Week 2 Exercises

## In-class exercises

1. We can run the script on a file (assuming we are in bash\_for\_bio/):

./scripts/week2/sam\_count.sh ./data/my\_file.bam

Try out the above script on ./data/CALU1\_combined\_final.sam.

Make sure you are in the top folder (bash\_for\_bio) when you do it.

1. Run run\_bwa.sh using:

./scripts/week2/run\_bwa.sh ./data/CALU1\_combined\_final.fasta

1. Run process\_data.R:

module load fhR  
Rscript process\_data.R input\_file="LUSC\_clinical.csv"  
module purge

1. Run process\_data.py:

module load fhPython  
python3 process\_file.py lusc\_file.csv  
module purge

## Homework Exercises

All exercises are required for the badge.

1. Copy the below script into a file called samtools\_count.sh. What does the script do?

#!/bin/bash  
module load SAMtools  
samtools view -c $1

How would we modify the above script to redirect the output to ${1}.counts.txt?

1. Make sure you can get scripts/week2/run\_bwa.sh to run on rhino.
2. Modify scripts/week2/run\_bwa.sh to
   1. Take an additional argument, a folder path
   2. Save the SAM file to this folder path

Hint 1: I recommend that you copy run\_bwa.sh into a new script and work from there. Put your code into the codeblock below:

Write an example to run your new version of the script below and run it:

For question 4, pick one language to answer.

4R. (R) Modify the below R script and save it to a file called scripts/week2/r\_csv\_script.R. It should also take an argument called $FILEPATH for read.csv():

library(tidyverse)  
csv\_file <- read\_csv("myfile.csv")  
summary(csv\_file)

How would you run this on the command line?

module load fhR  
-----  
module purge

How would you redirect the output of your script to a file?

4Py. (py) Modify the below Python script to be runnable and save it to scripts/week2/py\_csv\_script.py. Your new version should also take the first position argument (a file path) and process the file:

#| eval: false  
import pandas as pd  
csv\_file = pd.read\_csv("my\_file.csv")  
csv\_file.describe()

How would you run this on the command line?

module load fhPython  
-----  
module purge