Introduction to Python II Exercises 03 (sample program answers)

Remember that you can come out with a different way to solve the exercises

At the beginning while you are getting acquainted with programming and Python as a language your objective is to produce a suitable RESULT. As you get more experience, you will be able to apply your python knowledge to write elegant code. But for the time being focus on the results.

1

```
import urllib.request

response = urllib.request.urlopen('http://www.bbc.com/news')

html = response.read().decode()

#number of chars
print("Number of chars: ", len(html))

html = html.split('\n')

print("Number of Lines: ", len(html))

count = 0

for line in html:
    if "BBC" in line:
        count += 1

print("BBC is found {} times in the html source".format(count))

print("First: ", html[0])
print("Last: ", html[-1])  #the last line could be blank

print("Headers: ")
print(response.info())
```

2

```
import urllib.request
from bs4 import BeautifulSoup

url = 'http://www.meteomedia.com'

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

mysoup = BeautifulSoup(html, "html.parser")

# Retrieve all of the anchor tags
a_tags = mysoup('a')
no_of_a_tags = len(a_tags)

print("There are {} <a> tags in the page".format(no_of_a_tags))

# continues in next page
```

```
# Print the href from the anchor tag if it exists

for tag in a_tags:
   href =tag.get('href')
   if href != None:
      print(href)
```

3

```
import urllib.request

from bs4 import BeautifulSoup

url = 'http://www.groupce.com/python/html/thejourney.html'

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

soup = BeautifulSoup(html, "html.parser")

## Retrieve all of the required tags

a_tags = soup.findAll('a')
tr_tags = soup.findAll('tr')

print("There are {} <a> tags in the html source".format(len(a_tags)))
print("There are {} <tr> tags in the html source".format(len(tr_tags)))

print("-----------------------")
print("These are the <a> tags (qty={})".format(len(a_tags)))

for tag in a_tags:
    print(tag.prettify())

print("-----------------------")
print("These are the <tr> tags (qty={})".format(len(tr_tags)))

for tag in tr_tags:
    print(tag.prettify())

#continues in next page

print("-----------------------")

print("These are the <td> tags")
```

```
print("-----------------------")
print("These are the <td> tags")

for tag in tr_tags:

    td_tags = tag.findAll('td')

    for td_tag in td_tags:
        print(td_tag.prettify())

print("-----------------------")
print("These are the 2nd <td> tags in each <tr> tag")

for tag in tr_tags:
    td_tags = tag.findAll('td')

    #printing the second td tag  (it is a list of tags))
     print(td_tags[1])
```

4

```
import urllib.request
import json

url = 'http://www.groupce.com/python/json/json_comments.json'

response = urllib.request.urlopen(url)data = response.read().decode()

jsoninfo = json.loads(data)

total_sum = 0

for field in jsoninfo["comments"]:

      if field['name'][0] == 'A':
            print("Name: {} and count:{}".
                  format(field["name"],field["count"]))

            total_sum = total_sum + field["count"]
            print("Running sum: {}".format(total_sum))
```

5

```
import json

countrylist = [{"country":"Canada",  "capital":"Ottawa","population":"883,391"},
               {"country":"Bulgaria","capital":"Sofia", "population":"1,211,000"},
               {"country":"USA",     "capital":"Washington","population":"7,288,000"},
               {"country":"France",  "capital":"Paris", "population":"2,240,000"},
               {"country":"Romania","capital":"Bucharest","population":"1,883,425"

with open('capitals.json', 'w') as outfile:
    json.dump(countrylist, outfile, indent=4, sort_keys=True, separators=(',', ':'))

with open('capitals.json', 'r') as infile:
    jsondata = infile.read()
    jsoninfo = json.loads(jsondata)

for record in jsoninfo:
    print("Country: {}, Capital: {}, Population: {}".format(record["country"],
          record["capital"], record["population"]))
```

See next page for question no. 6

6

```
import urllib.request
from bs4 import BeautifulSoup

url = 'http://www.groupce.com/python/html/thejourney.html'
response = urllib.request.urlopen(url)
html = response.read()
htmlinfo = BeautifulSoup(html, "html.parser")

firstNumber = int(input('Enter a 1st integer between 1 and 5: '))
secondNumber = int(input('Enter a 2nd integer between 1 and 5: '))

#select all <tr> tags (table rows) and then choose the one corresponding to the
#First Number entered. #Once positioned in the right <tr> tag, select the first <a>tag.
# once the <a> tag is found, then get the href attribute of the <a> tag.

tr_tags = htmlinfo.findAll('tr')
selectedRow = tr_tags[firstNumber-1]
selectedTag = selectedRow.find('a')
selectedHref = selectedTag.get('href')

# Navigate to the url found in the href tag

url = selectedHref
response = urllib.request.urlopen(url)
html = response.read()
htmlinfo = BeautifulSoup(html, "html.parser")

#do the same logic as above to find the href for the corresponding row.

tr_tags = htmlinfo.findAll('tr')
selectedRow = tr_tags[secondNumber-1]
selectedTag = selectedRow.find('a')
selectedHref = selectedTag.get('href')

# Navigate to the url found in the href tag

url = selectedHref
response = urllib.request.urlopen(url)
html = response.read()
htmlinfo = BeautifulSoup(html, "html.parser")

# Get the contents of the title tag (done in two steps for clarity)

title_tag = htmlinfo.find('title')
title = title_tag.contents

#get the equation from the title by splitting it and getting the 2nd elem of the list
equation = title[0].split()[1]

#evaluate the arithmetic expression and print it
print("The result is: {}".format(eval(equation)))
```