

### BeautifulSoup (parse a website)

```
from bs4 import BeautifulSoup      import module

soup = BeautifulSoup(html,         create new object
"lxml")

S = soup.find('a')                 finds first <a> tags

S_all = soup.find_all('a')         finds all <a> tags, saves as
list

S_list = soup.find_all(['a',      finds all <a> and <b> tags
'b'])

S_all_tags =                       finds all tags <> and no text
soup.find_all(True)
```

### Requests

```
import requests                   import module

S = requests.get(url)             get html

S.content                         url content in object form
```

### QT (Graphical Interface)

### Selenium (automated web experience)

```
import selenium                   import module
```

### JSON (write or read from JSON file)

```
import json                       import module

J = json.dumps(data)              saves dictionary to JSON
format

J_sorted = json.dumps(data,       saves dictionary to JSON
sort_keys = True)                format but in the order
```

### Scrapy (web crawler)



By **jimmyray9**  
[cheatography.com/jimmyray9/](https://cheatography.com/jimmyray9/)

Not published yet.  
Last updated 15th May, 2017.  
Page 1 of 1.

Sponsored by **Readability-Score.com**  
Measure your website readability!  
<https://readability-score.com>