Efficacy of Data Science Education in K–12 Education

Training every student for the jobs of tomorrow is no easy task.
Education researchers, curriculum developers, and Federal agencies including the National Science Foundation (NSF) and Institute of Education Sciences (IES) have been supporting research to design data science education programs for over a decade. Critical resources are needed to expand and scale this research so that all students are prepared in foundational skills for big data, artificial intelligence, machine-learning, and other emerging technologies that will entirely reshape careers, civics, and personal life.

#1: Prepares students for automation-resistant jobs

- By 2030, while only 5% of jobs will be fully automated, **60% of all occupations will have constituent tasks automated,** meaning that "welders to mortgage brokers will work alongside rapidly-evolving machines" – everyone will need fluency in the inputs, outputs, and consumption of data from technology.[1]
- **Technological skill requirements will grow +55% by 2030**, outpacing social-emotional (+24%) or advanced cognitive skill (+8%) requirements while physical and manual skills are automated (-14%).[2]
- **85%+ of global companies will adopt "big data"** analytics by 2025, and most companies planned to "accelerate the digitalization of work processes" in a post-COVID economy.[3]
- Black Americans are most at-risk for automation, representing up to 97% of automation in some geographies – a combination of re-skilling and education will lead to the highest payoffs.[4]

#2: Increases equity in mathematics and STEM

- K-12 data science curricula interventions have been found to **increase student competence (self-efficacy), autonomy (agency), and relatedness (sense of belonging)** in mathematics and technology, especially for students who previously struggled in mathematics.[5]
- Increasing relevance, integrating technology, and reducing procedural repetition have been identified as key strategies for tackling well-documented "math anxiety," which has historically created STEM achievement barriers for disadvantaged students and others who do not identify with STEM subjects.[6] [7]
- A National Academies study identified introductory data science in K-12 as a **primary strategy for increasing representation and diversity in emerging technology**, with early exposure especially critical.[8]

#3: Accelerates achievement and builds technology confidence

- The unique combination of inquiry-based technology learning, cross-subject applications, and personalization to students' interests through open datasets are driving **stronger engagement, higher persistence, and faster achievement** than anticipated across data science programs.[9] [10] [11] [12]
- Data science curricula have increased technology confidence [13] [14] and real-world career connections in high school, with students pursuing internships and other digital career training after exposure.[15]
- Integrating data science lessons into traditional curriculum have **dramatically increased Advanced Placement (AP) scores where implemented**, boosting pass rates from 2% to 42% in some Title I schools.[16]

#4: Backed by a decade of research & development (R&D)

2010 – National Science Foundation (NSF) grant funds development of first K-12 Data Science course[17]
5 high school courses and 60+ other classroom resources now exist for immediate classroom use, supported by NSF grants and philanthropy.

2014 – National Science Foundation (NSF) grant funds development of first student data software[18]
K-12 students have learned Excel, R, Python, Tableau, SQL, and a number of real-world tools, using classroom software tools like CODAP as a launchpad.

2017 – Data Science Education Technology (DSET) Conference[19]
First national convening of K-12 education researchers for data science, held regularly since.

2019 – International Data Science in Schools Project (IDSSP)[20]
First learning framework for data science education, created by an international group of education researchers in statistics, computer science, and related disciplines.

2020 – National Summit for K-12 Data Science Education (DS4E)[21]
First national summit for K-12 data science education, co-hosted by Stanford University and University of Chicago, held regularly since.[22]

2020 – ASA / NCTM releases GAISE II: A Framework for Statistics & Data Science Education[23]
First U.S.-based K-12 learning framework modernized for era of data science education

2021 – Institute of Education Sciences (IES) Technical Working Group[24]
First Federal working group for data science education, convened by the National Center for Education Research (NCER).

2022 – National Academies of Sciences, Engineering, and Medicine (NASEM) National Workshop[25]
"Foundations of Data Science for Students in Grade K-12," cataloguing first decade of education research in the new discipline of K-12 Data Science.

Next steps recommended by IES

1. **Specify the Developmental Pathway**—more research to better articulate K-12 learning pathways for students, including in elementary (K-5) and middle (6-8).
2. **Improve Data Science Software**—best practices and increased accessibility for which data science software should be incorporated into K-12 instruction and when.
3. **Build Tools for Assessment**—development of classroom assessment tools to track student progress, ensure transfer credit to postsecondary education, and improve the existing evidence base.
4. **Integrate Equitably into Schooling and Systems**—ensure equitable access to and participation in high quality data science education for all learners, regardless of zip-code or local resources.
5. **Improve Implementation**—address several systematic barriers for scaling implementation, including insufficient financial resources, lack of teacher training, and misalignment in college requirements.

Selected References

[1] McKinsey Global Institute. 2018. "AI, automation, and the future of work: Ten things to solve for," June 1, 2018. https://www.mckinsey.com/featured-insights/future-of-work/ai-automation-and-the-future-of-work-ten-things-to-solve-for#part2

[2] Ibid.

[3] World Economic Forum. 2020. "The Future of Jobs Report." October 2020. https://www3.weforum.org/docs/WEF_Future_of_Jobs_2020.pdf

[4] McKinsey Global Institute, "The future of work in Black America," Oct 4, 2019. https://www.mckinsey.com/featured-insights/future-of-work/the-future-of-work-in-black-america

[5] Heinzman, Erica. 2022. "'I love math only if it's coding': A Case Study of Student Experiences in an Introduction to Data Science Course." Statistics Education Research Journal. 21(2), 5. https://doi.org/10.52041/serj.v21i2.43

[6] Rossnan, S. (2006). Overcoming math anxiety. Mathitudes, 1 (1), pages 1 of 4

[7] Ruedinger, Dylan; Cooper, Daniella Miriam; Minnis, Samantha; and Schaub, Gayle, "Fostering Data Literacy Teaching with Quantitative Data in the Social Sciences" (2022). Scholarly Papers and Articles. 75. https://scholarworks.gvsu.edu/library_sp/75

[8] National Academies of Sciences, Engineering, and Medicine. 2020. *Roundtable on Data Science Postsecondary Education: A Compilation of Meeting Highlights*. Washington, DC: The National Academies Press. https://doi.org/10.17226/25804.

[9] Matuk, C., DesPortes, K., Amato, A., Silander, M., Vacca, R., Vasudevan, V. & Woods, P.J. (2021). Challenges and opportunities in teaching and learning data literacy through art. In, Proceedings of the International Society for the Learning Sciences Conference, pp. 681-685. Online. [Full text]

[10] Higgins, Traci; Rubin, Andee; Mokros, Jan; Sagrans, Jacob; Ren-Mitchell, Ada. 2021. "When The Data Drive the Learning." In: R Helenius, E Falck (Eds.), Statistics Education in the Era of Data Science. Proceedings of the Satellite conference of the International Association for Statistical Education (IASE), Aug-Sept 2021. https://iase-web.org/documents/papers/sat2021/IASE2021%20Satellite%20136_RUBIN.pdf?1648765057.

[11] Stephanie C. Hicks & Rafael A. Irizarry (2018) A Guide to Teaching Data Science, The American Statistician, 72:4, 382-391, DOI: 10.1080/00031305.2017.1356747

[12] Erwin, R. W. (2015). Data Literacy: Real-World Learning Through Problem-Solving With Data Sets. American Secondary Education, 43(2), 18–26. http://www.jstor.org/stable/43694208

[13] Program Evaluation and Research Group (PERG), 2020. Data Clubs. NSF grant nos. DRL-1742255. September 2020. https://www.terc.edu/dataclubs/wp-content/uploads/sites/10/2021/09/Fall-2020-Data-Dispostions-Results.pdf

[14] Tucker, M., Shaw, S.T., Son, J.Y., & Stigler, J.W. (Accepted). Integrating R in a college statistics course improves student attitudes toward programming. Submitted to the Annual Meeting of the American Educational Research Association (Orlando, Florida, April 9-12, 2021).
(PDF) Integrating R in a College Statistics Course Improves Student Attitudes Toward Programming. Available from: https://www.researchgate.net/publication/348817542_Integrating_R_in_a_College_Statistics_Course_Improves_Student_Attitudes_Toward_Programming.

[15] Drozda et al., 2022. Zarek Drozda, Davis Johnstone, Brooke Van Horne: "Previewing the National Landscape of Data Science Implementation." Commissioned for the National Academy of Sciences. September 1, 2022.

[16] Skew the Script, Impact Report. November 2022. https://skewthescript.org/about

[17] NSF, Award #0962919. 2010. MOBILIZE: Mobilizing for Innovative Computer Science Teaching and Learning. https://www.nsf.gov/awardsearch/showAward?AWD_ID=0962919

[18] NSF, Award #1435470. 2014. Common Online Data Analysis Platform (CODAP). https://www.nsf.gov/awardsearch/showAward?AWD_ID=1435470

[19] Data Science Education Technology (DSET) Conference, Feb 15-17, 2017. https://codap.concord.org/dset/index.html

[20] IDSSP Curriculum Team, Curriculum Frameworks for Introductory Data Science, http://idssp.org/files/IDSSP_Frameworks_1.0.pdf ISBN: 978-0-646-80819-2

[21] National Summit for K-12 Data Science Education, 2020. "Bringing Math Class into the Digital Age." March 3, 2020. https://ed.stanford.edu/news/bringing-math-class-data-age

[22] Data Science 4 Everyone, National Commitments Events; 2021,22 https://www.datascience4everyone.org/commitments-directory

[23] Bargagliotti, Anna et al., "Pre-K–12 guidelines for assessment and instruction in statistics education II (GAISE II)," American Statistical Association (ASA) and National Council of Teachers of Mathematics (NCTM), November 2020. https://www.amstat.org/docs/default-source/amstat-documents/gaiseiiprek-12_full.pdf

[24] National Center for Education Research (NCER), 2021. "Catalyzing a New Field: Data Science Education in K-12," Technical Working Group Meeting summary. Institute of Education Sciences (IES), U.S. Department of Education. October 26, 2021. Washington, D.C. https://ies.ed.gov/ncer/whatsnew/techworkinggroup/pdf/DataScienceTWG.pdf

[25] National Academies of Sciences, Engineering, and Medicine (NASEM), 2022. *"Foundations of Data Science for Students in Grade K-12: A Workshop,"* September 13-14, 2022. Washington, D.C. https://www.nationalacademies.org/our-work/foundations-of-data-science-for-students-in-grades-k-12-a-workshop