

MODULE PRACTICE

Beautiful Soup Traversal

BeautifulSoup is a Python library used to parse and traverse an HTML page. BeautifulSoup can scrape webpage data and collect it in a form suitable for data analytics.

Beautiful Soup Object and Methods

```
# This line of code creates a BeautifulSoup object from
a webpage:

soup = BeautifulSoup(webpage.content, "html.parser")

# Within the `soup` object, tags can be called by name:

first_div = soup.div

# or by CSS selector:

all_elements_of_header_class = soup.select(".header")

# or by a call to `.find_all`:

all_p_elements = soup.find_all("p")
```

BeautifulSoup uses a parser to take in the content of a webpage. It provides tree traversal and advanced searching methods. It creates an object from the website contents.