**CSE-3024 Web Mining**

**Lab Assignment 4b (Using Scrapy)**

**Alokam Nikhitha**

**19BCE2555**

Question

**Experiment 4 b**

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

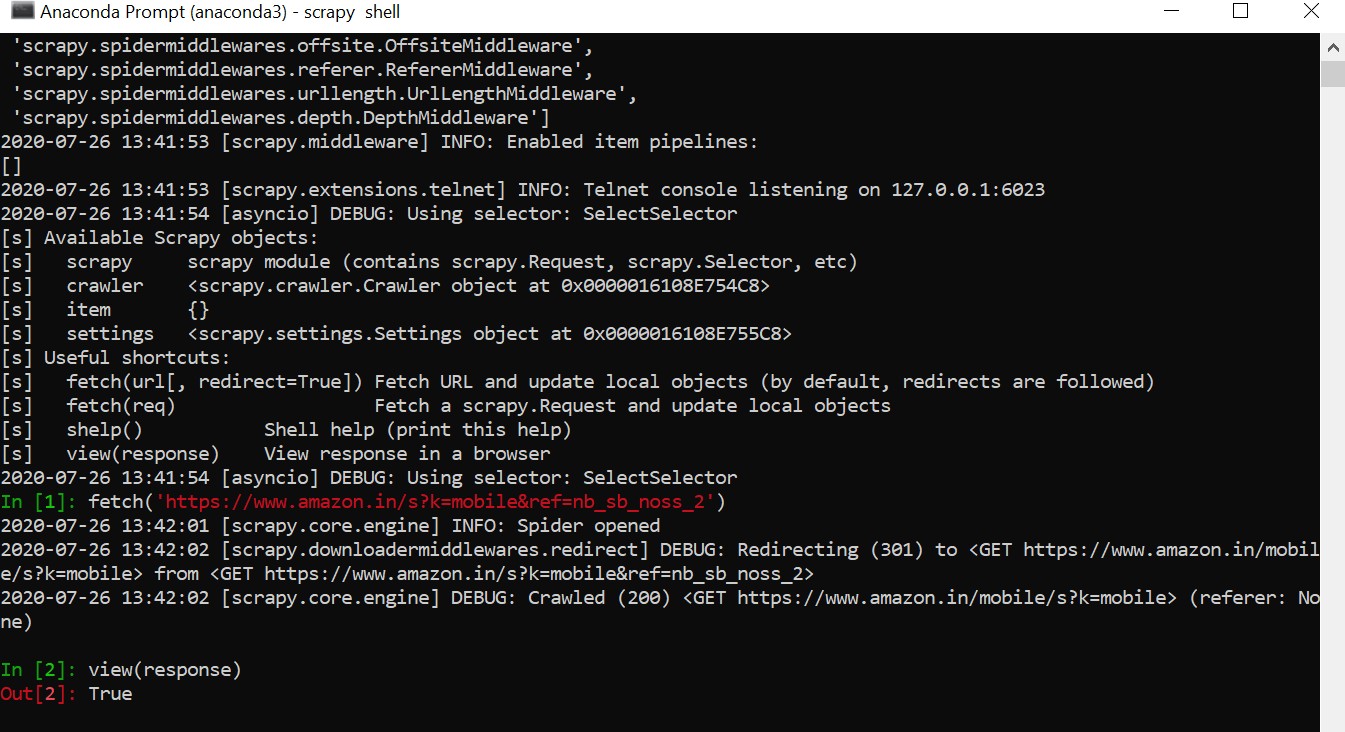
**Problem statement:**

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

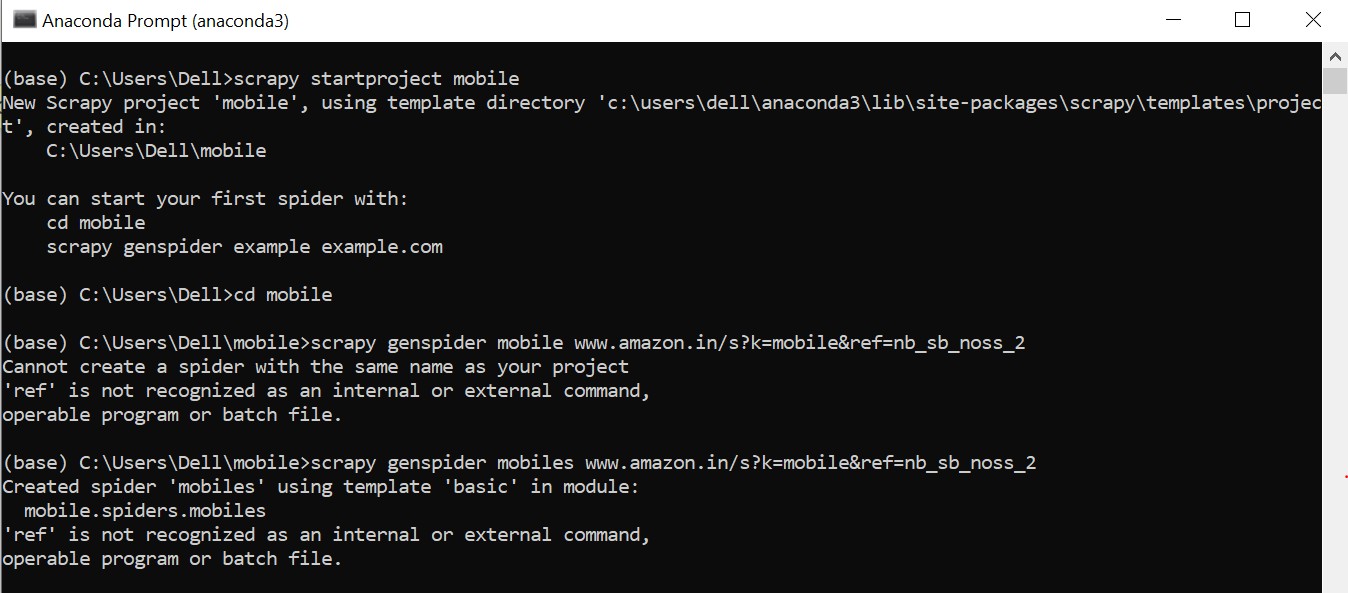
**Procedure:**

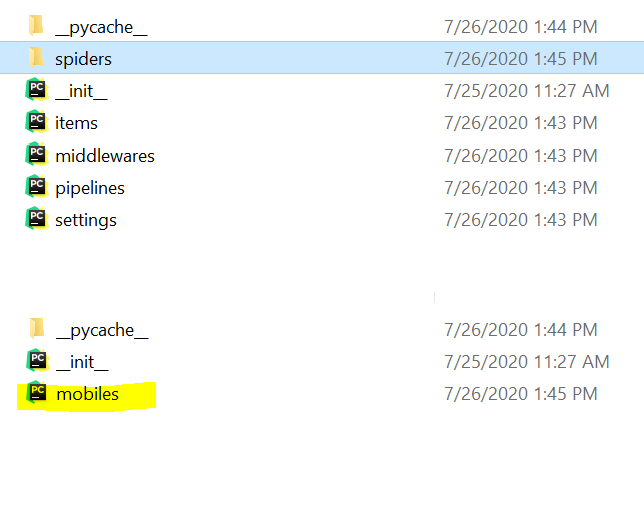
* + Firstly we install scrapy package with “pip install scrapy” in anaconda prompt
  + Later, we can start Shell by “scrapy shell”
  + Then Crawler run in the shell by use of the fetch and using view(response) to view fetched data.
  + An object should be created for the scrapper by “scrapy startproject mobile”
  + Create folder named “mobile”and move to that particular folder using command“cd mobile”
  + Create a python(.py) file inside the “spider” folder by using the command “scrapy genspider ..url..”
  + Here I scrapped data of amazons mobile as by product so the same url is pasted here.
  + Then python code is written in the file.
  + We can view the output in the terminal on typing “scrapy crawl ..name.. “ on teriminal
  + Finally it is exported as csv file using command “scrapy crawl mob -o data.csv”.

**Installing Scarpy in Anaconda .**



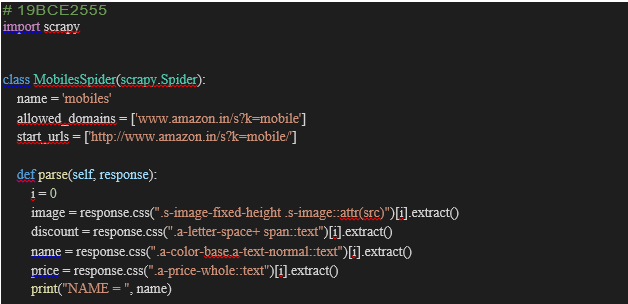
#creating scrapy project as name mobiles:

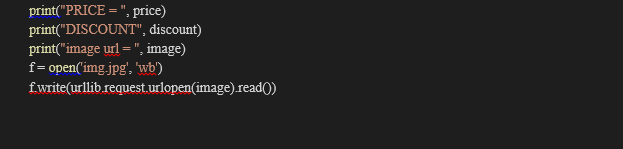




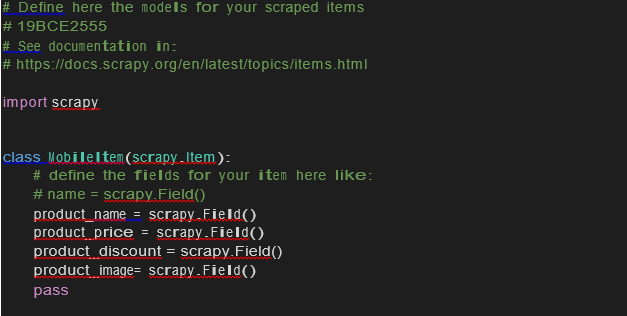
**Code:**

#mobiles.py:

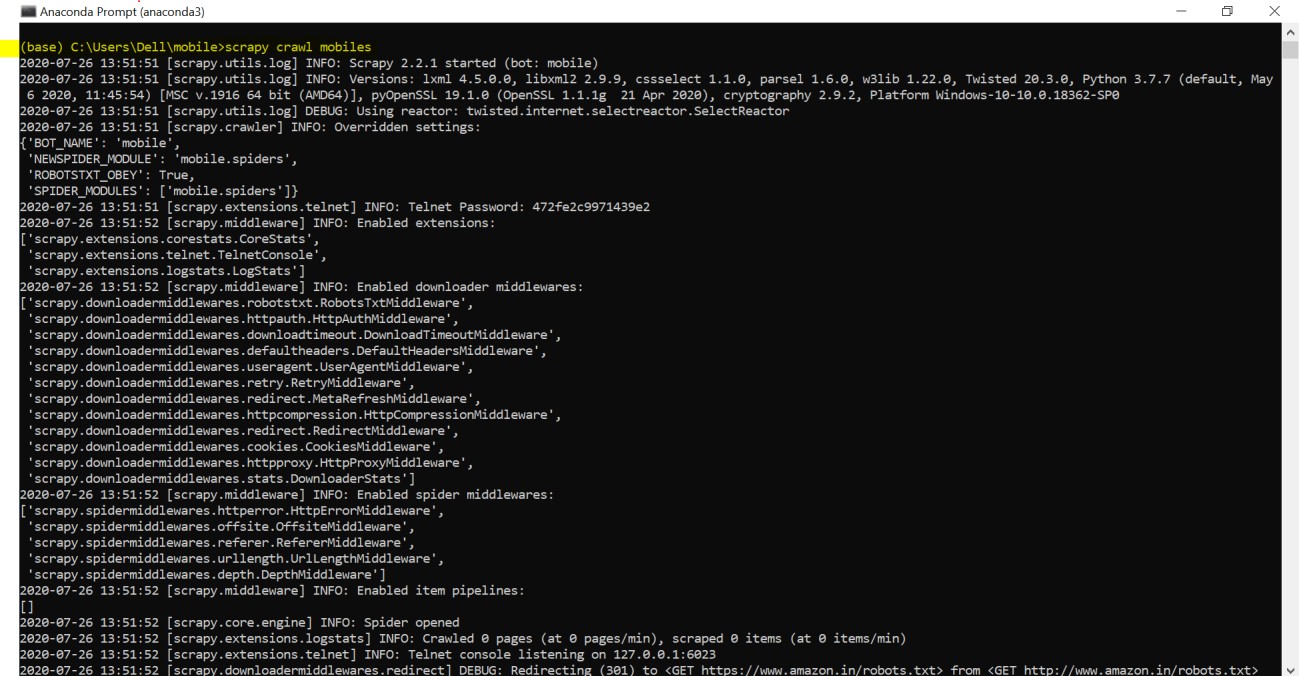


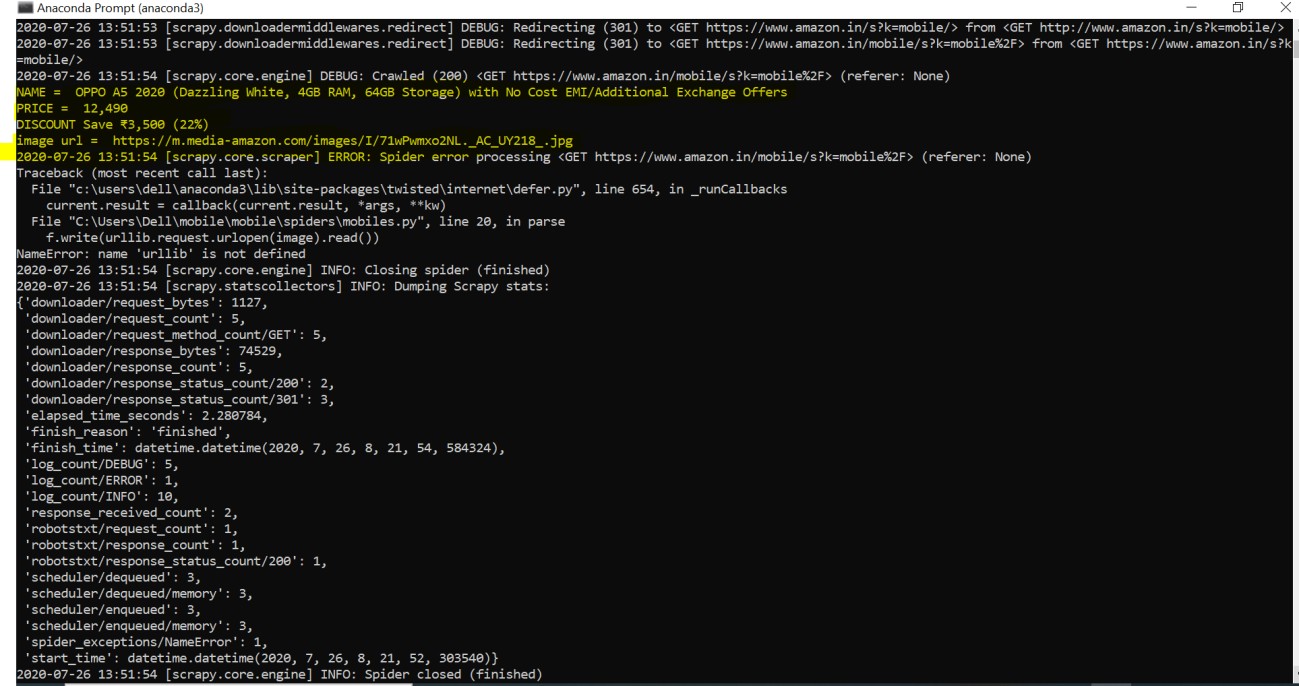


Items.py



**Output**



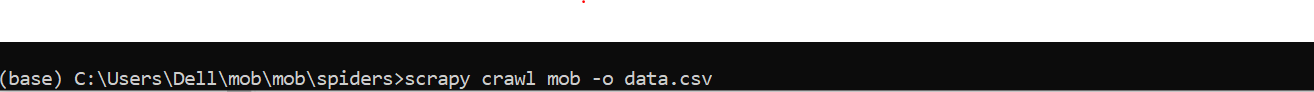
The basic information of the product is highlighted below 

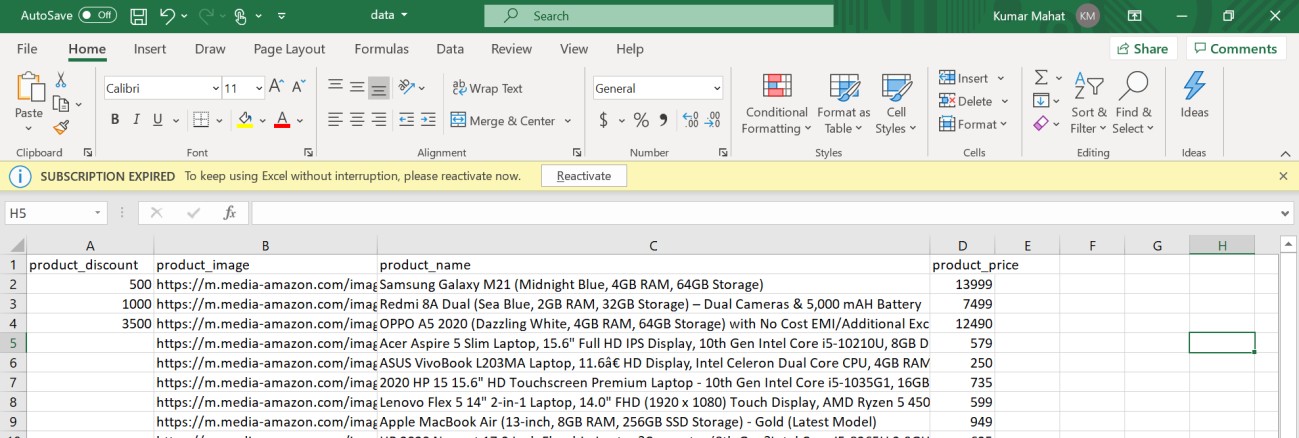
By clicking on the link extracted from the webpage we get the following image: Link: https://m.media-amazon.com/images/I/71wPwmxo2NL.\_AC\_UY218\_.jpg



**Results and Output**

#Exporting scrapped data as csv



We can see that the data is scrapped and it is dumped in excel sheet

