In [494…
```
import selenium
from selenium import webdriver
import pandas as pd
from selenium.webdriver.common.by import By
import warnings
warnings.filterwarnings('ignore')
import time
```

In [85]:
```
driver=webdriver.Chrome()
```

# 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.

In [ ]:
```
#Q1(1). First get the webpage https://www.shine.com/

driver.get('https://www.shine.com/')
```

In [ ]:
```
#Q1(2) Enter “Data Analyst” in “Job title, Skills” field and enter “Bangalore” in “ent

job_title=driver.find_element(By.CLASS_NAME,'form-control')
job_title.send_keys('Data Analyst')

location= driver.find_element(By.XPATH, '/html/body/div[1]/div[4]/div/div[2]/div[2]/di
location.send_keys('Bangalore')
```

In [ ]:
```
#Q1(3).Then click the search button

search=driver.find_element(By.CLASS_NAME,'searchForm_btnWrap_advance__VYBHN')
search.click()
```

In [266…
```
#Q1(4) Then scrape the data for the first 10 jobs results you get.

company_location=[]
location=driver.find_elements(By.XPATH,'//div[@class=" jobCard_jobCard_lists_item__YxF

for x in location:
    company_location.append(x.text)

company_name=[]
company=driver.find_elements(By.CLASS_NAME,'jobCard_jobCard_cName__mYnow')[0:10]

for x in company:
    company_name.append(x.text)

experience_required = []
experience_elements = driver.find_elements(By.XPATH, '//div[contains(@class, "jobCard_
```

```
for x in experience_elements:
    experience_required.append(x.text)
```

In [298…

```
#Q1(5)  Finally create a dataframe of the scraped data.

import pandas as pd
df=pd.DataFrame({'company_name':company_name,'company_location':company_location,'expe
df
```

Out[298]:

| | company_name | company_location | experience_required |
|---|---|---|---|
| **0** | quiscon biotech | Bangalore\n+17 | 0 to 4 Yrs |
| **1** | nina s hr consultancy | Bangalore | 5 to 10 Yrs |
| **2** | nina s hr consultancy | Bangalore\n+14 | 0 to 4 Yrs |
| **3** | skyleaf consultants | Bangalore | 4 to 6 Yrs |
| **4** | quiscon biotech | Bangalore\n+8 | 4 to 6 Yrs |
| **5** | acme services private limited | Bangalore | 6 to 9 Yrs |
| **6** | seven consultancy | Bangalore | 3 to 6 Yrs |
| **7** | alpine manpower services | Bangalore\n+15 | 0 to 4 Yrs |
| **8** | alpine manpower services | Bangalore\n+8 | 4 to 6 Yrs |
| **9** | nina s hr consultancy | Bangalore\n+6 | 0 to 3 Yrs |

# 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:

In [214…

```
driver=webdriver.Chrome()
```

In [215…

```
#Q2 1) First get the webpage https://www.shine.com/

driver.get('https://www.shine.com/')
```

In [218…

```
#Q2(2) Enter “Data scientist” in “Job title, Skills” field and enter “Bangalore” in “e

job_title=driver.find_element(By.CLASS_NAME,'form-control')
job_title.send_keys('Data Scientist')

location=driver.find_element(By.XPATH,'/html/body/div[1]/div[4]/div/div[2]/div[2]/div/
location.send_keys('bangalore')
```

In [227…
```
#Q2(3).Then click the search button

submit=driver.find_element(By.XPATH,'/html/body/div[1]/div[4]/div/div[2]/div[2]/div/fo
submit.click()
```

In [261…
```
#Q2(4) Then scrape the data for the first 10 jobs results you get.

job_location=[]
location=driver.find_elements(By.XPATH,'//div[@class=" jobCard_jobCard_lists_item__YxR

for x in location:
    job_location.append(x.text)


company_name=[]
company=driver.find_elements(By.XPATH,'//div[@class="jobCard_jobCard_cName__mYnow"]')[
for x in company:
    company_name.append(x.text)
```

In [297…
```
#Q2(5)  Finally create a dataframe of the scraped data.

df=pd.DataFrame({'company_name':company_name,'job_location':job_location})
df
```

Out[297]:

| | company_name | job_location |
|---|---|---|
| 0 | quiscon biotech | Delhi\n+6 |
| 1 | nina s hr consultancy | Delhi |
| 2 | nina s hr consultancy | Delhi |
| 3 | skyleaf consultants | Delhi |
| 4 | quiscon biotech | Delhi\n+6 |
| 5 | acme services private limited | Delhi\n+4 |
| 6 | seven consultancy | Delhi |
| 7 | alpine manpower services | Delhi |
| 8 | alpine manpower services | Delhi |
| 9 | nina s hr consultancy | Delhi |

# 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

In [267…
```
driver=webdriver.Chrome()
```

In [268…
```
#1. first get the web page https://www.shine.com/
driver.get('https://www.shine.com/')
```

In [269…
```
#2. Enter “Data Scientist” in “Skill, Designations, and Companies” field.
job_title=driver.find_element(By.XPATH,'/html/body/div[1]/div[4]/div/div[2]/div[2]/div
job_title.send_keys('Data Scientist')
```

In [271…
```
#3. Then click the search button
submit=driver.find_element(By.XPATH,'/html/body/div[1]/div[4]/div/div[2]/div[2]/div/fc
submit.click()
```

In [272…
```
#4. Then apply the location filter by checking the respective boxes
location=driver.find_element(By.XPATH,'/html/body/div[1]/div[1]/div[4]/div/div[1]/div/
location.click()

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

In [277…
```
#4. Then apply the salary filter by checking the respective boxes
salary=driver.find_element(By.XPATH,'/html/body/div[1]/div[1]/div[4]/div/div[1]/div/di
salary.click()

select_salary=driver.find_element(By.XPATH,'/html/body/div[1]/div[1]/div[4]/div/div[1]
select_salary.click()

# click on show result
show_result=driver.find_element(By.XPATH,'/html/body/div[1]/div[1]/div[4]/div/div[1]/c
show_result.click()
```

In [293…
```
#5. Then scrape the data for the first 10 jobs results you get.(job-location, company

company_name=[]
company=driver.find_elements(By.XPATH,'//div[@class="jobCard_jobCard_cName__mYnow"]')[
for x in company:
    company_name.append(x.text)


job_location=[]
location=driver.find_elements(By.XPATH,'//div[@class=" jobCard_jobCard_lists_item__YxR
for x in location:
    job_location.append(x.text)


experience_req=[]
experience=driver.find_elements(By.XPATH,'//div[@class=" jobCard_jobCard_lists_item__Y
for x in experience:
    experience_req.append(x.text)
```

In [ ]:
```
#6. Finally create a dataframe of the scraped data
df=pd.DataFrame({'company_name':company_name,'job_location':job_location,'experience_r
df
```

# Q4: Scrape data of first 100 sunglasses listings on flipkart.com. You have to scrape four attributes: 1. Brand 2. Product Description 3. Price

In [300…
```
driver=webdriver.Chrome()
```

In [301…
```
#1. Go to Flipkart webpage by url :https://www.flipkart.com/
driver.get('http://www.flipkart.com/')
```

In [305…
```
#2. Enter “sunglasses” in the search fieldwhere “search for products, brands and more”
search=driver.find_element(By.XPATH,'/html/body/div[1]/div/div[1]/div[1]/div[2]/div[2]
search.send_keys('Sunglasses')

submit_search=driver.find_element(By.XPATH,'/html/body/div[1]/div/div[1]/div[1]/div[2]
submit_search.click()
```

In [347…
```
#3. Now Scraping data of first 100 sunglasses listings on flipkart.com including data

brand_name=[]
item_price=[]
start=0
end=3
for page in range (start,end):

    brand=driver.find_elements(By.XPATH,'//div[@class="_2WkVRV"]')
    for x in brand:
        brand_name.append(x.text)

    price=driver.find_elements(By.XPATH,'//div[@class="_30jeq3"]')
    for x in price:
        item_price.append(x.text)


