

# 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 (*Questions*)

- What is the take-home message from the webinar? You can (and should) use R entirely for your data workflow?
  - Demonstate how R can be used for the data life cycle process
  - A good clear demo of this.
  - Show the endpoint first. The pie in the sky.
    - \* Example: SppDistMonProj, mobile aucistic recorder package
    - **look for other example**

- **meet on Wed to discuss**
- How do we want to collaborate on this? Through GitHub? If so, how? Separate branches or forked?
  - **GitHub is a good option**
  - **McCrea will work off master branch, Adam will fork**
  - **Keep work distinct in the repo**
- Do we want to use the SppDistMonProj package as the example? What other examples do we have?
- Should we attempt to include any class participation? If so, what format?
  - **Probably not enough time.**
- Should we focus on the tidyverse?
  - **Doesn't really apply**
- Version control?
  - **repo is available for the webinar, we can then mention GitHub and version control**
- Rmarkdown for generating our presentation slides? Flipbook is pretty cool for incrementally describing code.
  - **rmarkdown**
  - **Send template to Adam**
  - **McCrea will send project file dir template to Adam**

## 4 (*Topics Ideas*)

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

1. Explore your data
  - Import
    - **Add Adam to the SppDistMonProj AGOL**
    - **Adam will look into accessing Survey123 data using R**
  - 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
  - **Brent Frakes: writing products to ServCat**