Ai WEB SCRAPING -AMAZON CUSTOMER REVIEWS

INTRODUCTION:

Web scraping is a technique used to extract large amounts of data from websites and store it in your computer . This data can be later used for analysis,how to scrape reviews of a particular product from Amazon website using python

STEPS:

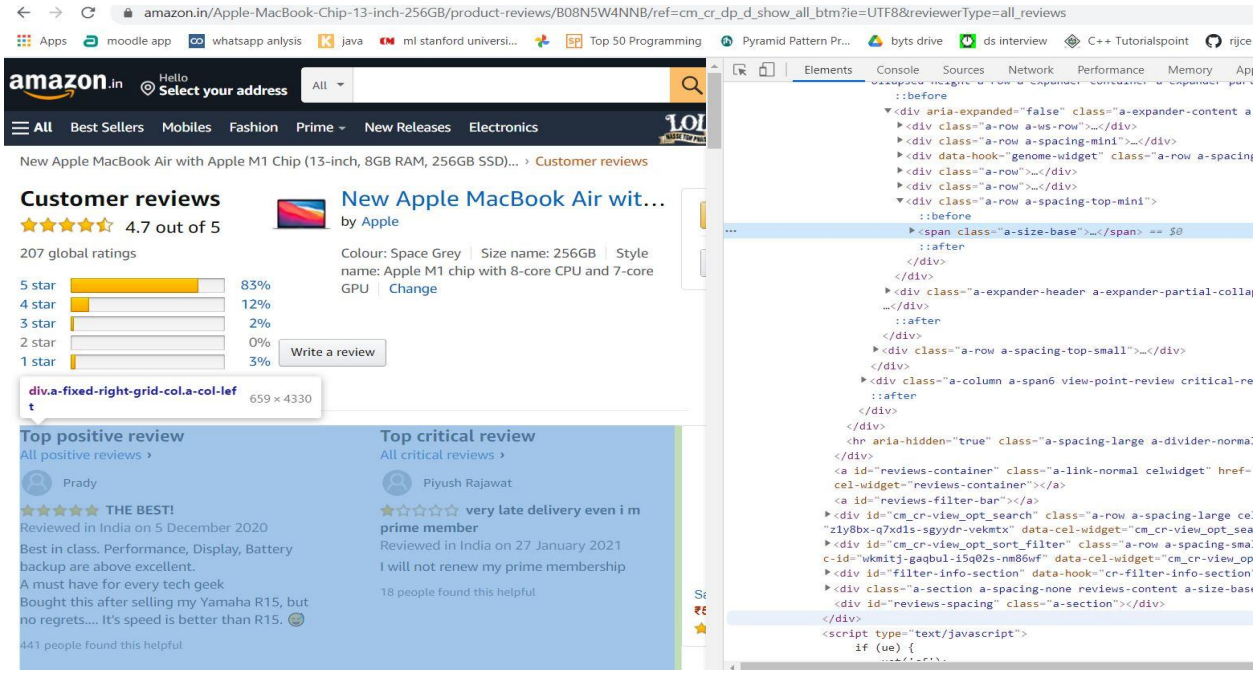
1)Get the Url of the page to be Scrapped.
2)Inspect the elements of the page and identify the Tags required.
3)Access the URL
4)Get the element from the Required Tags.

```
import requests
from bs4 import BeautifulSoup
```

the library ***requests*** is used to get the content from a web page. We send a request to the URL and we get a response. The response will contain a status code along with the web page content. ***BeautifulSoup*** converts the contents of a page into a proper format.

HEADER & COOKIES:

To find your headers and cookies, go to Amazon website and search for a particular product. Then right click any element and select Inspect (or use shortcut key Ctrl+Shift+I). From the Network tab, we can find headers and cookies.



Ai WEB SCRAPING -AMAZON CUSTOMER REVIEWS

Scraping Customer Review Links:

Beautiful Soup is a Python package for parsing HTML and XML documents. It creates a parse tree for parsed pages that can be used to extract data from HTML, which useful for web scraping

```
: link ='https://www.amazon.in/Apple-MacBook-Chip-13-inch-512GB/product-reviews/B08N5WRWNW/ref=cm_cr_dp_d_show_all_btm?ie=UTF8&reviewerType
=all_reviews'

: page = requests.get(link)

: page

: <Response [200]>

: page.content

: b'<!doctype html><html lang="en-in" class="a-no-js" data-19ax5a9jf="dingo"><!-- sp:feature:head-start -->\n<head><script>var aPageStart =
```

Customer reviews will be present in each page of the products.But these are just few.We want all the customer reviews for the products.so,we have to Scrape the “see all customer reviews” link and we first define a function which will go to pag of each and every product and extract all the customer reviews links of each product using corresponding HTML tags.

We use the below function to extract all the customer reviews in a List:

```
In [9]: names = soup.find_all('span',class_='a-profile-name')

In [10]: len(names)

Out[10]: 12

In [11]: cust_name =[]
         for i in range(0,len(names)):
           cust_name.append(names[i].get_text())

In [12]: cust_name

Out[12]: ['Anubhav Rai',
          'Sourabh Garg',
          'Anubhav Rai',
          'Sourabh Garg',
          'Dipesh',
          'Ramkrishna Pattnayak',
          'Sunny Raj Keshri',
          'Nitin K.',
          'Amazon Customer',
          'Vaneet G.',
          'Niraj',
          'Shreyans']

In [13]: cust_name.pop(1)

Out[13]: 'Sourabh Garg'
```

Ai WEB SCRAPING -AMAZON CUSTOMER REVIEWS

Savings Reviews in CSV file:

We have now scrapped all the reviews and further we have to save in a file in order to perform further analysis.

We convert the reviews list into a dictionary. Then import the pandas library and use it to convert the dictionary into a Data Frame.Then using **to-csv()** function we convert it into a csv file & store it in our computer.

```
In [49]: df['review description']=review_content
```

```
In [50]: df['review date']=review_date
```

```
In [51]: df['review title']=review_title
```

```
In [52]: df
```

Out[52]:

	Review given by	Ratings	review description	review date	review title
0	Anubhav Rai	1.0 out of 5 stars	Your browser does not support HTML5 video.\n...	29 June 2020	Lovely but An expensive Machine!!!!
1	Sourabh Garg	1.0 out of 5 stars	MacBook Pro is not repairable. This is espec...	10 October 2020	Think before you buy a MacBook Pro 2019
2	Dipesh	5.0 out of 5 stars	M1 goodness is here! Very powerful with pret...	28 December 2020	Pro MacBook for Pro users!
3	Ramkrishna Pattnayak	5.0 out of 5 stars	ConsCOST IS TOO HIGHCost Is High As per $ Pr...	30 July 2020	Higher spec model with core i5 10th gen proces...
4	Sunny Raj Keshri	5.0 out of 5 stars	Upgraded to this machine from macbook pro mi...	8 January 2021	Beginning of apple silicon era
5	Nitin K.	5.0 out of 5 stars	Apple really went all out with this M1 Macbo...	14 January 2021	Apple Silicon is much better than expected for...
6	Amazon Customer	1.0 out of 5 stars	My MacBook Pro suddenly stopped working afte...	4 March 2021	The laptop couldn't wait to see its first mont...
7	Vaneet G.	5.0 out of 5 stars	I am IT professional so I am giving the feed...	21 February 2021	Good product
8	Niraj	4.0 out of 5 stars	The device is functioning well and delivery ...	15 January 2021	Quality
9	Shreyans	5.0 out of 5 stars	Best performance with sturdy body. Less heat...	19 July 2020	Power and portability together

```
In [53]: df.to_csv("./review.csv", index =True)
```