**How to Submit batch jobs using the HPC**

1. Need to get files onto HPC open on demand profile
   1. [Documentation](https://hpcdocs.hpc.arizona.edu/storage_and_transfers) is good on this
   2. I used [Globus](https://hpcdocs.hpc.arizona.edu/storage_and_transfers/transfers/globus/) which was really easy with large files like genetics data
   3. Once it is installed it will appear on your bar
      1. A screenshot of a computer

         AI-generated content may be incorrect.
      2. Make sure it is turned on before use again
   4. Tip: make a subfolder with everything you need in it for the analysis. Makes writing the slurm job easier

A screenshot of a computer

AI-generated content may be incorrect.

1. Writing batch job Script
   1. First you need to write a batch job script in a text file. This file will include your request to the scheduling platform, and where to find the data and what code to run. An easy way to create this file is by using shell assess in the HPC on demand and using nano to write it.
   2. Step 1: Select the “Open in terminal“ button in the file you want this script to go in.

A screenshot of a computer

AI-generated content may be incorrect.

* 1. Step 2: Type “nano” and whatever else you want to name you nano created file, like “my\_job.slurm”. This creates a text file called “my\_job.slurm”
  2. Step 3: Write your batch script.
     1. There are plenty of details online about how to do this (See [this page](https://hpcdocs.hpc.arizona.edu/running_jobs/batch_jobs/intro/) on the website). I will paste an example of one I used below.

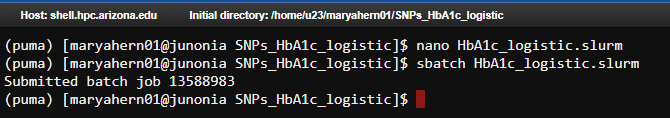
#!/bin/bash  
  
#SBATCH --job-name=R.SNPs.glucose.linear  
#SBATCH --ntasks=10  
#SBATCH --nodes=1  
#SBATCH --mem=5gb  
#SBATCH --time=96:00:00  
#SBATCH --partition=standard  
#SBATCH --account=skrishnan  
#SBATCH --output=output.glucose.linear.log  
  
module load R/4.4  
R -f glucose\_linear.batch.R

* 1. Step 4: Get out of nano
     1. Here you cannot use your mouse so you have to use key shortcuts to save and exit Nano. Hit:
        1. Ctrl+O (to write)
        2. Enter to save file name
        3. Ctrl+X to exit.
  2. Step 5: Now, back in the terminal for the folder of the files you want to run, simply type “sbatch my\_job.slurm”. Sbatch is the command to start it and my\_job.slurm can be replaced with whatever you named that nano file you created. The terminal should respond with something like “job created 134532” that number corresponds to your job number.
     1. If you want to double check that it worked, go to the “jobs” and select “active jobs”

A screenshot of a computer

AI-generated content may be incorrect.

This is what the output will look like to let you know that it is successfully submitted



#!/bin/bash

#SBATCH --job-name=R.SNPs.PLINK.prep.slurm

#SBATCH --ntasks=20

#SBATCH --nodes=10

#SBATCH --time=24:00:00

#SBATCH --partition=standard

#SBATCH --account=skrishnan

#SBATCH --output=output.SNPs.PLINK.prep.log

module load R/4.4

R -f Step-1\_Prep\_data\_PLINK.batch.R

#!/bin/bash

#SBATCH --job-name=R.SNPs.glucose.linear.finalQC

#SBATCH --ntasks=10

#SBATCH --nodes=5

#SBATCH --mem=5gb

#SBATCH --time=24:00:00

#SBATCH --partition=standard

#SBATCH --account=skrishnan

#SBATCH --output=output.glucose.linear.finalQC.log

module load R/4.4

R -f Glucose\_linear\_EN\_finalQC.batch.R

#!/bin/bash

#SBATCH --job-name=R.SNPs.a1c.linear.finalQC

#SBATCH --ntasks=10

#SBATCH --nodes=5

#SBATCH --mem=5gb

#SBATCH --time=24:00:00

#SBATCH --partition=standard

#SBATCH --account=skrishnan

#SBATCH --output=output.a1c.linear.finalQC.log

module load R/4.4

R -f HbA1c\_linear\_EN\_finalQC.batch.R

Notes

* Requesting enough memory
  + Core and CPU are interachangeable
  + N tasks is how you specify cores
  + Each node has 5 GB of data
  + If you are getting “OOM” execution codes this means out of memory. The easiest way to get more memory is to increase the ntasks