

Methodology:

I have used Selenium to load the website homepage because it is JavaScript-rendered. Use `time.sleep(5)` while scraping data from CNBC and `WebDriverWait` to ensure all articles load properly. I parsed the HTML using BeautifulSoup and extract title using `<h2>`, the URL from `href`, the description from `div.article__chapo` and Date either from the HTML(if available) or code execution date. In the last export data into excel file with the required field.

Learning Outcomes:

I have learned the use of Selenium with headless Chrome to scrape JavaScript-heavy websites. I also learned how to parse HTML using BeautifulSoup to extract specific elements from the web page. I used Webdriver Manager to avoid hardcoding paths. Additionally, Learned to save scraped data in Excel format using pandas and Improved debugging and inspection skills by exploring website HTML structure.