## Web-Scraping in Python with BeautifulSoup

#### Web Scraping

Web Scraping is a technique to fetch data and information from websites.

Everything you see on a webpage can be scraped.



Can be done in most programming languages, we'll use Python.

#### Is Web Scraping illegal?

- Violation of the Computer Fraud and Abuse Act (CFAA)
- Violation of the Digital Millennium Copyright Act (DMCA)
- ► Trespass to Chattel
- Copy right infringement
- ▶ Breach of contract

#### Web crawling

- A web crawler (also known as a web spider or webrobot).
- browses the World Wide Web in a methodical, It is a program or automated script which automated manner.
- Crawlers can also be used for automating maintenance tasks on a Web site.
- Such as checking links or validating HTML code.

#### Uses

- Web Crawlers
- ► E-Commerce price comparer
- ► Preparing dataset for ML model
  - Scraping Social Media Profiles
- Weather Data etc.

Purpose of project

Purpose of the project is to scrap data from flipcart mobile price list and store data in csv.

The data would be used for sentiment analysis.



#### Libraries, Workflow and Explanation

The project is basically divided into two parts :

(a) Scraping data from website.

(b) With the help of data sentiment analysis based on rating.

### Scraping Data from Flipcart

### ►<u>Library used:</u>

- ► BeautifulSoup
- requests
- CSV
- Pandas
- ▶ matplotlib.pyplot
- Seaborn
- , numpy

```
In [1]: import pandas as pd
import matplotlib.pyplot as plt
import csv
from bs4 import BeautifulSoup
import requests
import seaborn as sp
import numpy as np
```

### Csv lib:

the most common import and export format CSV (Comma Separated Values) format is tor spreadsheets and databases.

spreadsheet which having fields or columns In this project we used csv library for create So that their values get saved its respected name, price, rating, count\_rating

fieldnames=['name','price','rating','count\_rating'] file=open('flipkart2.csv','w',encoding='utf-8'

Create flipkart.csv file with written mode and orovides fields

### Requests lib:

- Requests will allow you to send HTTP/1.1 requests using Python
- In this project we request from flipkart link

### Link or url is

c7cc6ccdb202 3.Q0QIS4SPJNLH&ppt=browse&ppn=browse&ssid=b3fqp623lc00000016383692280 https://www.flipkart.com/mobiles/pr?sid=tyy%2C4io&otracker=clp metro expandable 5 3.metr store QOQIS4SPJNLH\_wp3&fm=neo%2Fmerchandising&iid=M\_c59a68bf-abb0-414d-a5a1oExpandable.METRO EXPANDABLE Shop%2BNow mobile-phones-09&p%5B%5D=facets.brand%255B%255D%3DPOCO

Requesting:

page=requests.get(url)

### BeautifulSoup:

Beautiful Soup is a Python package used for parsing HTML In this project : we parsing the code by html.parser(it will extract all html code of webpage)

soup=BeautifulSoup(page.content,'html.parser')
print(soup)

Code is :

2" rel="search" type="application/opensearchdescription+xml"/×\meta content="website" property="og:type"/><meta content="Flip kart.com" name="og\_site\_name" property="og:site\_name"/>"/>\link href="/apple-touch-icon-57x57.png" rel="apple-touch-icon" sizes tter:app:id:ipad"/><meta content="http://dl.flipkart.com/dl/home?" name="twitter:app:url:ipad"/><meta content="Flipkart" name khtml lang="en"xkheadxlink href="https://rukminim1.flixcart.com" rel="preconnect"/xlink href="//static-assets-web.flixcart. www/linchpin/fk-cp-zion/css/app.chunk.dd97f3.css" rel="stylesheet"/>cmeta content="text/html; charset=utf-8" http-equiv="cont ent-type"/×meta content="IE-Edge" http-equiv="X-UA-Compatible"/×meta content="102988293558" property="fb:page id"/×meta co ntent="658873552,624500995,100000233612389" property="fb:admins"/>cmeta content="noodp" name="robots"/>Link href="https://st atic-assets-web.flixcart.com/www/promos/new/20150528-140547-favicon-retina.ico" rel="shortcut icon"/><link href="/osdd.xml?v= ="57X57"/×link href="/apple-touch-icon-72X72.png" rel="apple-touch-icon" sizes="72X72"/×link href="/apple-touch-icon-114X11 4.png" rel="apple-touch-icon" sizes="114x114"/×link href="/apple-touch-icon-144x144.png" rel="apple-touch-icon" sizes="144x1 nt="Flipkart" name="al:ios:app\_name"/×umeta content="742044692" name="al:ios:app\_store\_id"/×umeta content="Flipkart" name="tw ="@flipkart" name="twitter;site"/>xmeta content="@flipkart" name="twitter;creator"/>xmeta content="Mobile Price List | Compar Flipkart app Free shipping & COD." name="twitter:description"/>xmeta content="in" name="twitter:app:country"/>xmeta conte com/www/linchpin/fk-cp-zion/css/app modules.chunk.94b5e7.css" rel="stylesheet"/>/link href="//static-assets-web.flixcart.com/ e}" name="twitter:app:url:iphone"/>xmeta content="Flipkart" name="twitter:app:name:ipad"/>xmeta content="742044692" name="twi 44"/>link href="/apple-touch-icon-57x57.png" rel="apple-touch-icon"/><meta content="app" name="twitter:card"/><meta content e Mobiles on Buy Online @ Flipkart" name="twitter:title"/><meta content="Shop for electronics, apparels &amp; more using our itter:app:name:iphone"/>≾meta content="742044692" name="twitter:app:id:iphone"/>≺meta content="http://dl.flipkart.com/dl/hom

#### code

► Numbers of pages you want to scraping in this we used 4 pages which contain 72 products.

There are 8 empty lists are used for store their values

product list store mobile name or product name. battery details, hd quality, reviews respectively rating, sound, hd, reviews lists store rating P price list store prices of products

```
number_of_pages=4
product=[]
p_price=[]
rating=[]
apps= []
os= []
hd= []
sound= []
reviews=[]
```

```
for i in range(1,number_of_pages):
    j=str(i)
```

Then we work on products division which contain all details of products we call them by using their class name \_3<mark>pLy-c row</mark>

```
content=soup.find_all('div',class_='_3pLy-c row')
                                    for item in content:
```

For each class of product we have to scraping to find details

```
item_name=item.find('div',class_='_4rR01T')
item_price=item.find('div',class_='_30jeq3_1_WHN1')
item_rating=item.find('div', attrs={'class':'_3LWZlK'})
specification= item.find('div', attrs={'class':'fMghE0'})
for item in content:
```

