

Assignment 1: Reproducibility, Workflow, Version Control

Tristen Townsend

1) Discussion Questions

Question

Why are reproducible practices becoming the norm in data analytics?

Answer: Reproducible practices are useful when one wants to be able to re-perform their analyses or allow for others to replicate analyses. This is especially helpful when analyses are complex or have many steps as it ensures data is analyzed and produces the same results whether it is done at different times or by different people.

Question

What are your previous experiences with data analytics, R, and Git? Include both formal and informal training.

Answer: I have experience performing data analysis using Excel and a little experience doing model verification which required cleaning up data and creating useful visuals. I took Applied Statistics for the Environment last semester where I was introduced to R. I have no experience with Git.

Question

Are there any components of the course about which you feel confident?

Answer: I feel a bit confident doing coding, GLMs, and some data visualization as I have done so in my previous R class. I also feel that I may be okay with troubleshooting because I've had to do that with programs frequently in the past.

Question

Are there any components of the course about which you feel apprehensive?

Answer: I'm a bit apprehensive about data wrangling because it seems like it could be a bit complicated and deciding which data to transform and how to do so can sometimes be tricky.

2) GitHub

Your Repository

Provide a link below to your course repository in GitHub. Make sure you have pulled all recent changes from the course repository (https://github.com/KateriSalk/Environmental_Data_Analytics) and that you have updated your course README file.

Answer: https://github.com/tristen0708/Environmental_Data_Analytics