



# Journal of eScience Librarianship

putting the pieces together: theory and practice

## R Workshops for Researchers:

# A Successful Partnership Between a Library and a Statistical Consulting Laboratory

**Jocelyn Boice**, Colorado State University, Fort Collins, CO, US, [jocelyn.boice@colostate.edu](mailto:jocelyn.boice@colostate.edu)

**Mara Sedlins**, Colorado State University, Fort Collins, CO, US, [mara.sedlins@colostate.edu](mailto:mara.sedlins@colostate.edu) 

**Julia L. Sharp**, Colorado State University, Fort Collins, CO, US, [julia.sharp@colostate.edu](mailto:julia.sharp@colostate.edu) 

---

### Abstract

Colorado State University Libraries has partnered with the university's Franklin A. Graybill Statistics and Data Science Laboratory to jointly offer workshops teaching R coding for academic research, while employing Statistics graduate students as instructors. This partnership has capitalized on the strengths of both units and has provided an opportunity to build campus relationships. The development, assessment, challenges, benefits, and evolution of this fruitful and ongoing collaboration are described with the intent that this information may assist in the creation or enhancement of services at other institutions.

---

**Received:** December 1, 2022 **Accepted:** June 23, 2023 **Published:** July 18, 2023

**Competing Interests:** The authors declare that they have no competing interests.

The *Journal of eScience Librarianship* is a peer-reviewed open access journal. © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0), which permits unrestricted use, distribution, and reproduction in any medium non-commercially, provided the original author and source are credited.

See <https://creativecommons.org/licenses/by-nc/4.0>.

♾️ OPEN ACCESS

## Introduction

Academic library support for research data management is varied, ranging from consultations on data-related questions, to workshops teaching relevant content and skills, to services for data visualization (Radecki and Springer 2020, 16-17). Often this support entails cooperation with other university units or outside institutions. For example, several universities have established internal collaborations for data support services involving their libraries and information technology departments, and in some cases, additional units (AAU and APLU 2021, 25; CRDDS, n.d.; NC State, n.d.; Wittenberg and Elings 2017). At Iowa State University, the library was among several university divisions that worked together to create a data repository and accompanying guidelines (AAU and APLU 2021, 15), while at Arizona State University, the library provided instruction on data management topics for research office employees and worked with the office to add pertinent subject matter to an online researcher training system (Harp and Ogborn 2019, 4). The New England Software Carpentry Library Consortium is an example of cooperation between unrelated organizations, with multiple academic institutions creating a consortium to share membership in The Carpentries, thereby providing access to research data management training for staff and distributing the monetary burden (Atwood et al. 2019, 3). Indeed, intra-institutional cooperation has been recognized as “a necessary condition for building effective research support services” generally (Bryant, Dortmund, and Lavoie 2020, 2) while collaborations of all kinds are seen as being “critical to the success of . . . RDS [research data services]” specifically (Ali et al. 2022, 16).

At Colorado State University (CSU), a valuable collaboration has evolved between the CSU Libraries (Libraries) and the CSU Franklin A. Graybill Statistics and Data Science Laboratory (Stat Lab). The two units cooperate closely to leverage existing graduate student expertise to teach R programming workshops for the approximately 33,000 students (Institutional Research 2022, 2) and 7,500 faculty and staff (53) on the Fort Collins campus, as well as multiple community partners (e.g., United States Department of Agriculture). In this article, the program and partnership at CSU is described.

## Program and Partnership Development

In 2016, the Libraries initiated a workshop series focused on data management “to fill a perceived gap in training in these topics in research education” at CSU (Magle 2018, 1). Some of the workshops were original presentations, while others were based on existing instructional materials, primarily Data Carpentry lessons such as “Data Analysis and Visualization in R for Ecologists” (Michonneau and Fournier 2022). The workshop series was branded Data and Donuts and marketed to the CSU community. The Libraries’ Data Management Specialist taught the sessions, with library staff providing classroom assistance for individual questions and technology troubleshooting.

One of the publicity efforts for the initial series of workshops caught the eye of the Director of the Stat Lab, a unit within the College of Natural Sciences that offers “general statistical consulting to researchers from every college at Colorado State University” via drop-in hours and scheduled meetings (Franklin A.

Graybill, n.d.). The Director recognized that there was a natural synergy between the library's endeavor and the existing pool of R expertise in the Statistics Department, a field which uses R extensively. The Director reached out to the Data Management Specialist to discuss opportunities for partnership, and as a result, in the spring of 2017, the two arranged to enlist graduate student volunteers from the Statistics Department, who are typically trained in R programming throughout their education with more focused training as they progress, to provide in-classroom support for the workshops.

Beginning in the fall 2017 semester, the workshops were expanded and reorganized into two tracks so that attendees could easily identify the workshops that required coding and those that did not. The Data and Donuts track incorporated sessions that focused primarily on non-coding topics, such as using spreadsheets, creating data management plans, and the concept of reproducible research. In contrast, Coding and Cookies (C&C) workshops focused on teaching coding skills relevant to data management for academic research, such as using R for data cleaning, analysis, and visualization.

After several semesters of successful workshops, discussions began about deepening the partnership between the Libraries and the Stat Lab. As part of this change, Statistics graduate students would teach the C&C workshops, as well as serve as classroom assistants. This modification was implemented in spring 2019.

Currently, the Libraries and the Stat Lab jointly lead the C&C program and share responsibility for the practical details of running the workshop series. The Libraries supplies the meeting space via physical classrooms or video conferencing software, as well as providing the registration system, refreshments, and support from the Libraries' communications staff to create, for example, visual art or flyers. The Stat Lab recruits workshop instructors from the Statistics graduate students and promotes the series through its communication channels. Both the Libraries and the Stat Lab provide assistants who help with technical difficulties and questions which arise during the lesson. Both also contribute to planning and adapting future sessions based on information derived from assessment activities. For example, participant attendance has been used to inform the frequency and types of the sessions. At the beginning of the collaboration, the introductory R Basics session was held only once per semester, but based on attendance and feedback from participants, it is now held twice per semester.

Although the offerings have varied somewhat over the years, each semester the C&C workshop series consists of approximately five 90-minute sessions, most of which are offered once, with the exception of R Basics. The fall 2021 schedule is shown in Table 1, and represents the typical topics covered, although a workshop on using Git for version control often replaces the RMarkdown workshop in the spring series.

**Table 1:** Fall 2021 C&C workshops

<b>Workshop</b>	<b>Date</b>
R Basics	September 7
R Basics	September 21
Tidy Data in R	October 12
Data Visualization using ggplot2	October 26
Reproducible Reports using RMarkdown	November 9

## **Instructors**

The instructor recruitment process begins when Statistics graduate students are sent an email one semester before the workshops are offered. Those interested in leading or facilitating workshops complete a Google form which asks them to identify the topics they are interested and experienced in, and to describe their qualifications. Graduate students are selected with an eye towards diversity, inclusivity, and experience. Though not a requirement to be an instructor for C&C, that experience may include opportunities offered to graduate students through the Department of Statistics, such as completing teaching-focused credits towards their degree or serving as an instructor for credit-bearing courses. During a semester, there are typically five to eight graduate students who lead or facilitate one workshop each, in order to broadly offer the experience to interested parties. The Stat Lab compensates the graduate students on a per-workshop basis.

The Stat Lab and Libraries provide overall guidance and coaching to the graduate students before and following each workshop. This support has included facilitation of practice sessions so instructors may become familiar with teaching spaces and technology, asking clarifying questions during the workshops, and offering feedback from observation directly following a session. Graduate students have also reached out informally for advice and insights from their peers, either from those who have previously led the workshop or from those who are involved with current workshops. Attendee feedback, when received, is provided to the instructors as well.

