**ECE 3822: Software Tools FOr eNGINEERS**

# HOMEWORK No. 1: Command Line PROGRAMMING

The goal of this homework is to familiarize you with some basic command line tools to manipulate large repositories of text files. All assignments will be executed on the large text database provided on the course web site.

The tasks are:

1. Customize your .profile and .bashrc. Create a directory in your home directory called “bin.” Place an executable in that directory, such a shellscript or a binary file (C program) that prints “hello world.” Modify your path in your .bashrc so that your personal bin directory is included in your standard path. Demonstrate that you can execute the binary in your bin directory by typing it from any directory on your system. Define an alias for a convenient command, such as “ece\_3822\_d” that provides a custom directory listing (e.g., “ls -la”). Demonstrate that this is available from any shell in your login session.
2. Count the total number of directories in the database. Count the total number of files. Count the number of sessions that have a first name that begins with “R” and a last name that begins with “S” and occurred between 2010 and 2013.
3. Count the number of files in the database that have the word “spike” occurring at least once. Call this subset “A”. Repeat this for files containing the word “seizure,” which we will refer to as subset B. Produce a histogram of the distribution of these words in subsets A and B, listing the most frequently occurring first and least frequently occurring last. Show both the absolute counts and the percentage (often referred to as the cumulative distribution).
4. For subset A, produce a histogram of all two-word sequences that occur in this subset of the database. (Hint: search for “Unix for Poets” and follow one of the examples :)