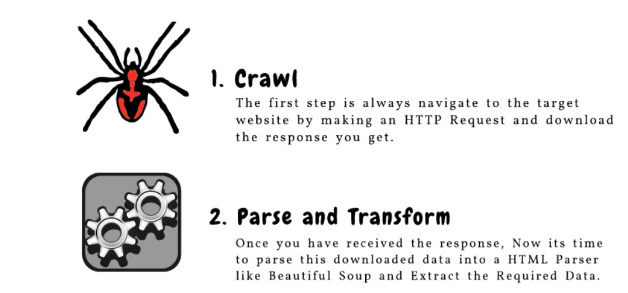
**WEB SCRAPPING**

Taking out data from the website, it helps in taking out the unstructured data from HTML into structured data in our spreadsheet**TWO COMPONENT OF WEB SCRAPPING**

**First install the BEAUTIFULSOUP extention in juyter or vs code, so that BEAUTIFULSOUP library work**

**Method BEAUTIFULSOUP to extract data from website**

Importing request or asking permission from website to extract data



**Importing BEAUTIFULSOUP library**



**Adding website URL from which we are extracting data**



**Sending an HTTP request to the URL**



**Checking if the request was successful (by default status code is 200) with FOR LOOP method**



**Once we had response from the website then we need to parser the HTML content**

**Here:-** parser means extracting data line by line

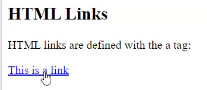


**Extracting information from the page**

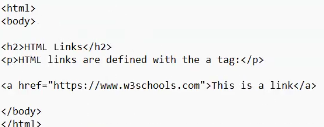
**Here:-** extracting the ANCHOR TAG(extracting all the links) ‘a’ stands for ANCHOR TAG.



**((((((( ANCHOR TAGS** => and convert this link into a msg and when ever we clcik on that msg link will get ooen



We can also write HTML code is notepad and can also open in browser.



By pressing CTRL+O to open the saved html notepad in browser and then we can click over msg in which link is saved to open the website**))))))))))))**

**Now printing the extracted link**

**Here**:-‘href’ is syntax for mentioning the links from website.



**If unable to retrieve the page data print (failed to retrieve the data)**

**Whole coding will look like this**

**INPUT**



**OUTPUT**

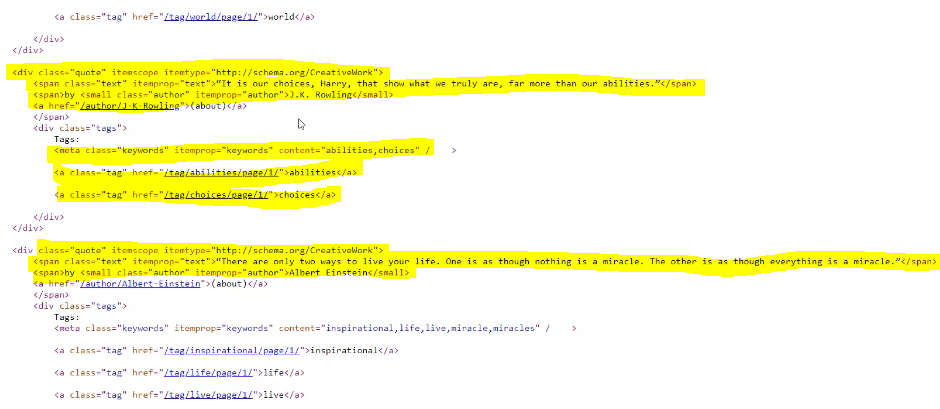


**Another example**

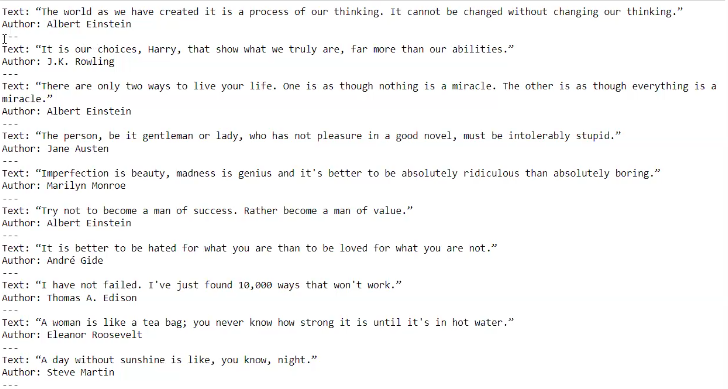
**Here:- we are taking out the quotations along with its author information, ‘span’ and ‘small’Is a tag like anchor tag**



Right click on website and click on ‘inspect’ to have website ‘html’ view

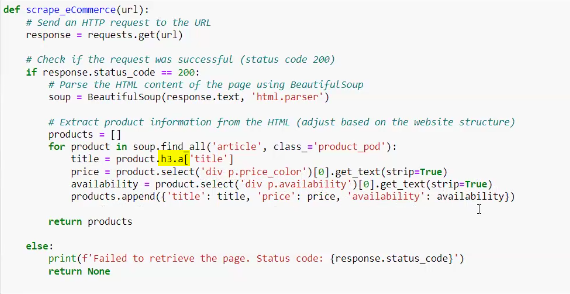


**Output**



**Example3**

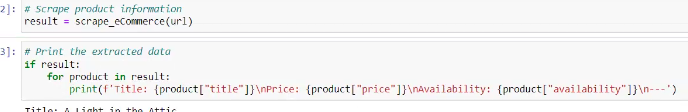
Extracting product information from HTML

**Here**:- h3 is heading of title

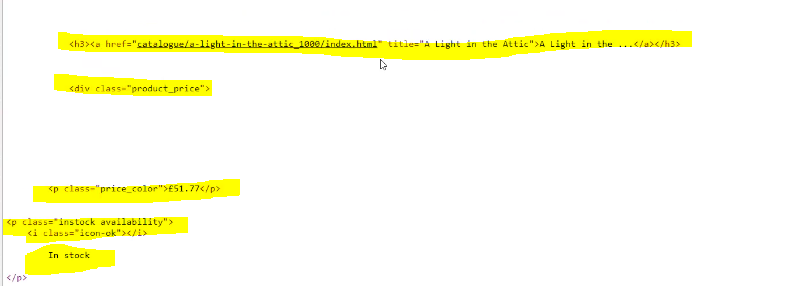
Passing website URL



Scrap project information and print extracted data



Html view



Output

