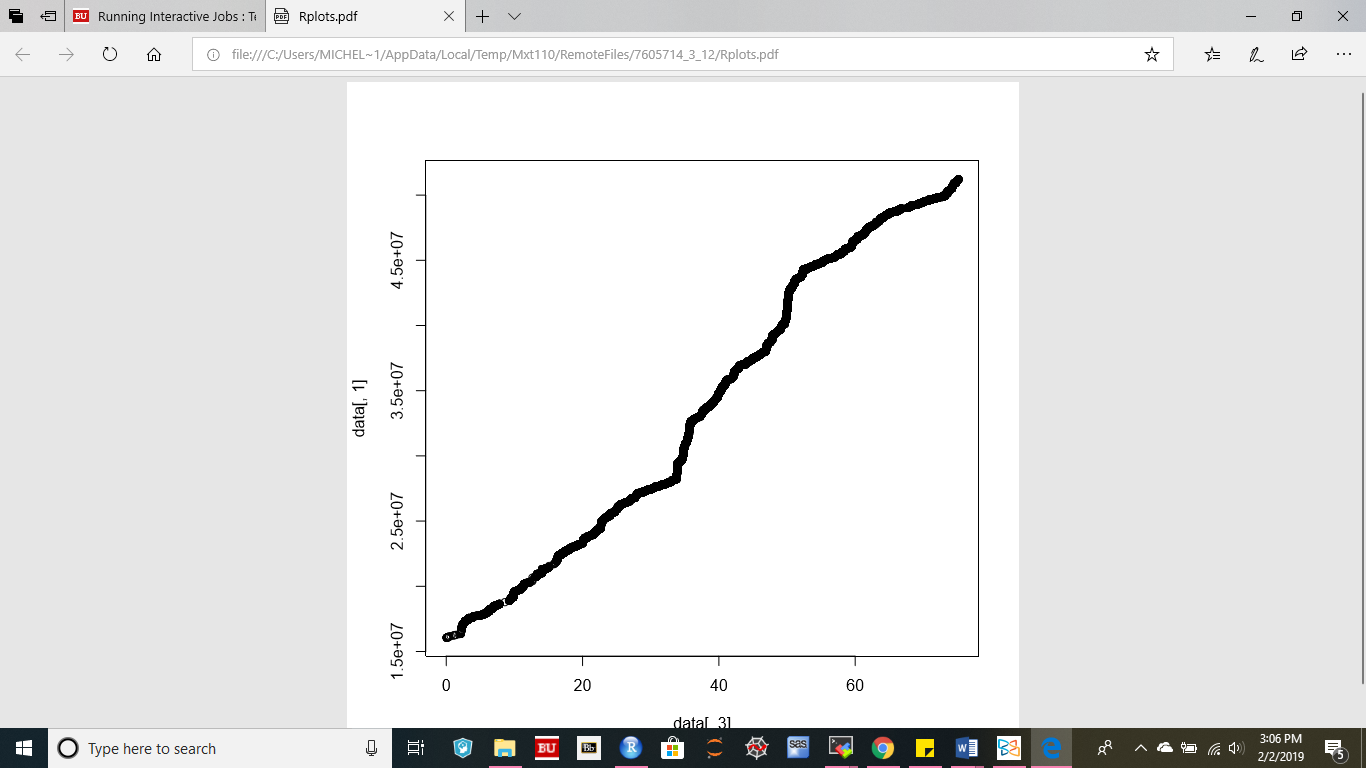
* **6 (command file - gedit)**
* #!/bin/bash -l  
    
  # All lines starting with "#" are comments  
  # All lines starting with "#$" are SGE qsub commands  
    
  # Uncomment and specify a project to use (Required for BUMC projects).  
  #$ -P bs859  
    
  # Give this job a name  
  #$ -N test  
    
  # Join standard output and error to a single file  
  #$ -j y  
    
  # Name the file where to redirect standard output and error  
  # #$ -o test.qlog  
    
    
  # Send an email when the job begins and when it ends running  
  # (b = when job begins, a = if job aborts, e = when job ends)  
  #$ -m bae  
    
    
  #echo $DATADIR  
  export DATADIR=/projectnb/bs859/data/1000G/1000GP\_Phase3   
  ls $DATADIR  
  cd $DATADIR  
  wc -l 1000GP\_Phase3.sample  
  wc -l genetic\_map\_chr22\_combined\_b37.txt   
  head -10 genetic\_map\_chr22\_combined\_b37.txt  
  tail -10 genetic\_map\_chr22\_combined\_b37.txt  
    
  # Load any modules you might want to use.  
  # e.g. module load R/3.2.3  
  module load R/3.2.3  
    
  # Run the commands you need to run for your job  
  R CMD BATCH test.R
* **6**
* [mchsu@scc1 class1]$ R --vanilla < test.R
* **(R script)**

data <- read.table("/projectnb/bs859/data/1000G/1000GP\_Phase3/genetic\_map\_chr22\_combined\_b37.txt", header = T)

plot(data[,3],data[,1])

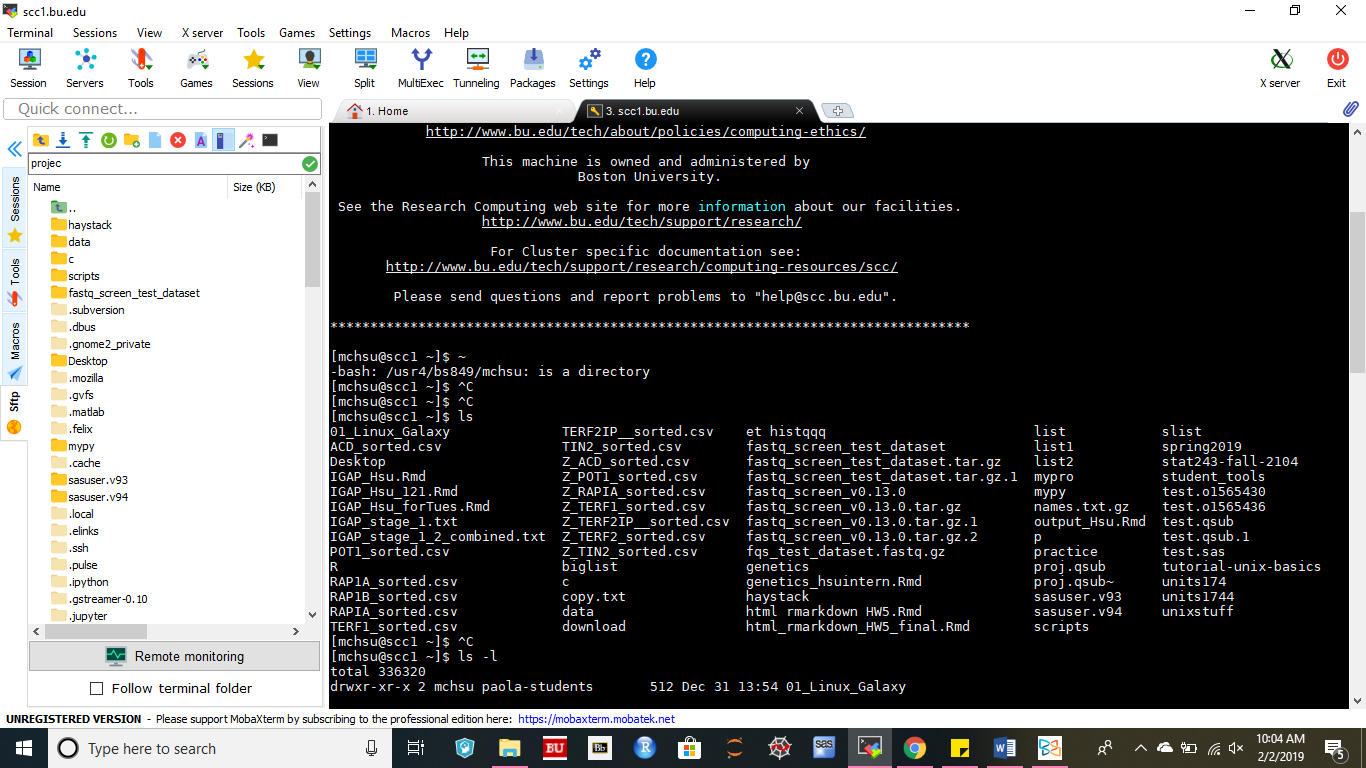
* **6 (interactively)**
* [mchsu@scc1 1000GP\_Phase3]$ qrsh -P bs859
* [mchsu@scc-pi3 class1]$ DATADIR=/projectnb/bs859/data/1000G/1000GP\_Phase3
* [mchsu@scc-pi3 class1]$ cd $DATADIR
* [mchsu@scc-pi3 class1]$ ls $DATADIR
* [mchsu@scc-pi3 1000GP\_Phase3]$ wc -l 1000GP\_Phase3.sample
* [mchsu@scc-pi3 1000GP\_Phase3]$ wc -l genetic\_map\_chr22\_combined\_b37.txt
* [mchsu@scc-pi3 1000GP\_Phase3]$ head -10 genetic\_map\_chr22\_combined\_b37.txt
* [mchsu@scc-pi3 1000GP\_Phase3]$ tail -10 genetic\_map\_chr22\_combined\_b37.txt
* **6 (qsub)**
* **[mchsu@scc-pi3 class1]$ cat test.o1795793 > HW1.log**
* 1000GP\_Phase3.sample  
  1000GP\_Phase3\_chr1.hap.gz  
  1000GP\_Phase3\_chr1.legend.gz  
  1000GP\_Phase3\_chr10.hap.gz  
  1000GP\_Phase3\_chr10.legend.gz  
  1000GP\_Phase3\_chr11.hap.gz  
  1000GP\_Phase3\_chr11.legend.gz  
  1000GP\_Phase3\_chr12.hap.gz  
  1000GP\_Phase3\_chr12.legend.gz  
  1000GP\_Phase3\_chr13.hap.gz  
  1000GP\_Phase3\_chr13.legend.gz  
  1000GP\_Phase3\_chr14.hap.gz  
  1000GP\_Phase3\_chr14.legend.gz  
  1000GP\_Phase3\_chr15.hap.gz  
  1000GP\_Phase3\_chr15.legend.gz  
  1000GP\_Phase3\_chr16.hap.gz  
  1000GP\_Phase3\_chr16.legend.gz  
  1000GP\_Phase3\_chr17.hap.gz  
  1000GP\_Phase3\_chr17.legend.gz  
  1000GP\_Phase3\_chr18.hap.gz  
  1000GP\_Phase3\_chr18.legend.gz  
  1000GP\_Phase3\_chr19.hap.gz  
  1000GP\_Phase3\_chr19.legend.gz  
  1000GP\_Phase3\_chr2.hap.gz  
  1000GP\_Phase3\_chr2.legend.gz  
  1000GP\_Phase3\_chr20.hap.gz  
  1000GP\_Phase3\_chr20.legend.gz  
  1000GP\_Phase3\_chr21.hap.gz  
  1000GP\_Phase3\_chr21.legend.gz  
  1000GP\_Phase3\_chr22.hap.gz  
  1000GP\_Phase3\_chr22.legend.gz  
  1000GP\_Phase3\_chr3.hap.gz  
  1000GP\_Phase3\_chr3.legend.gz  
