

Designing Data Science Workshops for Data-Intensive Scientific Research

Allison Theobold

Department of Mathematical Sciences, Montana State University
Bozeman, MT

allisontheobold@montana.edu

February 6, 2020

References

- Allaire, J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., Wickham, H., Cheng, J., Chang, W., and Iannone, R. (2019). *rmarkdown: Dynamic Documents for R*. R package version 1.16. Available at <https://CRAN.R-project.org/package=rmarkdown>.
- American Statistical Association Undergraduate Guidelines Workgroup (2014). *2014 curriculum guidelines for undergraduate programs in statistical science*. American Statistical Association, Alexandria, VA.
- Andelman, S. J., Bowles, C. M., Willig, M. R., and Waide, R. B. (2004). Understanding environmental complexity through a distributed knowledge network. *BioScience*, 54(3):240–246.
- Cobb, G. (2015). Mere renovation is too little too late: We need to rethink our undergraduate curriculum from the ground up. *The American Statistician*, 69(4):266–282.
- Ernest, M., Brown, J., Valone, T., and White, E. P. (2018). Portal project teaching database.
- Friedman, J. (2001). The role of statistics in the data revolution. *International Statistics Review*, 69:5–10.
- Green, J. L., Hastings, A., Arzberger, P., Ayala, F. J., Cottingham, K. L., Cuddington, K., Davis, F., Dunne, J. A., Fortin, M.-J., Gerber, L., and Neubert, M. (2005). Complexity in ecology and conservation: Mathematical, statistical, and computational challenges. *BioScience*, 55(6):501–510.
- Hampton, S. E., Jones, M. B., Wasser, L. A., Schildhauer, M. P., Supp, S. R., Brun, J., Hernandez, R. R., Boettiger, C., Collins, S. L., Gross, L. J., Fernandez, D. S., Budden, A., White, E. P., Teal, T. K., Labou, S. G., and Aukema, J. E. (2017). Skills and knowledge for data-intensive environmental research. *BioScience*, 67(6):546–557.

- Hernandez, R. R., Mayernik, M. S., Murphy-Mariscal, M. L., and Allen, M. F. (2012). Advanced technologies and data management practices in environmental science: Lessons from academia. *BioScience*, 62(12):1067–1076.
- Michonneau, F., Teal, T., Fournier, A., Seok, B., Obeng, A., Pawlik, A. N., Conrado, A. C., Woo, K., Lijnzaad, P., Hart, T., White, E. P., Marwick, B., Bolker, B., Jordan, K. L., Ashander, J., Dashnow, H., Hertweck, K., Cuesta, S. M., Becker, E. A., Guillou, S., Shiklomanov, A., Kluges, D., Odom, G. J., Jean, M., Mislan, K. A. S., Johnson, K., Jahn, N., Mannheimer, S., Pederson, S., Pletzer, A., Fouilloux, A., Switzer, C., Bahlai, C., Li, D., Kerchner, D., Rodriguez-Sanchez, F., Rajeg, G. P. W., Ye, H., Tavares, H., Leinweber, K., Peck, K., Lepore, M. L., Hancock, S., Sandmann, T., Hodges, T., Tirok, K., Jean, M., Bailey, A., von Hardenberg, A., Theobald, A., Wright, A., Basu, A., Johnson, C., Voter, C., Hulshof, C., Bouquin, D., Quinn, D., Vanichkina, D., Wilson, E., Strauss, E., Bledsoe, E., Gan, E., Fishman, D., Boehm, F., Daskalova, G., Tavares, H., Kaupp, J., Dunic, J., Keane, J., Stachelek, J., Herr, J. R., Millar, J., Lotterhos, K., Cranston, K., Direk, K., Tylen, K., Chatzidimitriou, K., Deer, L., Tarkowski, L., Chiapello, M., Burle, M.-H., Ankenbrand, M., Czapanskiy, M., Moreno, M., Culshaw-Maurer, M., Koontz, M., Weisner, M., Johnston, M., Carchedi, N., Burge, O. R., Harrison, P., Humburg, P., Pauloo, R., Peek, R., Elahi, R., Cortijo, S., Umashankar, S., Goswami, S., Yanco, S., Webster, T., Reiter, T., Pearse, W., and Li, Y. (2019). *datacarpentry/r-ecology-lesson: Data carpentry: Data analysis and visualization in R for ecologists*, July 2019.
- Nolan, D. and Temple Lang, D. (2010). Computing in the statistics curricula. *The American Statistician*, 64(2):97–107.
- RStudio Team (2015a). *RStudio Cloud*. RStudio, Inc., Boston, MA.
- RStudio Team (2015b). *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA.
- Strasser, C. A. and Hampton, S. E. (2012). The fractured lab notebook: Undergraduates and ecological data management training in the united states. *Ecosphere*, 3(12):1–18.
- Teal, T. K., Cranston, K. A., Lapp, H., White, E., Wilson, G., Ram, K., and Pawlik, A. (2015). Data carpentry: Workshops to increase data literacy for researchers. *International Journal of Digital Curation*, 10(1):135–143.
- The Carpentries (2019). <https://carpentries.org/>.
- Wickham, H. (2014). Tidy data. *The Journal of Statistical Software*, 59.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York.
- Wickham, H. (2017). *tidyverse: Easily Install and Load the ‘Tidyverse’*. R package version 1.2.1.