

Started on	Monday, 1 July 2024, 2:24 PM
State	Finished
Completed on	Monday, 1 July 2024, 2:26 PM
Time taken	1 min 39 secs
Marks	4.00/5.00
Grade	8.00 out of 10.00 (80%)

Question 1

Complete

Mark 3.00 out of 3.00

What is not an application of web scraping?

- ☐ a. Create specific collections of data such as email-addresses or photos
- ☐ b. Extract necessary information from a web page
- ☐ c. Collect sample of data from a web source which might change a value over a long period of time.
- ☒ d. Browsing the internet using a web browser.
- ☐ e. Automatically aggregate data from different sources to create a third party application

Question 2

Complete

Mark 0.50 out of 0.50

While trying to retrieve a web page for scraping data, you received "Access Denied" message from the server. Why do you think this is?

- ☐ a. Error in the python program
- ☐ b. Error in server
- ☒ c. Server does not support web scraping of particular resource, and has determined your requests not allowed, because they were coming from an automated script (robot) or some other reason.

Question 3

Complete

Mark 0.50 out of 0.50

Web page had some content when you look at the browser. However, the web scraping could not extract that content. What could be the reasons?

(More than one answer is allowed)

- ☒ a. Issues with locating the correct tag in your code
- ☒ b. Web browser load the content dynamically, and your source code did not retrieve the secondary resources and dynamic content.

Question 4

Complete

Mark 0.00 out of 1.00

The last trinket in Lecture Notes demonstrated web scraping Google search results. However, it only returns the search results from the first page.

Use your knowledge on HTTP and the web scraping to determine how to obtain the rest of the search results, which are in second and third pages etc.

You can load the search results in browser and observe the links to 2nd or 3rd pages to the result to identify this.

- ☐ a. begin=2
- ☒ b. Use get parameter page=2 etc.
- ☐ c. page=20
- ☐ d. start=2
- ☐ e. start=20

◀ 4.4 Web Scraping in Python

Jump to...

4.5 Introduction to XML with Python ▶

GET IN TOUCH

🏠 University of Moratuwa
Centre for Open & Distance Learning
CODL

☎ 011 308 2787/8

☎ 011 265 0301 ext. 3850,3851

✉ open@uom.lk

🌐 [University Website](#)

🌐 [CODL Website](#)

[Data retention summary.](#)