Using Deseq counts walk through

Step 1) Copy our scripts

Log in to the super computer
Cd into the new directory
Make a directory /scratch/User/yourusername/day7/
Copy the two scripts we have provided in
/scratch/Shares/public/sread2021/scripts/day7/

```
[maallen3@ip-172-31-38-192 ~]$ mkdir /scratch/Users/maallen3/day7/
[maallen3@ip-172-31-38-192 ~]$ cd /scratch/Users/maallen3/day7/

[maallen3@ip-172-31-38-192 day7]$ scp /scratch/Shares/public/sread2021/scripts/day7/* /scratch/Users/maallen3/day7/
[maallen3@ip-172-31-38-192 day7]$ ls
DESeq2_example.R  sr_deseq2.sbatch
[maallen3@ip-172-31-38-192 day7]$
```

Step 2) Edit Our scripts

In the R script Change your working directory
In the sbatch script change your email and error and output files
####Make sure the error and output directory exist before you run!!!!!!

```
[maallen3@ip-172-31-38-192 day7]$ vi DESeq2_example.R
[maallen3@ip-172-31-38-192 day7]$ vi sr_deseq2.sbatch
```

```
# For more information see: http://www.bioconductor.org/help/workflows/rnaseqGene/
library("tidyverse")
library("DESeq2")
#set working dir
workdir <- '/scratch/Users/username/day7/'
dir.create(workdir, showWarnings = FALSE)
setwd(workdir)
getwd()
outdir <- paste(workdir, 'deseqresults', '/', sep='') ##naming our outdir
dir.create(outdir, showWarnings = FALSE) ###creating the directory
```

becomes

```
# For more information see: http://www.bioconductor.org/help/workflows/rnaseqGen
e/
library("tidyverse")
library("DESeq2")
#set working dir
workdir <- '/scratch/Users/maallen3/day7//'
dir.create(workdir, showWarnings = FALSE)
setwd(workdir)
getwd()
outdir <- paste(workdir, 'deseqresults', '/', sep='') ##naming our outdir
dir.create(outdir, showWarnings = FALSE) ###creating the directory
```

```
#!/bin/bash
#SBATCH --job-name=featurecounts                          # Job name
#SBATCH --mail-type=ALL                        # Mail events (NONE, BEGIN, END, FAIL, ALL)
#SBATCH --mail-user=allenma@colorado.edu               # Where to send mail
#SBATCH --nodes=1                                      # Number of cores job will run on
#SBATCH --ntasks=4                                     # Number of CPU (processers, tasks)
#SBATCH --time=2:00:00                                 # Time limit hrs:min:sec
#SBATCH --partition compute                            # Job queue
#SBATCH --mem=4gb                                      # Memory limit
#SBATCH --output=/scratch/Users/maallen3/eofiles/%x_%j.out
#SBATCH --error=/scratch/Users/maallen3/eofiles/%x_%j.err
```

becomes

```
#!/bin/bash
#SBATCH --job-name=featurecounts                          # Job name
#SBATCH --mail-type=ALL                        # Mail events (NONE, BEGIN, END, FAIL, ALL)
#SBATCH --mail-user=username@email.edu         # Where to send mail
#SBATCH --nodes=1                                      # Number of cores job will run on
#SBATCH --ntasks=4                                     # Number of CPU (processers, tasks)
#SBATCH --time=2:00:00                                 # Time limit hrs:min:sec
#SBATCH --partition compute                            # Job queue
#SBATCH --mem=4gb                                      # Memory limit
#SBATCH --output=/scratch/Users/username/eofiles/%x_%j.out
#SBATCH --error=/scratch/Users/username/eofiles/%x_%j.err
```

Step 3) Submit the sbatch script to the queue
The sbatch script runs the R script… how?

```
[maallen3@ip-172-31-38-192 day7]$ sbatch sr_deseq2.sbatch
```

Step 4) If it works you will end up with a directory named `deseqresults` in your working directory. In the `deseqresults` directory you will end up with two files. One is a MA plot. The other is the results table.

```
[maallen3@ip-172-31-38-192 day7]$ ls
DESeq2_example.R  deseqresults  sr_deseq2.sbatch
[maallen3@ip-172-31-38-192 day7]$ cd deseqresults/
[maallen3@ip-172-31-38-192 deseqresults]$ ls
MA_ifn_vs_ctrl_DEA.pdf  my_results.csv
[maallen3@ip-172-31-38-192 deseqresults]$
```

```
[maallen3@ip-172-31-38-192 deseqresults]$ head my_results.csv
"","baseMean","log2FoldChange","lfcSE","stat","pvalue","padj"
"ISG15",5467.28192690661,4.02749093689695,0.112807239370585,35.6707026862339,1.12550897691298e-278,1.36017759859933e-275
"TNFRSF4",38.983654980864,-0.662165889429076,0.231044329100826,-2.86299221897849,0.00419660868982427,0.020700414700623
"B3GALT6",504.191142890681,-0.319333608451855,0.0949507871169556,-3.36303105878293,0.000770916898415572,0.00446996185536295
"PUSL1",308.2871202625,-0.410790230588219,0.106697964452489,-3.84983378299419,0.000118198009167511,0.000871748961248965
"CPTP",568.314931997898,-0.449697043154923,0.0979394702638105,-4.59145859944207,4.40158963059598e-06,4.22887535264358e-05
"DVL1",480.122153925568,-0.307277948704738,0.0879016038524443,-3.49568882445076,0.000472839767304362,0.00301772011430498
"LOC148413",396.45089057145,-0.21439574052623,0.0672310874348464,-3.18883926592621,0.00142845279829772,0.00803190192568932
"ATAD3B",1939.33985769226,-0.269039934371458,0.0854022437360979,-3.15024081472403,0.00163135933884124,0.00904949791929669
"ATAD3A",2510.8096246294,-0.243877012912492,0.0589796648424328,-4.13491532821074,3.55085530241766e-05,0.000291918954623928
```