⚡ EdSurge

OPINION | TEACHING & LEARNING

New Data Science Standards Are Needed for a Data-Filled World. Here's What We Propose.

By **Jo Boaler** and **Rob Gould**   Oct 11, 2021

whiteMocca / Shutterstock

For more than a century there have been calls to increase "statistical literacy." And such cries have grown even louder in recent years, with the rise of big data and fast computing.

The reality, though, is that an unacceptably high number of students fail algebra courses, which focus students on outdated methods and calculations performed by hand. These antiquated curricula have discouraged many students from persisting in STEM fields, and exacerbated inequities prevalent in the U.S.

A different approach to teaching mathematics is needed—one that develops data literacy for all students. Not only would such an approach be more relevant and increase student engagement, it has the potential to reduce the widespread vulnerability to misleading information shared via social media.

Research has shown that students are not being well prepared to be critical consumers of data and online resources, which has led to concerns about a threat to our democracy, which relies on voters' ability to sort truths from lies. On the other hand, the emerging field of data science, defined as a synthesis of statistics, mathematics and computer science, promises to provide students with powerful problem-solving strategies they will use in the workplace and their daily lives. And the need for people who can reason with data in almost all jobs in all sectors of the economy.

For K-12 educators today, this represents a challenge: how can teachers infuse in their young students an interest in the new discipline of data science.

But there has long been a missing piece: a lack of standards for data science. This situation continues even as schools and districts across the U.S. recognize the need for data literacy; that some state frameworks call for attention to data literacy (such as the 2021 California Mathematics Framework); and teachers across subject areas develop their own data lessons and courses.

Although data science is interdisciplinary, one possible home for data science standards is in mathematics standards, as there are important mathematical tools and methods that support data science. Another possibility is a separate set of standards that stand apart from mathematics—increasing the possibility to develop a truly interdisciplinary approach to developing students' data acumen. In either scenario the time seems ripe for planting a flag in the ground and offering ideas for the development of data literacy and data science through the grades. Such standards can prepare students as they move through middle school and high school and be complemented and deepened by a high school data science course, that some states and colleges now accept as an alternative to algebra 2.

At the high school level, teaching the synthesis of mathematical, statistical and computational thinking that make up data science can lead students not only to important and well-paying careers, but it can also eradicate the inequities built into the calculus pathway. In most districts in the U.S., high-achieving students engage in what is known as a "race to calculus," missing courses in middle school to get to the calculus pinnacle. Yet research shows that most students who take calculus in school repeat it or take a lower-level course in college.

The need to compress courses to reach calculus also means that most students are filtered out of the pathway in middle school, and the students chosen to go forward are disproportionately white and male. Data science provides a more equitable alternative to calculus that will not require middle school tracking, and will connect with students' daily lives and communities, raise awareness of social-justice issues, and appeal to broader groups of students.

This will not be a lower-level pathway, either, since data science is a rigorous discipline that is rich and important for many different college majors, inside STEM subjects and the humanities. The National Academy of Education recently called for high school courses that engage students in civic reasoning—focusing on exactly the mathematical content and practices set out in currently available data science courses. One example is the Mobilize Introduction to Data Science course that was jointly developed by UCLA and the Los Angeles Unified School District and Stanford's Youcubed: Explorations in Data Science.

In this associated publication we lay out a set of standards that build from the American Statistical Association's PreK-12 Guidelines for Assessment and Instruction in Statistics Education. One important quality of the standards is that, at every grade, they are enfolded within a data investigation cycle. Data science should not be taught as a set of disconnected methods but instead as an approach to problem-solving with data, highlighting mathematical content and practices. As students advance, they will actively engage in this problem-solving investigation cycle with increasing levels of

sophistication. Although each grade level lists important knowledge, the knowledge is linked and developed as part of a coherent whole. The data cycle we envisage looks like this:

Our goal in setting out these standards is not to claim that they are the only way to develop data literacy through the grades, but to raise awareness and to start or enrich conversations happening across the U.S. Some highlights of the data science standards we propose include students developing curiosity about events in their lives that can be considered with data, learning to pose their own statistical investigative questions on topics that interest and affect them, and confronting the ethical implications of data collection and analysis.

Establishing standards is only the first step, though. Much work still remains in preparing educators to teach data science, setting expectations for parents and securing the required resources. Organizations across the country are working together to spread awareness of the need for data science in schools (see for example, The Messy Data Coalition and youcubed's data science resources)), and online courses are being made to prepare teachers in the important knowledge and teaching pedagogy they will need (see for example our own YouCubed program. In addition, the American Statistical Association has a broad repository of teaching resources.

But standards are an important piece of the puzzle, and one that we hope will unlock further work and consideration–elevating a content area that is currently in its infancy, but may be the most important of all in the preparation of data literate citizens, empowered to navigate and understand their data filled worlds.

Jo Boaler is a Professor of Mathematics Education at Stanford University, and the faculty director of youcubed, an education resource that has reached over 230 million students.

Rob Gould is a professor of statistics at UCLA and director of the Center for Teaching Statistics.

Big Data   Teaching & Learning

Learn more about EdSurge operations, ethics and policies here. *Learn more about EdSurge supporters* here.