

Libraries used :-

Selenium :-

- Selenium is a web testing library. It is used to automate browser activities.

BeautifulSoup :-

- BeautifulSoup is a Python package for parsing HTML and XML documents. It uses parse trees that is helpful to extract the data easily.

Pandas :-

- Pandas is a library used for data manipulation and analysis. It is used to extract the data and store it in the desired format.

In [1]:

```
from bs4 import BeautifulSoup
from selenium import webdriver
import pandas as pd
```

In [2]:

```
driver = webdriver.Chrome("/Users/yashwantjangid/Documents/chromedriver")
```

In [3]:

```
driver.get("https://www.flipkart.com/search?q=smart+phones&sid=tyy%252C4io&as=or")
```

In [4]:

```
content = driver.page_source
# parse the html using beautiful soup and store in variable `soup`
soup = BeautifulSoup(content)
```

In []:

```
products=[] #List to store name of the product
SellingPrices=[] #List to SellingPrices of the product
OriginalPrice=[] #List to OriginalPrice of the product=[]
Product_Desc=[] #List to Product Description
Specifiactions= []# List to Product Specification
ratings=[]
```

Get List of all the links of Mobiles

In [5]:

```
links=[] #List to all product links
for a in soup.findAll('a',href=True, attrs={'class':'_31qSD5'}):

    links.append(a.get('href'))
```

In [85]:

```
import requests
from urllib.request import urlopen
import re
```

In []:

In [202]:

```

products=[] #List to store name of the product
SellingPrices=[] #List to SellingPrices of the product
OriginalPrice=[] #List to OriginalPrice of the product=[]
Product_Desc=[] #List to Product Description
Specifiactions= []# List to Product Specification
ratings=[] #List to store rating of the product
Description=[]# List to descriptions of product
imageUrls=[]
for l in links[:18]:
    link = "https://www.flipkart.com"+l
    driver.get(link)
    content1 = driver.page_source
    soup = BeautifulSoup(content1)
    #page = requests.get("https://www.flipkart.com"+l)
    #pagesoup = BeautifulSoup(page.text, 'html.parser')
    nameofpro = soup.find('span',attrs={'class':'_35KyD6'}).getText()
    SPrices=soup.find('div', attrs={'class':'_1vC40E _3qQ9m1'}).getText()
    OPrice=soup.find('div', attrs={'class':'_3auQ3N _1POkHg'}).getText()

    Desc=soup.find('div',attrs={'class':'_3la3Fn _1zZOAc'}).getText()
    images = soup.find('div',attrs={'_2_AcLJ'})
    # for image in images:
        # imageUrl.append(image['src'])
    tags = soup.body.find_all('div',attrs={'class':'_2_AcLJ','style':True})

    spec=soup.find('div',attrs={'class':'_3Rrcbo'}).getText()

    rating=soup.find('div', attrs={'class':'hGSR34'}).getText()
    products.append(nameofpro)
    SellingPrices.append(SPrices)
    OriginalPrice.append(OPrice)
    Description.append(Desc)
    Specifiactions.append(spec)
    ratings.append(rating)
    imageUrls.append(images['style'])

```

```
In [ ]:
```

```
In [204]:
```

```
df = pd.DataFrame({'Product Name':pd.Series(products),
                    'Original Price':pd.Series(OriginalPrice),
                    'Average Rating':pd.Series(ratings),
                    'Selling Price':pd.Series(SellingPrices),
                    'Specifiactions':pd.Series(Specifiactions),
                    'Description':pd.Series(Description),
                    'Product Link':pd.Series(links),
                    'Image Urls':pd.Series(imageUrls)
                    })

df.style.set_properties(subset=['Specifiactions'], **{'width': '100px'})

df.to_csv('Products1.csv', index=False, encoding='utf-8')
```

In [205]:

```
pd.read_csv('Products1.csv')
```

	Product Name	Original Price	Average Rating	Selling Price	Specifiactions	Description	Product Link	
0	Redmi Note 7 Pro (Space Black, 64 GB) (4 GB RAM)	₹15,999	4.6	₹13,999	GeneralIn The BoxHandset, SIM Ejector Tool, So...	\n\tWhether it's work or entertainment, the Re...	/redmi-note-7-pro-space-black-64-gb/p/itmfeqkx...	backg image
1	Redmi Note 7 Pro (Neptune Blue, 64 GB) (4 GB RAM) ...	₹15,999	4.6	₹13,999	GeneralIn The BoxHandset, SIM Ejector Tool, So...	\n\tWhether it's work or entertainment, the Re...	/redmi-note-7-pro-neptune-blue-64-gb/p/itmfeqkx...	backg image
2	Redmi Note 7 Pro (Nebula Red, 64 GB) (4 GB RAM)	₹15,999	4.6	₹13,999	GeneralIn The BoxHandset, SIM Ejector Tool, So...	\n\tWhether it's work or entertainment, the Re...	/redmi-note-7-pro-nebula-red-64-gb/p/itmferghu...	backg image
3	Redmi Note 7 Pro (Space Black, 128 GB) (6 GB RAM) ...	₹17,999	4.6	₹16,999	GeneralIn The BoxHandset, SIM Ejector Tool, So...	\n\tWhether it's work or entertainment, the Re...	/redmi-note-7-pro-space-black-128-gb/p/itmfdxz...	backg image
4	Redmi Note 7 Pro (Neptune Blue, 128 GB) (6 GB...	₹17,999	4.6	₹16,999	GeneralIn The BoxHandset, SIM Ejector Tool, So...	\n\tWhether it's work or entertainment, the Re...	/redmi-note-7-pro-neptune-blue-128-gb/p/itmffd...	backg image
5	Redmi Note 7 Pro (Nebula Red, 128 GB) (6 GB RAM)	₹17,999	4.6	₹16,999	GeneralIn The BoxHandset, SIM Ejector Tool, So...	\n\tWhether it's work or entertainment, the Re...	/redmi-note-7-pro-nebula-red-128-gb/p/itmffyqg...	backg image
6	Redmi Note 7S (Sapphire Blue, 32 GB) (3 GB RAM)	₹11,999	4.5	₹10,999	GeneralIn The BoxHandset, Charging Cable, SIM ...	NaN	/redmi-note-7s-sapphire-blue-32-gb/p/itmfgpbac...	backg image
7	Realme 3 (Dynamic Black, 64 GB) (4 GB RAM)	₹12,999	4.5	₹10,999	GeneralIn The BoxHandset, Adapter, Micro-USB C...	The Realme 3 redefines excellent performance a...	/realme-3-dynamic-black-64-gb/p/itmfe68wrbfnzq...	backg image
8	Realme 3 (Dynamic Black, 32 GB) (3 GB RAM)	₹10,999	4.5	₹8,999	GeneralIn The BoxHandset, Adapter, Micro-USB C...	The Realme 3 redefines excellent performance a...	/realme-3-dynamic-black-32-gb/p/itmfe68wrbfnzq...	backg image
9	Realme 3 Pro (Lightning Purple, 64 GB) (6 GB RAM) ...	₹17,999	4.5	₹15,999	GeneralIn The BoxHandset, Adapter, Micro-USB C...	Take photography a notch higher with the Realm...	/realme-3-pro-lightning-purple-64-gb/p/itmffmg...	backg image

	Product Name	Original Price	Average Rating	Selling Price	Specifiactions	Description	Product Link	
10	Realme 3 Pro (Lightning Purple, 128 GB) (6 GB...	₹19,999	4.5	₹16,999	GeneralIn The BoxHandset, Adapter, Micro-USB C...	Take photography a notch higher with the Realm...	/realme-3-pro-lightning-purple-128-gb/p/itmffm...	backg image
11	Redmi 6 (Black, 64 GB) (3 GB RAM)	₹10,499	4.4	₹7,499	GeneralIn The BoxHandset, Adapter, USB Cable, ...	NaN	/redmi-6-black-64-gb/p/itmfa94eqhsxwhy?pid=MO...	backg image
12	Redmi Note 7S (Onyx Black, 32 GB) (3 GB RAM)	₹11,999	4.5	₹10,999	GeneralIn The BoxHandset, Charging Cable, SIM ...	NaN	/redmi-note-7s-onyx-black-32-gb/p/itmfgpbq3wy...	backg image
13	Realme 3 Pro (Nitro Blue, 64 GB) (4 GB RAM)	₹15,999	4.5	₹13,999	GeneralIn The BoxHandset, Adapter, Micro-USB C...	Take photography a notch higher with the Realm...	/realme-3-pro-nitro-blue-64-gb/p/itmffmg3xyrgu...	backg image
14	Redmi 6 (Gold, 64 GB) (3 GB RAM)	₹10,499	4.4	₹7,499	GeneralIn The BoxHandset, Adapter, USB Cable, ...	Click away with the Redmi 6 - it is built ...	/redmi-6-gold-64-gb/p/itmfa94ek8ujmm5z?pid=MOB...	backg image
15	OPPO A3s (Purple, 64 GB) (4 GB RAM)	₹13,990	4.3	₹9,990	GeneralIn The BoxHandset, Adapter (5V-1A), Mic...	NaN	/oppo-a3s-purple-64-gb/p/itmfgfej4rzqsepsz?pid=...	backg image
16	Redmi 6 (Rose Gold, 64 GB) (3 GB RAM)	₹10,499	4.4	₹7,499	GeneralIn The BoxHandset, Adapter, USB Cable, ...	NaN	/redmi-6-rose-gold-64-gb/p/itmfa8bdjfsd5q3n4?pi...	backg image
17	Redmi 6 (Blue, 64 GB) (3 GB RAM)	₹10,499	4.4	₹7,499	GeneralIn The BoxHandset, Adapter, USB Cable, ...	NaN	/redmi-6-blue-64-gb/p/itmfa94eu8vg7zsa?pid=MOB...	backg image
18	NaN	NaN	NaN	NaN	NaN	NaN	/realme-3-pro-nitro-blue-128-gb/p/itmffmg3xyrg...	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN	/redmi-note-7s-ruby-red-32-gb/p/itmfgpbamgc77w...	NaN
20	NaN	NaN	NaN	NaN	NaN	NaN	/honor-9i-graphite-black-64-gb/p/itmexnp94kcvd...	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN	/redmi-go-blue-16-gb/p/itmfgqsaszhhyyy6?pid=MO...	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN	/redmi-note-5-pro-gold-64-gb/p/itmfa2fc3bknpqpk...	NaN
23	NaN	NaN	NaN	NaN	NaN	NaN	/honor-9n-sapphire-blue-32-gb/p/itmfa9pgsehv6nm...	NaN

In [206]:

```
import json
rows= [products,OriginalPrice,SellingPrices,Specifiactions,ratings,Description,
with open('data.json', 'w') as outfile:
    json.dump(rows, outfile) #Store data in json file
```

In [207]:

```
pd.read_json('data.json') # print json file
```

	0	1	2	3	4	5	6	
0	Redmi Note 7 Pro (Space Black, 64 GB) (4 GB RAM)	Redmi Note 7 Pro (Neptune Blue, 64 GB) (4 GB ...	Redmi Note 7 Pro (Nebula Red, 64 GB) (4 GB RAM)	Redmi Note 7 Pro (Space Black, 128 GB) (6 GB ...	Redmi Note 7 Pro (Neptune Blue, 128 GB) (6 GB...	Redmi Note 7 Pro (Nebula Red, 128 GB) (6 GB RAM)	Redmi Note 7S (Sapphire Blue, 32 GB) (3 GB RAM)	R
1	₹15,999	₹15,999	₹15,999	₹17,999	₹17,999	₹17,999	₹11,999	₹1
2	₹13,999	₹13,999	₹13,999	₹16,999	₹16,999	₹16,999	₹10,999	₹1
3	GeneralIn The BoxHandset, SIM Ejector Tool, So...	GeneralIn The BoxHandset, SIM Ejector Tool, So...	GeneralIn The BoxHandset, SIM Ejector Tool, So...	GeneralIn The BoxHandset, SIM Ejector Tool, So...	GeneralIn The BoxHandset, SIM Ejector Tool, So...	GeneralIn The BoxHandset, SIM Ejector Tool, So...	GeneralIn The BoxHandset, Charging Cable, SIM ...	G
4	4.6	4.6	4.6	4.6	4.6	4.6	4.5	4.
5	\n\tWhether it's work or entertainment, the Re...	\n\tWhether it's work or entertainment, the Re...	\n\tWhether it's work or entertainment, the Re...	\n\tWhether it's work or entertainment, the Re...	\n\tWhether it's work or entertainment, the Re...	\n\tWhether it's work or entertainment, the Re...	NA	TI
6	/redmi-note-7-pro-space-black-64-gb/p/itmfeqkx...	/redmi-note-7-pro-neptune-blue-64-gb/p/itmfeqkx...	/redmi-note-7-pro-neptune-blue-64-gb/p/itmferghu...	/redmi-note-7-pro-space-black-128-gb/p/itmfdxz...	/redmi-note-7-pro-neptune-blue-128-gb/p/itmffd...	/redmi-note-7-pro-neptune-blue-128-gb/p/itmffyqg...	/redmi-note-7s-sapphire-blue-32-gb/p/itmfgpbac...	/re

7 rows × 24 columns

*Dump Json into mongoDb using **PyMongo** *

In []:

```
import pymongo

myclient = pymongo.MongoClient("mongodb://localhost:27017/")
mydb = myclient["WebScrapingDb"] #Database name WebScrapingDb

x = mydb.insert_many(rows)
```

