Content Copyright by Pierian Data

Opening and Reading Files

So far we've discussed how to open files manually, one by one. Let's explore how we can open files programatically.

Review: Understanding File Paths

In [2]:
```
pwd
```

Out[2]:
```
'C:\\Users\\Marcial\\Pierian-Data-Courses\\Complete-Python-3-Bootcamp\\12-Advanced Python Modules'
```

Create Practice File

We will begin by creating a practice text file that we will be using for demonstration.

In [3]:
```
f = open('practice.txt','w+')
```

In [4]:
```
f.write('test')
f.close()
```

Getting Directories

Python has a built-in os module that allows us to use operating system dependent functionality.

You can get the current directory:

In [1]:
```
import os
```

In [6]:
```
os.getcwd()
```

Out[6]:
```
'C:\\Users\\Marcial\\Pierian-Data-Courses\\Complete-Python-3-Bootcamp\\12-Advanced Python Modules'
```

Listing Files in a Directory

You can also use the os module to list directories.

In [7]:
```
# In your current directory
os.listdir()
```

Out[7]:
```
['.ipynb_checkpoints',
 '00-Collections-Module.ipynb',
 '01-Datetime-Module.ipynb',
 '01-Opening-and-Reading-Files.ipynb',
 '02-Math-and-Random-Module.ipynb',
 '03-Python Debugger (pdb).ipynb',
 '04-Timing your code - timeit.ipynb',
 '05-Overview-of-Regular-Expressions.ipynb',
 '06-Unzipping-and-Zipping-Files.ipynb',
 '07-OS-Module.ipynb',
 '08-Advanced-Python-Module-Exercise',
 'comp_file.zip',
 'Example_Top_Level',
 'extracted_content',
 'new_file.txt',
 'new_file2.txt',
 'practice.txt']
```

In [8]:
```
# In any directory you pass
os.listdir("C:\\Users")
```

Out[8]:
```
['admin.DESKTOP-O64BPTC',
 'All Users',
 'Default',
 'Default User',
 'defaultuser0',
 'desktop.ini',
 'Marcial',
 'Public']
```

Moving Files

You can use the built-in **shutil** module to to move files to different locations. Keep in mind, there are permission restrictions, for example if you are logged in a User A, you won't be able to make changes to the top level Users folder without the proper permissions, more info

In [9]:
```
import shutil
```

In [10]:
```
shutil.move('practice.txt','C:\\Users\\Marcial')
```

Out[10]:
```
'C:\\Users\\Marcial\\practice.txt'
```

In [11]:
```
os.listdir()
```

Out[11]:
```
['.ipynb_checkpoints',
 '00-Collections-Module.ipynb',
 '01-Datetime-Module.ipynb',
 '01-Opening-and-Reading-Files.ipynb',
 '02-Math-and-Random-Module.ipynb',
 '03-Python Debugger (pdb).ipynb',
 '04-Timing your code - timeit.ipynb',
 '05-Overview-of-Regular-Expressions.ipynb',
 '06-Unzipping-and-Zipping-Files.ipynb',
 '07-OS-Module.ipynb',
 '08-Advanced-Python-Module-Exercise',
 'comp_file.zip',
 'Example_Top_Level',
 'extracted_content',
 'new_file.txt',
 'new_file2.txt']
```

In [12]:
```
shutil.move('C:\\Users\\Marcial\practice.txt',os.getcwd())
```

Out[12]:
```
'C:\\Users\\Marcial\\Pierian-Data-Courses\\Complete-Python-3-Bootcamp\\12-Advanced Python Modules\\practice.txt'
```

In [13]:
```
os.listdir()
```

Out[13]:
```
['.ipynb_checkpoints',
 '00-Collections-Module.ipynb',
 '01-Datetime-Module.ipynb',
 '01-Opening-and-Reading-Files.ipynb',
 '02-Math-and-Random-Module.ipynb',
 '03-Python Debugger (pdb).ipynb',
 '04-Timing your code - timeit.ipynb',
 '05-Overview-of-Regular-Expressions.ipynb',
 '06-Unzipping-and-Zipping-Files.ipynb',
 '07-OS-Module.ipynb',
 '08-Advanced-Python-Module-Exercise',
 'comp_file.zip',
 'Example_Top_Level',
 'extracted_content',
 'new_file.txt',
 'new_file2.txt',
 'practice.txt']
```

Deleting Files

NOTE: The os module provides 3 methods for deleting files:

- os.unlink(path) which deletes a file at the path your provide
- os.rmdir(path) which deletes a folder (folder must be empty) at the path your provide
- shutil.rmtree(path) this is the most dangerous, as it will remove all files and folders contained in the path. **All of these methods can not be reversed! Which means if you make a mistake you won't be able to recover the file. Instead we will use the send2trash module. A safer alternative that sends deleted files to the trash bin instead of permanent removal.**

Install the send2trash module with:

```
pip install send2trash
```

at your command line.

In [14]:
```
import send2trash
```

In [15]:
```
os.listdir()
```

Out[15]:
```
['.ipynb_checkpoints',
 '00-Collections-Module.ipynb',
 '01-Datetime-Module.ipynb',
 '01-Opening-and-Reading-Files.ipynb',
 '02-Math-and-Random-Module.ipynb',
 '03-Python Debugger (pdb).ipynb',
 '04-Timing your code - timeit.ipynb',
 '05-Overview-of-Regular-Expressions.ipynb',
 '06-Unzipping-and-Zipping-Files.ipynb',
 '07-OS-Module.ipynb',
 '08-Advanced-Python-Module-Exercise',
 'comp_file.zip',
 'Example_Top_Level',
 'extracted_content',
 'new_file.txt',
 'new_file2.txt',
 'practice.txt']
```

In [16]:
```
send2trash.send2trash('practice.txt')
```

In [17]:
```
os.listdir()
```

Out[17]:
```
['.ipynb_checkpoints',
 '00-Collections-Module.ipynb',
 '01-Datetime-Module.ipynb',
 '01-Opening-and-Reading-Files.ipynb',
 '02-Math-and-Random-Module.ipynb',
 '03-Python Debugger (pdb).ipynb',
 '04-Timing your code - timeit.ipynb',
 '05-Overview-of-Regular-Expressions.ipynb',
 '06-Unzipping-and-Zipping-Files.ipynb',
 '07-OS-Module.ipynb',
 '08-Advanced-Python-Module-Exercise',
 'comp_file.zip',
 'Example_Top_Level',
 'extracted_content',
 'new_file.txt',
 'new_file2.txt']
```

Walking through a directory

Often you will just need to "walk" through a directory, that is visit every file or folder and check to see if a file is in the directory, and then perhaps do something with that file. Usually recursively walking through every file and folder in a directory would be quite tricky to program, but luckily the os module has a direct method call for this called os.walk(). Let's explore how it works.

In [18]:
```
os.getcwd()
```

Out[18]:
```
'C:\\Users\\Marcial\\Pierian-Data-Courses\\Complete-Python-3-Bootcamp\\12-Advanced Python Modules'
```

In [19]:
```
os.listdir()
```

Out[19]:
```
['.ipynb_checkpoints',
 '00-Collections-Module.ipynb',
 '01-Datetime-Module.ipynb',
 '01-Opening-and-Reading-Files.ipynb',
 '02-Math-and-Random-Module.ipynb',
 '03-Python Debugger (pdb).ipynb',
 '04-Timing your code - timeit.ipynb',
 '05-Overview-of-Regular-Expressions.ipynb',
 '06-Unzipping-and-Zipping-Files.ipynb',
 '07-OS-Module.ipynb',
 '08-Advanced-Python-Module-Exercise',
 'comp_file.zip',
 'Example_Top_Level',
 'extracted_content',
 'new_file.txt',
 'new_file2.txt']
```

In [2]:
```
for folder , sub_folders , files in os.walk("Example_Top_Level"):

    print("Currently looking at folder: "+ folder)
    print('\n')
    print("THE SUBFOLDERS ARE: ")
    for sub_fold in sub_folders:
        print("\t Subfolder: "+sub_fold )

    print('\n')

    print("THE FILES ARE: ")
    for f in files:
        print("\t File: "+f)
    print('\n')

    # Now look at subfolders
```

```
Currently looking at folder: Example_Top_Level


THE SUBFOLDERS ARE:
         Subfolder: Mid-Example-One
         Subfolder: Mid-Example-Two


THE FILES ARE:
         File: Mid-Example.txt


Currently looking at folder: Example_Top_Level\Mid-Example-One


THE SUBFOLDERS ARE:
         Subfolder: Bottom-Level-One
         Subfolder: Bottom-Level-Two


THE FILES ARE:
         File: Mid-Level-Doc.txt


Currently looking at folder: Example_Top_Level\Mid-Example-One\Bottom-Level-One


THE SUBFOLDERS ARE:


THE FILES ARE:
         File: One_Text.txt


Currently looking at folder: Example_Top_Level\Mid-Example-One\Bottom-Level-Two


THE SUBFOLDERS ARE:


THE FILES ARE:
         File: Bottom-Text-Two.txt


Currently looking at folder: Example_Top_Level\Mid-Example-Two


THE SUBFOLDERS ARE:


THE FILES ARE:
```

Excellent, you should now be aware of how to work with a computer's files and folders in whichever directory they are in. Remember that the os module works for any oeprating system that supports Python, which means these commands will work across Linux,MacOs, or Windows without need for adjustment.