TASK 1: WEB SCRAPING

Introduction:
Web Scraping is a technique used to extract large amounts of data from websites automatically. This data is usually unstructured and is converted into structured form for analysis.

Objectives of Web Scraping:
• Collect real-time data from public websites
• Create custom datasets for analysis
• Reduce manual data collection effort

Tools Used:
• Python (Requests, BeautifulSoup, Scrapy)
• No-code tools: Octoparse, ParseHub

Process of Web Scraping:
1. Identify the target website and dataset
2. Study the HTML structure using browser developer tools
3. Send HTTP request to fetch web page
4. Parse HTML and extract required elements
5. Store the extracted data in CSV/Excel/JSON format

Python Code for Web Scraping:

```
import requests
from bs4 import BeautifulSoup
import csv

url = "https://example.com"
response = requests.get(url)

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

titles = soup.find_all("h2")

with open("output.csv", "w", newline="") as file:
    writer = csv.writer(file)
    writer.writerow(["Title"])

    for title in titles:
        writer.writerow([title.text.strip()])

print("Data scraped successfully")
```