

# BeautifulSoup and Requests

```
In [1]: from bs4 import BeautifulSoup
import requests
```

```
In [2]: url = 'https://www.scrapethissite.com/pages/forms/'
```

```
In [4]: page = requests.get(url)
```

```
In [6]: soup = BeautifulSoup(page.text, 'html')
```

```
In [7]: print(soup)
```

```
<!DOCTYPE html>
<html lang="en">
<head>
<meta charset="utf-8"/>
<title>Hockey Teams: Forms, Searching and Pagination | Scrape This Site | A
public sandbox for learning web scraping</title>
<link href="/static/images/scrapper-icon.png" rel="icon" type="image/png"/>
<meta content="width=device-width, initial-scale=1.0" name="viewport"/>
<meta content="Browse through a database of NHL team stats since 1990. Prac
tice building a scraper that handles common website interface components."
name="description"/>
<link crossorigin="anonymous" href="https://maxcdn.bootstrapcdn.com/bootstr
ap/3.3.5/css/bootstrap.min.css" integrity="sha256-MfvZlkHCEqatNoGiOXveE8FIw
MzZg4W85qfrfIFBfYc= sha512-dTfge/zgoMP7QbHy4gWMEGsbdsZeCXz7irItjcC3sPUFtf
0kuFbDz/ixG7ArTxmDjLXDmezHubeNikyKGVyQ==" rel="stylesheet"/>
<link href="https://fonts.googleapis.com/css?family=Lato:400,700" rel="styl
esheet" type="text/css"/>
<link href="/static/css/styles.css" rel="stylesheet" type="text/css"/>
<meta content="noindex" name="robots"/>
```

In [8]: `print(soup.prettify())`

```
<!DOCTYPE html>
<html lang="en">
  <head>
    <meta charset="utf-8"/>
    <title>
      Hockey Teams: Forms, Searching and Pagination | Scrape This Site | A public
      sandbox for learning web scraping
    </title>
    <link href="/static/images/scraper-icon.png" rel="icon" type="image/png"/>
  >
    <meta content="width=device-width, initial-scale=1.0" name="viewport"/>
    <meta content="Browse through a database of NHL team stats since 1990. Practice
    building a scraper that handles common website interface components." name="description"/>
    <link crossorigin="anonymous" href="https://maxcdn.bootstrapcdn.com/bootstrap/3.3.5/css/bootstrap.min.css" integrity="sha256-MfvZlkHCEqatNoGiOXveE8FIwMzZg4W85qfrfIFBfYc= sha512-dTfge/zgoMYpP7QbHy4gWMEGsbdsZeCXz7irItjcC3sPUFtf0kuFbDz/ixG7ArTxmDjLXDmezHubeNikyKGVyQ==" rel="stylesheet"/>
    <link href="https://fonts.googleapis.com/css?family=Lato:400,700" rel="stylesheet"/>
```

In [ ]: