#Q1: Write a python program to scrape data for “Data Analyst” Job position in “Bangalore” location. You have to scrape the job-title, job-location, company_name, experience_required. You have to scrape first 10 jobs data. This task will be done in following steps:

1. First get the webpage https://www.shine.com/ (https://www.shine.com/)
2. Enter “Data Analyst” in “Job title, Skills” field and enter “Bangalore” in “enter the location” field.
3. Then click the searchbutton.
4. Then scrape the data for the first 10 jobs results you get.
5. Finally create a dataframe of the scraped data. Note

In [2]:
```
!pip install selenium
```

```
Requirement already satisfied: selenium in c:\users\user\anaconda3\lib\site-p
ackages (4.13.0)
Requirement already satisfied: urllib3[socks]<3,>=1.26 in c:\users\user\anaco
nda3\lib\site-packages (from selenium) (1.26.16)
Requirement already satisfied: trio~=0.17 in c:\users\user\anaconda3\lib\site
-packages (from selenium) (0.22.2)
Requirement already satisfied: trio-websocket~=0.9 in c:\users\user\anaconda3
\lib\site-packages (from selenium) (0.10.4)
Requirement already satisfied: certifi>=2021.10.8 in c:\users\user\anaconda3
\lib\site-packages (from selenium) (2023.7.22)
Requirement already satisfied: attrs>=20.1.0 in c:\users\user\anaconda3\lib\s
ite-packages (from trio~=0.17->selenium) (22.1.0)
Requirement already satisfied: sortedcontainers in c:\users\user\anaconda3\li
b\site-packages (from trio~=0.17->selenium) (2.4.0)
Requirement already satisfied: idna in c:\users\user\anaconda3\lib\site-packa
ges (from trio~=0.17->selenium) (3.4)
Requirement already satisfied: outcome in c:\users\user\anaconda3\lib\site-pa
ckages (from trio~=0.17->selenium) (1.2.0)
Requirement already satisfied: sniffio in c:\users\user\anaconda3\lib\site-pa
ckages (from trio~=0.17->selenium) (1.2.0)
Requirement already satisfied: cffi>=1.14 in c:\users\user\anaconda3\lib\site
-packages (from trio~=0.17->selenium) (1.15.1)
Requirement already satisfied: exceptiongroup in c:\users\user\anaconda3\lib
\site-packages (from trio-websocket~=0.9->selenium) (1.1.3)
Requirement already satisfied: wsproto>=0.14 in c:\users\user\anaconda3\lib\s
ite-packages (from trio-websocket~=0.9->selenium) (1.2.0)
Requirement already satisfied: PySocks!=1.5.7,<2.0,>=1.5.6 in c:\users\user\a
naconda3\lib\site-packages (from urllib3[socks]<3,>=1.26->selenium) (1.7.1)
Requirement already satisfied: pycparser in c:\users\user\anaconda3\lib\site-
packages (from cffi>=1.14->trio~=0.17->selenium) (2.21)
Requirement already satisfied: h11<1,>=0.9.0 in c:\users\user\anaconda3\lib\s
ite-packages (from wsproto>=0.14->trio-websocket~=0.9->selenium) (0.14.0)
```

In [3]:
```
1 !pip install --upgrade selenium
```

```
Requirement already satisfied: selenium in c:\users\user\anaconda3\lib\site-p
ackages (4.13.0)
Requirement already satisfied: urllib3[socks]<3,>=1.26 in c:\users\user\anaco
nda3\lib\site-packages (from selenium) (1.26.16)
Requirement already satisfied: trio~=0.17 in c:\users\user\anaconda3\lib\site
-packages (from selenium) (0.22.2)
Requirement already satisfied: trio-websocket~=0.9 in c:\users\user\anaconda3
\lib\site-packages (from selenium) (0.10.4)
Requirement already satisfied: certifi>=2021.10.8 in c:\users\user\anaconda3
\lib\site-packages (from selenium) (2023.7.22)
Requirement already satisfied: attrs>=20.1.0 in c:\users\user\anaconda3\lib\s
ite-packages (from trio~=0.17->selenium) (22.1.0)
Requirement already satisfied: sortedcontainers in c:\users\user\anaconda3\li
b\site-packages (from trio~=0.17->selenium) (2.4.0)
Requirement already satisfied: idna in c:\users\user\anaconda3\lib\site-packa
ges (from trio~=0.17->selenium) (3.4)
Requirement already satisfied: outcome in c:\users\user\anaconda3\lib\site-pa
ckages (from trio~=0.17->selenium) (1.2.0)
Requirement already satisfied: sniffio in c:\users\user\anaconda3\lib\site-pa
ckages (from trio~=0.17->selenium) (1.2.0)
Requirement already satisfied: cffi>=1.14 in c:\users\user\anaconda3\lib\site
-packages (from trio~=0.17->selenium) (1.15.1)
Requirement already satisfied: exceptiongroup in c:\users\user\anaconda3\lib
\site-packages (from trio-websocket~=0.9->selenium) (1.1.3)
Requirement already satisfied: wsproto>=0.14 in c:\users\user\anaconda3\lib\s
ite-packages (from trio-websocket~=0.9->selenium) (1.2.0)
Requirement already satisfied: PySocks!=1.5.7,<2.0,>=1.5.6 in c:\users\user\a
naconda3\lib\site-packages (from urllib3[socks]<3,>=1.26->selenium) (1.7.1)
Requirement already satisfied: pycparser in c:\users\user\anaconda3\lib\site-
packages (from cffi>=1.14->trio~=0.17->selenium) (2.21)
Requirement already satisfied: h11<1,>=0.9.0 in c:\users\user\anaconda3\lib\s
ite-packages (from wsproto>=0.14->trio-websocket~=0.9->selenium) (0.14.0)
```

