This paper was commissioned for the *Workshop on Foundations of Data Science for Students in Grades K-12*, whose work was supported by the Valhalla Foundation. Opinions and statements included in the paper are solely those of the individual authors, and are not necessarily adopted, endorsed, or verified as accurate by the *Workshop Planning Committee on Foundations of Data Science for Students in Grades K-12*, the Board on Science Education, the Board on Mathematical Sciences and Analytics, the Computer Science and Telecommunications Board, or the National Academy of Sciences, Engineering, and Medicine.

Previewing the National Landscape of K-12 Data Science Implementation

Zarek Drozda, University of Chicago
Davis Johnstone, Florida State University
Brooke Van Horne, University of Michigan

Data is undeniably changing our world. The rise of "big" or complex data, increasing computational power, and global connectivity have given rise to the emergent field of data science, increasingly distinguished from related fields both in technique and scope (Donoho, 2017). These changes have also changed the nature of daily career, personal, and civic engagement. Stakeholders in education (National Research Council, 2012) and the popular press (Freakonomics, 2019) have advocated that *all* students should learn about modern data methods and uses. Data education researchers have expressed that we are in a "state of emergency" with the speed and scale of these technological transformations, especially given students now encounter data several times a day (National Center for Education Research, 2021). How has K-12 education responded? This paper will provide a preview of the national landscape of K-12 data science implementation through existing frameworks and policy, case studies on implementation models from across the United States, and educator experiences.

Content Frameworks

Content expectations across multiple school subjects in U.S. primary and secondary education already incorporate at least some learning about data collection, utilization, and analysis. Data-related concepts consistently appear in mathematics, science, computer science, and social studies across states. These existing standards may provide the building blocks or even partially comprise a *data science* education. However, stakeholder interviews cautioned the degree to which these concepts are actually taught during the school day may vary widely – meaning there is no guarantee students are necessarily learning the concepts listed here. Significant technological changes and the speed of digital transformation have also created a challenging moving target for curricular relevance. Recent efforts both domestically and internationally have attempted to articulate discrete data science learning experiences, primarily for secondary students. Both existing and new frameworks relevant to data science education include, but are not necessarily limited to:

- **Next Generation Science Standards (NGSS),** built from *A Science Framework for K-12 Science Education* (2011), outlines eight practices of science and engineering that are essential for all students to learn. NGSS standards expect students to demonstrate both topic knowledge and skills specific to each practice area concurrently. (NGSS, 2013, 1). Especially relevant practices of the NGSS framework include Practice 4 ("Analyzing and interpreting data") and Practice 5 ("Using Mathematics and Computational Thinking") (ibid), both included in Appendix F – Science and Engineering Practices. In K-2, students learn basic skills like "collecting, recording, and sharing observations." These practices are built consistently over time so that students are comfortable with detailed statistical analysis, the comparison of data sets for consistency, and the use of models to generate and analyze data" by high school (NGSS, 2013, 9). Notably, students are expected to **"use digital tools (e.g., computers) to analyze very large data sets for patterns and trends")** as early as Grades 6-8 in Practice 5.

- **Common Core State Standards for High School: Statistics and Probability** are an important building block for data science education. The standards consist of four main learning goals: 1) Interpreting Categorical and Quantitative Data, 2) Making Inferences and Justifying Conclusions, 3) Conditional Probability and the Rules of Probability, and 4) Using Probability to Make Decisions (Common Core, 2022). While data science and statistics are not necessarily synonymous; some data science curricula use these standards as a foundation. For example, the University of California, Los Angeles CenterX created their Introduction to Data Science (IDS) Curriculum to introduce students to "dynamic data analysis" through the lens of these four learning outcomes (University of California, Los Angeles, 2020). Importantly, UCLA also incorporated the **Computer Science Teachers Association's (CSTA) K-12 Computer Science Standards** into their curriculum (ibid).

- **K-12 Computer Science Framework** was developed in 2017 between a field-wide partnership including Code.org, the Computer Science Teachers Association, the National Math + Science Initiative, and many others (K-12 Computer Science Framework, 2016). The framework includes four sub-concepts outlining Data and Analysis from a Computer Science lens: 1) Data Collection 2) Data Storage 3) Visualization and Transformation and 4) Inference and Models. The K-12 standards in Computer Science emphasize the management, treatment, and computational efficiency for analyzing data, and equip students to think about accuracy of computer models as a function of the quality and quantity of data – especially in the context of prediction and automation of data analysis (i.e. machine-learning). The standards also encourage students to think carefully about how data is presented, and how visual presentation can influence conclusions.

- **College, Career & Civic Life (C3) Framework for Social Studies Standards** was created in 2013 to enhance the quality of social studies instruction and increase focus on civic education. The C3 framework mentions "data" over 53 times throughout the standards. Data appears in a few distinct contexts: 1) as one component of historical evidence (along with primary sources, documents, text, images, and other artifacts), 2) as domain-relevant information for geography or economics and 3) in the context of geospatial technology and geographic information systems (GIS) (National Council for Social Studies, 2013). In the formal standards, data is expected to be integrated into learning pathways on the National Economy, Geographic Representations, Communicating & Critiquing Conclusions, as well as companion frameworks for introductory high school courses in Psychology and Sociology.

- **Pre-K–12 Guidelines for Assessment and Instruction in Statistics Education II (GAISE II):** in 2020 the American Statistical Association and the National Council of Teachers of Mathematics created an updated version of GAISE I, originally published in 2005 and revised in 2007. The new document centers around the idea that data is of growing importance in all facets of society, and recommends "statistical literacy for all" (Bargagliotti et al. 2020, 3). The paper establishes a two-dimensional framework consisting of the statistical problem-solving process applied to three different levels of statistical literacy (A, B, and C) (Bargagliotti et al., 2020, 13-19). GAISE II is very similar to the NGSS in that it provides concrete practices and concurrently aims to develop a deeper theoretical understanding of statistics.

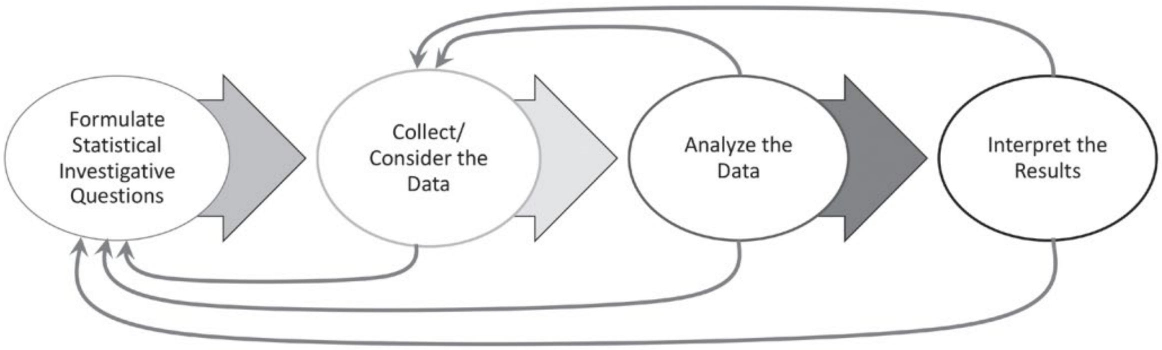

Figure 1: GAISE II's statistical problem-solving process (Bargagliotti et al. 2020, 13).

- **"Data Science K-10 Big Ideas,"** developed by Jo Boaler of Stanford's Graduate School of Education and Rob Gould of the University of California, Los Angeles, provides an accompanying extension of the GAISE II report to provide guidance for data science and data literacy prior to an advanced high school course. The guidelines also help teachers connect data science concepts to content they are already teaching in their classrooms, through curated tasks and "Data Talks" for a younger audience (ibid). Some early curriculum, including **youcubed**®'s ecosystem of courses, lesson plans, and K-5 activities, take inspiration from GAISE II (Stanford Graduate School of Education, 2022).

- **International Data Science in Schools Project (IDSSP)** was a cross-country and cross-sector collaboration—including the International Statistical Institute and Google—to outline a vision for a distinct data science training in secondary education, completed in September 2019 (Burr, 2022). The international working group created two curriculum frameworks for data science education, which they intend to be compatible with a variety of delivery formats (classroom, self-learning, etc.):
    1. A pre-calculus course of some 200-240 hours for students in their last two years of high school.
    2. A curriculum to teach teachers from a variety of backgrounds how to teach Data Science to students (ibid).

