BIO 410/510 Final Project Workflow

Please submit both the "cleaned-up, report-worthy"" html AND the "all work shown, echo = TRUE" files to Canvas by 9 am Wed 12/5

Points: 40

Final project objective

The objective of the final project will be to complete a fully reproducible workflow that uses data to address your chosen question. The project must illustrate all of the following tasks:

- Some form of data access / reading into R
- Integration of multiple datasets to address the question
- Data tidying preparation using tidyr, including data joins
- Use of dplyr to manipulate and summarize the data in relevant ways
- Initial data visualization with ggplot2
- Final, publication-worthy visualization with ggplot2
- RMarkdown writeup, with final submission as both the .Rmd file and a nicely formatted as a PDF or html document that includes code and results
- Overall clean and clear presentation of the workflow, code, and explanation

I. Introduction (Literature review)

The literature review will be around 5 well cited paragraphs that do the following:

1. Introduce the problem and explain why

- Set the stage for the problem

- Put the concept and question into context

- Lots of big-picture citations (such as reviews) in the first paragraph

2. Past work and data available on the project

- Who has addressed this problem, and what did they do it?

- What are the data available to address this problem?

- How has the data available and/or methods changed recently?

3. Purpose of the study

- Further refine your approach (e.g., what data will you combine, how will you address the question)

- Justify why this is needed now (e.g., visualization to test a new dimension of the question or better convey an old one)

4. Hypotheses/questions

- List these clearly and in a logical order

- Make hypotheses directionally using predictions (e.g. "I predict N will reduce plant diversity" rather than "I predict N will change plant diversity")

II. Approach/Methods

- Provide a description of the data sources, including who collected the data and how it was collected

- Describe how you will use the data to address your questions

- Include R chunks with data import, data tidying, and any important preliminary visualizations (for the "cleaned up" version these will be in non-exported R code chunks; they should be shown in the "all work shown" version)

III. Results

- Report the key findings (but save the interpretation and contextualization for the discussion)

- Provide evidence to answer all questions or hypotheses posed in the introduction

- Include R chunks that contain data summaries and visualizations

- Have these chunks export captioned, publication-worthy figures

IV. Discussion

The discussion will be around 3-5 well-cited paragraphs that do the following:

1. Provide a big overview or summary of consequences of strongest results (1 paragraph)

2. Expand thoughts on results and hypothesis (2-4 paragraphs)

3. Discuss the limitations of the approach a/o potential future directions (1 paragraph)

4. Summarize the conclusion or take home message (1-2 paragraphs)

Goals for this section include:
- Address hypotheses (are they resolved?), questions, aims, and/or limitations
- Discuss the consequences or implications of results
- Make comparisons with previous findings (support or contradict)
- Announce your study's contribution to the current field
- Draw broader conclusions
- The comprehensive literature review should be in the introduction, here it is necessary to link to past work but no need to introduce too many new citations

V. Literature cited

Please follow MLA citation guidelines. This will include in-text, parenthetical citations and a literature cited section at the end. Aim for at least 10 peer-reviewed cited works.

This is a nice reference on scientific writing: Turbek, Sheela P., Taylor M. Chock, Kyle Donahue, Caroline A. Havrilla, Angela M. Oliverio, Stephanie K. Polutchko, Lauren G. Shoemaker, and Lara Vimercati. "Scientific Writing Made Easy: A Step-by-Step Guide to Undergraduate Writing in the Biological Sciences." The Bulletin of the Ecological Society of America 97, no. 4 (October 2016): 417–26. https://doi.org/10.1002/bes2.1258.