

Draft Title: A Code-Based Data Workflow Using R

This webinar provides a brief introduction into using R to develop a code-based workflow for ecological survey data. Participants will learn about the advantages of using a scripting language like R for moving data along the data life cycle. We will walk through an example of using R to manage scientific data.

1 Webinar Information

Presenters: McCrea Cobb and Adam Smith

When:

Location:

Additional resources: GitHub repository

2 Outline

1. Quick review of the data life cycle steps
2. Overview
3. Advantages of a code-based data workflow
 - Documented
 - Reproducible
 - Replicable
 - More efficient
 - Less error-prone
4. Describe the typical steps:
 - Develop scripts using packaged functions
 - Wrap scripts in custom functions
 - Develop a package to house, document, and reuse custom functions
5. Run through an example

3 (*Topics Ideas*)

The list below is taken from <https://learn.datacamp.com/courses/working-with-data-in-the-tidyverse>:

1. Explore your data
 - Import
 - Dealing with missing values
 - Summarize data (glimpse and skim)
2. Tame your data

- Cast column types
 - Recode values
 - Select variables
 - Tame variable names
 - Rename variables to convention (janitor package)
3. Tidy your data
 4. Archiving

4 Resources

Resources for Teaching R

- DataCamp: Tidyverse Course
- learnr package
- RStudio teaching resources

R Resources

- R cheatsheets

Other

- How to share your data with a statistician