In [1]:
```
!pip install selenium
```
```
Requirement already satisfied: selenium in c:\users\asus\anaconda3\lib\site-packages (4.8.0)
Requirement already satisfied: trio~=0.17 in c:\users\asus\anaconda3\lib\site-packages (from selenium) (0.22.0)
Requirement already satisfied: trio-websocket~=0.9 in c:\users\asus\anaconda3\lib\site-packages (from selenium) (0.9.2)
Requirement already satisfied: certifi>=2021.10.8 in c:\users\asus\anaconda3\lib\site-packages (from selenium) (2021.10.8)
Requirement already satisfied: urllib3[socks]~=1.26 in c:\users\asus\anaconda3\lib\site-packages (from selenium) (1.26.9)
Requirement already satisfied: outcome in c:\users\asus\anaconda3\lib\site-packages (from trio~=0.17->selenium) (1.2.0)
Requirement already satisfied: async-generator>=1.9 in c:\users\asus\anaconda3\lib\site-packages (from trio~=0.17->selenium) (1.10)
Requirement already satisfied: idna in c:\users\asus\anaconda3\lib\site-packages (from trio~=0.17->selenium) (3.3)
Requirement already satisfied: exceptiongroup>=1.0.0rc9 in c:\users\asus\anaconda3\lib\site-packages (from trio~=0.17->selenium) (1.1.0)
Requirement already satisfied: sortedcontainers in c:\users\asus\anaconda3\lib\site-packages (from trio~=0.17->selenium) (2.4.0)
Requirement already satisfied: cffi>=1.14 in c:\users\asus\anaconda3\lib\site-packages (from trio~=0.17->selenium) (1.15.0)
Requirement already satisfied: attrs>=19.2.0 in c:\users\asus\anaconda3\lib\site-packages (from trio~=0.17->selenium) (21.4.0)
Requirement already satisfied: sniffio in c:\users\asus\anaconda3\lib\site-packages (from trio~=0.17->selenium) (1.2.0)
Requirement already satisfied: pycparser in c:\users\asus\anaconda3\lib\site-packages (from cffi>=1.14->trio~=0.17->selenium) (2.21)
Requirement already satisfied: wsproto>=0.14 in c:\users\asus\anaconda3\lib\site-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\asus\anaconda3\lib\site-packages (from urllib3[socks]~=1.26->selenium) (1.7.1)
Requirement already satisfied: h11<1,>=0.9.0 in c:\users\asus\anaconda3\lib\site-packages (from wsproto>=0.14->trio-websocket~=0.9->selenium) (0.14.0)
```

In [656…
```
import pandas as pd
import selenium
from selenium import webdriver
import warnings
warnings.filterwarnings('ignore')
from selenium.webdriver.common.by import By
import requests
import time
from selenium.common.exceptions import NoSuchElementException, StaleElementReferenceException, ElementNotInteractableException
import re
from selenium.common.exceptions import ElementNotInteractableException
from bs4 import BeautifulSoup
import requests
from selenium.webdriver.common.keys import Keys
```

In [ ]:
```
#1. Write a python program which searches all the product under a particular product from www.amazon.in. The
#product to be searched will be taken as input from user. For e.g. If user input is ‘guitar’. Then search for
#guitars
```

In [3]:
```
driver=webdriver.Chrome(r"chromedriver.exe")
```

In [4]:
```
driver.get("https://www.amazon.in/")
```

In [6]:
```

val=input("enter your item: ")
```
```
enter your item: guitars
```

In [ ]:
```
#2. In the above question, now scrape the following details of each product listed in first 3 pages of your search
#results and save it in a data frame and csv. In case if any product has less than 3 pages in search results then
#scrape all the products available under that product name. Details to be scraped are: "Brand
#Name", "Name of the Product", "Price", "Return/Exchange", "Expected Delivery", "Availability" and
#“Product URL”. In case, if any of the details are missing for any of the product then replace it by “-“.
```

In [ ]:
```
driver=webdriver.Chrome(r"chromedriver.exe")
driver.get("https://www.amazon.in/")
```