For each product we scrap their name ,price , rating ,specification by using of their class name a each division of respected class we have same details for all products

```
class=" 4rR01T">POCO M2 Pro (Two Shades of Black, 64 GB)</div>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     class=" 4rR01T">POCO M2 Pro (Two Shades of Black, 64 GB)</div>
                                                                                                                                                                                                                  Pro (Green and Greener, 64 GB)</div>
                                                                                                              class=" 4rR01T">POCO M2 Pro (Out of the Blue, 64 GB)</div>
                                                                                                                                                                                                                                                                                                                                                                               Silver, 128 GB)</div>
                                                                                                                                                                                                                                                                                                                                                                                                  Silver, 128 GB)</div>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4rR01T">POCO M2 Pro (Out of the Blue, 64 GB)</div>
                                                                                                                                                                                                                                                                                                                                                          (Predator Black, 256 GB)</div>
                                                                                                                                                                                                                                                                                                                                       (Predator Black, 128 GB)</div>
                                                                                                                                                                                                                                                                                                                                                                                                                  (Predator Black, 128 GB)</div>
                                                                                                                                                                                                                                                                                                                  Pro 5G (Power Black, 64 GB)</div>
                                                                                                                                                                                                                                     Pro 5G (Cool Blue, 64 GB)</div>
                                                  class="_4rR01T">POCO X3 Pro (Steel Blue, 128 GB)</div>
                                                                                                                                                                                                                                                                                                                                                                                                                                      (Graphite Black, 256 GB)</div>
(Steel Blue, 64 GB)</div>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  class="_4rR01T">POCO X3 Pro (Steel Blue, 128 GB)</div>
                                                                                                                                                                           Pro 5G (Yellow, 128 GB)</div>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      class=" 4rR01T">POCO M3 Pro 5G (Yellow, 128 GB)</div>
                                                                                                                                                        Pro 5G (Yellow, 64 GB)</div>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   class=" 4rR01T">POCO M3 Pro 5G (Yellow, 64 GB)</div>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     class="_4rR01T">POCO C31 (Shadow Gray, 32 GB)</div>
                                                                                               C31 (Shadow Gray, 32 GB)</div>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (Shadow Gray, 64 GB)</div>
                                      (Shadow Gray, 64 GB)</div>
                                                                                                                                                                                                (Arctic Blue, 64 GB)</div>
                                                                                                                                                                                                                                                         (Matte Black, 64 GB)</div>
(Arctic Blue, 32 GB)</div>
                                                                                                                                     (Matte Black, 32 GB)</div>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (Royal Blue, 64 GB)</div>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Royal Blue, 32 GB)</div>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (Matte Black, 32 GB)</div>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         <div class=" 4rR01T">POCO C3 (Arctic Blue, 64 GB)</div>
                 (Royal Blue, 32 GB)</div>
<div class="_4rR@1T">POCO C31 (Royal Blue, 64 GB)</div>
<div class=" 4rR@1T">POCO C31 (Royal Blue, 32 GB)</div>
                                                                                                                                                                                                                                                                                               (Lime Green, 32 GB)</div>
                                                                                                                                                                                                                                                                                                                                                                             (Gunmetal
                                                                                                                                                                                                                                                                                                                                                                                                (Gunmetal
                                                                                                                                                                                                                                                                                                                                                                                                                      F3 GT
                                                                                                                                                                                                                                                                                                                                          GT
                                                                                                                                                                                                                                                                                                                                                             GT
                                                                                                                                                                                                                                                                                                                                                                                GT
                                                                                                                                                                                                                                                                                                                                                                                                   GT
                 class="_4rR01T">POCO C31
class=" 4rR01T">POCO C31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               class=" 4rR01T">POCO C31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       class="_4rR01T">POCO C31
                                                                                                                                                           M3
                                                                                                                                        63
                                                                                                                                                                              M3
                                                                                                                                                                                                                                                                                                                                                                                                                                           F1
                                                                                                                                                                                                                                                                                                                                                                                                                                                            F1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                class=" 4rR01T">POCO C3
                                                                                                                                                                                                   80
                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                     MM
                                                                                                                                                                                                                      Z
Z
                                                                                                                                                                                                                                       MM
                                                                                                                                                                                                                                                                               63
                                                                                        class=" 4rR01T">POCO
                                                                                                                                                                                                                                                                                                                                                                                                                                                     class=" 4rR01T">POCO
                                                                                                                                class=" 4rR01T">POCO
                                                                                                                                                    class=" 4rR01T">POCO
                                                                                                                                                                      class="_4rR01T">POCO
class="_4rR01T">POCO
class="_4rR01T">POCO
                                                                                                                                                                                                                                                                    class="_4rR01T">POCO
                                                                                                                                                                                                                                                                                                             class=" 4rR01T">POCO
                                                                                                                                                                                                                                                                                                                                                   class="_4rR01T">POCO
class="_4rR01T">POCO
class="_4rR01T">POCO
                                                                                                                                                                                                                                                                                                                                                                                                              class="_4rR01T">POCO
                                                                                                                                                                                                                               class=" 4rR01T">POCO
                                                                                                                                                                                                                                                                                                                               class=" 4rR01T">POCO
                                                                                                                                                                                                                                                  class=" 4rR01T">POCO
                                                            <div
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        <di>vip>
                                                                               vib>
                                                                                                  vib>
                                                                                                                     <div
                                                                                                                                                                              <div
                                                                                                                                                                                                                                                            vib>
                                                                                                                                                                                                                                                                                                  <div
                                                                                                                                                                                                                                                                                                                     <div>
                                                                                                                                                                                                                                                                                                                                                             <div
                                                                                                                                                                                                                                                                                                                                                                                                  vib>
                                                                                                                                                                                                                                                                                                                                                                                                                      vib>
                                                                                                                                                                                                                                                                                                                                                                                                                                                           vib>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               <div>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  <di>vip>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      <div>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             <di>vip>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                vib>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   vib>
                                                                                                                                        <div>
                                                                                                                                                            <div
                                                                                                                                                                                                   vib>
                                                                                                                                                                                                                      vib>
                                                                                                                                                                                                                                       <div
                                                                                                                                                                                                                                                                                <di>vip>
                                                                                                                                                                                                                                                                                                                                       vib>
                                                                                                                                                                                                                                                                                                                                                                                <div>
                                                                                                                                                                                                                                                                                                                                                                                                                                         <div>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          vib>
```

