DataScience4
everyone

Beyond Borders 2024:

Primary and Secondary Data Science Education Around the World

Sean Sukol, *Author*
Shea Stripling, *Designer*
Data Science 4 Everyone

Introduction

Data Science 4 Everyone is a coalition advancing data science education so that every K-12 student is equipped with the data skills they need to succeed in our modern world. We believe that equitable access to data science education opens doors to higher education, high-paying careers, an engaged community, and a thriving democracy.

This report extends our November 2022 publication "Who Else is Teaching Data Science in K-12" with a deeper investigation into data science and data literacy education, as well as advocacy efforts in six countries around the world. It also includes brief descriptions of exciting recent developments for data science education in several more countries. To understand and report on education systems outside our own context, we reached out to over 60 organizations and individuals across 14 countries; interviewed more than a dozen subject-matter experts; and researched countless curricula, classroom resources, news and journal articles, policy documents, conference presentations, and more. While it is impossible to provide a comprehensive analysis of data science education around the world in these few pages, we aim to highlight the policy, research, curricular, funding, and advocacy activity that we think are most applicable to readers and data science education advocates in the US.

...many other countries have already established data science and data literacy education as a national priority and undertaken education reforms to prepare their students and citizens for a more competitive, data-rich, and

The US has only just begun to lay the foundation for data science education with surveys of the current education landscape and investigations into the implication of AI for our society. But our peers and competitors have already made data science, machine learning, and AI required components of their K-12 curricula; issued learning progressions and other national standards for the study of data science; leveraged public-private partnerships and established national grant programs to reach students in remote and underserved communities; and collaborated with industry leaders to write textbooks and fund research into the best ways to teach data science.

Our reports on data science education and education advocacy in Canada, the UK, Germany, China, India, and New Zealand each begin with a brief description of the country's education system to orient US readers who may be unfamiliar with them. We briefly describe relevant national initiatives and policies, especially recent education reforms, and highlight organizations that deliver or advocate for data science education in primary and secondary schools. We also provide information on tertiary education initiatives, research into teaching and curricula, and industry concerns where appropriate. Policies and projects in Australia, Colombia, Estonia, Finland, Israel, Japan, the Netherlands, Poland, Singapore, and South Korea receive brief descriptions. We include citations, links, and quotes by experts familiar with the respective local contexts throughout the report, and each section ends with a list of items for further reading.

Our writing on data science education initiatives, policies and programs is descriptive. We do not prescribe or argue for any one method of curricular integration, policy advocacy, or teaching in this report. Rather, we recognize the decentralized nature of education policymaking in the US and hope to showcase the many policy and education options we see abroad with the hope that readers will be inspired by those most applicable and effective in their own contexts.

As the primary author, I wish to express my profound gratitude to our many collaborators in the US and abroad who have shared their insight, experience, and feedback. My early professional career as a teacher in China, Spain, and the US was a humbling and formative experience, and I developed an intimate appreciation for the everyday challenges of educators around the world. Through my recent work with US state legislatures and agencies, I have also developed an understanding of the ways in which the availability of data—and our capacity to interpret data—impact our laws, policies, society, and livelihoods. This report underscores the urgent need in the US to make data science and data literacy education a fundamental component of our K-12 system. It can be done, and our future and fortune depend on it.

Sean Sukol
Data and Policy Analyst
Data Science 4 Everyone
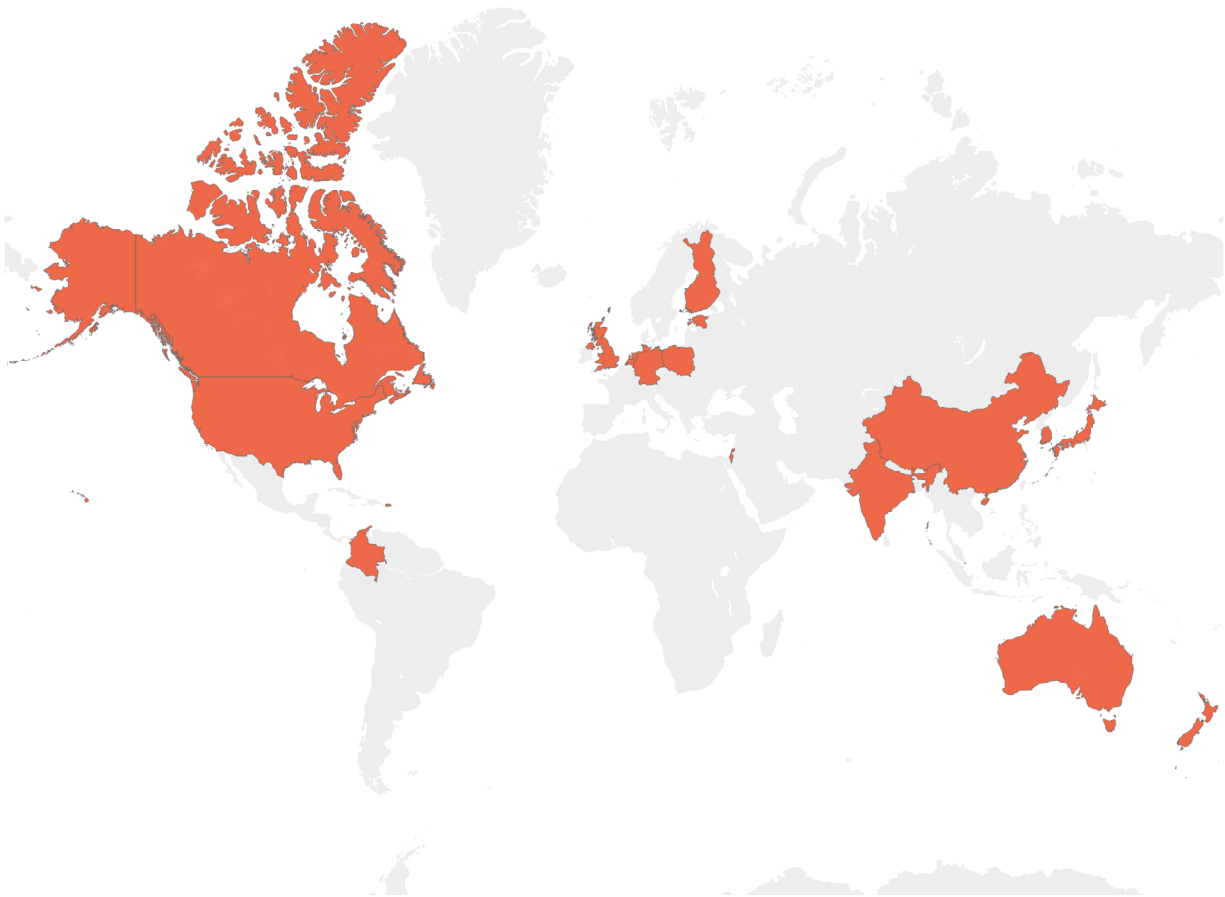
The Need for Data Science Education in the United States

The data boom of the 21st century has arrived, and it is increasingly machine-driven: there will soon be over 200 zettabytes—200 trillion gigabytes—of data in public and private cloud storage in 2025, up from 44 zettabytes in 2022.[1] Over 40% of internet data was machine-generated in 2020, and the explosive growth in popularity of generative AI platforms in 2023 is sure to accelerate the trend. The immense quantity and velocity of data has created new opportunities and challenges, and our everyday tasks at home, work, and school now commonly require us to understand and interpret variables, observations, and statistics. Simple decisions are becoming overwhelmingly complex and often require increasingly sophisticated data analysis techniques with ever-larger datasets. Recent breakthroughs in artificial intelligence, currently led by US companies, represent only the latest application of the power of digital data at scale.

Yet businesses across industries have struggled to keep up, and the workforce data literacy skills gap is now a well-documented problem. Over 50% of US and UK leaders believe that their organizations have a data literacy skills gap, and only 5% of organizations believe that they are "fully data literate."[2] Recruiters report that data skills and data literacy are already the most in-demand skill set for entry-level positions, and they are anticipated to increase in importance in the coming years.[3] The Bureau of Labor Statistics expects employment of data scientists to grow by 35% between 2022 and 2032, making it the third-fastest growing occupation in the entire US economy.[4] And even as many organizations do not provide any or only limited data skills training for their employees, fewer than half of US universities have a specific initiative to teach students the data skills they will need to compete in an increasingly competitive and globalized economy.[5]

The issue is just as acute at the primary and secondary school levels. While the need for data skills has increased at home and in the workplace, K-12 students have been losing ground on the essential learning skills necessary to develop them. Over the last decade, scores in data analysis, statistics, and probability on the mathematics exam of the National Assessment of Educational Progress (NAEP), commonly referred to as the "nation's report card," dropped

10 points for grade 4 students and 17 points for grade 8 students, equivalent to a loss of between one and two years of learning.[6] Math teachers generally reported a lower emphasis on data literacy instruction during this period, despite the fact that 43% of children in the US believe they are not learning enough about how to analyze and use data, and 9% believe they are learning nothing about it at all.[7]

> "Data is now more valuable than oil, and can be refined anywhere at any time to extract its value – though only if the individual is equipped with the right skills to do so. The wealth of nations will only become more dependent on our education systems to impart the relevant technical skills, problem-solving techniques, and quantitative foundations for this new world. The extraordinary opportunities that today's data-driven technologies afford everyone will go to those who realize it the fastest, and those who ensure the greatest participation in them."
>
> **Zarek Drozda**
> **Executive Director**
> **Data Science 4 Everyone**

Worst of all, we are lagging behind our peers. The US has trailed the average of other Organisation for Economic Co-operation and Development (OECD) nations on the Programme for International Student Assessment (PISA) since math was first tested in 2003. And though the US saw a comparatively small decrease in reading exam scores during the COVID-19 pandemic, US math scores for the 2022 exam were among the lowest ever recorded by the PISA.[8] Policymakers, teachers, and parents in the US have finally begun to turn their attention to the work that must be done to improve data science and data literacy education. But many nations with similarly large, diverse, and globalized

1   TechJury, "How Much Data is Created Every Day?" TechJury, Accessed February 11, 2024, Techjury.net - How Much Data is Created Every Day?.
2   "The State of Data Literacy 2023" (DataCamp: DataCamp, 2023), 14, Datacamp - The State of Data Literacy 2023, page 14.
3   "Bridging the Data Literacy Gap Between Academia and Workforce," Tableau Blog, Accessed February 11, 2024, Tableau - Bridging the Data Literacy Gap Between Academia and the Workforce.
4   U.S. Bureau of Labor Statistics, Occupational Outlook Handbook: Data Scientists and U.S. Bureau of Labor Statistics, and Occupational Outlook Handbook: Fastest Growing Occupations.
5   The Great Data Literacy Gap (Seattle: Tableau, 2021), 7, Forrester - The Great Data Literacy Gap, page 7.
6   "Data Science is Vital to Student Success: So Why Are Outcomes Going Down?" (Chicago: DS4E, 2023), DS4E's 2023 NAEP Report.
7   Transforming Education Report (London: World's Largest Lesson, 2023), 15, World's Largest Lesson - Transforming Education, page 15.
8   "U.S. students show no improvement in PISA global assessment," Axios, December 5, 2023, Axios - U.S. students' math scores plunge in global education assessment.

economies have already adopted national policies, established institutions, and invested funding to position themselves as global leaders in analytics, data science, machine learning, and AI research.

The risks associated with under-investment are high. In a recent *Cyber Defense Review* position, "The Impending Data Literacy Crisis Among Military Leaders," Lieutenant Colonel Andrew G. Farina argues that "the potential for AI to drastically change the speed of decisions, and thereby the speed of war, will be revolutionary. Unfortunately, without a focused effort to improve military leaders' understanding of the data science field, commanders will lack trust in these technologies or, far worse, will over-rely on amoral machines to make decisions for them."[1] Several defense officials have made urgent calls for quickly improving the data literacy of US army commanders, including proposals for required examinations[2] and new curriculum reform efforts at military academies such as West Point.[3]

> Emerging technologies fueled by data are accelerating an unprecedented global talent race, and the champions of the digital frontier will be entirely determined by the ability of national education systems to adapt accordingly.

The US must establish a data science and data literacy education pipeline from K-12 education to our workforce, universities, and government. While we have been encouraged by the recent work of federal, state, and local policymakers to improve the state of data science and data literacy education in the US, we have only just begun. Inattention to our data skills risks our leadership in AI development and our long-run economic competitiveness and national security. Emerging technologies fueled by data are accelerating an unprecedented global talent race, and the champions of the digital frontier will be entirely determined by the ability of national education systems to adapt accordingly.

Jurisdiction	2022 Mathematics PISA mean score
Singapore	575
Macao	552
Taiwan	547
Hong Kong	540
Japan	536
South Korea	527
Estonia	510
Switzerland	508
Canada	497
Netherlands	493
Ireland	492
Belgium	489
Denmark	489
Poland	489
United Kingdom	489
Australia	487
Austria	487
Czech Republic	487
Slovenia	485
Finland	484
Latvia	483
Sweden	482
New Zealand	479
Germany	475
Lithuania	475
France	474
Hungary	473
Spain	473
International Average (OECD)	472
Portugal	472
Italy	471
Vietnam	469
Norway	468
Malta	466
United States	465

1   "The Impending Data Literacy Crisis Among Military Leaders," page 92, Fall 2022, The Cyber Defense Review.
2   Erik Davis, "The Need to Train Data-Literate U.S. Army Commanders." War on the Rocks, Oct 17, 2023.
3   Meghan Dower-Rogers, "West Point Center for Data Analysis and Statistics Holds Data Literacy Training Event for Army Leaders," West Point, Aug 2, 2023.

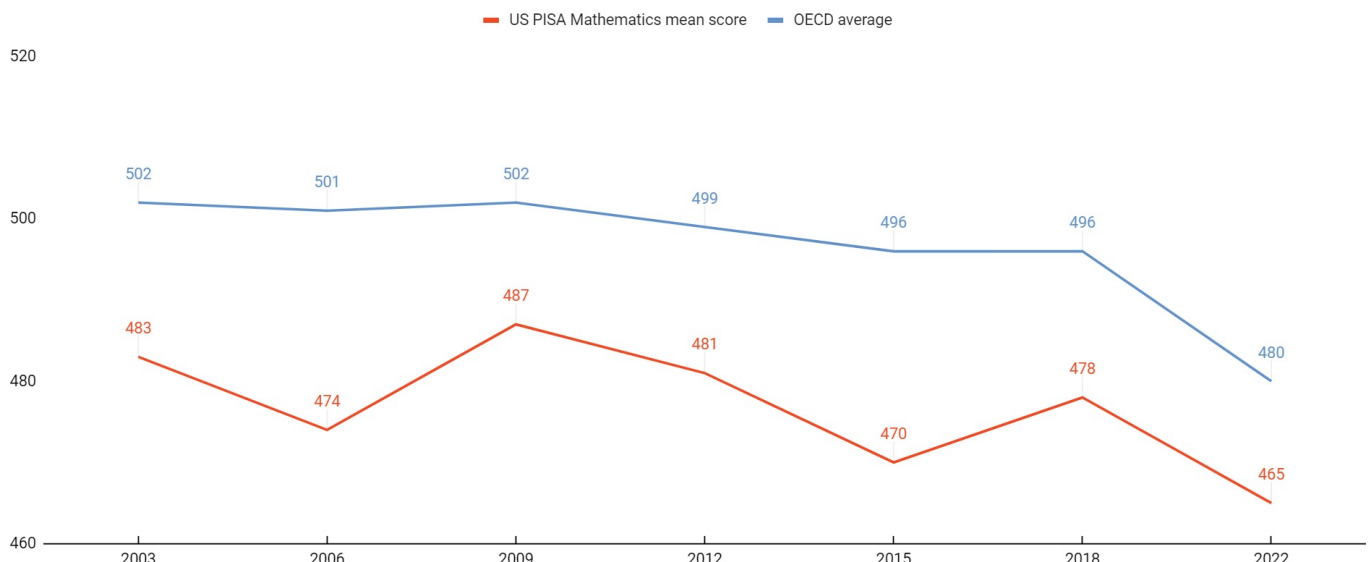# What is Data Science and Data Literacy Education?

As the need for data science and data literacy skills has increased, many organizations have developed their own working definitions of data science and data literacy for their specific application, industry, or education context. For the purposes of this report, we define data science education in broad terms so that the many conversations between policymakers, researchers, parents, teachers, and students can align to discuss data science and data literacy education in the US.

Data literacy refers to the ability to read, understand, interpret, and communicate data or claims derived from data. In the context of K-12 education, data literacy involves teaching students how to iteratively question data and representations of data critically, including understanding their origins, limitations, and potential biases. It also involves teaching students how to communicate data effectively through visualizations and other means.

Data science education refers to instruction focused on systematic processes, analytical techniques, and utilization of appropriate technologies to gain knowledge from data. In the context of K-12 education, it equips students with introductory problem-solving skills necessary to collect, consider, analyze, interpret, and communicate data to answer investigative questions. It draws upon tools and methods across the curriculum in areas including methodological fields such as mathematics, statistics, and computer science, and disciplines in which they are applied, such as life sciences, social sciences, digital humanities, and others. This may include exploratory data analysis, visualization, modeling, and data ethics.

Data science education programs may utilize instruction and technology tools in:

- analytics
- operations research
- data acquisition and management
- data secuirty and privacy
- the ethical use of data
- social impacts and professional practices of statistics and data science
- applied statistics
- predictive modeling
- statistical methods and modeling
- statistical problem-solving processes
- artificial intelligence
- machine learning
- data literacy
- exploratory data analysis
- data quality and representation
- data visualization

Legend: ▬ US PISA Mathematics mean score ▬ OECD average

Year	US PISA Mathematics mean score	OECD average
2003	483	502
2006	474	501
2009	487	502
2012	481	499
2015	470	496
2018	478	496
2022	465	480

Canada

Canada

Primary and secondary education in Canada falls under provincial jurisdiction, and the decentralized function of the Canadian education system makes it especially salient for comparison to data science education in the US. Canadian public primary and secondary schools are managed by elected local school boards and deliver curricula based on education standards set by the provincial ministries of education.[1] The government's central statistics office, Statistics Canada, runs several programs to promote data science learning and upskilling among public sector employees, including the Data Science Network for the Federal Public Service and the Data Science Fellowship Program.[2]

CanCode, an initiative of the 2017 Innovation and Skills Plan, funds K-12 learning and teacher professional development programs with a focus on digital skills development, including coding and data analytics.[3] CanCode was extended for the second time in 2021, and by 2022, CanCode-funded programs reached over 4.5 million students and 220,000 teachers. The third iteration will distribute an additional 80 million Canadian dollars to at least 29 projects across Canadian provinces and online.[4]

Among the first recipients of CanCode funding was Kids Code Jeunesse, "a bilingual Canadian charity determined to give every Canadian child access to digital skills education, with a focus on girls and underserved communities."[5] Kids Code Jeunesse has trained over 22,000 educators and reached more than 700,000 children through initiatives such as the Algorithm and Data Literacy Project, a partnership between the United Nations Educational, Scientific and Cultural Organization (UNESCO), the Canadian Commission for UNESCO, and Kids Code Jeunesse's parent organization, Digital Moment.[6] The Algorithm and Data Literacy Project's goal is to "empower kids to exercise critical thinking in how they engage online, and to become proactive, creative users and makers rather than passive consumers." Digital Moment has published a series of materials through the project to help students and educators improve their data literacy skills and understand AI. The AI Primer, Data

Literacy Discussion Guide, Educator's Guide, and student workshops produced by the Algorithm and Data Literacy Project emphasize ethics, algorithmic bias and discrimination, and data privacy. The Algorithm and Data Literacy project is currently developing a project-based learning module that introduces students to AI-driven facial recognition technology with exercises on visual perceptions and representations of human emotion.

