

Assignment 1: Introduction

Elizabeth Good

OVERVIEW

This exercise accompanies the introductory material in Environmental Data Analytics.

Directions

1. Rename this file `<FirstLast>_A01_Introduction.Rmd` (replacing `<FirstLast>` with your first and last name).
2. Change “Student Name” on line 3 (above) with your name.
3. Work through the steps, **creating code and output** that fulfill each instruction.
4. Be sure to **answer the questions** in this assignment document.
5. When you have completed the assignment, **Knit** the text and code into a single PDF file.
6. After Knitting, submit the completed exercise (PDF file) to the appropriate assignment section on Sakai.

1) Finish setting up R Studio

Install TinyTex

Now, run this code cell the same way. This will install “tinytex” – a helper app that allows you to knit your markdown documents into professional quality PDFs.

Set your default knit directory

This setting will help deal with relative paths later on... - From the Tool menu, select **Global Options** - Select the RMarkdown section - In the “Evaluate chunks in directory”, set the option to “Project” (If you don’t see this option, try restarting RStudio.)

2) Discussion Questions

Enter answers to the questions just below the `>Answer:` prompt.

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

Answer: My first experience with R was my senior year of college when I took an introductory statistics class to fulfill the prerequisite for Duke. It was very basic background work with simple functions for statistical tools. I took Duke’s Applied Statistical Modeling for Environmental Management class last year which involved a significant R component including more complex functions for data manipulation, visualization, and modeling. Over the summer I worked for

the US EPA doing a data analysis project comparing air quality sensor data to air monitoring data, relying totally on R to do the analyses. I gained a lot of experience with more complex functions for data manipulation and visualization, including writing my own functions and loops. There was a strong emphasis on the reproducibility of the code. This is also where I got my first experience with Git and working collaboratively on code. Informally, I started doing data analysis in college as a chemistry major, working mostly in Excel to make plots and do error calculations.

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

Answer: I'm relatively confident in my basic R skills at the moment, but I'm very interested in learning new and more efficient ways to get to an end goal.

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

Answer: While I used Git for my work over the summer, I feel like I just knew enough to get by; I never really understood how it was working. I'm slightly apprehensive about using Git for this class. I'm also not very confident with large data sets in complex formats or with variables of various classes.

3) GitHub

Provide a link below to your forked course repository in GitHub. Make sure you have pulled all recent changes from the course repository and that you have updated your course README file, committed those changes, and pushed them to your GitHub account.

Answer: https://github.com/goodelizabeth/EDE_Fall2023.git

4) Knitting

When you have completed this document, click the **knit** button. This should produce a PDF copy of your markdown document. Submit this PDF to Sakai.