

## Workshop Overview

CEREO typically offers Data or Software Carpentry workshops once per semester, covering basic computational skills for data management, analysis, or research computing. The goal of this offering was slightly different: We wanted to provide hands-on instruction for higher level R programming techniques targeted at graduate students in the sciences. In contrast to Data and Software Carpentry workshop offerings, this workshop expected basic knowledge of R programming prior to the workshop, took place fully over Zoom, and spanned five mornings instead of two full days. The curriculum was developed by the instructors for a prior version in November 2020, with some topics adapting material published by the Carpentries, and other topics being developed from scratch. We limited the attendance to ~20 participants due to the online format. All attendees were from WSU, and we were able to include participants from multiple campuses.

The workshop took place from March 15th – 19th, 2021. We invited attendees to join us for all five days of instruction, but did not require this. Broad topics by day included: 1) Git/GitHub intro + data manipulation and visualization with tidyverse R packages, 2) R Markdown, 3) basic modeling with R, 4) workflow management with the drake package, and 5) an overview of making a free personal research website with GitHub. While attendance to every lesson wasn't required, we did build upon (and reuse) skills learned throughout the week so that participants could see how techniques could be put into practice. For example, Git and GitHub were taught at the start of the workshop and integrated into each day's lesson plan to a varying degree. The workshop syllabus can be found in Appendix B.

The workshop was taught by a team of five instructors with the additional assistance of four helpers. Helpers generally had previous experience with the techniques or packages being taught and were present to troubleshoot technical problems, monitor the Zoom call, and manage the shared notetaking document (a Google Doc). To manage instructor and helper communications throughout the week of the workshop we created a Slack workspace used by the full team. Like similar CEREO workshops, the lessons were generally composed of tutorials that included hands-on exercises with live coding demonstrations. Challenge questions were interspersed throughout most tutorials, allowing participants to practice throughout the workshop

*Table 1. Summary of Workshop Demographics based on survey responses*

| <b>Number of:</b>             |    |
|-------------------------------|----|
| Registrants                   | 21 |
| Male                          | 7  |
| Female                        | 6  |
| Prefer not to say             | 3  |
| Graduate student              | 17 |
| Faculty                       | 1  |
| Postdoc                       | 2  |
| Undergraduate student         | 0  |
| Other positions (e.g., staff) | 1  |
| WSU Depts represented         | 9  |
| WSU Colleges represented      | 3  |
| WSU campuses represented      | 3  |

*Subsection totals may not sum to 21 if not all participants responded to demographic questions. Data were gathered from multiple sources with different response rates.*

and gain confidence in their ability to apply the techniques being taught. We also provided many of the scripts in a format that would be accessible online to the participants after the workshop, allowing them to revisit the same material afterwards.

We sold out all 20 spots in the workshop. For those attendees that provided gender information, male and female attendees were nearly equally represented. Most attendees were graduate students, but staff, post-docs, and faculty were also represented (Table 1). Members of 9 separate WSU departments (1 non-WSU), three colleges (1 non-WSU), and three campuses (1 non-WSU), were represented. For participants by college, see Table A1. Approximate attendance counts in order of workshop day were: 18, 12, 14, 10, and 11.

We also provided a one-hour pre-workshop Zoom call the first day of the workshop to assist participants with installing the software if they needed help. In addition to this we sent out a PDF describing software needed in advance, some installation tips, and ways to check if difficult software had been installed successfully. None of the participants attended the pre-workshop installation call and very few installation problems arose, although some participants did not finish installations before the workshop.

## **Summary of Participant Feedback**

### *Survey Results*

Links to confidential surveys were sent to the workshop participants immediately before and after the workshop. The surveys were created by the instruction team to replace surveys typically used for Carpentries workshops, but which were not available for this non-Carpentry offering. The goal was to assess participant satisfaction and solicit feedback for improving future iterations of this workshop. Pre-workshop survey response rates were higher than post-workshop (76% vs. 38%). An overview of survey findings is given in Table 2.

Most pre-workshop respondents (88%) had previous experience in a programming language, and all of those had previous experience with R. However, only 19% of the respondents had previous experience with a statistical analysis program that used a graphic user interface (GUI), such as JMP or SPSS. Respondents expressed a desire to build efficiency, comfort, and understanding of R pre-workshop; learning git was also a common motivation for attendance. Post-workshop survey results indicated that participants appreciated a variety of the topics they experienced, though data manipulation was mentioned most frequently. On average, attendees found the workshop very useful (avg. rating 9.25/10) and would be very likely to recommend it to someone else (9.75/10). These scores were highest for the attendees with previous R experience (usefulness: 9.71/10; recommend: 10/10), which was an expectation for attendance. Additionally, respondents appreciated the helpers' and instructors' knowledge and dedication, along with the synthesis of topics. Respondents did not coalesce around common suggestions for workshop improvement, but examples included more git/GitHub background and in-person instruction. This is a change from last time (when respondents mentioned being short on time, moving quickly, etc.), which may be a result of our choice to lengthen workshop days to allow for slower pacing.

### *In-workshop Feedback*

Participants were also asked to provide simple short answer feedback about the workshop at the end of each day. Response rates varied day-to-day, and sometimes highlighted different

strengths of the workshop than those in the post-workshop survey. Suggestions for improvement here commonly mentioned pacing and providing more context for the lesson topics. Participants frequently mentioned GitHub, R Markdown, and the modeling topics when discussing useful things they had learned.

*Table 2. Overview of survey responses*

|                                                                                                                                                                                                                   |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Overview                                                                                                                                                                                                          |
| Response rates (76%) were higher than post-survey response rates (38%)                                                                                                                                            |
| Pre-survey responses (n = 16)                                                                                                                                                                                     |
| 88% of pre-workshop respondents came in with previous experience programming and 88% had prior experience with R.                                                                                                 |
| 19% of pre-workshop respondents had prior experience with a GUI-based statistical analysis program.                                                                                                               |
| Most pre-workshop respondents considered themselves beginners (56%) or intermediate-level users (25%) with R.                                                                                                     |
| Pre-workshop respondents mentioned improving efficiency (25%), increasing their comfort/understanding with R (31%), and learning about reproducibility (19%) or Git/GitHub (19%) as common reasons for attending. |
| Post-survey responses (n = 8)                                                                                                                                                                                     |
| Respondents rated the workshop as very useful (9.25 / 10 for usefulness on average) and would be very likely to recommend the workshop to someone else (9.75 / 10 on average).                                    |
| Respondents mentioned the knowledge and dedication of the workshop instructors and helpers as one of its strengths.                                                                                               |
| Modeling was one of the topics mentioned most often as being useful, followed by sections of the lesson on data manipulation and workflow management.                                                             |
| Respondents did not coalesce around common suggestions for workshop improvement, but examples included more git/GitHub background and in-person instruction.                                                      |

## Appendices

### Appendix A. Table

Table A1. Participants by College

| College                                                       | Count     |
|---------------------------------------------------------------|-----------|
| College of Agricultural, Human, and Natural Resource Sciences | 8         |
| Murrow College of Communication                               | 0         |
| College of Arts and Sciences                                  | 5         |
| College of Veterinary Medicine                                | 0         |
| Not listed or non-WSU                                         | 1         |
| College of Education                                          | 0         |
| Voiland College of Engineering and Architecture               | 7         |
| <b>Grand Total</b>                                            | <b>21</b> |

## Appendix B. Workshop syllabus

### Day 1: Brief intro to Git/GitHub. Intro to Data Manipulation and Visualization with the tidyverse

- Git/GitHub
  - Set up Git & create a repository
  - Set up GitHub & connect your local Git repository
- tidyverse
  - Select(), filter(), piping, mutate()
  - Pivot\_wider(), pivot\_longer()
  - Basic ggplot syntax
- Pushing to GitHub

### Day 2: Intro to R Markdown

- Introduction to the markdown language
  - Formatting: Headers, super/subscripts, bold/italics
  - Inserting images, figures, equations
  - Practice reading, analyzing, and plotting data
  - Methods for making tables
  - Exporting the results as HTML

### Day 3: Model Building with R

- Building models
  - Building a model and assessing fit (predictions vs data)
  - Assessing model residuals
  - Incorporating interactions with continuous and categorical variables
  - Including polynomial terms
  - GAMs
- Purrr + broom for model building
  - Nested dataframes/list-columns
  - Run models iteratively using map functions
  - Combining map with glance to extract model parameters
  - Visualize model parameters

### Day 4: Workflow management in R with drake

- Introduction to workflow management
- Introduction to functions
- Introduction to drake
  - File structure
  - Splitting a workflow into functions
  - Writing core drake scripts
  - Organizing a full analysis plan
  - Editing or recovering a workflow

### Day 5: Intro to making a website with Git/GitHub

- GitHub Pages background
- Finding and cloning templates
- Editing template documents locally and online
- Customizing website layout and formatting
- Uploading images and HTML files