    next_page=driver.find_element(By.XPATH,'/html/body/div/div/div[3]/div[1]/div[2]/di
    next_page.click()
    time.sleep(3)
```

In [349…
```
#4. Finally create a dataframe of the scraped data
df=pd.DataFrame({'brand_name':brand_name,'item_price':item_price})
df[0:100]
```

Out[349]:

| | brand_name | item_price |
|---|---|---|
| 0 | OAKLEY | ₹7,352 |
| 1 | OAKLEY | ₹7,752 |
| 2 | SRPM | ₹204 |
| 3 | PIRASO | ₹239 |
| 4 | Elligator | ₹179 |
| ... | ... | ... |
| 95 | ROYAL SON | ₹664 |
| 96 | ROYAL SON | ₹854 |
| 97 | GANSTA | ₹228 |
| 98 | ROYAL SON | ₹899 |
| 99 | Ray-Ban | ₹10,616 |

100 rows × 2 columns

# Q5: Scrape 100 reviews data from flipkart.com for iphone11 phone.

In [350…
```
driver=webdriver.Chrome()
```

In [353…
```
#1. Go to Flipkart webpage
driver.get('https://www.flipkart.com/apple-iphone-11-black-64-gb/p/itm4e5041ba101fd?pi
```

In [387…
```
#2. As shown in the page you have to scrape the tick marked attributes. These are: 1.

product_rating=[]
review_summary=[]
Full_review=[]

start=0
end=3
for page in range (start,end):


    rating=driver.find_elements(By.XPATH,'//div[@class="_3LWZlK _1BLPMq"]')
    for x in rating:
        product_rating.append(x.text)


    Reviews=driver.find_elements(By.XPATH,'//p[@class="_2-N8zT"]')
    for x in Reviews:
        review_summary.append(x.text)


    Fullreview=driver.find_elements(By.XPATH,'//div[@class="t-ZTKy"]')
    for x in Fullreview:
```

```
        Full_review.append(x.text)

    next_button=driver.find_element(By.XPATH,'/html/body/div[1]/div/div[3]/div/div/div
    next_button.click()
    time.sleep(3)
```

In [388…

```
#4. Finally create a dataframe of the scraped data
df=pd.DataFrame({'product_rating':product_rating,'review_summary':review_summary,'Full
df
```

Out[388]:

| | product_rating | review_summary | Full_review |
|---|---|---|---|
| **0** | 5 | Classy product | Photos super |
| **1** | 5 | Terrific | Very very good |
| **2** | 5 | Terrific purchase | Value for money 🤩 |
| **3** | 5 | Classy product | Camera is awesome\nBest battery backup\nA perf... |
| **4** | 5 | Wonderful | This is amazing at all |
| **5** | 5 | Just wow! | Perfect Product!! |
| **6** | 5 | Worth every penny | Feeling awesome after getting the delivery of ... |
| **7** | 5 | Perfect product! | V Good all |
| **8** | 5 | Best in the market! | Good Camera |
| **9** | 5 | Fabulous! | Super🔥 and good performance 👌❤️ |
| **10** | 5 | Must buy! | It's really awesome |
| **11** | 5 | Great product | Purple is best |
| **12** | 4 | Worth the money | Camera is just wow 👌👌 |
| **13** | 5 | Must buy! | Go for iPhone 11 , if confused between iPhone ... |
| **14** | 5 | Brilliant | Excellent Phone. |
| **15** | 5 | Terrific purchase | Value for money 🖤🖤 |
| **16** | 5 | Brilliant | very good camera quality |
| **17** | 5 | Fabulous! | It's very good battery life and display and vi... |
| **18** | 5 | Excellent | NYC |
| **19** | 5 | Best in the market! | Damn this phone is a blast . Upgraded from and... |
| **20** | 5 | Best in the market! | Such an awesome experience with iPhone 11 awes... |
| **21** | 5 | Awesome | Awesome Phone. Battery backup top-notch... |
| **22** | 5 | Worth every penny | It is better to buy iPhone 11 over iPhone 12 i... |
| **23** | 5 | Wonderful | Excellent Fabulous Adorable Iphone 11 Value fo... |
| **24** | 5 | Must buy! | happy ❤️ |
| **25** | 5 | Brilliant | Best phone |
| **26** | 5 | Brilliant | Perfect iPhone on this budget!! Camera and the... |
| **27** | 5 | Perfect product! | Battery backup is extraordinary, camera is dec... |
| **28** | 5 | Worth every penny | iPhone is delivered on time. Display is great ... |
| **29** | 5 | Classy product | Outstanding performance this phone |

# Q6: Scrape data forfirst 100 sneakers you find whenyou visit flipkart.com and search

# for “sneakers” inthe search field.

In [390…
```
driver=webdriver.Chrome()
```

In [391…
```
#1. Go to Flipkart webpage
driver.get('https://www.flipkart.com/search?q=sneakers&otracker=search&otracker1=searc
```

In [407…
```
#2. Scraping attributes of each sneaker: Brand and Price
brand_name=[]
Product_price=[]

