Zomato Data Analysis Using Python

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

```
dataframe = pd.read_csv("Zomato data .csv")
print(dataframe.head())
```

```
                    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
```

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

dataframe['rate']=dataframe['rate'].apply(handleRate)
print(dataframe.head())
```

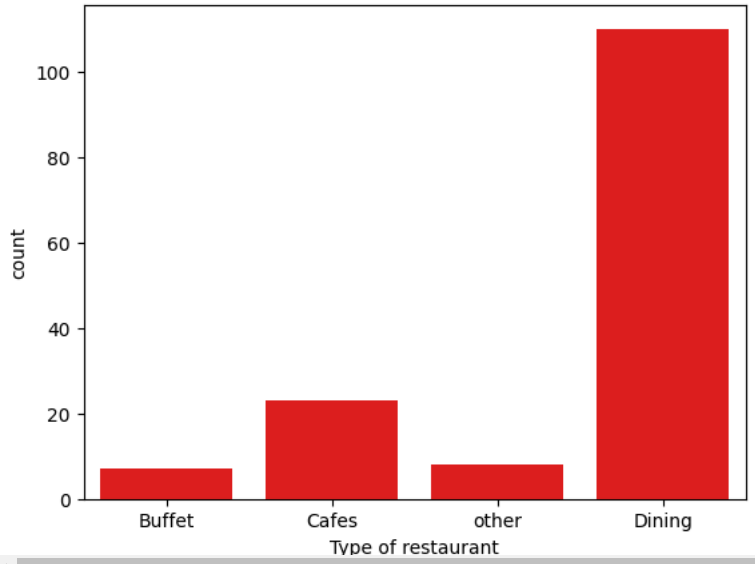
```
                    name online_order book_table  rate  votes  \
0                  Jalsa          Yes        Yes   4.1    775
1         Spice Elephant          Yes         No   4.1    787
2        San Churro Cafe          Yes         No   3.8    918
3  Addhuri Udupi Bhojana           No        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

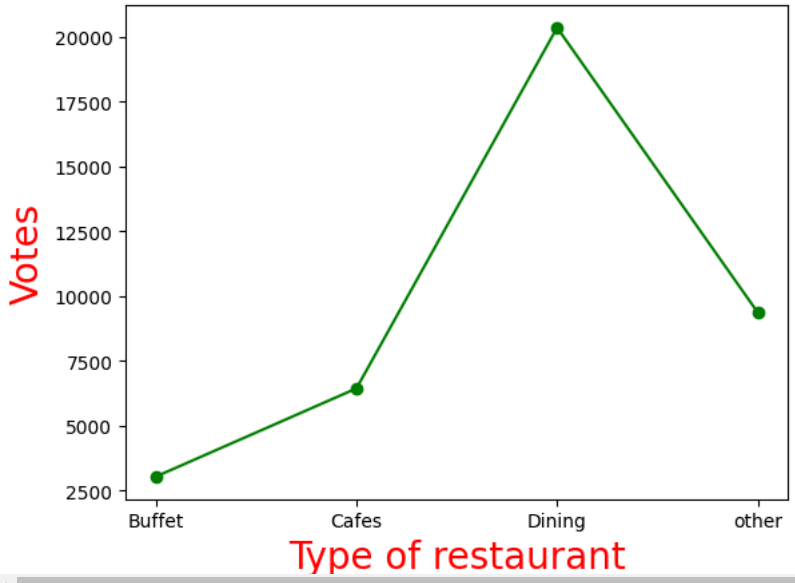
```
sns.countplot(x=dataframe['listed_in(type)'], color='red')
plt.xlabel("Type of restaurant")

```

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




```
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)
```

```
Text(0, 0.5, 'Votes')
```

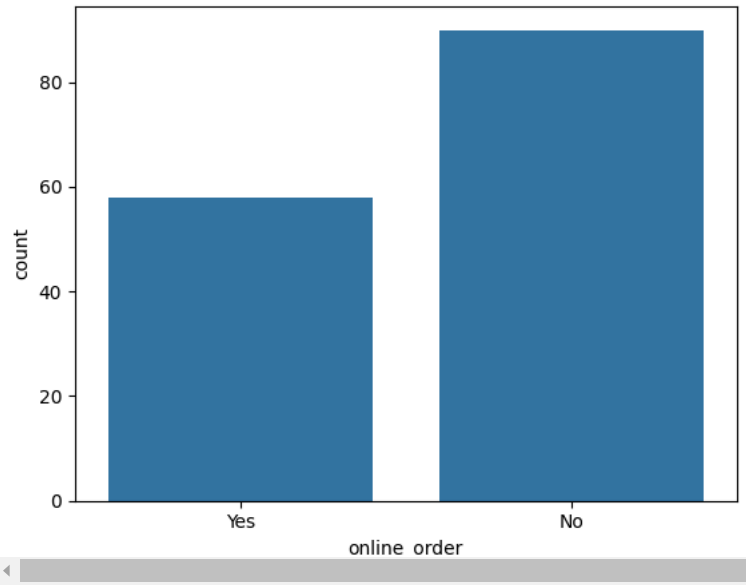



```
max_votes = dataframe['votes'].max()
restaurant_with_max_votes = dataframe.loc[dataframe['votes'] == max_votes, 'name']

print("Restaurant(s) with the maximum votes:")
print(restaurant_with_max_votes)
```

