**CSE-3024 Web Mining**

**Lab Assignment 4**

**Alokam Nikhitha**

**19BCE2555**

Question

**Experiment 4 (28.01.2022)**

1. Use BeautifulSoup or Scrapy to crawl any one of the E-commerce website 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

**Problem statement:**

To Crawl any of the E-commerce website and extract the data from the page like Name, Price, Discount and image of the Product.

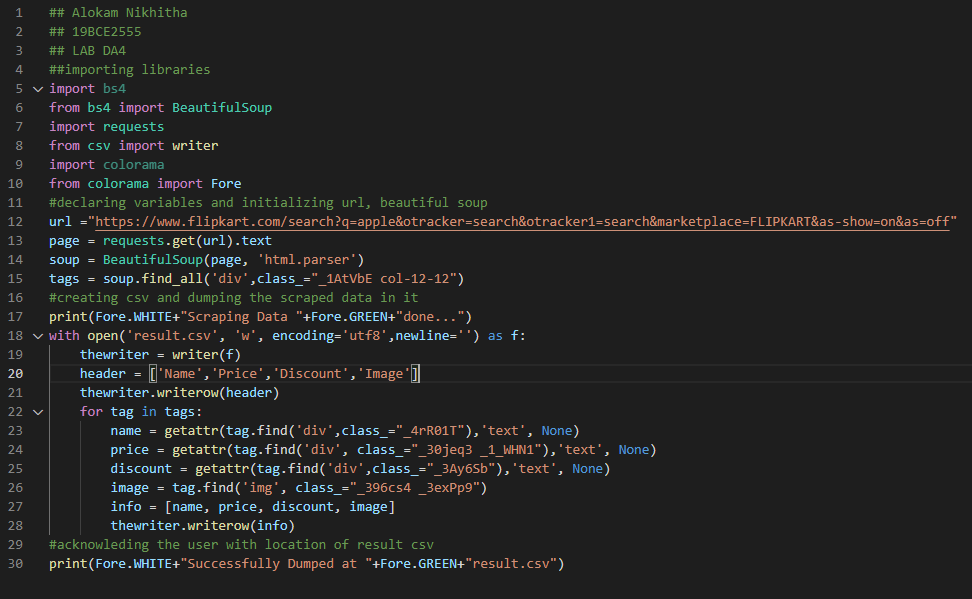
**Procedure:**

* We will Firstly import our libraries which are necessary in order to scrap the data from the website.
* Later we will declare the variables and also we will initialize the URL and also the beautifulsoup.
* Later we will create the file result.csv and dump the scrapped data into it. Here we will make a made the header with Name, Price, Discount and Image.
* Later we will collect the data of the product and add it to the CSV file.
* On running the python file the results.csv file will be created with scrapped data in it.
* In the result.csv file, the data of the product i.e, Name, Price, Discount and Image will be displayed .The image is returned in the form of a link.

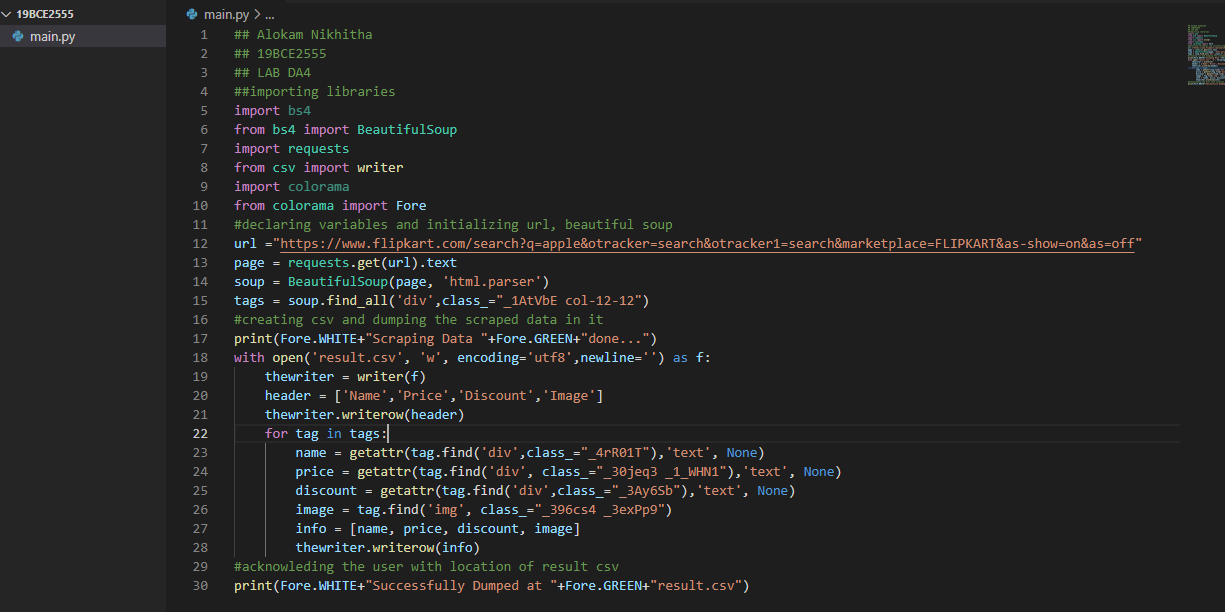
**URL of the website from which we are scrapping the data :**

