Back

Reading Web Data From Python

Quiz, 8 questions

loading   https://d3njjcbhbojbot.cloudfront.net/web/images/icons/loading.gif

Load Error!

1  
point

1.

Which of the following Python data structures is most similar to the value returned in this line of Python:



1

x = urllib.urlopen('http://www.py4inf.com/code/romeo.txt')

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX



file handle



dictionary



list



socket



regular expression

1  
point

2.

In this Python code, which line actually reads the data?



1

2

3

4

5

6

7

8

9

10

import socket

mysock = socket.socket(socket.AF\_INET, socket.SOCK\_STREAM)

mysock.connect(('www.py4inf.com', 80))

mysock.send('GET http://www.py4inf.com/code/romeo.txt HTTP/1.0\n\n')

while True:

data = mysock.recv(512)

if ( len(data) < 1 ) :

break

print data

mysock.close()

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX



mysock.recv()



socket.socket()



mysock.close()



mysock.connect()



mysock.send()

1  
point

3.

Which of the following regular expressions would extract the URL from this line of HTML:



1

<p>Please click <a href="http://www.dr-chuck.com">here</a></p>

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX



href="(.+)"



href=".+"



http://.\*



<.\*>

1  
point

4.

In this Python code, which line is most like the open() call to read a file:



1

2

3

4

5

6

7

8

9

10

import socket

mysock = socket.socket(socket.AF\_INET, socket.SOCK\_STREAM)

mysock.connect(('www.py4inf.com', 80))

mysock.send('GET http://www.py4inf.com/code/romeo.txt HTTP/1.0\n\n')

while True:

data = mysock.recv(512)

if ( len(data) < 1 ) :

break

print data

mysock.close()

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX



mysock.connect()



import socket



mysock.recv()



mysock.send()



socket.socket()

1  
point

5.

Which HTTP header tells the browser the kind of document that is being returned?



Content-Type:



Metadata:



HTML-Document:



ETag:



Document-Type:

1  
point

6.

What should you check before scraping a web site?



That the web site returns HTML for all pages



That the web site only has links within the same site



That the web site allows scraping



That the web site supports the HTTP GET command

1  
point

7.

What is the purpose of the BeautifulSoup Python library?



It builds word clouds from web pages



It animates web operations to make them more attractive



It allows a web site to choose an attractive skin



It optimizes files that are retrieved many times



It repairs and parses HTML to make it easier for a program to understand

1  
point

8.

What ends up in the "x" variable in the following code:



1

2

3

html = urllib.urlopen(url).read()

soup = BeautifulSoup(html)

x = soup('a')

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX



A list of all the anchor tags (<a..) in the HTML from the URL



True if there were any anchor tags in the HTML from the URL



All of the externally linked CSS files in the HTML from the URL



All of the paragraphs of the HTML from the URL



I, **Pradeep Sathyamurthy**, understand that submitting work that isn’t my own may result in permanent failure of this course or deactivation of my Coursera account. Learn more about Coursera’s Honor Code

Submitting...Submit error! Please try again.Submit QuizSubmit Quiz

--> <img height="1" width="1" style="display:none" src="https://www.facebook.com/tr?id=946401778754875&amp;ev=PageView&amp;noscript=1">

Confirm Navigation

Are you sure you want to leave this page?

Stay on this Page  Leave this Page

Confirm Navigation

Are you sure you want to leave this page?

Stay on this Page  Leave this Page