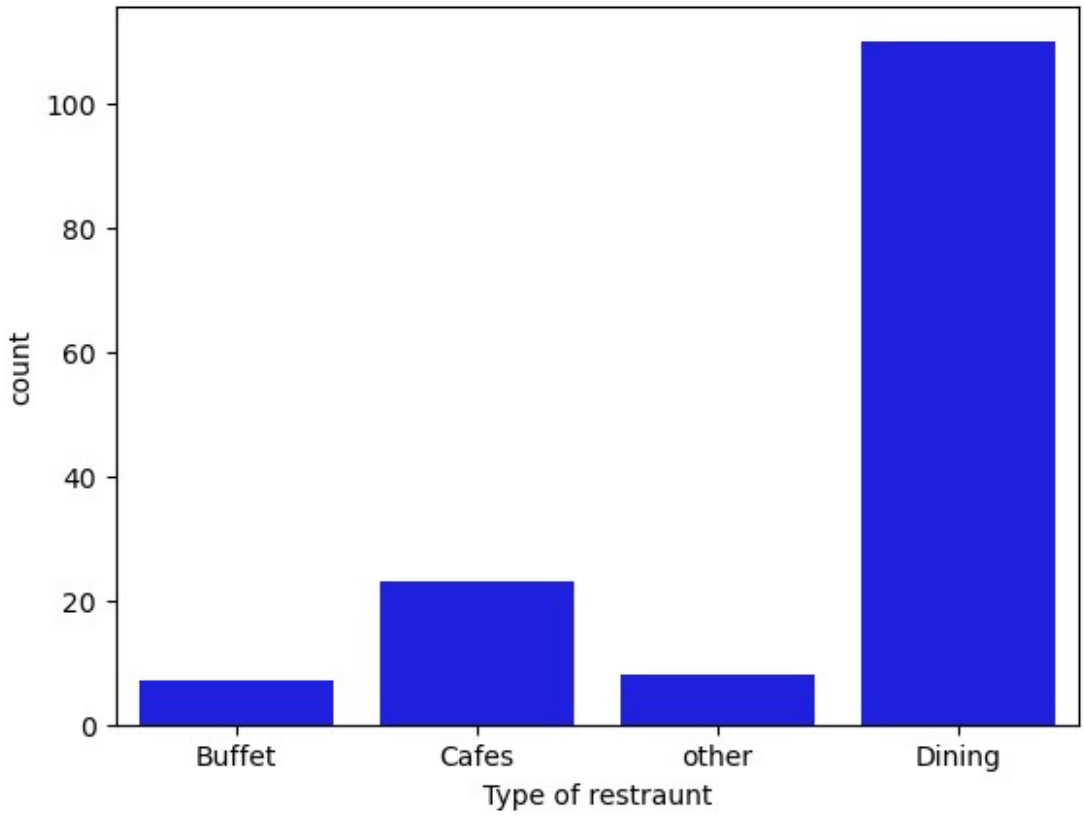
```
                     name online_order book_table  rate  votes  \
0                   Jalsa          Yes        Yes   4.1    775
1          Spice Elephant          Yes         No   4.1    787
2         San Churro Cafe          Yes         No   3.8    918
3  Addhuri Udupi Bhojana           No    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

What types of restraint do the majority of coustomer order from?

```
sns.countplot(x = data['listed_in(type)'] , color = 'b')
plt.xlabel('Type of restraunt')
```

