NAME : POOJA MAHENDRA MALUSARE

PROJECT NAME : TAB Web Scraping

Project Guide : Prof. Sameer Warsolkar

1.Aim:

The aim of a web scraping project focused on tablet phones is to automatically extract detailed information about various tablet phone models from multiple websites. This includes specifications, prices, reviews, and availability. The goal is to compile this data into a structured format for comparison, market analysis, and consumer insights, ensuring legal and ethical compliance throughout the process.

2.Objectives:

➢ Data Extraction:

- Automate the retrieval of detailed specifications, prices, reviews, and availability of various tablet phone models from multiple e-commerce and review websites.

➢ Data Cleaning and Formatting:

- Clean and standardize the extracted data to ensure consistency and accuracy, converting it into a structured format like CSV, JSON, or a database.

➢ Data Integration:

- Integrate the scraped data with existing datasets or databases to enhance the overall dataset, providing a comprehensive view of the tablet phone market.

➢ Analysis and Reporting:

- Analyze the compiled data to identify trends, price comparisons, consumer preferences, and market insights, generating reports to support decision-making.

➢ Automation and Maintenance:

- Develop automated scripts to regularly update the dataset with new data, ensuring the information remains current and relevant.

➢ Compliance and Ethics:

- Ensure that the web scraping process adheres to legal and ethical standards, including respecting website terms of service and privacy policies. Implement mechanisms to handle any potential issues, such as CAPTCHAs or website blocking.

3. OUTLINE:

From this site, we are going to grab the following information:

1. Tab_Names
2. E-side
3. shipping_cost
4. Tab_price

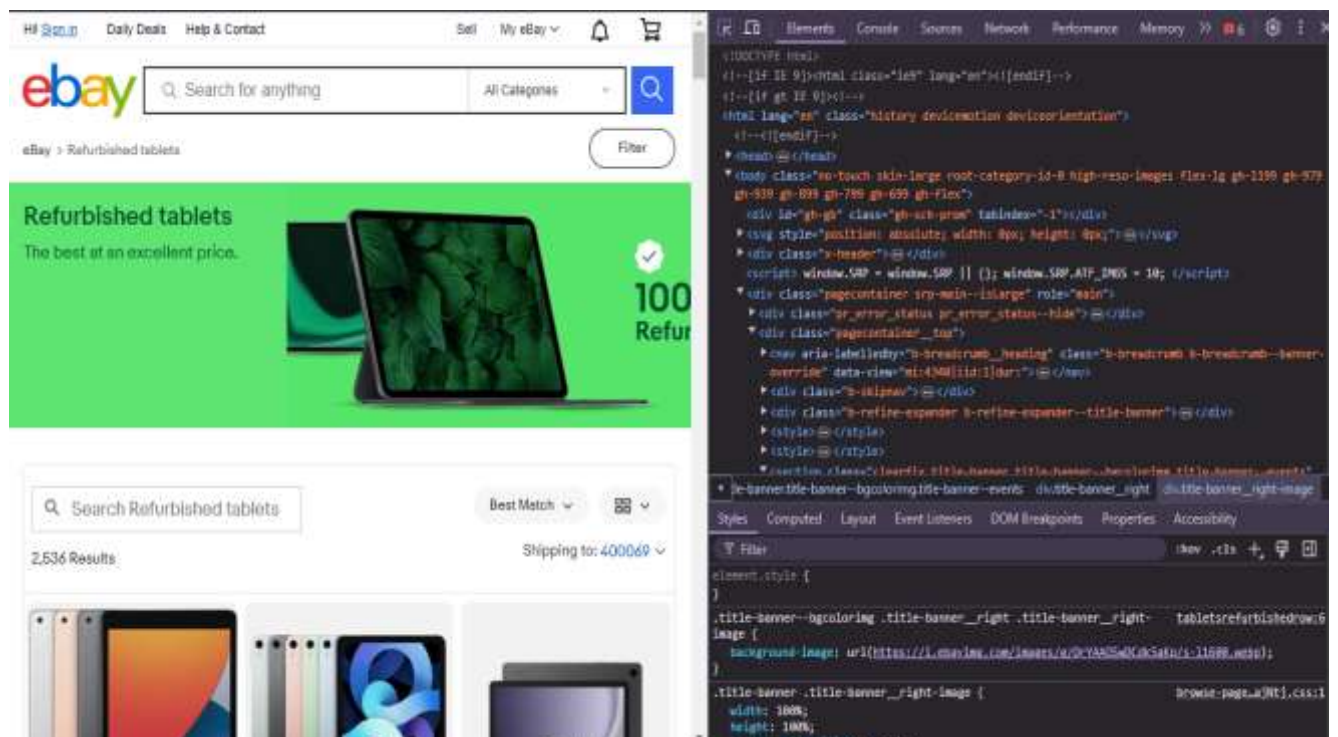# 4. STEPS:

Choose the Website and Webpage URL:

➢ The first step is to select the website you want to scrape. We will try to extract data of Tab.

1. Inspect the website:

➢ Now the next step is to understand the website structure.
Understand what the attributes of the elements that are of your
interest are. Right click on the
website to select "Inspect". This will open HTML code. Use the inspector
tool to see the name of all the elements to use in the code.

. Installing the important libraries:

Python has several web scrapping libraries. We will use the following libraries:

➢ Requests – for making HTTP requests to website

➢ BeautifulSoup – for parsing the HTML code

➢ Pandas – for storing the scraped data in data frame

3. Write the Python source code:

We'll write the main python code. The code will perform the following steps:

➢ Using requests to send an HTTP GET requests

➢ Using BeautifulSoup to parse the HTML code

➢ Extracting the required data from the HTML code

➢ Store the information in a pandas DataFrame

4. Exporting the extracted data:

We'll export the data as a CSV file. We will use the pandas library. We'll use the pandas library.

5. Benefits:

➢ Access to valuable data for analysis or research.

➢ Automation of data collection, saving time and effort.

➢ Stay up to date with change changes on the target websites.

6. Risk:

➢ Legal issues related to web scraping.

➢ Technical challenges due to website changes.

5. Tab web scraping coding command steps

1. Accessing Tab website: .

```
[1]: import requests
     from bs4 import BeautifulSoup
     import pandas as pdbbb
```

```
[2]: page = requests.get("https://www.ebay.com/e/row/tabletsrefurbishedrow")
     soup = BeautifulSoup(page.content, 'html.parser')
     soup
```

```
[2]: <!DOCTYPE html>
<!--[if IE 9]><html class="ie9" lang="en"><![endif]--><!--[if gt IE 9]><!--><html lang="en"><!--<![endif]--><head><meta charset="utf-8"/><meta content
="Refurbished tablets | eBay.  The best at an excellent price." name="twitter:title"><meta content="@eBay" name="twitter:site"/><meta content="http
s://ir.ebaystatic.com/cr/v/c1/ebay-logo-1-1200x1200-margin.png" name="twitter:image"/><meta content="website" property="og:type"/><meta content=" The
best at an excellent price." name="description"/><meta content="https://www.ebay.com/e/row/tabletsrefurbishedrow" property="og:url"/><meta content="no
index" name="robots"/><meta content=" The best at an excellent price." property="og:description"/><meta content="https://ir.ebaystatic.com/cr/v/c1/eba
y-logo-1-1200x630-margin.png" property="og:image"/><meta content=" The best at an excellent price." name="twitter:description"/><meta content="1026282
13125203" property="fb:app_id"/><meta content="summary" name="twitter:card"/><meta content="eBay" property="og:site_name"/><link href="https://i.ebayi
mg.com" rel="preconnect"/><meta content="unsafe-url" name="referrer"/><link href="https://ir.ebaystatic.com" rel="preconnect"/><meta content="34E98E6F
27109BE1A9DCF19658EEEE33" name="msvalidate.01"/><meta content="Refurbished tablets | eBay.  The best at an excellent price." property="og:title"/><met
a content="acf32e2a69cbc2b0" name="y_key"/><title>Refurbished tablets | eBay.  The best at an excellent price.</title><meta content="8kHr3jd3Z43q1ovwo
0KVgo_NZKIEMjthBxti8m8fYTg" name="google-site-verification"/><meta content="6e11485a66d91eff" name="yandex-verification"/><!-- NGMARS SIGNATURE --><li
nk href="//i1.ebayimg.com" rel="dns-prefetch"/><link crossorigin="" href="//i.ebayimg.com" rel="preconnect"/><link crossorigin="" href="//i1.ebayimg.c
om" rel="preconnect"/><link crossorigin="" href="//ir.ebaystatic.com" rel="preconnect"/><link crossorigin="" href="//secureir.ebaystatic.com" rel="pre
connect"/><meta content="width=device-width, initial-scale=1, minimum-scale=1" name="viewport"/><link href="//ir.ebaystatic.com" rel="dns-prefetch"/><
link href="//secureir.ebaystatic.com" rel="dns-prefetch"/><link href="//i.ebayimg.com" rel="dns-prefetch"/><link href="//rover.ebay.com" rel="dns-pref
etch"/><script>$ssgST=new Date().getTime();</script><link href="https://ir.ebaystatic.com/rs/c/browse-page-desktop-pr-h5wajNtj.css" rel="stylesheet"/>
<link href="https://ir.ebaystatic.com/rs/v/vncih4uck22i1nu1dll4knjunv2.css?prcr=B1:N" rel="stylesheet" type="text/css"/><!--H$s0-4--><noscript data-ma
```

2. Using BeautifulSoup :


```
[2]: page = requests.get("https://www.ebay.com/e/row/tabletsrefurbishedrow")
     soup = BeautifulSoup(page.content, 'html.parser')
     soup
```

```
[2]: <!DOCTYPE html>
<!--[if IE 9]><html class="ie9" lang="en"><![endif]--><!--[if gt IE 9]><!--><html lang="en"><!--<![endif]--><head><meta charset="utf-8"/><meta content
="Refurbished tablets | eBay.  The best at an excellent price." name="twitter:title"><meta content="@eBay" name="twitter:site"/><meta content="http
s://ir.ebaystatic.com/cr/v/c1/ebay-logo-1-1200x1200-margin.png" name="twitter:image"/><meta content="website" property="og:type"/><meta content=" The
best at an excellent price." name="description"/><meta content="https://www.ebay.com/e/row/tabletsrefurbishedrow" property="og:url"/><meta content="no
index" name="robots"/><meta content=" The best at an excellent price." property="og:description"/><meta content="https://ir.ebaystatic.com/cr/v/c1/eba
y-logo-1-1200x630-margin.png" property="og:image"/><meta content=" The best at an excellent price." name="twitter:description"/><meta content="1026282
13125203" property="fb:app_id"/><meta content="summary" name="twitter:card"/><meta content="eBay" property="og:site_name"/><link href="https://i.ebayi
mg.com" rel="preconnect"/><meta content="unsafe-url" name="referrer"/><link href="https://ir.ebaystatic.com" rel="preconnect"/><meta content="34E98E6F
27109BE1A9DCF19658EEEE33" name="msvalidate.01"/><meta content="Refurbished tablets | eBay.  The best at an excellent price." property="og:title"/><met
a content="acf32e2a69cbc2b0" name="y_key"/><title>Refurbished tablets | eBay.  The best at an excellent price.</title><meta content="8kHr3jd3Z43q1ovwo
0KVgo_NZKIEMjthBxti8m8fYTg" name="google-site-verification"/><meta content="6e11485a66d91eff" name="yandex-verification"/><!-- NGMARS SIGNATURE --><li
nk href="//i1.ebayimg.com" rel="dns-prefetch"/><link crossorigin="" href="//i.ebayimg.com" rel="preconnect"/><link crossorigin="" href="//i1.ebayimg.c
om" rel="preconnect"/><link crossorigin="" href="//ir.ebaystatic.com" rel="preconnect"/><link crossorigin="" href="//secureir.ebaystatic.com" rel="pre
connect"/><meta content="width=device-width, initial-scale=1, minimum-scale=1" name="viewport"/><link href="//ir.ebaystatic.com" rel="dns-prefetch"/><
link href="//secureir.ebaystatic.com" rel="dns-prefetch"/><link href="//i.ebayimg.com" rel="dns-prefetch"/><link href="//rover.ebay.com" rel="dns-pref
etch"/><script>$ssgST=new Date().getTime();</script><link href="https://ir.ebaystatic.com/rs/c/browse-page-desktop-pr-h5wajNtj.css" rel="stylesheet"/>
<link href="https://ir.ebaystatic.com/rs/v/vncih4uck22i1nu1dll4knjunv2.css?prcr=B1:N" rel="stylesheet" type="text/css"/><!--H$s0-4--><noscript data-ma
```

3. Accessing the Tab_Names:

4. Accessing the Tab_Prices:

5. Accessing the Tab_shipping:

```
[8]: shipping_tags = soup.select(".s-item__shipping.s-item__logisticsCost")
     shipping = []
     for tag in shipping_tags:
         text = tag.get_text()
         cleaned_text = text.replace('$', '').strip()   # Remove the dollar symbol and extra whitespace
         shipping.append(cleaned_text)

     print("shipping:", shipping)
```

shipping: ['30.41 shipping', '34.86 shipping', '66.43 shipping', '79.38 shipping', '91.58 shipping', '26.48 shipping', '28.64 shipping', '32.18 shippin
g', '41.41 shipping', '28.64 shipping', '28.37 shipping', '41.06 shipping', '31.17 shipping', '36.17 shipping', '28.94 shipping', '38.08 shipping', '86.4
5 shipping', '64.79 shipping', '65.25 shipping', '35.22 shipping', '28.49 shipping', '59.07 shipping', '31.59 shipping', '30.85 shipping', '36.94 shippin
g', '29.22 shipping', '87.22 shipping', '28.76 shipping', '34.83 shipping', '32.47 shipping', '31.88 shipping', '76.32 shipping', '28.93 shipping', '38.3
7 shipping', '35.71 shipping', '42.44 shipping', '65.99 shipping', '67.02 shipping', '67.31 shipping', '28.93 shipping', '59.50 shipping', '29.20 shippin
g', '28.59 shipping', '30.12 shipping', '27.12 shipping', '29.99 shipping', '30.32 shipping', '31.03 shipping', '28.31 shipping', '30.99 shipping', '31.1
7 shipping', '35.84 shipping', '27.25 shipping', '38.28 shipping', '38.90 shipping', '34.92 shipping', '38.32 shipping', '31.23 shipping', '33.86 shippin
g', '32.44 shipping']

```
[9]: len(shipping)
```

[9]: 60

6. Accessing the Tab_E-Side :-

```
[10]: eBay_Refurbished_tags=soup.select(".s-item__certified-refurbished.s-item__certified-refurbished--isLarge")
      eBay_Refurbished=[]
      for i in eBay_Refurbished_tags:
          eBay_Refurbished.append(i.get_text())
      print(eBay_Refurbished)
```

['eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurb
ished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay
Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished',
'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbi
shed', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay
Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished',
'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbi
shed', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished', 'eBay Refurbished']

```
[11]: len(eBay_Refurbished)
```

[11]: 60

7.Importing the pandas and creating DataFrame

```
[12]:  import pandas as pd
       tab = pd.DataFrame({
           "Tab_Name": name,
           "E-side": eBay_Refurbished,
           "shipping_cost": shipping,
           "Tab_price":price
       })
       tab
```

[12]:		Tab_Name	E-side	shipping_cost	Tab_price
	0	Apple iPad 8 (8th Gen) - (2020 Model) - 32GB ...	eBay Refurbished	30.41 shipping	169.99
	1	Apple iPad Air 4 (2020) 10.9" 64GB 256GB All C...	eBay Refurbished	34.86 shipping	320.99 to 401.99
	2	Samsung Galaxy Tab A9+ SM-X210N WIFI 128GB Gra...	eBay Refurbished	66.43 shipping	159.99
	3	Apple iPad Pro 11-inch (4th Gen) (2022) M2 Chi...	eBay Refurbished	79.38 shipping	599.00
	4	Microsoft Surface Pro 9 13" Touch Tablet, Inte...	eBay Refurbished	91.58 shipping	868.00
	5	CTL Chromebook Touch Tablet Tx1 32GB Chrome Ta...	eBay Refurbished	26.48 shipping	36.95
	6	Apple iPad 5th Gen 9.7" 32GB 128GB Gray Gold S...	eBay Refurbished	28.64 shipping	108.99 to 137.99
	7	Google Pixel Tablet 128GB Porcelain WiFi Excel...	eBay Refurbished	32.18 shipping	229.99
	8	Apple iPad Pro 3 M1 (2021) 11" 128GB 256GB 512...	eBay Refurbished	41.41 shipping	542.99 to 746.99

8. Converting and Storing DataFrame in the form of a CSV file and opening the file in application:

```
[13]:  tab.to_csv("tab_Report_Scraping.csv",index=False)
```

```
[14]:  Read=pd.read_csv("tab_Report_Scraping.csv")
       Read
```

[14]:		Tab_Name	E-side	shipping_cost	Tab_price
	0	Apple iPad 8 (8th Gen) - (2020 Model) - 32GB ...	eBay Refurbished	30.41 shipping	169.99
	1	Apple iPad Air 4 (2020) 10.9" 64GB 256GB All C...	eBay Refurbished	34.86 shipping	320.99 to 401.99
	2	Samsung Galaxy Tab A9+ SM-X210N WIFI 128GB Gra...	eBay Refurbished	66.43 shipping	159.99
	3	Apple iPad Pro 11-inch (4th Gen) (2022) M2 Chi...	eBay Refurbished	79.38 shipping	599.00
	4	Microsoft Surface Pro 9 13" Touch Tablet, Inte...	eBay Refurbished	91.58 shipping	868.00
	5	CTL Chromebook Touch Tablet Tx1 32GB Chrome Ta...	eBay Refurbished	26.48 shipping	36.95
	6	Apple iPad 5th Gen 9.7" 32GB 128GB Gray Gold S...	eBay Refurbished	28.64 shipping	108.99 to 137.99
	7	Google Pixel Tablet 128GB Porcelain WiFi Excel...	eBay Refurbished	32.18 shipping	229.99
	8	Apple iPad Pro 3 M1 (2021) 11" 128GB 256GB 512...	eBay Refurbished	41.41 shipping	542.99 to 746.99

6.Conclusion

In conclusion, a web scraping project focused on tablet phones aims to systematically collect detailed and up-to-date information on various models from multiple online sources. By automating the extraction, cleaning, and integration of data, the project ensures a comprehensive and structured dataset that can be used for market analysis, price comparison, and consumer insights. Through adherence to legal and ethical standards, the project not only enhances data collection efficiency but also supports informed decision-making in the tablet phone market.

THANK YOU