

## Workshop Overview

Data Carpentry workshops introduce computational skills needed for data management and analysis. Participants are guided through the steps of the data lifecycle through hands-on exercises beginning with data cleaning and organization and through data visualization in R or Python. We required no prior computational experience for this workshop, which was open to WSU students and staff in addition to non-WSU participants.

We hosted the [Data Carpentry workshop](#) in October of 2019. Participants were taught how to organize data in spreadsheets, clean data in spreadsheets using OpenRefine, and manage and visualize data with R. Compared with our last offering of Data Carpentry, we did not run concurrent breakout groups for both R and Python languages due to space limitations. Instead, we focused on R programming. SQL was also dropped in favor of coverage of more advanced topics in R. The workshop syllabus can be found in Appendix B.

The workshop was team taught between four instructors with the assistance of six helpers. The helpers have previous experience with the software being taught and were present to troubleshoot participants' problems as they arose. The workshops are generally composed of tutorials that include hands-on exercises with live coding demonstrations. Challenge questions interspersed throughout each tutorial allow participants to practice throughout the workshop 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.

Because we hosted a self-organized Carpentry workshop there was not an administration fee to pay. We had a total of 32 registrants in this workshop, with a higher proportion of female attendees than male. Most participants were graduate students, although postdoctoral researchers, undergraduates, faculty, and other career stages were represented (Table 1). Of those attendees from WSU, members from 12 departments took part in the workshop. For participants by college, see Table A1. Twenty-five participants attended day one, and 22 attended day two.

*Table 1. Summary of Data Carpentry Workshop Demographics*

<b>Number of:</b>	
Registrants	32
Male	9
Female	17
Graduate student	24
Faculty	1
Postdoc	5
Undergraduate student	1
Other positions	1
Depts represented	12
Colleges represented	6
WSU campuses represented	3

*Subsection totals may not sum to 32 if not all participants responded to demographic questions. Data were gathered from multiple sources with different response rates.*

## 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 Carpentries to assess the effectiveness of their workshops by measuring participant responses before and after participating in Software Carpentry. Pre-workshop survey response rates were higher than post-workshop (84% vs. 34%). An overview of survey findings is given in Table 2.

Just above 50% of pre-workshop respondents had previous experience in a programming language, while 59% had previous experience with a statistical analysis program with a graphical user interface. Post-workshop survey results indicated that 100% of respondents felt they gained immediately applicable knowledge from the workshop. On average, participants said post-workshop that they were likely (93%) to recommend the workshop to a friend or colleague. Additionally, respondents appreciated the helpers' and instructors' attentiveness and the usefulness of the material taught.

*Table 2. Overview of survey responses*

	Pre-survey response rates (84%) were higher than post-survey response rates (34%)
	Around half of pre-workshop respondents came in with some previous experience programming, and all post-workshop respondents left feeling that they had gained immediately usable knowledge.
	More respondents had prior experience with a GUI-based statistical analysis program than with a programming language.
	Awareness of Data Carpentry came primarily from emails or flyers, followed mostly by advisors, friends or colleagues, and newsletters or websites.
	Respondents found the content to be relevant and the instructors/helpers to be attentive and knowledgeable.
	Respondents noted that there is a lot of content provided in the two-day format. Allowing for slower pace and more time to experiment with R were some suggestions provided.
	Respondents said that they felt comfortable learning in the workshop environment.

## Appendices

### Appendix A. Table

Table A1. Participants by College

College	Count
College of Agricultural, Human, and Natural Resource Sciences	13
Murrow College of Communication	7
College of Arts and Sciences	6
College of Veterinary Medicine	2
Not listed or non-WSU	2
College of Education	1
Voiland College of Engineering and Architecture	1
<b>Grand Total</b>	<b>32</b>

## **Appendix B. Workshop syllabus**

### **Data Organization**

- Organizing data in Excel/spreadsheets
- Data cleaning with OpenRefine

### **Programming in R**

- Working with objects, vectors, and data frames
- Reading data and working with factors
- Intro to dplyr, including count(), n(), and row\_number()
- Intro to tidyr, including left\_join(), right\_join(), inner\_join(), full\_join(), and Reduce() from base R
- Visualizing data with ggplot2, including investigating data distributions (ggridges) and exploring a dataset (ggrepel, ggpubr, gganimate, viridis)