```
Text(0.5, 0, 'Type of restraunt')
```



CONCLUSION - So here we can see that maxmium people likes to eat in a dining

Average spending on each order?

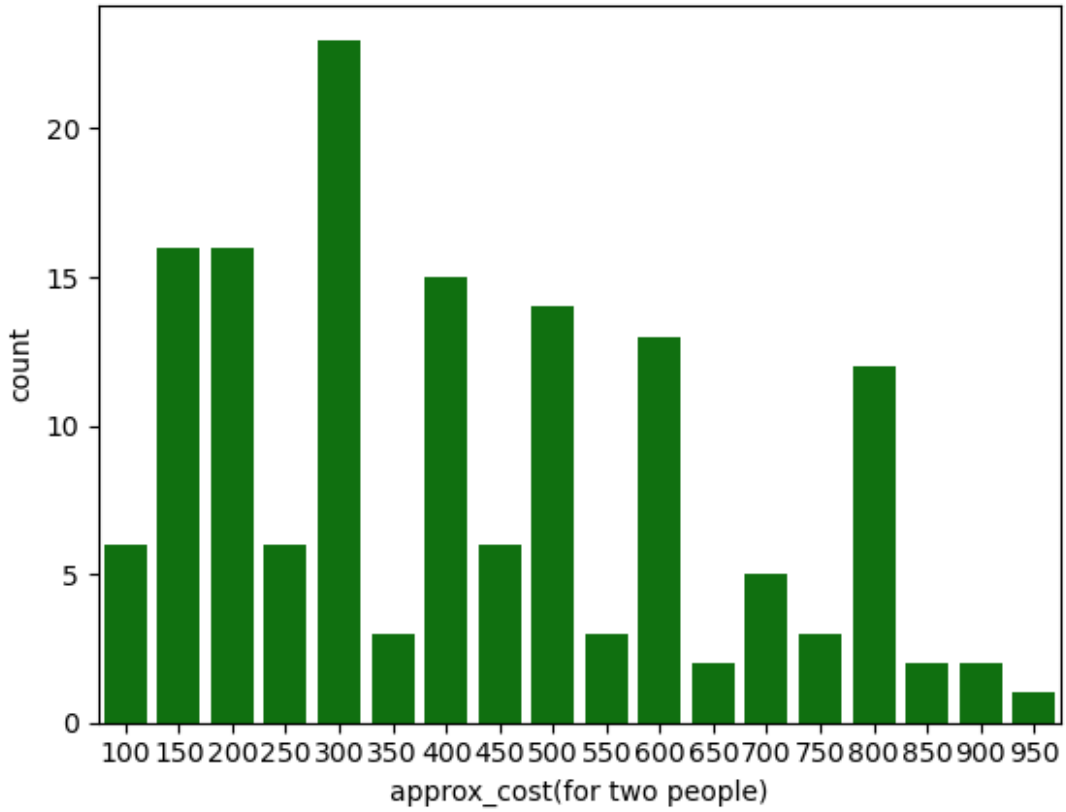
```
data.head()
```

	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

Average spending on each order?

```
df = data['approx_cost(for two people)']
sns.countplot(x = df , color = 'g')
```

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

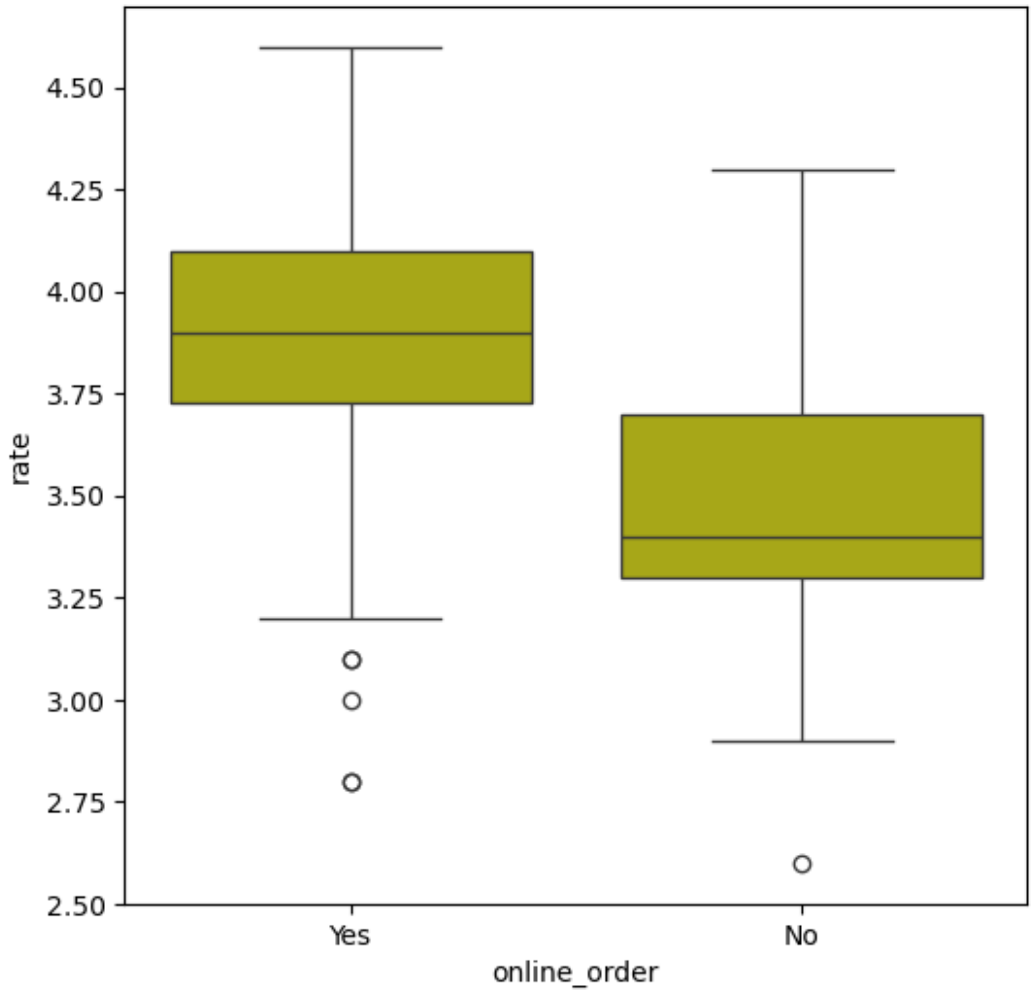


CONCLUSION - The majority of people preffer restraunt with an cost of rupees 300

Which mode offline or online has ecieved the max rating?

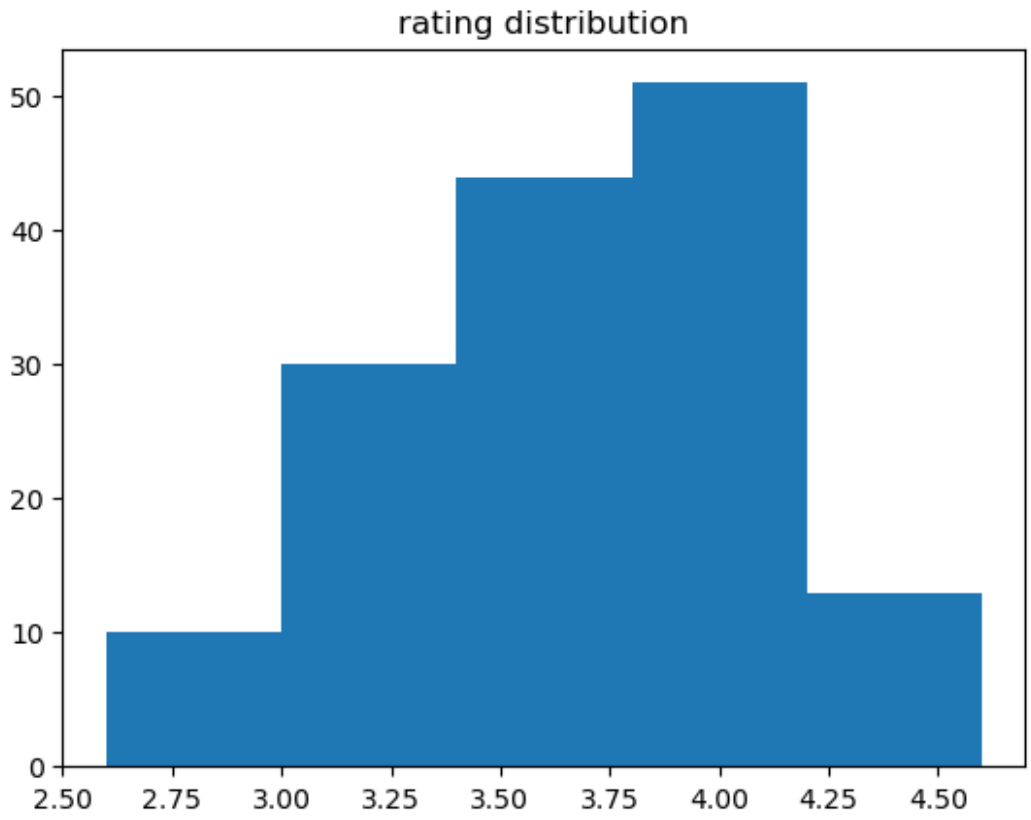
```
plt.figure(figsize = (6,6))
sns.boxplot(x = 'online_order' , y = 'rate' , data = data , color =
'y')
```

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



Rating of majority of restraunt have received ??

