ECON446 :Applications of Cloud Computing and Big Data

Week 4, Lecture 4 | Selenium Cont. & APIs

Sam Borghese

Tuesday, April 23rd, 2024

1. More about Selenium
2. API Review and Wrappers
3. Mock APIs
4. Mock API Examples

More about Selenium

Selenium

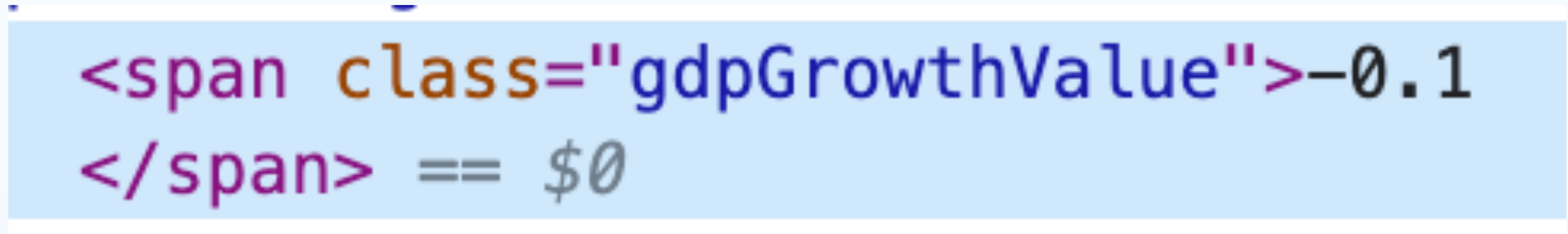
Automate Web Browsers

Free Open source automated testing framework used to validate web applications across different browsers and platforms

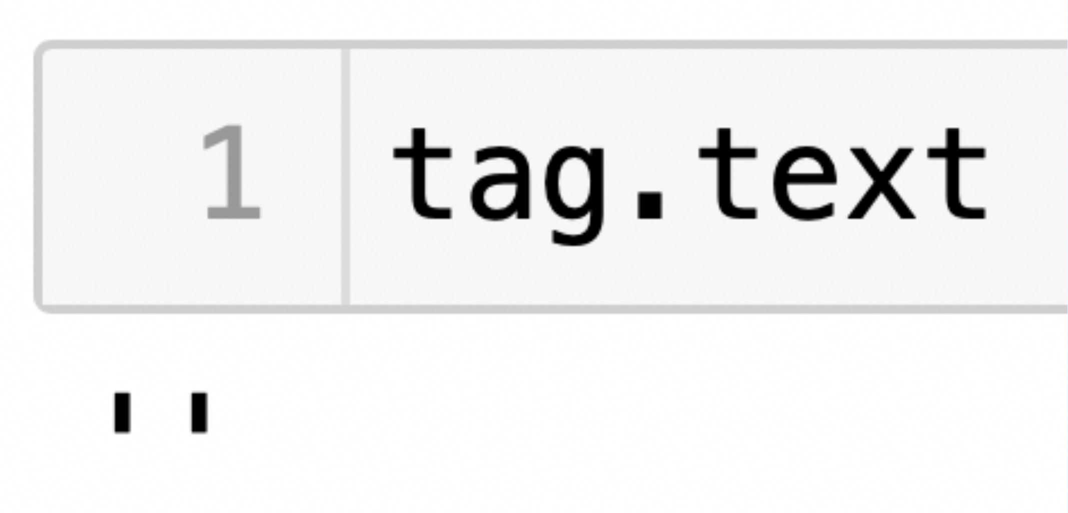
The name Selenium comes **from a joke made by Huggins in an email, mocking a competitor named Mercury**, saying that you can cure mercury poisoning by taking selenium supplements.

Note 1.) .get_attribute method

Sometimes the Text attribute does not work and you have to use .get_attribute



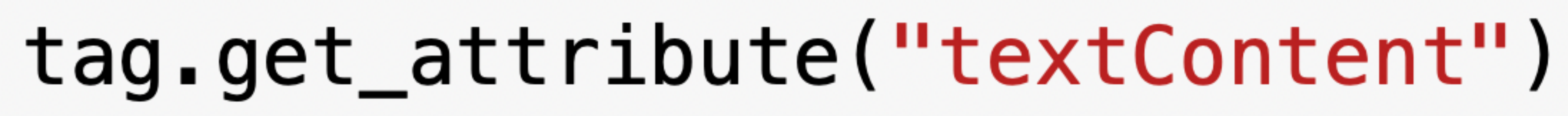
```
<span class="gdpGrowthValue">-0.1
</span> == $0
```



```
tag.text
```

```
''
```

