Assignment 1: Introduction

Nikki Egna

OVERVIEW

This exercise accompanies the introductory material in Environmental Data Analytics.

Directions

1. Change "Student Name" on line 3 (above) with your name.
2. Work through the steps, **creating code and output** that fulfill each instruction.
3. Be sure to **answer the questions** in this assignment document.
4. When you have completed the assignment, **Knit** the text and code into a single PDF file.
5. After Knitting, submit the completed exercise (PDF file) to the dropbox in Sakai. Add your last name into the file name (e.g., "Salk_A03_Introduction.Rmd") prior to submission.

The completed exercise is due on Tuesday, January 14th at 1:00 pm.

1) Discussion Questions

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

   Answer: I began learning R in my sophomore year of my undergraduate experience at the University of Southern California. What started as frustration quickly blossomed into fascination. Through conducting real-world experiments and analyzing the results using R, I recognized the power and potential of technology as it relates to ecological research. In a subsequent on-campus research-based internship with the Wrigley Institute as part of the San Pedro Ocean Timeseries (SPOT) project, I put my knowledge of R to work. I organized the oceanographic data collected from a CTD device, and calculated parameters such as partial pressure of CO2. I also created graphs in R displaying the trends of partial pressure values as they changed with depth, dissolved oxygen levels, and salinity levels. I then ran statistical analyses on the significance of those trends, which the Wrigley Institute ultimately used as preliminary data for its upcoming publication on the SPOT initiative. After graduating USC, I accepted a job as a Research Associate at San Diego Zoo Global. I specifically worked with the Reticulated Giraffe Conservation Program and was primarily responsible for the project's data analysis. I wrote a script that organizes and analyzes the results of our citizen science platform wildwatchkenya.org, where volunteers assist in classifying the images from our gridded network of trail cameras in northern Kenya. I implemented a plurality algorithm that converts the multiple citizen scientist classifications into one single aggregated answer, so each photo has a consensus on the species, number of individuals, and behaviors exhibited. I am currently writing a paper assessing the accuracy of these citizen scientists as compared to experts. I also assisted the PI with his Ph.D. work on the analysis of several sets of aerial survey data of giraffe. I ran statistical analyses on his results, created maps showing the flight paths of the aerial surveys, and created graphical visualizations of the difference between the counts from the aerial surveys versus the published counts from the IUCN. Lastly, I currently hold a research assistantship position at Duke with Dr. John Poulsen's lab, which studies tropical ecology of tree communities and wildlife in Gabon, Africa. My role is to use R to clean up the collected tree growth data from Gabon and model carbon storage as part of their National Resource Inventory initiative.

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

   Answer: I feel confident in that I have experience in R and data analysis in general. My main goal is to improve on my skills that I already have and learn new things along the way as well. I am very excited to be learning in a classroom setting, as most of my skills in R have been self-taught up until this point.

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

   Answer: I have not used Github before, as it just seemed like a hurtle and I never put in the effort to tackle it. I am really looking forward to becoming familiar with this tool, and to improve on my organization skills as a result. I also can use improvement on my visualization skills. I have rudimentary knowledge of graphing and plotting, but I hope to be able to take my skills to the next level and be able to create some really nice looking products.

2) 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.

Answer: https://github.com/nikki-egna/Environmental_Data_Analytics_2020