```
plt.hist(data['rate'],bins = 5)
plt.title("rating distribution")
plt.show()
```



CONCLUSION - The majority restraunt recieved rating from 3.5 to 4

Most ordered food online , average spending on each order??

average order spending

```
data.head()
```

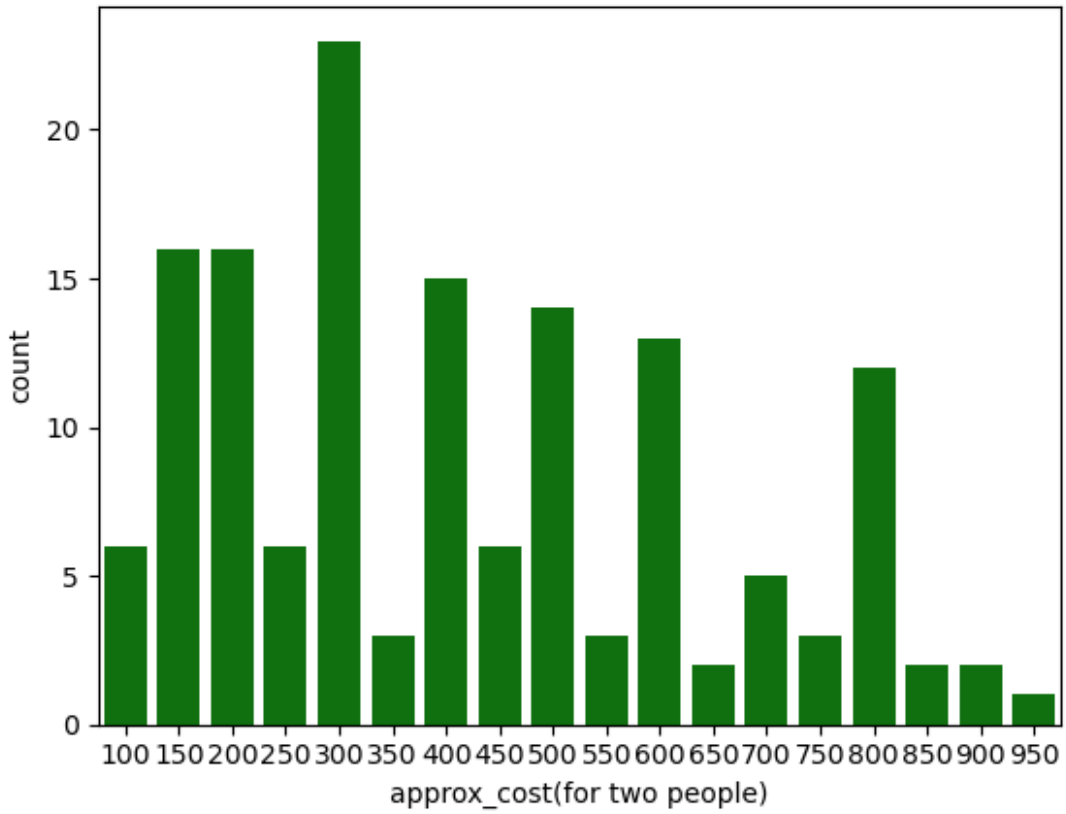
	name	online_order	book_table	rate	votes	\
0	Jalsa	Yes	Yes	4.1	775	
1	Spice Elephant	Yes	No	4.1	787	
2	San Churro Cafe	Yes	No	3.8	918	
3	Addhuri Udupi Bhojana	No	No	3.7	88	
4	Grand Village	No	No	3.8	166	

	approx_cost(for two people)	listed_in(type)
0	800	Buffet
1	800	Buffet
2	800	Buffet

