NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS

Teaching Data Science in High School: Enhancing Opportunities and Success

Ensuring that all students have the mathematical experiences necessary to increase their opportunities for personal and professional success is essential. Data science is a rigorous, engaging, and practical field of study and can be a significant part of a high school student's mathematical experience. Knowledge of data science is important, and a data science course should be accepted as a high school mathematics course that can be used for credit towards graduation, provided the course includes or builds on previous, substantive student work with essential concepts, knowledge, skills, and habits of mind in mathematics and statistics, as described in Catalyzing Change (NCTM, 2018).

Introduction

Data science captures the complexity of data and data methods that have arisen with advances in technology. It is a collaborative science that uses complex data and methodology to identify and explain trends and patterns in a particular context. Since data science bridges disciplines, it should be taught across the curriculum. Because of the structure of most high schools, data science courses will likely reside in mathematics departments.

The continual and rapid increase in the capacity of technology to collect, organize, and manage data compels educators to prepare students to live in this world "awash in data" (Erickson, 2017). Students should become familiar with data science, a field that is quickly changing and evolving. Every second of every day, the world creates enough data to fill 50 new libraries of Congress (Domo, n.d.).

It is essential that students understand data so that they can comprehend the massive amounts of information that they encounter on a regular basis, and is available at their fingertips. High-quality experiences working with data expose students to new and different kinds of content that can energize and motivate them and enable them to see many uses for mathematics to make sense of the world around them.

Declarations

- All students should have the opportunity to take four years of high school mathematics, and data science content should be available to all students in order to complete their high school mathematics graduation requirement.

  Catalyzing Change recommends "that high schools should ensure that all students enroll in a mathematics course every year in high school, and complete four years of high school mathematics,

including a mathematics or statistics course during their last year of high school." (p.83) The Essential Concepts should be covered within the first two and no longer than the first three years of high school. In the final years, high schools can offer a variety of courses, including those that contain Data Science content.

- A high school data science course merits mathematics credit if it includes substantive student work with essential concepts, including those from Functions, Quantitative Literacy, Visualizing and Summarizing Data, Statistical Inference, and Probability (NCTM, 2018).

  The mathematics should include attention to
  - *quantitative reasoning, making and defending data-based conclusions, and assessing risks,*
  - *describing functions using a variety of representations,*
  - *identifying key features of functions represented graphically and associating these features with equivalent symbolic representations.*

  Students' experiences with statistics should include
  - *the statistical investigative process (Bargagliotti, 2020), which provides a template for statistical thinking when working with data,*
  - *visualizing distributions of one or more variables to reveal essential qualities of the context from which the data were collected,*
  - *sampling and approximating sampling distributions of a sample statistic formed from repeated random samples of a given size drawn from a population to develop intuition about sampling variability, leading to an understanding of concepts such as margin of error,*
  - *randomization and the role it plays in study design, where random sampling supports generalizability, and random assignment supports investigations of causality.*

  This work should have an emphasis on modeling and applying appropriate statistical analytical models and measures of goodness-of-fit of those models to understand the underlying context that produced the data.

- A high school data science course merits mathematics credit if it includes substantive student work with skills students develop from their understanding of the essential concepts.

  Skills relevant to mathematics and statistics related to data involve identifying, sorting and analyzing data characteristics to categorize and to formulate and test predictions based on data; employing appropriate technological tools in visualizing, modeling, and constructing meaningful representations using data; using mathematics and data to make decisions and form conclusions; communicating conclusions (or critiquing others' conclusions) effectively, precisely and with sound justification and argumentation.
  See the appendix for a list of skills.

- A high school data science course merits mathematics credit if it includes substantive student work with habits of mind in mathematics and statistics.

  > *Student work should demonstrate scaffolded learning of mathematics/statistics concepts with a habits of mind framework that supports recognizing patterns, decomposing problems, thinking algorithmically, doing-undoing problems, building rules to represent functions, and abstracting from computation. A student's knowledge should go beyond content to include skills such as communication, collaboration, critical thinking, and creativity in engaging with the various components of mathematics/statistics including data collection, data processing, data visualization, data analysis, data validation, and prediction with data.* See the appendix for a list of the habits of mind.

- Students should have access to mathematical action technology within and out of school to support their mathematical and statistical work in any high school mathematics course they choose to take.

  > *Technology can make a difference in mathematics teaching and learning (NCTM, 2018; Sacristan, 2021), enabling students to build conceptual understanding and engaging them in problems of real interest (Drijvers et al. 2010). Mathematical action technology—technology dedicated to mathematics (Dick & Hollebrands, 2011) - enables more students to have access to more mathematics, allows them to visualize mathematical ideas, and gives them agency to choose solution pathways that make sense to them. The technology can support the development of conceptual understanding, enable students to connect multiple representations and carry out mathematical work, give students space to concentrate on mathematical ideas and enable students to gain confidence and expand their horizons by easily making and testing conjectures (Burrill, 2023).*

- A high school data science course involves significant content knowledge and skills. A data science course is a valuable resource for students in learning how to appreciate and understand the world around them.

  > *People are likely to use statistical reasoning and thinking about data in their daily lives. All major corporations, organizations, and institutions have a need to analyze large amounts of data. As consumers and citizens, people need to make sense of the data they receive about their personal finances, health care, education, and politics. Data science can also be used to make informed decisions as people collectively attempt to address complex social issues and identify possible solutions.*

Appendix: Skills and Habits of Mind

A high school data science course merits mathematics credit if it builds the following skills and habits of mind.

Skills: How students engage cognitively with mathematics and data.

Students should do the following:
Use mathematics and data to make logical and informed decisions.
Interact with relevant contexts through rich and accessible data.
Reason deductively and inductively with data.
Formulate and test predictions based on finding, sorting, characterizing, and analyzing mathematical and statistical models.
Recognize which mathematical strategies and tools are efficient in a given data situation.
Develop flexible and creative problem solving through data-driven processes.
Visualize, model, and construct multiple representations for authentic and data-rich situations while making connections among representations.
Justify conclusions and critique the reasoning of others through data investigations.
Communicate effectively and precisely through a data lens.
Tackle ethical and social issues through data collection/consideration, data analysis, and communication of results.
Work independently as well as in teams to ask meaningful questions and make logical and data-informed decisions.

Habits of mind: The dispositions that shape a student's identity.

Students should develop the habits of mind that allow them to develop the following mindsets.
Be willing to be wrong in search of the truth.
Be open to challenging questions and being challenged.
Exhibit curiosity about the stories in data and mathematical relationships.
Develop a mindset for persistence, challenge, and for seeing failure as an opportunity to refine and elaborate.
Appreciate statistical/mathematical models as ways to answer questions and understand the underlying context of a problem or situation. Appreciate the meaningful attributes that data and mathematical models can show about a situation.
Believe that mathematics and statistics can be used meaningfully to make informed decisions.
Develop confidence to move from being data consumers to becoming data producers and analyzers.
Be willing to question, analyze, and challenge the accepted meaning of statistical and mathematical models.
Become risk takers while engaging with relevant models and data.
Recognize the importance of understanding risk and its role in informed decision making, knowing that every decision will have benefits and costs that need to be considered in making the decision.

References

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D. (2020). *Pre-K-12 guidelines for assessment and instruction in statistics education (GAISE) report II*. American Statistical Association and National Council of Teachers of Mathematics.

Brown, T., Franklin, C., & Sierra, L. (2021). What is risk know-how? *Significance*, *18*(6), 30–33.

Burrill, G., Funderburk, J., Byer, B., & Gorsuch, R. (2023). Using technology to explore the wage gap. *Mathematics Teacher: Learning and Teaching Mathematics PK–12. 116*(5), 378-386. https://doi.org/10.5951/MTLT.2022.0348

Domo. (n.d.). Data Never Sleeps 6. Retrieved from https://www.domo.com/learn/infographic/data-never-sleeps-6

Erickson, T. (2019). Awash with Data: An introduction to Data Science and data moves with CODAP. https://codap.xyz/awash/

Gould, R. (2021). Toward data-scientific thinking. *Teaching Statistics*, *43*(1), 11–22. https://doi.org/10.1111/test.12267

NCTM (2018). *Catalyzing Change in High School Mathematics*. National Council of Teachers of Mathematics.