In [4]:
```
1 import selenium
```

In [5]:
```
1 import pandas as pd
```

In [6]:
```
1 from selenium import webdriver
```

In [7]:
```
1 import warnings
```

In [8]:
```
1 warnings.filterwarnings('ignore')
```

In [9]:
```
1 from selenium.webdriver.common.by import By
```

In [10]:
```
import time
```

In [11]:
```
driver=webdriver.Chrome()
```

In [12]:
```
#opening the website on atometed chromdriver.
driver.get(" https://www.shine.com/")
```

In [13]:
```
time.sleep(10)
```

In [15]:
```
# entering designation and location required
designation=driver.find_element(By.XPATH,"/html/body/div[1]/div[4]/div/div
designation.send_keys('Data Analyst')
```

In [16]:
```
location=driver.find_element(By.ID,"id_loc")
location.send_keys("Bangalore")
```

In [17]:
```
search=driver.find_element(By.XPATH,"/html/body/div[1]/div[4]/div/div[2]/d
search.click()
```

In [18]:
```
job_title=[]
job_location=[]
company_name=[]
experince_required=[]
```

In [19]:
```
#scraping job title from given page.
title_tags=driver.find_elements(By.XPATH,'//h2[@itemprop="name"]')
for i in title_tags[0:10]:
    title=i.text
    job_title.append(title)
#scraping location form given page.
location_tags=driver.find_elements(By.XPATH,'//div[@class=" jobCard_jobCar
for i in location_tags[0:10]:
    location=i.text
    job_location.append(location)

#scraping companny name from given page.
company_tags=driver.find_elements(By.XPATH,'//div[@class="jobCard_jobCard_
for i in company_tags[0:10]:
    company=i.text
    company_name.append(company)
#scraping experince required from given page.

experince_tags=driver.find_elements(By.XPATH,'//div[@class=" jobCard_jobCa
for i in experince_tags[0:10]:
    experince=i.text
    experince_required.append(experince)

```

In [20]:
```
print(len(job_title),len(job_location),len(company_name),len(experince_rec
```

```
10 10 10 10
```

In [24]:
```
import pandas as pd
df=pd.DataFrame({'title':job_title,'Location':job_location,'Compnany_name
df
```

Out[24]:

	title	Location	Compnany_name	Experince_required
0	Technical Data Analyst	Bangalore\n+9	krishna consultancy.. hiring for co...	4 to 9 Yrs
1	KYC Client Data Analyst	Bangalore	jpmorgan	4 to 6 Yrs
2	opening data analyst - Bangalore	Bangalore	provizor it services pvt. ltd.	2 to 5 Yrs
3	Senior Data Analyst	Bangalore\n+8	harjai computers private limited	0 to 2 Yrs
4	Data Warehouse _ Manager For bangalore	Bangalore	boyen haddin consulting and technol...	3 to 6 Yrs
5	Data Warehouse _Manager	Bangalore	boyen haddin consulting and technol...	3 to 6 Yrs
6	Data Modeler _Bangalore	Bangalore	boyen haddin consulting and technol...	3 to 6 Yrs
7	Data Modeler _Bangalore	Bangalore	boyen haddin consulting and technol...	3 to 6 Yrs
8	Hiring for -Data Modeler - Bangalore	Bangalore	boyen haddin consulting and technol...	3 to 6 Yrs
9	for Data Modeler _Bangalore	Bangalore	boyen haddin consulting and technol...	3 to 6 Yrs

In [25]:
```
driver.close()
```

Q2:Write a python program to scrape data for “Data Scientist” Job position in“Bangalore” location. You have to scrape the job-title, job-location, company_name. You have to scrape first 10 jobs data. This task will be done in following steps:

1. First get the webpage https://www.shine.com/ (https://www.shine.com/)
2. Enter “Data Scientist” in “Job title, Skills” field and enter “Bangalore” in “enter thelocation” field.
3. Then click the search button.
4. Then scrape the data for the first 10 jobs results you get.
5. Finally create a dataframe of the scraped data.

In [26]:
```
import selenium
import warnings
import pandas as pd
import time
from selenium import webdriver
from selenium.webdriver.common.by import By
warnings.filterwarnings('ignore')
```

In [31]:
```
#opening the website on aoutomated chrome driver.
driver=webdriver.Chrome()
driver.get(" https://www.shine.com/")
```

In [32]:
```
time.sleep(5)
```

In [34]:
```
#enteering required designation and location required.
designation=driver.find_element(By.ID,"id_q")
designation.send_keys("Data Scientist")
location=driver.find_element(By.XPATH,'/html/body/div[1]/div[4]/div/div[2]
location.send_keys("Bangalore")
```

In [36]:
```
search=driver.find_element(By.XPATH,"/html/body/div[1]/div[4]/div/div[2]/d
search.click()
```

In [37]:
```
job_title=[]
job_location=[]
company_name=[]
experince_required=[]
```

In [38]:
```
#scraping job title from given page.
title_tags=driver.find_elements(By.XPATH,'//h2[@itemprop="name"]')
for i in title_tags[0:10]:
    title=i.text
    job_title.append(title)
#scraping location form given page.
location_tags=driver.find_elements(By.XPATH,'//div[@class=" jobCard_jobCar
for i in location_tags[0:10]:
    location=i.text
    job_location.append(location)

#scraping companny name from given page.
company_tags=driver.find_elements(By.XPATH,'//div[@class="jobCard_jobCard_
for i in company_tags[0:10]:
    company=i.text
    company_name.append(company)
#scraping experince required from given page.

experince_tags=driver.find_elements(By.XPATH,'//div[@class=" jobCard_jobCa
for i in experince_tags[0:10]:
    experince=i.text
    experince_required.append(experince)

```

In [39]:
```
print(len(job_title),len(job_location),len(company_name),len(experince_rec
```

```
10 10 10 10
```

In [40]:
```
import pandas as pd
df=pd.DataFrame({'title':job_title,'Location':job_location,'Compnany_name
df
```

Out[40]:

	title	Location	Compnany_name	Experince_required
0	Senior Data Scientist	Bangalore\n+7	nustar technologies india private l...	5 to 10 Yrs
1	Data Scientist	Bangalore	mercede	14 to 16 Yrs
2	Data Scientist	Bangalore	the fashion cosmo	0 to 4 Yrs
3	SR. DATA SCIENTIST	Bangalore	te connectivity india pvt ltd	5 to 7 Yrs
4	Senior Data Scientist- ML Engineer	Bangalore	sap india pvt ltd	5 to 7 Yrs
5	Data Scientist Hiring Fresher and Experience	Bangalore\n+17	kavya staffing solutions	0 to 4 Yrs
6	Data Scientist III	Bangalore	inmobi	7 to 9 Yrs
7	Data Scientist	Bangalore\n+6	quiscon biotech	0 to 1 Yr
8	Senior Data Scientist, AWS	Bangalore	amazon	10 to 12 Yrs
9	Data Scientist Recruitment	Bangalore\n+17	kavya staffing solutions	0 to 4 Yrs

In [41]:
```
driver.close()
```

Q3: In this question you have to scrape data using the filters available on the webpage You have to use the location and salary filter. You have to scrape data for “Data Scientist” designation for first 10 job results. You have to scrape the job-title, job-location, company name, experience required. The location filter to be used is “Delhi/NCR”. The salary filter to be used is “3-6” lakhs The task will be done as shown in the below steps:

1. first get the web page https://www.shine.com/ (https://www.shine.com/)
2. Enter “Data Scientist” in “Skill, Designations, and Companies” field.
3. Then click the search button.
4. Then apply the location filter and salary filter by checking the respective boxes
5. Then scrape the data for the first 10 jobs results you get.
6. Finally create a dataframe of the scrapeddata.

In [73]:
```
import selenium
import warnings
import pandas as pd
import time
from selenium import webdriver
from selenium.webdriver.common.by import By
warnings.filterwarnings('ignore')
```

In [74]:
```
#opening the website on aoutomated chrome driver.
driver=webdriver.Chrome()
driver.get(" https://www.shine.com/")
```

In [75]:
```
time.sleep(5)
```

In [81]:
```
#enteering required designation required.
designation=driver.find_element(By.ID,"id_q")
designation.send_keys("Data Scientist")

```

In [82]:
```
search=driver.find_element(By.XPATH,"/html/body/div[1]/div[4]/div/div[2]/d
search.click()
```

In [83]:
```
filter_by=driver.find_element(By.XPATH,'/html/body/div[1]/div[1]/div[4]/di
filter_by.click()
```

In [84]:
```
location = driver.find_element(By.XPATH,'/html/body/div[1]/div[1]/div[4]/d
location.send_keys("Delhi")
```

In [85]:
```
show_results=driver.find_element(By.XPATH,'/html/body/div[1]/div[1]/div[4
show_results.click()
```

In [66]:
```
filter_by=driver.find_element(By.XPATH,'/html/body/div[1]/div[1]/div[4]/di
filter_by.click()
```

In [67]:
```
salary=driver.find_element(By.XPATH,'/html/body/div[1]/div[1]/div[4]/div/d

salary.send_keys("3-6,Lakhs")
```

In [68]:
```
show_results=driver.find_element(By.XPATH,'/html/body/div[1]/div[1]/div[4
show_results.click()
```

In [69]:
```
job_title=[]
job_location=[]
company_name=[]
experince_required=[]
```

In [70]:
```
#scraping job title from given page.
title_tags=driver.find_elements(By.XPATH,'//h2[@itemprop="name"]')
for i in title_tags[0:10]:
    title=i.text
    job_title.append(title)
#scraping location form given page.
location_tags=driver.find_elements(By.XPATH,'//div[@class=" jobCard_jobCar
for i in location_tags[0:10]:
    location=i.text
    job_location.append(location)

#scraping companny name from given page.
company_tags=driver.find_elements(By.XPATH,'//div[@class="jobCard_jobCard_
for i in company_tags[0:10]:
    company=i.text
    company_name.append(company)
#scraping experince required from given page.

experince_tags=driver.find_elements(By.XPATH,'//div[@class=" jobCard_jobCa
for i in experince_tags[0:10]:
    experince=i.text
    experince_required.append(experince)

```

In [71]:
```
print(len(job_title),len(job_location),len(company_name),len(experince_rec
```

```
10 10 10 10
```

In [72]:
```
import pandas as pd
df=pd.DataFrame({'title':job_title,'Location':job_location,'Compnany_name
df
```

Out[72]:

	title	Location	Compnany_name	Experince_required
0	Data Scientist (ML/AI)	Pune	gfl recruitment private limited	0 to 1 Yr
1	Data Scientist (ML/AI) Pune	Pune	gfl recruitment private limited	0 to 1 Yr
2	Data analyst / Data scientist	Mumbai City	gfl recruitment private limited	0 to 2 Yrs
3	Senior Data Scientist	Bangalore\n+7	nustar technologies india private l...	5 to 10 Yrs
4	Data Scientist	Australia\n+10	vaishnavi services	14 to 24 Yrs
5	Data Scientist	Bangalore	mercede	14 to 16 Yrs
6	Data Scientist	Pune	bhs corrugated india private limite...	4 to 7 Yrs
7	Data Scientist	Bangalore	the fashion cosmo	0 to 4 Yrs
8	Data Scientist Mumbai	Mumbai City	morpheus human consulting pvt ltd	2 to 7 Yrs
9	Associate Divisional Manager Data Scientist Ch...	Chennai	talent leads hr solutions pvt ltd	7 to 12 Yrs

Q4: Scrape data of first 100 sunglasses listings on flipkart.com. You have to scrape four attributes: 6. Brand 7. ProductDescription 8. Price The attributes which you have to scrape is ticked marked in the below image. To scrape the data you have to go through following steps:

1. Go to Flipkart webpage by url :https://www.flipkart.com/ (https://www.flipkart.com/)
2. Enter “sunglasses” in the search fieldwhere “search for products, brands and more” is written and click the search icon
3. After that you will reach to the page having a lot of sunglasses. From this page you can scrap the required data as usua
4. After scraping data from the first page, go to the “Next” Button at the bottom other page , then click on it.
5. Now scrape data from this page as usual
6. Repeat this until you get data for 100sunglasses.

In [17]:
```
import selenium
import warnings
import pandas as pd
import time
from selenium import webdriver
from selenium.webdriver.common.by import By
warnings.filterwarnings('ignore')
```

In [18]:
```
#opening the website on aoutomated chrome driver.
driver=webdriver.Chrome()
driver.get("https://www.flipkart.com/")
time.sleep(3)
```

In [19]:
```
 #Entering sunglasses in search box-
sunglasses=driver.find_element(By.CLASS_NAME,"Pke_EE")
sunglasses.send_keys('sunglasses')


```

In [20]:
```
# Closing the pop up-
pop_up=driver.find_element(By.XPATH,'//button[@class="_2iLD__"]')
pop_up.click()

```

In [21]:
```

#Clicking on Search Button-
search=driver.find_element(By.CLASS_NAME,"L0Z3Pu")
#search.click()
```

In [5]:
```

```

In [22]:
```
#Creating Empty List for different attributes-
brand=[]
prod_desciption=[]
price=[]
time.sleep(3)
```