"https://www.flipkart.com/search?q=apple&otracker=search&otracker1=search&marketplace=FLIPKART&as-show=on&as=off"

**Code:**

****

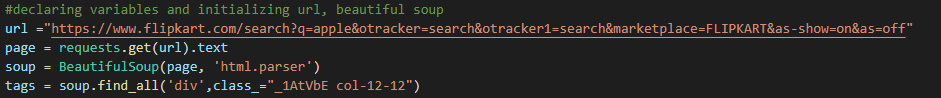
**Code Snippets and Outputs:**

****

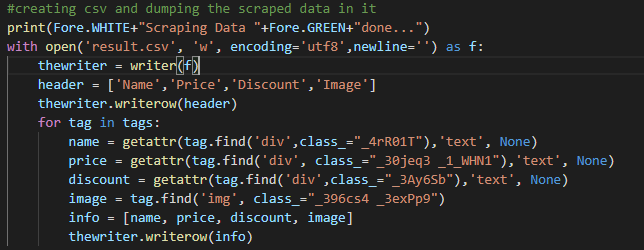
The python file is stored in the folder named ‘19BCE2555’.



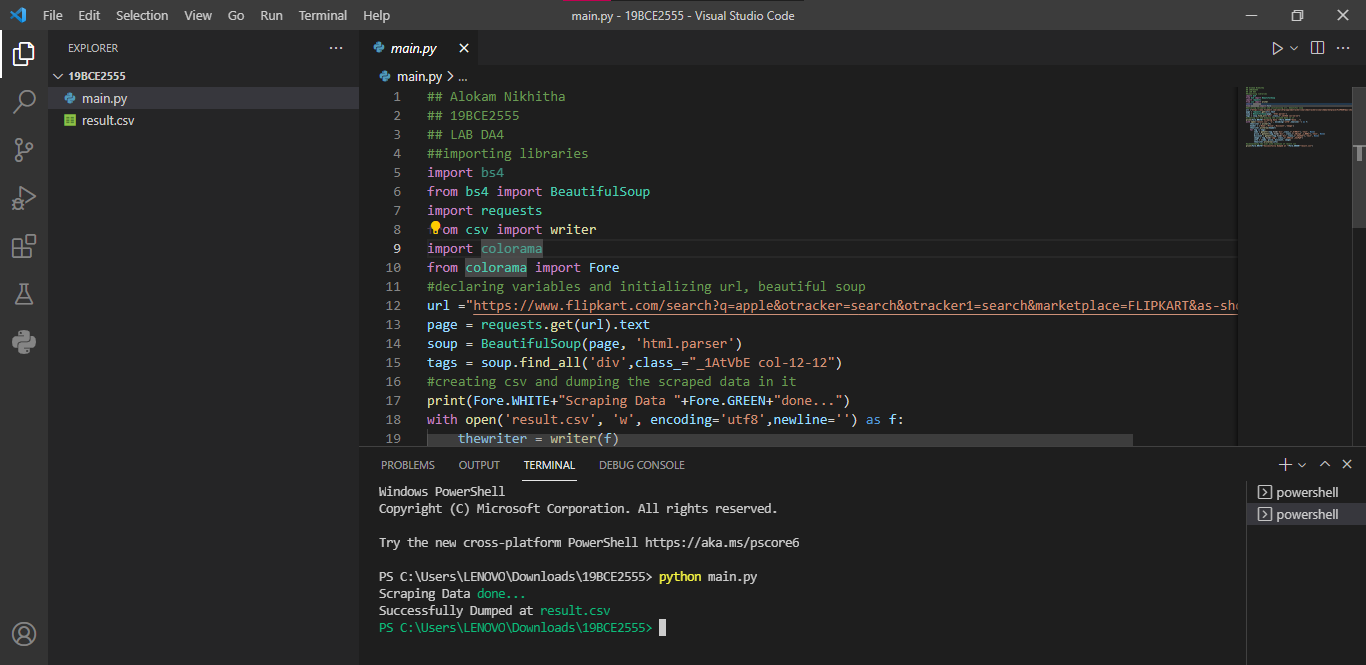
Here we are importing the necessary libraries inorder to Scrap the data.



Here we are declaring the variables and also we are Initializing the URL (here url is from ‘Flipkart’ site). And also initializing the beautifulsoup.

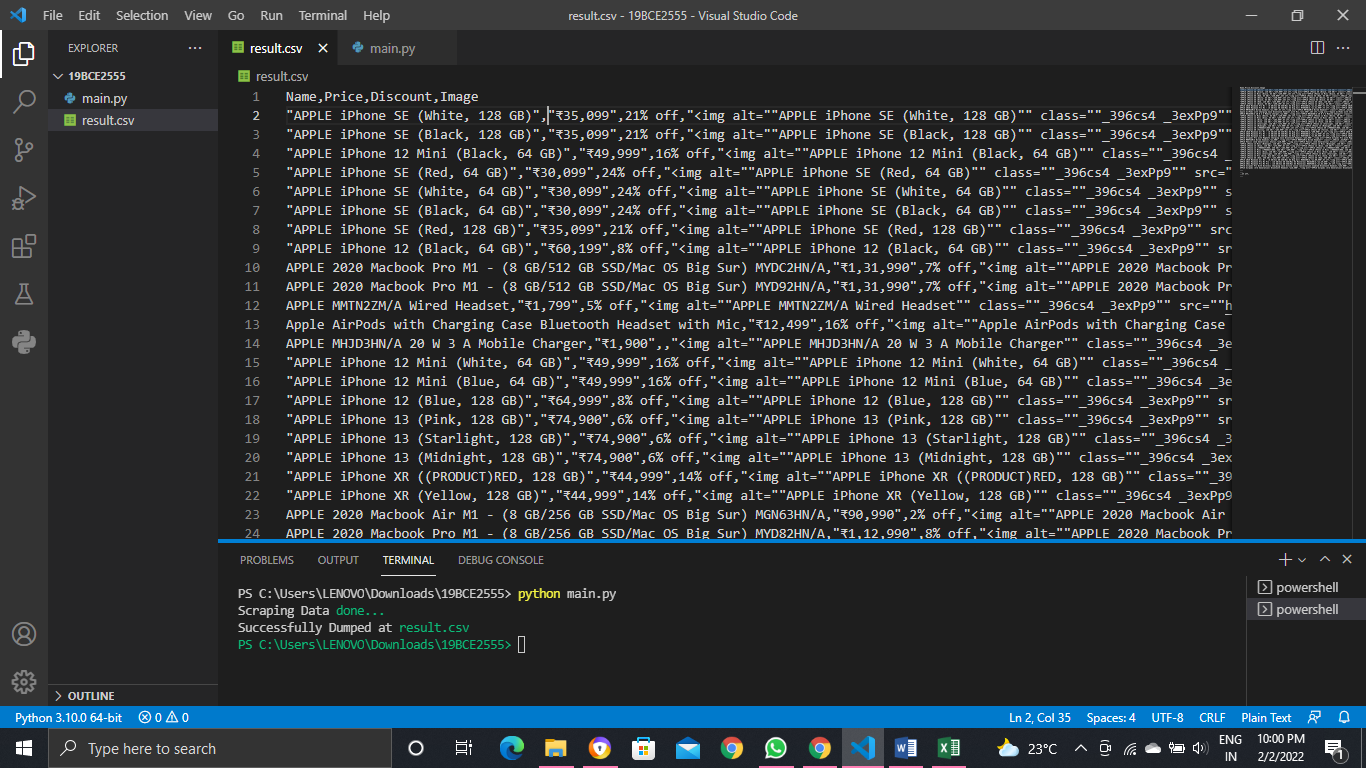


Here we are creating CSV file and dumping the scrapped data in it . Here we are opening the file result.csv and dumping the scrapped data into it. Here we made the header with Name, Price, Discount and Image. Later we are collecting the data of the product and adding it to the CSV file.

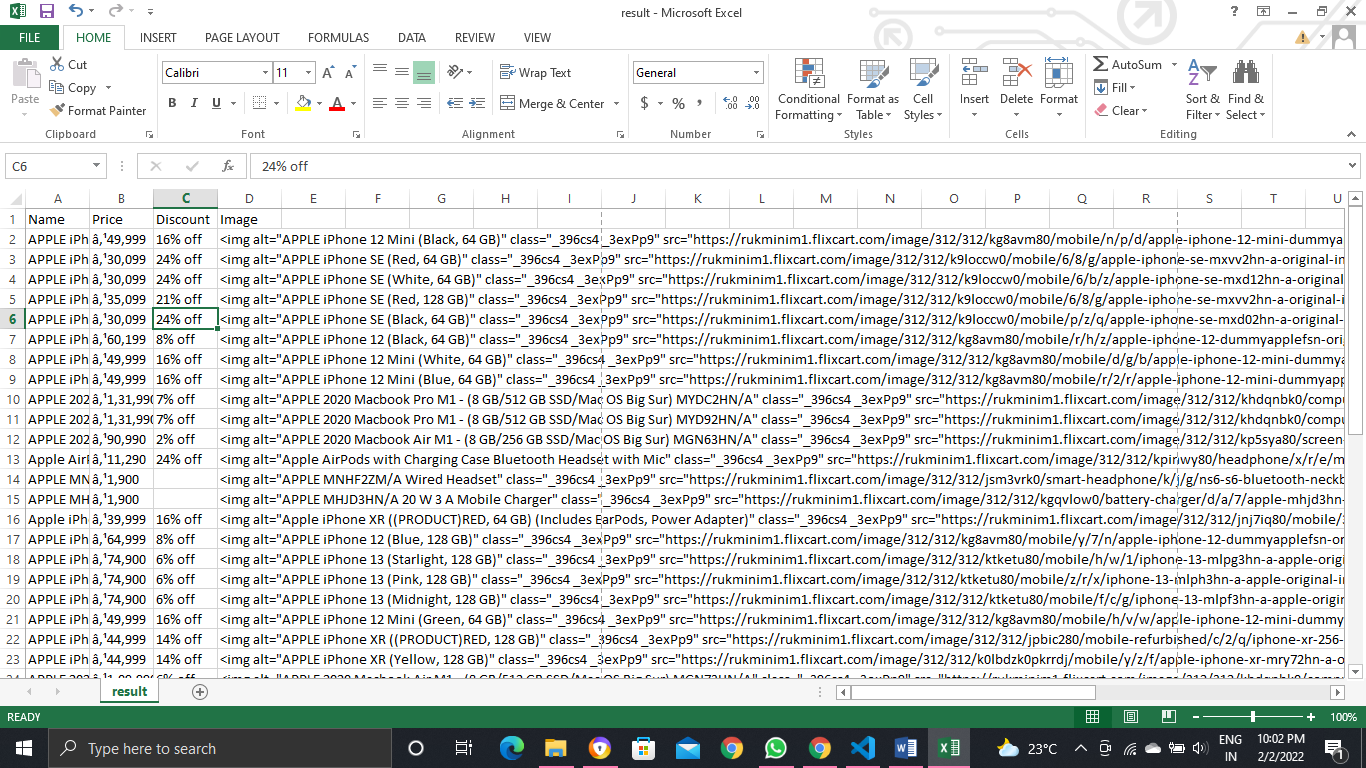


Here we are running the code using the command

‘python main.py’ in the terminal. Here the “result.csv” file got created in the folder after running the code.

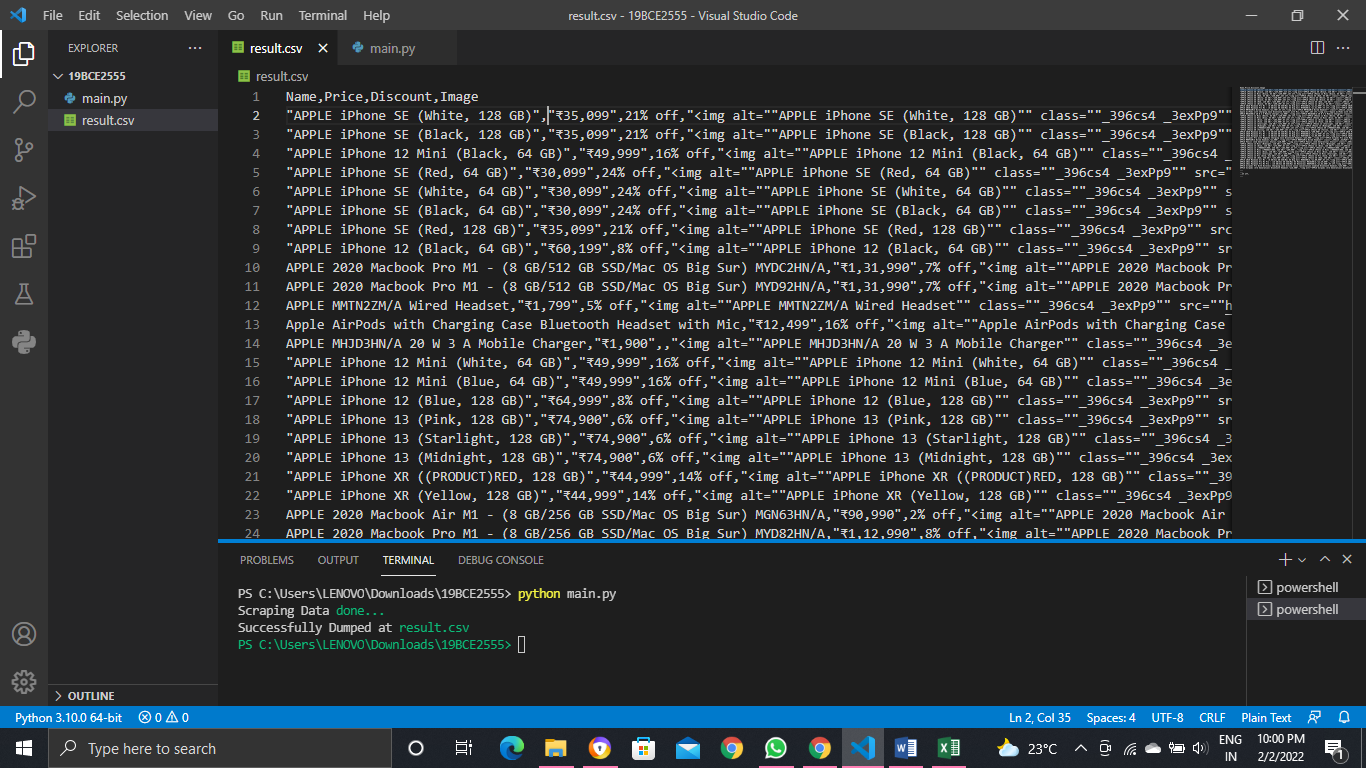


In the result.csv we can see the data that is scrapped from the Flipkart website.



Here is the Excel sheet in which we collected the scrapped data. The name of the product, price and Discount are collected. The image is returned in the form of a link.

**Results and Output**



Here is the list of data that is being scrapped from the website and dumped into result.csv file and which is formed in the folder of the code after running the code using beautifulsoup.



This is the view of Excel file in which the data is being stored.