Class Task: From HTML to Insight – Scrape, Clean & Analyze!

# Objective

Students will:  
- Scrape a new website   
- Clean and organize the data using Pandas  
- Perform basic analysis  
- Answer reflective questions

# Instructions

1. Choose any structured website to scrape (not BooksToScrape).  
 Examples: quotes.toscrape.com, news.ycombinator.com, webscraper.io/test-sites  
2. Use Inspect Element to find tag and class names of desired elements (quote, author, price, etc.)  
3. Use Python with BeautifulSoup and Requests to scrape the data.  
4. Store your data in a Pandas DataFrame.  
5. Apply at least two cleaning steps (e.g., remove special characters, trim whitespace, filter data).  
6. Perform the analysis questions listed below.  
7. Submit your code and answers.

# Data Analysis Questions

1. What’s the most frequent item?  
 (e.g., Most quoted author, repeated product, most common rating)  
  
2. What’s the average of a numerical column?  
 (e.g., average product price, average quote length)  
  
3. Are there any missing or unusual values?  
 (e.g., Empty titles, null authors, unknown prices?)  
  
4. Visualize one insight using matplotlib or pandas.plot()  
 (Optional but encouraged)  
  
5. Your Reflection (write answers):  
 a. What challenges did you face while scraping this website?  
 b. Was the HTML easy or hard to navigate?  
 c. How can this scraped data be useful in real life?