In [7]:
```
product=driver.find_element(By.ID,"twotabsearchtextbox")

product.send_keys('guitar')
```

In [8]:
```
search=driver.find_element(By.XPATH,"/html/body/div[1]/header/div/div[1]/div[2]/div/form/div[3]/div/span")
search.click()
```

In [16]:
```
product_urls=[]
start=0
end=3
for page in range(start,end):
    url=driver.find_elements(By.XPATH,'//a[@class="a-link-normal s-underline-text s-underline-link-text s-link-style a-text-normal"]')
    for i in url[0:30]:
        product_urls.append(i.get_attribute("href"))
    next_button=driver.find_elements(By.XPATH,'//a[@class="s-pagination-item s-pagination-next s-pagination-button s-pagination-separator"]')

    time.sleep(3)
```

In [35]:
```
Brand=[]
for url in product_urls:
    driver.get(url)
    time.sleep(2)

    try:
        brand=driver.find_element(By.XPATH,'//*[@id="productOverview_feature_div"]/div/table/tbody/tr[1]/td[2]/span')
        Brand.append(brand.text)
    except NoSuchElementException:
        Brand.append("-")
```

In [39]:
```
Price=[]
for url in product_urls:
    driver.get(url)
    time.sleep(2)

    try:
        price=driver.find_element(By.XPATH,'//span[@class="a-price-whole"]')
        Price.append(price.text)
    except NoSuchElementException:
        Price.append("-")
```

In [42]:
```
Return_Exchange=[]
for url in product_urls:
    driver.get(url)
    time.sleep(3)

    try:
        return_exchange=driver.find_element(By.XPATH,'//span[@class="a-declarative"]/div/a')
        Return_Exchange.append(return_exchange.text)
    except NoSuchElementException:
        Return_Exchange("-")
```

In [47]:
```
Expected_delivery=[]
for i in product_urls:
    driver.get(url)
    time.sleep(3)

    try:
        expected_delivery=driver.find_element(By.XPATH,'//span[@class="a-text-bold"]')
        Expected_delivery.append(expected_delivery.text)
    except NoSuchElementException:
        Expected_delivery("-")
```

In [55]:
```
Availability=[]
for i in product_urls:
    driver.get(url)
    time.sleep(3)

    try:
        availability=driver.find_element(By.XPATH,'//span[@class="a-size-medium a-color-success"]')
        Availability.append(availability.text)
    except NoSuchElementException:
        Availability("-")
```

In [57]:
```
print(len(Brand),len(Price),len(Return_Exchange),len(Expected_delivery),len(Availability),len(product_urls))
```
```
90 90 90 90 90 90
```

In [58]:
```
df=pd.DataFrame({'BRAND':Brand,'PRICE':Price,'RETURN_EXCHANGE':Return_Exchange,'EXPECTED_DELIVERY':Expected_delivery,'AVAILABILITY':Availability,'PRODUCT_URL'

df
```

Out[58]:

	BRAND	PRICE	RETURN_EXCHANGE	EXPECTED_DELIVERY	AVAILABILITY	PRODUCT_URL
0	Kadence	6,499	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/sspa/click?ie=UTF8&spc=M...
1	Kadence	11,999	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/sspa/click?ie=UTF8&spc=M...
2	Kadence	5,003	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/sspa/click?ie=UTF8&spc=M...
3	VAULT	4,084	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/sspa/click?ie=UTF8&spc=M...
4	JUAREZ	1,999	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/Juarez-Acoustic-Cutaway-...
...	...	...	...	...	...	...
85	Henrix	2,849	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/Henrix-Acoustic-Guitar-D...
86	YAMAHA	7,899	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/Yamaha-Acoustic-Guitar-N...
87	REVEL	2,399	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/Acoustic-Plectrums-Begin...
88	Giusen	3,791	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/sspa/click?ie=UTF8&spc=M...
89	Guitar Bro	8,999	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/sspa/click?ie=UTF8&spc=M...

90 rows × 6 columns

In [60]:
```
df.to_csv('guitars.csv')
```

In [62]:
```
df=pd.read_csv('guitars.csv')
```

In [63]:
```
df
```

Out[63]:

	Unnamed: 0	BRAND	PRICE	RETURN_EXCHANGE	EXPECTED_DELIVERY	AVAILABILITY	PRODUCT_URL
0	0	Kadence	6,499	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/sspa/click?ie=UTF8&spc=M...
1	1	Kadence	11,999	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/sspa/click?ie=UTF8&spc=M...
2	2	Kadence	5,003	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/sspa/click?ie=UTF8&spc=M...
3	3	VAULT	4,084	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/sspa/click?ie=UTF8&spc=M...
4	4	JUAREZ	1,999	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/Juarez-Acoustic-Cutaway-...
...	...	...	...	...	...	...	...
85	85	Henrix	2,849	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/Henrix-Acoustic-Guitar-D...
86	86	YAMAHA	7,899	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/Yamaha-Acoustic-Guitar-N...
87	87	REVEL	2,399	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/Acoustic-Plectrums-Begin...
88	88	Giusen	3,791	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/sspa/click?ie=UTF8&spc=M...
89	89	Guitar Bro	8,999	7 days Replacement	March 10 - 13	In stock	https://www.amazon.in/sspa/click?ie=UTF8&spc=M...

90 rows × 7 columns

In [ ]:
```
#3. Write a python program to access the search bar and search button on images.google.com and scrape 10
#images each for keywords ‘fruits’, ‘cars’ and ‘Machine Learning’, ‘Guitar’, ‘Cakes’.
```

In [64]:
```
driver=webdriver.Chrome(r"chromedriver.exe")
```

In [65]:
```
driver.get("https://images.google.com/")
```

In [67]:
```
keywords=driver.find_element(By.CLASS_NAME,"gLFyf")
keywords.send_keys("cake")
```

In [68]:
```
search=driver.find_element(By.XPATH,"/html/body/div[1]/div[3]/form/div[1]/div[1]/div[1]/button/div")
search.click()
```

In [75]:
```
for _ in range(10):
    driver.execute_script("window.scrollBy(0,100)")

images=driver.find_elements(By.XPATH,'//img[@class="rg_i Q4LuWd"]')

img_urls=[]
img_data=[]
for image in images:
    source=image.get_attribute('src')
    if source is not None:
        if(source[0:4]=='http'):
            img_urls.append(source)

for i in range(len(img_urls)):
    if i>10:
        breakBy.XPATH,

    print("Downloading {0} of {1} images" .format(i,10))
    response=requests.get(img_urls[i])
    file=open(r"C:\Users\ASUS\Downloads\Fliprobo"+str(i)+".jpg", "wb")
    file.write(response.content)
```
```
Downloading 0 of 10 images
Downloading 1 of 10 images
Downloading 2 of 10 images
Downloading 3 of 10 images
Downloading 4 of 10 images
Downloading 5 of 10 images
Downloading 6 of 10 images
Downloading 7 of 10 images
Downloading 8 of 10 images
Downloading 9 of 10 images
Downloading 10 of 10 images
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
Input In [75], in <cell line: 14>()
     14 for i in range(len(img_urls)):
     15     if i>10:
---> 16         breakBy.XPATH,
     18     print("Downloading {0} of {1} images" .format(i,10))
     19     response=requests.get(img_urls[i])

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

In [82]:
```
df= pd.DataFrame({"IMG_URLS":img_urls})

