PROJECT SPECIFICATIONS

The goal of this assignment is to assess your understanding of the use of the shell and open source tools to effectively report on and visualise data large datasets. You will be assessed on the clarity and quality of your shellscript(s) to examine and report on the data. While the efficiency of your shellscript will not be assessed, you should take care to avoid any excessive slow practices.

You are welcome to undertake the project on your home or laptop computers. Please note, however, that all materials submitted for marking must be working on a CSSE Linux computer by the due date.

Task 2:

The Department of Computer Science and Software Engineering runs its own small web-server, named secure.csse.uwa.edu.au, to support teaching related applications. As with most web-servers, each request is logged, one request per line, and each request's fields include: the requesting IP address, data and time of the request, the URL requested, the web-server's integer return code (indicating success or error), and the number of bytes transferred.

- The text file secure_access_log-20180506 provides the access logfile for a recent and typical week of activity (caution, file is 24MB).
- Don't forget, you can select smaller datasets (subsets) by using head and tail.
- This task asks you to develop at least 3 distinct graphical representations of the data in the logfile. Each representation must employ a different visualisation (chart) type. Only one representation may be a 'simple' one, such as a histogram showing the distribution of bytes delivered. The other visualisations should present some more insightful information, such as any URLs that are 'trending' across the week, or more meaningful descriptions of the locations from which requests are made.
- For this first task, you'll probably find it easiest to develop three distinct shellscripts, or three distinct shell functions in one shellscript, producing three distinct plots. Each shellscript, or shell function, should produce its own plot, which you may produce in three, or in just one, HTML webpage.