

Framing and Broadening Understanding of Research Norms in a Second Course in Data Science

Tori Ellison, University of Illinois Urbana-Champaign

Proposed Data Science E-Book: <https://exploration.stat.illinois.edu/>

Website: <http://tori-ellison.com>



Abstract

How can we encourage students to think critically and invite them to “be a part of the conversation” when it comes to the research norms surrounding data science? We’ve developed and published online content for an undergraduate second course in data science that is framed upon building a deep awareness of the myriad of different decisions that can be made when completing a beginning-to-end data science analysis. Through engaging real-world examples and datasets, we explore how these different decisions can lead to different answers to the same research question. With this in mind, we frame and embed our course with discussions and exercises which encourage students to think critically about the research norms that surround how we answer questions with data.

We introduce the first section of this text as an example of how educators might frame a data science or statistics course in such a way. We introduce targeted questions that the students answer in projects and peer-evaluation surveys that encourage these critical thinking skills as well as assess them.

This content is free and accessible to all students and includes units on data wrangling, linear regression, feature selection including cross-validation, logistic regression, simulating sampling distributions, and statistical inference.

References

American Statistical Association. "ASA Ethical Guidelines for Statistical Practice. 2018."

Angwin, Julia, et al. "Machine bias risk assessments in criminal sentencing." ProPublica, May 23 (2016).

Meyer, David. "Amazon Reportedly Killed an AI Recruitment System Because It Couldn't Stop the Tool from Discriminating Against Women." Fortune, 10 Oct. 2018, <https://fortune.com/2018/10/10/amazon-ai-recruitment-bias-women-sexist/>.

O’neil, C. "Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy: Crown; 2016." Google Scholar Google Scholar Digital Library Digital Library.

Saltz, Jeffrey S., Neil I. Dewar, and Robert Heckman. "Key concepts for a data science ethics curriculum." Proceedings of the 49th ACM technical symposium on computer science education. 2018.

Tractenberg, Rochelle E., et al. "Using ethical reasoning to amplify the reach and resonance of professional codes of conduct in training big data scientists." Science and Engineering Ethics 21 (2015): 1485-1507.