df[0:10]
```

Out[82]:

	IMG_URLS
0	https://encrypted-tbn0.gstatic.com/images?q=tb...
1	https://encrypted-tbn0.gstatic.com/images?q=tb...
2	https://encrypted-tbn0.gstatic.com/images?q=tb...
3	https://encrypted-tbn0.gstatic.com/images?q=tb...
4	https://encrypted-tbn0.gstatic.com/images?q=tb...
5	https://encrypted-tbn0.gstatic.com/images?q=tb...
6	https://encrypted-tbn0.gstatic.com/images?q=tb...
7	https://encrypted-tbn0.gstatic.com/images?q=tb...
8	https://encrypted-tbn0.gstatic.com/images?q=tb...
9	https://encrypted-tbn0.gstatic.com/images?q=tb...

In [ ]:
```
#4. Write a python program to search for a smartphone(e.g.: Oneplus Nord, pixel 4A, etc.) on www.flipkart.com
#and scrape following details for all the search results displayed on 1st page. Details to be scraped: “Brand
#Name”, “Smartphone name”, “Colour”, “RAM”, “Storage(ROM)”, “Primary Camera”,
#“Secondary Camera”, “Display Size”, “Battery Capacity”, “Price”, “Product URL”. Incase if any of the
#details is missing then replace it by “- “. Save your results in a dataframe and CSV.
```

In [152…
```
driver=webdriver.Chrome(r"chromedriver.exe")
driver.get("https://www.flipkart.com/")
```

In [153…
```
keywords=driver.find_element(By.XPATH,"/html/body/div[1]/div/div[1]/div[1]/div[2]/div[2]/form/div/div/input")
keywords.send_keys('Oneplus Nord')
```

In [154…
```
search=driver.find_element(By.XPATH,"/html/body/div[1]/div/div[1]/div[1]/div[2]/div[2]/form/div/button")
search.click()
```

In [156…
```
urls=[]
url=driver.find_elements(By.XPATH,'//a[@class="_1fQZEK"]')

for i in url:

    urls.append(i.get_attribute("href"))

print(len(urls))
```
```
24
```

In [137…
```
Brand_Name=[]
for url in urls:
    driver.get(url)
    time.sleep(2)

    try:
        brand_name=driver.find_element(By.XPATH,'//span[@class="B_NuCI"]')
        Brand_Name.append(brand_name.text)
    except NoSuchElementException:
        Brand_Name.append("-")
```

In [157…
```
Smartphone_Name=[]
for url in urls:
    driver.get(url)
    time.sleep(2)

    try:
        smartphone_name=driver.find_element(By.XPATH,'//span[@class="B_NuCI"]')
        Smartphone_Name.append(smartphone_name.text)
    except NoSuchElementException:
        Smartphone_Name.append("-")
```

In [167…
```
Price=[]
for url in urls:
    driver.get(url)
    time.sleep(2)

    try:
        price=driver.find_element(By.XPATH,'//div[@class="_30jeq3 _16Jk6d"]')
        Price.append(price.text)
    except NoSuchElementException:
        Price.append("-")
```

In [170…
```
Colour=[]
for url in urls:
    driver.get(url)
    time.sleep(2)

    try:
        colour=driver.find_element(By.XPATH,'/html/body/div[1]/div/div[3]/div[1]/div[2]/div[7]/div/div/div[4]/div/div[2]/div/div[1]/table/tbody/tr[4]/td[2]')
        Colour.append(colour.text)
    except NoSuchElementException:
        Colour.append("-")
```

In [181…
```
Storage=[]
for url in urls:
    driver.get(url)
    time.sleep(5)

    try:
        storage=driver.find_element(By.XPATH,'/html/body/div[1]/div/div[3]/div[1]/div[2]/div[6]/div[1]/div/div[2]/ul/li[1]')
        Storage.append(storage.text)
    except NoSuchElementException:
        Storage.append("-")
```

In [189…
```
Primary_Camera=[]
for url in urls:
    driver.get(url)
    time.sleep(4)

    try:
        primary_camera=driver.find_element(By.XPATH,'/html/body/div[1]/div/div[3]/div[1]/div[2]/div[6]/div[1]/div/div[2]/ul/li[3]')
        Primary_Camera.append(primary_camera.text)
    except NoSuchElementException:
        Primary_Camera.append("-")
```

In [193…
```
Display_Size=[]
for url in urls:
    driver.get(url)
    time.sleep(5)

    try:
        display_size=driver.find_element(By.XPATH,'/html/body/div[1]/div/div[3]/div[1]/div[2]/div[6]/div[1]/div/div[2]/ul/li[2]')
        Display_Size.append(display_size.text)
    except NoSuchElementException:
        Display_Size.append("-")
```

In [199…
```
Battery_Capacity=[]
for url in urls:
    driver.get(url)
    time.sleep(3)

    try:
        battery_capacity=driver.find_element(By.XPATH,'/html/body/div[1]/div/div[3]/div[1]/div[2]/div[7]/div/div/div[4]/div/div[2]/div[1]/div[7]/table/tbody/tr
        Battery_Capacity.append(battery_capacity)
    except NoSuchElementException:
        Battery_Capacity.append("-")
```

In [202…
```
print(len( Brand_Name),len(Smartphone_Name),len(Price),len(Colour),len(Storage),len(Primary_Camera),len(Display_Size),len(Battery_Capacity),len(urls))
```
```
24 24 24 24 24 24 24 24 24
```

In [207…
```
df1=pd.DataFrame({'BRAND':Brand_Name,'NAME':Smartphone_Name,'PRICE':Price,'COLOUR':Colour,'STORAGE':Storage,'CAMERA':Primary_Camera,'DISPLAY':Display_Size,'BA

df1
```

Out[207]:

	BRAND	NAME	PRICE	COLOUR	STORAGE	CAMERA	DISPLAY	BATTERY	URLS
0	OnePlus Nord CE 2 Lite 5G (Blue Tide, 128 GB) ...	OnePlus Nord CE 2 Lite 5G (Blue Tide, 128 GB) ...	₹19,990	-	6 GB RAM | 128 GB ROM	64MP Rear Camera	16.74 cm (6.59 inch) Display	-	https://www.flipkart.com/oneplus-nord-ce-2-lit...
1	OnePlus Nord CE 2 Lite 5G (Black Dusk, 128 GB)...	OnePlus Nord CE 2 Lite 5G (Blue Tide, 128 GB) ...	₹21,999	-	-	-	-	-	https://www.flipkart.com/oneplus-nord-ce-2-lit...
2	-	OnePlus Nord CE 2 Lite 5G (Black Dusk, 128 GB)...	₹19,675	-	-	-	-	-	https://www.flipkart.com/oneplus-nord-ce-2-lit...
3	OnePlus Nord CE 2 Lite 5G (Black Dusk, 128 GB)...	-	-	-	-	-	-	-	https://www.flipkart.com/oneplus-nord-ce-2-lit...
4	OnePlus Nord CE 2 Lite 5G (Blue Tide, 128 GB) ...	OnePlus Nord CE 2 Lite 5G (Black Dusk, 128 GB) ...	₹21,999	-	-	-	-	-	https://www.flipkart.com/oneplus-nord-ce-2-lit...
5	OnePlus Nord CE 5G (Charcoal Ink, 128 GB) (8 ...	OnePlus Nord CE 5G (Charcoal Ink, 256 GB) (12 GB...	₹24,990	-	-	-	-	-	https://www.flipkart.com/oneplus-nord-ce-5g-ch...
6	OnePlus Nord 2T 5G (Gray Shadow, 128 GB) (8 G...	OnePlus Nord 2T 5G (Gray Shadow, 128 GB) (8 G...	₹28,999	-	-	-	-	-	https://www.flipkart.com/oneplus-nord-2t-5g-gr...
7	OnePlus Nord 2T 5G (Jade Fog, 128 GB) (8 GB RAM)	OnePlus Nord 2T 5G (Jade Fog, 128 GB) (8 GB RAM)	₹28,999	-	-	-	-	-	https://www.flipkart.com/oneplus-nord-2t-5g-ja...
8	OnePlus Nord 2 5G (Gray Sierra, 128 GB) (8 GB...	OnePlus Nord 2 5G (Gray Sierra, 128 GB) (8 GB...	₹29,999	Gray Sierra	8 GB RAM | 128 GB ROM	50MP Rear Camera	16.33 cm (6.43 inch) Display	<selenium.webdriver.remote.webelement.WebEleme...	https://www.flipkart.com/oneplus-nord-2-5g-gra...
9	-	-	-	-	-	-	-	-	https://www.flipkart.com/oneplus-nord-ce-2-lit...
10	OnePlus Nord CE 5G (Blue Void, 256 GB) (12 GB...	OnePlus Nord CE 5G (Blue Void, 256 GB) (12 GB...	₹27,900	-	-	-	-	-	https://www.flipkart.com/oneplus-nord-ce-5g-bl...
11	OnePlus Nord CE 5G (Charcoal Ink, 256 GB) (12...	OnePlus Nord CE 5G (Charcoal Ink, 256 GB) (12...	₹27,900	-	-	-	-	-	https://www.flipkart.com/oneplus-nord-ce-5g-ch...
12	-	-	-	-	-	-	-	-	https://www.flipkart.com/oneplus-nord-ce-2-lit...
