

Assignment 1: Reproducibility, Workflow, Version Control

Lindsay Roth

OVERVIEW

This exercise accompanies the lessons in Environmental Data Analytics (ENV872L) on reproducibility, workflow, and version control.

Directions

1. Change “Student Name” on line 3 (above) with your name.
2. Use the lesson as a guide. It contains code that can be modified to complete the assignment.
3. Work through the steps, **creating code and output** that fulfill each instruction.
4. Be sure to **answer the questions** in this assignment document. Space for your answers is provided in this document and is indicated by the “>” character. If you need a second paragraph be sure to start the first line with “>”. You should notice that the answer is highlighted in green by RStudio.
5. When you have completed the assignment, **Knit** the text and code into a single PDF file. You will need to have the correct software installed to do this (see Software Installation Guide) Press the **Knit** button in the RStudio scripting panel. This will save the PDF output in your Assignments folder.
6. After Knitting, please submit the completed exercise (PDF file) to the dropbox in Sakai. Please add your last name into the file name (e.g., “Salk_A01_Reproducibility.pdf”) prior to submission.

The completed exercise is due on Thursday, 17 January, 2018 before class begins.

1) Discussion Questions

Question

Why are reproducible practices becoming the norm in data analytics?

Answer: Reproducible practices are becoming the norm in data analytics because it is important to be able to go back to your own work or someone elses work and be able to understand what was done so that you can use those methods in your future work.

Question

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

Answer: I have taken the Statistics class at the Nicholas School with John Poulsen using R. Before that I had no other experience with R, data analytics, coding, or Git.

Question

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

Answer: I feel confident in my abilities to understand the grammar of R code and be able to use the various functions in R to help me solve problems.

Question

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

Answer: I am a little apprehensive about connecting R Studio with the cloud and using GitHub to manage my assignments. Additionally, I would like to work on my trouble shooting skills to be able to fix any issues I am having independently.

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/lhr12/Environmental_Data_Analytics