MA5111 – Linux/Unix Exercises

1. File and Directory Commands
- Use pwd to print out your current directory to make sure of where you are
- Use mkdir to create directories data, proj1, and proj2
- Oops! - one too many, remove proj2
- Change into data and create files seqs1.txt and seqs2.txt using the touch command

2. File Creation, Editing, Copying, and Moving
Use nano to add 5 sequences to seqs1.txt and 3 to seqs2.txt, each sequence should have a sequence name preceded by a '>' char and then some nucleotides on the following line, for example:

>seq1
CAAACCCGTG
>seq2
CCCGCGCCTC

- Use cat to read your files in the terminal
- Use output redirection to append seqs1.txt and seqs2.txt to a new file called allseqs.txt
- Copy allseqs.txt to the proj1 directory
- Change into the proj1 directory and rename (move) the allseqs.txt file to myseqs.txt
- Try deleting the proj1 dir, what happens?

3. File Processing

Use wc to count the number of lines in myseqs.txt
Use head to view the first 6 lines of myseqs.txt
Use tail to view the last 4 lines of myseqs.txt
Use the diff command to compare seqs1.txt to myseqs.txt, output the results to a file called changes.txt, then try diff -y to show a side by side comparison u
Use the more and grep commands to output sequence names in myseqs.txt
Use the more and grep commands to count the sequences in myseqs.txt
Use grep to find any sequences containing CG dinucleotides in myseqs.txt
Use sed to convert all the T's in myseqs.txt to U's

If you have finished with these, try out some of the other commands from the notes / cheatsheet on Blackboard.