

Session 3 Exercises

Exercise Set #5

Navigate yourself to the `Bootstrappers_Unix_Workshop` folder which you downloaded for the course. Change directories to the `Session_3` folder.

View a portion of the file `Biology100_Grades.csv`

What was the highest average in the class?

What was the lowest average in the class?

How many students passed the class?

How many provisionally passed?

How many failed?

How many students had an average over 90?

How many students had an average under 75?

Create a new file for each Status but with only the Student ID and Average separated by a tab. Name these files `Biology100_Pass.tab`, `Biology100_PPass.tab`, and `Biology100_Fail.tab`.

Exercise Set #6

You have the sequences for several CHIP-seq peaks. View a portion of the file `Cohesion_Peaks.fasta`.

How many sequences are there?

How many nucleotides are in each sequence?

How many ZNF143 peaks are there? How many CTCF peaks are there?

How many RAD21 peaks are there?

Create a new file with just the CTCF peak headers and sequences titled
`CTCF_Peaks.fasta`

How many of the original peaks contain the 5-mer "TTCCC"?

Create new file with just the "TTCCC" containing peaks and their headers
titled `TTCCC_Peaks.fasta`

View a portion of the file `CalculusI_Grades.csv`

What was the mean average for Calculus I?

What was the mean average for Biology 100?

How many students took both Biology 100 and Calculus I?

How many students passed both classes?

How many students passed at least one class?

How many students failed both classes?

How many students failed at least one of the classes but passed the
other?

Replace all of the commas with tabs for `CalculusI_Grades.csv` and `Biology100_Grades.csv`. Change the extension of each of these files to `.tab`