

# BIO 410/510 Final Project: Topic Identification

**Due date update: Please submit to Canvas before class on Monday Oct 8**

## **Final project objective**

To help you look ahead, the objective of the final project will be to complete a fully reproducible workflow that uses data to address your chosen environmental 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 PDF document that includes code and results
- Overall clean and clear presentation of the workflow, code, and explanation

## **Final project topic selection**

We will be building toward the final project throughout the class. At this stage I need enough information from you to tailor the data sources we cover and the examples we use in class. A full project proposal, with questions and identified data sources, will be due later in the fall. For now, please submit a .Rmd file saved as “LASTNAME\_ProjectTopic.Rmd”. In addition to being uploaded as FinalProjectTopicID.pdf, this document is available on Canvas as an .Rmd file as an example of the format. Include the following:

### **If you do not have your own data:**

- 1) General topic of interest (e.g., climate, species conservation, evolution, genetic diversity, hydrology, human demography, etc)
- 2) Potential questions within that topic area (e.g., how do climate forecasts differ regionally? How does environmental change influence species composition?, how does phylogenetic diversity vary by system?, etc)
- 3) Potential datasets within that topic area (e.g., specify relevant government databases, academic data repositories, etc)

### **If you do have your own data:**

- 1) The questions you would like to answer with your data
- 2) The structure of the dataset (give as much information about the data as the str() function would return)