#3. for page range
start=0
end=3
for page in range (start,end):

#4. scrape first 100 brand and price of shoes
    brand=driver.find_elements(By.XPATH,'//div[@class="_2WkVRV"]')
    for x in brand:
        brand_name.append(x.text)

    price=driver.find_elements(By.XPATH,'//div[@class="_30jeq3"]')
    for x in price:
        Product_price.append(x.text)

    next_button=driver.find_element(By.XPATH,'/html/body/div/div/div[3]/div[1]/div[2]/
    next_button.click()
    time.sleep(3)
```

In [410…
```
#5. Finally create a dataframe of the scraped data
df=pd.DataFrame({'brand_name':brand_name,'Product_price':Product_price})
df[0:100]
```

Out[410]:

| | brand_name | Product_price |
|---|---|---|
| 0 | Layasa | ₹399 |
| 1 | Sparx | ₹659 |
| 2 | BRUTON | ₹299 |
| 3 | Nobelite | ₹299 |
| 4 | DUCATI | ₹287 |
| ... | ... | ... |
| 95 | Peelu | ₹719 |
| 96 | Layasa | ₹399 |
| 97 | New Balance | ₹7,699 |
| 98 | New Balance | ₹5,000 |
| 99 | Skechers | ₹4,649 |

100 rows × 2 columns

# Q7: Go to webpage 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:

In [413...
```
driver=webdriver.Chrome()
```

In [415...
```
#1. Go to amazon webpage
driver.get('https://www.amazon.in/')
```

In [421...
```
#2. type 'laptop' on searchbar
search=driver.find_element(By.XPATH,'/html/body/div[1]/header/div/div[1]/div[2]/div/fo
search.send_keys('laptop')

#3. clicking on submit button
submit=driver.find_element(By.XPATH,'/html/body/div[1]/header/div/div[1]/div[2]/div/fo
submit.click()

#4. clicked on i7 filter
submit_i7=driver.find_element(By.XPATH,'/html/body/div[1]/div[2]/div[1]/div[2]/div/div
submit_i7.click()
```

In [431...
```
#5. scrape first 10 laptops titles
laptop_title=[]
title=driver.find_elements(By.XPATH,'//h2[@class="a-size-mini a-spacing-none a-color-b
for x in title:
    laptop_title.append(x.text)
```

```
#6. scrape first 10 laptops ratings
laptop_ratings=[]
ratings=driver.find_elements(By.XPATH,'//span[@class="a-size-base puis-normal-weight-t
for x in ratings:
    laptop_ratings.append(x.text)

#7. scrape first 10 laptops price
laptop_price=[]
price=driver.find_elements(By.XPATH,'//span[@class="a-price-whole"]')[0:10]
for x in price:
    laptop_price.append(x.text)
```

In [432…

