Done

Bienvenue Thomas AWOUNFOUET de Using Python to Access Web Data

Your answer is correct, score saved.

Extracting Data from JSON

In this assignment you will write a Python program somewhat similar to http://www.py4e.com/code3/json2.py. The program will prompt for a URL, read the JSON data from that URL using **urllib** and then parse and extract the comment counts from the JSON data, compute the sum of the numbers in the file and enter the sum below:

We provide two files for this assignment. One is a sample file where we give you the sum for your testing and the other is the actual data you need to process for the assignment.

- Sample data: http://py4e-data.dr-chuck.net/comments_42.json (Sum=2553)
- Actual data: http://py4e-data.dr-chuck.net/comments_1921192.json (Sum ends with 97)

You do not need to save these files to your folder since your program will read the data directly from the URL. **Note:** Each student will have a distinct data url for the assignment - so only use your own data url for analysis.

Data Format

The data consists of a number of names and comment counts in JSON as follows:

```
{
  comments: [
    {
      name: "Matthias"
      count: 97
    },
    {
      name: "Geomer"
      count: 97
    }
    ...
  ]
}
```

The closest sample code that shows how to parse JSON and extract a list is json2.py. You might also want to look at geoxml.py to see how to prompt for a URL and retrieve data from a URL.

Sample Execution

```
$ python3 solution.py
Enter location: http://py4e-data.dr-chuck.net/comments_42.json
Retrieving http://py4e-data.dr-chuck.net/comments_42.json
Retrieved 2733 characters
Count: 50
Sum: 2...
```

Turning in the Assignment

Enter the sum from the actual data and your Python code below:

Sum: 2297 (ends with 97) Submit Assignment

Python code:

```
import json
import urllib.request, urllib.parse, urllib.error


url = input('Enter - ')
print('1_Retrieving', url, type(url)) #class str
uh = urllib.request.urlopen(url)
print('2_uh', uh, type(uh)) #class 'http.client.HTTPResponse'
data = uh.read()
print("3_uh.read() or data", data, type(data)) #calss bytes
data = data.decode()
print("4_data decode", data, type(data)) #class str
info = json.loads(data)
print("5_info", info, type(info), '5_len_info', len(info)) #class dict
print('6_comments',info['comments'][2]['count']) # just as way to recover smth
print('7_comments type', type(info['comments']), '7_comments len', len(info['comments']))
x = 0
for item in (info['comments']) :
    print('count', item['count'], type(item['count']))
    x = x + item['count']
print(x)
```