The growing importance of reproducibility and responsible workflow in the data science and statistics curriculum

Aneta Piekut
Univ. of Sheffield

Colin Rundel
Duke University

Micaela Parker
ADSA

Nicholas Horton
Amherst College

Rohan Alexander
Univ. of Toronto
Welcome!

Host and Moderator: Leigh Johnson  
Capital University
What's new in the journal?

**Latest Article Link**

**Playmeans: Inclusive and Engaging Data Science through Music**
Davit Khachatrian
Published online: 30 Nov 2022

**Collections Link**

**Collections**
Browse article collections from Journal of Statistics Education.

Teaching Simpson's Paradox, Confounding, and Causal Inference
Created 06 Sep 2022 | Last updated 08 Sep 2022
Edited by Nicholas J. Horton
SCRATCH to R: Toward an Inclusive Pedagogy in Teaching Coding

Shu-Min Liao
Amherst College

January 17, 2023 at 4:00 PM EST
Consortium for the Advancement of Undergraduate Statistics Education

Important Dates:

Posters & Beyond
- January 29, 2023 (will receive feedback)
- March 5, 2023 (final abstracts due)

Birds-of-a-Feather Discussions
- Proposals due May 6, 2023

Conference runs June 1 - 3, 2023 (with workshops beginning May 30th) at the Penn Stater Hotel in State College, PA

https://www.causeweb.org/cause/uscots/uscots23
Consortium for the Advancement of Undergraduate Statistics Education

https://www.causeweb.org/usproc/
Submission Deadline: June 23, 2023

Reading Group

Article Link
Zoom Registration Link
Dec. 14, 2023 12:00 - 1:00 PM
Aneta Piekut is Senior Lecturer in Quantitative Social Sciences at the Sheffield Methods Institute, University of Sheffield. She is an associate editor of the *Journal of Statistics and Data Science Education*. 
Colin Rundel is an associate professor of the practice in Statistical Science at Duke University. His research interests include applied spatial statistics with an emphasis on Bayesian statistics and computational methods. Colin was a guest editor of the special issue on teaching reproducibility and responsible workflow.
Micaela Parker is the Founder and Executive Director of the Academic Data Science Alliance (ADSA), https://academicdatascience.org. Micaela was a guest editor of the special issue on teaching reproducibility and responsible workflow.
Nicholas J. Horton is Beitzel Professor of Technology and Society (Statistics and Data Science) at Amherst College. He is the editor of the *Journal of Statistics and Data Science Education* and is the co-chair of the National Academies Committee on Applied and Theoretical Statistics (CATS).
Rohan Alexander is an assistant professor at the University of Toronto jointly appointed in the Faculty of Information and the Department of Statistical Sciences. He is an associate editor of the *Journal of Statistics and Data Science Education*.
“The growing importance of reproducibility and responsible workflow in the data science and statistics curriculum” (Nicholas J. Horton, Rohan Alexander, Aneta Piekut, Colin Rundel, https://doi.org/10.1080/26939169.2022.2141001)


“LUSTRE: An online data management and student project resource” (John Towse, Rob Davies, Ellie Ball, Rebecca James, Ben Gooding, and Matthew Ivory, https://www.tandfonline.com/doi/full/10.1080/26939169.2022.2118645)
“Teaching for Large-Scale Reproducibility Verification” (Lars Vilhuber, Hyuk Harry Son, Meredith Welch, David N. Wasser, and Michael Darisse, https://www.tandfonline.com/doi/full/10.1080/26939169.2022.2074582)


“A Journey from Wild to Textbook Data to Reproducibly Refresh the Wages Data from the National Longitudinal Survey of Youth Database” (Dewi Amaliah, Dianne Cook, Emi Tanaka, Kate Hyde, and Nicholas Tierney, https://www.tandfonline.com/doi/full/10.1080/26939169.2022.2094300)

“Approachable case studies support learning and reproducibility in data science: An example from evolutionary biology” (Luna L. Sanchez Reyes and Emily Jane McTavish, https://www.tandfonline.com/doi/full/10.1080/26939169.2022.2099487)
1. What insights have you gained from these papers?
Questions for the panel

1. What insights have you gained from these papers?
2. What next steps are you going to take in response to these papers?
Questions for the panel

1. What insights have you gained from these papers?
2. What next steps are you going to take in response to these papers?
3. Audience Q&A
Questions for the panel

1. What insights have you gained from these papers?
2. What next steps are you going to take in response to these papers?
3. Audience Q&A
4. Closing thoughts: data and code sharing
As of September 1, 2022, all submissions to the *Journal of Statistics and Data Science Education* require a “Data Availability Statement” which outlines how code and data underlying a paper have been made available.

See [https://www.tandfonline.com/journals/ujse21](https://www.tandfonline.com/journals/ujse21) for more info