

Assignment 1: Introduction

Yilin Zhong

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”

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: I previously had taken two class that works with R. The first is ENV710 Applied Data Analysis for Environmental Social Science, and the second course is Biostatistics and Epidemiology for Global Health I (GLHLTH 705). In both class, we work with comparison of means between two groups using t-test, and various Chi-Square Tests. We also use Z-test for comparison of proportions between two groups. In ENV710, we work with Linear Models and OLS Assumptions. We learned to use R to make histograms, bar graph, and box-plots. However, I don't have any experience with Git.

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

Answer: The components of the course that I feel confident doing is making scatter plots, box-plots, histograms, and bar graphs. I also feel confident on using ggplots. Furthermore, I have some experience on GLMs with my pervious course in R. Moreover, I am condifent on making tables and using R to determine summary statistics, such as mean, median, Standarad Deviation, and Range.

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

Answer: For the rest of the course materials, I haven't learn from my previous courses, which I feel apprehensive on. For instance, time series analysis, stationarity, seasonality tests, spatial data analysis, data scraping, web scraped data, API, DOM parsing. I am looking forward to learn these new techniques.

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/Yilin157/EDA-Spring2023.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.