  1000GP\_Phase3\_chr4.hap.gz  
  1000GP\_Phase3\_chr4.legend.gz  
  1000GP\_Phase3\_chr5.hap.gz  
  1000GP\_Phase3\_chr5.legend.gz  
  1000GP\_Phase3\_chr6.hap.gz  
  1000GP\_Phase3\_chr6.legend.gz  
  1000GP\_Phase3\_chr7.hap.gz  
  1000GP\_Phase3\_chr7.legend.gz  
  1000GP\_Phase3\_chr8.hap.gz  
  1000GP\_Phase3\_chr8.legend.gz  
  1000GP\_Phase3\_chr9.hap.gz  
  1000GP\_Phase3\_chr9.legend.gz  
  genetic\_map\_chr10\_combined\_b37.txt  
  genetic\_map\_chr11\_combined\_b37.txt  
  genetic\_map\_chr12\_combined\_b37.txt  
  genetic\_map\_chr13\_combined\_b37.txt  
  genetic\_map\_chr14\_combined\_b37.txt  
  genetic\_map\_chr15\_combined\_b37.txt  
  genetic\_map\_chr16\_combined\_b37.txt  
  genetic\_map\_chr17\_combined\_b37.txt  
  genetic\_map\_chr18\_combined\_b37.txt  
  genetic\_map\_chr19\_combined\_b37.txt  
  genetic\_map\_chr1\_combined\_b37.txt  
  genetic\_map\_chr20\_combined\_b37.txt  
  genetic\_map\_chr21\_combined\_b37.txt  
  genetic\_map\_chr22\_combined\_b37.txt  
  genetic\_map\_chr2\_combined\_b37.txt  
  genetic\_map\_chr3\_combined\_b37.txt  
  genetic\_map\_chr4\_combined\_b37.txt  
  genetic\_map\_chr5\_combined\_b37.txt  
  genetic\_map\_chr6\_combined\_b37.txt  
  genetic\_map\_chr7\_combined\_b37.txt  
  genetic\_map\_chr8\_combined\_b37.txt  
  genetic\_map\_chr9\_combined\_b37.txt  
  2505 1000GP\_Phase3.sample  
  43415 genetic\_map\_chr22\_combined\_b37.txt  
  position COMBINED\_rate(cM/Mb) Genetic\_Map(cM)  
  16051347 9.6640973708 0  
  16052618 9.7078062447 0.0123386217370137  
  16053624 9.7138922111 0.0221107973013803  
  16053659 9.716343506 0.0224508693240903  
  16053758 9.707808785 0.0234119423938053  
  16054713 9.5801157234 0.0325609529096523  
  16054960 9.5805112114 0.0349273391788681  
  16055070 6.2364537599 0.0356133490924571  
  16055171 6.2367174178 0.0362432575516549  
  51183017 0.7630419388 75.0889612921482  
  51196164 0.7605385596 75.0989600925913  
  51199851 0.7584997972 75.1017566813436  
  51211392 0.7581924384 75.1105069802751  
  51212875 0.7588786842 75.1116323973638  
  51216564 0.7590146945 75.1144324025718  
  51217134 0.7593957753 75.1148652581637  
  51222100 0.7655011432 75.1186667368409  
  51223637 0.7767175209 75.1198605516705  
  51229805 0 75.1198605516705  
  /share/pkg/r/3.2.3/install/lib64/R/bin/BATCH: line 60: test.Rout: Permission denied
* **6 -batch**
* [mchsu@scc-pi3 class1]$ qsub -P bs859 proj2.qsub
* Your job 1795793 ("test") has been submitted
* **6 plot**

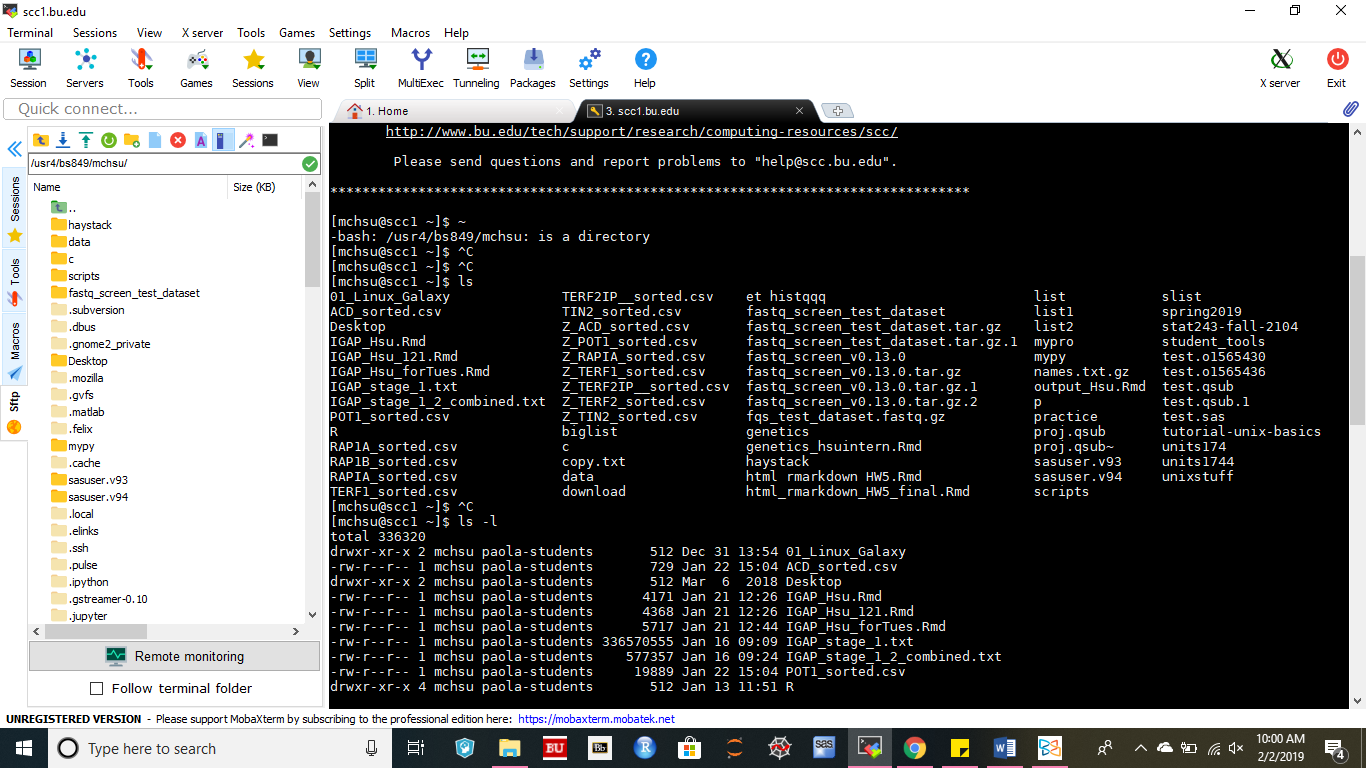


Additional questions:

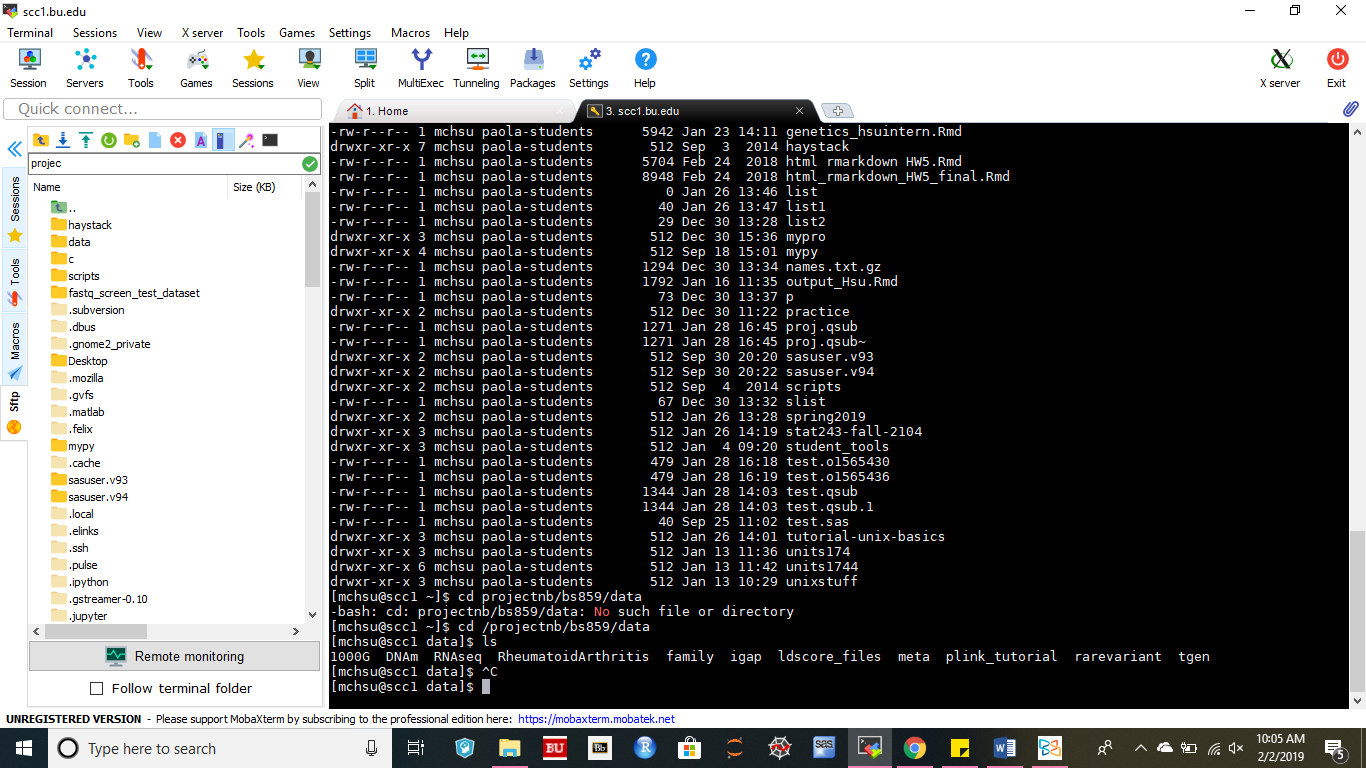
4. Locate the following important locations, and display a listing of the files in each.

a) Your login directory ~username;

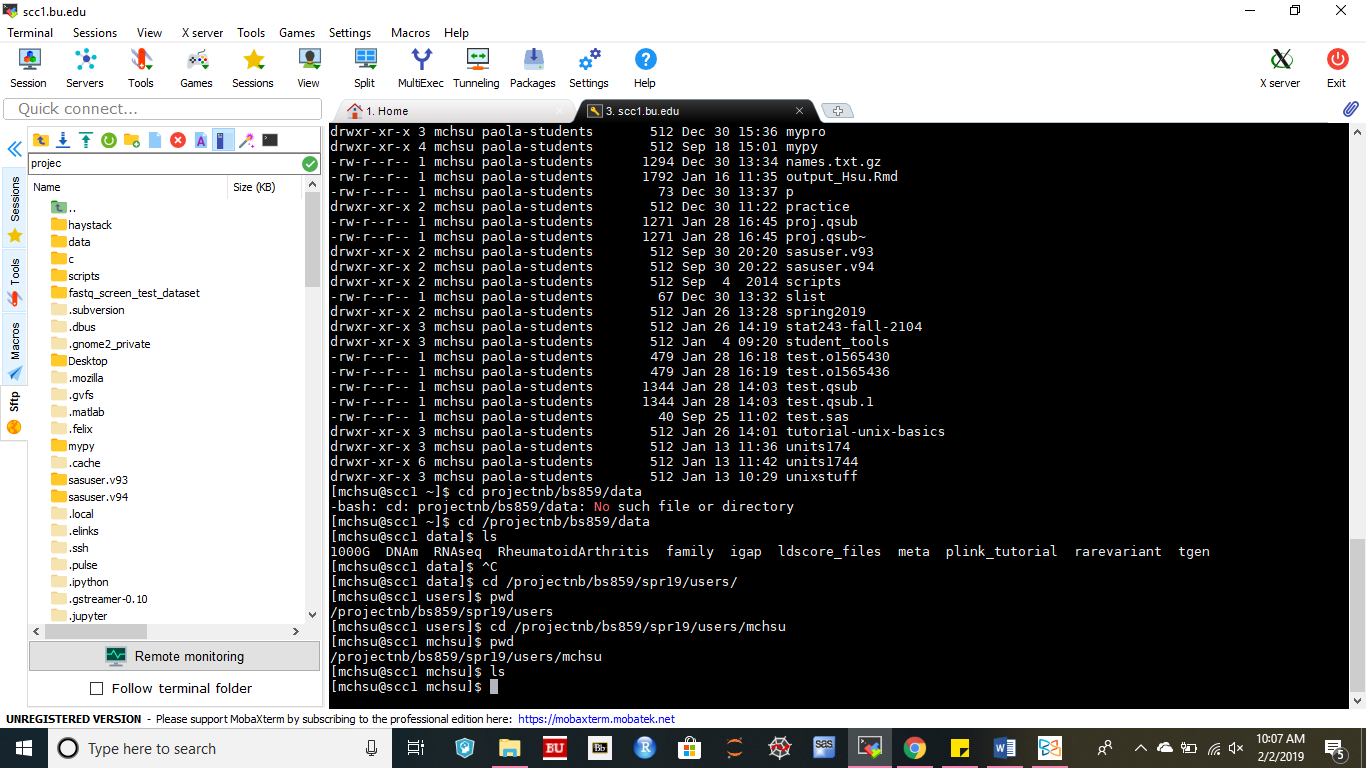




b) The Data Directory where many of the BS859 data sets are stored: /projectnb/bs859/data/

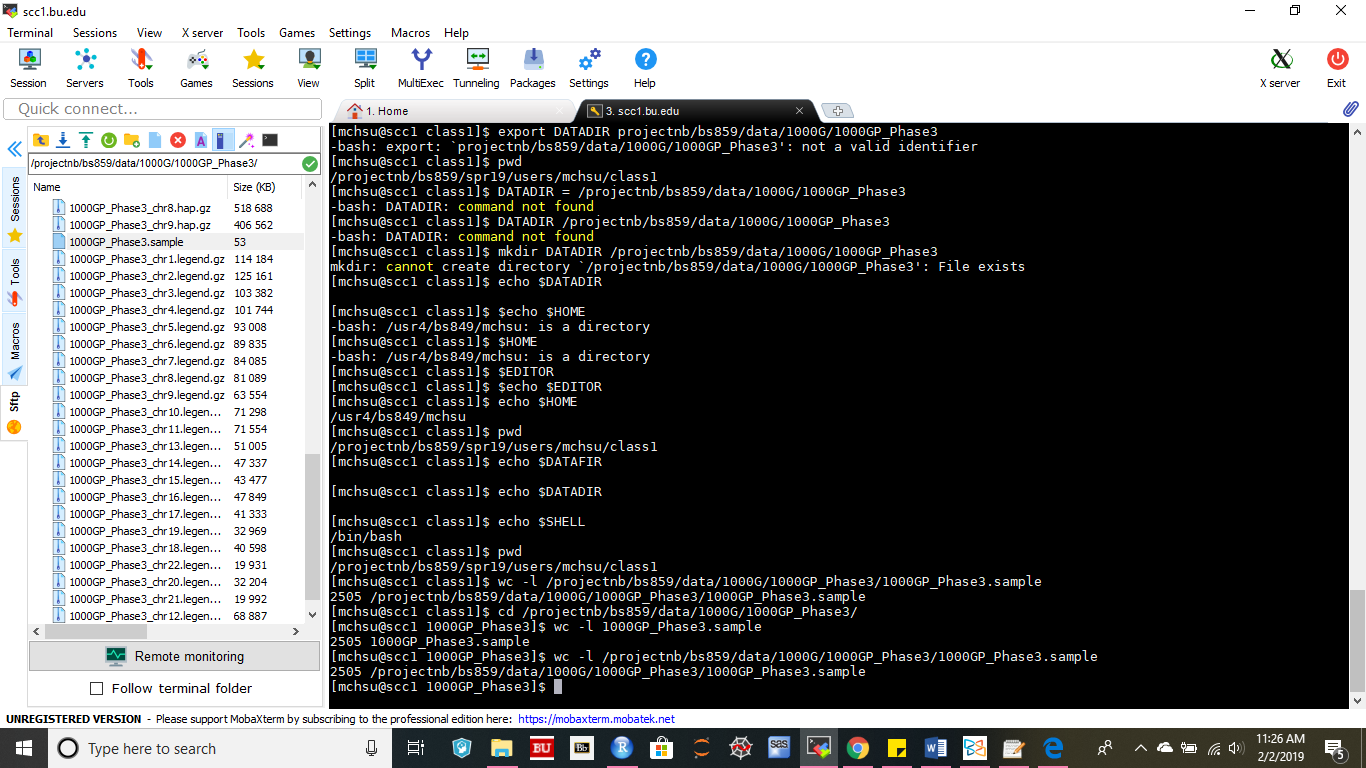


c) Your Directory in the BS859 Project Directory: /projectnb/bs859/spr19/users/username

 (currently, it has no file.)

5.

c) Write a command that will tell you how many lines are in the file called 1000GP\_Phase3.sample in the 1000GP\_Phase3 directory. Run the command. How many lines are there? This is the number of samples (individuals) in the 1000G Phase 3 data



There are 2505 lines in this dataset.