**Lab Experiment 1**

**Data Collection using Web Scraping**

**Objective:**

To understand and practice the basics of web scraping using Python's Beautiful Soup library for extracting data from web pages.

**Prerequisites Library :** Requests, BeautifulSoup, Pandas, CSV

**Step 1 :** Import Required Libraries

**Step 2:** Send a GET Request to the Website (choose any website of your choice)

**Step 3:** Parse the HTML Content

**Step 4:** Inspect the Web Page

**Step 5:** Extract Data Using Beautiful Soup

**Step 6:** Convert the extracted data into a pandas dataframe

**Step 7:** Save the content into a CSV

**Note:** Always check the website's `robots.txt` file to ensure your scraping complies with the site's terms of service.

Link to the official Beautifulsoup library documentation to know all the various functions and its input arguments - [Click Here](https://www.crummy.com/software/BeautifulSoup/bs4/doc/)

**Submission Guidelines:**

1. Submit the assignment as a .ipynb (Jupyter Notebook) file or .py script. **Include a .pdf export** of the notebook / script. ( [Click Here](https://www.analyticsvidhya.com/blog/2024/08/ipynb-files-to-pdf/) If you do not know how to export PDF from .ipynb notebook)
2. Include comments in your code to explain your logic and approach.
3. Use Markdown cells for detailed explanations, if necessary, especially for complex code blocks
4. Ensure that all code cells have been executed and the outputs are visible.
5. Do not clear the outputs before submission.
6. Submit the assignment via the designated Google Classroom link.
7. Submit your work before the deadline.
8. Late submissions may incur penalties
9. Avoid Plagiarism; ensure the work you submit is your own.