div>	ivy divy divy divy	/div> /div> div> div> /div> /div>	(/div> (/div> (/div> (/div> (/div> (/div> (/div> /div> /div> /div> /div>	/div> /div> /div> div> /div> /div> /div>
₹9,499 ₹20,999₹14,999 <td>499</td> <td>/&gt;666 /&gt;666 /&gt;666 /&gt;666</td> <td>730,999728,999728,999728,999721,99973,49973,49973,49973,499<td>9996, 9996, 9996, 9996, 4996, 4996,</td></td>	499	/>666 />666 />666 />666	730,999728,999728,999728,999721,99973,49973,49973,49973,499 <td>9996, 9996, 9996, 9996, 4996, 4996,</td>	9996, 9996, 9996, 9996, 4996, 4996,
1 WHN1">₹ 1 WHN1">₹ 1 WHN1">₹	WHN1"	1_WHN1">₹13,9 1_WHN1">₹14,4 1_WHN1">₹9,49 1_WHN1">₹8,49 1_WHN1">₹8,49 1_WHN1">₹8,99	1 WHN1"> 1 W	MHN1
_30jeq3 _30jeq3 _30jeq3	J.D.D.D.D.D.D.	_30jeq3 _30jeq3 _30jeq3 _30jeq3 _30jeq3 _30jeq3	30jeq3 30jeq3 30jeq3 30jeq3 30jeq3 30jeq3 30jeq3	
class=" class=" class="	S S S	Class="- class="- class="- class="- class="-	classe"." classe"." classe". classe". classe". classe	lass lass lass lass lass lass
<pre><div <="" <div="" pre=""></div></pre>	<pre></pre>	<pre><div <div="" <div<="" td=""><td><pre></pre></td><td>00000000</td></div></pre>	<pre></pre>	00000000

## For specifications for div class name = "\_1xgFaf" contains all details related to products as shown

<div class="fMghE0">4 GB RAM | 64 GB ROM | Expandable Upto 512 GB D">16.59 cm (6.53 inch) HD+ Display h Lithium-ion Polymer Batteryli>class="rgWa7D">MediaTek Helio G35 Processorli>class="rgWa7D">1 Year Warranty for</ur> Handset, 6 Months for Accessories

<div class="fMghEO">3 GB RAM | 32 GB ROM | Expandable Upto 512 GB D">16.59 cm (6.53 inch) HD+ Display h Lithium-ion Polymer Battery/1i>Lass="rgWa7D">MediaTek Helio G35 Processor/1i>Lithium-ion Polymer Battery

Handset, 6 Months for Accessories

<div class="fMghEO">4 GB RAM | 64 GB ROM | Expandable Upto 512 GB</rr> D">16.59 cm (6.53 inch) HD+ Display h Lithium-ion Polymer Batteryli class="rgWa7D">MediaTek Helio G35 Processorlithium-ion Polymer Battery</ur> Handset, 6 Months for Accessories

<div class="fMghE0">8 GB RAM | 128 GB ROM | Expandable Upto 1 TBD">16.94 cm (6.67 inch) Full HD+ Display/li>48MP + 8MP + 2MP + 2MP | 20MP Front Camera/li>/li>/li> tiple Hands-free Voice Assistantsli>class="rgWa7D">One Year Warranty for Handset, 6 Months for Accessories/li> a7D">5160 mAh Lithium-ion Polymer Batteryli>class="rgWa7D">Qualcomm Snapdragon 860 Processorli>class="rgWa7D">Mul

D">16.94 cm (6.67 inch) Full HD+ Display <div class="fMghEO">6 GB RAM | 64 GB ROM | Expandable Upto 512 GB</rr>

Extract details from specification division

Like :extract its sound details , its hd details , its os detail so one ...

```
for each in specification:
    col=each.find_all('li', attrs={'class':'rgWa7D'})
    app=col[0].text
    os_= col[1].text
    hd_= col[2].text
    sound_= col[3].text
#if item name is not Mone.
```

#### \*

# We made a dataframe that will store all values or details columns wise

#### Output:

		22				
Product_Name	Supported_apps	sound_system	so	Resolution	Price	Rating
POCO C31 (Royal Blue, 64 GB)	4 GB RAM   64 GB ROM   Expandable Upto 512 GB	5000 mAh Lithium-ion Polymer Battery	16.59 cm (6.53 inch) HD+ Display	13MP + 2MP + 2MP   5MP Front Camera	₹9,499	4.4
OCO C31 (Royal Blue, 32 GB)	3 GB RAM   32 GB ROM   Expandable Upto 512 GB	5000 mAh Lithium-ion Polymer Battery	16.59 cm (6.53 inch) HD+ Display	13MP + 2MP + 2MP   5MP Front Camera	₹8,499	4.3
CO C31 (Shadow Gray, 64 GB)	4 GB RAM   64 GB ROM   Expandable Upto 512 GB	5000 mAh Lithium-ion Polymer Battery	16.59 cm (6.53 inch) HD+ Display	13MP + 2MP + 2MP   5MP Front Camera	₹9,499	4.4
POCO X3 Pro (Steel Blue, 128 GB)	8 GB RAM   128 GB ROM   Expandable Upto 1 TB	5160 mAh Lithium-ion Polymer Battery	16.94 cm (6.67 inch) Full HD+ Display	48MP + 8MP + 2MP + 2MP   20MP Front Camera	₹20,999	4.4
OCO M2 Pro (Two Shades of Black, 64 GB)	6 GB RAM   64 GB ROM   Expandable Upto 512 GB	5000 mAh Lithium-ion Polymer Battery	16.94 cm (6.67 inch) Full HD+ Display	48MP + 8MP + 5MP + 2MP   16MP Front Camera	₹14,999	4,4
CO C31 (Shadow Gray, 32 GB)	3 GB RAM   32 GB ROM   Expandable Upto 512 GB	5000 mAh Lithium-ion Polymer Battery	16.59 cm (6.53 inch) HD+ Display	13MP + 2MP + 2MP   5MP Front Camera	₹8,499	4.3
POCO M2 Pro (Out of the Blue, 64 GB)	4 GB RAM   64 GB ROM   Expandable Upto 512 GB	5000 mAh Lithium-ion Polymer Battery	16.94 cm (6.67 inch) Full HD+ Display	48MP + 8MP + 5MP + 2MP   16MP Front Camera	₹13,999	4.3
OCO C3 (Matte Black, 32	3 GB RAM I 32 GB ROM I	5000 mAh Li-ion Polymer	16.59 cm (6.53 inch)	13MP + 2MP + 2MP I 5MP		•
	POCO C31 (Royal Blue, 32 GB) POCO C31 (Shadow Gray, 64 GB) POCO X3 Pro (Steel Blue, 128 GB) POCO M2 Pro (Two Shades of Black, 64 GB) POCO C31 (Shadow Gray, 32 GB) POCO C31 (Shadow Gray, 32 GB) POCO C31 (Matte Black, 32		3 GB RAM   32 GB ROM   Expandable Upto 512 GB 4 GB RAM   64 GB ROM   Expandable Upto 512 GB 8 GB RAM   128 GB ROM   Expandable Upto 1 TB 6 GB RAM   64 GB ROM   Expandable Upto 512 GB 3 GB RAM   32 GB ROM   Expandable Upto 512 GB 4 GB RAM   64 GB ROM   Expandable Upto 512 GB 3 GB RAM   64 GB ROM   Expandable Upto 512 GB 3 GB RAM   64 GB ROM   Expandable Upto 512 GB 3 GB RAM   64 GB ROM   Expandable Upto 512 GB 3 GB RAM   64 GB ROM   Expandable Upto 512 GB 3 GB RAM   64 GB ROM   Expandable Upto 512 GB 3 GB RAM   64 GB ROM   Expandable Upto 512 GB 3 GB RAM   64 GB ROM   Expandable Upto 512 GB 3 GB RAM   64 GB ROM   64 GB ROM	3 GB RAM   32 GB ROM   5000 mAh Lithium-ion Expandable Upto 512 GB ROM   5000 mAh Lithium-ion Expandable Upto 512 GB ROM   5160 mAh Lithium-ion Expandable Upto 512 GB RAM   128 GB ROM   5160 mAh Lithium-ion Expandable Upto 512 GB ROM   5000 mAh Lithium-ion Polymer Battery 3 GB RAM   32 GB ROM   5000 mAh Lithium-ion Polymer Battery 3 GB RAM   32 GB ROM   5000 mAh Lithium-ion Polymer Battery 3 GB RAM   32 GB ROM   5000 mAh Lithium-ion Polymer Battery 3 GB RAM   32 GB ROM   5000 mAh Lithium-ion Polymer Battery 3 GB RAM   32 GB ROM   5000 mAh Lithium-ion Polymer Battery 3 GB RAM   32 GB ROM   5000 mAh Lithium-ion Polymer Battery	SGB RAM   32 GB ROM           5000 mAh Lithium-ion Polymer Battery         16.59 cm (6.53 inch) Polymer Battery           4 GB RAM   128 GB ROM           5000 mAh Lithium-ion Polymer Battery         16.59 cm (6.53 inch) Polymer Battery           8 GB RAM   128 GB ROM           5160 mAh Lithium-ion Polymer Battery         16.94 cm (6.67 inch) Polymer Battery           6 GB RAM   64 GB ROM           5000 mAh Lithium-ion Polymer Battery         16.94 cm (6.67 inch) Polymer Battery           6 GB RAM   128 GB ROM           5000 mAh Lithium-ion Polymer Battery         16.94 cm (6.67 inch) Polymer Battery           6 GB RAM   64 GB ROM           5000 mAh Lithium-ion Polymer Battery         16.94 cm (6.67 inch) Polymer Battery           7 GB RAM   64 GB ROM           5000 mAh Lithium-ion Polymer Battery         16.94 cm (6.67 inch) Polymer Battery           8 GB RAM   64 GB ROM           5000 mAh Lithium-ion Polymer Battery         16.59 cm (6.53 inch) Polymer Battery	3 GB RAM   32 GB ROM           5000 mAh Lithium-ion Polymer Battery         16.59 cm (6.53 inch) Pront Camera         13MP + 2MP + 2MP + 2MP   5MP Pront Camera           4 GB RAM   12 GB ROM           5000 mAh Lithium-ion Polymer Battery         16.59 cm (6.53 inch) Proposer Battery         13MP + 2MP + 2MP + 2MP   5MP Pront Camera           8 GB RAM   128 GB ROM           5160 mAh Lithium-ion Polymer Battery         16.94 cm (6.67 inch) Prisplay         48MP + 8MP + 2MP + 2MP + 2MP   5MP Pront Camera           6 GB RAM   64 GB ROM           5000 mAh Lithium-ion Polymer Battery         16.94 cm (6.67 inch) Prisplay         48MP + 8MP + 5MP + 2MP + 2MP   5MP Pront Camera           7 GB RAM   64 GB ROM           5000 mAh Lithium-ion Polymer Battery         16.94 cm (6.67 inch) Prisplay         13MP + 2MP + 5MP + 2MP   5MP Pront Camera           8 GB RAM   64 GB ROM           5000 mAh Lithium-ion Polymer Battery         16.94 cm (6.67 inch) Prisplay         13MP + 2MP + 5MP + 2MP   5MP Pront Camera           9 GB RAM   64 GB ROM           5000 mAh Lithium-ion Polymer Battery         16.94 cm (6.67 inch) Prisplay         13MP + 2MP + 5MP + 2MP   7MP   7MP + 2MP   7MP   7MP + 2MP   7MP   7MP + 2MP   7MP   7MP + 2MP   7MP   7

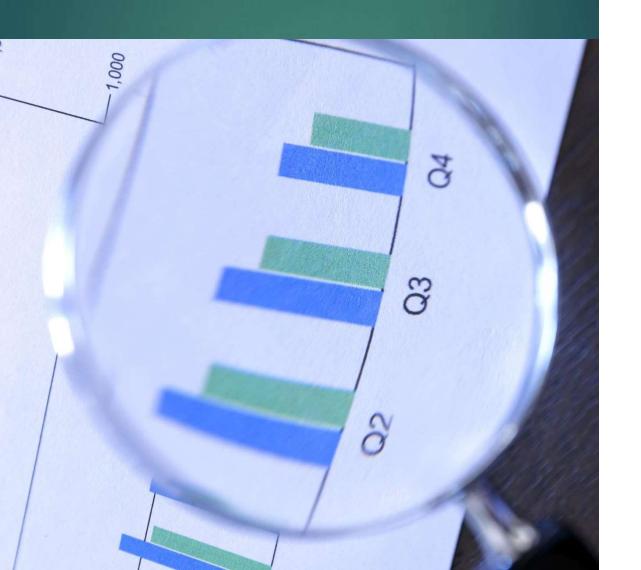
Totals Products in 4 pages are 72 and 7 columns

Now we got data and arrange them in csv file which contains all details related product

71 SAMSUNG Galaxy A03 Core 2 GB (Blue, 32 GB) Ex

72 rows × 7 columns

#### Sentiment Analysis



#### In [13]:

from textblob import TextBlob
from wordcloud import WordCloud
import re
import nltk
from nltk.corpus import stopwords
from nltk.tokenize import RegexpTokenizer
from nltk.stem.porter import PorterStemmer

### <u>Library used:</u>

- TextBlob
- WordCloud
- ©
- nltk
- Stopwords
- RegexpTokenizer
- PorterStemmer

matplot

## TextBlob and WordCloud

- TextBlob is a Python library for processing textual
- used for representing text data in which the size Word Cloud is a data visualization technique of each word indicates its frequency or importance.

## For plot word cloud we use code:

```
wordCloud = WordCloud(width = 500,height =300,random_state = 21,max_font_size =119).generate(allWords)
allWords = ''.join([twts for twts in df['Product_Name']])
                                                                                                                                          plt.imshow(wordCloud,interpolation = "bilinear")
```

In this word get cloud for samsung mobile data

## this image is plot using of matplotlib Output of wordcloud

Gala F22 Denimoglestiata

- In this we a function which tokenize the word with respect to word +
- ► And make pattern

for not take stop

Words

And clean the text

Return that all text

```
def getCleanedText(text):
    text=text.lower()
    tokenizer=RegexpTokenizer(r'\w+')
    ps=PorterStemmer()
    en_stop=set(stopwords.words('english'))
    tokens=tokenizer.tokenize(text)
    new_tokens=[token for token in tokens if token not in en_stop]
    stemmed_tokens=[ps.stem(tokens) for tokens in new_tokens]
    clean_text="".join(stemmed_tokens)
    return clean_text
```

# Apply clean text function in column

df['Price']=df['Price'].apply(getCleanedText)

4

Product_Name		Supported_apps	sound_system	so	Resolution	Price	Rating
SAMSUNG Galaxy F22 6 GB RAM   128 GB ROM   (Denim Blue, 128 GB) Expandable Upto 1 TB	6 GB RAM   128 G Expandable U	B ROM   pto 1 TB	6000 mAh Lithium-ion Battery	16.26 cm (6.4 inch) HD+ Display	48MP + 8MP + 2MP + 2MP   13MP Front Camera	14	4.3
SAMSUNG Galaxy F22 6 GB RAM   128 GB ROM   (Denim Black, 128 GB) Expandable Upto 1 TB	6 GB RAM   128 GE Expandable Up	S ROM   sto 1 TB	6000 mAh Lithium-ion Battery	16.26 cm (6.4 inch) HD+ Display	48MP + 8MP + 2MP + 2MP   13MP Front Camera	14	4.3
SAMSUNG Galaxy F12 (Sky 4 GB RAM   64 GB ROM   Blue, 64 GB) Expandable Upto 512 GB	4 GB RAM   64 GE Expandable Upto	3 ROM   512 GB	6000 mAh Lithium-ion Battery	16.55 cm (6.515 inch) HD+ Display	48MP + 5MP + 2MP + 2MP   8MP Front Camera	499	4.2
SAMSUNG Galaxy F12 4 GB RAM   64 GB ROM   (Celestial Black, 64 GB) Expandable Upto 512 GB	4 GB RAM   64 GB Expandable Upto	ROM   512 GB	6000 mAh Lithium-ion Battery	16.55 cm (6.515 inch) HD+ Display	48MP + 5MP + 2MP + 2MP   8MP Front Camera	499	4.2
SAMSUNG Galaxy F12 (Sea 4 GB RAM   64 GB ROM   Green, 64 GB) Expandable Upto 512 GB	4 GB RAM   64 GB   Expandable Upto 5	2 GB	6000 mAh Lithium-ion Battery	16.55 cm (6.515 inch) HD+ Display	48MP + 5MP + 2MP + 2MP   8MP Front Camera	499	4.2
ı		1	:	•		:	į
SAMSUNG Galaxy A03 Core 2 GB RAM   32 GB ROM   (Black, 32 GB) Expandable Upto 1 TB	2 GB RAM   32 G Expandable U	B ROM   pto 1 TB	5000 mAh Li-ion Battery	16.51 cm (6.5 inch) HD+ Display	8MP Rear Camera   5MP Front Camera	7	4.5
SAMSUNG Galaxy A70s 8 GB RAM   128 GB ROM   (Prism Crush White, 128 GB) Expandable Upto 512 GB	8 GB RAM   128 GE Expandable Upto	512 GB	4500 mAh Lithium Ion Battery	17.02 cm (6.7 inch) Full HD+ Display	64MP + 5MP + 8MP   32MP Front Camera	19	4.3
SAMSUNG Galaxy M32 6 GB RAM   128 GB ROM   (Black, 128 GB) Expandable Upto 1 TB	6 GB RAM   128 GB Expandable Up	ROM	6000 mAh Battery	16.26 cm (6.4 inch) Full HD+ Display	64MP + 8MP + 2MP + 2MP   20MP Front Camera	730	4.3

Then we normalized the rating value as true

►As it show value for 4.2 rating we get normalize value of 0.541667 ► And similar to others

new\_df['Rating'].value\_counts(normalize=True)

0.541667 0.375000 0.083333 4.3

Name: Rating, dtype: float64

## Now it turn of accuracy

## (for accuracy we design a model)

```
X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.2,random_state=10)
                                                                                                         from sklearn.feature_extraction.text import CountVectorizer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          print("Train Accuracy", RF model.score(X train, y train))
                                                from sklearn.model_selection import train_test_split
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        print("Test Accuracy", RF_model.score(X_test,y_test))
from sklearn.ensemble import RandomForestClassifier
                                                                                                                                                                                                            X_cv=cv.fit_transform(df['Product_Name']).toarray()
                                                                                                                                                           cv=CountVectorizer(ngram_range=(1,2))
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RF_model=RF.fit(X_train,y_train)
                                                                                                                                                                                                                                                                                                                                                                                                                                 RF=RandomForestClassifier()
                                                                                                                                                                                                                                                                                                                       y=df['Rating']
```

Train Accuracy 1.0 Test Accuracy 1.0

## Thank you

