Imani Martinez

Lab Assignment 3

1. Give the wget command to download the Lab3 folder on GitHub to your computer.

wget https://raw.githubusercontent.com/raw-lab/BINF2111/main/data/Lab3.tar.gz

2. Give the command to extract and unzip the Lab3 folder.

gunzip Lab3.tar.gz

3. Give the command to unzip lab3_EFMCounts.csv.gz (found in the Lab3 folder you just extracted/unzipped).

tar -xvf Lab3.tar

gunzip Lab3/lab3_EFMCounts.csv.gz

4. Use the printf command to answer the following questions in a text file. Each answer should be on a new line.

What is your name?
How are you doing today?
When is your birthday?
Is there anything you are still confused about?

printf "Imani\nGood\nJuly 8th\nNo\n" answers.txt

5. Give the command to print all but the 3rd, 4th, and 5th columns of lab3_EFMCounts.csv into a new file.

cut -f 3,4,5 --complement lab3_EFMCounts.csv lab3_EFM.txt

6. Give the command(s) to count the number of images that are of high quality in lab3_EFMCounts.csv. Hint: Image quality occurs in column 7.

cut -f 7 lab3_EFMCounts.csv | grep -c "high"

7. Provide two unique commands to convert lab3_EFMCounts.csv to a TSV. Both commands cannot use the same base command (use sed & tr, sed & awk, etc. instead of sed & sed).

sed ‘s/\t/ ,/g’ lab3_EFMCounts.csv > lab3_EFMCounts.tsv

cat lab3_EFMCounts.csv | tr -s ',' '\t' > lab3_EFMCounts.tsv

8. Using lab3_EFMCounts.csv OR the TSV you created in question 7, find the top 10 images with the highest counts. Hint: Counts are in column 6.

sort -k 6 lab3_EFMCounts.csv
sort -t "," -nr -k6 lab3_EFMCounts.csv

9. Give the command(s) to write a bash script that, when ran, will print out "Hello World". Be sure that the file ends in .sh!

Use nano to create a bash script

#!/bin/bash

echo ‘Hello World’

Then in command line

bash hello_world.sh

10. Using any text editor, give the steps/commands to create a new file, write "This is the answer to lab 3, question 10.", and save and exit. You are not allowed to use printf or echo.

nano to create a bash script

#!/bin/bash

cat << EOF

This is the answer to lab 3, question 10.

EOF

In my terminal I then ran bash question10.sh

BONUS I (2 pts): Find a way to convert lab3_EFMCounts.csv to SQL.

csvsql --insert --tables lab3_efmcounts --dialect sqlite lab3_EFMCounts.csv > lab3_EFMCounts.sql

BONUS II (2 pts): Convert any nucleotide FASTA into amino acids using any command(s) you want.

transeq -sequence input_nucleotide.fna -outseq output_protein.faa