Bio 4342 Lab Homework 1:

Unix problems. Do each problem by yourself or with a neighbor. Some of the steps ask you to do something on your computer, for these feel free to experiment on you mac but write the correctly formatted command on this sheet. Or if you want to can open this document in word and fill it in.

1. Do the following:
    a. Execute the following 2 commands (can you explain what they are doing?)
        cd

        cp ../Shared/tostudents/ ./*

    b. Create a subdirectory called *temp*

    c. There is a file in your directory called *practice.txt*. Make a copy of *practice.txt* inside the *temp* directory

    d. Make another copy of the practice file in *temp* with the name *prac.txt*

    e. Change the current working directory to *temp*

    f. List all the files in the directory *temp*

    g. Find the contents of the file *prac.txt* and *practice.txt*. Are they the same? There are two commands to look at the contents of a file what are they?

    h. Change the current working directory to you home directory

    i. Execute the following commands and compare how they work with respect to files, directories and the contents of directories:
        ls

        ls practice.txt

        ls temp

        *ls **

    j. Clean up by removing the temp directory and all files within

2. Try all of problem 1 again in a different order (write out the new commands below):
    Try doing the steps in this order: *a,b,e,c,d,h,f,g*. How do steps *d,c,f* and *g* change?

3. Lets do some more simple commands but with more files:
   a. Set your current directory to your home directory

   b. Make a new directory called *lots*

   c. List out all the files in */Users/Shared/lots*. There are almost 1000 files. The list will probably scroll by too quickly you can scroll back in your terminal window to view them.

   d. Each one is a small file. These files have only a little text inside each one but we can pretend they are the real sequence files generated with the sequencing machines. They have names like *jep71c01.b1* or *jep71c01.g1*. These names reflect the plate "jep71" the well "c01" and if they are forward "b1" or reverse "g1" reads.

   e. How would you copy all the files in the lots directory inside *Shared/* to the directory you made in step b above?

   f. Make a subdirectory call *some*

   g. How would you copy only those files that start with k from *lots* into *some*.

   h. Lets say you find that the sequencing machine had a problem and well C5 reads were defective. How would you delete all the files that came from this well?

      i. How would you delete all the reactions from all the A row from all the plates?

      ii. How would you delete all the reactions from plates in the ktt series?

      iii. How would you delete all the reactions from plate lqq82 but only from column 8?

   i. If you have been working in your computer you may want to clean up by deleting all files inside the lots folder and deleting the lots folder itself. The commands to do this are:

4. If you completed step 1 there is a file in your home directory called *unknown.pep*. It is a peptide sequence in GENPEPT file format.
   a. Can you tell the organism this that this sequence came from?

   b. Can you tell what protein this is?

5. There is another file in your home directory called unknown.asn1.

       a. Can you tell the organism that this sequence if from?

       b. Can you tell what the protein is from the information in the file?

       c.What might you do to really know for sure what this protein is?

6. There is another file in your home directory called *unknown.fasta*. It is a peptide sequence in fasta format.

       a. Can you tell the organism that this sequence came from?

       b. Can you tell what the protein is?