

Exploring Data Science Education: From Tutorials to Assessment

Proposal: Graduation with Distinction

Evan Dragich

supervised by Dr. Mine Çetinkaya-Rundel, PhD.

Spring 2023

Background

An essential component of any educational research is a validated, relevant instrument to measure students' learning outcomes. Whether to award college credit like the College Board's AP and CLEP exams, to measure students' previously-held misconceptions, or just as a tool upon which to evaluate educational interventions, such assessments have been developed and statistically analyzed for a wide range of subjects such as Spanish, Psychology, Chemistry, and Calculus (Godfrey and Jagesic 2016; Solomon et al. 2021; Mulford 1996; Epstein 2013).

In the field of statistics, previous work on measuring students' reasoning skills led to the development of the Comprehensive Assessment of Outcomes in Statistics (CAOS). The revised CAOS 4, comprising 40 multiple-choice items on a variety of commonly-taught first-semester introductory concepts, was first administered in 2005 and allowed instructors to measure whether their courses were successfully resulting in their desired learning outcomes. However, many instructors noted that their findings reflected a much lower understanding than expected, specifically regarding the topics of data visualization and data collection (Delmas et al. 2007). However, a key feature of the CAOS was the lack of hard computation nor need to recall specific formulas or definitions, allowing greater accessibility for a variety of statistics-adjacent uses.

In 2022, as more universities begin to support the emerging field of data science via specific courses, concentrations, or even majors, there is a need to measure students' learning outcomes in these introductory classes analogously to the subjects named above (Swanstrom, n.d.). Lacking a clearly-defined scope, empirical studies of so called "data science" curricula suggest that the field can be thought of an augmentation of traditional statistical modeling concepts, with emphases on computing, data visualization and manipulation, as well as a consideration of ethics and the role data plays in society (Zhang and Zhang 2021).

Methodology

The goal of this work is to create and validate a 25-40 question multiple-choice assessment that instructors of introductory data science courses can use to obtain standardized, objective measurements of their students' success in learning the course's desired outcomes.

To create such an assessment, the research team has drafted candidate items on a variety of topics and gathered "think-aloud" interview feedback from university professors from a variety of backgrounds and teach introductory data science courses. The team will then analyze the feedback on each item, considering modifications or deletion, aiming to integrate the variety of perspectives gained from the interviews.

Once a draft has been made to incorporate any edits from the professor stage, the assessment will be administered to students in the form of focus groups and/or traditional written educational assessment (ie. a typical online test).

As well, to supplement the assessment work, we plan to finalize the `{dsbox}` R package for release on CRAN. The package consists of a series of 10 introductory data science tutorials built with the interactive `{learnr}` package which teach many of the concepts that appear on the assessment.

Data Access and Technology Needs

The interview transcripts and notes from the professor interviews has been collected and is securely stored on a private Github with other research team members. Data collection from students will occur similarly in the case of focus groups, or via Qualtrics in the case of traditional written administration. No special technology needs are anticipated.

Ethical Implications

In Spring 2022, the previous work involving interviews with professors was approved under Duke IRB protocol #2022-0399. This protocol was granted with exemption #2, related to the educational assessment nature of the research and that participants could not be linked to their responses nor reasonably put at risk of harm.

We are in the process of applying for a new IRB protocol to conduct focus groups and/or pilot assessments with introductory data science (STA199) students, which will ideally occur before the end of the fall semester.

While there will be no direct benefit to our participants for participating in this research, the creation of such a standardized assessment without mathematical or computing pre-requirements will improve accessibility to introductory data science courses for future generations of students, particularly those of historically underserved populations.

Timeline

Including previous work:

January 2022 - April 2022

Existing items reviewed by newest research team member and coalesced into single, cohesive, browseable assessment in the form of a Quarto book. Final {dsbox} tutorial written, and package was successfully edited to pass all necessary checks.

April 2022 - May 2022

Three 2-hour interviews conducted with faculty members to obtain detailed feedback on each item, as well as brainstorm general ideas on which topics and themes to include.

May 2022

Universally-agreed upon edits (typos, rendering formatting issues, tense disagreements) were amended and published.

September 2022 - October 2022

Research team meeting once a week to integrate professor feedback and arrive at pilot assessment ready for student use. IRB proposal for student focus groups/individual assessment drafted and subsequently submitted.

November 2022 - December 2022

Assessment administered to STA199 students. {dsbox} package finalized and submitted to CRAN by the end of the semester.

January 2023 - February 2023

Student feedback discussed among research group, and final edits made to assessment. Depending on timing and IRB, assessment administered as a “pre” test to STA199 students.

March 2023

Thesis written, including a description of work on {dsbox} package, as well as a summary of work on assessment, specifically the process of iteratively reworking items and identifying salient takeaways from the variety of feedback received. Depending on status of earlier Spring 2023 student administration, a statistical analysis of the instrument regarding subscales and/or validity may be conducted or written as a registered report.

Proposed target conference or publication

We aim to publish the final assessment and associated paper in the Journal of Statistics and Data Science Education (JSDSE), and/or present at the United States Conference on Teaching Statistics (USCOTS)

Bibliography

- Çetinkaya-Rundel, Mine, and Victoria Ellison. 2021. “A Fresh Look at Introductory Data Science.” *Journal of Statistics and Data Science Education* 29 (sup1): S16–26.
- Delmas, Robert, Joan Garfield, Ann Ooms, and Beth Chance. 2007. “ASSESSING STUDENTS’ CONCEPTUAL UNDERSTANDING AFTER A FIRST COURSE IN STATISTICS.” *STATISTICS EDUCATION RESEARCH JOURNAL* 6 (2): 28–58.
- Epstein, Jerome. 2013. “The Calculus Concept Inventory-Measurement of the Effect of Teaching Methodology in Mathematics.” *Notices of the American Mathematical Society* 60 (8): 1018–27.
- Godfrey, Kelly E., and Sanja Jagesic. 2016. *Validating College Course Placement Decisions Based on CLEP Exam Scores: CLEP Placement Validity Study Results. Statistical Report.* College Board.
- Mulford, DouglasRobert. 1996. “An Inventory for Measuring College Students’ Level of Misconceptions in First Semester Chemistry.” *Unpublished Master’s Thesis, Purdue University, IN.*
- Solomon, Erin D., Julie M. Bugg, Shaina F. Rowell, Mark A. McDaniel, Regina F. Frey, and Paul S. Mattson. 2021. “Development and Validation of an Introductory Psychology Knowledge Inventory.” *Scholarship of Teaching and Learning in Psychology* 7: 123–39.
- Swanstrom, Ryan. n.d. “Data Science Colleges and Universities.” <http://datascience.community/colleges>.
- Zhang, Zhiyong, and Danyang Zhang. 2021. “What Is Data Science? An Operational Definition Based on Text Mining of Data Science Curricula.” *Journal of Behavioral Data Science* 1 (1): 1–16.