SECTION 14: ADVANCED PYTHON MODULES, 2 hours 23 minutes, 13 sections

13/13 Advanced Python Module Puzzle - Solution

- **-> this is the solutions to the previous problem**
  - -> we have a puzzle and solutions notebook
  - -> he's imported shell utilities and unpacket archives
  - -> we are unzipping a file
  - -> in the arguments, he gives the name of a file, where we want to extract it to and zip
  - -> he does through the file structure, in which case we can see the folders et al
  - -> then he uses the `with open` keywords to load the content
  - -> we have 5 folders with random text files <- latin ipsum
  - -> we have a telephone number in each of these
  - -> we want to iterate through each of the txt files, iterate through them and search for the telephone
  - -> import re
  - -> we are searching for the pattern
  - -> testing it on a single string
  - -> then re.findall() <- this is what we are using to find the test string in the txt file

- **-> then we are making a function to search for a pattern**
  - -> def search(file, pattern=...):
  - -> we are opening the file
  - -> and then reading the text of the file
  - -> we are then using an if block
  - -> opening the file, looking for a pattern
  - -> then creating the search function
  - -> walking through all the files
  - -> going through the files <- there are 5 folders
  - **-> so we are using the os module**
  - **-> like Bash, but in Python**
  - -> we are iterating through the different folders
  - -> the full current location, plus the extracted content folder
  - -> providing the full file path
  - -> we are iterating through each of the files
  - -> in this case, he is storing the file path in the name of a variable
  - -> we want to make sure we don't get path errors
  - -> results.append <- we are doing a walk through the entire folder
  - -> grabbing the file path, performing a search function

```
In [1]: import shutil

In [2]: shutil.unpack_archive('unzip_me_for_instructions.zip','','zip')

In [3]: 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 search for a telephone number.
Good luck!

In [4]: import re

In [5]: pattern = r'\d{3}-\d{3}-\d{4}'

In [8]: test_string = 'here is a phone number 123-123-1234'

In [9]: re.findall(pattern,test_string)
Out[9]: ['123-123-1234']

In [ ]: 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)

In [11]: import os

In [13]: results = []

In [14]: os.getcwd()+'\\extracted_content'
Out[14]: 'C:\\Users\\Marcial\\Pierian-Data-Courses\\Complete-Python-3-Bootcamp\\12-Advanced Python Modules\\08-Advanced-Python-Module-Exercise\\extracted_content'

In [ ]: 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 [22]: results
         None,
         None,
         None,
         None,
         <_sre.SRE_Match object; span=(1062, 1074), match='719-266-2837'>,
         None,
         None,
         None,
```

- -> and then taking a file and searching for the pattern
- -> returning a telephone number
- -> for f in files
- -> we are calling the search function on each of the files, to make sure we don't get any file path errors
- -> results.append
- **-> we are doing a walk through the entire content folder**
- -> then performing the search function which he's defined
- -> this takes the file and searches for a pattern
- -> then he uses for r in results
- -> for r in results:
- -> print the group
- -> this returns the telephone number, in this example
- -> this is a number you can call which will sing you a song
- -> the main thing we have to do is extract the telephone number