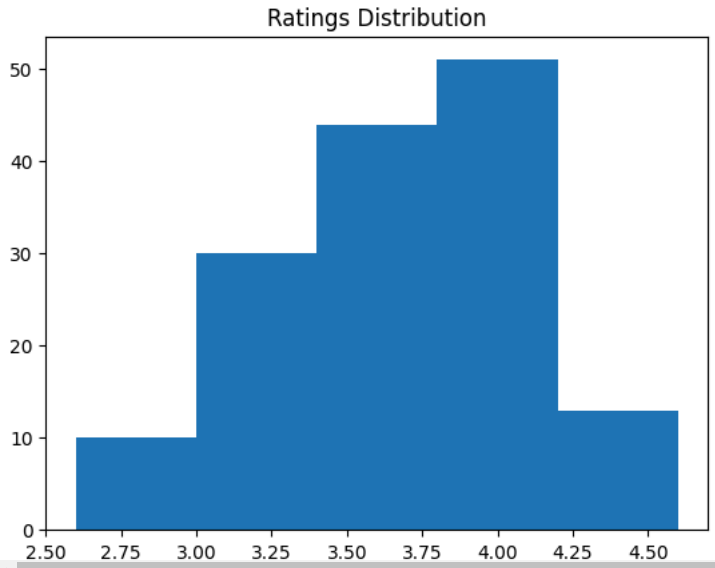
```
Restaurant(s) with the maximum votes:
38    Empire Restaurant
Name: name, dtype: object
```

```
sns.countplot(x=dataframe['online_order'])
```

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




```
1 plt.hist(dataframe['rate'],bins=5)
2 plt.title("Ratings Distribution")
3 plt.show()
4
```

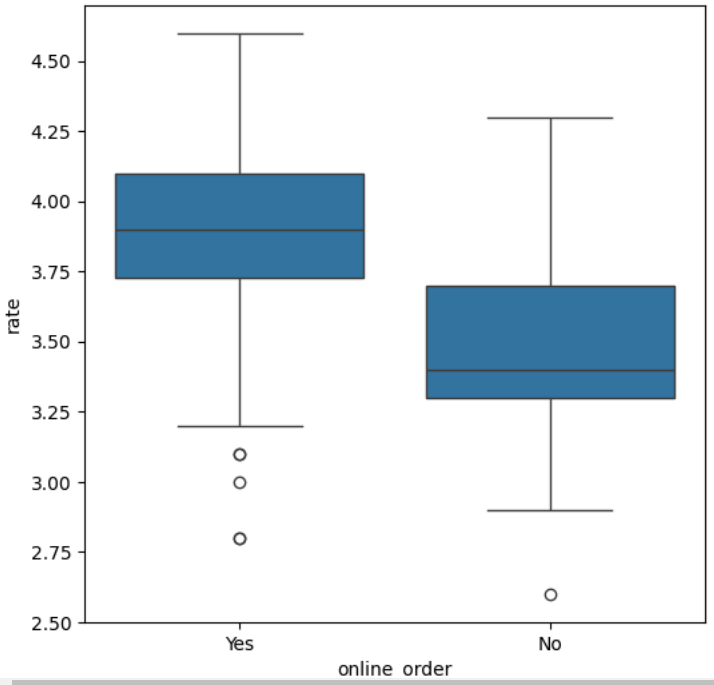



```
1 couple_data=dataframe['approx_cost(for two people)']
2 sns.countplot(x=couple_data)
```

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


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

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




```
1 pivot_table = dataframe.pivot_table(index='listed_in(type)', columns='online_order', aggfunc='size', fill_value=0)
2 sns.heatmap(pivot_table, annot=True, cmap="YlGnBu", fmt='d')
3 plt.title("Heatmap")
4 plt.xlabel("Online Order")
5 plt.ylabel("Listed In (Type)")
6 plt.show()
```