**Kulvir Singh**

**19BCE2074**

**Experiment 3 (22.02.2020)**

*Question 1*

***Problem Statement :***

Use beautiful soup to crawl any one of the E-commerce websites of your choice and perform the same. The following information needs to be extracted from the page: (Choose any one product: e.g. laptop, smartphone … etc.)

a)Product name

b)Product price

c)Product discount

d)Product image

***Procedure :***

Tools Used: Beautiful Soup(bs4), request packages, python, ide to run the program (Google Colab) for execution of various commands

Steps for Implementation:

1. Install the bs4, request packages (pip install bs4, request)

2. Store the URL to be scraped as a string

3. use request.get to get data from URL

4. use html parser for extracting the data as html

5. use find and accurate tags to extract the required data

6. run the code in the colab terminal

***Code :***

!pip install requests

!pip install bs4

import requests

from bs4 import BeautifulSoup

webpageurl = "https://www.flipkart.com/hp-pavilion-gaming-ryzen-5-quad-core-3550h-8-gb-1-tb-hdd-windows-10-home-4-gb-graphics-nvidia-geforce-gtx-1650-15-ec0101ax-laptop/p/itma1af6bf593dc8?pid=COMFSFNVDXG74QXR&lid=LSTCOMFSFNVDXG74QXRY8FRH2&marketplace=FLIPKART&srno=b\_1\_22&otracker=hp\_rich\_navigation\_6\_1.navigationCard.RICH\_NAVIGATION\_Electronics~Gaming\_PK7I48M403OR&fm=organic&iid=a4a75fb6-a7c4-406f-8829-6f0897fc37fc.COMFSFNVDXG74QXR.SEARCH&ppt=browse&ppn=browse&ssid=tri79mziu80000001614107514263"

data = requests.get(webpageurl)

soup = BeautifulSoup(data.text, 'html.parser')

product\_name = soup.find('span', {'class': 'B\_NuCI'})

product\_price = soup.find('div', {'class': '\_30jeq3 \_16Jk6d'})

product\_discount = soup.find('div', {'class': '\_3Ay6Sb \_31Dcoz'})

product\_image = soup.find('div', {'class': 'CXW8mj \_3nMexc'})

print("\n\n\*\*\*\*\*\*\*\*\*\*\*OUTPUT\*\*\*\*\*\*\*\*\*\*\*")

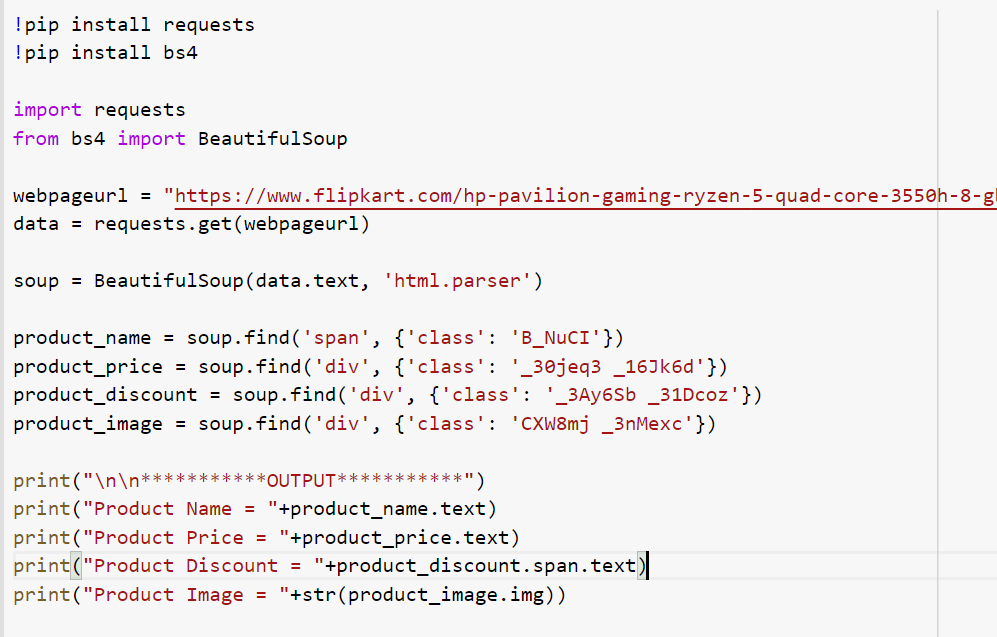
print("Product Name = "+product\_name.text)

print("Product Price = "+product\_price.text)

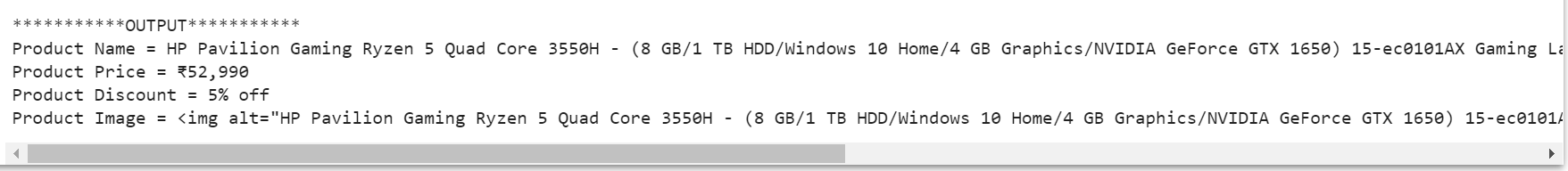
print("Product Discount = "+product\_discount.span.text)

print("Product Image = "+str(product\_image.img))

***Code Screenshot:***



***Output Screenshots :***



Output as text

\*\*\*\*\*\*\*\*\*\*\*OUTPUT\*\*\*\*\*\*\*\*\*\*\*

Product Name = HP Pavilion Gaming Ryzen 5 Quad Core 3550H - (8 GB/1 TB HDD/Windows 10 Home/4 GB Graphics/NVIDIA GeForce GTX 1650) 15-ec0101AX Gaming Laptop  (15.6 inch, Black, 2.04 kg)

Product Price = ₹52,990

Product Discount = 5% off

Product Image = img alt="HP Pavilion Gaming Ryzen 5 Quad Core 3550H - (8 GB/1 TB HDD/Windows 10 Home/4 GB Graphics/NVIDIA GeForce GTX 1650) 15-ec0101AX Gaming Laptop" class="\_396cs4 \_2amPTt \_3qGmMb" src="//img1a.flixcart.com/www/linchpin/fk-cp-zion/img/placeholder\_fcebae.svg"/