In [25]:
```
start=0
end=3
for page in range(start,end):
    brands=driver.find_elements(By.XPATH,'//div[@class="_2WkVRV"]')
    for i in brands[0:100]:
        brand.append(i.text)
    Product_desc=driver.find_elements(By.XPATH,'//a[@class="IRpwTa"]')
    for i in Product_desc[0:100]:
        prod_desciption.append(i.text)
    prices=driver.find_elements(By.XPATH,'//div[@class="_30jeq3"]')
    for i in prices[0:100]:
        price.append(i.text)
    next_button=driver.find_elements(By.XPATH,'//a[@class="_1LKTO3"]')
```

In [26]:
```
#checking length for all attributes-
print(len(brand),len(prod_desciption),len(price))
```

```
120 120 120
```

In [27]:
```
# Creating Dataframe-
df=pd.DataFrame({'Brand':brand[0:100],'Prod_desc':prod_desciption[0:100],
df
```

Out[27]:

	Brand	Prod_desc	Price
0	VINCENT CHASE	UV Protection Wayfarer Sunglasses (59)	₹479
1	VINCENT CHASE	Polarized, UV Protection Round Sunglasses (50)	₹609
2	SRPM	UV Protection Wayfarer Sunglasses (50)	₹149
3	Elligator	UV Protection Aviator, Wayfarer Sunglasses (54)	₹159
4	ROADWAY	UV Protection Retro Square Sunglasses (Free Size)	₹394
...	...	...	...
95	ROADWAY	UV Protection Wayfarer Sunglasses (Free Size)	₹299
96	PROVOGUE	UV Protection, Gradient Rectangular, Wayfarer ...	₹475
97	ROYAL SON	UV Protection Rectangular, Retro Square Sungla...	₹449
98	Elligator	UV Protection Wayfarer Sunglasses (53)	₹149
99	Eyewearlabs	Polarized, UV Protection Retro Square Sunglass...	₹1,299

100 rows × 3 columns

In [28]:
```
1 driver.close()
```

Q5: Scrape 100 reviews data from flipkart.com for iphone11 phone. You have to go the link: https://www.flipkart.com/apple-iphone-11-black-64-gb/productreviews/itm4e5041ba101fd?pid=MOBFWQ6BXGJCEYNY&lid=LSTMOBFWQ6BXGJCEYNYZXSHRJ&market (https://www.flipkart.com/apple-iphone-11-black-64-gb/productreviews/itm4e5041ba101fd?pid=MOBFWQ6BXGJCEYNY&lid=LSTMOBFWQ6BXGJCEYNYZXSHRJ&market) place=FLIPKAR As shown in the above page you have to scrape the tick marked attributes. These are:

1. Rating
2. Review summary
3. Full review
4. You have to scrape this data for first 100reviews

In [1]:
```
1
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
Cell In[1], line 2
      1 # Activating the chrome browser
----> 2 driver=webdriver.Chrome()

NameError: name 'webdriver' is not defined
```

In [2]:
```
1 import selenium
2 import warnings
3 import pandas as pd
4 import time
5 from selenium import webdriver
6 from selenium.webdriver.common.by import By
7 warnings.filterwarnings('ignore')
```

In [3]:
```
1 # Activating the chrome browser
2 driver=webdriver.Chrome()
```

In [4]:
```
1 # Opening the homepage-
2 driver.get("https://www.flipkart.com/apple-iphone-11-black-64-gb-includes-
```

In [5]:
```
#checking length for all attributes-
urls=[]
short_review=[]
complete_review=[]
stars=[]
time.sleep(2)
```

In [6]:
```
#scraping 10 pages url
url_1 = driver.find_elements(By.XPATH,"//a[@class='ge-49M _2Kfbh8']")
for i in url_1:
    urls.append(i.get_attribute('href'))
url_2 = driver.find_elements(By.XPATH,"//a[@class='ge-49M']")
for i in url_2:
    urls.append(i.get_attribute('href'))
time.sleep(4)

```

In [7]:
```

for i in urls:
    driver.get(i)
    #for scrapping the number of stars
    for j in driver.find_elements(By.XPATH,"//div[@class='col _2wzgFH K0kl
        stars.append(j.text)
    #for scrapping the short review
    for k in driver.find_elements(By.XPATH,"//p[@class='_2-N8zT']"):
        short_review.append(k.text)
    #for scrapping the complete review
    for l in driver.find_elements(By.XPATH,"//div[@class='t-ZTKy']/div/div
        complete_review.append(l.text)
```

In [8]:
```
print(len(stars),len(short_review),len(complete_review))
```

```
100 100 100
```

In [9]:
```
df=pd.DataFrame({'Rating':stars,'Review_summary':short_review,'Full_summar
df
```

Out[9]:

	Rating	Review_summary	Full_summary
0	5	Classy product	Camera is awesome\nBest battery backup\nA perf...
1	5	Must buy!	It’s really awesome
2	5	Worth every penny	Feeling awesome after getting the delivery of ...
3	5	Best in the market!	Good Camera
4	5	Terrific	Very very good
...	...	...	...
95	5	Must buy!	My dream is complete just mind-blowing phone c...
96	5	Wonderful	Excellent Fabulous Adorable Iphone 11 Value fo...
97	5	Highly recommended	Thanks Flipkart For this amazing deal! I had a...
98	5	Best in the market!	Finally got it 😙
99	5	Classy product	Outstanding performance this phone

100 rows × 3 columns

In [10]:
```
driver.close()
```

Q6: Scrape data forfirst 100 sneakers you find whenyou visit flipkart.com and search for “sneakers” inthe search field. You have to scrape 3 attributes of each sneaker:

1. Brand
2. ProductDescription
3. Price As shown in the below image, you have to scrape the above attributes.

In [11]:
```
# Activating the chrome browser
driver=webdriver.Chrome()
```

In [12]:
```
# Opening the homepage-
driver.get("https://www.flipkart.com/")
```

In [18]:
```
# Entering sneakers in search box-
sneakers=driver.find_element(By.CLASS_NAME,"Pke_EE")
sneakers.send_keys('sneakers')

pop_up=driver.find_element(By.XPATH,'//button[@class="_2iLD__"]')
pop_up.click()

#Clicking on Search Button-
search=driver.find_element(By.CLASS_NAME,"L0Z3Pu")
search.click()
```

In [19]:
```
#Creating Empty List for different attributes-
brand=[]
description=[]
price=[]
```

In [20]:
```
# Scraping Data for different Attributes-
start=0
end=4
for page in range(start,end):#for loop for scrapping 4 page
    brands=driver.find_elements(By.CLASS_NAME,'_2WkVRV')#scraping brands r
    for i in brands:
        brand.append(i.text)#appending the text in Brand list
    prices=driver.find_elements(By.XPATH,"//div[@class='_30jeq3']")# scrap
    for i in prices:
        price.append(i.text)
    desc=driver.find_elements(By.XPATH,'//a[@class="IRpwTa" or @class="IRp
    for i in desc:
        description.append(i.text)

    nxt_button=driver.find_elements(By.XPATH,"//a[@class='_1LKTO3']")#scra
    try:
        driver.get(nxt_button[1].get_attribute('href'))#getting the link j
    except:
        driver.get(nxt_button[0].get_attribute('href'))
```

In [21]:
```
print(len(brand),len(description),len(price))
```

```
160 160 160
```

In [22]:
```
# Creating Dataframe-
df=pd.DataFrame({'Brand':brand,'Prod_desc':description,'Price':price})
df
```

Out[22]:

	Brand	Prod_desc	Price
0	Nobelite	Sneakers For Men	₹279
1	Krors	Mens Breathable Black Sneakers Running Shoes F...	₹439
2	BRUTON	Modern Trendy Shoes Sneakers For Men	₹229
3	aadi	Sneakers For Men	₹299
4	Layasa	Sneakers For Men	₹249
...	...	...	...
155	Magnolia	Trendy Stylish Chunky Sneakers For Men	₹340
156	PUMA	BMW MMS Drift Cat Delta Sneakers For Men	₹3,439
157	Rambo	SHOTGUN Sneakers For Men	₹699
158	RED TAPE	Sneaker Casual Shoes For Men | Soft Cushion In...	₹1,619
159	Sparx	SM 734 Sneakers For Men	₹715

160 rows × 3 columns

In [23]:
```
driver.close()
```

Q7: Go to webpage https://www.amazon.in/ (https://www.amazon.in/) Enter “Laptop” in the search field and then click the search icon. Then set CPU Type filter to “Intel Core i7” as shown in the below image: After setting the filters scrape first 10 laptops data. You have to scrape 3 attributes for each laptop:

1. Title
2. Ratings
3. Price

In [24]:
```
# Activating the chrome browser
driver=webdriver.Chrome()
```

In [25]:
```
# Opening the homepage-
driver.get("https://www.amazon.in/")
```

In [26]:
```
# Entering laptop in search box-
laptop=driver.find_element(By.XPATH,"//div[@class='nav-search-field ']//ir
laptop.send_keys('Laptop')

#Clicking on Search Button-
search=driver.find_element(By.XPATH,"/html/body/div[1]/header/div/div[1]/d
search.click()
```

In [28]:
```
# code to filter i7 CPu-
cpu_filter1=driver.find_element(By.XPATH,"/html/body/div[1]/div[1]/div[1],
cpu_filter1.click()
```

