Assignment 1: Reproducibility, Workflow, Version Control

Wanchen(Connie) Xiong

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 practice is becoming the norm because it is important to establish such practice so that someone else can use the practice to obtain the same result. It enhances efficiency in processing data.

Question

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

> Answer: I used R and Git in a previous internship to do data management for the department that I worked for. I mainly used R markdown to standardize word documents. And I used Git to keep my progress up-to-date with my coworkers. Therefore, I used R in a formal setting before.

Question

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

> Answer: I am confident about navigating myself through the program itself in terms of using its basic functions, especially those related to R markdown. And I feel confident in communicating my concerns and struggles with my instructors in this class; I will make sure I ask questions when I am stuck and receive feedback and help.

Question

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

Answer: I am not confident in finding solutions to address with errors that come up. And I want to learn more about the fundamentals about R coding. I am also not sure about how classmates are going to help each other in this kind of setting since I feel like we are working individually most of the time. It would be helpful if we have a problem session, even just for an hour or two each week for students to bring their questions to discuss with instructors or among each other.

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/conxiong/Environmental_Data_Analytics.git I have updated my course README file.