

Find and Find_all

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

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

```
In [3]: page = requests.get(url)
```

```
In [4]: soup = BeautifulSoup(page.text, 'html')
```

```
In [5]: 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/zgoMYpP7QbHy4gWMEGsbdsZeCXz7irItjcC3sPUFtf
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 [6]: soup.find('div')
```

```
Out[6]: <div class="container">
<div class="col-md-12">
<ul class="nav nav-tabs">
<li id="nav-homepage">
<a class="nav-link hidden-sm hidden-xs" href="/">

        Scrape This Site
    </a>
</li>
<li id="nav-sandbox">
<a class="nav-link" href="/pages/">
<i class="glyphicon glyphicon-console hidden-sm hidden-xs"></i>
        Sandbox
    </a>
</li>
<li id="nav-lessons">
<a class="nav-link" href="/lessons/">
<i class="glyphicon glyphicon-education hidden-sm hidden-xs"></i>
        Lessons
    </a>
</li>
<li id="nav-faq">
<a class="nav-link" href="/faq/">
<i class="glyphicon glyphicon-flag hidden-sm hidden-xs"></i>
        FAQ
    </a>
</li>
<li class="pull-right" id="nav-login">
<a class="nav-link" href="/login/">
        Login
    </a>
</li>
</ul>
</div>
</div>
```

In [7]: `soup.find_all('div')`

Out[7]:

```
[<div class="container">
  <div class="col-md-12">
    <ul class="nav nav-tabs">
      <li id="nav-homepage">
        <a class="nav-link hidden-sm hidden-xs" href="/">
          
            Scrape This Site
          </a>
        </li>
        <li id="nav-sandbox">
          <a class="nav-link" href="/pages/">
            <i class="glyphicon glyphicon-console hidden-sm hidden-xs"></i>
              Sandbox
            </a>
          </li>
          <li id="nav-lessons">
            <a class="nav-link" href="/lessons/">
              <i class="glyphicon glyphicon-education hidden-sm hidden-xs"></i>
                Lessons
              </a>
            </li>
          </ul>
        </div>
      </div>
    ]
```

```
In [8]: soup.find_all('div', class_ = 'col-md-12')
```

```

Out[8]: [<div class="col-md-12">
  <ul class="nav nav-tabs">
    <li id="nav-homepage">
      <a class="nav-link hidden-sm hidden-xs" href="/">
        
          Scrape This Site
        </a>
      </li>
      <li id="nav-sandbox">
        <a class="nav-link" href="/pages/">
          <i class="glyphicon glyphicon-console hidden-sm hidden-xs"></i>
            Sandbox
          </a>
        </li>
      <li id="nav-lessons">
        <a class="nav-link" href="/lessons/">
          <i class="glyphicon glyphicon-education hidden-sm hidden-xs"></i>
            Lessons
          </a>
        </li>
      <li id="nav-faq">
        <a class="nav-link" href="/faq/">
          <i class="glyphicon glyphicon-flag hidden-sm hidden-xs"></i>
            FAQ
          </a>
        </li>
      <li class="pull-right" id="nav-login">
        <a class="nav-link" href="/login/">
          Login
        </a>
      </li>
    </ul>
  </div>,
  <div class="col-md-12">
    <h1>
      Hockey Teams: Forms, Searching and Pagination
      <small>25 items</small>
    </h1>
    <hr/>
  </div>,
  <div class="col-md-12">
    <p class="lead">
      Browse through a database of NHL team stats since 1990. Practice building a scraper that handles common website interface components.
      Take a look at how pagination and search elements change the URL as you browse. Build a web scraper that can conduct searches and paginate through the results.
    </p>
    <hr/>
  </div>,
  <div class="col-md-12 text-center">
    <form action="/pages/forms/" class="form form-inline" method="GET">
      <label for="q">
        Search for Teams:
      </label>
      <input class="form-control" id="q" name="q" placeholder="Search for Teams" type="text">
    </form>
  </div>
  </div>]

```

```
type="text"/>
<input class="btn btn-primary" type="submit" value="Search"/>
</form>
<hr/>
</div>,
<div class="col-md-12 text-center text-muted">
    Lessons and Videos © Hartley Brody 2023
</div>]
```

```
In [9]: soup.find_all('p', class_='lead')
```

```
Out[9]: [<p class="lead">
```

Browse through a database of NHL team stats since 1990. Practice building a scraper that handles common website interface components.

Take a look at how pagination and search elements change the URL as you browse. Build a web scraper that can conduct searches and paginate through the results.

```
</p>]
```

```
In [11]: soup.find('p', class_='lead').text.strip()
```

```
Out[11]: 'Browse through a database of NHL team stats since 1990. Practice building a
scraper that handles common website interface components.\n
Take a look at how pagination and search elements change the URL as you browse. Build a web scraper that can conduct searches and paginate through the results.'
```

```
In [12]: soup.find_all('th')
```

```
Out[12]: [<th>
            Team Name
        </th>,
        <th>
            Year
        </th>,
        <th>
            Wins
        </th>,
        <th>
            Losses
        </th>,
        <th>
            OT Losses
        </th>,
        <th>
            Win %
        </th>,
        <th>
            Goals For (GF)
        </th>,
        <th>
            Goals Against (GA)
        </th>,
        <th>
            + / -
        </th>]
```

```
In [13]: soup.find('th').text.strip()
```

```
Out[13]: 'Team Name'
```

```
In [ ]:
```