

Python Discord Bot

Day 07

All about Web-scraping

Contents

Exercise 01:	\$scrape	3
Exercise 02:	\$currency	4
Exercise 03:	\$job	6
Exercise 04:	\$movie	9

Exercise 01: \$scrape

	Difficulty:	***
<u> </u>	Objectives:	Use a command to instruct the discord bot to scrape the first name, last name and username of the first user in the first table from selected website.
P	Skills required:	 Beautiful soup (Find next) Inspect Html Python requests get method Python string format method
	Resources:	Selected website for web scraping: https://webscraper.io/test-sites/tables For your reference: https://beautiful-soup-4.readthedocs.io/en/latest/#find-all-next-and-find-next https://realpython.com/python-f-strings/#python-f-strings-the-pesky-details
8	Search keywords:	 Python requests get method Implementing Web Scraping in Python with BeautifulSoup BeautifulSoup find next
<u></u>	Command required:	\$scrape

Code snippet:

```
Import requests
From bs4 import BeautifulSoup
...

output = f"First Name = {firstname} \nLast Name = {lastname} \nUsername = {username}"
```

⊘ Output example:

```
Code4Teens BOT Today at 18:30

First Name = Mark

Last Name = Otto

Username = @mdo
```

Exercise 02: \$currency

	Difficulty:	***
<u> </u>	Objectives:	Use a command to instruct the discord bot to scrape the exchange rate of US Dollar from the selected website. [] You must use web scraping to scrape the Currency Name (USD Big) and the Exchange Rate (We Buy Rate).
P	Skills required:	 Beautiful soup (Going down & Going sideways) Inspect Html Python requests get method Python string format method Python Split and Join
a	Resources:	Selected website for web scraping: https://mtradeasia.com/main/daily-exchange-rates/ For your reference: https://beautiful-soup-4.readthedocs.io/en/latest/#going-sideways
P	Search keywords:	 Python requests get method Implementing Web Scraping in Python with BeautifulSoup Python BeautifulSoup Navigating tree sideways String split and join in python
<u></u>	Command required:	\$currency

Code snippet:

```
Import requests
From bs4 import BeautifulSoup
...

output = "Currency: {}\nExchange Rate (MYR):
{}".format(update_currency, rate)
```

Output example:



Exercise 03: \$job

	Difficulty:	***
	Objectives:	Use a command to instruct the discord bot to scrape a List of All Job title and Company name in the selected webpage and output the result in an embed message form.
	Skills required:	 Web scraping Loops List & tuples Join function Zip function
	Resources:	Selected website for web scraping: https://my.wobbjobs.com/jobs For your reference: https://realpython.com/beautiful-soup-web-scraper-python/ https://www.w3schools.com/python/ref func zip.asp https://www.pluralsight.com/guides/extracting-data-html-beautifulsoup
8	Search keywords:	 Scraping data with multiple values within same class Join list with newline character Combining list with matching index
<u></u>	Command required:	\$job

Code snippet:

```
Import requests
From bs4 import BeautifulSoup

#Scrape all available Jobs in the webpage & store the result in a List

#Scape all the company names of the available jobs & store the result in a List
...

#Combine these 2 list with matching index
...

#Add a new line character to separate each line of result
...

#Output
Output = "Job: " + joblist + "Company: " + company_list

#Put the output in an embed message form
...
```

⊘ Output example:



Code4Teens BOT Today at 8:23 PM

Job List:

Job: Call Center Specialist - Outbound (Verification Specialist) , Company: Verity Intelligence Sdn Bhd

Job: People & Culture Intern, Company: WOBB

Job: Senior Associate (PR & Communications), Company: Perspective Strategies

Job: Sales Executive, Company: Virtual Peddlers

Job: Sales Consultant , Company: Life n Fitness

Job: Finance & Operations Intern , Company: WOBB

Job: Business Development Intern (Corporate Solutions), Company: WOBB

Job: Sales Client Service Representative-China Support , Company: Think Markets Pty. Ltd

Job: Financial Sales Consultant (China Support), Company: Think Markets Pty. Ltd

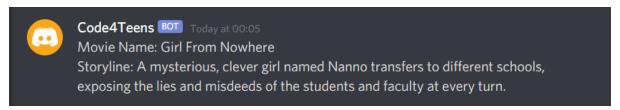
Exercise 04: \$movie

<u> </u>	Difficulty:	$\wedge \wedge \wedge \wedge \wedge$
	Objectives:	Use a command to instruct the discord bot to scrape the storyline of the movie.
1	Skills required:	 String split or String slice Python requests get method Python string format method Beautiful soup (Extract URL & Going down) Inspect Html Error Handling
	Resources:	Selected website for web scraping: https://www.imdb.com/find?q={}&s=tt&ttype=ft&ref">https://www.imdb.com/find?q={}&s=tt&ttype=ft&ref">https://www.imdb.com/find?q={}&s=tt&ttype=ft&ref">https://www.imdb.com/find?q={}&s=tt&ttype=ft&ref">https://beautiful-soup- 4.readthedocs.io/en/latest/#quick-start https://beautiful-soup-4.readthedocs.io/en/latest/#going-down
D	Search keywords:	 Python requests get method Python string format method Implementing Web Scraping in Python with BeautifulSoup Extract links from webpage using BeautifulSoup
<u></u>	Command required:	\$movie [name] a) Your bot must able to handle the error handling part: response when no movie name was inserted.

Code snippet:

```
Import requests
From bs4 import BeautifulSoup
...
# Error Handling
...
# Search movie name via the search URL
...
# Extract link of the first movie from the results
...
# Scrape the storyline by using the new movie link
...
output = f"Movie Name: {movie_name} \nStoryline: {storyline}"
```

⊘ Output example:



O Error Handling Example:

