## Libraries used :-

### Selenium :-

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

### BeautifulSoup :-

> - Beautiful Soup is a Python package for parsing HTML and XML documents. It
s parse trees that is helpful to extract the data easily.

### Pandas :-

> - Pandas is a library used for data manipulation and analysis. It is used tc
ct the data and store it in the desired format.

In [81]:

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

In [11]:

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

In [13]:

```
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

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

In [82]:

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

In [119]:

```
import re
Image_Url=[] # list to image url
#bs = BeautifulSoup(html, 'html.parser')
soup = BeautifulSoup(content)
for a in soup.findAll('a',href=True, attrs={'class':'_31qSD5'}):
    name=a.find('div', attrs={'class':'_3wU53n'})
    SPrices=a.find('div', attrs={'class':'_1vC4OE _2rQ-NK'})
    OPrice=a.find('div', attrs={'class':'_3auQ3N _2GcJzG'})
    spec=a.find('ul',attrs={'class':'vFw0gD'})
    images = a.findAll('img')
    for image in images:
        Image_Url.append(image['src'])


    rating=a.find('div', attrs={'class':'hGSR34 _2_KrJI'})
    products.append(name.text)
    SellingPrices.append(SPrices.text)
    OriginalPrice.append(OPrice)
    Specifiactions.append(spec.text)
    ratings.append(rating)
```

In [122]:

```
df = pd.DataFrame({'Product Name':pd.Series(products[:23]),
                   'Original Price':pd.Series(OriginalPrice[:23]),
                   'Selling Price':pd.Series(SellingPrices[:23]),
                   'Specifiactions':pd.Series(Specifiactions[:23]),
                   'Image Url':pd.Series(Image_Url[:23]),
                  'Rating':pd.Series(ratings[:23])})
df.style.set_properties(subset=['Specifiactions'], **{'width': '100px'})

df.to_csv('products3.csv', index=False, encoding='utf-8')
```

In [129]:

```
pd.read_csv('products3.csv')
```

| | Product Name | Original Price | Selling Price | Specifiactions | |
|---|---|---|---|---|---|
| 0 | Redmi Note 7 Pro (Space Black, 64 GB) | ₹15,999 | ₹13,999 | 4 GB RAM I 64 GB ROM I Expandable Upto 256 GB1... | https://rukminim1.flixcart.com/image/3 |
| 1 | Redmi Note 7 Pro (Space Black, 64 GB) | ₹15,999 | ₹13,999 | 4 GB RAM I 64 GB ROM I Expandable Upto 256 GB1... |  |
| 2 | Redmi Note 7 Pro (Neptune Blue, 64 GB) | ₹15,999 | ₹13,999 | 4 GB RAM I 64 GB ROM I Expandable Upto 256 GB1... | //img1a.flixcart.com/www/linchpin/fk-cp |
| 3 | Redmi Note 7S (Sapphire Blue, 32 GB) | ₹11,999 | ₹10,999 | 3 GB RAM I 32 GB ROM I Expandable Upto 256 GB1... | https://rukminim1.flixcart.com/image/3 |
| 4 | Realme 3 (Dynamic Black, 64 GB) | ₹12,999 | ₹10,999 | 4 GB RAM I 64 GB ROM I Expandable Upto 256 GB1... |  |
| 5 | Redmi Note 7S (Onyx Black, 32 GB) | ₹11,999 | ₹10,999 | 3 GB RAM I 32 GB ROM I Expandable Upto 256 GB1... | //img1a.flixcart.com/www/linchpin/fk-cp |
| 6 | Realme 3 (Dynamic Black, 32 GB) | ₹10,999 | ₹8,999 | 3 GB RAM I 32 GB ROM I Expandable Upto 256 GB1... | https://rukminim1.flixcart.com/image/3 |
| 7 | Realme 3 Pro (Lightning Purple, 128 GB) | ₹19,999 | ₹16,999 | 6 GB RAM I 128 GB ROM I Expandable Upto 256 GB... |  |
| 8 | Realme 3 Pro (Lightning Purple, 64 GB) | ₹17,999 | ₹15,999 | 6 GB RAM I 64 GB ROM I Expandable Upto 256 GB1... | //img1a.flixcart.com/www/linchpin/fk-cp |
| 9 | Redmi 6 (Blue, 64 GB) | ₹10,499 | ₹7,499 | 3 GB RAM I 64 GB ROM I Expandable Upto 256 GB1... | https://rukminim1.flixcart.com/image/3 |
| 10 | Redmi Note 7S (Ruby Red, 32 GB) | ₹11,999 | ₹10,999 | 3 GB RAM I 32 GB ROM I Expandable Upto 256 GB1... |  |
| 11 | Redmi 6 (Black, 64 GB) | ₹10,499 | ₹7,499 | 3 GB RAM I 64 GB ROM I Expandable Upto 256 GB1... | //img1a.flixcart.com/www/linchpin/fk-cp |
| 12 | OPPO A3s (Purple, 64 GB) | ₹13,990 | ₹9,990 | 4 GB RAM I 64 GB ROM I Expandable Upto 256 GB1... | https://rukminim1.flixcart.com/image/3 |
| 13 | Redmi 6 (Gold, 64 GB) | ₹10,499 | ₹7,499 | 3 GB RAM I 64 GB ROM I Expandable Upto 256 GB1... |  |
| 14 | Redmi 6 (Rose Gold, 64 GB) | ₹10,499 | ₹7,499 | 3 GB RAM I 64 GB ROM I Expandable Upto 256 GB1... | //img1a.flixcart.com/www/linchpin/fk-cp |
| 15 | Realme 3 Pro (Nitro Blue, 128 GB) | ₹19,999 | ₹16,999 | 6 GB RAM I 128 GB ROM I Expandable Upto 256 GB... | https://rukminim1.flixcart.com/image/3 |
| 16 | Redmi Note 5 Pro (Gold, 64 GB) | ₹16,999 | ₹11,999 | 6 GB RAM I 64 GB ROM I Expandable Upto 128 GB1... |  |
| 17 | Infinix Smart 3 Plus (Sapphire Cyan, 32 GB) | ₹7,999 | ₹6,999 | 2 GB RAM I 32 GB ROM I Expandable Upto 256 GB1... | //img1a.flixcart.com/www/linchpin/fk-cp |
| 18 | Honor 9i (Graphite Black, 64 GB) | ₹19,999 | ₹8,999 | 4 GB RAM I 64 GB ROM I Expandable Upto 128 GB1... | https://rukminim1.flixcart.com/image/3 |
| 19 | Redmi Go (Blue, 16 GB) | ₹5,999 | ₹4,799 | 1 GB RAM I 16 GB ROM I Expandable Upto 128 GB1... |  |

| | Product Name | Original Price | Selling Price | Specifiactions | |
|---|---|---|---|---|---|
| 20 | Redmi Note 6 Pro (Red, 64 GB) | ₹15,999 | ₹11,999 | 4 GB RAM I 64 GB ROM I Expandable Upto 256 GB1... | //img1a.flixcart.com/www/linchpin/fk-cp |
| 21 | Honor 9N (Sapphire Blue, 32 GB) | ₹13,999 | ₹8,499 | 3 GB RAM I 32 GB ROM I Expandable Upto 256 GB1... | https://rukminim1.flixcart.com/image/3 |
| 22 | Honor 9i (Aurora Blue, 64 GB) | ₹19,999 | ₹8,999 | 4 GB RAM I 64 GB ROM I Expandable Upto 128 GB1... |  |

In [136]:

In [138]:

```
import json
rows= [products[:23],OriginalPrice[:23],SellingPrices[:23],Specifiactions[:23],
with open('data.json', 'w') as outfile:
    json.dump(rows, outfile) #Store data in json file
```

In [139]:

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

| | 0 | 1 | |
|---|---|---|---|
| 0 | Redmi Note 7 Pro (Space Black, 64 GB) | Redmi Note 7 Pro (Space Black, 64 GB) | Redmi Note 7 64 GB) |
| 1 | ₹15,999 | ₹15,999 | ₹15,999 |
| 2 | ₹13,999 | ₹13,999 | ₹13,999 |
| 3 | 4 GB RAM I 64 GB ROM I Expandable Upto 256 GB1... | 4 GB RAM I 64 GB ROM I Expandable Upto 256 GB1... | 4 GB RAM I 6 Expandable U |
| 4 | https://rukminim1.flixcart.com/image/312/312/j... | ... | //img1a.flixcar cp-zion/i... |
| 5 | 4.6 | 4.6 | 4.5 |

6 rows × 23 columns

**Dump Json into mongoDb using PyMongo **

In [ ]:

```
import pymongo

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

mydict = { "name": "John", "address": "Highway 37" }
x = mydb.insert_many(rows)
```

In [ ]: