

# Notes on Hannah Gates R Project

by Zoi Goebel

## Github Repository:

- I know the error that you might have run into. Github works weird when you have files in folders, but if you do git add while the file is contained in a folder, it should create that folder in your git repository. It could be that you tried to push those files to your repository before creating folders to contain them in and that might have been what was causing your error although I'm unsure.
- You don't need transpose.awk but I think it is a residual file.

## R Markdown File

- Sort SNP data by SNP\_ID
  - I think that the data already comes sorted by SNP\_ID so this step is unnecessary but I understand why you did it.
- Clean and merge the SNP position data with the genotype data
  - You arrange by SNP\_ID again, which is unnecessary here.
- Sort the transposed maize data by SNP\_ID
  - Make sense to sort here since it would allow for the data to line up
  - I also like how you checked the number of rows in both files
- Parsing through chromosome
  - I like how you made functions for this part of your code, because it makes it easier on the eyes
- Visualization
  - I like how you combined both species to make comparison easier to visualize, I wish I would have done this with my data, but I'm not the best at knowing how to visualize my data.

Summary: Overall your program completes the task it was designed to do, there might be some repetitive code in there but that isn't a big deal. You did a good job understanding stipulations regarding merging files together and your visualization code looks well done. I liked the incorporation of functions for the bigger tasks of your code, because it made it easier to understand and read.