```
#8. Finally create a dataframe of the scraped data
df=pd.DataFrame({'laptop_title':laptop_title,'laptop_ratings':laptop_ratings,'laptop_p
df
```

Out[432]:

| | laptop_title | laptop_ratings | laptop_price |
|---|---|---|---|
| **0** | Acer Nitro 5 12th Gen Intel Core i7-12650H Gam... | 4.1 | 1,04,990 |
| **1** | Lenovo [SmartChoice] IdeaPad Slim 3 Intel Core... | 3.3 | 62,990 |
| **2** | ASUS Vivobook 15, Intel Core i7-12650H 12th Ge... | 4.0 | 59,990 |
| **3** | Dell Inspiron 5430 13th Gen Laptop, Intel i7-1... | 3.2 | 86,249 |
| **4** | MSI GF63 Thin, Intel Core i7-11800H, 40CM FHD ... | 4.5 | 70,990 |
| **5** | Lenovo [SmartChoice] IdeaPad Slim 3 Intel Core... | 3.3 | 62,990 |
| **6** | ASUS Creator Series Vivobook 14X OLED (2023), ... | 3.5 | 85,990 |
| **7** | Dell Inspiron 5630 13th Gen Laptop, Intel Core... | 4.5 | 89,990 |
| **8** | ASUS TUF Gaming F15 (2023) 90WHr Battery, Inte... | 4.2 | 1,15,990 |
| **9** | ASUS Creator Series Vivobook 14X OLED (2023), ... | 4.3 | 85,990 |

# Q8: Write a python program to scrape data for Top 1000 Quotes of All Time

In [436…

```
driver=webdriver.Chrome()
```

In [437…

```
#1. First get the webpagehttps://www.azquotes.com
driver.get('http://www.azquotes.com/')
```

In [439…

```
#2. Click on TopQuotes
click_on_top_quotes=driver.find_element(By.XPATH,'/html/body/div[1]/div[1]/div[1]/div/
click_on_top_quotes.click()
```

In [491…

```
#3. Than scrap a) Quote b) Author c) Type Of Quotes
start=0
end=3
for page in range (start,end):
    quotes_types=[]
    types=driver.find_elements(By.XPATH,'//div[@class="tags"]')[0:200]
    for x in types:
```

```
        quotes_types.append(x.text)

    top_quotes=[]
    quotes=driver.find_elements(By.XPATH,'//a[@class="title"]')[0:200]
    for x in quotes:
        top_quotes.append(x.text)

    Author_list=[]
    types=driver.find_elements(By.XPATH,'//div[@class="author"]')[0:200]
    for x in types:
        Author_list.append(x.text)

    next_click=driver.find_element(By.XPATH,'/html/body/div[1]/div[2]/div/div/div/div[
    next_click.click()
    time.sleep(3)
```

In [492…

