**Log File Parsing**

Author : Kalyana Sundaram N

Problem statement:

In a log file, there would be details about servers. We want to parse that log file according to servername. That is, the log file will be in this format:

Timestamp,servername,applicationName,productName,comments

This log might contain information about many servers, that is , servername field might contain server1,server2,etc. through out the log file.

We wanted to get all the server1 records or lines in the log file and then create a new file with the name server1.txt and save those records.

Solution:

I have used python to solve the above problem.

I have created 4 py scripts to do this.

* Pyt.py – this will actually do the requirement
* Gen.py – This will generate input csv log files.
* Modul.py – This will be utilized to do some functions
* Basictest.py – this will be utilized to do unit testing.

Workaround :

* Copy all the four scripts into a folder in your system.
* Run gen.py- Asks for 3 inputs.
  + It will ask for path, where the input files would be generated and placed. Give the same path as where you had saved those 4 scripts.
  + The script again will ask for how many input files ,user wants to generate. Give any integer number.
  + The script again will ask for how many lines , user wants to print in the input file. Give any integer. Note that, default string is a 3 line text. If yougive 10 as your input, 3\*10 =30 lines will be printed in the input file.
* The input files will be generated in the path given by you. **Note: Please include slash (/ or \) at the end of your input path.**
* Run pyt.py – Asks for two inputs
  + Input 1: It will ask for a dir where the output files generated and should be placed. Give the same dir as where the script resides.
  + Input 2: Once the execution is done, it will ask for whether to proceed with Unit testing results or not. Providing Y or y will run unit testing and give you results.

How to verify data in output file

If input file contains following data:

01/01/2019 00:00:00,server1,application1,product1,comment1

01/01/2019 00:00:00,server2,application2,product2,comment2

Output should be below:

File : Server1.txt with content “01/01/2019 00:00:00,server1,application1,product1,comment1”

File: Server2.txt with content “01/01/2019 00:00:00,server2,application2,product2,comment2”

Above two files should be in the path given by the user at the start.