

Module 11

The Week Ahead

In this module, you'll first automate a web browser to **scrape**, or extract, data about the National Aeronautics and Space Administration (NASA) Mars exploration. You'll then visualize and analyze the data.

With all the information that the internet makes available, people and businesses have access to an overwhelming amount of data. But, that data isn't always formatted in a way that makes it easy to analyze.

Enter web scraping. Web scraping automates the tedious tasks of extracting online data for analysis. Instead of a person manually visiting each website, copying the data, and then pasting that data into a file, a web-scraping script automatically performs all those actions and, if necessary, organizes the data for analysis. That is, instead of us manually collecting and formatting the data, web scraping automates the process for us.

Web scraping has a wide range of applications. Worldwide, large companies use web scraping to assess their reputations or to track the online presence of their competitors. People can also use web scraping on a smaller scale for personal projects. For example, web scraping can simplify the process of collecting the current news about a particular subject.

What You Will Learn

By the end of this module, you'll be able to:

- Describe basic HyperText Markup Language (HTML) elements.
- Explain how websites use HTML to structure a webpage.
- Explain how Cascading Style Sheets (CSS) uses the `class` and `id` attributes to style HTML elements and identify the components of a webpage for scraping.
- Use BeautifulSoup and Splinter to both automate a web browser and scrape data.
- Visualize and analyze scraped data by using Python tools.

Planning Your Schedule

Here's an overview of the lessons and assignments that you'll cover in this module. You can use the time estimates to help pace your learning and plan your schedule.

- Introduction (15 minutes)
- Getting Started with Using Web Scraping Tools (1 hour)
- How HTML and CSS Build a Webpage (2 hours)
- Styling HTML Elements with CSS (2 hours)
- Using Chrome Developer Tools (30 minutes)
- Automate a Web Browser and Perform a Web Scrape (2 hours)
- More Adventures in Web Scraping (1 hour)

© 2022 edX Boot Camps LLC