13	OnePlus Nord 2 5G (Blue Haze, 128 GB) (8 GB RAM)	OnePlus Nord 2 5G (Blue Haze, 128 GB) (8 GB RAM)	₹29,799	-	-	-	-	-	https://www.flipkart.com/oneplus-nord-2-5g-blu...
14	OnePlus Nord 2 5G (Blue Haze, 256 GB) (12 GB ...	OnePlus Nord 2 5G (Blue Haze, 256 GB) (12 GB ...	₹33,998	Blue Haze	12 GB RAM | 256 GB ROM	50MP Rear Camera	16.33 cm (6.43 inch) Display	<selenium.webdriver.remote.webelement.WebEleme...	https://www.flipkart.com/oneplus-nord-2-5g-blu...
15	-	OnePlus Nord CE 2 5G (Gray Mirror, 128 GB) (6...	₹23,999	Gray Mirror	6 GB RAM | 128 GB ROM	50MP Rear Camera	16.33 cm (6.43 inch) Display	<selenium.webdriver.remote.webelement.WebEleme...	https://www.flipkart.com/oneplus-nord-ce-2-5g-...
16	OnePlus Nord CE 5G (Silver Gray, 128 GB) (8 G...	OnePlus Nord CE 5G (Silver Gray, 128 GB) (8 G...	₹24,999	-	-	-	-	-	https://www.flipkart.com/oneplus-nord-ce-5g-si...
17	OnePlus Nord CE 2 5G (Gray Mirror, 128 GB) (8...	OnePlus Nord CE 5G (Blue Void, 128 GB) (6 GB ...	₹22,999	-	-	-	-	-	https://www.flipkart.com/oneplus-nord-ce-5g-bl...
18	OnePlus Nord CE 2 5G (Bahama Blue, 128 GB) (8...	OnePlus Nord CE 2 5G (Gray Mirror, 128 GB) (8...	₹24,990	Gray Mirror	8 GB RAM | 128 GB ROM	-	16.33 cm (6.43 inch) Display	<selenium.webdriver.remote.webelement.WebEleme...	https://www.flipkart.com/oneplus-nord-ce-2-5g-...
19	OnePlus Nord CE 5G (Silver Ray, 128 GB) (6 GB...	OnePlus Nord CE 2 5G (Bahama Blue, 128 GB) (8...	₹24,999	Bahama Blue	8 GB RAM | 128 GB ROM	64MP Rear Camera	16.33 cm (6.43 inch) Display	<selenium.webdriver.remote.webelement.WebEleme...	https://www.flipkart.com/oneplus-nord-ce-2-5g-...
20	OnePlus Nord CE 5G (Charcoal Ink, 128 GB) (6 ...	OnePlus Nord CE 5G (Silver Ray, 128 GB) (6 GB...	₹24,499	-	-	-	-	-	https://www.flipkart.com/oneplus-nord-ce-5g-si...
21	OnePlus Nord CE 5G (Blue Void, 128 GB) (6 GB ...	OnePlus Nord CE 5G (Charcoal Ink, 128 GB) (6 ...	₹20,999	-	-	-	-	-	https://www.flipkart.com/oneplus-nord-ce-5g-ch...
22	OnePlus Nord CE 5G (Silver Ray, 256 GB) (12 G...	OnePlus Nord CE 5G (Silver Ray, 256 GB) (12 G...	₹27,990	-	-	-	-	-	https://www.flipkart.com/oneplus-nord-ce-5g-si...
23	OnePlus Nord 2T 5G (Jade Fog, 256 GB) (12 GB ...	OnePlus Nord 2T 5G (Jade Fog, 256 GB) (12 GB ...	₹32,990	Jade Fog	12 GB RAM | 256 GB ROM	50MP Rear Camera	17.02 cm (6.7 inch) Display	<selenium.webdriver.remote.webelement.WebEleme...	https://www.flipkart.com/oneplus-nord-2t-5g-ja...

In [208…
```
df1.to_csv('oneplus_nord.csv')
```

In [ ]:
```
#5. Write a program to scrap geospatial coordinates (latitude, longitude) of a city searched on google maps
```

In [263…
```
driver=webdriver.Chrome(r"chromedriver.exe")
driver.get('https://www.google.co.in/maps/')
```

In [264…
```
keywords=driver.find_element(By.XPATH,'//input[@class="tactile-searchbox-input"]')
keywords.send_keys('Delhi')
```

In [265…
```
search=driver.find_element(By.XPATH,'//button[@class="mL3xi"]')
search.click()
```

In [268…
```
 try:

    url_string = driver.current_url

    print("URL Extracted: ", url_string)

    lat_lng = re.findall(r'@(.*)data',url_string)

    if len(lat_lng):

        lat_lng_list = lat_lng[0].split(",")

        if len(lat_lng_list)>=2:

            lat = lat_lng_list[0]

            lng = lat_lng_list[1]

        print("Latitude = {}, Longitude = {}".format(lat, lng))


except Exception as e:


    lat_lng = re.findall( r'@(.*)data', url_string )
```
```
URL Extracted:  https://www.google.co.in/maps/place/Delhi/@28.6443981,76.8130253,10z/data=!3m1!4b1!4m6!3m5!1s0x390cfd5b347eb62d:0x37205b715389640!8m2!3d28.704
0592!4d77.1024902!16zL20vMDlmMDc
Latitude = 28.6443981, Longitude = 76.8130253
```

In [ ]:
```
#7. Write a python program to scrape the details for all billionaires from www.forbes.com. Details to be scrapped:
#"Rank", "Name", "Net worth", "Age", "Citizenship", "Source", "Industry".
```

In [379…
```
driver=webdriver.Chrome(r"chromedriver.exe")
driver.get('https://www.forbes.com/')
```

In [327…
```
lists=driver.find_element(By.XPATH,'//svg[@class="hamburger_svg__fs-icon hamburger_svg__fs-icon--hamburger"]')
lists.click()
```

In [328…
```
sub_list=driver.find_element(By.XPATH,'//div[@class="mpBfVZz3"]')
sub_list.click()
```

In [329…
```
billionaires=driver.find_element(By.XPATH,'//li[@class="TjJgrPSg _2bNo56RE secondary"]')
billionaires.click()
```

In [330…
```
Rank=[]
rank=driver.find_elements(By.XPATH,'//div[@class="rank"]')
for i in rank:
    ranks=i.text
    Rank.append(ranks)
```

In [335…
```
Names=[]
name=driver.find_elements(By.XPATH,'//div[@class="personName"]')
for i in name:
    names=i.text
    Names.append(names)
```

In [340…
```
Net_Worth=[]
net_worth=driver.find_elements(By.XPATH,'//div[@class="netWorth"]')
for i in net_worth:
    net_worths=i.text
    Net_Worth.append(net_worths)
```

In [384…
```
Age=[]
age=driver.find_elements(By.XPATH,'//div[@class="age"]//div')
for i in age:
    ages=i.text
    Age.append(ages)
```

In [344…
```
Citizenship=[]
citizenship=driver.find_elements(By.XPATH,'//div[@class="countryOfCitizenship"]')
for i in citizenship:
    citizenships=i.text
    Citizenship.append(citizenships)
```

In [345…
```
Source=[]
source=driver.find_elements(By.XPATH,'//div[@class="source-column"]')
for i in source:
    sources=i.text
    Source.append(sources)
```

In [381…
```
Industry=[]
industry=driver.find_elements(By.XPATH,'//div[@class="category"]//div')
for i in industry:
    industries=i.text
    Industry.append(industries)
```

In [382…
```
print(len(Rank),len(Names),len(Net_Worth),len(Age),len(Citizenship),len(Source),len(Industry))
```
```
200 200 200 200 200 200 200
```

In [385…
```
df=pd.DataFrame({"RANKS":Rank,"NAMES":Names,"NET_WORTH":Net_Worth,"AGE":Age,"CITIZENSHIP":Citizenship,"SOURCES":Source,"INDUSTRY":Industry})

