Python_Jobs_Scraping

January 13, 2025

[ ]:
```
# pip install bs4
```

[2]:
```
from bs4 import BeautifulSoup
import requests
```

[4]:
```
url = 'https://realpython.github.io/fake-jobs/'

response = requests.get(url)
```

[5]:
```
data = BeautifulSoup(response.content, 'html.parser')
```

[6]:
```
# Scraping one job
job_title = data.find(class_='title is-5')

print(job_title.text)
```

```
Senior Python Developer
```

[7]:
```
#scraping all jobs title
all_job_title = data.find_all(class_='title is-5')

print(all_job_title[3].text)
```

```
Fitness centre manager
```

[10]:
```
# Scraping element by id
container = data.find(id='ResultsContainer')
```

[11]:
```
# Scraping element by class
job_divs = container.find_all(class_='column is-half')
```

[17]:
```
# Printing job titles
for i, job in enumerate(job_divs[:10]):
    title = job.find(class_='title')
    print(title.text)
```

```
Senior Python Developer
Energy engineer
Legal executive
Fitness centre manager
```

```
Product manager
Medical technical officer
Physiological scientist
Textile designer
Television floor manager
Waste management officer
```

[14]:

```python
# Extracting Images through class
for i, job in enumerate(job_divs[:5]):
    image = job.find(class_='image')
    print(image)
```

```
<figure class="image is-48x48">
<img alt="Real Python Logo" src="https://files.realpython.com/media/real-python-
logo-thumbnail.7f0db70c2ed2.jpg?__no_cf_polish=1"/>
</figure>
<figure class="image is-48x48">
<img alt="Real Python Logo" src="https://files.realpython.com/media/real-python-
logo-thumbnail.7f0db70c2ed2.jpg?__no_cf_polish=1"/>
</figure>
<figure class="image is-48x48">
<img alt="Real Python Logo" src="https://files.realpython.com/media/real-python-
logo-thumbnail.7f0db70c2ed2.jpg?__no_cf_polish=1"/>
</figure>
<figure class="image is-48x48">
<img alt="Real Python Logo" src="https://files.realpython.com/media/real-python-
logo-thumbnail.7f0db70c2ed2.jpg?__no_cf_polish=1"/>
</figure>
<figure class="image is-48x48">
<img alt="Real Python Logo" src="https://files.realpython.com/media/real-python-
logo-thumbnail.7f0db70c2ed2.jpg?__no_cf_polish=1"/>
</figure>
```

[15]:

```python
# Storing extracted data
dict_data = {}

for job in job_divs:
    title = job.find(class_='title').text
    sub_title = job.find(class_='subtitle').text
    location = job.find(class_='location').text.strip()
    date = job.find(class_='is-small').text.strip()

    dict_data = {
        'title': title,
        'subtitle': sub_title,
        'location': location,
        'date': date
    }
```

[ ]: