**Zomato Data Analysis Using Python**

The following libraries were used to analyse Zomato’s Database.

Numpy– with Numpy arrays, complex computations are executed quickly, and large calculations are handled efficiently.

Matplotlib– It has a wide range of features for creating high-quality plots, charts, histograms, scatter plots, and more.

Pandas– The library simplifies the loading of data frames into 2D arrays and provides functions for performing multiple analysis tasks in a single operation.

Seaborn– It offers a high-level interface for creating visually appealing and informative statistical graphics.

Database used: [Zomato\_data](file:///C:\Users\vikki\Downloads\Zomato%20data%20.csv)

1. **IMPORT REQUIRED LIBRARIES FOR ANALYSIS:**
2. **IMPORT DATASET TO JUPYTER NOTEBOOK:**

**SYNTAX:**

*# Lets import the required libraries for analysis*

Import pandas as pd

import numpy as np

import seaborn as sns

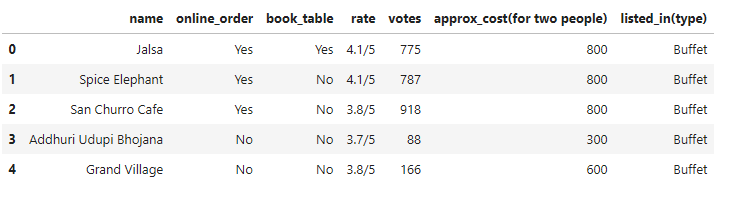
import matplotlib.pyplot as plt

*# The dataframe is named as Zomato*

Zomato = pd.read\_csv("Zomato\_data.csv")

Zomato.head()

**OUTPUT:**



1. **The datatype of rate is converted to float and the denominator is eliminated**.

**SYNTAX:**

# let’s convert the data type of the “rate” column to float and remove the denominator.

def handleRate(value):

value = str(value).split('/')

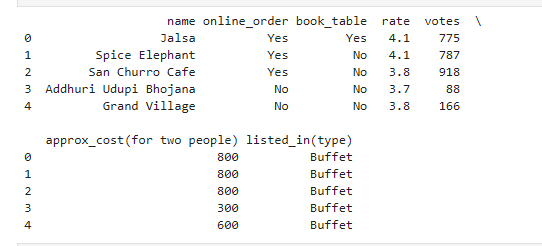
value = value[0];

return float(value)

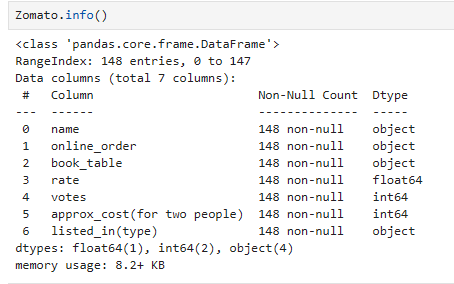
Zomato['rate'] = Zomato['rate'].apply(handleRate)

print(Zomato.head())

**OUTPUT:**



1. UNDERSTAND THE SUMMARY OF THE DATAFRAME:



1. **FOCUS ON THE TYPE OF DINING THAT THE RESTAURANT HOLDS :**

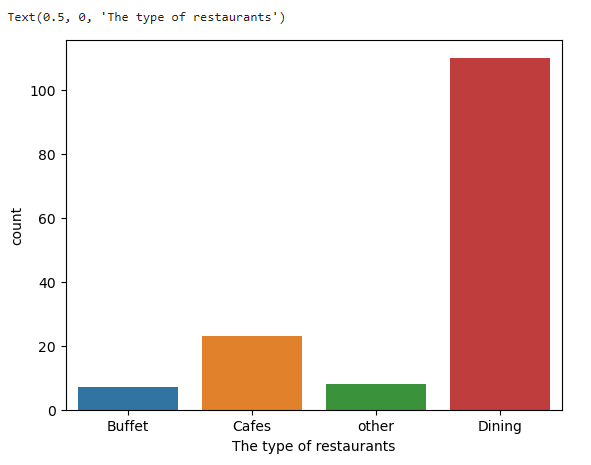
**SYNTAX:**

# Let's analyse the type of restaurants present in the dataset:

sns.countplot(x= Zomato['listed\_in(type)'])

plt.xlabel("The type of restaurants")

**OUTPUT:**



**On observation, it is evident that – Most of the restaurants offer dining compared to Buffet and Cafes.**

1. **TYPE OF RESTAURANTS PREFERRED BY INDIVIDUALS:**

**SYNTAX:**

grouped\_data = Zomato.groupby('listed\_in(type)')['votes'].sum()

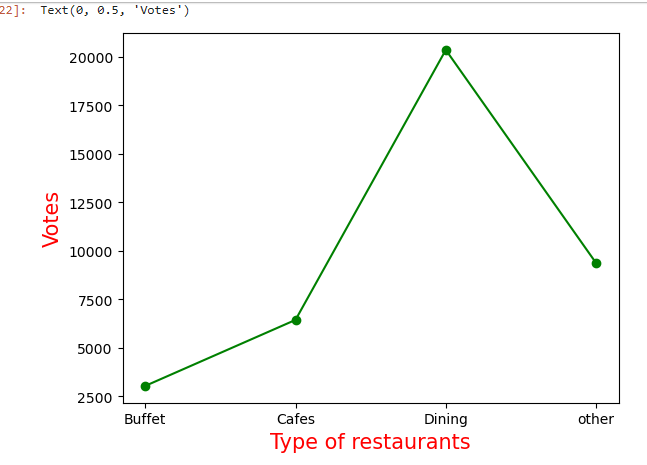
result = pd.DataFrame({'votes': grouped\_data})

plt.plot(result, c= "green", marker = "o")

plt.xlabel("Type of restaurants", c="red", size = 15)

plt.ylabel("Votes", c = "red", size = 15)

**OUTPUT:**

****

* On observation, the dining restaurants are most preferred by individuals.

1. **DETERMINE THE RESTAURANT THAT RECEIVED THE MAXIMUM NUMBER OF VOTES?**

**SYNTAX:**

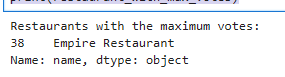
max\_votes = Zomato['votes'].max()

restaurant\_with\_max\_votes = Zomato.loc[Zomato['votes'] == max\_votes, 'name']

print("Restaurants with the maximum votes:")

print(restaurant\_with\_max\_votes)

**OUTPUT:**



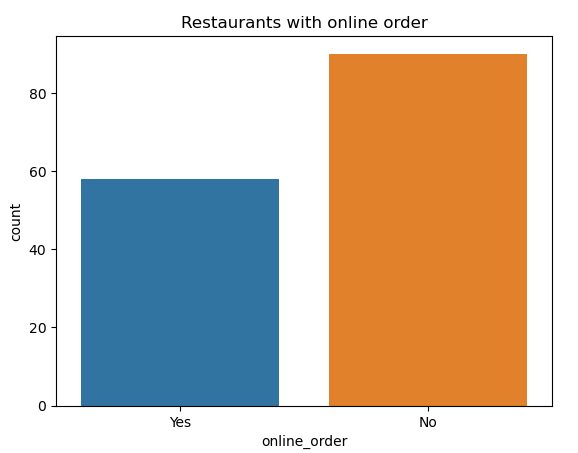
1. **EXPLORE THE COLUMN – ONLINE ORDER**

**SYNTAX:**

sns.countplot(x=Zomato['online\_order'])

plt.title("Restaurants with online order")

**OUTPUT:**

****

* Majority of the restaurants does not accept online orders.

1. **EXPLORE COLUMN – RATE:**

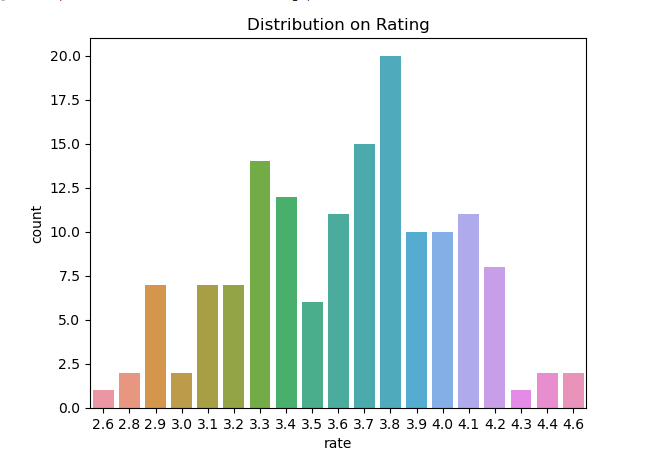
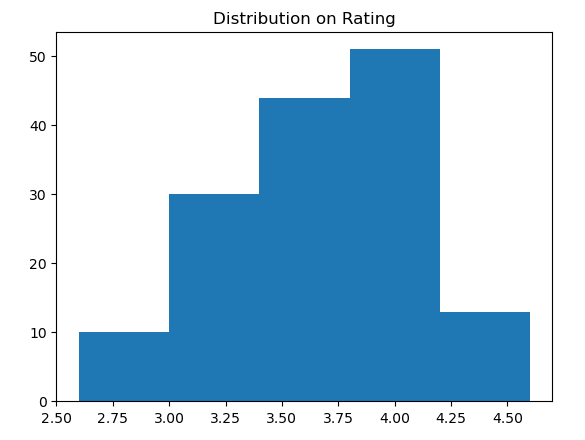
**SYNTAX:**

plt.hist(Zomato['rate'], bins = 5)

plt.title("Distribution on Rating")

plt.show()

**OUTPUT:**

** **

* Most of the restaurants hold the rating between 3.5 – 4.

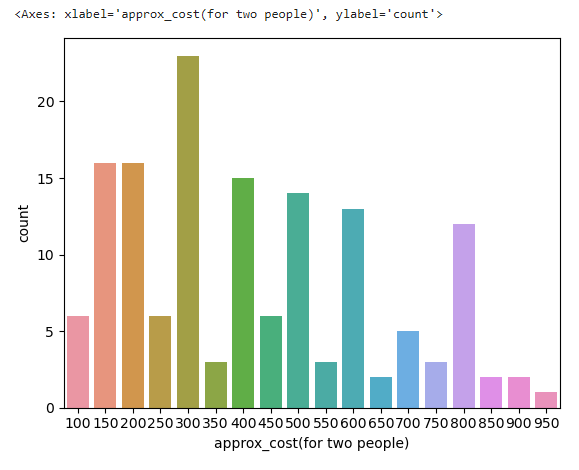
1. **EXPLORE COLUMN - APPROXIMATE COST FOR TWO PEOPLE**

**SYNTAX**

**couple\_data = Zomato['approx\_cost(for two people)']**

**sns.countplot(x= couple\_data)**

**OUTPUT**

****

* The approximate cost for two people being 300 is the most preferred among customers.

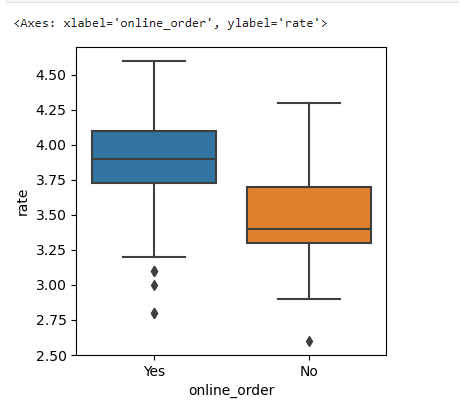
1. **EXPLORE RATING BETWEEN ONLINE AND OFFLINE ORDERS:**

**SYNTAX:**

plt.figure(figsize = (6,6))

sns.boxplot(x = 'online\_order', y = 'rate', data = Zomato)

**OUTPUT:**

****

* Offline orders received lesser ratings compared to online orders.

1. **COMPARSION BETWEEN DINING RESTAURANTS AND CAFES.**

**SYNTAX:**

pivot\_table = Zomato.pivot\_table(index='listed\_in(type)', columns='online\_order', aggfunc='size', fill\_value=0)

sns.heatmap(pivot\_table, annot=True, cmap="YlGnBu", fmt='d')

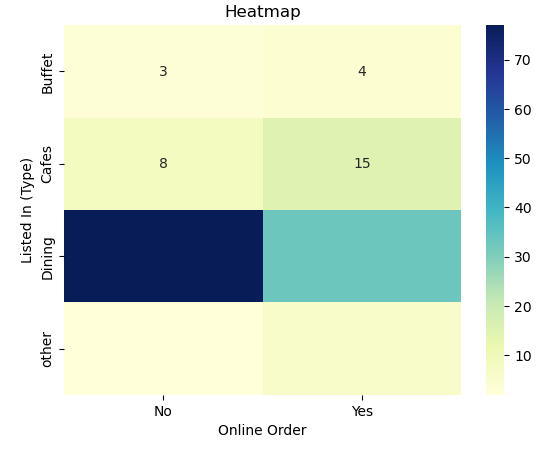
plt.title("Heatmap")

plt.xlabel("Online Order")

plt.ylabel("Listed In (Type)")

plt.show()

**OUTPUT:**

****

* Dining restaurants receive offline orders, whereas cafes receive more online orders comparatively. In clear, customers would love to take away from the restaurants but are comfortable in placing online orders on cafes.