Short read workshop 2023 -Day 6

Authors: Georgia Barone & Rutendo Sigauke

Part 1- Learning R in RStudio

**** Before beginning this worksheet, make sure RStudio is downloaded on your local computer**

1. Log in to AWS
2. cd into the sr2023 repository and run a "git pull"
3. Open/complete the Learning_R.R script on the sr2023 GitHub day06 repository under scripts

** Learning_R.R is an R script, so make sure to open it in RStudio

```
[GEFB2000@ip-172-31-29-36 ~ $ cd sr2023/
[GEFB2000@ip-172-31-29-36 ~/sr2023 $ git pull
```

Part 2- Writing an R script to submit on a supercomputer

1. cd into /scratch/Users/<your_username>/
2. Use the mkdir command to make a folder called "workshop-day6"
3. Inside your workshop-day6 folder, make the folders "scripts", "eofiles", & "results"

```
[GEFB2000@ip-172-31-29-36 ~/sr2023 $ cd /scratch/Users/GEFB2000/
[GEFB2000@ip-172-31-29-36 /scratch/Users/GEFB2000 $ mkdir workshop-day6
[GEFB2000@ip-172-31-29-36 /scratch/Users/GEFB2000/workshop-day6 $ mkdir scripts eofiles results
```

4. Use the rsync or scp command to copy Learning_R_submit_aws.R and Submit_Rscript.sbatch from /Users/ <your_username> /sr2023/day06/scripts, into the scripts directory in /scratch/Users/<your_username>/workshop-day6 you just made.

5. Go back to /scratch/Users/<your_username>/workshop-day6. Enter the scripts directory you made and use vim to open and edit the Learning_R_submit_aws.R file. Add your own working directory path (it should be something along the lines of: /scratch/Users/<your_username>/workshop-day6/results).

```
## Learning R : example for submitting in sbatch
## Author: Taylor Jones (2022), Rutendo Sigauke (2023)

##############################################################################
## In this script we plan to submit the R script as an SBATCH job           ##
## This is useful for when you have a computationally intensive job that     ##
## requires more resources than on your personal computer.                   ##
## For example, counting reads in a bam file over gene features for RNA-seq. ##
##############################################################################

##############################################################################
# Set your working directory first
workdir <- '/path/to/working/directory/'
setwd(workdir)
getwd()
```

6. Look through the rest of the script before saving and exiting vim, to make sure you know what the code is doing and where your output will be saving to.

7. Now use vim to open/edit the Submit_Rscript.sbatch file. This is the sbatch script we will be using to submit our code to the supercomputer. Edit the script by adding your eofiles path, email, and path to Rscript .

8. Once you are happy with your Learning_R_submit_aws.R and Submit_Rscript.sbatch scripts, submit Submit_Rscript.sbatch.

9. If the script worked, copy mtcars.csv & mtcars_mpg_wt_scatterplot.png to your local computer to view.