```
3                            300          Buffet
4                            600          Buffet
```

```
most_ordered = data['approx_cost(for two people)']
sns.countplot (x = most_ordered , color = '#008000')
```

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



CONCLUSION - the majority of couples prefer restraunt with the approximate cost of 300 rupees

Which mode offline or online has recieved the max rating

```
data.head()
```

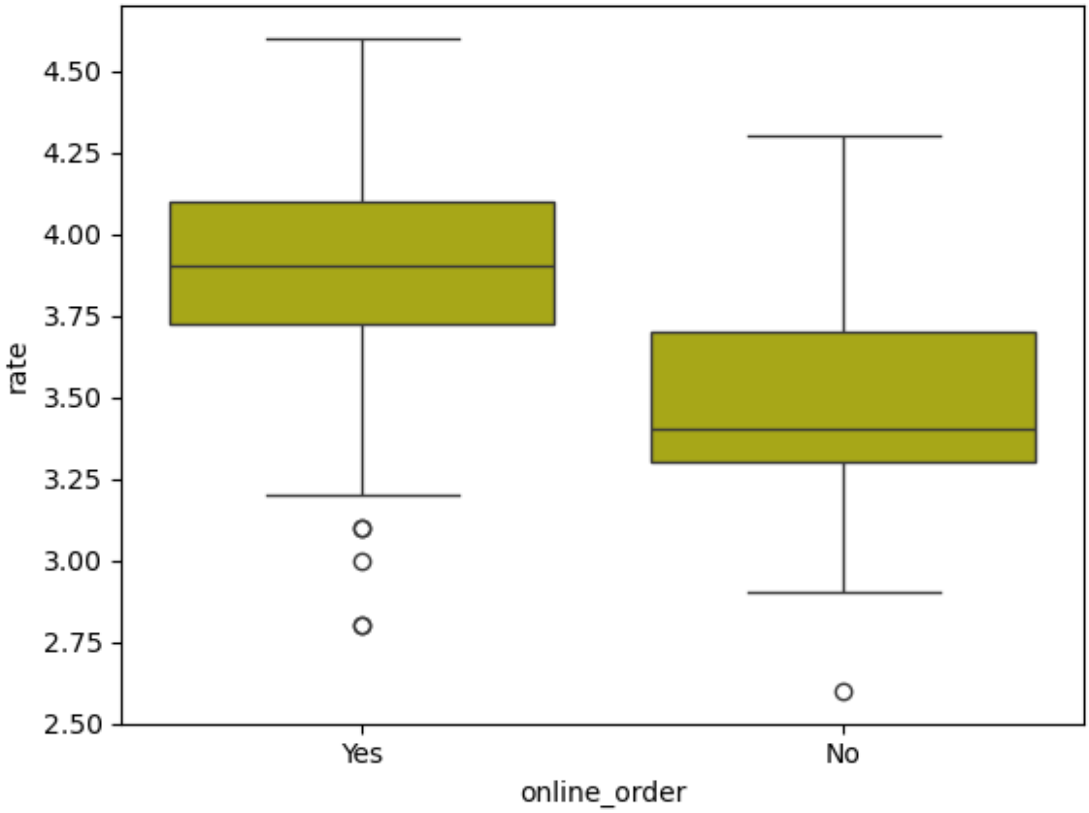
```
                    name online_order book_table  rate  votes  \
0                  Jalsa          Yes        Yes   4.1    775
1         Spice Elephant          Yes         No   4.1    787
2        San Churro Cafe          Yes         No   3.8    918
3  Addhuri Udupi Bhojana           No        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

```
plt.figure(figure = (6,6))
sns.boxplot(x = 'online_order' , y = 'rate' , data = data , color =
'y')
```

