Notes on Catherine Batchelder's R Project

by Zoi Goebel

Github repository:

- No folders to contain the MD file
- No readme file
    - Even if there isn't much to write, having a readme file allows others who analyze your github to know information about your project that you found important to convey.

R Markdown File

- Creating new files for each chromosome (maize)
    - You could have arranged the files before running them in the for loop, which might make your code run faster since it wouldn't be organizing them for each iteration of the loop.
    - Although now that I think about it, it breaks the code down into smaller chunks for the program to work on, which could make it faster. I'm not sure which method works best, but with this large of a data set your method probably works better.
- SNPS per Chromosome
    - Combining the two data sets might be confusing later on for data analysis when trying to select for teosinte or maize during data visualization since the way the data becomes transposed makes it hard to retain the group data.
- Data Visualization
    - There are comments missing in this section, which makes it hard to understand what each line of code is being used for.
    - case_when function is nice, I wish I used this for my code

Summary of notes: From what I can tell the code seems to be well written and concise. It made me realize that my code might be a bit bloated with unnecessary chunks of code, but I used a different route to make my data processing less heavy. Overall well documented other than the lack of a markdown file and the missing comments at the visualization step of the coding.