



Finding Your Way into Data Science Education as a Science Teacher

Matthias Fischer, Cody Pritchard, Zhen Xu & Joshua Rosenberg

To cite this article: Matthias Fischer, Cody Pritchard, Zhen Xu & Joshua Rosenberg (2025) Finding Your Way into Data Science Education as a Science Teacher, *The Science Teacher*, 92:6, 49-55, DOI: [10.1080/00368555.2025.2558519](https://doi.org/10.1080/00368555.2025.2558519)

To link to this article: <https://doi.org/10.1080/00368555.2025.2558519>



Published online: 19 Nov 2025.



Submit your article to this journal [↗](#)



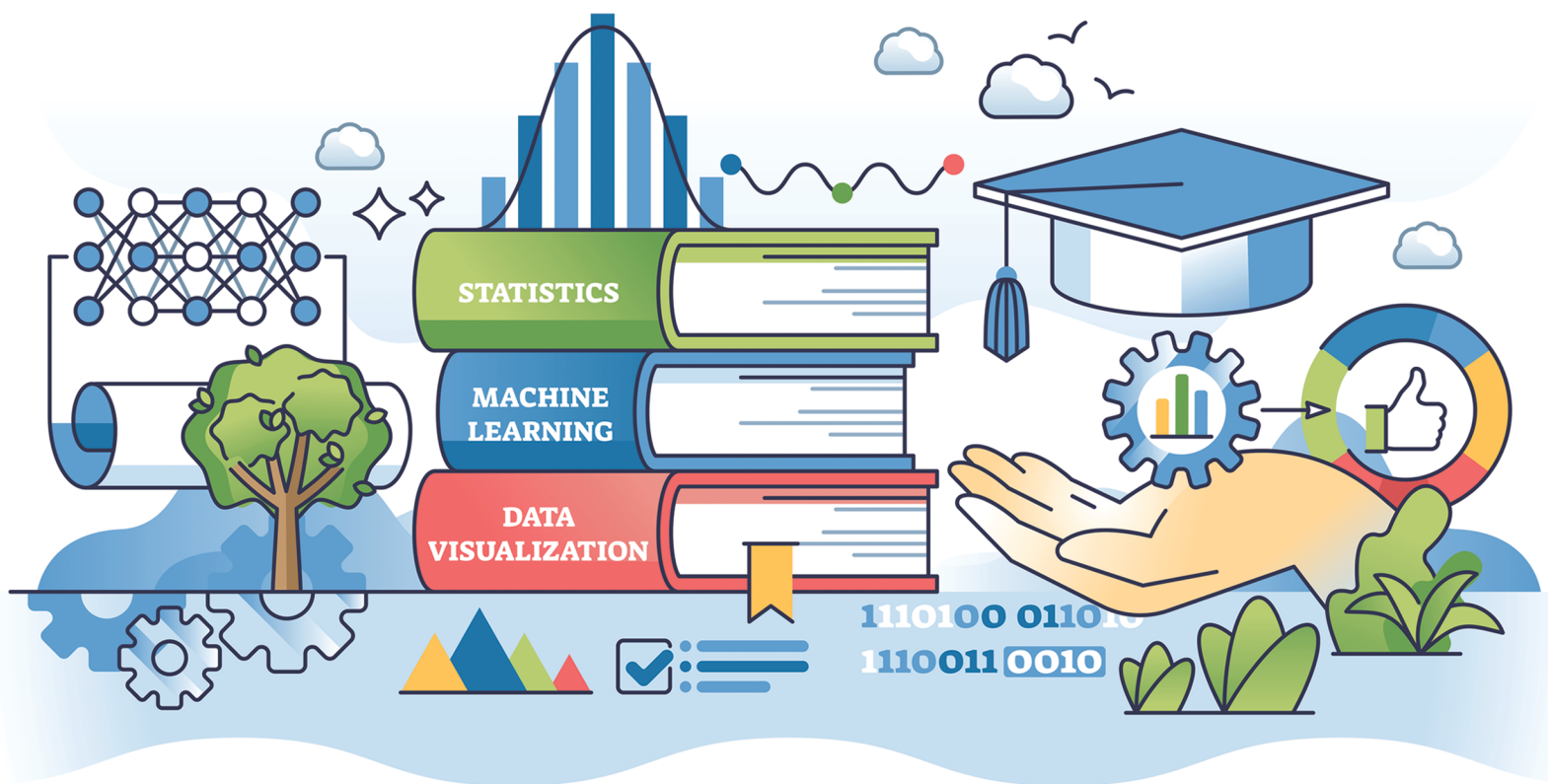
Article views: 37



View related articles [↗](#)



View Crossmark data [↗](#)



Finding Your Way into Data Science Education as a Science Teacher

MATTHIAS FISCHER , CODY PRITCHARD , ZHEN XU , AND JOSHUA ROSENBERG 

ABSTRACT

As data becomes an increasingly pervasive force in our daily lives, equipping students with data science-related skills and knowledge has never been more urgent. Consequently, data science is gradually being incorporated into school curricula. However, science teachers not only express concerns about their preparedness to teach data science but often feel hesitant to approach these topics altogether. With this paper, we aim to support those science teachers without backgrounds in data science by addressing three major challenges: (1) the lack of training to teach data science, (2) the difficulty of designing relevant and engaging data science lessons, and (3) the challenge of teaching data science with limited technological resources. To achieve this, we formed a unique collaboration between a science teacher with no formal training to teach data science and three data science education experts. Through this approach, we aim to translate the experts' knowledge into practical support, enabling any science teacher to take their first steps in supporting their students in developing data science-related skills and knowledge, even when starting from square one.

Keywords: Data science education; science teacher; professional development

As data becomes an increasingly pervasive force in our daily lives, equipping students with data science-related skills and knowledge has never been more urgent. Thus, educational leaders and researchers are advocating for the inclusion of data science in K-12 curricula (National Center for Education Research 2021; Salman 2024). Furthermore, the *Framework for K-12 Science Education* (National Research Council 2012) and the *Next Generation Science Standards* (NGSS Lead States 2013)—as well as the standards for mathematical practice in the *Common Core State Standards*—emphasize integrating data analysis and interpretation as a core science and engineering practice, along with other practices that are closely related to data science, in particular using mathematics and computational thinking and developing and using models.

In addition to these practices, the crosscutting concepts in the *Framework* also connect with student sensemaking in learning data science. In particular, the crosscutting concepts of patterns, scale, proportion, and quantity are highly relevant. One goal in analyzing data is to help students identify and interpret patterns across contexts, while understanding scale, proportion and quantity is essential when working with numerical data.

The challenges involved in this integration are often understated. Science teachers not only express concerns about their preparedness to teach data science but often feel hesitant to approach these topics at all (Miller 2022; Shreiner and Dykes 2021). Recognizing how daunting this task can seem, we decided to step back and address fundamental questions that educators may have. Where can teachers find helpful resources to learn how to effectively incorporate data science into their classrooms? How can they create relevant and engaging data science lessons? And, for those with technology and/or funding constraints, what accessible solutions exist to help them teach data science?

With this paper, we aim to assist science teachers without a background in teaching data science by addressing these challenges. To this end, we formed a unique and fruitful partnership between three experts in data science education and a science teacher who has received no formal training in data science education. To help the novice get started with data science education, the three experts shared their knowledge in a focus group interview, which the novice moderated. Thanks to the discursive nature of the focus group interview (and the previous preparatory meetings), the article's findings, summarized by the novice, represent a co-construction of the experts' expertise and the novice's growing understanding — a collaboration from which both parties benefited. Through this approach, we hope to translate the experts' expertise into practical support that enables novices to take their first steps in teaching data science after being sent back to square one.

Challenges and recommendations

Before addressing the three key challenges outlined above, we want to establish a common understanding of what data science and data science education are. Data science is a domain that

involves mathematics and statistics, coding and computer science, and subject-area expertise (such as in particular areas of science, like in the life sciences). Thus, it has many connections to the science and engineering practices that are a core part of the NGSS — especially the practices of analyzing and interpreting data, using mathematics and computational thinking, and planning and carrying out investigations; in this way, data science can serve as a vehicle to support students' engagement in these (and other) science practices. Following from this definition, we define data science education using the definition from *DataScience4Everyone* (n.d.):

Data science education refers to instruction focused on the systematic processes, analytical techniques, computational methods, and utilization of appropriate technologies to gain knowledge and insight from data. In the context of K-12 education, it equips students with the introductory skills and problem-solving that are necessary to collect or consider, analyze, model, interpret, and communicate data to answer investigative questions.

First challenge: Teaching data science without any data science-related training

The definition indicates that data science education is an interdisciplinary field and not limited to specific subjects. Consequently, teachers can leverage their scientific knowledge when teaching data science. In data science instruction, as presented in this article, students frequently engage in projects that involve addressing questions using either self-collected or teacher-provided datasets. The technologies used for data collection, as well as the mathematical and computational methods employed for data analysis, are often also found in the evaluation of experiments in science education — a skill that every science teacher possesses. Furthermore, several resources are available to help them start teaching data science, requiring little prior knowledge. Thus, a lack of knowledge in statistics or coding is not a barrier to taking the first steps in data science.

If teachers feel unsure of what data is, they can start their journey by watching introductory videos. *DataScience4Everyone* offers several short videos that provide an initial understanding of data science and teaching data science. Another starting point is journal articles that describe classroom activities. *Science and Children*, *Science Scope*, and *The Science Teacher* offer several articles worth exploring — e.g., Garner and Rosenberg (2023). Additionally, we recommend various writings that foster an understanding of teaching data science (see Table 1).

Another way to familiarize yourself with data science is to take on the role of a student. There are numerous high-quality, free lessons and curricula available online waiting to be used for this purpose. Teachers can use these resources both in the classroom and for self-directed learning. Table 2 entails various websites offering such activities. If teachers want to dive deeper and acquire specific competencies—such as coding languages—*DataScience4Everyone* offers an overview of different “crash courses.” Other valuable

TABLE 1

Important white papers and reports about data science education.

- NCTM, NSTA, ASA, NCSS, and CSTA. 2024. *Data Science. A Joint Position of NCTM, NSTA, ASA, NCSS, and CSTA.* https://static.nsta.org/pdfs/PositionStatement_JointDataScience.pdf
- Lee, Victor R., and Michelle H. Wilkerson. 2018. *Data use by middle and secondary students in the digital age: A status report and future prospects.* Washington, D.C.: National Academy of Sciences Engineering, and Medicine, Board on Science Education. Committee on Science Investigations and Engineering Design for Grades 6–12. https://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=1635&context=itls_facpub
- Sukol, S. (2024). *State of the Field: Data Science and Data Literacy Education in US K-12.* DataScience4Everyone. https://www.datascience4everyone.org/_files/ugd/0a9d2b_f183f7139980484a9816319a99393bc9.pdf

TABLE 2

Overview of online available free resources to teach data science.

Websites	Type of resource				
	Introductory videos	Preplanned lessons and curricula	Professional development opportunities for data science-related competencies	Conferences & (recorded) webinars	Data sets
DataScience4Everyone	X	X	X	X	X
CAUSE		X		X	X
Youcubed		X			X
The Concord Consortium		X		X	X
CODAP		X			X
Data Nuggets		X			X
Kaggle					X
Data Classroom					X

sources for professional development are conferences and (recorded) webinars. We have included several suggestions in Table 2.

After developing an understanding of data science education, one may wonder what to keep in mind when preparing and instructing a data science class. During the focus group meeting, three key points emerged regarding this question:

- **Context matters:** Start by finding a question that you want to answer through collecting and analyzing data. This question sets the context for the lesson, especially if it sparks your students' curiosity and/or if they find the question meaningful.
- **Foster an open and flexible understanding of the nature of data:** Give students the opportunity to discover the versatile

nature of data. Data is not limited to numbers. There are countless possibilities for collecting data.

- **Allow time and space for your students to go in a new direction:** Give students the freedom to lead their own inquiries. Allow them to discover their own approach to exploring patterns, visualizing data, and using their own language before formally introducing scientific terms, the underlying mathematics, and data methods. Students could do this individually or in small groups.

At this point, teachers might wonder how they could create a data science lesson without getting lost in the lack of structure that an open approach to data science education entails. Three points to keep in mind emerged from the focus group interview:

1. *Large data sets can be overwhelming—so start small(er)*: Start by considering the variables available and the questions that can be answered using those variables. This can be an open inquiry. What do you notice or wonder about the data set? Grouping students to explore a larger data set together can also be a helpful strategy.
2. *Allow students to explore visualizations*: Once students have identified their questions and corresponding variables, encourage them to explore visualizations using the available software. This part will require some light direct instruction to familiarize students with the program/platform.
3. *Keep students focused on their questions*: As students work with their data sets and create visualizations, it is essential for teachers to redirect them back to the initial questions. Does this visualization provide insight or an answer to the questions? Is it visually clear and effective?

While this advice provides a good starting point, there is, obviously, more to keep in mind. As a next step, teachers could explore the writings (see [Table 3](#)) that address various questions, such as: How should I structure data science lessons? How can I authentically integrate science education with data science education? How should I approach and use existing data sets? How can I formatively assess the data-related practices my students are using in their lessons? What prompts can I use to engage my students in open-ended inquiry with data and data visualizations? In these writings, teachers will also encounter further inspiring teaching examples.

Second challenge: Designing relevant and engaging data science lessons

Teachers may wonder how their lessons can be relevant (and engaging) to their students. Three points from our group meeting spoke about ways to achieve this, as follows:

- *Avoid making assumptions about your students' interests*: Don't be afraid to ask them directly. Interest surveys are helpful to collect data on common interests.
- *Find data related to students' interests and community*: See if you/your students can collect data related to questions emerging from their interests/community. The aforementioned open understanding of the nature of data will help your students to understand that data can be collected on anything.
- *Teach abstract concepts with concrete examples*: If you want to teach abstract concepts like outliers, measures of center, and different visualizations, then using data from your classroom or school is helpful. While this may not always seem interesting, it is concrete and easy to discuss and can eventually lead to more abstract thinking.

This advice once again highlights the importance of using contexts that capture the students' interest. Research tells us that this can only be achieved if teachers consider and incorporate

cultural and sociopolitical dimensions of their lessons (Lee, Wilkerson, and Lanouette 2021). Therefore, selecting a topic that is relevant to students at their stage of life and within their communities is essential. For example, those teaching about ecosystem processes can choose a topic (and select data) from or about their local ecosystem, such as water quality or biodiversity data. Those teaching about chemistry could choose local water quality and corresponding data. And those teaching the physical sciences could use examples related to climate science and local energy production and consumption.

Another example can be seen in an activity explained by Cody. His high school students collected data on college tuition over the past 15 years for colleges they were interested in attending. Using scatter plots and linear regression models, students predicted the cost of tuition a few years from now, when they expect to be in college. However, it is important to keep in mind that while this lesson might have been relevant to Cody's students, it may not be relevant to other groups of students. Ultimately, it all comes down to knowing your students and allowing them to select topics that interest them.

A key aspect of teaching data science is providing students with opportunities to analyze and interpret data visualizations. There are several resources available to teachers for engaging students in this process. *The New York Times*, in collaboration with *The Learning Network*, features a "What's Going On in This Graph" resource with hundreds of data visualizations and specific questions to engage students in this process daily. Furthermore, the article by Wilkerson et al. (2025) provides a toolkit and a focused discussion framework for engaging students in these investigations.

The experts also acknowledged the challenge of finding meaningful topics with sufficient data that allow learning specific skills. One option is to use free online lessons. Some of these lessons are specifically designed to capture common interests, for example, "Coral Bleaching and Climate Change" by *Data Nuggets*. Other options are websites that offer free data sets. [Table 2](#) provides resources to help teachers find data sets aligned with their students' interests and the competencies they want to teach. Lastly, teachers could create their own data set, for example, by asking *ChatGPT*.

Third challenge: Teaching data science without technological resources

One problem that data science educators often face is a lack of technology, including hardware and software. Therefore, we also discussed how a data science lesson with limited access to technology would look.

In the first scenario, only tablets, laptops, smartphones, and Wi-Fi are available. If a teacher is, for example, planning to collect data with students, there are various free tools available online. Instead of purchasing new sensors, smartphone sensors can be used through different apps. The experts as well as the novice named several apps they use regularly to collect data:

TABLE 3

Literature to answer specific questions regarding data science education.

Specific questions	Literature suggestion
How should I structure a data science lesson?	Lee, Hollylynne S., Gemma F. Mojica, Emily Thrasher, and Peter Baumgartner. 2022. "Investigating Data Like a Data Scientist: Key Practices and Processes." <i>Statistics Education Research Journal</i> , 21 (2): 1–23. https://doi.org/10.52041/serj.v21i2.41
How can I authentically link science education and data science education?	Hammett, Amy, and Chad Dorsey. 2020. "Messy Data, Real Science." <i>The Science Teacher</i> 87 (8): 40–48. https://doi.org/10.1080/00368555.2020.12293530 Kjelvik, Melissa K. and Elizabeth H. Schultheis. 2019. "Getting Messy with Authentic Data: Exploring the Potential of Using Data from Scientific Research to Support Student Data Literacy." <i>CBE - Life Sciences Education</i> , 18 (2): 1–8. https://doi.org/10.1187/cbe.18-02-0023
How do I approach and use already existing data sets? What do I need to consider before using a data set for my class?	Sagrans, Jacob, Janice Mokros, Christine Voyer, and Meggie Harvey. 2022. "Data Science Meets Science Teaching." <i>The Science Teacher</i> 89 (3): 64–69. https://doi.org/10.1080/00368555.2022.12293671 Weintraub, Naomi, and Kristin Hunter-Thomson. 2022. "Teacher's Toolkit: Leverage the Data You Have: Increase Creativity and Confidence in Science." <i>Science Scope</i> 45 (5): 14–20. https://doi.org/10.1080/08872376.2022.12291475
What is the role of variability in data science activities?	Hunter-Thomson, Kristin. 2022. "Interdisciplinary Ideas: Why Is Variability Worth the Teaching Challenge?" <i>Science Scope</i> 45 (3): 8–13. https://doi.org/10.1080/08872376.2022.12291452
How can I formatively assess the data-related practices my students are using in their lessons?	Bargagliotti, Anna, Christine Franklin, Pip Arnold, Rob Gould, Sheri Johnson, Leticia Perez, and Denise A. Spangler. 2020. <i>Pre-K–12 Guidelines for assessment and instruction in statistics education II (GAISE II)</i> . Alexandria, VA: American Statistical Association and National Council of Teachers of Mathematics. https://www.amstat.org/docs/default-source/amstat-documents/gaiseiiprek-12_full.pdf Thrasher, E., Hollylynne Lee, Bruce Graham, Matthew Grossman, and Gemma Mojica. 2024. "Making sense of data visualizations: A toolkit for supporting student discussions." <i>Statistics Teacher</i> , Fall 2024 issue. https://www.statisticsteacher.org/files/2024/12/ThrasherMakingSenseofDataVisualizations01062025.pdf
What prompts can I use to engage my students in open-ended inquiry with data and data visualizations?	Wilkerson, Michelle H., John Kim, Hollylynne S. Lee, David J. Stokes, and Matthew Ferrell. 2025. "How Teachers Envision Using Data Visualization Discussion Tasks in Classroom Instruction." <i>International Journal of Science and Mathematics Education</i> , 1–35. https://doi.org/10.1007/s10763-024-10521-y The New York Times, and The Learning Network. n.d. "What's going on in this graph? Graphs, maps and charts from The Times—and an invitation to students to discuss them live." The New York Times. Accessed July 9, 2025. https://www.nytimes.com/column/whats-going-on-in-this-graph

- *The Phyphox mobile device app*: Depending on your smart-phone, you can collect data on acceleration, light intensity, magnetic fields, GPS, and more. The data collected is transferred into a spreadsheet.
- *Viana.NET (Computer)/Viana (iOS)*: With this software, you can analyze filmed movements (e.g., the free fall of an object) and obtain data on the distance covered or the current speed at certain points in time.

- *The iNaturalist/Seek apps*: You can take pictures of animals or plants, and the app will help identify what you see through image recognition. By uploading the pictures to the underlying database, you can contribute to the citizen science project (see Garner and Rosenberg 2023).

Additionally, whether teachers are using an existing data set or collecting new data, there are various online analysis tools

available free of charge. The experts highlighted the following tools, although this list is not exhaustive:

- *CODAP*: A web-based app for analyzing and visualizing data.
- *Dataflow*: A web-based app for programming, data processing, and data graphing.
- *JASP*: An easy-to-use statistical software for performing statistical tests on your data.
- *Google Sheets*: A web-based spreadsheet program to format, organize, and calculate.

Usually, there are how-to guides on the websites of these analysis tools. Alternatively, tutorials are often available on online video platforms. These options will help you to get to know and use these tools. Further suggestions for analysis tools and their use can be found in Pimentel et al.'s (2022) and Rosenberg et al.'s (2020) work. Additionally, *CAUSE* and *The Concord Consortium* offer a resource library with analysis tools.

In the second scenario, the experts were asked for their advice on “unplugged” data science activities. The unanimous opinion was that this is no easy task. However, they pointed out that data can also be collected using paper and pencil – again, building on an open understanding of the nature of data. For example, teachers (or students) can collect data on different features in their classroom, such as height, shoe size, eye color, etc. Then, students can find their own way of representing and interpreting the data before being introduced to more traditional methods of visualization. As one might have noticed, there is little difference between an “unplugged” data science activity and a statistics class activity. Therefore, reviewing statistics education textbooks might also be helpful.

Further considerations on the safety and inclusivity of data science activities

Most data science activities do not inherently pose risks to students. However, several important aspects should be considered. For instance, when data is collected through scientific experiments or place-based outdoor experiences, the same safety precautions applied in science classes must be followed. Additionally, teachers should evaluate the extent to which issues such as data privacy are considered and addressed in the data collection process. It may be necessary for teachers to provide brief guidance on these or related topics, such as appropriate online behavior.

Finally, we would like to share some thoughts on how data science activities can be designed to promote inclusivity in the classroom. A key first step is allowing students to develop their own research questions. This approach creates space for their lives, communities, and interests to be reflected in the (data science) classroom. Building on these questions and the resulting projects, self-directed learning can take place, with teachers acting as learning companions. One advantage of such projects is

the naturally occurring differentiation in complexity, as the questions can be addressed in various ways and through different analytical methods.

Data science activities can also be made accessible for English Language Learners using tools such as *ChatGPT*, *DeepL*, or *Google Lens*. These tools can help translate assignments, data collected or provided in students' native languages, and even the final project outcomes. Additionally, we emphasize that differentiation strategies commonly used in science education—such as those outlined by Whitworth and Sneed (2024)—are equally relevant and applicable to data science instruction.

Conclusion

In light of three critical challenges science teachers might face when integrating data science into their classrooms, we support and encourage science teachers without a background in teaching data science to take these on. To do so, we formed a partnership between three experts and a science teacher without any background in data science to co-construct knowledge and to collate the specific resources, strategies, and tools included in this paper. We would like to point out the encouraging fact that the experts benefited from the science education-specific knowledge of the novice; for example, they learned about new apps for collecting data. Our goal for this article is achieved if teachers, especially those who have never taught data science, feel empowered to take on the challenge of finding their own way into data science education. Our hopes are exceeded if these teachers do not feel like they are being sent back to square one when preparing their first data science lesson.

ACKNOWLEDGEMENTS

We would like to thank the Hanns-Seidel Foundation for funding this research from funds of the Federal Ministry of Education and Research (BMBF) in Germany. In the preparation of this manuscript, ChatGPT (Version 4o mini) was used to improve and smoothen the language.

DISCLOSURE STATEMENT

Cody Pritchard has done contractual work for the non-profit *DataScience4Everyone*.

REFERENCES

- DataScience4Everyone. DataScience4Everyone. n.d. “Frequently Asked Questions.” Accessed January 27, 2025. <https://www.datascience4everyone.org/faqs>
- Garner, Amanda V., and Joshua Rosenberg. 2023. “Utilizing iNaturalist to Support Place-Based Learning and Data Analysis.” *Science Scope* 46 (7): 54–60. <https://doi.org/10.1080/08872376.2023.12315907>.
- Lee, Victor R., Michelle H. Wilkerson, and Kathryn Lanouette. 2021. “A Call for a Humanistic Stance Toward K–12 Data Science Education.” *Educational Researcher* 50 (9): 664–672. <https://doi.org/10.3102/0013189X211048810>.
- Miller, Katherine M. 2022. “Developing Pedagogical Content Knowledge for STEM Integration Through Data Literacy: A Case Study of High School Science Teachers.” PhD diss., University of Pennsylvania. <https://repository.upenn.edu/entities/publication/a8e4097b-ca87-48d9-8400-4d16e06a9fb5>.

- National Center for Education Research. 2021. *Catalyzing a New Field: Data Science Education in K-12. Technical Working Group Summary*. Washington, D.C: Institute of Education Sciences. <https://ies.ed.gov/ncser/2025/01/catalyzing-new-field-data-science-education-k-12>.
- National Research Council. 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/13165>.
- NGSS Lead States. 2013. *Next Generation Science Standards: For States, By States*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/18290>.
- Pimentel, Daniel R., Nicholas J. Horton, and Michelle H. Wilkerson. 2022. "Tools to Support Data Analysis and Data Science in K-12 Education." <https://www.nationalacademies.org/documents/embed/link/LF2255DA3DD1C41C0A42D3BEF0989ACAEC3053A6A9B/file/DD2EB10E0294DBA9ECC3AB74BD35576C5FFB34D76352?noSaveAs=1>
- Rosenberg, Joshua, Alex Edwards, and Bodong Chen. 2020. "Getting Messy with Data." *The Science Teacher* 87 (5): 30–34. https://doi.org/10.2505/4/tst20_087_05_30.
- Salman, Javeria. 2024. "Calculating the Value of Data Science Classes." The Hechinger Report, March 14. <https://hechingerreport.org/calculating-the-value-of-data-science-classes/>
- Shreiner, Tamara L., and Bradford M. Dykes. 2021. "Visualizing the Teaching of Data Visualizations in Social Studies: A Study of Teachers' Data Literacy Practices, Beliefs, and Knowledge." *Theory & Research in Social Education* 49 (2): 262–306. <https://doi.org/10.1080/00933104.2020.1850382>.
- Whitworth, Brooke A., and Amy Sneed. 2024. "Powerful Practices for the Differentiated Science Classroom." *The Science Teacher* 91 (6): 47–56. <https://doi.org/10.1080/00368555.2024.2404231>.
- Wilkerson, Michelle H., John Kim, Hollylynn S. Lee, David J. Stokes, and Matthew Ferrell. 2025. "How Teachers Envision Using Data Visualization Discussion Tasks in Classroom Instruction." *International Journal of Science and Mathematics Education*: 1–35. <https://doi.org/10.1007/s10763-024-10521-y>.

© 2025 National Science Teaching Association

Matthias Fischer (fim@englisches-institut.eu) (ORCID: 0000-0002-3065-9231) recently finished his PhD in Science Education at the University of Education Heidelberg. His research interests include school dropout and homelessness in the classroom. In January, he started his teacher training in Germany to become a high school teacher of physics and mathematics. **Cody Pritchard** (ORCID: 0009-0006-9080-8573) is a 2nd-year PhD student at the University of Tennessee, Knoxville, studying data science education, student privacy and digital rights, and curriculum development. **Zhen Xu** (ORCID: 0000-0002-5080-7684) is a Postdoc Research Associate at the University of Tennessee, Knoxville. **Joshua Rosenberg** (jrosenb8@utk.edu) (ORCID: 0000-0003-2170-0447) is an Associate Professor of STEM education and a Faculty Fellow for the Center for Enhancing Education in Mathematics and Sciences at the University of Tennessee, Knoxville. He studies the use of data in education, especially in STEM education contexts.