Module 5.2: Learn - Coding

5.2 Coding

Combine multiple genes that worked the best to make a predicted sex for each sample

Using the data tables for all the genes we have studied in the CCLE, assess which ones you think work best for predicting the sex chromosome complement.

To do this, ask yourself:

1. Which genes showed the best separation between cell lines from individuals that reported their sex to be male versus female?
2. How much evidence would I require to conclude that the cell line has a Y chromosome?
    - High expression of all the chrY genes we studied?
    - A majority of the genes?
    - Even one of the genes?
3. Do I need to see high expression of XIST in addition to chrY gene expression to conclude that the cell line has a chrY?

Add predicted sex to a sample information table

Once you determine what you think is a good method, implement that method across all the cell lines in the CCLE and output your results as a comma-separated values (csv) file using any and all of the coding tools you have learned. We are leaving this assignment very open ended because we want you to try your hand at

coding without any template code to work off of because this is what real research is all about– applying knowled
something new to solve a problem.

If you have never written code on your own, here is a good way to go about it:

1. Open a new Rmd and save it in a directory you can easily find again
2. At the top of the Rmd, start by writing down the goal of your code in plain language to make sure you kno
   at the end
3. Identify and describe the input files you will be reading data from
4. Decide and describe what output files you want and what you want them to contain (such as csv files cont
   predicted genetic sex, and whatever else you think is important to know)
5. Write out the steps you will take in plain language to go from that input to that output
6. Once you have your steps, write chunks of code for each step with a description of that step above each c

Remember that we will be grading on effort and progress not on perfect completion. Do your best and ask your c
instructors for help if you reach a problem you can't solve along the way.