ID 2090 : Introduction to Scientific Computing
Summer 2021
Assignment – 1

Instructions:

1. Each question carries 5 marks.
2. Make a consolidated report of your submission as a tar file named after your rollnumber eg., ME20B001-Assn1.tar. When we untar this file, it should create a folder MM20B001-Assn1 which should contain your stuff including any images, screenshots, scripts, PDF files etc., Please try this yourself once and confirm before you upload.
3. Upload the tar file on moodle as per the deadline for the assignment. Late submissions are accepted but with a small penalty proportional to the delay. I shall decide the extent of this based on the actual submissions of the class.

[1] Write a script that generates a backup filename need to be created by a program with a unique name containing the time stamp when the backup was invoked. Run the script couple of times in quick succession and confirm that the file names are not repeated. Your script should also check if such a file does not exist and then use the command "touch" to create it.

Expected output: A script that can be executed to generate the empty files with a unique name each time it is run.

Application: During a project work, we need to archive our work time to time. While there is git available for program codes, one can also use the above mentioned strategy to create backup folders that can be moved out to another machine for safety.

[2] There is file by name "figs.tar.gz" in the datasets folder of google drive for this course. Download it to your machine, gunzip and untar it to see a set of folders and sub folders that contain some images. These images have wrong file extensions. Write a script that can be executed from the folder "figs" which will fix the file extensions to have them correctly indicating the type of image file. A list of "mv" commands manually entered is not acceptable.

Expected output: A script that can be executed from the "figs" folder and will generate a text output log about the file extensions being changed and how it is confirmed to be correct after change.

Application: We should not trust the filename for its contents. Also we may need to rename a whole bunch of files as required by certain software (eg., imagej to create a video out of a set of images) that expects the files to be named in a recognizable sequence for batch processing.

[3] There is a file by name "apache2.tar.gz" in the datasets folder of google drive for this course. Download it to your machine, gunzip and untar it to see a set of log files. Each of these files has access log of the apache web server in a certain format that includes the IP address of the machine that connected, the date time stamp, the URL requested and so on. Using these files, generate a date wise statistics of number of unique IP addresses that visited the server. You are welcome to choose a range for the dates that is reasonable as per your wish.

Expected output: A script that can be run in the "apache2" folder and gives a text output that contains two columns, namely, the date and the number of unique IP addresses. The script and the output file need to be submitted as part of the assignment.

Application: Scientific software (such as OpenFOAM, for example) create a bunch of folders with various output files in those. These folders are typically named after the time step by the program. We need to be able to visit some of those files to extract certain information from those files for analysis.

--- end of assignment---