

Web scraping example

Version 1.0, 2024/08/07

This notebook provides an example about the use of web scraping to retrieve price data.

Specifically, it uses the test website [Books to scrape](#) to demonstrate how to gather data from a website, including navigation through the various pages. We used the browser inspection functions to analyze the HTML structure of the webpages and define the paths for data extractions. The techniques and style used here are only an example, you could achieve the same result in many different ways. Also the structure is specific for this website, you will need to adapt this code for others.

There are additional resources you may explore in order to get some practice with different website architectures before applying your efforts to real websites:

- [Web Scraper test sites](#)
- [web-scraping.dev](#)

You can run this notebook on your own environment, provided you install the Python packages listed in the requirements.txt file. You can also run this notebook on a cloud environment like Google Colab. The minimum version of Python tested with this notebook is 3.9, we suggest at least 3.10 or newer.

Please be considerate in your web scraping operations and always include a delay between calls to avoid overloading the source website.