

Call for papers: Teaching reproducibility and responsible workflow (Journal of Statistics and Data Science Education)

Modern statistics and data science utilizes an iterative data analysis process to solve problems and extract meaning from data in a reproducible manner. Models such as the PPDAC (Problem, Plan, Data, Analysis, Conclusion (<https://dataschools.education/about-data-literacy/ppdac-the-data-problem-solving-cycle>)) Cycle have been widely adopted in many pre-secondary classrooms. The importance of the data analysis cycle has also been described in guidelines for statistics majors (<https://www.amstat.org/asa/education/Curriculum-Guidelines-for-Undergraduate-Programs-in-Statistical-Science.aspx>), undergraduate data science curricula (<http://dstf.acm.org>), and data science courses (<https://r4ds.had.co.nz/introduction.html>).

The National Academies of Science, Engineering, and Medicine's 2018 "Data Science for Undergraduates" consensus study (<https://nas.org/envisioningds>) identified the importance of workflow and reproducibility as a component of data acumen needed in our graduates. The NASEM report reiterated that "documenting, incrementally improving, sharing, and generalizing such workflows are an important part of data science practice owing to the team nature of data science and broader significance of scientific reproducibility and replicability." They also noted that reproducibility and workflow raised important questions about the ethical conduct of science. These reports identify the need for students to have multiple experiences with the entire data analysis cycle.

However, many challenges exist:

1. technologies are rapidly evolving
2. few faculty were trained in the use of these methods
3. best practices have not been clearly identified
4. insufficient vetted and inclusive curricular materials are available
5. accounting for student heterogeneity and broadening participation
6. many aspects of student understandings in this area are unknown

To highlight work in this important and developing area, the *Journal of Statistics and Data Science Education* is inviting submission of papers related to "Teaching reproducibility and responsible workflow" to appear in a forthcoming issue.

Sample topics (non-exhaustive)

- Teaching workflows and workflow systems
- Fostering reproducible analysis
- Promoting reproducibility as a component of replicability and scientific conduct
- Developing and implementing documentation and code standards
- Incorporating source code (version) control systems

- Supporting collaboration
- Integrating ethics
- Conducting effective formative and summative assessment

Submissions at all levels of education (primary through graduate programs and continuing education) and disciplines (social sciences, digital humanities, and STEM) are encouraged.

Timetable

- May 2021 (call for submissions)
- September 1, 2021 (call for reviewers)
- September 15, 2021 (deadline for submissions via the *Journal of Statistics and Data Science Education* submission site <https://mc.manuscriptcentral.com/ujse> (<https://mc.manuscriptcentral.com/ujse>), please select the “Teaching reproducibility and workflow” option)
- July 2022 (proposed publication date)

Papers received after September are in scope and will be considered as regular submissions.

About the journal

The *Journal of Statistics and Data Science Education* is an open-access peer-reviewed journal with no author fees that is published by Taylor and Francis and the American Statistical Association. Articles accepted for publication are promptly made available online and featured on the journal website (<https://www.tandfonline.com/toc/ujse21/current>).

Questions about submissions or the timeline? Please contact Nicholas Horton (Amherst College, JSDSE Incoming Editor).

Last updated May 18, 2021