multiverse

Data Science Unit 4

Accessing Data From the Internet

Before we start…

- ➔ Make sure you are comfortable
- ➔ Have water and maybe a strong coffee handy
- ➔ If you need a break… take it!
- ➔ If you need a stretch – please go ahead!
- ➔ Please mute yourselves if you are not talking
- ➔ Have your video on at all times

…and let's get started!

In this session we will…

1. **Define** APIs and Web Scraping
2. **Investigate** APIs and see how they work
3. **Build** a web scraper using BeautifulSoup.

Private and Confidential

multiverse

Web Services and API

What is an API?

Private and Confidential

Web API

Private and Confidential

Famous APIs

Private and Confidential

Famous APIs

Private and Confidential

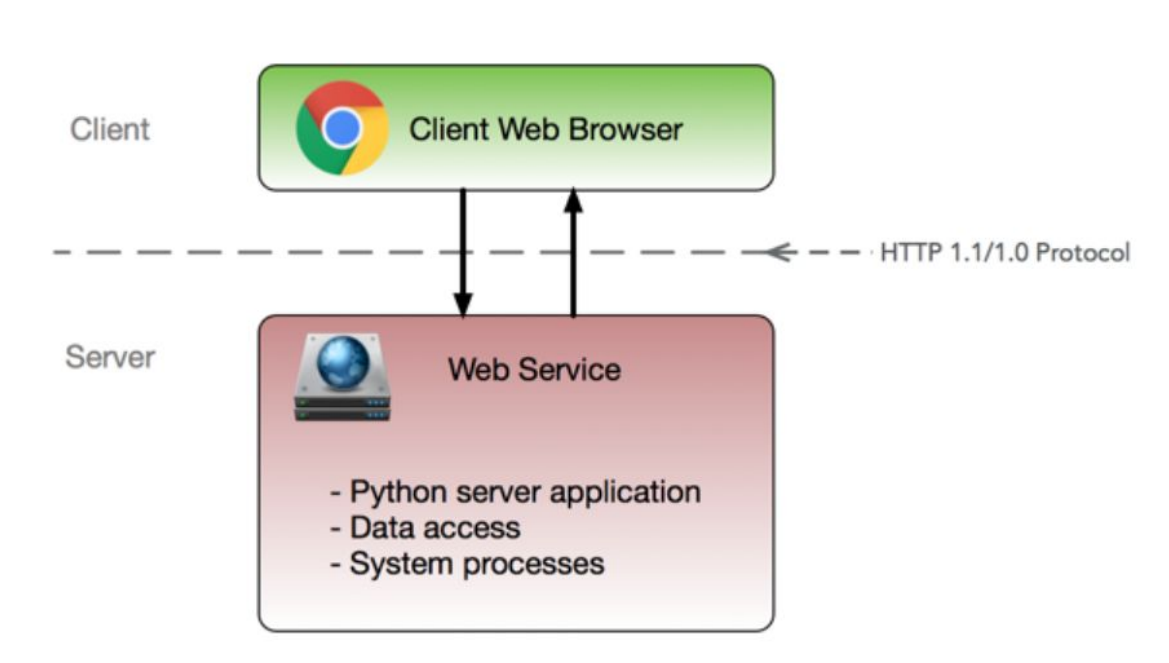
Making API Calls

Private and Confidential

HTTP

Private and Confidential

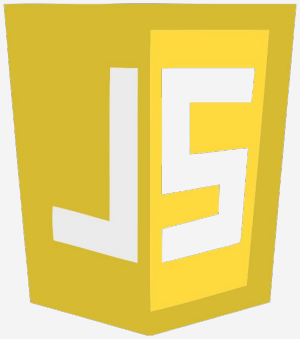
HTTP

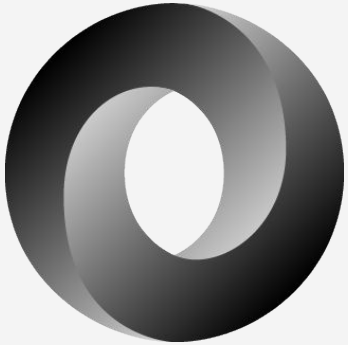


Private and Confidential

JSON

JavaScript





Private and Confidential

multiverse

Let’s Practice

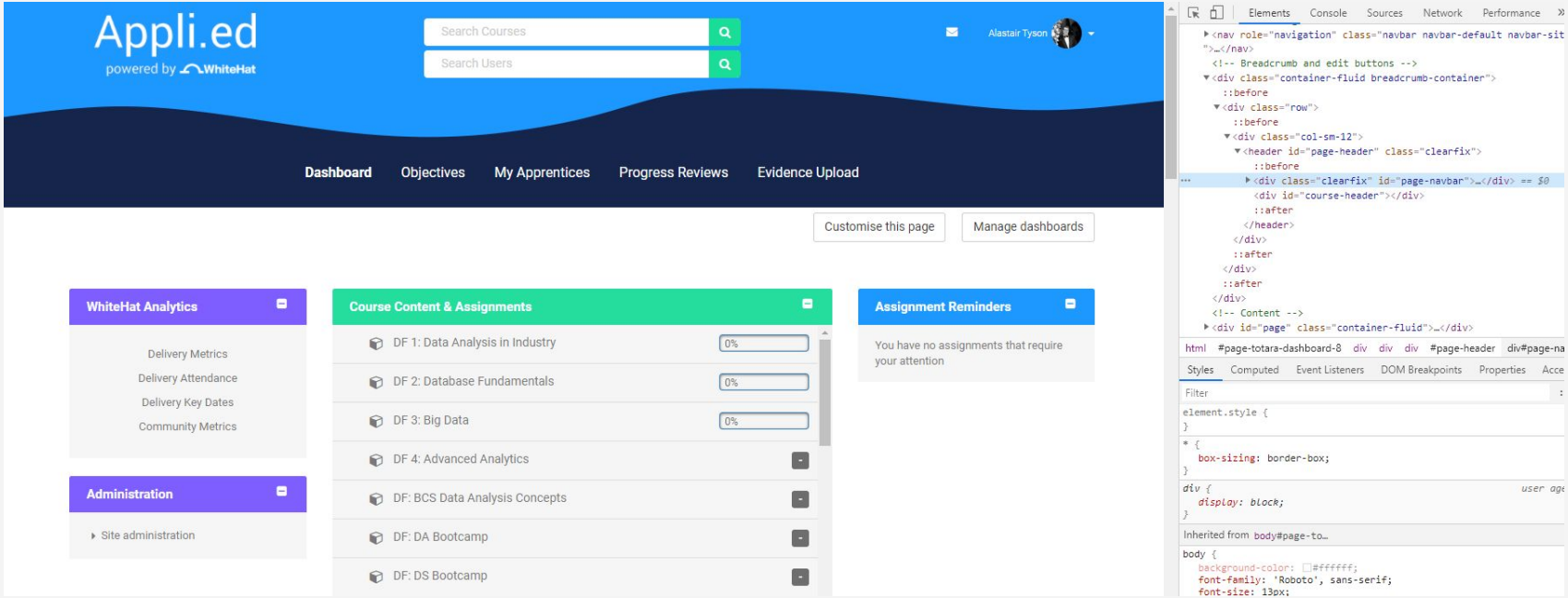
multiverse

Web Scraping

BeautifulSoup

Private and Confidential

HTML



Private and Confidential

HTML

 <h1> <a> <p> <ul>

Private and Confidential

Along came a Spider…

```
# convert this into a soup object
soup = BeautifulSoup(html.text, 'html.parser')
```

```
# print the restaurant names
soup.find_all(name='span', attrs={'class':'rest-row-name-text'})
```

```
# for each element you find, print out the restaurant name
for entry in soup.find_all(name='span', attrs={'class':'rest-row-name-
    print(entry.text)
```


Private and Confidential

Warning

Private and Confidential

multiverse

Let’s Practice