ASUS LAPTOPS WEB SCRAPPING

Project for python module

Institute Name- Itvedant Education Pvt. Ltd

Name- Aniruddha Mangesh Ambre

Email Address- aniruddhaambre49@gmail.com

Date of Submission- 01 October 2023

- **Description:**

Asus laptops web scraping project involves extracting data from websites like flipkart by automating the process using scripts or software. This data can be used for various purposes, such as data analysis, research, reporting, or populating databases.

This database contains tables as follows:

1. Laptop names
2. Discounted price
3. Original price

❖ Project Steps:

A. Identify Target Websites:

Select the websites from which you plan to extract data. Make sure these websites either explicitly allow web scraping or have a robots.txt file that authorizes scraping.

B. 2. Define Data Requirements:

Specify the precise information you intend to extract. This may encompass text, images, hyperlinks, product specifications, prices, customer reviews, and so forth.

C. Choose a Web Scraping Tool/Programming Language:

Choose whether you'll utilize a web scraping library such as BeautifulSoup (for Python), Scrapy, or opt for a headless browser automation tool like Puppeteer (for JavaScript).

D. Develop Web Scraping Code:

Develop code to navigate the designated website(s), pinpoint pertinent HTML elements, and retrieve the targeted information. Address any pagination, authentication, or CAPTCHA hurdles if they arise.

E. Data Storage:

Determine the method and location for storing the gathered data. Standard choices encompass CSV files, databases (either SQL or NoSQL), or cloud-based storage services.

ASUS LAPTOPS WEB SCRAPPING

- **Project Benefits:**

1) Access to valuable data for analysis or research.
2) Automation of data collection, saving time and effort.
3) Stay up-to-date with changes on the target websites.

- **Project Risks:**

1) Legal issues related to web scraping.
2) Technical challenges due to website changes.

Asus laptops web scraping coding command:

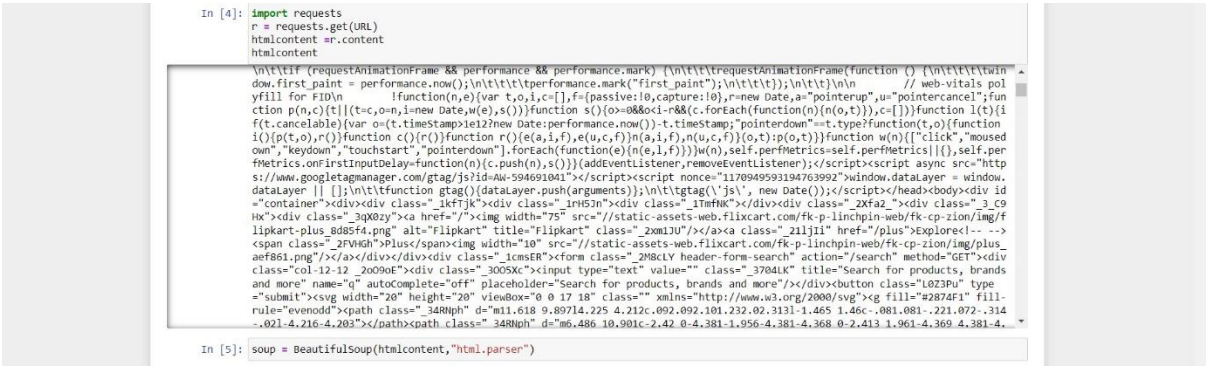
1. Accessing flipkart web site by using import requests URL and BeautifulSoup library.

```
In [2]: import numpy as np

In [ ]: #Accessing flipkart web site by using import requests URL and BeautifulSoup library.

In [3]: import requests
        from bs4 import BeautifulSoup
        import pandas as pd
        URL = "https://www.flipkart.com/computers/laptops/~asus-gaming-laptops21/pr?sid=6bo%2Cb5g&ctx=eyJjYXJkQ29udGV4dCI6eyJhdHRyaWJ1dGV
```

- Output



```
In [4]: import requests
        r = requests.get(URL)
        htmlcontent =r.content
        htmlcontent
```

\n\t\tif (requestAnimationFrame && performance && performance.mark) {\n\t\t\trequestAnimationFrame(function () {\n\t\t\t\twin
dow.first_paint = performance.now();\n\t\t\t\tperformance.mark("first_paint");\n\t\t\t\t});\n\t\t}\n\n   // web-vitals pol
yfill for FID\n   !function(n,e){var t,o,i,c=[],f={passive:!0,capture:!0},r=new Date,a="pointerup",u="pointercancel";fun
ction p(n,c){t||(t=c,o=n,i=new Date,w(e),s())}function s(){o>=0&&o<i-r&&(c.forEach(function(n){n(o,t)}),c=[])}function l(t){i
f(t.cancelable){var o=(t.timeStamp>1e12?new Date:performance.now())-t.timeStamp;"pointerdown"==t.type?function(t,o){function
i(){p(t,o),r()}function c(){r()}function r(){e(a,i,f),e(u,c,f)}n(a,i,f),n(u,c,f)}(o,t):p(o,t)}}function w(n){["click","mouseo
own","keydown","touchstart","pointerdown"].forEach(function(e){n(e,l,f)})}w(n),self.perfMetrics=self.perfMetrics||{},self.per
fMetrics.onFirstInputDelay=function(n){c.push(n),s()}}(addEventListener,removeEventListener);</script><script async src="http
s://www.googletagmanager.com/gtag/js?id=AW-594691041"></script><script nonce="1170949593194763992">window.dataLayer = window.
dataLayer || [];\n\t\tfunction gtag(){dataLayer.push(arguments)};\n\t\tgtag(\'js\', new Date());</script></head><body><div id
="container"> <div class="_3_C9
Hx"> <img width="75" src="//static-assets-web.flixcart.com/fk-p-linchpin-web/fk-cp-zion/img/f
lipkart-plus_8d85f4.png" alt="Flipkart" title="Flipkart" class="_2xm1JU"/>Explore<!-- -->
Plus<img width="10" src="//static-assets-web.flixcart.com/fk-p-linchpin-web/fk-cp-zion/img/plus_
aef861.png"/> <form class="_2M8clY header-form-search" action="/search" method="GET"><div
class="col-12-12 _2oO9oE"> <input type="text" value="" class="_3704LK" title="Search for products, brands
and more" name="q" autoComplete="off" placeholder="Search for products, brands and more"/> <button class="L0Z3Pu" type
="submit"><svg width="20" height="20" viewBox="0 0 17 18" class="" xmlns="http://www.w3.org/2000/svg"><g fill="#2874F1" fill-
rule="evenodd"><path class="_34RNph" d="m11.618 9.89714.225 4.212c.092.092.101.232.02.313l-1.465 1.46c-.081.081-.221.072-.314
-.02l-4.216-4.203"></path><path class="_34RNph" d="m6.486 10.901c-2.42 0-4.381-1.956-4.381-4.368 0-2.413 1.961-4.369 4.381-4.

```
In [5]: soup = BeautifulSoup(htmlcontent,"html.parser")
```

ASUS LAPTOPS WEB SCRAPPING

2. Access Asus laptops name by using this class.

```
In [ ]: #Access Asus laptops name by using this class.
```

```
In [6]: Laptop = soup.find_all("div",class_ ="_4rR01T")
        Laptop
```

Out[6]:

```
[<div class="_4rR01T">ASUS Vivobook Pro 15 OLED For Creator Ryzen 5 Hexa Core 5600HS - (16 GB/512 GB SSD/Windows 11 Home/4 G...
</div>,
 <div class="_4rR01T">ASUS TUF Gaming F15 - AI Powered Gaming Core i5 11th Gen 11260H - (8 GB/512 GB SSD/Windows 11 Home/4 G...
</div>,
 <div class="_4rR01T">ASUS TUF Gaming F15 with 90WHr Battery Core i5 11th Gen 11400H - (16 GB/512 GB SSD/Windows 11 Home/4 G...
</div>,
 <div class="_4rR01T">ASUS ROG Strix G15 Ryzen 7 Octa Core 6800H - (16 GB/512 GB SSD/Windows 11 Home/4 GB Graphics/NVIDIA Ge...
</div>,
 <div class="_4rR01T">ASUS ROG Strix G15 (2022) Ryzen 7 Octa Core AMD R7-6800H - (16 GB/1 TB SSD/Windows 11 Home/4 GB Graphi...
</div>,
 <div class="_4rR01T">ASUS Vivobook Pro 15 For Creator, Ryzen 5 Hexa Core 5600HS - (8 GB/512 GB SSD/Windows 11 Home/4 GB Gra...
</div>,
 <div class="_4rR01T">ASUS ROG Strix SCAR 15 (2022) Core i9 12th Gen 12900H - (32 GB/1 TB SSD/Windows 11 Home/16 GB Graphics...
</div>,
 <div class="_4rR01T">ASUS Vivobook 16X (2023) For Creator, Intel H-Series Core i7 12th Gen 12650H - (16 GB/1 TB SSD/Windows...
</div>,
 <div class="_4rR01T">ASUS Vivobook Pro 15 For Creator, Ryzen 5 Hexa Core 5600HS - (8 GB/512 GB SSD/Windows 11 Home/4 GB Gra...
</div>,
 <div class="_4rR01T">ASUS Vivobook Pro 15 For Creator, Ryzen 7 Octa Core 5800HS - (16 GB/1 TB SSD/Windows 11 Home/4 GB Grap...
</div>,
 <div class="_4rR01T">ASUS ROG Zephyrus G15 (2022) with 90Whr Battery Ryzen 7 Octa Core 6800HS - (16 GB/1 TB SSD/Windows 11 ...
</div>,
 <div class="_4rR01T">ASUS TUF Gaming F15 - AI Powered Gaming Core i5 11th Gen 11260H - (16 GB/512 GB SSD/Windows 11 Home/4 ...
</div>,
 <div class="_4rR01T">ASUS ROG Strix G15 Advantage Edition with 90Whr Battery Ryzen 9 Octa Core 5980HX - (16 GB/1 TB SSD/Win...
</div>,
 <div class="_4rR01T">ASUS Vivobook Pro 15 For Creator, Ryzen 5 Hexa Core 5600HS - (16 GB/512 GB SSD/Windows 11 Home/4 GB Gr...
</div>,
 <div class="_4rR01T">ASUS ROG Strix G16 (2023), Intel Core i5 13th Gen 13450HX - (16 GB/1 TB SSD/Windows 11 Home/6 GB Graph...
</div>,
 <div class="_4rR01T">ASUS ROG Strix G15 Ryzen 9 Octa Core AMD R9-5900HX - (16 GB/512 GB SSD/Windows 10 Home/4 GB Graphics/N...
</div>,
 <div class="_4rR01T">ASUS TUF Gaming A15 (2023) with 90WHr Battery Ryzen 7 Octa Core 7735HS - (16 GB/512 GB SSD/Windows 11 ...
```

- Output

```
In [9]: list_Laptop=[]
        for i in range(0,len(Laptop)):
            list_Laptop.append(Laptop[i].get_text())
            #print(Mobile_names[i].get_text())
        print(Laptop)
```

```
12900H - (32 GB/1 TB SSD/Windows 11 Home/16 GB Graphics...</div>, <div class="_4rR01T">ASUS Vivobook 16X (2023) For Creator,
Intel H-Series Core i7 12th Gen 12650H - (16 GB/1 TB SSD/Windows...</div>, <div class="_4rR01T">ASUS Vivobook Pro 15 For Crea
tor, Ryzen 5 Hexa Core 5600HS - (8 GB/512 GB SSD/Windows 11 Home/4 GB Gra...</div>, <div class="_4rR01T">ASUS Vivobook Pro 15
For Creator, Ryzen 7 Octa Core 5800HS - (16 GB/1 TB SSD/Windows 11 Home/4 GB Grap...</div>, <div class="_4rR01T">ASUS ROG Zep
hyrus G15 (2022) with 90Whr Battery Ryzen 7 Octa Core 6800HS - (16 GB/1 TB SSD/Windows 11 ...</div>, <div class="_4rR01T">ASU
S TUF Gaming F15 - AI Powered Gaming Core i5 11th Gen 11260H - (16 GB/512 GB SSD/Windows 11 Home/4 ...</div>, <div class="_4r
R01T">ASUS ROG Strix G15 Advantage Edition with 90Whr Battery Ryzen 9 Octa Core 5980HX - (16 GB/1 TB SSD/Win...</div>, <div c
lass="_4rR01T">ASUS Vivobook Pro 15 For Creator, Ryzen 5 Hexa Core 5600HS - (16 GB/512 GB SSD/Windows 11 Home/4 GB Gr...</div
>, <div class="_4rR01T">ASUS ROG Strix G16 (2023), Intel Core i5 13th Gen 13450HX - (16 GB/1 TB SSD/Windows 11 Home/6 GB Grap
h...</div>, <div class="_4rR01T">ASUS ROG Strix G15 Ryzen 9 Octa Core AMD R9-5900HX - (16 GB/512 GB SSD/Windows 10 Home/4 GB
Graphics/N...</div>, <div class="_4rR01T">ASUS TUF Gaming A15 (2023) with 90WHr Battery Ryzen 7 Octa Core 7735HS - (16 GB/512
GB SSD/Windows 11 ...</div>, <div class="_4rR01T">ASUS Vivobook 16X (2023) For Creator, Intel H-Series Core i7 12th Gen 12650
H - (16 GB/512 GB SSD/Windo...</div>, <div class="_4rR01T">ASUS Vivobook Pro 15 For Creator, Ryzen 7 Octa Core 5800HS - (16 G
B/512 GB SSD/Windows 11 Home/4 GB Gr...</div>, <div class="_4rR01T">ASUS Vivobook Pro 15 For Creator, Ryzen 5 Hexa Core 5600H
S - (16 GB/1 TB SSD/Windows 11 Home/4 GB Grap...</div>, <div class="_4rR01T">ASUS TUF Gaming F15 Core i5 12th Gen 12500H - (1
6 GB/512 GB SSD/Windows 11 Home/4 GB Graphics/NVIDIA G...</div>, <div class="_4rR01T">ASUS Vivobook Pro 15 For Creator, Ryzen
7 Octa Core 5800HS - (16 GB/512 GB SSD/Windows 11 Home/4 GB Gr...</div>, <div class="_4rR01T">ASUS TUF Gaming F15 - AI Powere
d Gaming Core i5 11th Gen 11400H - (8 GB/512 GB SSD/Windows 11 Home/4 G...</div>, <div class="_4rR01T">ASUS ROG Strix G16 (20
23) with 90WHr Battery Intel HX-Series Core i9 13th Gen 13980HX - (16 GB/1 TB SS...</div>]
```

ASUS LAPTOPS WEB SCRAPPING

3. Access the pricing during a sale by using this class.

```
In [ ]: #Access the pricing during a sale by using this class.
```

```
In [7]: Laptop_Prices = soup.find_all(class_="_30jeq3 _1_WHN1")
        Laptop_Prices
```

```
Out[7]: [<div class="_30jeq3 _1_WHN1">₹69,990</div>,
         <div class="_30jeq3 _1_WHN1">₹57,990</div>,
         <div class="_30jeq3 _1_WHN1">₹61,990</div>,
         <div class="_30jeq3 _1_WHN1">₹84,990</div>,
         <div class="_30jeq3 _1_WHN1">₹88,990</div>,
         <div class="_30jeq3 _1_WHN1">₹53,990</div>,
         <div class="_30jeq3 _1_WHN1">₹2,10,490</div>,
         <div class="_30jeq3 _1_WHN1">₹88,990</div>,
         <div class="_30jeq3 _1_WHN1">₹53,990</div>,
         <div class="_30jeq3 _1_WHN1">₹73,990</div>,
         <div class="_30jeq3 _1_WHN1">₹1,15,990</div>,
         <div class="_30jeq3 _1_WHN1">₹61,990</div>,
         <div class="_30jeq3 _1_WHN1">₹1,14,990</div>,
         <div class="_30jeq3 _1_WHN1">₹65,990</div>,
         <div class="_30jeq3 _1_WHN1">₹94,990</div>,
         <div class="_30jeq3 _1_WHN1">₹1,04,990</div>,
         <div class="_30jeq3 _1_WHN1">₹99,990</div>,
         <div class="_30jeq3 _1_WHN1">₹99,990</div>,
         <div class="_30jeq3 _1_WHN1">₹69,990</div>,
         <div class="_30jeq3 _1_WHN1">₹67,990</div>,
         <div class="_30jeq3 _1_WHN1">₹85,990</div>,
         <div class="_30jeq3 _1_WHN1">₹64,990</div>,
         <div class="_30jeq3 _1_WHN1">₹57,990</div>,
         <div class="_30jeq3 _1_WHN1">₹2,09,990</div>]
```

- Output

```
In [8]: list_Laptop_Prices=[]

        for i in range(0,len(Laptop_Prices)):
            list_Laptop_Prices.append(Laptop_Prices[i].get_text())

            #print(Mobile_names[i].get_text())
        print(list_Laptop_Prices)
```

```
['₹69,990', '₹57,990', '₹61,990', '₹84,990', '₹88,990', '₹53,990', '₹2,10,490', '₹88,990', '₹53,990', '₹73,990', '₹1,15,990',
'₹61,990', '₹1,14,990', '₹65,990', '₹94,990', '₹1,04,990', '₹99,990', '₹99,990', '₹69,990', '₹67,990', '₹85,990', '₹64,990',
'₹57,990', '₹2,09,990']
```

4. Then access the before discount prices in detail by using this class.

```
In [ ]: #Then access the before discount prices in detail by using this class.
```

```
In [21]: discount = soup.find_all(class_="_3I9_wc _27UcVY")
         discount
```

```
Out[21]: [<div class="_3I9_wc _27UcVY">₹<!-- -->86,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->75,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->89,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->1,17,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->1,35,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->76,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->3,59,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->1,18,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->76,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->1,02,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->1,95,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->80,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->1,97,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->74,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->1,25,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->1,36,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->1,21,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->1,22,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->88,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->85,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->1,03,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->90,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->75,990</div>,
          <div class="_3I9_wc _27UcVY">₹<!-- -->2,51,990</div>]
```

- Output

```
In [22]: list_discount_Prices=[]

         for i in range(0,len(discount)):
             list_discount_Prices.append(discount[i].get_text())

             #print(Mobile_names[i].get_text())
         print(list_discount_Prices)
```

```
['₹86,990', '₹75,990', '₹89,990', '₹1,17,990', '₹1,35,990', '₹76,990', '₹3,59,990', '₹1,18,990', '₹76,990', '₹1,02,990', '₹1,9
5,990', '₹80,990', '₹1,97,990', '₹74,990', '₹1,25,990', '₹1,36,990', '₹1,21,990', '₹1,22,990', '₹88,990', '₹85,990', '₹1,03,99
0', '₹90,990', '₹75,990', '₹2,51,990']
```

ASUS LAPTOPS WEB SCRAPPING

5. Then arrange properly in column all above data by using import pandas library.

```
In [ ]: #Then arrange properly in column all above data by using import pandas library.
```

```
In [24]: import pandas as pd
df=pd.DataFrame({"Laptop_Names":list_Laptop,"Discounted Prices":list_Laptop_Prices,"Before sale":list_discount_Prices})
df
```

Out[24]:

	Laptop_Names	Discounted Prices	Before sale
0	ASUS Vivobook Pro 15 OLED For Creator Ryzen 5 ...	₹69,990	₹86,990
1	ASUS TUF Gaming F15 - AI Powered Gaming Core i...	₹57,990	₹75,990
2	ASUS TUF Gaming F15 with 90WHr Battery Core i5...	₹61,990	₹89,990
3	ASUS ROG Strix G15 Ryzen 7 Octa Core 6800H - (...	₹84,990	₹1,17,990
4	ASUS ROG Strix G15 (2022) Ryzen 7 Octa Core AM...	₹88,990	₹1,35,990
5	ASUS Vivobook Pro 15 For Creator, Ryzen 5 Hexa...	₹53,990	₹76,990
6	ASUS ROG Strix SCAR 15 (2022) Core i9 12th Gen...	₹2,10,490	₹3,59,990
7	ASUS Vivobook 16X (2023) For Creator, Intel H-...	₹88,990	₹1,18,990
8	ASUS Vivobook Pro 15 For Creator, Ryzen 5 Hexa...	₹53,990	₹76,990
9	ASUS Vivobook Pro 15 For Creator, Ryzen 7 Octa...	₹73,990	₹1,02,990
10	ASUS ROG Zephyrus G15 (2022) with 90Whr Batter...	₹1,15,990	₹1,95,990
11	ASUS TUF Gaming F15 - AI Powered Gaming Core i...	₹61,990	₹80,990
12	ASUS ROG Strix G15 Advantage Edition with 90Wh...	₹1,14,990	₹1,97,990
13	ASUS Vivobook Pro 15 For Creator, Ryzen 5 Hexa...	₹65,990	₹74,990
14	ASUS ROG Strix G16 (2023), Intel Core i5 13th ...	₹94,990	₹1,25,990
15	ASUS ROG Strix G15 Ryzen 9 Octa Core AMD R9-59...	₹1,04,990	₹1,36,990
16	ASUS TUF Gaming A15 (2023) with 90WHr Battery ...	₹99,990	₹1,21,990
17	ASUS Vivobook 16X (2023) For Creator, Intel H-...	₹99,990	₹1,22,990
18	ASUS Vivobook Pro 15 For Creator, Ryzen 7 Octa...	₹69,990	₹88,990
19	ASUS Vivobook Pro 15 For Creator, Ryzen 5 Hexa...	₹67,990	₹85,990

s page

6. Then store above all data in CSV file and again open this file in application.

```
In [ ]: # Then store above all data in CSV file and again open this file in application.
```

```
In [26]: df.to_csv("Flipcart_Laptop LaptopNames and Prices.csv",index = False)
```

```
In [27]: df1 = pd.read_csv("Flipcart_Laptop LaptopNames and Prices.csv")
df1
```

Out[27]:

	Laptop_Names	Discounted Prices	Before sale
0	ASUS Vivobook Pro 15 OLED For Creator Ryzen 5 ...	₹69,990	₹86,990
1	ASUS TUF Gaming F15 - AI Powered Gaming Core i...	₹57,990	₹75,990
2	ASUS TUF Gaming F15 with 90WHr Battery Core i5...	₹61,990	₹89,990
3	ASUS ROG Strix G15 Ryzen 7 Octa Core 6800H - (...	₹84,990	₹1,17,990
4	ASUS ROG Strix G15 (2022) Ryzen 7 Octa Core AM...	₹88,990	₹1,35,990
5	ASUS Vivobook Pro 15 For Creator, Ryzen 5 Hexa...	₹53,990	₹76,990
6	ASUS ROG Strix SCAR 15 (2022) Core i9 12th Gen...	₹2,10,490	₹3,59,990
7	ASUS Vivobook 16X (2023) For Creator, Intel H-...	₹88,990	₹1,18,990
8	ASUS Vivobook Pro 15 For Creator, Ryzen 5 Hexa...	₹53,990	₹76,990
9	ASUS Vivobook Pro 15 For Creator, Ryzen 7 Octa...	₹73,990	₹1,02,990
10	ASUS ROG Zephyrus G15 (2022) with 90Whr Batter...	₹1,15,990	₹1,95,990
11	ASUS TUF Gaming F15 - AI Powered Gaming Core i...	₹61,990	₹80,990
12	ASUS ROG Strix G15 Advantage Edition with 90Wh...	₹1,14,990	₹1,97,990
13	ASUS Vivobook Pro 15 For Creator, Ryzen 5 Hexa...	₹65,990	₹74,990
14	ASUS ROG Strix G16 (2023), Intel Core i5 13th ...	₹94,990	₹1,25,990
15	ASUS ROG Strix G15 Ryzen 9 Octa Core AMD R9-59...	₹1,04,990	₹1,36,990
16	ASUS TUF Gaming A15 (2023) with 90WHr Battery ...	₹99,990	₹1,21,990
17	ASUS Vivobook 16X (2023) For Creator, Intel H-...	₹99,990	₹1,22,990
18	ASUS Vivobook Pro 15 For Creator, Ryzen 7 Octa...	₹69,990	₹88,990

s page