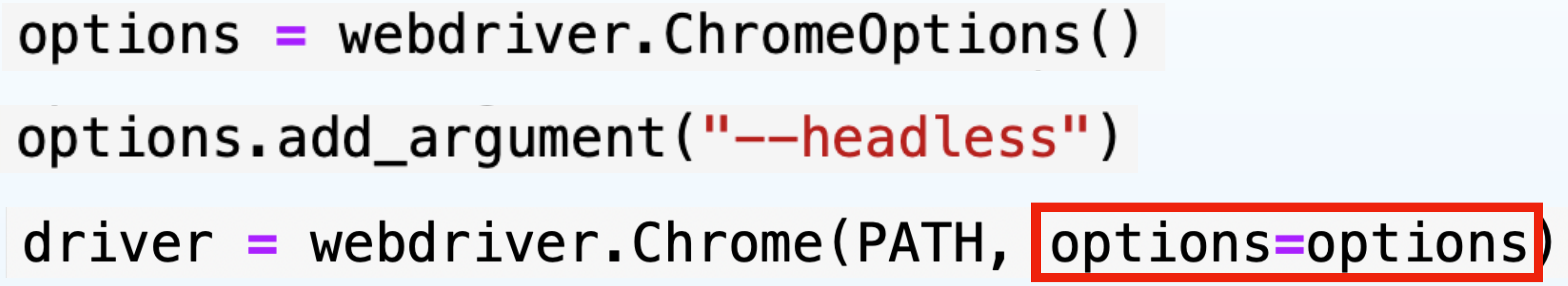
textContent will give you the associated text



```
tag.get_attribute("textContent")
```

Note 2.) Most useful “option” is —headless

Headless allows your code to run without windows popping up.

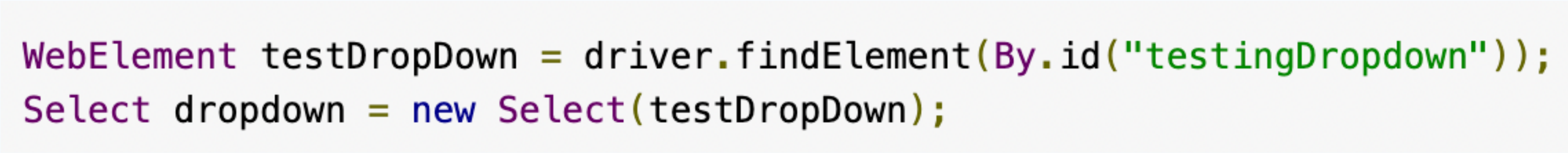


```
options = webdriver.ChromeOptions()

options.add_argument("--headless")

driver = webdriver.Chrome(PATH, options=options)
```

Set specific options

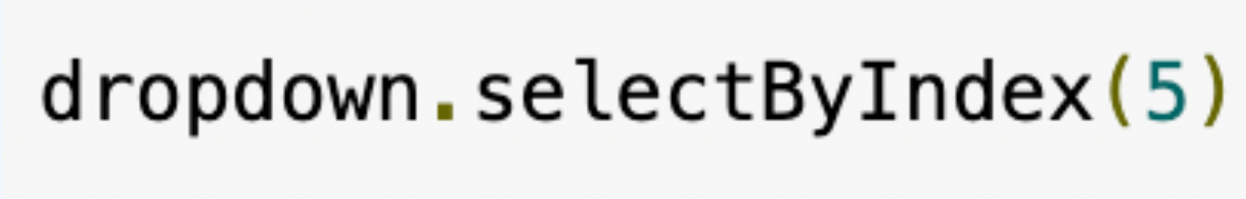
Note 3.) Dealing with Drop Down Menus

Each Javascript element has it’s own way to be interacted with



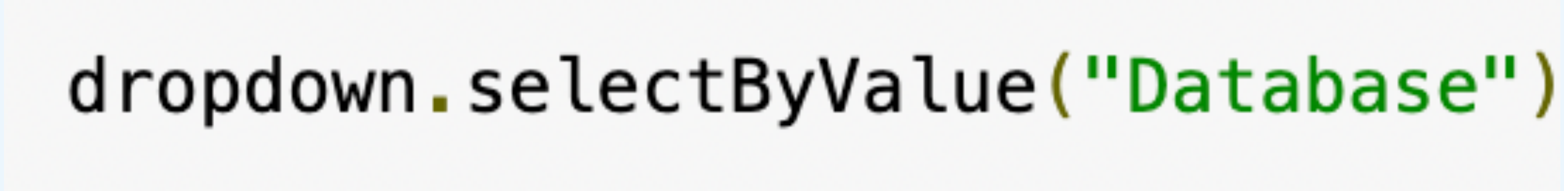
```
WebElement testDropDown = driver.findElement(By.id("testingDropdown"));
Select dropdown = new Select(testDropDown);
```

selectByIndex - Used to select an option based on Index number



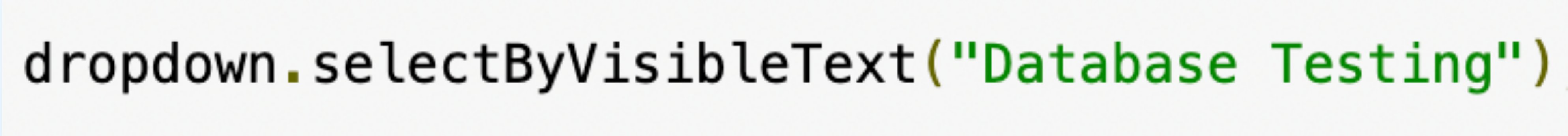
```
dropdown.selectByIndex(5)
```

selectByValue - Used to select based on its ‘value attribute’



```
dropdown.selectByValue("Database")
```

selectByVisibleText -Used to select an option based on text over the option



```
dropdown.selectByVisibleText("Database Testing")
```

Scrape information from Eurostat

Example 1 :

Show why Selenium is necessary for scraping Eurostat data and then use it pull the official Euro GDP number.

API Review & API Wrappers

RESTful API AKA Web API

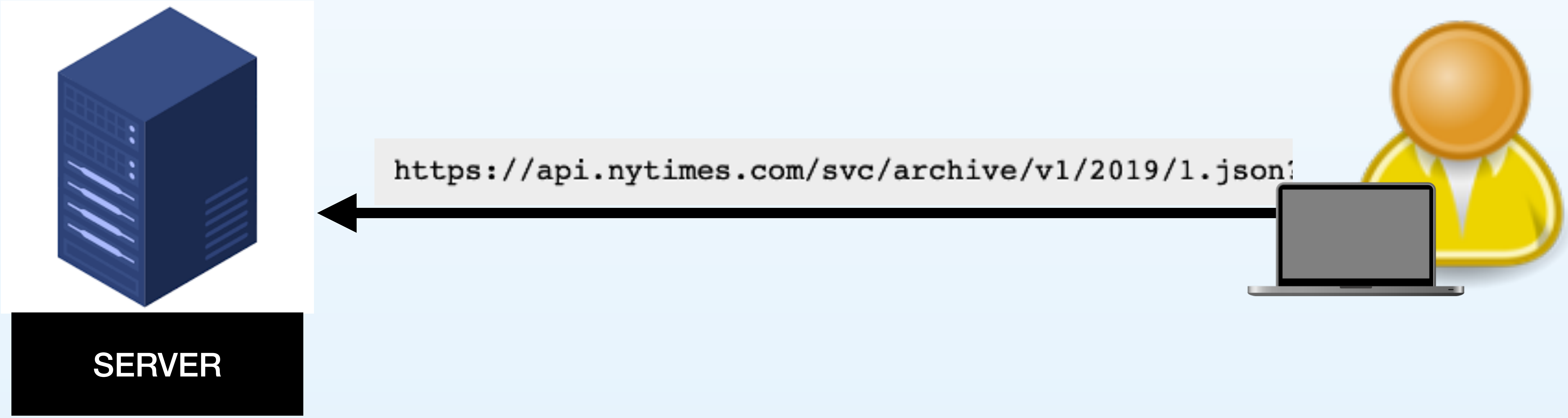
Application Programming Interface : “Allows computers to talk to one another”

REST : Set of architectural constraints
Utilizes HTTP to send data in several formats

RESTful API AKA Web API

Calling an API sends a request for data in the form of a URL

“API Call”

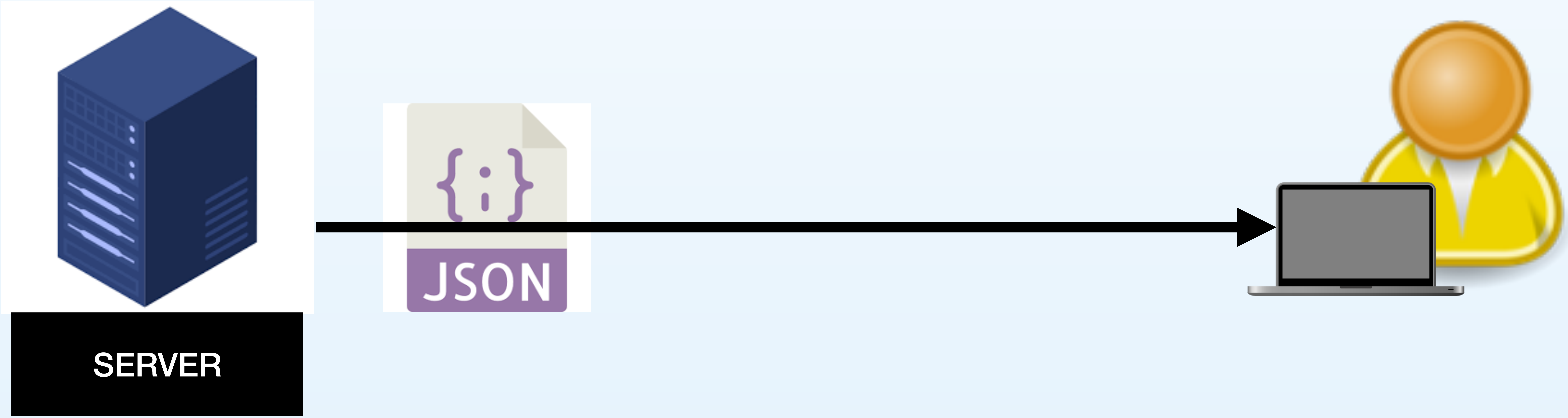


```
https://api.nytimes.com/svc/archive/v1/2019/1.json
```

RESTful API AKA Web API

Calling an API sends a request for data in the form of a URL

The server will prepare the requested data and send it to the user



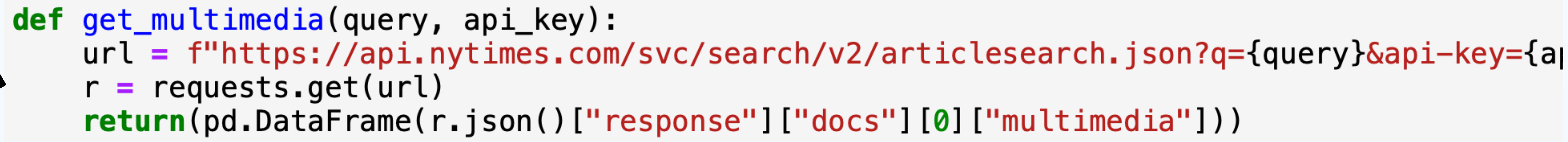
API Wrappers

Streamlined Process for calling APIs

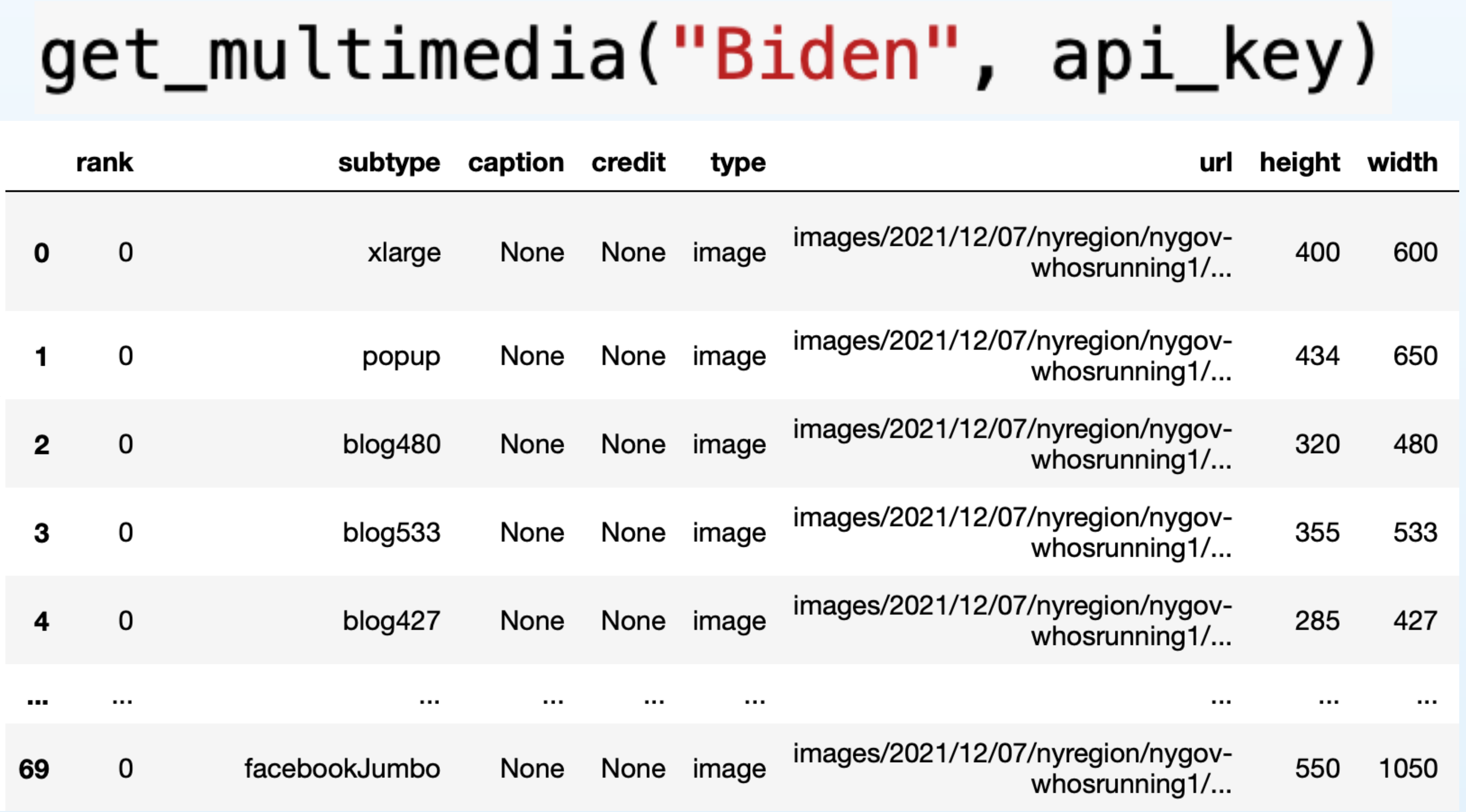
Creates a package used inside a specific programming language that queries the API using cleanly designed functions

API Wrappers

Functions are much easier in building loops and functions as opposed to creating new queries



```
def get_multimedia(query, api_key):
    url = f"https://api.nytimes.com/svc/search/v2/articlesearch.json?q={query}&api-key={a
    r = requests.get(url)
    return(pd.DataFrame(r.json()["response"]["docs"][0]["multimedia"]))
```



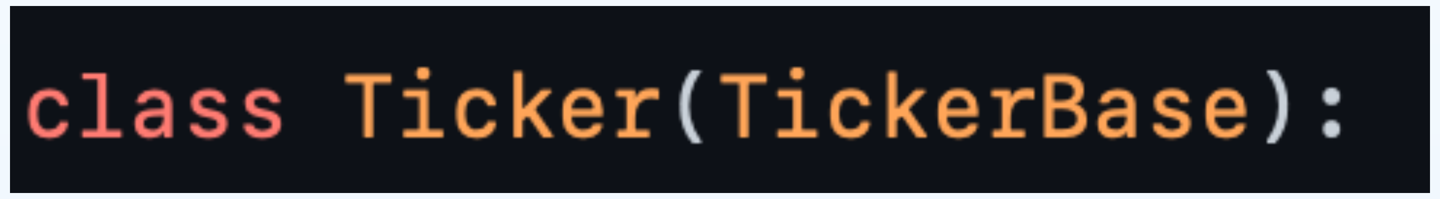
```
get_multimedia("Biden", api_key)
```

	rank	subtype	caption	credit	type	url	height	width
0	0	xlarge	None	None	image	images/2021/12/07/nyregion/nygov-whosrunning1/...	400	600
1	0	popup	None	None	image	images/2021/12/07/nyregion/nygov-whosrunning1/...	434	650
2	0	blog480	None	None	image	images/2021/12/07/nyregion/nygov-whosrunning1/...	320	480
3	0	blog533	None	None	image	images/2021/12/07/nyregion/nygov-whosrunning1/...	355	533
4	0	blog427	None	None	image	images/2021/12/07/nyregion/nygov-whosrunning1/...	285	427
...	...	...	...	...	...	...	...	...
69	0	facebookJumbo	None	None	image	images/2021/12/07/nyregion/nygov-whosrunning1/...	550	1050

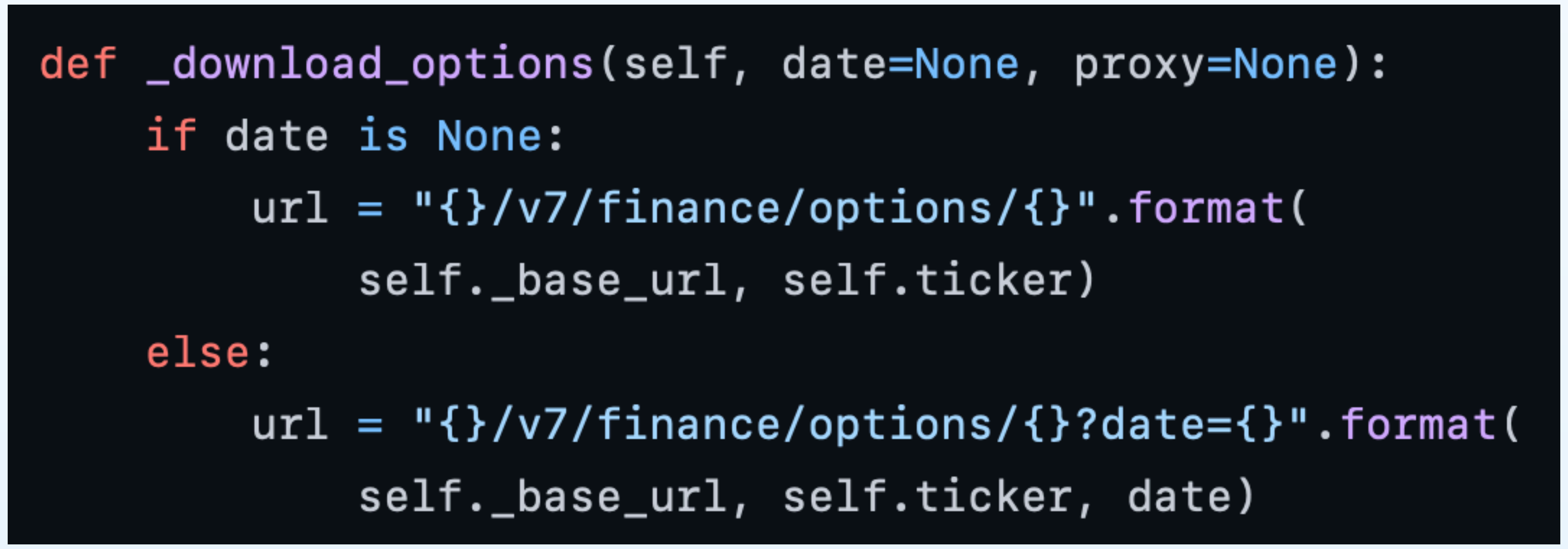
API Wrapper Examples

yfinance

GitHub Documentation



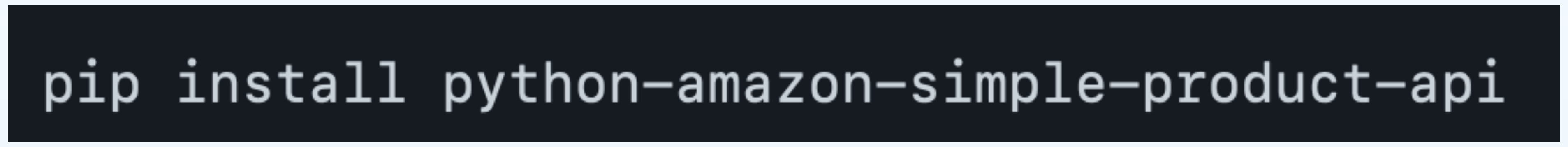
```
class Ticker(TickerBase):
```



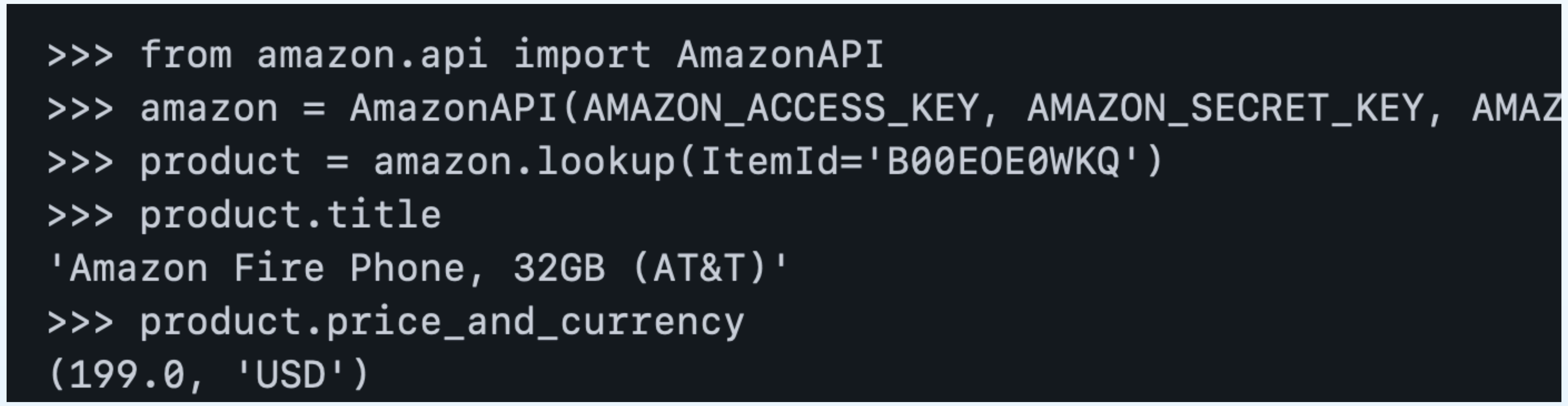
```
def _download_options(self, date=None, proxy=None):
    if date is None:
        url = "{}/v7/finance/options/{}".format(
            self._base_url, self.ticker)
    else:
        url = "{}/v7/finance/options/{}?date={}".format(
            self._base_url, self.ticker, date)
```

AmazonAPI

Query Example



```
pip install python-amazon-simple-product-api
```



```
>>> from amazon.api import AmazonAPI
>>> amazon = AmazonAPI(AMAZON_ACCESS_KEY, AMAZON_SECRET_KEY, AMAZ
>>> product = amazon.lookup(ItemId='B00EOE0WKQ')
>>> product.title
'Amazon Fire Phone, 32GB (AT&T)'
>>> product.price_and_currency
(199.0, 'USD')
```

List of Python Wrappers : https://github.com/realpython/list-of-python-api-wrappers

Mock API

Mock API

Web Scraping/Crawling that is disguised in an API Wrapper

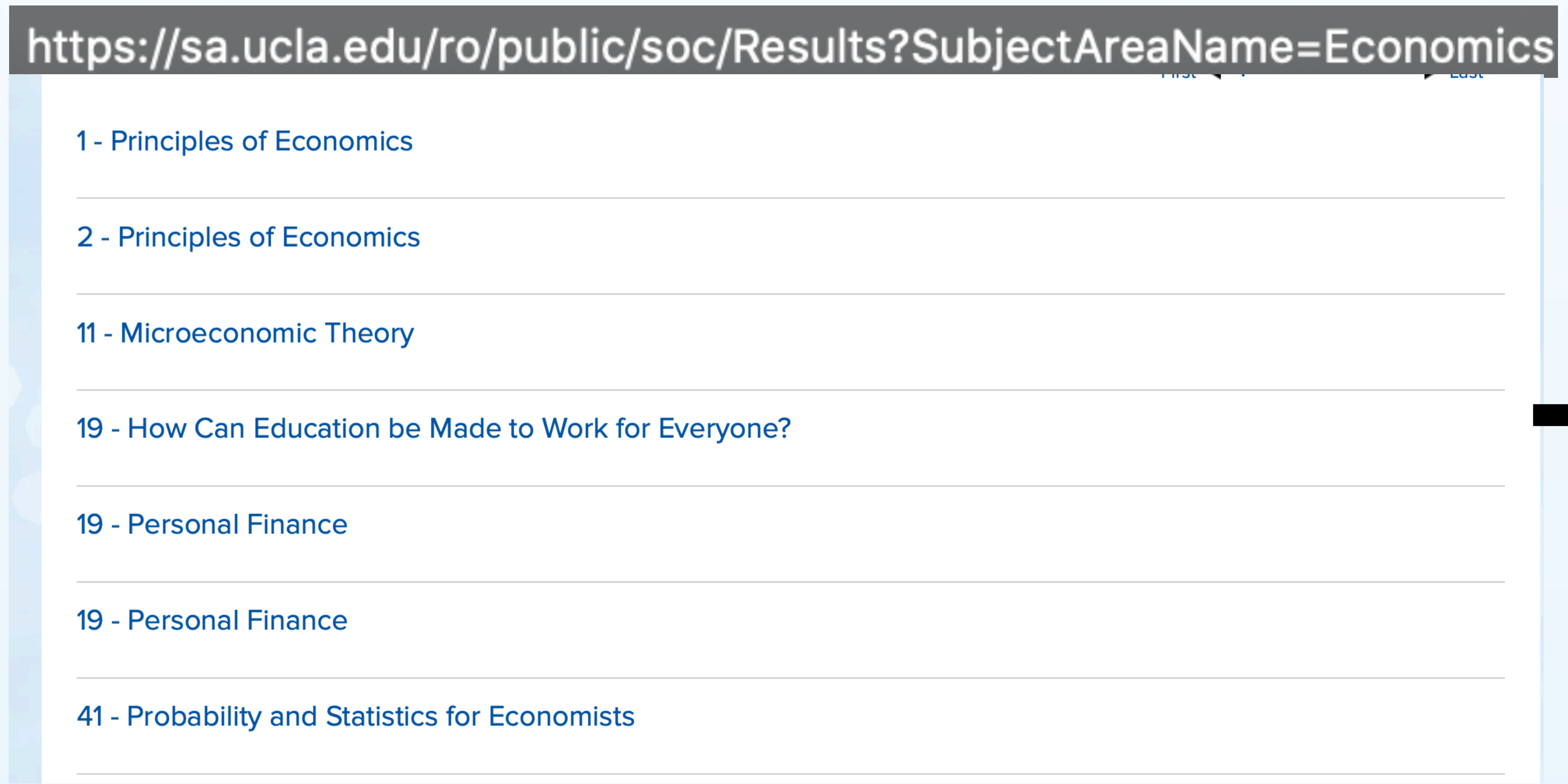
Pulls data from a website that does NOT have an API

Has easy to use functions/methods

May not be authorized by the webpage

Mock API : Example

UCLA Registrar does not have an API



https://sa.ucla.edu/ro/public/soc/Results?SubjectAreaName=Economics

1 - Principles of Economics

2 - Principles of Economics

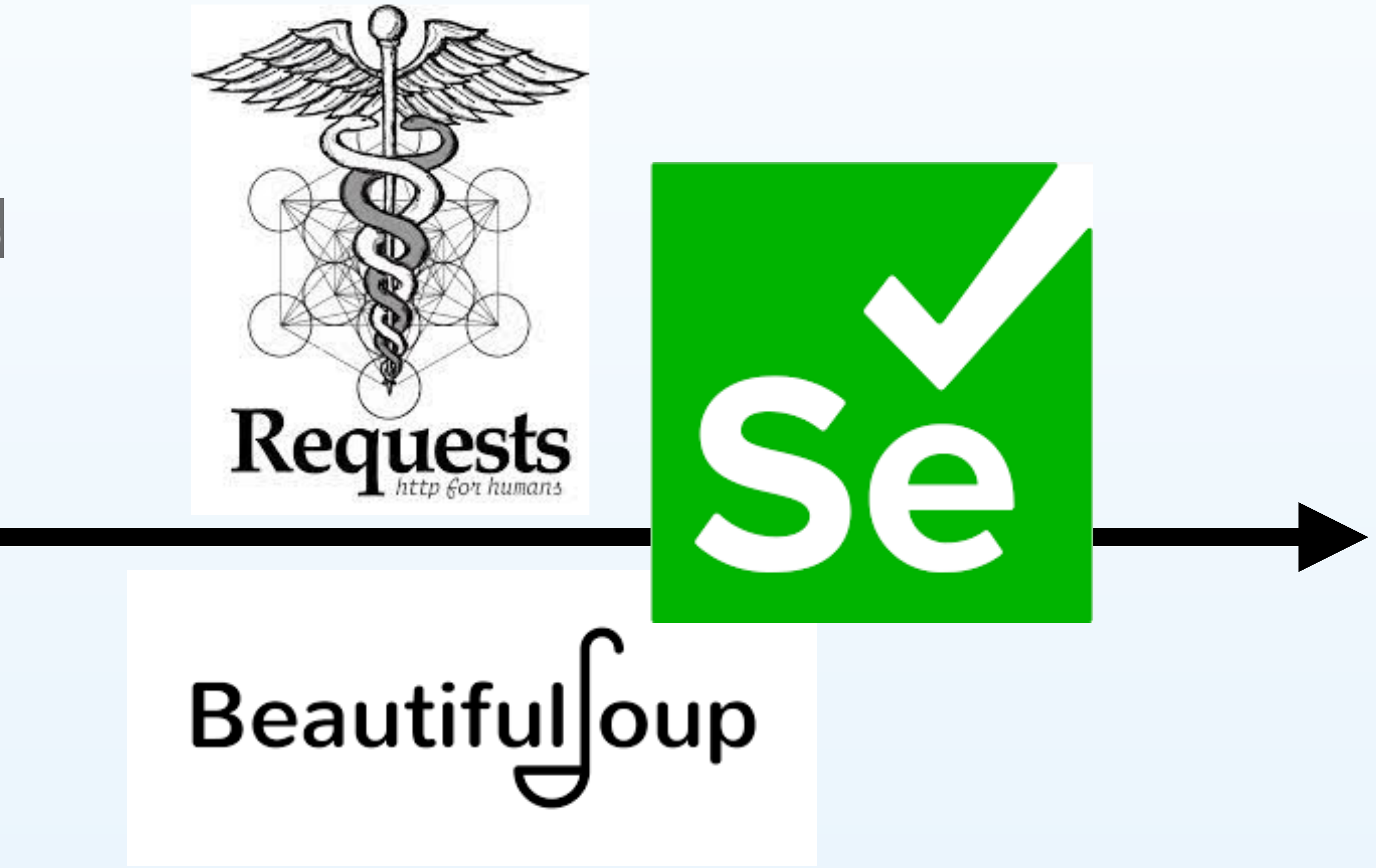
11 - Microeconomic Theory

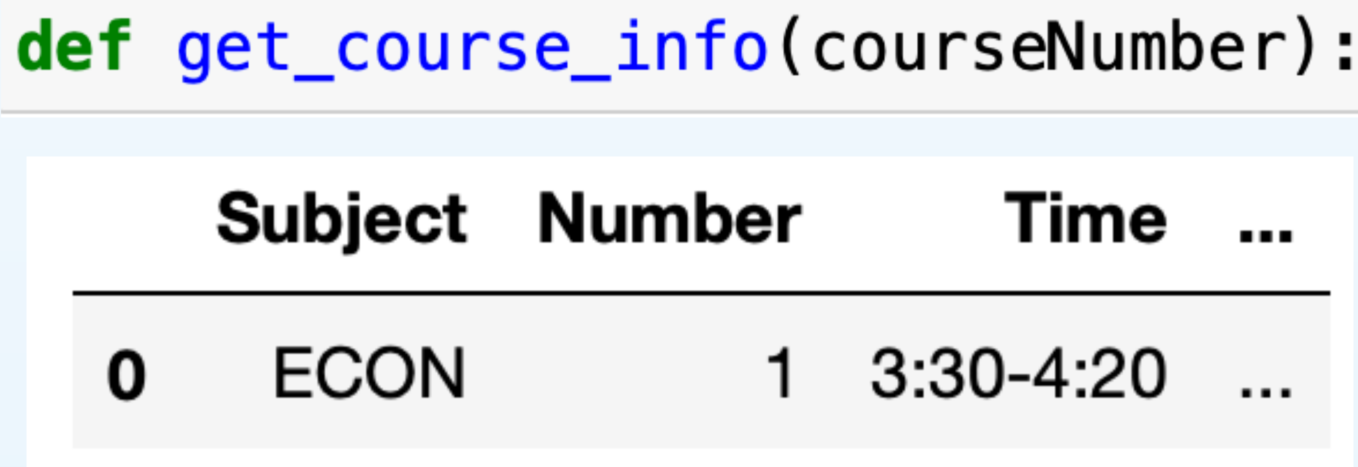
19 - How Can Education be Made to Work for Everyone?

19 - Personal Finance

19 - Personal Finance

41 - Probability and Statistics for Economists





```
def get_course_info(courseNumber):
```

	Subject	Number	Time	...
0	ECON	1	3:30-4:20	...

Mock API Example

Scrape information from Eurostat

Example 2 :

Use the code from earlier in this slide and make a mock API. When passed a code, it should return the associated value from Eurostat.