Zomato

February 15, 2026

1 Zomato Data Analysis Project

1.1 Importing Required Libraries

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

1.2 Loading the Data and Creating Dataframe

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

[3]:
```
dataframe
```

[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
..                     …          …        …    …    …
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]
```

1.3 Data Cleaning and Data preprocessing

[4]:

```
# 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
```

[5]:

```
dataframe.info()
```

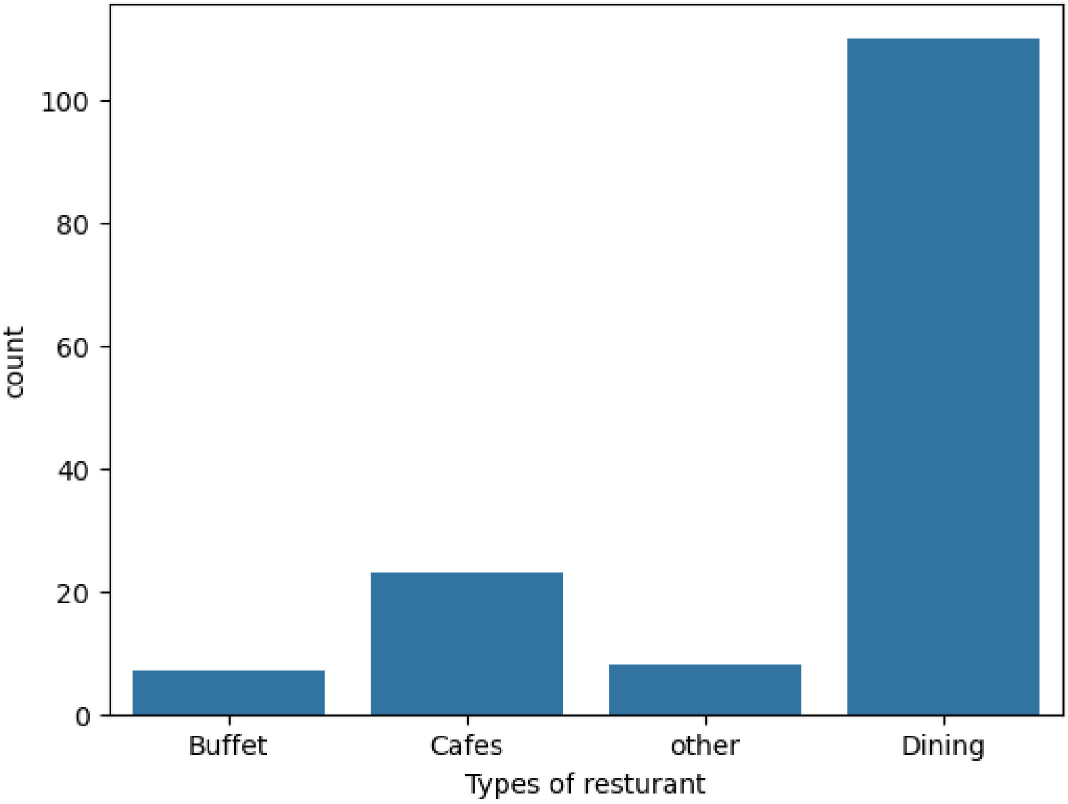
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 148 entries, 0 to 147
Data columns (total 7 columns):
 #   Column                       Non-Null Count  Dtype
---  ------                       --------------  -----
 0   name                         148 non-null    object
 1   online_order                 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
```

```
[6]: # Types of Resturant
     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("Types of resturant")
```

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



[8]:

```
dataframe.head()
```

[8]:

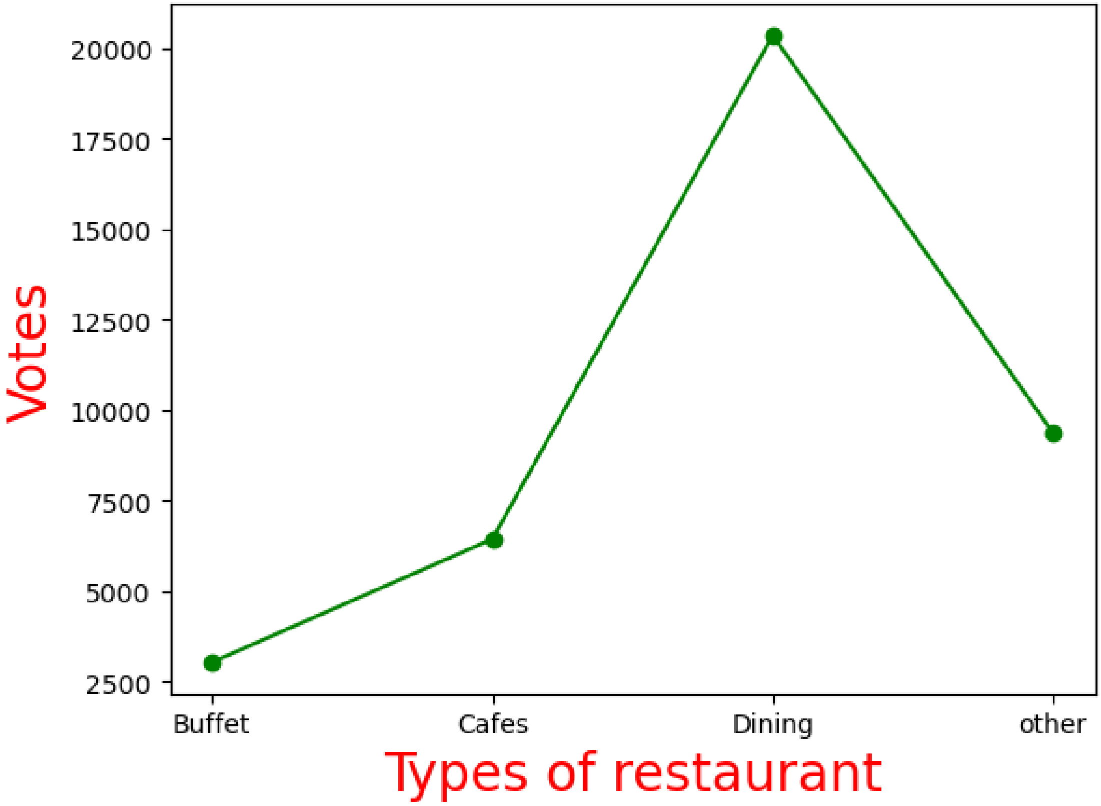
```
                     name online_order book_table  rate  votes  \
0                   Jalsa          Yes        Yes   4.1    775
1          Spice Elephant          Yes         No   4.1    787
2         San Churro Cafe          Yes         No   3.8    918
3   Addhuri Udupi Bhojana           No   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

[10]:

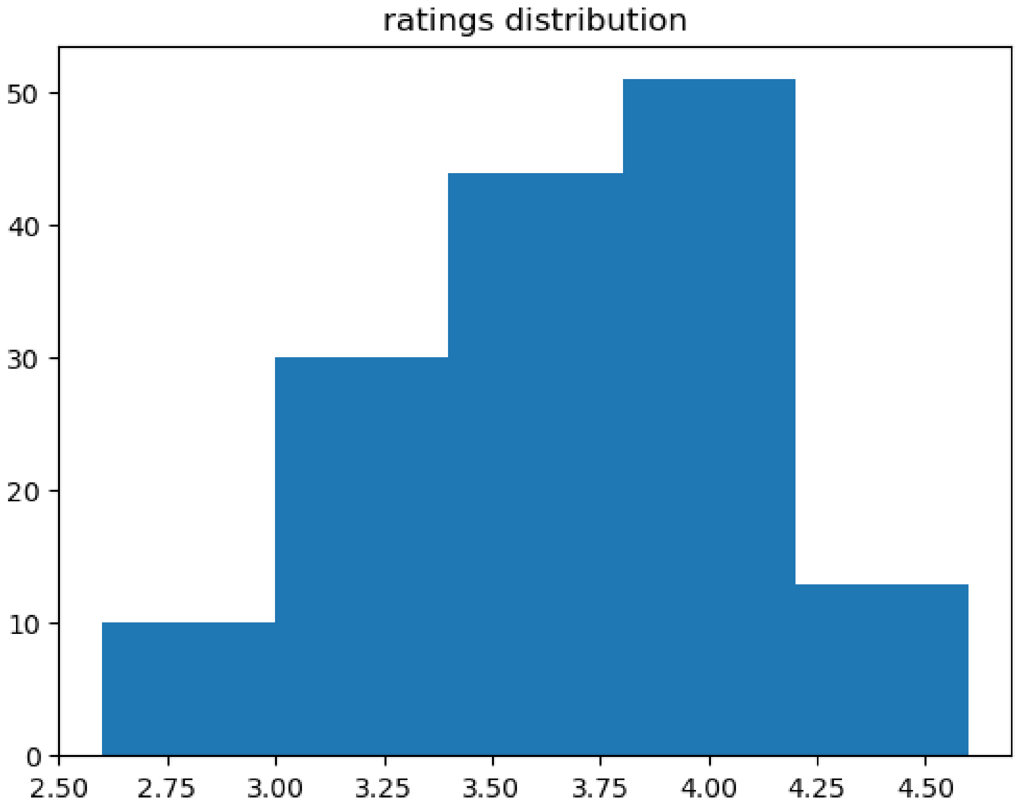
```
grouped_data= dataframe.groupby('listed_in(type)')['votes'].sum()
result = pd.DataFrame({'votes': grouped_data})
plt.plot(result , c='green',marker = 'o')
plt.xlabel('Types of restaurant', c = 'red',size = 20)
plt.ylabel('Votes',c = 'red' , size = 20)
```

[10]:

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

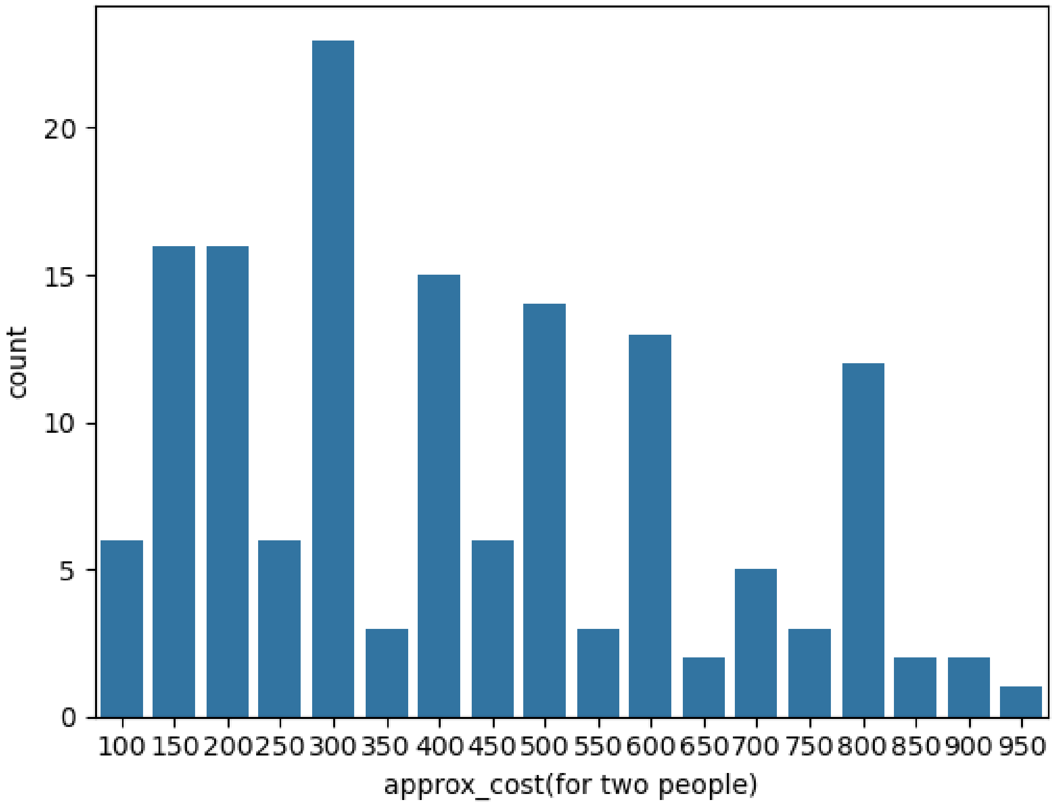


```
[12]: plt.hist(dataframe['rate'],bins = 5)
      plt.title("ratings distribution")
      plt.show()
```



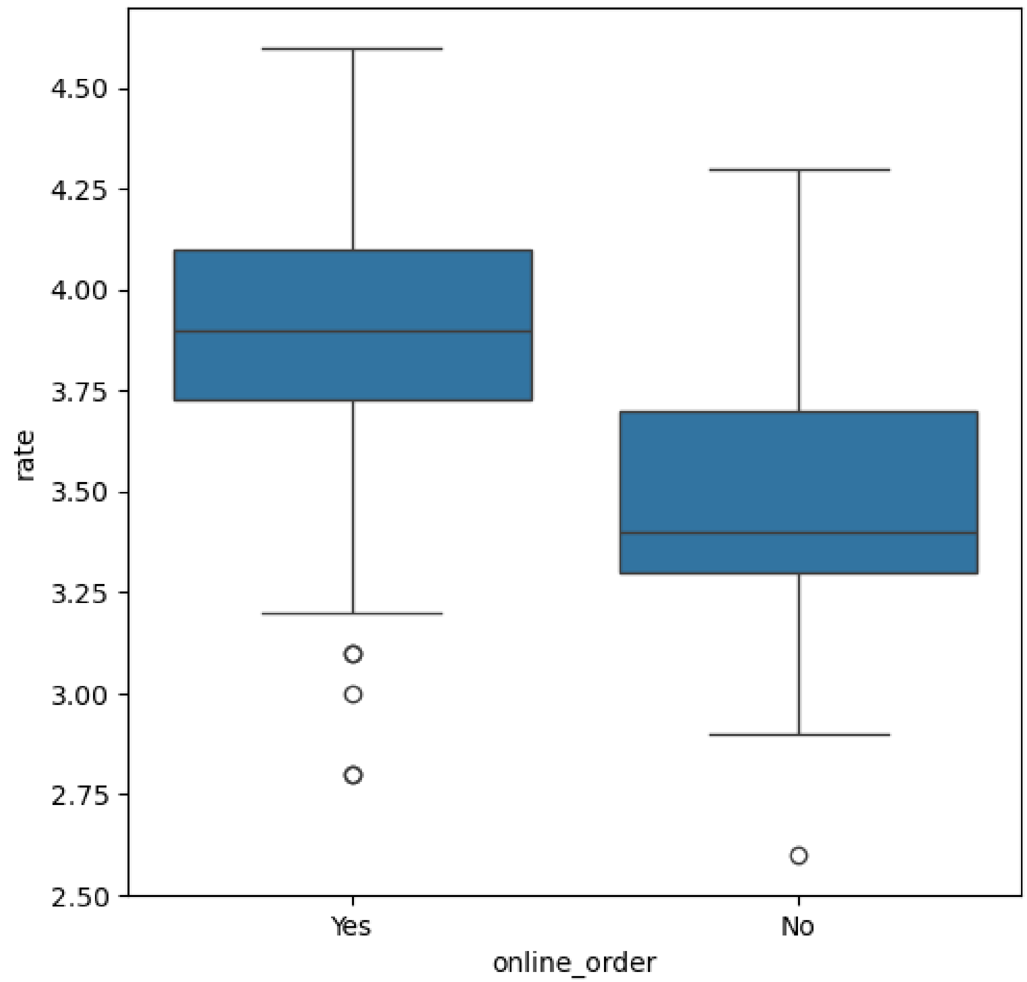
```
[14]: # Average order spendings by couples
      couple_data = dataframe['approx_cost(for two people)']
      sns.countplot(x=couple_data)
```

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

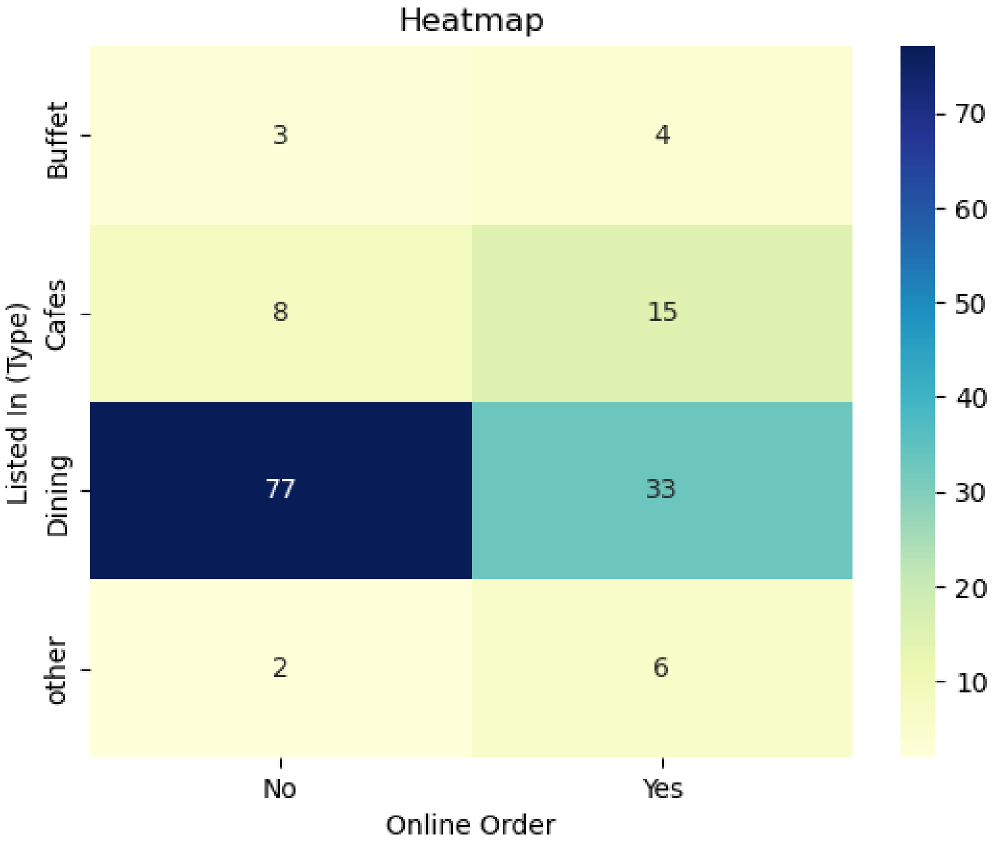


[15]:
```
# Which mode recevied maximum rating
plt.figure(figsize = (6,6))
sns.boxplot(x = 'online_order', y = 'rate' , data = dataframe)
```

[15]:
```
<Axes: xlabel='online_order', ylabel='rate'>
```



```
[16]: 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()
```



1.4 Conclusion

- **Ques 1.** What type of restaurant do the majority of the customers order from? **Answer:** Majority of the resturant falls in dining category.
- **Ques 2.** How many votes has each type of restaurant received from customers? **Answer:** Dining resturants has received maximum number of votes.
- **Ques 3.** What are the ratings that the majority of the resturants have received? **Answer:** The majority resturants received ratings from 3.5 to 4.
- **Ques 4.** Zomato has observed that most couples order of their food online. What is their average spending on each order? **Answer:** The majority of couples prefer resturants with an approx: cost of 300 rupees.
- **Ques 5.** Which mode(online or offline)has received the maximum rating? **Answer:** offline order received lower rating in comparison to online order.
- **Ques 6.** Which type of resturant received more offline orders,so that Zomato can provide customers with some good offers? **Answer:** Dining resturants primarily accept offline orders, whereas cafes primarily receive online orders. this suggests that clients prefer orders in person at resturants, but prefer online ordering at cafes.