Importing Libraries

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

Create the data frame

```
dataframe = pd.read_csv(r"C:\Users\sulta\OneDrive\Desktop\Zomato\
Zomato data .csv")
print(dataframe)
```

```
                        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]
```

```
dataframe
```

```
                        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]
```

Convert the data type of column - rate

```
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
```

```
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
```

Question-1 What type of restaurent do the majority of customers order from?

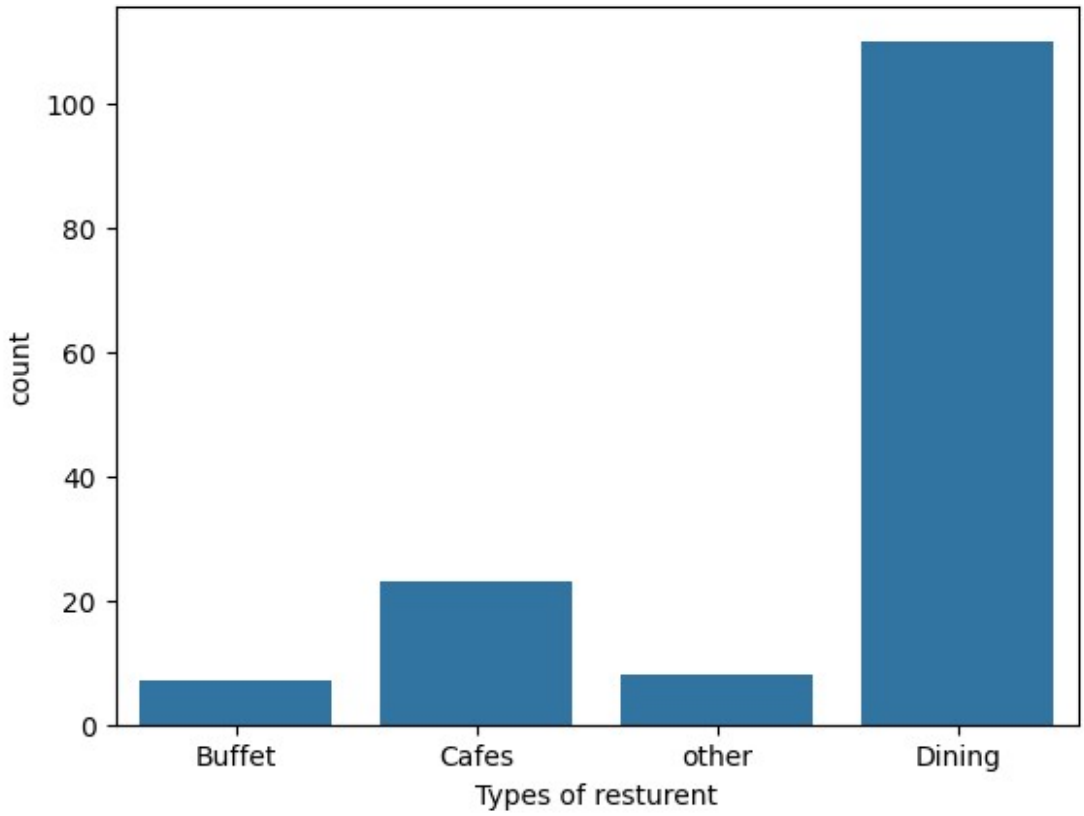
```
dataframe.head()
```

```
                  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  
```

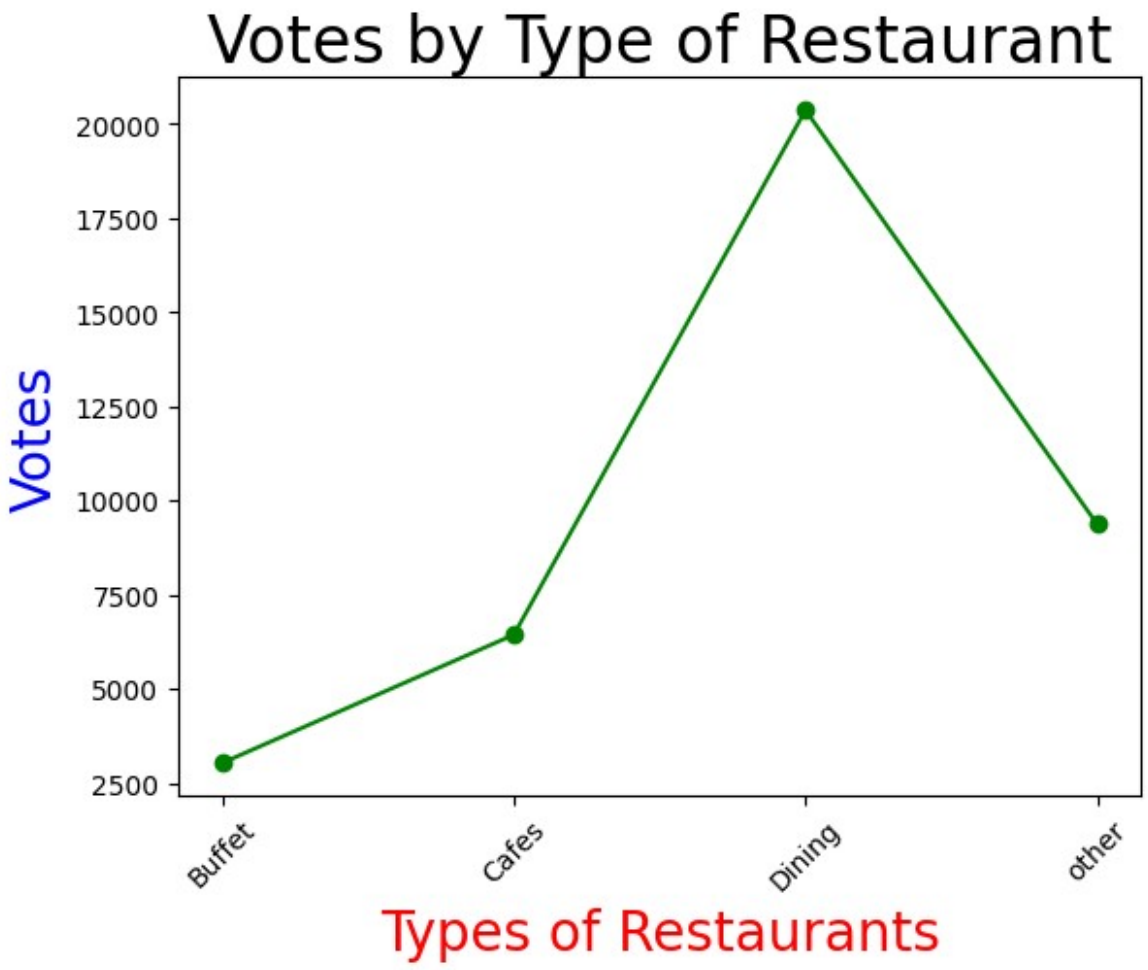
```
sns.countplot(x=dataframe['listed_in(type)'])
#plt.hist(cleaned_data['rate_cleaned'], bins=5, edgecolor='black')
plt.xlabel("Types of resturent")
```

```
Text(0.5, 0, 'Types of resturent')
```



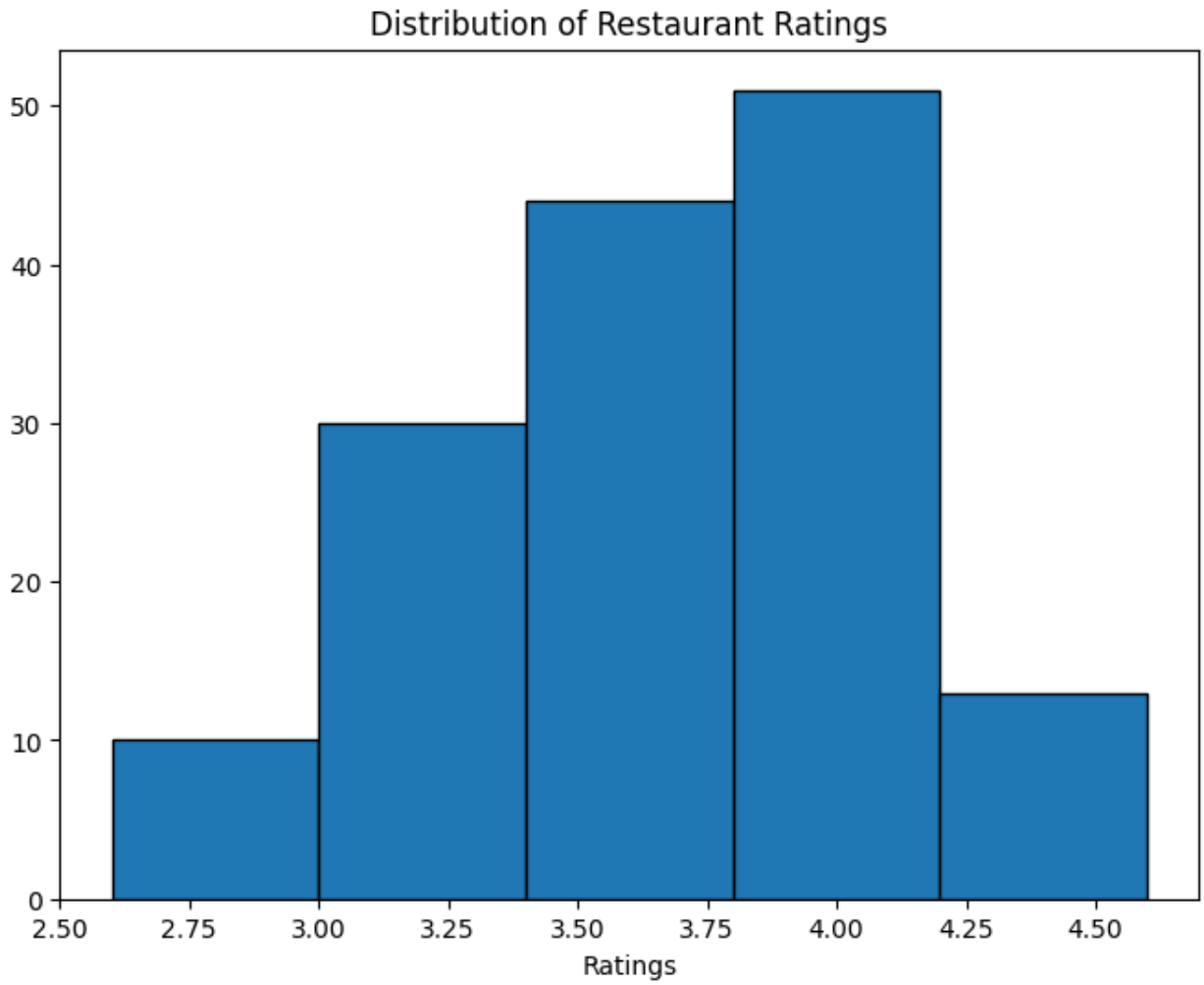
Question-2 How many votes has each type of resturent received from customers?

```
grouped_data = dataframe.groupby('listed_in(type)')['votes'].sum()
result = pd.DataFrame({'votes': grouped_data}).reset_index()
plt.plot(result['listed_in(type)'], result['votes'], c="green",
marker="o")
plt.xlabel("Types of Restaurants", c="red", size=20)
plt.ylabel("Votes", c="blue", size=20)
plt.title("Votes by Type of Restaurant", size=24)
plt.xticks(rotation=45)
plt.show()
```



Question-3 What type of ratings that the majority of restaurants have received?

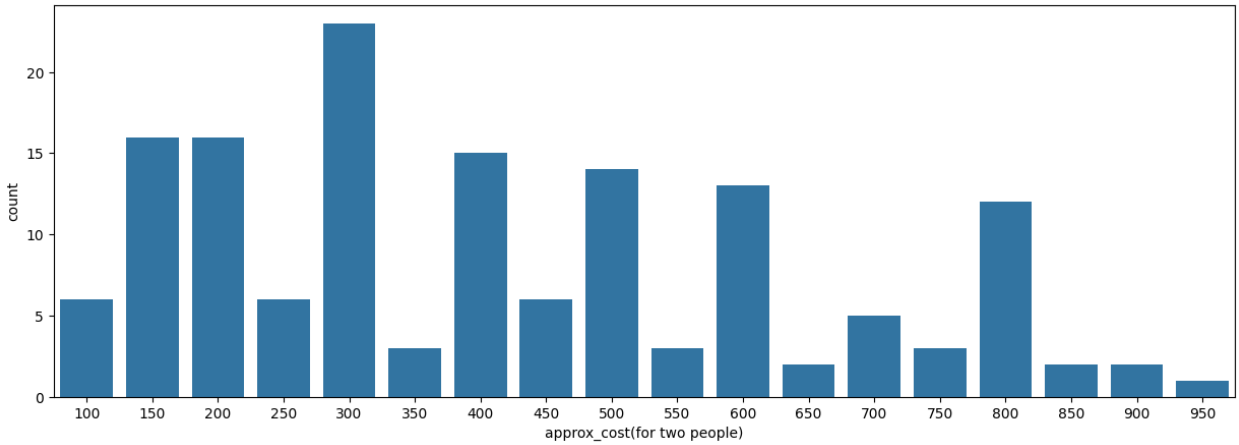
```
dataframe['rate_cleaned'] = dataframe['rate'].str.extract(r'(\d+\.\
d+)').astype(float)
cleaned_data = dataframe.dropna(subset=['rate_cleaned'])
plt.figure(figsize=(8, 6))
plt.hist(cleaned_data['rate_cleaned'], bins=5, edgecolor='black')
plt.title('Distribution of Restaurant Ratings')
plt.xlabel('Ratings')
plt.show()
```



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

```
couple_data=dataframe['approx_cost(for two people)']
plt.figure(figsize=(15,5))
sns.countplot(x=couple_data)
```

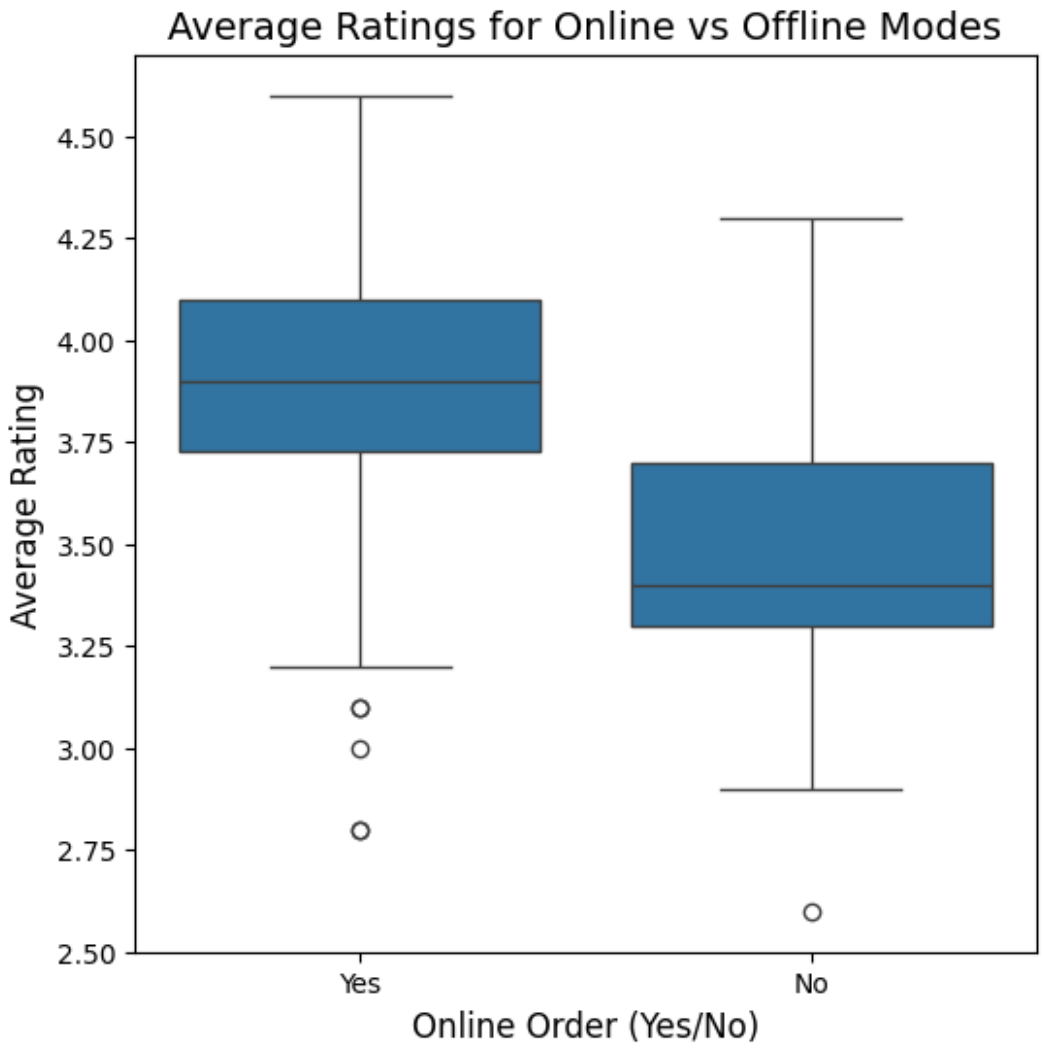
```
<Axes: xlabel='approx_cost(for two people)', ylabel='count'>
```



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

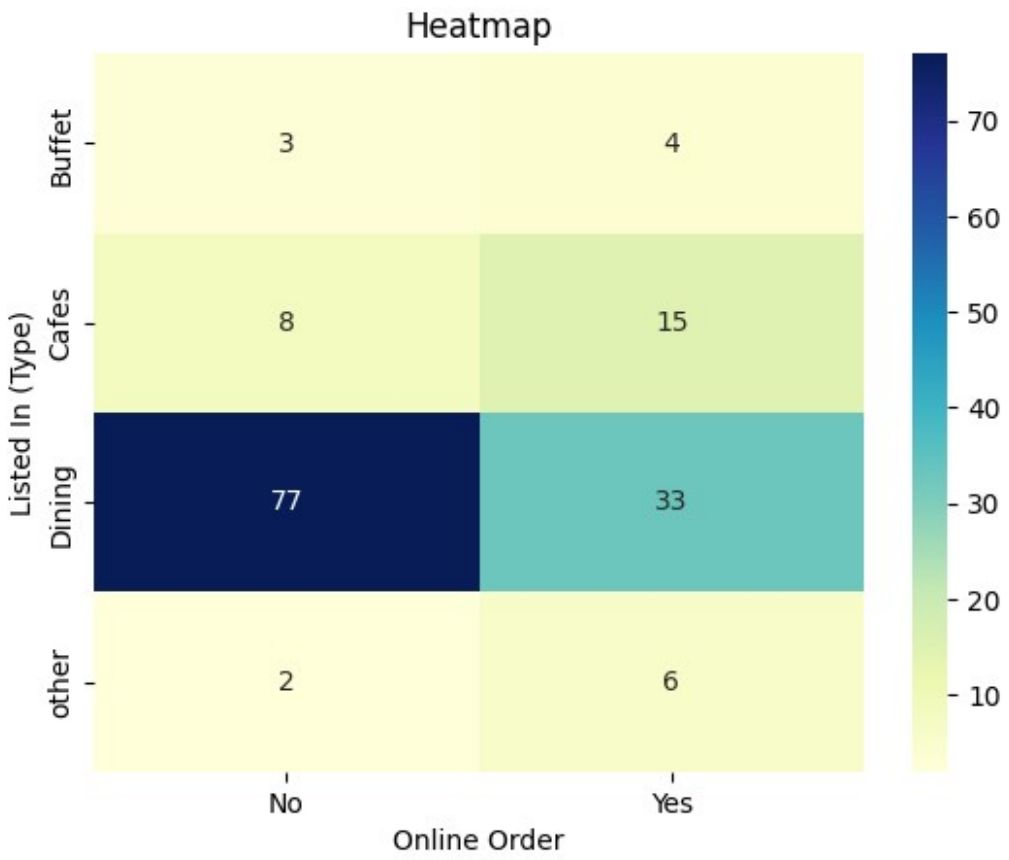
```
plt.figure(figsize = (6,6))
plt.title('Average Ratings for Online vs Offline Modes', fontsize=14)
sns.boxplot(x = 'online_order', y = 'rate', data = dataframe)
plt.xlabel('Online Order (Yes/No)', fontsize=12)
plt.ylabel('Average Rating', fontsize=12)
```

```
Text(0, 0.5, 'Average Rating')
```



Questiom-6 Which type of resturent received more offline orders so that Zomato can customers with some good offers?

```
pivot_table = dataframe.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.ylabel("Listed In (Type)")
plt.show()
```


Heatmap
Buffet
Cafes
Dining
other
Listed In (Type)
3
4
8
15
77
33
2
6
No
Yes
Online Order
70
60
50
40
30
20
10