"In today's information-driven world, familiarity with data is crucial, and data science is the discipline of learning from data. It aids students in applying computational thinking processes and tools, honing critical thinking skills and helping informed decision-making, not just in the classroom but in our increasingly digital world. At Callysto, we develop curriculum-aligned data science resources for teachers to utilize in their classrooms, allowing students to learn these invaluable skills, and how they can be useful in a wide variety of domains and applications."
Byron Chu
Callysto Program Mananger

The module, *Machine Intelligente, dis-moi comment je me sens?* ("Intelligent Machine, tell me how I Feel?") focuses on teaching students about how issues of bias, sample, biometric data, and cultural differences in the expression of emotion are inherent in artificial intelligence and machine learning technology.[7] In 2021, the Algorithm and Data Literacy Project reached 529 pre-service teachers through nine universities across six provinces.[8]

The current round of CanCode funding also supports Callysto, a nonprofit program that provides a free Jupyter Notebook-based online learning platform that brings data science skills into grades 5-12 classrooms across Canada. Callysto trains teachers to integrate data science in their lessons and provides student-friendly data science modules that align with curricula requirements. Callysto also

1   Government of Canada, "Understand the Canadian education system," EduCanada, accessed February 24, 2024, EduCanada -Understand the Canadian Education System.
2   Statistics Canada, "Data Science," Statistics Canada, accessed February 24, 2024, Statistics Canada - Data Science Centre.
3   Innovation, Science and Economic Development Canada, "Canada's National Digital and Data Consultations," ISED - ISDE Canada, accessed February 24, 2024, Innovation, Science and Economic Development Canada - CanCode.
4   Innovation, Science and Economic Development Canada, "Funded CanCode Initiatives," ISED - ISDE Canada, accessed February 24, 2024, Innovation, Science and Economic Development Canada - CanCode 3.0 Project Description.
5   Kids Code Jeunesse, "About - Kids Code Jeunesse," accessed February 24, 2024, Kids Code Jeunesse - About.
6   Algorithm Literacy, "Algorithm Literacy," accessed February 24, 2024, The Algorithm and Data Literacy Project.
7   Lotfi Gharbi, Learning Specialist, Digital Moment, email message to the author, February 23, 2024.
8   Kids Code Jeunesse, "KCJ Annual Report 2021," Issuu, page 24, accessed February 24, 2024, Kids Code Jeunesse - Annual Report 2020-2021.

facilitates free in-person and online data science hacka-thons for students. During these events, students analyze datasets in order to answer questions and respond to challenges with support from the organization's staff, who review their analyses and final reports. To date, Callysto has reached 130,000 students and 5,000 teachers.[1]

"More and more, the way in which this data is processed and used by artificial intelligence is hidden behind the interface of the tools we use."
Lotfi Gharbi
Learning Specialist, Digital Moment

Provincial governments have already begun adopting curricula and learning standards that integrate data science education topics. British Columbia's recently redesigned science curriculum includes a "processing and analyzing data and information" learning strand for students in grades K-10.[2] Early learners in grades 1 and 2 "sort and classify data and information using drawings, pictographs and provided tables"; "compare observations with predictions through discussion"; and "identify simple patterns and connec-tions." In grades 3 and 4, students "use tables, simple bar graphs, or other formats to represent data and show simple patterns and trends" and "compare results with predicti-ons, suggesting possible reasons for findings." In grades 9 and 10, students begin applying true data science learning concepts including:

- evaluate their methods and experimental conditions, including identifying sources of error or uncertainty, confounding variables, and possible alternative expla-nations and conclusions;
- describe specific ways to improve their investigation methods and the quality of the data
- evaluate the validity and limitations of a model or ana-logy in relation to the phenomenon modeled;
- demonstrate an awareness of assumptions, question information given, and identify bias in their own work and secondary sources;
- consider the changes in knowledge over time as tools and technologies have developed;
- connect scientific explorations to careers in science;
- exercise a healthy, informed skepticism, and use scien-tific knowledge and findings to form their own investi-gations and to evaluate claims in secondary sources;
- consider social, ethical, and environmental implications of the findings from their own and others' investigati-ons; and
- critically analyze the validity of information in seconda-ry sources and evaluate the approaches used to solve problems.

1   Callysto Project, accessed February 24, 2024, Callysto - What is Callysto?.
2   British Columbia Ministry of Education, "Science K-10 Curricular Competencies," accessed February 24, 2024, Province of British Columbia - Science K - 10 Curricular Competencies.

The Ontario Ministry of Education recently updated their province's math and science and technology curricula for grades 1–8, replacing learning standards from 2005 and 2007, respectively.[1] The new curricula sets standards for learning about data, including "how to collect, organize, display, and analyze data to make convincing arguments"; "how to determine when data is being misrepresented"; the creation of "infographics to tell a story using data"; and how to "make connections between the use of data and understanding the chance that something might happen."[2] Coding concepts are introduced beginning in grade one and increase in complexity by grade band.[3] Grades 1-3 learn about "algorithms, decomposing problems, and testing, debugging, and refining solutions"; grades 4-6 are introduced to "input, storage, processing, and output"; and grades 7-8 are instructed in "project management and program design" and "controlling large systems in action." Students are also required to learn about the impact and practical applications of emerging technologies, including AI, "on everyday life, including skilled trades," and the interdisciplinary application of multiple content areas to solve real-world problems.

Alberta has introduced computational thinking topics, including coding and algorithms, into its ongoing K-12 curriculum renewal, which has been underway since 2019.[4] Under the new curricula, grade four students will "investigate the role of data and evidence in science" and "examine design processes and their application in solving problems"; this is expanded on in grade five by applying "the design process in creating computational artifacts, including code"; and grade six students will "examine abstractions, coding structures and the impact of computers and technology." Updates to the grades 4-6 science curricula will be implemented in classrooms across Alberta beginning with the 2024 school year, and the province plans to begin drafting new curricula for grades 7 through 12 after receiving feedback on K-6 revisions and implementation pilots.[5]

1   CBC News, "Ontario elementary math curriculum going back to 'back to basics'," CBC News, published June 26, 2020, CBC - How Ontario's new math curriculum goes way beyond back-to-basics.
2   Government of Ontario, "New Math Curriculum for Grades 1-8," Ontario.ca, accessed February 24, 2024, Province of Ontario - New math curriculum for Grades 1-8.
3   Ontario Ministry of Education, "Key Changes to Strand A in the Science and Technology Curriculum," accessed February 24, 2024, Province of Ontario - Key Changes — Science and Technology, Grades 1–8.
4   Government of Alberta, "Science curriculum overview," Alberta.ca, accessed February 24, 2024, Province of Alberta - Highlights of the new elementary science curriculum.
5   Government of Alberta, "How we got here: Development of the K-12 curriculum," Alberta.ca, accessed February 24, 2024, Province of Alberta - How we got here.

Further Reading on Data Science Education in Canada

- [Data Science Network for the Federal Public Service](#)

- [Data Science Fellowship Program](#)

- [CanCode](#)

- [Kids Code Jeunesse](#)

- [The Algorithm and Data Literacy Project](#)

- [Callysto](#)

- [Explore British Columbia's Curriculum](#)

- [Ontario Curriculum and Resources](#)

- [LearnAlberta](#)

The United Kingdom

The United Kingdom

The United Kingdom is a sovereign nation in which the power to decide certain policies, including on education, are not reserved by the UK Parliament but rather "devolved" among its member countries: England, Scotland, Northern Ireland, and Wales.[1] This report focuses on developments in curricular research in England and a multilevel government partnership in Scotland, though certain national policies are also relevant, in particular the National Data Strategy, which was published in 2020 in response to a growing need for data skills across the national economy.

The Royal Society, the UK's national science academy, has driven significant research on data science curricula and teaching in compulsory education. The Royal Society's work is complemented by a constellation of government programs supporting research, analysis, and collaboration in postsecondary data science education and training including the Centre for Data Ethics and Innovation, the Data Skills Taskforce, the Alliance for Data Science Professionals, and the Alan Turing Institute. The Institute was founded in 2015 as the UK's national organization for the study of data science and officially adopted the study of AI in 2017. The Institute provides knowledge transfer and upskilling programming such as data study groups, enrichment internships, research fellowships, training for data science and AI educators, and a free online learning platform.[2] The Office for National Statistics has also partnered with public service broadcaster BBC and the Micro:bit Educational Foundation to release a "playground survey" that uses cross-curricular activities for primary school students to provide "a practical approach to understanding the value of data and how data is used in today's world."[3]

In 2018, the Royal Society's Advisory Committee on Mathematics Education published a review of England's primary and secondary curricula to "identify the extent to which data science (or elements of it) exist within computing, mathematics, science, geography, history and other data-rich subjects."[4] The review found that history and geography curricula at the primary level contained no mention of the application of mathematics or the use and analysis of data.[5] However, many data science topics were already included in the mathematics, computing, and science primary school curricula, such as:

- recording findings and interpreting and presenting data using bar charts, pictograms, tables, scatter graphs, and line graphs;
- solving questions using information presented in scaled bar charts and pictograms and tables;
- distinguishing between discrete and continuous data;
- using graphical representation of data to solve problems;
- calculating the mean of a dataset and creating graphs with multiple variables;
- solving problems by decomposing them into smaller parts;
- basic algorithmic skills such as sequence, selection, and repetition in computer programs;
- working with variables and various forms of input and output;
- analyzing, evaluating, and presenting data and information using a variety of software;
- applying mathematical knowledge to answer questions by collecting, presenting, and analyzing data;
- reporting and presenting findings from enquiries, including conclusions, causal relationships, and explanations.

"Data literacy shouldn't be an Excel training course — it should be creative and engaging, showing learners how they can use data for social good to make a real difference. The problems learners solve don't need to be dry or dull — instead they can pretend to be secret agents finding the evil villain's secret lair using datasets to narrow down the possibilities!"
Kate Farrell
Director of Curriculum and Professional Learning
Data Education in Schools

Mathematics, computing, and science curricula at the secondary level were likewise found to substantively integrate data science topics, such as:
- measures of central tendency and spread;
- representing data using frequency tables, bar charts, pie charts, bar charts, and box plots;
- describing simple mathematical relationships between two variables in observational and experimental contexts and illustrating using scatter graphs;

1   UK Parliament, "The School Curriculum in England: Frameworks and Content," Commons Library, accessed February 24, 2024, Introduction to devolution in the United Kingdom.
2   The Alan Turing Institute, Turing 2.0: Changing the World for the Better with Data Science and AI, page 48 (2023) and The Turing Online Learning Platform.
3   BBC, "Microbit Playground Survey," accessed February 28, 2024, BBC Mircobit Playground Survey.
4   Royal Society, "The Integration of Data Science in the primary and secondary curriculum" (London: Royal Society, 2018), accessed February 24, 2024, The Integration of Data Science into the Primary and Secondary Curriculum (2018).
5   Royal Society, *The Integration of Data Science in the primary and secondary curriculum* (London: Royal Society, 2018), 9, accessed February 24, 2024, The Integration of Data Science into the Primary and Secondary Curriculum.

- designing, using and evaluating computational abstractions that model the state and behavior of real-world problems and physical systems;
- understanding several key algorithms that reflect computational thinking;
- appropriate use of data structures;
- database management systems and data querying;
- explaining data in relation to predictions and hypotheses;
- evaluating data for potential sources of random and systematic error; and
- identifying further questions arising from their results.

In contrast to the primary level, secondary school social sciences curricula were found to include several data science topics, including Geographic Information Systems (GIS), communicating geographical information through numerical and quantitative skills, and the integration of data analysis into empirical/experimental psychological research methods in General Certificate of Secondary Education (GSCE) psychology.[1] However, the review also identified barriers to the full integration of data science education including necessary changes to curricula, teacher professional development, and the availability of real-world datasets appropriate for students. At the primary level in particular, the review noted a lack of detail in content on algorithms and the presentation and analysis of information and data using technology, and recommended the "application of mathematics and use of data in specific areas of the science curriculum to enhance subject understanding" and "meaningful data in computing when applying technology in the analysis and presentation of data."[2]

The Royal Society in 2019 recognized a sharp increase in demand for employees with data science skills across industries and published a report indicating that significant changes to curricula would be necessary to strengthen England's data science talent pipeline.[3] In 2020, the Advisory Committee launched the Mathematical Futures Program to consider how England's compulsory education system would need to change in the following 20 years in order to ensure that students would develop the mathematics skills, including data science, that "will be needed for citizens and society to thrive in the future."[4] The Advisory Committee's September 2023 discussion paper calls for systemic changes to mathematics education that integrate concepts from statistics, data science, and computer science into existing curricula, and the program is set to issue its recommendations and report in 2024.[5]

Though their systems share many similarities, education in the UK is a devolved responsibility of each of the member countries, and even regional governance can play a significant role in advancing data science education. The 2017 Edinburgh and South East Scotland City Region Deal, a public-private partnership between the UK and Scottish governments, institutions of higher education, and local businesses, represents £1.3 billion of development program

1   Royal Society, "The Integration of Data Science in the Primary and Secondary Curriculum" (London: Royal Society, 2018), accessed February 24, 2024, The Integration of Data Science into the Primary and Secondary Curriculum, pp. 10-13.
2   Royal Society, "The Integration of Data Science in the Primary and Secondary Curriculum" (London: Royal Society, 2018), accessed February 24, 2024, The Integration of Data Science into the Primary and Secondary Curriculum, pp. 14-16 .
3   Royal Society, "Dynamics of Data Science," accessed February 24, 2024, Dynamics of data science skills (2019).
4   Royal Society, "Mathematical Futures," accessed February 24, 2024, Mathematical Futures programme: A new approach to Mathematics and Data Education (2023), pp. 2.
5   Royal Society, "New Approach to Mathematics and Data Education" (London: Royal Society, 2023), accessed February 24, 2024, Discussion Paper: A new approach to mathematics and data education (2023).

spending over 15 years.[1] The Deal includes £25 million for the development of career pathways in key sectors and up to £661 million for a Data-Driven Innovation program to "establish the region as the Data Capital of Europe."[2] The program is hosted primarily by the University of Edinburgh and intends to assist 100,000 people in obtaining a data-science certification or qualification by 2033.[3]

The program's Data Skills Gateway includes discrete initiatives for data science education and upskilling in schools, colleges, universities, and the workplace. Data Education in Schools, the Gateway's K-12 initiative, provides materials for the study of the National Progression Award in Data Science, a national course by the Scottish Qualifications Authority.[4] The course first became available in 2019 and is available for both schools and colleges.[5] The Award requires competencies in fundamental data science concepts and "data citizenship," including data literacy and social responsibility, with the option to study statistics, computer programming and machine learning in more depth.[6] Data Education in Schools also provides professional development workshops and classroom and lesson plan resources for data science teachers.[7]

1   Edinburgh and South East Scotland City Region, "ESESCR Deal Document" (Edinburgh: ESESCR, 2018), accessed February 24, 2024, Edinburgh and Southeast Scotland City Region Deal (2018), page 11.
2   Edinburgh and South East Scotland City Region, "ESESCR Deal Document" (Edinburgh: ESESCR, 2018), 12, accessed February 24, 2024, Edinburgh and Southeast Scotland City Region Deal (2018), page 11.
3   DDI, "Annual Review 2022," Readymag, accessed February 24, 2024, 2022 Annual Review, page 7.
4   Data Education in Schools, "NPA Data Science," accessed February 24, 2024, NPA Data Science.
5   Scottish Qualifications Authority, "National Progression Awards," accessed February 24, 2024, NPA Data Science SCQF level 4, 5 and 6.
6   Data Education in Schools, "Educator's Guide" (October 2020), accessed February 24, 2024, Educators' Guide to the National Progression Award (2020).
7   Data Education in Schools, "Supporting Teachers and Learners," accessed February 24, 2024, Professional Learning.

Further Reading on Data Science Education in the United Kingdom

- [National Data Strategy](#)

- [Data Skills Taskforce](#)

- [Alliance for Data Science Professionals](#)

- [The Alan Turing Institute](#)

- [The Royal Society Advisory Committee on Mathematics Education](#)

- [Mathematical Futures programme](#)

- [Edinburgh and South East Scotland City Region Deal](#)

- [Data Skills Gateway](#)

- [Data Education in Schools](#)

- [National Progression Award in Data Science](#)

China

China

The Chinese public education system is more centralized than that of the US, where education standards and curricula are the domain of state boards of education, local school boards, and other regional entities. The Chinese Ministry of Education sets education policy and teacher certification requirements, decides on curricula and course materials, and provides statistics and analysis of education outcomes at the national level.[1] Provincial education authorities are then responsible for using Ministry guidance in the development of local curriculums and implementation plans, which must in turn be submitted to the Ministry for approval. National education policies in China since 2017 reflect ambitious goals for data science education with an emphasis on AI research, development, and commercialization. China has quickly leveraged its centralized education authority and public-private partnerships to get a head start on its peer nations by establishing a data science education pipeline from primary school through postgraduate education.

The Chinese State Council released the *New Generation Artificial Intelligence Development Plan* in 2017 with the goal of making China a "global AI innovation center" by 2030.[2] Among the plan's major initiatives is the "construction of a multi-layer AI education program" that includes coursework at primary and secondary schools.[3] In response, the Ministry in 2018 made "Data and Computing" and "Information System and Society" mandatory curriculum modules for high school students. These modules deemphasize instruction on basic computing skills that are typical of information technology courses and focus instead on introducing AI concepts through programming languages and simple algorithms.[4] Students also use project-based learning to develop computer applications and learn skills fundamental to the collection, processing, analysis, and visualization of data.

Later in 2018, an AI textbook for secondary school students (and perhaps the world's first AI textbook in general), *Fundamentals of Artificial Intelligence*, was published by East China Normal University in collaboration with Sense-Time Group, an AI software development company.[5] The development of *Fundamentals* included input from high school teachers, and it was quickly implemented in the curriculum of an AI education pilot program in about 40 high schools. *Fundamentals* was soon followed by the publication of an AI textbook for junior high school, developed by a partnership between Northwest Normal University, the National Center for Educational Technology and AI, and intelligent speech developer iFlytek.[6]

In 2019, the Office of the State Council directed significant reforms in high school mathematics curricula in order to deemphasize exam-oriented education, among other goals.[7] Several previously mandatory topics in geometry, algebra, and calculus were made optional and removed from the *Gaokao*, China's standard college entrance exam,

1   Ministry of Education of the People's Republic of China, "What We Do," accessed February 24, 2024, PRC Ministry of Education - What We Do.
2   Embassy of the People's Republic of China in Finland, "China Science and Technology Newsletter No. 17" , page 4, accessed February 24, 2024, China Science and Technology Newsletter No. 17, page 4.
3   Asia Pacific Foundation of Canada, "Nurturing the Next Generation AI Workforce: A Snapshot of AI Education," accessed February 24, 2024, Nurturing the Next-Generation AI Workforce: A Snapshot of AI Education in China's Public Education System.
4   GETChina Insights, "China's Governmental Strategy and School-Enterprise Cooperaton on AI Education, Medium, 2019, China's Governmental Strategy and School-Enterprise Cooperation on AI Education.
5   China Daily, "China's AI education pushes forward," *China Daily*, June 11, 2018, accessed February 24, 2024, China Daily - First AI textbook for high school students released.
6   Asia Pacific Foundation of Canada, "Nurturing the Next Generation AI Workforce: A Snapshot of AI Education," accessed February 24, 2024, Asia-Pacific Foundation of Canada - Nurturing the Next-Generation AI Workforce: A Snapshot of AI Education in China's Public Education System.
7   State Council of the People's Republic of China, " Notice of the State Council on Issuing the White Paper on Standardization of Artificial Intelligence", June 19, 2019, accessed February 24, 2024, Office of the State Council - Guiding Opinions of the General Office of the State Council on Promoting the Reform of Educational Methods in Ordinary High Schools in the New Era.

including three-view diagrams, proofs, number sequences, and derivatives.[1] Statistics and probability topics such as random sampling, estimating population from samples, independent and dependent events, and frequency distributions were moved to textbooks earlier in the secondary school sequence, ensuring students will encounter those topics sooner than previous cohorts.

"Building data literacy is not just about familiarizing the students with data, but also equipping them with the principles and practices of scientific methods - that of being able to think, reason, and act based on empirical observations. For me, learning data science is humbling. Identifying causality in a sea of data illuminates to me that Truth needs to be not only examined, but contextualized."
Dan "April" Feng
Chief Operating Officer of Ameelio

Additionally, an entire semester's worth of topics related to data science was added to high school textbooks and the *Gaokao,* including an introduction to big data, distributions of discrete random variables, normal distributions, linear regression with two variables, and predictive models. In general, the new curriculum emphasized applied math and the use of real-world data. For example, students learned about estimating populations using samples through case studies that asked them to estimate the number of tanks in Germany between WWI and WWII and the prevalence of obesity in American office workers. The Beijing *Gaokao* in 2021 and again in 2022 included questions related to statistical analysis, hypothesis testing, probability, and data analysis.

The paramount importance of the *Gaokao* in the Chinese education system cannot be overstated: unlike the holistic admissions systems common among US universities, the *Gaokao* represents the single criterion for admission for Chinese students. Almost twelve million students took the 2022 *Gaokao*; fewer than 7,000 enrolled in China's top two universities, and fewer than 60% of Chinese students enrolled in tertiary education in 2023.[2] Undergraduate data science programs are the bridge between the skills taught

in secondary school and postgraduate study or employment in data science fields. A central feature of the Ministry's strategy has been the remarkably rapid expansion of university-level courses of study in data science topics. In 2016, only three higher education institutions in China offered the Ministry-approved undergraduate major "Data Science and Big Data Technology"; by 2018, that number rose to 250.[3] In 2019, the Ministry licensed 35 universities to offer AI as a newly approved undergraduate major; by 2023, 150 colleges and universities offered AI-related programs.[4] China's national policies emphasize multi-stream investment in AI research and development, including funding for university research centers through the "Bases and Talents" programs and public-private partnerships such as regional "guidance funds."[5]

1   Feng, April, "Memo: Data Science in Chinese Schools, Google Docs, accessed February 24, 2024, April Feng - Memo: Data Science in Chinese High Schools.
2   CGTN, "Explainer: Why the Gaokao is China's most important exam," accessed February 24, 2024, CGTN - Explainer: Why the Gaokao is China's most important exam.
3   Jilong Zhang, Anna Fu, Hao Wang, and Shenqin Yin - The Development of Data Science Education in China from the LIS Perspective, page 5, The Development of Data Science Education in China from the LIS Perspective.
4   APEC Business Advisory Council - Artificial Intelligence in APEC: Progress, Preparedness, and Priorities, page 21, Artificial Intelligence in APEC.
5   Center for Security and Emerging Technology, Georgetown University, Chinese Public AI R&D Spending: Provisional Findings, Georgetown University, Publication Date, accessed February 24, 2024, Center for Security and Emerging Technology - Chinese Public AI R&D Spending: Provisional Findings, page 10.

Further Reading on Data Science Education in China

- [Full Translation: China's 'New Generation Artificial Intelligence Development Plan' (2017)](#)

- [Global Times: - Introducing AI courses as part of China's compulsory education syllabus a matter of time: observers](#)

- [The National Security Commission on Artificial Intelligence: - Final Report](#)

India

India

The education system in India is complex and, in general, is jointly regulated by the central government and the various state ministries of education, though there are many local and regional exceptions. Formal education is compulsory for children ages 6-14.[1] The National Curriculum Framework for School Education 2023 was designed by the central government Ministry of Education's National Steering Committee for the Development of National Curriculum Frameworks, formed in 2021.[2] The Central Board of Secondary Education (CBSE) has also played an important role in academic standards and curricula. India recently began the largest program of education reform since establishing a fundamental right to education for children in 2009 and has partnered with private business in the development of curricula that integrate data science and computer programming as core components of primary and secondary school education.[3]

"Recognizing the critical importance of data science and computer science, India, with CBSE's endorsement as a skill subject, is making significant strides in mainstreaming these disciplines in K-12 education. Qubits complements these efforts through comprehensive curriculum, teacher training and robust platform support."
De Paul Kannamthanam
Founder and CEO, Qubits Learning

The adoption of National Education Policy 2020 established the government's goal of developing India's leadership role in fields including "artificial intelligence, machine learning, and data science."[4] The CBSE responded in 2021 by introducing optional 12-hour computer programming modules in grades 6-8 and data science modules in grades 8-12.[5] The grade eight data science module introduces students to qualitative and quantitative data, data visualizations, AI, and data science career pathways.[6] Grade nine students learn the basics of data analysis including data loss and footprints; variables and univariate and multivariate analysis; data collection, sources of data, and an introduction to big data; plots and histograms; and data governance.[7] Statistics is the primary focus of grade 10, which introduces frequency tables, measures of central tendency, frequency distributions, bias, probability, the central limit theorem, percentiles, and z-scores.[8] Grade 11 students learn to make the distinction between correlation and causation and about randomization; sampling bias; observational studies and control groups; searching for data online and XML; and finally begin coding with R and RStudio.[9] In the final module, grade 12 students learn about univariate and multivariate analysis; decision trees; nearest neighbor searches; simple and multiple linear regression; machine learning; Visual Studio Code; and Python.[10] Beginning with grade nine, each module also dedicates a chapter to ethics in data science.

Teaching materials for the modules were created in collaboration with Microsoft India, which had previously worked

1   National Center for Education Statistics, "Digest of Education Statistics: 2015 - Appendix A8," accessed February 24, 2024, National Center for Education Statistics - The Education System in India.
2   Press Information Bureau, "Ministry of Education forms a National Steering Committee for the Development of National Curriculum Frameworks," accessed February 24, 2024, Ministry of Education forms a National Steering Committee for the Development of National Curriculum Frameworks.
3   BBC News, "UK's cyber security challenge revealed," accessed February 24, 2024, BBC News - India launches children's right to education.
4   Ministry of Education, *National Education Policy 2020* (New Delhi: Government of India, 2020), page 15, accessed February 24, 2024, National Education Policy 2020, page 15.
5   The Hindu, "CBSE introduces coding, data sciences as skill modules," *The Hindu,* June 8, 2021, accessed February 24, 2024, The Hindu - CBSE introduces coding, data sciences as skill modules.
6   Central Board of Secondary Education, *Class VIII Data Science Student Handbook* (New Delhi: CBSE, 2022), accessed February 24, 2024, CBSE and Microsoft - Data Science Grade VIII Student Handbook.
7   Central Board of Secondary Education, *Class IX Data Science Student Handbook* (New Delhi: CBSE, 2022), accessed February 24, 2024, CBSE and Microsoft - Data Science Grade IX Student Handbook.
8   Central Board of Secondary Education, *Class X Data Science Student Handbook* (New Delhi: CBSE, 2022), accessed February 24, 2024, CBSE and Microsoft - Data Science Grade X Student Handbook.
9   Central Board of Secondary Education, *Class XI Data Science Student Handbook* (New Delhi: CBSE, 2022), accessed February 24, 2024, CBSE and Microsoft - Data Science Grade XI Student Handbook.
10   Central Board of Secondary Education, *Class XII Data Science Student Handbook* (New Delhi: CBSE, 2022), accessed February 24, 2024, CBSE and Microsoft - Data Science Grade XII Student Handbook.
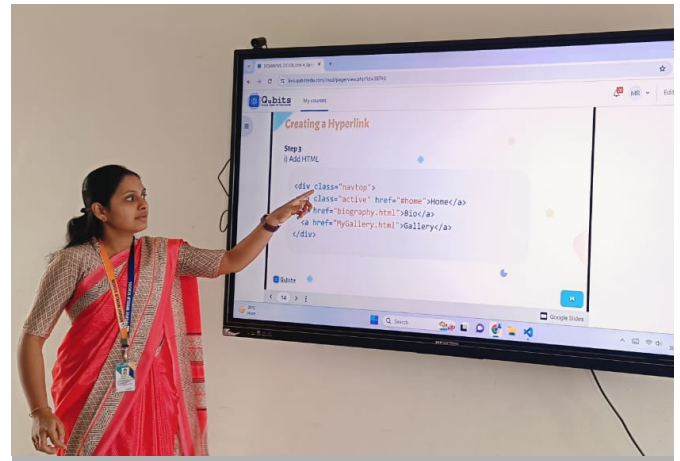
with the CBSE in 2019 to launch a teacher training and classroom technology integration program that promoted cloud computing and AI learning in grades 8-10.[1] The CBSE also collaborated with IBM in 2020 to integrate AI education into the curricula of grades 11 and 12 in approximately 200 schools, and with Intel in 2021 to launch an online plat orm for students grades eight and up to learn skills relating to the development and implementation of AI.[2] The engagement of technology giants in the Indian education system may be in part a result of the Enactment of Companies Act, 2013, which established a "corporate social responsibility" requirement for certain companies to annually contribute 2% of their annual net profit to social welfare activities.[3]

Other initiatives began introducing data science concepts to schools even before the CBSE turned its focus to the development of data science curricula. Atal Innovation Mission, launched by the central government in 2016 to promote a national "culture of innovation and entrepreneurship," has worked with private businesses to establish over 10,000 "Tinkering Labs" in schools that promote STEM-focused learning and provide students with access to electronics and robotics equipment.[4]

Atal Tinkering Labs released learning modules on AI in 2020 that introduce students to fundamental data science concepts. The "Let's Learn Artificial Intelligence Base Module" contains units on AI basics, including ethics and AI's current and future impact on society; machine learning; data, datasets, databases, data visualization, and big data; statistics and probability; algorithms; and coding using Scratch and Python.[5] The "Step-Up Module" extends the Base Module's introduction to AI with project-based learning activities on robotics and image processing.[6] Qubits Learning, an international learning platform and curriculum developer, works with schools to introduce algorithmic and computational thinking at all grade levels through a project-based learning approach. The curriculum focuses on programming, artificial intelligence, and data science with the aim of equipping students with the foundational skills necessary to build a career in computer science.[7]

"If we want to create an equal and just future for everyone, we must focus on making data science available for all. Data has the power to shake up the systems, so let's ensure that this power is shared with those who want to be heard."
Mayuri Dhumal
Data Values Advocate, Global Partnership for Sustainable Development Data

1   Microsoft News, "CBSE, Microsoft collaborate on AI learning for schools," *Microsoft News*, February 16, 2022, accessed February 24, 2024, Microsoft - CBSE And Microsoft join hands to build up capacity for AI Learning for Schools.
2   IBM News Room, "CBSE, IBM collaborate to integrate AI curriculum in school curriculum," *IBM News Room*, July 16, 2020, accessed February 24, 2024, Notification: AI Student Community.
3   Corporate Social Responsibility Portal, accessed February 24, 2024, National CSR Portal - CSR in India.
4   Atal Innovation Mission, "Overview," accessed February 24, 2024, Government of India - Atal Innovation Mission Overview.
5   Atal Innovation Mission, *Let's Learn AI: Base Module* (New Delhi: Atal Innovation Mission, 2022), accessed February 24, 2024, ATAL Innovation Mission - Let's Learn Artificial Intelligence Base Module.
6   Atal Innovation Mission, *Let's Learn AI: StepUp Projects* (New Delhi: Atal Innovation Mission, 2022), accessed February 24, 2024, ATAL Innovation Mission - Let's Learn Artificial Intelligence Step-Up Module.
7   Qbits Learning, "Future Ready CS Curriculum," (Powerpoint, 2024), accessed February 24, 2024, International Powerpoint Presentation.

The number of data scientists in India has grown rapidly since 2017, and by 2022, almost 30% of the world's data scientists were living and working there, according to an industry-wide survey conducted by data science community platform Kaggle.[1] The growth in demand for data scientists in the private sector has led to the swift expansion of data science education in India, especially at the tertiary level and among online learning platform businesses. In 2020, the central Ministry of Education announced "the world's first ever online B.Sc. degree in programming and Data Science" at the Indian Institute of Technology Madras, noting that the government expected the data science field to generate an estimated 11.5 million jobs by 2026.[2] The program was established in response to the government's concern over the number of students pursuing data science education abroad, and it is intended to improve opportunities for students and working professionals regardless of location or academic background.[3] A report by an online education platform predicted striking growth in India's data science education business, from about $200 million in 2023 to almost $1.4 billion in 2028, driven largely by an increase in labor market demands as India captures an increasingly large percent of the global analytics market.[4]

1   Kaggle, *Kaggle State of Machine Learning and Data Science Report 2022* (Location: Kaggle, 2022), page 11, accessed February 24, 2024, Google Cloud - Kaggle DS and ML Survey 2022, page 11.
2   Ministry of Education, *PR-IITM* (New Delhi: Ministry of Education, 2022), accessed February 24, 2024, Union HRD Minister virtually launches World's first-ever Online B.Sc. Degree in Programming and Data Science.
3   Indian Institute of Technology Madras, "Data Science Program," accessed February 24, 2024, Indian Institute of Technology Madras - Data Science.
4   Business Standard, "India's data science education market to reach $1.39 bn by 2028: Report," *Business Standard*, December 27, 2021, accessed February 24, 2024, Business Standard - India's data science education market to reach $1.39 bn by 2028: Report.

Further Reading on Data Science Education in India

- [National Curriculum Framework for School Education 2023](#)

- [Atal Innovation Mission](#)

- [Qubits Learning](#)

- [Indian Institute of Technology Madras - Data Science](#)

Germany

Germany

The education system in Germany resembles the US in some aspects, though a major difference comes after students enter primary school. At the secondary level, German students must select from a variety of education pathways that generally prepare them for either vocational training or tertiary education. While education policy is set primarily at the state level and required curricula can vary between the sixteen *Länder* – federal states – there is also significant cooperation among the states in the development and review of learning standards.[1] The study of data and probability (*Stochastik*) is integrated into national standards that set an obligatory framework for state standards which is actively researched by a professional association of teachers and educators.[2] However, many German educators and education researchers argue too little classroom time is devoted to *Stochastik* and more attention should be paid to issues of data literacy.[3]

"Germany has an active and growing group of probability and statistics educators doing research and development to support schools and teachers in implementing a modern concept of statistics and data literacy. Some flagship projects are trying out more advanced ideas related to AI and data science in experimental classrooms."
**Prof. Dr. Rolf Biehler, Paderborn University
Co-Director, ProDaBi-Project**

Concerned educators and researchers are joined by several initiatives in their belief that data literacy must be a focus of the German education system. The *Stifterverband für die Deutsche Wissenschaft,* a 3,500 member association that provides funding for the promotion of the sciences, published a joint Data Literacy Charter in 2021 whose signatories "express their common understanding of data literacy in the sense of comprehensive data literacy and its overall import-

ance in educational processes."[4] Dorothee Bär, then-State Minister in the Federal Chancellery for Digitalization, said of her support for the Charter that "data literacy is one of the ABCs of the digital age."[5]

TrainDL is a policy and research project that "aims to develop educational concepts for data literacy and artificial intelligence competencies and embed them in teacher and school education" by bringing researchers and decision makers together to exchange experiences, transfer knowledge, and set a common policy agenda.[6] Begun in 2021 and funded by the European Union's Erasmus+ program for education, training, youth, and sport, the project is organized jointly by 11 partner organizations in Germany, Lithuania, and Austria, including the German Informatics Society; the Lithuanian National Education Agency; and the Austrian Federal Ministry of Education, Science, and Research, among others.[7] The project organized online and in-person workshops and sessions on policy, AI and data literacy in education, and teacher training. Additionally, they invited representatives from enterprise software developer SAP, Code.org, the European Commission, and UNESCO to address attendees at the 2024 Summit on Teaching Data Literacy and AI Competencies, which also included workshops on policy dialogue and teaching materials.[8]

TrainDL has released a series of reports, research articles, and visualizations in three discrete "clusters": policy, education, and evaluation. TrainDL's foundational policy research – including the identification of relevant policies in Germany, Austria and Lithuania, and a review of the curricula of several German states to determine the extent to which they included material on data literacy and AI – culminated in the development of an interactive online policy monitor.[9] The monitor's primary feature is a map that responds to user-selected filters on policy dimensions, including target audience and policy type, providing for a visual window into the adoption of AI and data literacy education policies in Germany, Lithuania, and Austria.[10] TrainDL has also released a set of draft policy recommendations and implementation mechanisms, including that:

1   Eurydice, *Bildungswesen: Structures of Education and Training Systems in Europe* (Brussels: Eurydice, 2022), page 240, accessed February 24, 2024, The Education System in the Federal Republic of Germany 2019/2020, page 240.
2   Stochastik in der Schule," accessed February 24, 2024, Stochastik in der Schule.
3   Rolf Biehler, discussion with the author, January 2024.
4   Stifterverband, *Data Literacy Charter* (Berlin: Stifterverband, 2022), accessed February 24, 2024, Data Literacy Charter.
5   Stifterverband, "Charta Data Literacy," accessed February 24, 2024, Data Literacy Charter.
6   "train-dl.eu," accessed February 24, 2024, TrainDL.
7   "About Train-DL Consortium," train-dl.eu, accessed February 24, 2024, TrainDL - Consortium.
8   "Events," train-dl.eu, accessed February 24, 2024, TrainDL - Events.
9   "Train-DL Deliverable 1.1," train-dl.eu, accessed February 24, 2024, D1.1 - Policy Research Summary and D5.1 - Research of Policies/Curricula.
10   "traindl-policymonitor.ocg.at," accessed February 24, 2024, TrainDL - Policy Monitor.

- "people should be familiarized with the basic concepts of AI and acquire basic data skills at an early stage of school education (latest secondary I), because AI and data are already affecting many aspects of people's lives"; and that
- "teacher trainings need to be designed and implemented to provide the confidence to teach these topics whilst also making clear that teachers do not need to have an all-encompassing AI and data literacy knowledge."[1]

Several organizations in Germany advocate for education in data science topics by bringing them directly into classrooms as optional modules. *KI macht Schule* ("AI Goes to School,") has been promoting machine learning and AI education with a focus on ethical and social implications since 2020.[2] Composed largely of volunteers working in local groups across eight cities, the program brings courses on AI into secondary school classrooms and has also created modules on specialized topics such as AI in medicine, transportation, and art. Recognizing that AI and machine learning are typically taught as highly technical subjects at the university level, *KI macht Schule* plans lessons that require minimal prior knowledge and seeks to reach students who otherwise might not have opportunities to learn about AI or machine learning in the classroom.[3]

The Computational and Mathematical Modeling Program (CAMMP) was founded in 2011 to raise public awareness "of the importance of mathematics and simulation science for society as a whole."[4] CAMMP's half-day workshops for secondary school students on topics such as image classi-

fication and machine learning, activity recognition and AI, climate data, spam filters, and Google's search algorithms, emphasize the application of mathematics and the analysis of data needed to answer questions and address problems in everyday life.

Projekt Data Science und Big Data in Der Schule (ProDaBi) – Project Data Science and Big Data at School, administered through Paderborn University in the state of North Rhine-Westphalia, has worked since 2018 to develop data science "teaching units for secondary level and professional development courses for teachers in collaboration with teachers and facilitators."[5] The project evolved from the Paderborn Symposium on Data Science Education at School Level 2017, which was initiated and funded in part by the *Deutsche Telekom Stiftung*, a large corporate foundation that supports youth STEM education and that continues to provide support for ProDaBi (T-Mobile US is a subsidiary of *Deutsche Telekom*).[6] ProDaBi uses contemporary research on best practices for the development of data science curricula to design learning modules for grades 5-12. It also conducts and implements its own design-based curricular research in regional computer science classrooms. Current modules include:

- a year-long project focused on data exploration, machine learning, and conducting a data science project;
- a data science and artificial intelligence module (with and without programming);

1   "Train-DL Deliverable 5.3," train-dl.eu, accessed February 24, 2024, TrainDL - D5.3 - Development of Policy and Curricula Recommendations (CS Teachers), pp. 6-7.
2   "Vision," ki-macht-schule.de, accessed February 24, 2024, KI macht Schule - Vision.
3   "Diversity," ki-macht-schule.de, accessed February 24, 2024, KI macht Schule - Diversity.
4   "The idea behind CAMMP," cammp.online, accessed February 24, 2024, CAMMP - About Us.
5   "prodabi.de," accessed February 24, 2024, ProDaBi-Project.
6   "Data Science School," Telekom-Stiftung.de, accessed February 24, 2024, Deutsche Telekom Stiftung - Data Science in School.

- a data awareness module that explores data on both the individual and societal levels;
- several projects centered on collecting and analyzing environmental data using sensors;
- an exploration of location data to help students understand where, how, and why data is collected and processed; and
- a module focused on teaching nutrition with food data cards and understanding machine learning using decision trees.

ProDaBi collaborates with the Concord Consortium and implements CODAP in several learning modules, while advanced students with programming experience utilize Python-based Jupyter Notebooks. ProDaBi also offers self-study and professional development courses that introduce computer science teachers to state curriculum-aligned teaching units and materials developed in collaboration with regional teachers.

In addition to its curricular research and development work, the project also organizes workshops, lectures, and an ongoing Colloquium on Data Science and Artificial Intelligence in School.[1] Currently in its third phase, the project is turning its focus to the integration of data and AI literacy in STEM education and the inclusion of data on topical subjects such as climate change.

---

1   "Colloquium," prodabi.de, accessed February 24, 2024, ProDaBi - Colloquium.

Further Reading on Data Science Education in Germany

- *Stifterverband für die Deutsche Wissenschaft:* Data Literacy Charter

- TrainDL

- TrainDL-Interactive Policy Monitor

- Computational and Mathematical Modeling Program (CAMMP)

- CAMMP List of Publications and Talks

- *KI Macht Schule*

- Projekt Data Science und Big Data in Der Schule (ProDaBi)

- ProDaBi Publications

- *Deutsche Telekom Stiftung:* Data Science in school

- Paderborn Symposium on Data Science Education at School Level 2017: The Collected Extended Abstracts

- Computing Education Research Group: Publications

New Zealand

New Zealand

The country of New Zealand – called *Aotearoa* in Māori – is a bicultural unitary state whose national government is responsible for many aspects of primary and secondary school education curricula and standards, which are implemented by local school boards.[1] The National Curriculum is composed of the English-medium New Zealand Curriculum and *Te Marautanga o Aotearoa*, the Māori-medium curriculum; this report addresses the New Zealand Curriculum.

Statistics is a key focus of education in New Zealand primary and secondary schools and has been a learning strand at all levels of compulsory education since 1992.[2] The New Zealand curriculum makes a distinction between mathematics ("the exploration and use of patterns and relationships in quantities, space, and time") and statistics ("the exploration and use of patterns and relationships in data") and states that they "are related but different ways of thinking and of solving problems."[3]

Education in New Zealand is currently in a period of substantial change: the Ministry of Education is in the process of a "Curriculum Refresh," and many schools are required to begin using the new mathematics and statistics curriculum at the start of 2024.[4] Additionally, the New Zealand Qualifications Authority is working in collaboration with the Ministry and others in the education sector to replace existing standards for the National Certificate of Educational Achievement, the secondary school qualification and a prerequisite for university enrollment.[5]

In 2021, an Expert Advisory Panel on Mathematics and Statistics was convened at the Royal Society *Te Apārangi* to "provide advice to the Ministry of Education on refreshing the English-medium Mathematics and Statistics learning area of the New Zealand Curriculum."[6] In their report, the Panel identified three trends indicating that "the future of mathematics and statistics requires computational thinking":

- "the growth in computational resources, both in the capacity of computers to handle and process data, and in the expansion of computer-intensive methods in mathematics and statistics";
- "the exponential growth in available data, and the rapid change in access and type of available data (big data)"; and
- "the increase in interest, understanding, and place of artificial intelligence (AI) in society."[7]

"The amount of data being created and consumed is increasing exponentially, as are associated digital technologies, and teaching students new computational ways of thinking about data and modelling to be used alongside statistical thinking will support them to understand and interpret modern data."
Dr. Anna Fergusson
Lecturer, Department of Statistics,
Faculty of Science, University of Auckland

Statistics education in New Zealand is not limited to basic probability and measures of central tendency as in some US curricula. Students learn about "identifying problems that can be explored by the use of appropriate data, designing investigations, collecting data, exploring and using patterns and relationships in data, solving problems, and communicating findings" in addition to "interpreting statistical information, evaluating data-based arguments, and dealing with uncertainty and variation."[8] The Curriculum Refresh goes even further by enumerating concepts foundational for the development of data science skills that students are expected to understand through a multicultural lens:

1   The Legislative Framework and NZC: 2023 Dec update, PDF file, accessed February 24, 2024, New Zealand Ministry of Education - The Legislative Framework and the New Zealand Curriculum Learning Area Structure and The Encyclopedia of New Zealand - Biculturalism.
2   Forbes, Sharleen. "An algorithmic approach to target selection for social network analysis." *Statistical Analysis and Data Mining: The ASA Data Science Journal* 7, no. 6 (2014): 443-464, accessed February 24, 2024, The coming of age of statistics education in New Zealand, and its influence internationally.
3   "What is mathematics and statistics?," The New Zealand Curriculum, accessed February 24, 2024, The New Zealand Curriculum Online - Mathematics and statistics.
4   "The New Zealand Curriculum," accessed February 24, 2024, Ministry of Education - The New Zealand Curriculum.
5   "Review of Achievement Standards," NCEA (National Certificate of Educational Achievement), accessed February 24, 2024, NCEA Education - Review of Achievement Standards.
6   "Pāngarau: Mathematics and Tauanga: Statistics in Aotearoa New Zealand," Royal Society Te Apārangi, accessed February 24, 2024, Royal Society Te Apārangi - Pāngarau Mathematics and Tauanga Statistics in Aotearoa New Zealand: About the Panel.
7   "Pāngarau Mathematics and Tauanga Statistics in Aotearoa New Zealand: Advice on refreshing the English-medium Mathematics and Statistics learning area of the New Zealand Curriculum," Expert Advisory Panel, Royal Society Te Apārangi, September 2021, accessed February 24, 2024, Pāngarau Mathematics and Tauanga Statistics in Aotearoa New Zealand.
8   "Learning area structure," The New Zealand Curriculum, accessed February 24, 2024, New Zealand Curriculum Online: Mathematics and Statistics.
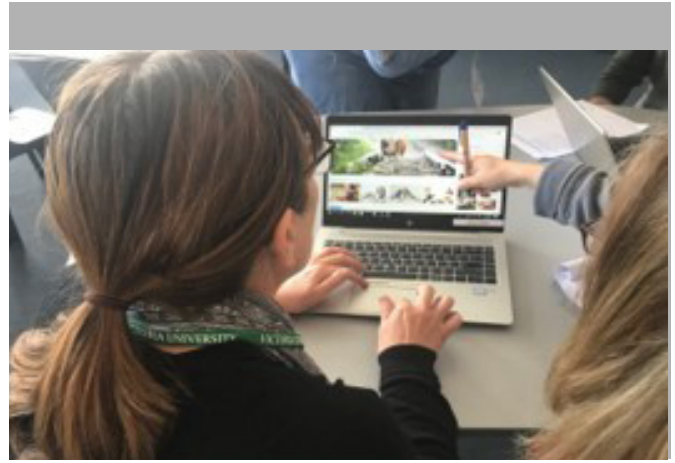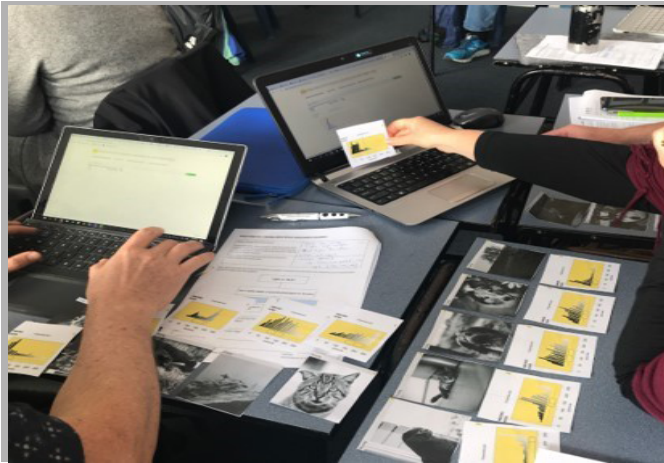
- "The world is full of patterns and structures that we use mathematics and statistics to understand."
- "The world is characterized by change and variation that we use mathematics and statistics to understand."
- "Mathematical and statistical logic and reasoning enable us to identify and explain relationships and to justify conclusions."
- "The interface between *mātauranga Māori* and *mātauranga* mathematics and statistics offers opportunities for insights that uphold the integrity of each knowledge system."
- "Mathematics and statistics have a continuous, evolving human history."[1]

Institutions and professional organizations are significant contributors to New Zealand's strong history of statistics education and ongoing effort to integrate data science into compulsory education. The New Zealand Statistical Association's Education Committee published a statement in 2021 that included an overview of international trends in data science education and curriculum development, an analysis of New Zealand's own learning standards, and recommended changes to statistics and data science education at the senior secondary level.[2] The Committee followed up in 2022 with specific recommendations for the Curriculum Refresh that acknowledged the increasing importance of evidence-based thinking for citizens and suggested the inclusion of "ethics, algorithmic bias, new data visualizations, [and] indigenous data sovereignty" in the new curriculum.[3]

The University of Auckland's Department of Statistics is the birthplace of the R programming language, which was developed for statistical analysis and is currently used by more than 2 million statisticians and data scientists, placing it among the most popular programming languages in the world.[4] The department's faculty have also been enormously influential in data science and statistics education.[5]

The University and faculty also lead the CensusAtSchool New Zealand project, a voluntary survey and curriculum-based learning activity that introduces primary and secondary school students to data science, probability, and statistics concepts with real data applicable to – and generated by – the students themselves. Over 30,000 students annually take part in the project and answer questions about their school, opinions, and activities.[6] Some survey data is made available to teachers at a classroom-level, while anonymized datasets are available for download and analysis in statistical software or directly via iNZight, a data visualization platform that utilizes R and was originally designed for use in New Zealand secondary schools.[7] The project also publishes and maintains a list of curriculum-aligned statistics teaching resources for students of every grade level.[8] CensusAtSchool New Zealand is part of an international education project established by the Royal Statistical Society in the UK; other participant countries include Australia, Canada, Ireland, Japan, South Africa, Korea, and the US.[9]

1   "Te Mātaiaho The Refreshed New Zealand Curriculum: The learning areas of Mātaiaho Mathematics and statistics," accessed February 24, 2024, Te Mātaiaho The Refreshed New Zealand Curriculum page 28.
2   "NZSA Statistics and data science at CL7_8," accessed February 24, 2024, Statement from the New Zealand Statistical Association Education Committee.
3   "NZSA Curriculum Refresh & Statistics," accessed February 24, 2024, Curriculum Refresh and Statistics - specific changes: A curriculum in statistics for changing times.
4   "Why R programming language still rules Data Science?" Project Pro, accessed February 24, 2024, Why R programming language still rules Data Science?.
5   "Contributors to R," R Project, accessed February 24, 2024, The R Project for Statistical Computing - Contributors.
6   "CensusAtSchool New Zealand," accessed February 24, 2024, CensusAtSchool New Zealand 2023 Survey.
7   "About Inzight," Inzight, accessed February 24, 2024, About iNZight.
8   "Resources," CensusAtSchool New Zealand, accessed February 24, 2024, CensusAtSchool New Zealand - Resources for Teaching Statistics.
9   "About," CensusAtSchool New Zealand, accessed February 24, 2024, CensusAtSchool New Zealand - About.

Further Reading on Data Science Education in New Zealand

- New Zealand Ministry of Education - Refreshing The New Zealand Curriculum

- Royal Society *Te Apārangi*

- *Pāngarau* Mathematics and *Tauanga* Statistics in *Aotearoa* New Zealand

- New Zealand Statistical Association Education Committee

- Statistics and Data Science Educator

- University of Auckland - Statistics Education Research

- CensusAtSchool New Zealand

- iNZight

Additional Countries

Australia

A 2023 report by the Australian Mathematical Sciences Institute and the Statistical Society of Australia said that "increasing student engagement with intermediate and higher-level mathematics subjects [...] in senior secondary school is seen as vital to grow the future data science workforce."[1]

Colombia

According to reports by workforce development platform Correlation One, which also runs the Data Science for All fellowship program, "Colombia has been investing heavily in Data Science & Analytics training for its citizens" since 2019, including by "training more than 15,000 Colombians in Data Science."[2]

Estonia

Among the highest performers on the PISA 2022 mathematics exam, Estonia became the first country to pilot the curriculum of the Computer-Based Math project in 2013. Computer-Based math's curriculum emphasizes computational thinking and the use of computers in primary and secondary school mathematics education."[3]

Finland

Finland's university students ranked first globally in data science skills in the 2021 issue of Coursera's annual Global Skills Report, which stated that "a recent push by its government and the University of Helsinki to teach 1% of the world basic AI skills."[4]

Israel

The Israeli Ministry of Education's Data Analysis High School Program provides grades 10-12 students across 88 high schools with eight to twelve hours per week of instruction on data analysis, visualizations, data ethics, statistics, and AI. The program emphasizes real-world data and integrates industry exports into classrooms as mentors.[5]

Japan

Recognizing the value of data science skills across all disciplines, the Japanese Ministry of Education, Culture, Sports, Science and Technology founded the Japan Inter-University Consortium for Mathematics, Data Science and AI Education was founded in 2017 to "to develop and spread the effective educational environment of this subject for every student, irrespective of his/her field of specialization, in higher education."[6]

"As we look to the future, data storytelling will become an indispensable skill in the job market, necessitating its integration into K-12 education today. Teaching our students how to craft compelling narratives from data will equip them to lead with insight and innovate in a data-driven world."
Shamsudeen Mustafa
Co-Founder & CEO at Correlation One

The Netherlands

The 2019 Dutch Digitisation Strategy 2.0 acknowledges that "the Netherlands must invest in people with specialised knowledge of fields such as AI and data science."[7]

Poland

A recent report on education in Poland notes that, regarding curricula in grades 4-8, "key skills to be developed by informatics contains a section on algorithmic thinking, problem-solving using computers, computational thinking, and game-based learning."[8] A 2019 memorandum signed by the Polish Ministries of Digital Affairs; Investment and Development; Science and Higher Education; and Entrepreneurship and Technology "calls for coordinated actions in order to equip citizens (not only students but also adults) in knowledge in the field of data science."[9]

1   "AMSI-SSA Data Science Review November 2023," Statistical Society of Australia, accessed February 24, 2024, Australian Mathematical Sciences Institute-Statistical Society of Australia - Data Science Review, page 4.
2   "Unleashing Colombia's Potential in the Data Economy," (Correlation One, 2023), 9.
3   "Computer-Based Math in Estonia," Wired, February 19, 2013, Wired - Math Rebels Invade Estonia With Computerized Education.
4   "Coursera: Finland world's top data science nation," University of Helsinki, accessed February 24, 2024, Coursera: Finland is the world's top data science nation.
5   "About Us," Data Analysis Program, Isreaeli Ministry of Education, accessed February 24, 2024, Data Analysis High School Program.
6   About Us," Mathematics Interdisciplinary Research Consortium, University of Tokyo, accessed February 24, 2024, Japan Inter-University Consortium for Mathematics, Data Science and AI Education.
7   "English Version of the Dutch Digitalisation Strategy 2.0," Nederland Digitaal, accessed February 24, 2024, English version of the Dutch Digitalisation Strategy 2.0.
8   Panskyi, T., Korzeniewska, E., Serwach, M. et al. New realities for Polish primary school informatics education affected by COVID-19. Educ Inf Technol 27, 5005–5032 (2022), New realities for Polish primary school informatics education affected by COVID-19.
9   Ponikowska, M. (2020). Vocational education and training for the future of work: Poland. Cedefop ReferNet thematic perspectives series. Page 28, Vocational education and training for the future of work: Poland. Cedefop ReferNet thematic perspectives series.

Singapore
The Singapore Ministry of Education's 2024 Pre-University Higher 1 mathematics course syllabus, "designed for students who intend to pursue further studies in business and social science courses," suggests that students should be exposed to applications of mathematical concepts including topics such as correlation and regression analysis, normal distributions and probability, and hypothesis testing.[1] The syllabus emphasizes the use of computers in mathematics and encourages the development of computational thinking among secondary students.

South Korea
The South Korean Ministry of Education released a policy document in 2020 announcing the rollout of an AI subject curriculum that includes instruction on "programming, basic principles of AI, AI utilization, and AI ethics" with plans to complete K-12 implementation nation-wide by 2025.[2]

1   Ministry of Education, Singapore, *2024 Pre-University H1 Mathematics* (Singapore: Ministry of Education, 2024), 16, Singapore Ministry of Education Curriculum Planning and Development Division - Syllabus 8865.
2   Asia Pacific Foundation of Canada, *AI K-12 Education Report,* (Vancouver: Asia Pacific Foundation of Canada, 2021), 14, Asia Pacific Foundation of Canada - Talent for the Future: AI Integration for K-12 Students in Canada and South Korea, page 14.

Further Reading on Data Science Education in Additonal Countries

- International Data Science in Schools (IDSSP) Project

- United Nations Educational, Scientific and Cultural Organization (UNESCO) - K-12 AI curricula: a mapping of government-endorsed AI curricula

- UNESCO: Beijing Consensus on Artificial Intelligence and Education: Outcome document of the International Conference on Artificial Intelligence and Education

- Saman Rizvi, Jane Waite, Sue Sentance: Artificial Intelligence teaching and learning in K-12 from 2019 to 2022: A systematic literature review

- Ben-Zvi, Dani, Katie Makar, and Joan Garfield: International Handbook of Research in Statistics Education

- Donoho, David: 50 Years of Data Science

Canada has seen the introduction of data science topics in several provincial curricula.

Colombia has trained more than 15,000 people in data science since 2019.

The UK has invested £661 to establish the "Data Capital of Europe."

The Netherlands established the Dutch Digitisation Strategy 2.0 in 2019.

Germany is producing significant research in Data Science education and policy.

Israel has a data analysis program in 88 high schools.

India introduced data science learning modules in grades 8-12.

Australia is researching ways to grow their data science workforce.

New Zealand is the birthplace of the R programming language.

Finland ranked first globally in data science skills in a recent report.

Estonia first piloted the Computer-Based Math curriculum in 2013.

Poland has called for action to equip students and adults with data science skills.

China released a textbook on AI for high school students in 2018.

Singapore secondary school math curricula include a focus on computational thinking.

South Korea is implementing an AI subject curriculum in K-12.

Japan founded the Inter-University Consortium for Mathematics, Data Science and AI Education in 2017.

DataScience4
everyone

March 2024