## **Adapting for Online Learning**

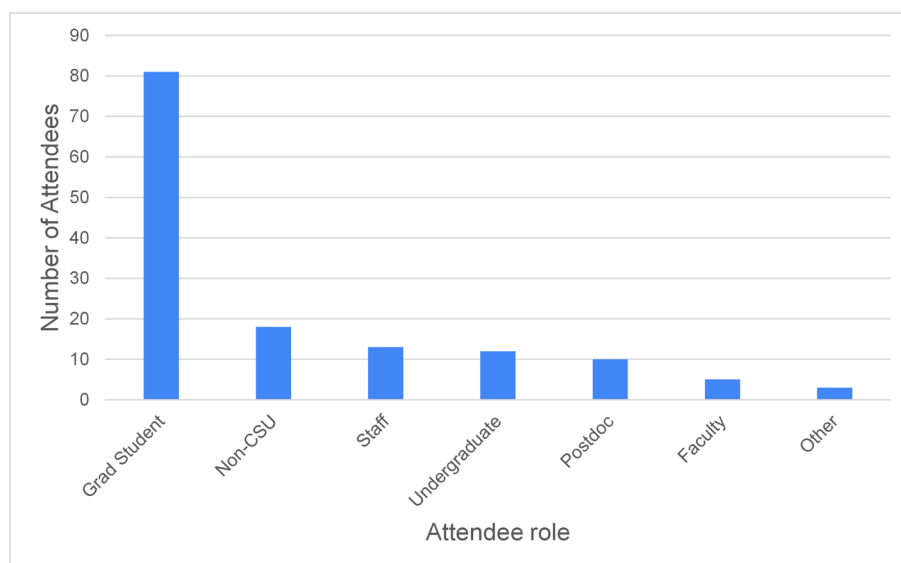
All ventures face challenges, and for the C&C program, the switch to online learning associated with COVID-19 was an unexpected one. Several workshops scheduled for spring 2020 were cancelled, and



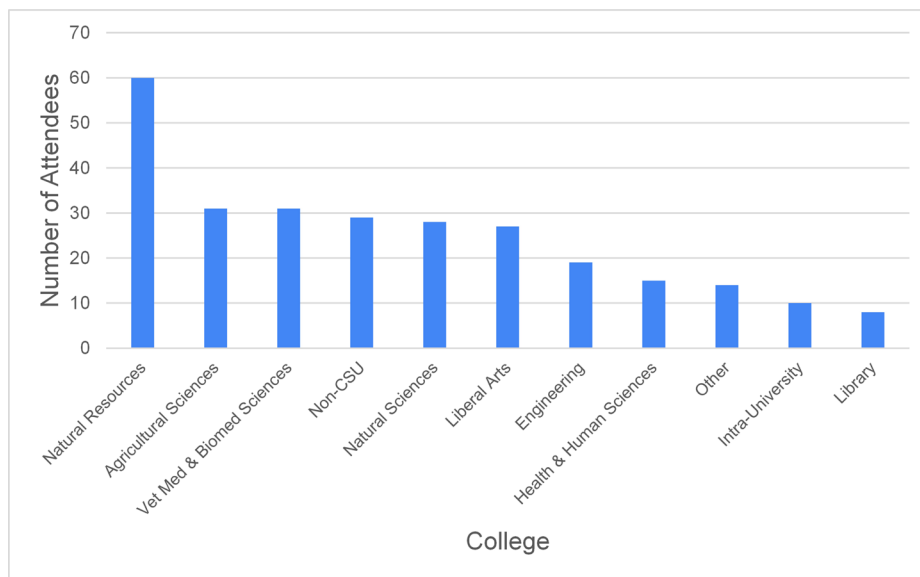
instead of leading those workshops, Statistics graduate students developed instructional materials and recorded videos to deliver the lessons virtually with a flipped-classroom format. The videos presented the main workshop content, and accompanying live sessions were adapted to include related example exercises so attendees could practice their skills after having watched the videos. Early in the pandemic, the live sessions were only offered virtually but have since transitioned to a hybrid model (i.e., in-person and online). By recording the content and making it available regardless of participation in the live sessions, the materials are now accessible to a broader audience than before. Videos in the most popular playlist, R Basics, have received more than 600 combined views (Colorado State University Libraries, n.d.). The challenge of switching to virtual coding workshops is by no means unique, and Chiewphasa and Moeller (2021) and Plomp, Tsang, and Martinez Lavanchy (2022) offer in-depth discussions of the process at their institutions.

## Evaluation and Impact

Registration and attendance data are routinely collected to facilitate the operation of the R workshops. Since the workshops are open to the public, these data indicate whether the attendees are affiliated with CSU, the colleges and departments represented, and the individual's role at the university (such as faculty or student), if any. This information helps organizers track participation trends and determine which units or roles are being served or overlooked by the workshops. For instance, the majority of attendees are CSU graduate students, but faculty and staff also participate, as do a small number of people unaffiliated with CSU such as employees of nearby government agencies (see Figure 1). Attendees come from all colleges in the university (see Figure 2), although the most prevalent represented colleges vary from semester to semester. Some individuals participate in multiple workshops, while others attend just one. The trends described above are based on unique attendees.



**Figure 1:** C&C workshop unique attendees by role at CSU, 2019 - 2022

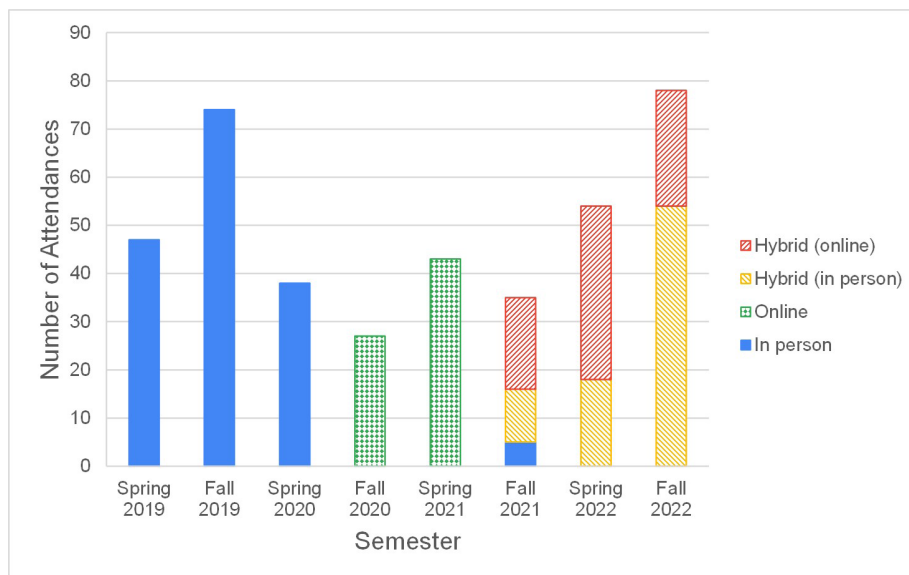


**Figure 2:** C&C workshop unique attendees by CSU college affiliation, 2019 – 2022

The Libraries and the Stat Lab have used this information, along with knowledge of research practices across the university, to guide outreach and publicity efforts. For example, the prevalence of graduate students at the workshops suggests a particular need within this population, which aligns with an observation by Dawn Paschal, then the Assistant Dean supervising data management services at the Libraries, who noted that “graduate students . . . are often the ones managing data in labs . . . They are the researchers of tomorrow” (Zuniga 2018, 33). Accordingly, the Stat Lab has advertised the workshops during presentations to graduate student audiences, such as incoming student orientation, and the Data Management Specialist highlights the workshops during guest lectures at graduate level courses. In another instance, when registration data suggested that the College of Veterinary Medicine and Biomedical Sciences was underrepresented among workshop participants given the amount and nature of research conducted within the College, new communication venues were pursued in an attempt to reach this population, such as placing a notice in the College’s weekly announcement email.

Along with registration and attendance data, the C&C workshop series is assessed using an online post-workshop survey. This voluntary survey asks attendees about their comfort level with executing a variety of activities in R both prior to and after the session, their perceptions of the instructor’s performance, and their impressions about the content and pace of the lesson. Similar questions ask about the recently implemented videos for pre-workshop viewing. The survey also solicits general comments. The response rate for the survey has been insufficient for generalizable analysis, but organizers have found the feedback helpful in understanding the varying backgrounds of attendees and how these backgrounds contribute to their

workshop experiences. As noted previously, survey responses are also shared with the session instructor as part of the training and development process. Since beginning in 2017, the C&C workshop series has offered more than 40 instruction sessions with over 500 total attendances. Figure 3 shows attendance numbers by semester and mode of instruction for the past four years. The workshops have been a consistently popular and visible service and have filled an important training gap for a variety of audiences who want to learn R, including students, faculty, staff, and community members.



**Figure 3:** C&C workshop attendance by semester and instruction modality, 2019 - 2022

### Benefits of the Partnership

The partnership between the Libraries and the Stat Lab has succeeded because it benefits everyone involved in the C&C workshop series. For the Libraries, the partnership has eased the instruction demands on the Data Management Specialist, who transitioned from the role of both organizer and teacher to the role of joint coordinator. This allowed the Data Management Specialist to focus on additional projects of importance to the Libraries and the CSU community. For the Stat Lab, participation in the C&C workshops has resulted in increased visibility on campus. The partnership has also presented an opportunity for the Stat Lab to begin branching out from its core consultative services to offer hands-on training to the university. The Stat Lab had long wanted to pursue such a program but was hindered by a lack of infrastructure and support, which the Libraries was able to provide. For the graduate student instructors, the workshops are an opportunity to develop their teaching skills while building on their existing subject knowledge. They learn how to teach attendees who have varying levels of expertise and how to adapt to new lesson plans and a new classroom environment. Graduate student instructors often choose to return for multiple semesters and can note this

experience on their resumes. Monetary compensation is an added benefit. Lastly, workshop participants learn from knowledgeable instructors who regularly use R for academic work. Attendees can ask questions at the time they occur and obtain real-time coding support and answers informed by the instructors' practical experience. One recent participant stated that "Coding and Cookies helped me a lot. It . . . was valuable because it helped show me what that program [R] was capable of . . . I can apply what I've learned" while another called the workshop they attended "the perfect introductory class as it . . . allowed everyone to learn at their own pace."

## Conclusion

In summary, the partnership between the Library and the Stat Lab has efficiently used existing resources and expertise, distributing the work of planning and implementing an R workshop series across two campus entities. Each entity is essential to the success of the program and contributes according to its strengths, deepening campus relationships at the same time. For others looking to begin a collaboration such as this, the first steps are to assess user needs, identify groups and individuals with the knowledge and skills to develop and deliver workshop materials, and ascertain who would be able to provide logistical support. Potential partners may vary. At CSU, the Stat Lab and Statistics graduate students were a natural fit with the Libraries, but institutions may find that other groups or populations such as undergraduate students, administrative professionals, or faculty are better suited to assist. Once potential partners are identified, it's time to reach out and start building connections.

---

## Acknowledgements

The authors would like to thank Dr. Tobin Magle for helping to establish the Libraries and Stat Lab partnership.

## Competing Interests

The authors declare that they have no competing interests.

---

## References

AAU and APLU (Association of American Universities and Association of Public & Land-Grant Universities). 2021. *Guide to Accelerate Public Access to Research Data*.

<https://www.aplu.org/library/guide-to-accelerate-access-to-public-data/file>.

Ali, Ibraheem, Thea Atwood, Renata Curty, Jimmy Gaphery, Tim McGeary, Jennifer Muilenberg, and Judy Ruttenberg. 2022. *Research Data Services Partnerships*. Washington, DC: Association of Research Libraries and Canadian Association of Research Libraries.

<https://doi.org/10.29242/report.rdspartnerships2022>.

- Atwood, Thea P., Andrew T. Creamer, Joshua Dull, Julie Goldman, Kristin Lee, Lora C. Leligdony, and Sarah K. Oelker. 2019. "Joining Together to Build More: The New England Software Carpentry Library Consortium." *Journal of eScience Librarianship* 8(1): e1161. <https://doi.org/10.7191/jeslib.2019.1161>.
- Bryant, Rebecca, Annette Dortmund, and Brian Lavoie. 2020. *Social Interoperability in Research Support: Cross-Campus Partnerships and the University Research Enterprise*. Dublin, OH: OCLC Research. <https://doi.org/10.25333/wyrd-n586>.
- Chiewphasa, Ben B., and Anna K. Moeller. 2021. "Reflections from Transitioning Carpentries Workshops Online." *Journal of eScience Librarianship* 10(4): e1217. <https://doi.org/10.7191/jeslib.2021.1217>.
- Colorado State University Libraries. n.d. "Coding and Cookies: Automating data cleaning and analysis using R." Accessed March 31, 2023. <https://libguides.colostate.edu/coding-cookies/r-basic>.
- CRDDS (Center for Research Data & Digital Scholarship). n.d. "What We Do." Accessed November 16, 2022. <https://www.colorado.edu/crdds/whatwedo>.
- Franklin A. Graybill Statistics and Data Science Laboratory. n.d. "Statistical Consulting at CSU." Accessed July 1, 2022. <https://statlab.colostate.edu/>.
- Harp, Matthew R., and Matt Ogborn. 2019. "Collaborating Externally and Training Internally to Support Research Data Services." *Journal of eScience Librarianship* 8(2): e1165. <https://doi.org/10.7191/jeslib.2019.1165>.
- Institutional Research, Planning, and Effectiveness. 2022. *Fact Book 2021-2022*. Fort Collins: Colorado State University. [http://irpe-reports.colostate.edu/pdf/fbk/2022/Full\\_FactBook\\_2021-22.pdf](http://irpe-reports.colostate.edu/pdf/fbk/2022/Full_FactBook_2021-22.pdf).
- Magle, Tobin. 2018. "Library Workshop Attendees Demographics." Unpublished report, February 2, 2018. PDF file.
- Michonneau, Francois, and Auriel Fournier, lesson maintainers. 2022. "Data Analysis and Visualization in R for Ecologists." Last modified October 19, 2022. <https://datacarpentry.org/R-ecology-lesson/>.
- NC State University Libraries. n.d. "Research Facilitation Service." Accessed November 16, 2022. <https://www.lib.ncsu.edu/rfs>.
- Plomp, Esther, Emmy Tsang, and Paula M. Martinez Lavanchy. 2022. "Taking the TU Delft Carpentries Workshops Online." *Journal of eScience Librarianship* 11(1): e1221. <https://doi.org/10.7191/jeslib.2022.1221>.
- Radecki, Jane, and Rebecca Springer. 2020. *Research Data Services in US Higher Education*. New York: Ithaka S+R. <https://doi.org/10.18665/sr.314397>.
- Wittenberg, Jamie, and Mary Elings. 2017. "Building a Research Data Management Service at the University of California, Berkeley: A Tale of Collaboration." *IFLA Journal* 43(1): 89-97. <https://doi.org/10.1177/0340035216686982>.
- Zuniga, Heidi. 2018. "Developing Better Data Management Support at Colorado State University Through Collaboration." *NCURA Magazine* 50(2): 32-35. [https://www.ncura.edu/Portals/0/Docs/Magazine/2018/MarchApril2018\\_NCURA\\_Magazine.pdf](https://www.ncura.edu/Portals/0/Docs/Magazine/2018/MarchApril2018_NCURA_Magazine.pdf).