In [29]:
```
#Creating Empty List for different attributes-
Title=[]
price=[]
Ratings=[]
Rating=[]
```

In [30]:
```
start=0
end=1
for page in range(start,end):
    title=driver.find_elements(By.XPATH,"//h2[@class='a-size-mini a-spacir
    for i in title[0:10]:
        Title.append(i.text)
    prices=driver.find_elements(By.XPATH,'//span[@class="a-price-whole"]')
    for i in prices[0:10]:
        price.append(i.text)
```

In [31]:
```
titles_rat=driver.find_elements(By.XPATH,"//div[@class='a-row a-size-small
for i in titles_rat:
    Rating.append(i.get_attribute("aria-label"))
rating=[]
for i in range(0,len(Rating))[0:20]:
    if i == 0 or  i/2 == i//2:
        rating.append(Rating[i][0:3])
```

In [32]:
```
#checking length for all attributes-
print(len(Title),len(price),len(rating))
```

```
10 10 10
```

In [33]:
```
# Creating Dataframe-
df=pd.DataFrame({'Title':Title,'Price':price,'Ratings':rating})
df
```

Out[33]:

	Title	Price	Ratings
0	Dell 14 Laptop, Intel Core i5-1135G7 Processor...	36,490	3.6
1	Dell 15 Laptop, Intel Core i5-1135G7 Processor...	34,990	3.7
2	Acer Aspire Lite 11th Gen Intel Core i3 Premiu...	51,990	3.9
3	HP Laptop 15s, Intel Celeron N4500, 15.6 inch(...	52,199	3.4
4	Acer Aspire Lite AMD Ryzen 5 5500U Premium Thi...	50,490	3.9
5	Acer Aspire Lite 11th Gen Intel Core i3-1115G4...	27,990	3.9
6	Xiaomi Notebook Ultra Max 11th Gen Intel Core ...	28,990	3.9
7	HP Chromebook X360 Intel Celeron N4120 14 inch...	35,990	4.2
8	AXL VayuBook Laptop 14.1 Inch FHD IPS Display ...	17,990	3.9
9	HP Laptop 15s, 12th Gen Intel Core i3-1215U, 1...	26,990	3.2

In [34]:
```
driver.close()
```

Q8: Write a python program to scrape data for Top 1000 Quotes of All Time. The above task will be done in following steps:

1. First get the webpagehttps://www.azquotes.com/ (https://www.azquotes.com/)
2. Click on TopQuotes3. Than scrap a) Quote b) Author c) Type Of Quotes

In [35]:
```
# Let's import all required libraries
import selenium
from selenium import webdriver
import pandas as pd
import warnings
warnings.filterwarnings("ignore")
import time
from selenium.webdriver.common.by import By
from selenium.webdriver.common.action_chains import ActionChains
```

In [36]:
```
# Activating the chrome browser
driver=webdriver.Chrome()
```

In [37]:
```
# Opening the homepage-
driver.get('https://www.azquotes.com/')
```

In [38]:
```
# Click on Top Quotes
search_button=driver.find_element(By.XPATH,"/html/body/div[1]/div[1]/div[1
search_button.click()
```

In [39]:
```
# Details

Quote=[]
Author=[]
type_of_quote=[]

Details=[i.text.split('\n') for i in driver.find_elements(By.XPATH,"//div[

for i in Details:
    Quote.append(i[0])
    Author.append(i[1])
    type_of_quote.append(i[2])
```

In [40]:
```
data=pd.DataFrame({'Quotes':Quote,'Author':Author,'Type of quote':type_of_
data
```

Out[40]:

	Quotes	Author	Type of quote
0	The essence of strategy is choosing what not t...	Michael Porter	Essence, Deep Thought, Transcendentalism
1	One cannot and must not try to erase the past ...	Golda Meir	Inspiration, Past, Trying
2	Patriotism means to stand by the country. It d...	Theodore Roosevelt	Country, Peace, War
3	Death is something inevitable. When a man has ...	Nelson Mandela	Inspirational, Motivational, Death
4	You have to love a nation that celebrates its ...	Erma Bombeck	4th Of July, Food, Patriotic
...	...	...	...
95	When the going gets weird, the weird turn pro.	Hunter S. Thompson	Music, Sports, Hunting
96	When a train goes through a tunnel and it gets...	Corrie Ten Boom	Trust, Encouraging, Uplifting
97	If you think you are too small to make a diffe...	Dalai Lama	Inspirational, Funny, Change
98	God doesn't require us to succeed, he only req...	Mother Teresa	Success, God, Mother
99	Change your thoughts and you change your world.	Norman Vincent Peale	Inspirational, Motivational, Change

100 rows × 3 columns

In [41]:
```
driver.close()
```

Q9: Write a python program to display list of respected former Prime Ministers of India(i.e. Name, Born-Dead, Term of office, Remarks) from https://www.jagranjosh.com/ (https://www.jagranjosh.com/). This task will be done in following steps:

1. First get the webpagehttps://www.jagranjosh.com/ (https://www.jagranjosh.com/)
2. Then You have to click on the GK option
3. Then click on the List of all Prime Ministers of India
4. Then scrap the mentioned data and make theDataFrame.

In [42]:
```
# Activating the chrome browser
driver=webdriver.Chrome()
```

In [46]:
```
# Opening the homepage -
url = "https://www.jagranjosh.com/"
driver.get(url)
```

In [47]:
```
# do click in search button
gk_btn = driver.find_element(By.XPATH,'/html/body/div/header/nav/div/div/
gk_btn.click()
```

In [49]:
```
pm_btn = driver.find_element(By.XPATH,'//*[@id="popluarGK"]/ul/li[2]/a')
pm_btn.click()
```

In [50]:
```
Name= []
Born_Dead = []
Terms_of_office = []
Remarks = []
```

In [51]:
```
#name
names = driver.find_elements(By.XPATH,'//div[@class="table-box"]/table/tbo
for name in names:
    Name.append(name.text)


#borndead
born_dead = driver.find_elements(By.XPATH,'//div[@class="table-box"]/table
for i in born_dead:
    Born_Dead.append(i.text)


#terms_of_office
terms_of_office = driver.find_elements(By.XPATH,'//div[@class="table-box"]
for i in terms_of_office:
    Terms_of_office.append(i.text.replace('\n', ''))


#Remarks
remarks = driver.find_elements(By.XPATH,'//div[@class="table-box"]/table/t
for i in remarks:
    Remarks.append(i.text)
```

In [52]:
```
len(Name),len(Born_Dead),len(Terms_of_office),len(Remarks)
```

Out[52]: (19, 19, 19, 19)

In [53]:
```
df=pd.DataFrame({"Names":Name,"Born-Dead":Born_Dead,"Term of office":Terms
                 "Remarks":Remarks})
df
```

Out[53]:

	Names	Born-Dead	Term of office	Remarks
0	Jawahar Lal Nehru	(1889–1964)	15 August 1947 to 27 May 196416 years, 286 days	The first prime minister of India and the long...
1	Gulzarilal Nanda (Acting)	(1898-1998)	27 May 1964 to 9 June 1964,13 days	First acting PM of India
2	Lal Bahadur Shastri	(1904–1966)	9 June 1964 to 11 January 19661 year, 216 days	He has given the slogan of 'Jai Jawan Jai Kisa...
3	Gulzari Lal Nanda (Acting)	(1898-1998)	11 January 1966 to 24 January 196613 days	-
4	Indira Gandhi	(1917–1984)	24 January 1966 to 24 March 197711 years, 59 days	First female Prime Minister of India
5	Morarji Desai	(1896–1995)	24 March 1977 to 28 July 1979 2 year, 126 days	Oldest to become PM (81 years old) and first t...
6	Charan Singh	(1902–1987)	28 July 1979 to 14 January 1980170 days	Only PM who did not face the Parliament
7	Indira Gandhi	(1917–1984)	14 January 1980 to 31 October 19844 years, 291...	The first lady who served as PM for the second...
8	Rajiv Gandhi	(1944–1991)	31 October 1984 to 2 December 19895 years, 32 ...	Youngest to become PM (40 years old)
9	V. P. Singh	(1931–2008)	2 December 1989 to 10 November 1990343 days	First PM to step down after a vote of no confi...
10	Chandra Shekhar	(1927–2007)	10 November 1990 to 21 June 1991223 days	He belongs to Samajwadi Janata Party
11	P. V. Narasimha Rao	(1921–2004)	21 June 1991 to 16 May 19964 years, 330 days	First PM from South India
12	Atal Bihari Vajpayee	(1924-2018)	16 May 1996 to 1 June 199616 days	PM for shortest tenure
13	H. D. Deve Gowda	(born 1933)	1 June 1996 to 21 April 1997324 days	He belongs to Janata Dal
14	Inder Kumar Gujral	(1919–2012)	21 April 1997 to 19 March 1998 332 days	------
15	Atal Bihari Vajpayee	(1924-2018)	19 March 1998 to 22 May 2004 6 years, 64 days	The first non-congress PM who completed a ful...
16	Manmohan Singh	(born 1932)	22 May 2004 to 26 May 2014 10 years, 4 days	First Sikh PM
17	Narendra Modi	(born 1950)	26 May 2014 - 2019	4th Prime Minister of India who served two con...
18	Narendra Modi	(born 1950)	30 May 2019- Incumbent	First non-congress PM with two consecutive ten...

In [54]:
```
1 driver.close()
```

In [ ]:
```
1
```

In [ ]:
```
1
```