

**ECOL 597 W**  
**Practical and Reproducible Data Science for EEB**  
**Final Project Instructions**

Your final product for this course will be an independent research analysis of a project of your choosing. You can use data from a rotation, your past or present research, or a publicly available dataset. The goal is to demonstrate the key things we've learned in this class: how to wrangle, display, and analyze data. Here is the relevant timeline:

**1. Choose a project**

- a. Start brainstorming, and think about what data you'd like to analyze and how you'll go about it.
- b. The **week of November 11** email me a brief description of your idea.

**2. Do the work**

- a. Your final project should have the following elements:
  - i. A scientific question/goal
  - ii. Between 1 and 5 display elements (i.e. beautiful figures)
  - iii. At least one appropriately used statistical test
  - iv. Reasonable interpretation of your data.
  - v. A well-organized Github repo for your final project which includes, at a minimum, a Readme, your data, and script(s).

**3. Peer review (5% of final grade)**

- a. The week of December 9 you will be required to review someone else's final project code. Your project does not have to be done by this date, but it should have all the required elements in 2.a
  - i. **Reviewer:** Fork and clone your partner's repo. Review their code, making sure it works. If there's anything you don't understand, or think they should change, add some notes to their script and/or a separate document
    1. **IMPORTANT REMINDER: MAKE YOUR CHANGES IN A BRANCH.**
  - ii. **Submit a pull request** to your partner when you're done with review.
    1. (Optional tip: to avoid Github headaches, give them a heads up you're about to look at their code so they can hold off on making changes while you review).

**4. Submit your final project (10% of final grade)**

- a. You will submit to the class repo **a single markdown or quarto document** detailing everything in part (a).
  - i. Name your document "lastname\_final.qmd." Include a link to your github repo in the document.
  - ii. **CREATE A BRANCH**
  - iii. Put your doc in the folder ECOL\_596W\_2024/final\_projects/
  - iv. Submit a pull request.

## Evaluation:

Here are some thing's I'll be looking at:

- i. A scientific question/goal**
  - a. Do you provide context for the project? Where were the data collected and how? Do you provide enough info for someone to understand the data presented?
- ii. Between 1 and 5 display elements (i.e. beautiful figures)**
  - a. Are the data presented in a logical and clear way?
  - b. Are there any clear errors? (e.g., a broken axis on a bar plot, unlabeled axes)
  - c. Are the figures well-formatted (i.e., clear and large axis labels, color blind friendly, aesthetically pleasing, and interpretable color scales)?
  - d. Are there figure legends?
- iii. At least one appropriately used statistical test**
  - a. Did you make good analysis choices? Are the tests appropriate for the data?
  - b. Are the results nicely displayed?
- iv. Reasonable interpretation of your data.**
  - a. Can the reader learn some science from your project?
- v. A well-organized Github repo for your final project which includes, at a minimum, a Readme, your data, and script(s).**
  - a. Are your data reproducible and interpretable to an outsider?