**Step 1: Import Required Libraries**

import requests # To fetch the webpage

from bs4 import BeautifulSoup # To parse HTML

import pandas as pd # To store data in tabular format

* requests: Sends HTTP requests to fetch web pages.
* BeautifulSoup: Extracts specific data from HTML.
* pandas: Used to create a table (DataFrame) and save data to a file.

**Step 2: Define Website URL**

url = "http://books.toscrape.com/"

* We will **scrape book details** from this page.

**Step 3: Fetch and Parse the Webpage**

response = requests.get(url) # Send request

soup = BeautifulSoup(response.text, "html.parser") # Parse HTML

* Fetches the **HTML content** of the page.
* BeautifulSoup(response.text, "html.parser") parses the content into a **structured format**.

**Step 4: Find All Book Listings**

books = soup.find\_all("article", class\_="product\_pod")

* Each book is inside an <article> tag with the class "product\_pod".
* find\_all() gets **all book elements** in a list.

**Step 5: Create an Empty List**

book\_data = []

* We will store **book details (title, price, availability)** in this list.

**Step 6: Loop Through Books & Extract Data**

for book in books:

title = book.h3.a["title"] # Get book title

price = book.select\_one(".price\_color").text # Get book price

availability = book.select\_one(".availability").text.strip() # Get availability status

book\_data.append([title, price, availability]) # Store in list

* **Extracts book details:**
  + book.h3.a["title"]: Gets the **title** from the <a> tag inside <h3>.
  + book.select\_one(".price\_color").text: Gets **price** from <p class="price\_color">.
  + book.select\_one(".availability").text.strip(): Gets **availability status**.
* Appends the details to the book\_data list.

**Step 7: Convert List to Pandas DataFrame**

df = pd.DataFrame(book\_data, columns=["Title", "Price", "Availability"])

* **Converts scraped data into a table** (DataFrame) with columns: "Title", "Price", "Availability".

**Step 8: Save Data to CSV**

df.to\_csv("books.csv", index=False)

* **Saves the scraped data as a CSV file** (books.csv).
* index=False prevents writing row numbers.