

Web scraping with Python using Requests and BeautifulSoup

Objective: To visit the public website (like Reuters) which has a description of Fortune 500 companies, and their board members and extract details(name,profile etc) about the board members of these companies.

Example :

Apple: (<https://www.reuters.com/markets/companies/AAPL.OQ>)

Amazon: (<https://www.reuters.com/markets/companies/AMZN.OQ/>)

What you have to do in this project, is you need to scrape data about the company founders and the current board members, and store it in the form of a table with two columns with headings Member and Job Profile.

Steps:

- 1.Go to the above website and try to identify the tags in which the targeted text is stored
- 2.Connect to the URLs. using requests.
3. Store the response that you get in a response object.
4. Use BeautifulSoup to parse the html that is sent back.
5. Using the various functions present in BeautifulSoup traverse the html tree
6. Extract the targeted tags using find or find_all functions
7. Use the .text() and .strip() functions to extract the text from the tags
8. Create a DataFrame using pandas.
9. Save your DataFrame to a csv file using the .to_csv() function.