df
```

Out[385]:

	RANKS	NAMES	NET_WORTH	AGE	CITIZENSHIP	SOURCES	INDUSTRY
0	1.	Elon Musk	$219 B	50	United States	Tesla, SpaceX	Automotive
1	2.	Jeff Bezos	$171 B	58	United States	Amazon	Technology
2	3.	Bernard Arnault & family	$158 B	73	France	LVMH	Fashion & Retail
3	4.	Bill Gates	$129 B	66	United States	Microsoft	Technology
4	5.	Warren Buffett	$118 B	91	United States	Berkshire Hathaway	Finance & Investments
...	...	...	...	...	...	...	...
195	192.	Marcel Herrmann Telles	$10.3 B	72	Brazil	beer	Food & Beverage
196	197.	Leon Black	$10 B	70	United States	private equity	Finance & Investments
197	197.	Joe Gebbia	$10 B	40	United States	Airbnb	Technology
198	197.	David Geffen	$10 B	79	United States	movies, record labels	Media & Entertainment
199	197.	Yu Renrong	$10 B	56	China	semiconductors	Manufacturing

200 rows × 7 columns

In [ ]:
```
#8. Write a program to extract at least 500 Comments, Comment upvote and time when comment was posted
#from any YouTube Video.
```

In [414…
```
driver=webdriver.Chrome(r"chromedriver.exe")
driver.get('https://www.youtube.com/')
```

In [364…
```
video=driver.find_elements(By.XPATH,'//a[@class="yt-simple-endpoint inline-block style-scope ytd-thumbnail"]')
video.click()
```

In [533…
```
for _ in range(10):
    driver.execute_script("window.scrollBy(0,10000)")
```

In [628…
```
Comments=[]
comments=driver.find_elements(By.ID,"content-text")
for i in comments[0:500]:
    comment=i.text
    Comments.append(comment)
```

In [624…
```
Time=[]
times=driver.find_elements(By.ID,"header-author")
for i in times[0:500]:
    text=i.text
    name = text.split('\n')[1]
    time_pattern = r'\d{1,2}:\d{2}\s*(AM|PM)'
    time_matches = re.findall(time_pattern, text)
    if time_matches:
        time = time_matches[1]
    else:
        time = ''
    Time.append((name,time))
```

In [630…
```
print(len(Time),len(Comments))
```
```
500 500
```

In [631…
```
df=pd.DataFrame({"COMMENTS":Comments,"TIME":Time})

df
```

Out[631]:

	COMMENTS	TIME
0	Kaunsa scene aapko sabse fav laga? Mera toh sa...	(2 days ago,)
1	No one can beat Kunal at being a teacher	(1 day ago,)
2	No one can beat Kunal at being a teacher and A...	(4 hours ago,)
3	I was missing old Ashish, and now he is back. ...	(8 hours ago,)
4	When a video is needed to backward again and a...	(2 hours ago,)
...	...	...
495	mera psycho wala part	(44 minutes ago,)
496	KEEP ANIMALS SAFE THIS HOLI \n\nHAPPY HOLI WIT...	(11 hours ago,)
497	Psycho dost sabse accha tha...	(26 minutes ago,)
498	Purana wala Ashish is back	(1 minute ago,)
499	And your holy dressup scene is the best ever	(3 minutes ago,)

500 rows × 2 columns

In [ ]:

In [ ]:

In [ ]:

In [ ]: