Data Science

A Joint Position of NCTM, NSTA, ASA, NCSS, and CSTA

Our Position

Data science bridges disciplines and thus should be introduced and taught across the curriculum in K-12 schools to help develop informed users of data. Data science captures the complexity of data and data methods that have arisen with advances in technology, including breakthroughs in artificial intelligence. It is a collaborative science that uses complex data and methods to explain trends and patterns with a critical piece being its interdisciplinary nature. K-12 education plays the critical role of scaffolding students' experiences in addressing complex data sets. All subjects in school should recognize the contribution of data to their discipline and take curricular approaches that integrate data with disciplinary lessons where appropriate.

Introduction

Data can be numbers, counts, and measurements but also images, video, sounds, or words. Students need to grapple with the fast-changing nature of data, yet a gap exists between the concepts taught in math and the data skills needed by other disciplines. All learners need opportunities to develop data literacy, knowledge, and skills.

Different disciplines may have different concepts of data and have different analysis techniques that have been developed for their specific needs. However, the approach to working through a problem and answering questions with data are consistent themes in practices used across all disciplines (Reynante et al., 2020; National Academies of Sciences, Engineering, and Medicine, 2023). Data science tools and practices can help students learn with and through data to make meaning of phenomena and issues across disciplines. A data-enabled approach affords students an opportunity to bring their personal interests, histories, cultural identities, and self into classrooms where learning is connected to *their* real world (e.g., Louie et al., 2021; V. Lee, Wilkerson, & Lanouette, 2021).

Building from the practices of data scientists who use data to solve real world problems, H. Lee et al. (2022) proposed that a data investigation process could provide students with rich experience with data across disciplines. A close look at standards and frameworks across disciplines show similar approaches to the use of data. Disciplinary guidance provides direction for allowing all students to be data scientists. For example, the *Guidelines for Assessment and Instruction in Statistics Education II* report (American Statistical Association, 2020) speaks to this specifically with statistics and mathematics, while the *College, Career, and Civics* (C3) framework addresses this in social studies (National Council for the Social Studies, 2013). For computing, Grillenberger and Romeike (2018) provide a comprehensive theoretical foundation for data literacy.

Declarations or Guiding Principles

1. **Data science is contextual and interdisciplinary.** Across disciplines, data science offers a common approach: framing the problem, generating questions, planning solutions, collecting data, analysis, visualization, and communication. Data science draws on concepts from mathematics, statistics, computing, as well as content knowledge of the domain within which a problem is situated. Given this, data science offers tools and approaches that can be utilized across disciplines to allow students to pose interesting questions, investigate the available data, analyze and answer questions by finding meaning in the data, and then communicate the message in the data.

2. **Data science is an investigative process.** Asking a compelling question that necessitates data to answer it is at the heart of data science. Each discipline has a set of robust and reliable discipline-specific techniques that are then used to analyze data and interpret and communicate the results. This process is iterative and questioning is key at guiding each point in the investigative process. While the terminology may differ across disciplines, the basic concept is the same: data science is grounded in inquiry.

3. **Data science understandings and experiences are for everyone.** Experiences with making sense of data should start in elementary school and progress as what students understand about the world becomes more complex. Making sense of and understanding the data must be the focus rather than the mechanics of computations and calculations. Appropriate grade-level technology should be used to work with the data at hand, especially as the problems increase in complexity. Data science is important for all students due to the increased importance of data-driven decision making and the data deluge in our daily lives.

4. **Data science educators must develop and practice ethical uses of data.** In preparation for engaging data towards civically responsible ends, students must learn to question the sources of data, such as individuals, devices, or systems that generate or contribute to its creation and track the various transformations and processes applied to data, such as cleaning, aggregation, or analysis. Ethical application of data science requires considering the ways data are collected and represented, how data are analyzed, and how findings are disseminated. The provenance of data is fundamental to understanding data quality; we must be honest in how the data were collected and assumptions that were made during collection. Working with data requires transparency regarding interpretation of the data, including methods used, limitations, possible sources of error, and algorithmic biases.

References

American Statistical Association (ASA). 2020. Pre-K–12 Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report II: A Framework for Statistics and Data Science Education. Alexandria, VA: ASA. https://www.amstat.org/asa/files/pdfs/GAISE/GAISEIIPreK-12_Full.pdf.

Grillenberger, A., & Romeike, R. (2018). Developing a theoretically founded data literacy competency model. In *Proceedings of the 13th Workshop in Primary and Secondary Computing Education* (pp. 1-10).

Lee, H. S., Mojica, G. M., Thrasher, E. P., & Baumgartner, P. (2022). Investigating data like a data scientist: Key practices and processes. *Statistics Education Research Journal*, 21(2), https://doi.org/10.52041/serj.v21i2.41.

Lee, V. R., Wilkerson, M. H., & Lanouette, K. (2021). A call for a humanistic stance toward K–12 data science education. *Educational Researcher*, 50(9), 664-672.

Louie, J., Stiles, J., Fagan, E., Roy, S., & Chance, B. (2021). Data investigations to further social justice inside and outside of STEM. *Connected Science Learning*, 3(1).

National Academies of Sciences, Engineering, and Medicine. (2023). *Foundations of Data Science for Students in Grades K-12: Proceedings of a Workshop*. Washington, DC: The National Academies Press. https://doi.org/10.17226/26852.

Reynante, B.M., Selbach-Allen, M.E. & Pimentel, D.R. (2020). Exploring the promises and perils of integrated STEM through disciplinary practices and epistemologies. *Science & Education 29*, 785–803. https://doi.org/10.1007/s11191-020-00121-x.

National Council for the Social Studies (NCSS). 2013. The College, Career, and Civic Life (C3) Framework for Social Studies State Standards. Silver Spring, MD: NCSS. www.socialstudies.org/standards/c3