```
#4. Finally create a dataframe of the scraped data
df=pd.DataFrame({'top_quotes':top_quotes,'Author_list':Author_list,'quotes_types':quot
df
```

Out[492]:

| | top_quotes | Author_list | quotes_types |
|---|---|---|---|
| **0** | A man is not old until regrets take the place ... | John Barrymore | Love, Inspirational, Life |
| **1** | A part of kindness consists in loving people m... | Joseph Joubert | Love, Kindness, Women |
| **2** | When you start to develop your powers of empat... | Susan Sarandon | Kindness, Power, Imagination |
| **3** | Judge a man by his questions rather than his a... | Voltaire | Wisdom, Clever, Men |
| **4** | If you make a sale, you can make a living. If ... | Jim Rohn | Inspirational, Business, Investment |
| **...** | ... | ... | ... |
| **95** | All things are difficult before they are easy. | Thomas Fuller | Inspirational, Motivational, Positive |
| **96** | People ask the difference between a leader and... | Theodore Roosevelt | Leadership, Differences, People |
| **97** | The thing that is really hard, and really amaz... | Anna Quindlen | Confidence, Letting Go, Being Yourself |
| **98** | Love is the ability and willingness to allow t... | Wayne Dyer | Love, Anxiety, Self Improvement |
| **99** | Sooner or later, those who win are those who t... | Paul Tournier | Positive, Sports, Confidence |

100 rows × 3 columns

# 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/.

In [495…
```
driver=webdriver.Chrome()
```

In [496…
```
#1. First get the webpagehttps://www.jagranjosh.com/
driver.get('http://www.jagranjosh.com/')
```

In [497…
```
#2. Then You have to click on the GK option
click_gk=driver.find_element(By.XPATH,'/html/body/div/header/nav/div/div/div[3]/ul/li[
click_gk.click()
```

In [498…
```
#3. Then click on the List of all Prime Ministers of India
click_PM_list=driver.find_element(By.XPATH,'/html/body/div[1]/div/div/div[2]/div/div[1
click_PM_list.click()
```

In [ ]:
```
# no class name XPATH FOUND FOR ANY Name, Born-Dead,Term of office, Remarks
```

# Q10: Write a python program to display list of 50 Most expensive cars in the world (i.e.Car name and Price) from https://www.motor1.com/

In [544…
```
driver=webdriver.Chrome()
```

In [545…
```
#1. First get the webpage https://www.motor1.com/
driver.get('https://www.motor1.com/')
```

In [547…
```
#2. Then You have to type in the search bar '50 most expensive cars'
type_50_most_expensive_cars=driver.find_element(By.XPATH,'/html/body/div[10]/div[2]/di
type_50_most_expensive_cars.send_keys('50 most expensive cars')
```

In [548…
```
#3. click on submit
submit=driver.find_element(By.XPATH,'/html/body/div[10]/div[2]/div/div/div[3]/div/div/
submit.click()
```

In [549…
```
#4. click on 50 most expensive carsin the world
click_tab=driver.find_element(By.XPATH,'/html/body/div[10]/div[9]/div/div[1]/div/div/d
click_tab.click()
```

In [552…
```
#5. scraping car name(but price was not scrapable)
car_name=[]
name=driver.find_elements(By.XPATH,'//h3[@class="subheader"]')
for x in name:
    car_name.append(x.text)
```

In [557…
```
#6. Finally create a dataframe of the scraped data
df=pd.DataFrame({'car_name':car_name})
df
```

Out[557]:

| | car_name |
|---|---|
| **0** | Aston Martin Valour |
| **1** | McLaren Elva |
| **2** | Czinger 21C |
| **3** | Ferrari Monza |
| **4** | Gordon Murray T.33 |
| **5** | Koenigsegg Gemera |
| **6** | Zenvo TSR-S |
| **7** | Hennessey Venom F5 |
| **8** | Bentley Bacalar |
| **9** | Hispano Suiza Carmen Boulogne |
| **10** | Bentley Mulliner Batur |
| **11** | Deus Vayanne |
| **12** | SSC Tuatara |
| **13** | Lotus Evija |
| **14** | Aston Martin Vulcan |
| **15** | Delage D12 |
| **16** | Ferrari Daytona SP3 |
| **17** | McLaren Speedtail |
| **18** | Rimac Nevera |
| **19** | Pagani Utopia |
| **20** | Pininfarina Battista |
| **21** | Gordon Murray T.50 |
| **22** | Lamborghini Countach |
| **23** | Mercedes-AMG Project One |
| **24** | Zenvo Aurora |
| **25** | Aston Martin Victor |
| **26** | Hennessey Venom F5 Roadster |
| **27** | Koenigsegg Jesko |
| **28** | Aston Martin Valkyrie |
| **29** | W Motors Lykan Hypersport |
| **30** | McLaren Solus |
| **31** | Lamborghini Sian |
| **32** | Koenigsegg CC850 |

| | car_name |
|---|---|
| **33** | Bugatti Chiron Super Sport 300+ |
| **34** | Lamborghini Veneno |
| **35** | Bugatti Bolide |
| **36** | Pininfarina B95 Speedster |
| **37** | Bugatti Mistral |
| **38** | Pagani Huayra Imola |
| **39** | Bugatti Divo |
| **40** | SP Automotive Chaos |
| **41** | Pagani Codalunga |
| **42** | 777 Hypercar |
| **43** | Mercedes-Maybach Exelero |
| **44** | Bugatti Centodieci |
| **45** | Bugatti Chiron Profilée |
| **46** | Rolls-Royce Sweptail |
| **47** | Bugatti La Voiture Noire |
| **48** | Rolls-Royce Boat Tail* |
| **49** | Rolls-Royce La Rose Noire Droptail |
| **50** | Most Expensive Cars In The World |

# ----------END----------

In [ ]: