Advanced Modules Exercise Solutions

It's time to test your new skills, this puzzle project will combine multiple skills sets, including unzipping files with Python, using os module to automatically search through lots of files.

Your Goal

This is a puzzle, so we don't want to give you too much guidance and instead have you figure out things on your own.

There is a .zip file called 'unzip_me_for_instructions.zip', unzip it, open the .txt file with Python, read the instructions and see if you can figure out what you need to do!

If you get stuck or don't know where to start, here is a guide/hints

Step 1: Unzipping the File

We can easily use the shutil library to extract and unzip the contents of the .zip file

```
In [2]: import shutil
```

```
In [3]: shutil.unpack_archive('unzip_me_for_instructions.zip','','zip')
```

Step 2: Read the instructions file

Let's figure out what we need to do, open the instructions.txt file.

```
In [7]: with open('extracted_content/Instructions.txt') as f:
            content = f.read()
            print(content)
```

```
Good work on unzipping the file!
You should now see 5 folders, each with a lot of random .txt files.
Within one of these text files is a telephone number formated ###-###-####
Use the Python os module and regular expressions to iterate through each file, open it, and s
earch for a telephone number.
Good luck!
```

Step 3: Regular Expression to Find the Link

There are many approaches to take here, but since we know we are looking for a phone number, there should be a digits in the form ###-###-####, so we can easily create a regex expression for this and test it. Once its tested and working, we can figure out how to run it through all the txt documents.

In [9]:

```
import re
```

In [10]:

```
pattern = r'\d{3}-\d{3}-\d{4}'
```

In [11]:

```
test_string = "here is a random number 1231231234 , here is phone number formatted 123-123-12
```

In [14]:

```
re.findall(pattern,test_string)
```

Out[14]:

```
['123-123-1234']
```

Step 4: Create a function for regex

Let's put this inside a function that applies it to the contents of a .txt file, this way we can apply this function to all the txt files in the extracted_content folder.

In [15]:

```
def search(file,pattern= r'\d{3}-\d{3}-\d{4}'):
    f = open(file,'r')
    text = f.read()

    if re.search(pattern,text):
        return re.search(pattern,text)
    else:
        return ''
```

Step 5: OS Walk through the Files to Get the Link

Now that we have a basic function to search through the text of the files, let's perform an os.walk through the unzipped directory to find the links hidden somewhere in one of the text files.

In [16]:

```
import os
```

In [17]:

```
results = []
for folder , sub_folders , files in os.walk(os.getcwd()+"\\extracted_content"):

    for f in files:
        full_path = folder+'\\'+f

        results.append(search(full_path))
```

In [18]:

```
for r in results:
    if r != '':
        print(r.group())
```

```
719-266-2837
```

Excellent work! More information on this phone number:

- https://www.npr.org/2011/12/21/144069758/callin-oates-the-hotline-you-dont-need-but-might-call-anyway
- https://twitter.com/CallinOates