Mathematics and Society
University of Isfahan
http://math-sci.ui.ac.ir

Journal of Mathematics and Society
ISSN (print): 2345-6493, ISSN (on-line): 2345-6507
Vol. 10 No. 3 (2025), pp. 99-117.

http://www.ui.ac.ir

IMPLEMENTING DATA SCIENCE EDUCATION IN SCHOOLS: OPPORTUNITIES AND CHALLENGES

REZA MOETI, ABOLFAZL RAFIEPOUR* AND MOHAMMAD REZA FADAEE

Abstract. The vast volume and storage of data in today's world have significantly increased public access to data. However, this trend poses challenges, such as the misinterpretation of data, which complicates decision-making in all fields that rely on data. On the other hand, the demand for skilled professionals in this field is exceedingly high. Data scientists and data analysts hold some of the highest-paying jobs globally. Furthermore, students are also part of this society, and their need to analyze data, both as citizens and as the future workforce, is a pressing issue. This need has drawn attention to integrating data science education into schools. The aim of this study is to explore the fundamental challenges of teaching data science in schools, focusing on curriculum design, data utilization, and technology. The definition of data science plays a crucial role in examining these challenges. A key aspect of data science is discovering patterns and generating knowledge using data, which requires proficiency in disciplines such as mathematics, statistics, and computer science. Considering this definition, curriculum changes could particularly target these three subjects: mathematics, statistics, and computer science. Among these, the most significant changes could be implemented in school mathematics.

Keywords: Data Science Education, School Mathematics, Curriculum, Technology, Data

Article Type: Research Paper.

Communicated by Saeid Maghsoudi.

*Corresponding author.

Received: 16-10-2024, Accepted: 05-01-2025, Published Online: 09-07-2025.

Cite this article: R. Moeti, A. Rafiepour and M. R. Fadaee,Implementing Data Science Education in Schools: Opportunities and Challenges, *Journal of Mathematics and Society*, **10** no. 3 (2025) 99–117. http://dx.doi.org/10.22108/msci.2025.143079.1704

.

1. Introduction

Today, we are experiencing a major revolution in the field of data [12], with the enormous volume of data, including Werner Herzog's analogy about the global data flow, becoming both a challenge and an opportunity in the era of big data and data science [38]. The proper understanding of data and the ability to interpret it has now become a necessity, as decision-making in business, politics, and society relies on data-driven evidence [11]. Citizens must understand the role of data in their lives and acquire skills such as data collection, analysis, and visualization [5]. Data science has been recognized as one of the top professions globally, and the demand for specialists in this field is rapidly growing. This situation highlights the importance of teaching data science in schools. Data science education at the K-12 level should enhance essential data-related skills and prepare students to play an effective role in society and critically engage with data [6, 39]. Despite the high significance of data science in education, research on how to integrate it into curricula is still ongoing. The goal is to provide a comprehensive and accessible program for all students, enabling them to become effective users and creators of data. In this regard, curriculum planning must address technological challenges, data-driven topics, and the foundations of data science [2, 8, 28]. This study focuses on data science education in schools, examining various definitions of the field and addressing the key challenges of teaching data science in schools, including curriculum, data, and technology.

2. Main Results

In 2001, William Cleveland introduced the term "data science," describing it as a combination of statistics and large-scale computing [7]. Following the economic recovery in the United States in 2011 and the launch of related programs, this field expanded further [39]. Data science is identified as an interdisciplinary field that integrates concepts and methods from statistics, computer science, mathematics, and information science [10, 32, 34]. Some definitions emphasize the analysis of data and discovering useful information using mathematical and statistical methods,[18] while others describe data science as a systematic process for collecting, organizing, and analyzing data [24]. Loukides' definition, widely regarded among researchers, describes data science as the use of computer-based systems to analyze vast amounts of data and extract actionable knowledge [27]. Overall, data science is a broad field dedicated to data collection, processing, and analysis to generate knowledge. The significance of defining data science raises the question of what priorities data science education in schools should focus on. Research suggests that school-level data science education should start with data and expand by leveraging modern technologies to enhance access to data.

Data science has had a significant impact on contemporary sciences, introducing new ways of thinking about data and computation. Many universities have developed educational programs related to data science, and proposals are now being made to incorporate it into elementary and secondary

http://dx.doi.org/10.22108/msci.2025.143079.1704

education [22]. Teaching data science equips students with skills such as calculating simple statistics, selecting appropriate visualizations, and interpreting data results [21]. However, many high school students struggle with analyzing data and making data-driven decisions, often focusing on details while overlooking broader patterns and trends [20]. According to Finzer, if the curriculum from at the K-12 level is not revised, society will face a shortage of data scientists compared to its future needs [10]. Therefore, teaching data science in elementary and secondary education is essential.

The integration of data science into school education is expanding as part of the global movement "Data Science for All", aiming to prepare students to use data science skills in life and the workforce [25, 37]. These efforts emphasize revising curricula to make data science a core component of the learning process. Research shows that the best data science learning experiences occur when students have the opportunity to engage with real-world data relevant to their lives [31]. In this context, curricula are evolving. For example, the state of California has introduced data science courses as an alternative to traditional mathematics courses in high schools[37]. Some states have developed certification programs for teachers, enabling them to incorporate data science content into other subjects or teach it as a standalone course [20]. International projects such as the "Global Data Science Project in Schools" have provided frameworks for teaching data science to students and preparing teachers for this role. These projects highlight the data learning cycle as the central aspect of data science and strive to present data science in an accessible and comprehensible manner to a broad range of students [15].

The American Statistical Association, in its latest guidelines, has emphasized the importance of data science and updated curricula from k-12 to address the needs of a data-driven world [2]. Alongside these efforts, organizations and platforms like "YouCubed" have created online courses for teachers to familiarize them with data science concepts and provide them with essential resources [20]. Ultimately, the development of data science curricula in schools can enhance students' capabilities and prepare them for the challenges of the modern world. Technology and data serve as two foundational pillars of these programs, with countries striving to use them to transform students into informed and data-literate citizens.

Data, as an integral part of our daily lives, is being generated on a massive and rapidly increasing scale. It is predicted that the volume of digital data will double by the end of 2024. This growth stems from various sources, such as Internet of Things (IoT) devices, social media platforms, and websites, which contribute to the production of big data and create new business opportunities [31, 13]. Authentic data, consisting of real quantitative and qualitative information from real-life phenomena, plays a critical role in data science education. Using authentic data, as opposed to manipulated data, provides a more genuine learning experience [19]. Although teachers face challenges such as time constraints and the complexity of data analysis, working with real data can enhance students' motivation

[33]. Multidimensional data, such as images and videos, offer opportunities for deeper understanding aligned with the needs of data science [37]. The American Statistical Association emphasizes the development of data-handling skills across different educational levels. At the initial stages, students should understand the concept of data and use it to answer questions. At more advanced levels, they perform complex analyses and recognize the importance of ethical data management [2]. However, big data remains overlooked in many school curricula. Since such data is primarily generated electronically, integrating technology education and data analysis into the curriculum is essential.

Given technological advancements, students should become familiar with digital tools for performing statistical and scientific analyses. Technology should be explicitly integrated into the curriculum of mathematics and statistics so that students know which concepts should be learned manually and which should be learned using technology. These decisions should be based on developmental research in science education and learning mathematics. In today's digital society, where vast amounts of data are generated, teachers must use various tools to process and analyze data. These tools include statistical software such as R, Python, and tools like CODAP and TinkerPlots, which allow students to analyze data without needing programming skills [30]. These tools are suitable for teaching fundamental data science concepts such as filtering data, grouping, summarizing, calculating, and merging data. In educational examples, data filtering is used to analyze patterns and uncover hidden patterns in data. [9]. The use of appropriate technologies in teaching data science helps students acquire the essential skills needed to understand and analyze data, enabling them to apply these skills in real-world problems.

3. **Conclusions**

Data science is universal and its learning is recommended for all citizens. It can be considered a necessity for every individual in society or in any job position. In this article, we reviewed research conducted in the field of school-level data science education and examined the key challenges in this area. Most research on school-level data science education emphasizes curriculum, data, and technology. Data science has gained significant attention in the curriculum of universities, standards organizations, and curriculum planners at both higher education and school levels. The Guidelines for Assessment and Instruction in Statistics Education (GAISEII) is one of the most prominent documents frequently cited in research. This document, with a particular focus on data science, has been made available to curriculum planners and teachers. Some countries, based on their curricular conditions, have begun developing various curricula in data science. These efforts, aligned with the diverse definitions of data science, have attracted attention within school mathematics. Understanding how to teach data science skills remains a challenge for specialists and researchers in this field. Some studies have recommended introducing data science education from elementary school and even preschool.

http://dx.doi.org/10.22108/msci.2025.143079.1704

Thus, both teachers and students require tools and training for the effective use of data. Selecting appropriate data and technology is also a significant challenge in this field. Almost all studies agree on the need to use real and reliable data. Moreover, both programming and non-programming tools are considered based on educational levels. Research indicates that data science can play an important role in demonstrating the applications of mathematics in society.

References

[1] I. O. Adisa, D. Herro, O. Abimbade and G. Arastoopour Irgens, Engaging elementary students in data science practices, *Information and Learning Sciences*, **125** no. 7/8 (2024) 513–544.

[2] A. Bargagliotti, C. Franklin, P. Arnold, R. Gould, S. Johnson, L. Perez and D. Spangler, *Pre-K-12 Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report II: A Framework for Statistics and Data Science Education*, American Statistical Association and National Council of Teachers of Mathematics, 2020.

[3] B. Baumer, A data science course for undergraduates: thinking with data, *Amer. Statist.*, **69** no. 4 (2015) 334–342.

[4] R. Biehler, D. Frischemeier, C. Reading and j. M. Shaughnessy, Reasoning about data, In: D. Ben-Zvi, K.Makar and J. Garfield, editors, *International handbook of research in statistics education*, Cham, Springer, (2017) 139–192.

[5] W. Burr, F. Chevalier, C. Collins, A. L. Gibbs, R. Ng and C. J. Wild, Computational skills by stealth in introductory data science teaching, *Teach Stat.*, **43** (2021) S34–S51.

[6] G. Cattaneo, G. Micheletti, M. Glennon, C. La Croce and C. Mitta, *The European data market monitoring tool: Key facts & figures, first policy conclusions, data landscape and quantified stories: d2.9 Final study report*, European Commission, B-1049 Brussels, 2020.

[7] W. S. Cleveland, Data science: An action plan for expanding the technical areas of the field of statistics, *International Statistics Review*, **69** no. 1 (2001) 21–26.

[8] J. Engel, Statistical literacy for active citizenship: a call for data science education, *Statistics Education Research Journal*, **16** no. 1 (2017) 44–49.

[9] T. Erickson, M. Wilkerson, W. Finzer and F. Reichsman, Data Moves, *Technology Innovations in Statistics Education*, **12** no 1 (2019).

[10] W. Finzer, The data science education dilemma, *Technology Innovations in Statistics Education*, **7** no. 2 (2013) 1–9.

[11] D. Frischemeier, R. Biehler, S. Podworny, and L. Budde, A first introduction to data science education in secondary schools: Teaching and learning about data exploration with CODAP using survey data, *Teach. Stat.*, **43** (2021) S182–S189.

[12] S. Gutta, Data science: The 5 V's of big data, (2020), Available at: https://medium.com/analytics-vidhya/the-5-vsof-big-data-2758bfcc51d (accessed 28 July 2024).

[13] S. Hedges and K. Given, Addressing confirmation bias in middle school data science education, *Found. Data Sci.*, **5** no. 2 (2023) 223–243.

[14] K. Holstein, B. M. McLaren and V. Aleven, Co-designing a real-time classroom orchestration tool to support teacher – AI complementarity, *Journal of Learning Analytics*, **6** no. 2 (2019) 27–52.

[15] IDSSP Curriculum Team. *Curriculum Frameworks for Introductory Data Science*, 2019.

[16] C. E. Konold, R. Coulter and A. Feldman, Engaging students with data, *Learning & Leading with Technology*, **28** no. 3 (2000) 50–55.

[17] D. Kadijevich, *Interactive displays: Use of interactive charts and dashboards in education*, In: A. Tatnall (Ed.), Encyclopedia of education and information technologies, Cham, Switzerland, Springer, 2019.

[18] D. Kadijevich and M. Stephens, Modern statistical literacy, data science, dashboards, and automated analytics and its applications, *Teaching of Mathematics*, **23** no. 1 (2020) 71–80.

[19] M. K. Kjelvik and E. H. Schultheis, Getting messy with authentic data: Exploring the potential of using data from scientific research to support student data literacy, *CBE Life Sci. Educ.*, **18** (2019) 1–8.

[20] T. Lamar and J. Boaler, The importance and emergence of K-12 data science, *Phi Delta Kappan*, **103** (2021) 49–53.

[21] N. Lazar and C. Franklin, The Big Picture: Preparing Students for a Data-centric World, *CHANCE*, **28** (2015) 43–45.

[22] V. R. Lee and V. Delaney, Identifying the Content, Lesson Structure, and Data Use Within Pre-collegiate Data Science Curricula, *J. Sci. Educ. Techno.* **31** (2022) 81–98.

[23] H. S. Lee and T. Harrison, Trends in Teaching Advanced Placement Statistics: Results from a National Survey, *Journal of Statistics and Data Science Education*, **29** no.3 (2021) 317–327.

[24] V. R. Lee, D. R. Pimentel, R. Bhargava, and C. D'Ignazio, Taking data feminism to school: A synthesis and review of pre-collegiate data science education projects, *British Journal of Educational Technology*, **53** no. 5 (2022) 1096–1113.

[25] V. R. Lee, M. H. Wilkerson and K. Lanouette, A call for a humanistic stance toward K–12 data science education, *Educational Researcher*, **50** no. 9 (2021) 664–672. https://doi.org/10.3102/0013189X211048810

[26] M. Liston, A. M. Morrin, T. Furlong and L. Griffin, Integrating Data Science and the Internet of Things Into Science, Technology, Engineering, Arts, and Mathematics Education Through the Use of New and Emerging Technologies, *Front. Educ.*, **7** (2022) 757866.

[27] M. Loukides, *What Is Data Science?* Cambridge, MA: O'Reilly Press, 2012.

[28] W. Martinez and D. LaLonde, Data science for everyone starts in kindergarten: strategies and initiatives from the American statistical association, *Harvard Data Science Review [Internet]*, **2** no. 3 (2020). https://hdsr.mitpress.mit.edu/pub/wkhg4f7a

[29] R. Moeti, A. Rafiepour and M. R. Fadaee, The role of technology in school data science, *In Thirteenth Congress of the European Society for Research in Mathematics Education (CERME13)*, no. 24, Alfréd Rényi Institute of Mathematics, ERME, 2023.

[30] L. Muñiz-Rodríguez, L. J. Rodríguez-Muñiz and A. Alsina, Deficits in the Statistical and Probabilistic Literacy of Citizens: Effects in a World in Crisis, *Mathematics*, **8** no. 11 (2020) 1872.

http://dx.doi.org/10.22108/msci.2025.143079.1704

[31] G. Nadzinski, B. Gerazov, S. Zlatinov, T. Kartalov, M. M. Dimitrovska, H. Gjoreski, R. Chavdarov, Z. Kokolanski, I. Atanasov, J. Horstmann, U. Sterle and M. Gams, Data Science and Machine Learning Teaching Practices with Focus on Vocational Education and Training, *Informatics in Education*, **22** no. 4 (2023) 671–690.

[32] D. Rose, *Data Science: Create Teams That Ask the Right Questions and Deliver Real Value*, New York: Apress, 2016.

[33] E. H. Schultheis and M. K. Kjelvik, Data Nuggets:Bringing real data into the classroom to unearth students' quantitative and inquiry skills,*American Biology Teacher*, **77** no. 1 (2015) 19–29.

[34] I. Song and Y. Zhu, Big data and data science: What should we teach? *Expert Systems*, **33** no. 4 (2016) 364–373.

[35] R. Thompson and G. Arastoopour Irgens, Data Detectives: A data science program for middle grade learners, *Journal of Statistics and Data Science Education*, **30** no. 1 (2022) 29–38.

[36] M. Vahidi-Asl, 50 Years of Data Science, *Mathematical Culture and Thought*, **43** no. 1 (2024) 273–327. [In Persian]

[37] T. Weiland and A. Sundrani, Opportunities for K-8 Students to Learn Statistics Created by States' Standards in the United States, *Journal of Statistics and Data Science Education*, **30** no.2 (2022) 165–178.

[38] M. B. Wilson and C. Turner, editors. *The philosophy of Werner Herzog*, Lexington Books, Washington, DC, 2020.

[39] M. S. Zakaria, Data science education programmes in Middle Eastern institutions: A survey study, *IFLA Journal*, **49** no. 1 (2023) 157–179.

Reza Moeti

Department of Mathematics Education, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman Kerman, Iran

Mahani Math Center, Afzalipour Research Institute, Shahid Bahonar University of Kerman Kerman, Iran

Email: rezamoeti@math.uk.ac.ir

Abolfazl Rafiepour

Department of Mathematics Education, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman Kerman, Iran Email: Rafiepour@uk.ac.ir

Mohamad Reza Fadaee

Department of Mathematics Education, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman Kerman, Iran

Email: m_fadaii@uk.ac.ir