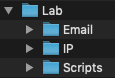
Reading and Writing Files

In this lab, you will be using Python’s file manipulation modules and Regex to learn how to extract and write information. It is assumed that you will be using Python 3 for this assignment. You will need the *data.txt* and *data2.txt* files to do this lab.

Task 1

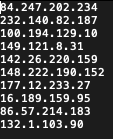
Use the os module to create a script that creates the following file structure:



Call the script directory.py and drag it into the *Scripts* folder. Drag *data.txt* and *data2.txt* into the *Lab* folder.

Task 2

The file *data.txt* has a secret string in it. The secret string follows the format **NNS.AAANN**. Where **N = a number**, **S = a special character** and **A = is an uppercase character**. Use Regex and the re module to find out the value of the secret string.



*IP Text File*

1. What Regex expression did you use? **[0-9]{2}[^\w]\.[A-Z]{3}[0-9]{2}**
2. What was the secret string? **12@.ERT45**

Task 3

Write a python script that extracts all the IP addresses from *data.txt* and writes them to a single *.txt*file in the IP folder. Make sure to use the re module and **Regex** to extract the numbers. The IP address takes the form **NNNN.NNNN.NNNN.NNNN** where **N** is a number. Note that each IP address may have 1-4 numbers per octet (per grouping separated by a period). Save the script in the *Scripts* folder and call it ip.py

1. What Regex expression did you use to extract the numbers? **[0-9]{1,3}\.[0-9]{1,3}\.[0-9]{1,3}\.[0-9]{1,3}**
2. What regex search function did you use? **findall()**

Task 3

Write a script named emails.py that writes a python script for each entry in *data.txt* and *data2.txt*. The scripts should be generated in the *Email* folder. Use fileinput to read both of the files. Use **Regex** and re module to extract the emails and first names. Each python script should be labeled email\_<total line number being read>.py. Additionally, create a shelve file in the Scripts folder called **emails** that stores all the emails and uses the corresponding first name as a key. This shelve file will be used by all the generated python scripts. Save the shelve file and email.py to your *Scripts* folder. Use pprint to write the array arr. The generated python files should look like this:

import shelve

email\_db = shelve.open(‘../Scripts/emails.db’)

arr = [<email>, <first\_name>]

print(email\_db[‘<first\_name>’],arr[0])

Running the generated python scripts should print two emails. Make sure those emails are the same.



1. How many files were created? **30 (note some students might answer 31. This is because 31 lines are read but no file is created for one of them)**
2. What Regex expression did you use to extract emails? **[^@\s]+@[^\.]+\.[a-z]+**

Zip your *Lab* folder and submit to blackboard.