**Approach**

* Sent an HTTP request using the **Requests** library to fetch the HTML content from the target URL.
* Parsed the HTML using **BeautifulSoup** to extract:
  + Headings (h1, h2)
  + Paragraphs (p)
  + Lists (li)
  + Table data (tr, td), grouping table cells into rows.
* Stored the extracted data into multiple output files:
  + **Two CSV files** for different data formats:
    1. Extract\_Text\_Data.csv – combined data in a structured table.
    2. Extract\_Table\_Data.csv – table data only.
  + **Three JSON files**:
    1. Product\_Information.JSON– combined Book Title ,Price, Stock Availability, and Button text.
    2. Form\_Information.JSON – combined Field name , Input type and Default values.
    3. Video\_Link.JSON– link of video.

**Tools Used**

* **Python**
* **Requests** – fetch webpage content
* **BeautifulSoup** – parse HTML and extract elements
* **CSV module** – save structured data in CSV files
* **JSON module** – store extracted data in JSON files

**Challenges Faced**

* Handling different HTML structures for headings, lists, tables, and form elements.
* Extracting **default values** from form fields that may not always have a value.
* Ensuring proper grouping of table rows when saving to CSV and JSON.
* Cleaning and formatting text to remove extra spaces or newline characters.