  The working group emphasized that all students—both those intending to further their studies in data science and those intending to pursue other careers—should understand and appreciate how data interacts with their daily lives, and should use this knowledge to make more informed decisions (ibid). The IDSSP is more explicitly focused on a "discrete" data science education than the Common Core, NGSS, and GAISE II, which address data science through the lens of other subjects.

- **Launch Years Data Science Course Framework.** Led by the University of Texas at Austin's Charles A. Dana Center, the Launch Years initiative aims to transform secondary math education to better suit the needs of students and align to more diverse post-secondary pathways (The University of Texas Dana Center, 2022). The course framework for data science focuses on articulating design principles for other curriculum providers, including active learning, authenticity, interdisciplinary connections, effective communication, and incorporation of technology (The University of Texas Dana Center, 2021). The framework also gives high-level guidance on student learning outcomes, organizing topics thematically ("collect and manage data") rather than discrete skills (i.e. data merging) (ibid). To date, the data science course framework accompanies a preparatory college math course that encompasses quantitative reasoning, statistics, and algebra, as well as a modern Algebra 2 course that includes data science principles (ibid).

Commonalities: both subject-specific and discrete *data science* learning frameworks attempt to clearly demonstrate *the value* of data to students, typically through example for younger students and then explicitly by high school (e.g. measuring information to enable calculations in mathematics, making valid and reliable claims in science, corroborate or gain otherwise inaccessible information in history). Conversely, all frameworks also attempt to outline possible limitations of data and the questions it can answer, mediated by measurement error (NGSS, P4, 9-12) and sample bias (ibid), representativeness of a population (GAISE II, IV-B), limits in complexity as controlled by models and computational power (K-12 CS, 9-12), and constraints in study design and existing biases (Dana Center). Given the strong potential for relevance and student engagement associated with data science learning, explicitly proving the value of data science to students up-front may be of particular importance, with "room" in existing standards to do so.

In STEM subjects (e.g. math, science, CS), the authors identified 21 shared foundational data concepts articulated across all subjects in existing frameworks. Students either gain methods, basic knowledge, or brief exposure to these concepts, which include data collection, data tables, basic measures of variability (mean, median, mode), data visualization (dot plots, histograms, box plots), data types (qualitative vs. categorical), functions and model fit, and sample selection. Additional concepts shared across at least three or more frameworks include data collection from or with modern technology devices, Bayesian probability, outliers in data, random sampling, and data preparation or "cleaning." A full list of these concepts and their appearance can be found in the appendix. These concepts may not fully comprise a *data science* education in their current form, yet their consistency suggests promise for an interdisciplinary approach across existing K-12 school subjects.

Major Differences: differences quickly emerge beyond a short list of particular competencies. We identify three major points of comparison, with many others possible:

<u>Age of Introduction:</u> even within the 21 shared data-related concepts, the age at which students first encounter topics differs dramatically across subjects. In one particularly divergent example, students may encounter histograms as early as K-2 in

computer science (K-12 CS, K-2), but may not do so until either 7th grade or even high school in mathematics (CCSS, Statistics & Probability), instead generating simple bar graphs manually in earlier grades (CCSS, 3). In another example, students may differentiate between qualitative and qualitative data types as early as 2nd grade in a science context (NGSS, P5, K-2) or in statistics (GAISE, A), but may would not gain distinctions between character vs. numeric data until at least middle school in computer science (K-12 CS, 6-8). As a generalization, science standards are generally most ambitious in attempting early introduction for a wide variety of data-related concepts (data collection, statistical question formation, data visualizations, etc.), followed by computer science, and then mathematics. Given GAISE does not guide by numeric grade-levels, the timing of additional statistics education is highly contextual.

While these differences may be less meaningful in the long-term, other critical concepts likely appear too late and with too little repetition for building data acumen that lasts. Correlation vs. Causation appears quite late in mathematics (CCSS, High School Statistics & Probability) (*likely an optional elective), in computer science (K-12 CS) (*a brief treatment in predictions), and even statistics as an explicit concept (GAISE II, C). Science teachers seem to be the lone wolves creating this intuition in middle school (NGSS, P4 6-8) in existing frameworks. Modeling and Model Fit is also granted little attention until late in high school, including in science (NGSS, PF, 9-12) and in mathematics (CCSS, High School Statistics & Probability) – with an earlier treatment appearing in computer science (K-12 CS, 6-8). Finally, Bayesian / probabilistic thinking, important for both everyday use of data and modern algorithms, is similarly given late treatment in both science (NGSS, P4 9-12) and mathematics (CCSS, High School Statistics & Probability). As the GAISE II states, "sound statistical skills take time to develop. They cannot be honed to the level needed in the modern world through one high school course." As data science education takes form, researchers and policymakers will need to identify the critical topics that are more consistent across subjects in elementary and middle school.