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



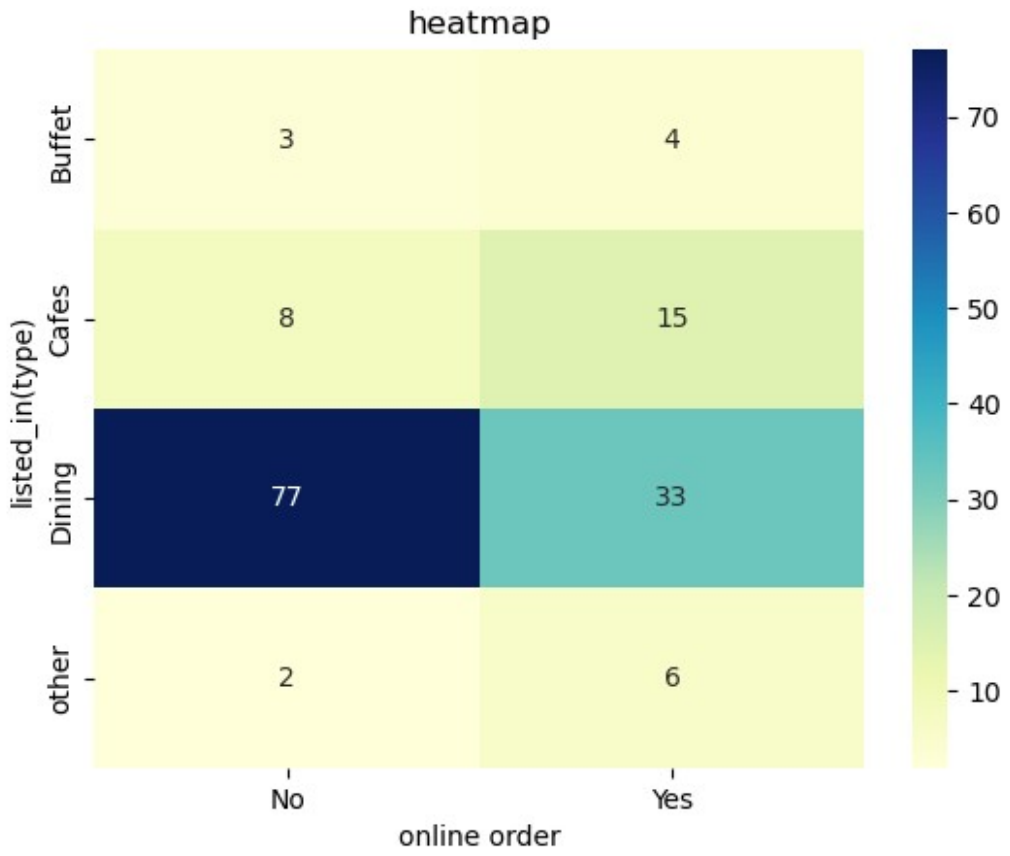
CONCLUSION - offline order recieved lower rating in comparision to online order

which type of restraunt received more offline order so that Zomato can give coustomer with some good offer

```
pivot_table = data.pivot_table(index = 'listed_in(type)' , columns =
'online_order' , aggfunc = 'size' , fill_value=0)
sns.heatmap(pivot_table , annot = True , cmap = 'YlGnBu', fmt = 'd')
plt.title('heatmap')
plt.xlabel('online order')
plt.yorder('Listed In (Type)')
plt.show()
```

```
---------------------------------------------------------------------------
-----
AttributeError                            Traceback (most recent call
last)
Cell In[223], line 5
      3 plt.title('heatmap')
      4 plt.xlabel('online order')
----> 5 plt.yorder('Listed In (Type)')
      6 plt.show()

AttributeError: module 'matplotlib.pyplot' has no attribute 'yorder'
```



CONCLUSION - Dining restraunt primarily accept offline orders , whereas cafes primarily recieve online order. this suggests that client prefer order in person at restraunt , but prefer online ordering at cafe.