Competencies (a progression) vs. Cycles (iterative): pedagogical distinctions in data-related concepts differ drastically between existing subject frameworks and dedicated data science guidelines, mostly a function of focus. CCSS Mathematics, NGSS Appendix F Practices, and K-12 CS frameworks all approach data-related concepts through general buckets of topics that are reinforced across time (i.e. repeat by each grade-level). In contrast, both GAISE II and discrete data science frameworks (IDSSP, "Big Ideas in K-10", and the Dana Center's course framework) heavily emphasize iteration, cycles, and "habits of mind" related to the study of data, which may be re-introduced every class. The latter group frequently cites a "data cycle," arranged differently but in each case emphasizing the study of data to be a discovery process rather than distinct and separate practices. The IDSSP emphasizes that educators should "get students, habitually, to ask questions worrying about quality and applicability" of data for a given question. Similarly, GAISE II emphasizes that "anyone who uses data – be more than just data crunchers. They should be data problem solvers who interrogate the data and utilize questioning throughout… to make decisions with confidence." Data science may be best conceived of as a process, rather than a subject or knowledge set.

Spectrum of Theory vs. Practice: mathematics & statistics frameworks provide theory for the treatment of data, while computer science provides mostly practice in the forms of technology and computational training. Both appear necessary but insufficient for early data science frameworks, which combine them concurrently, along with methodological complexity that can be customized to a given problem, use-case, or grade-level. For two examples, the K-12 CS Framework offers little training for how to handle outliers in data; while CCSS Mathematics gives limited direction on how to apply conditional probability beyond re-arranging formal probability expressions. In contrast, NGSS guidelines fall near the middle of the spectrum, both introducing formal theory (measures of variability; measurement error, correlation vs. causation) while also encouraging analysis of "data to define an optimal operational range," "to refine a problem statement or the design of a proposal object," or the "impact of new data on a working explanation," showing how theory can be applied for particular problems (NGSS, P4 / P5). NGSS has particular strength in that it identifies use-cases, encourages practice, and imparts theory to build scientific habits – similar to the process-oriented guidelines outlined in IDSSP. Unfortunately, these strengths risk obscurity as an Appendix in NGSS rather than as an explicit course, semester, or module in the school day. Stakeholder interviews revealed early curricula are not necessarily prioritizing all 8 practices, but may instead pick one or two.

What May be Missing: stakeholder interviews revealed several topics or pedagogical strategies that may be critical for a meaningful *data science* education, which may be excluded from both the subject-level building blocks and even discrete data science frameworks. We briefly outline a few here:

Technology as Critical: despite the promise of "unplugged" computer science experiences (Battal et al., 2021), this review found that data science experiences will likely need to rely on computer devices and accessible data software. Teacher interviews consistently revealed that student engagement appeared to be a function of 1) the opportunity to learn through technology 2) the ability to find and select datasets of personal interest online and 3) the ability to convert data into authentic digital communication formats. Moreover, both the IDSSP ("uses modern technology, a prerequisite for developing any significant capabilities in Data Science") and NCTM's *Catalyzing Change* report ("technology is driving changes that should be reflected in the high school mathematics curriculum...students should have opportunities to use dynamic interactive technology in all content domains") recommend technology as an integral part of modern curriculum.

Prior learning frameworks (CCSS, GAISE II, NGSS) briefly mention technology with relatively high-level guidance on its use, and CCSS still recommends graphing calculators for modern data analysis, rather than computer-based or online tools. While encouraging technology integration, GAISE in particular mentions that "access to technology varies across school districts" and instead opts to outline methods to "incorporate appropriate uses of technology into statistical activities" when possible. Rather than matching technology integration to local resources, a prior goal should be to guarantee modern technology access (both computers and sufficient internet bandwidth), prior to program implementation. Simplifying content or removing technology may do a disservice to students in the long-term, especially as it relates to any digital literacy or media literacy aspects of data science. Moreover, COVID-related changes have increased device and bandwidth access for students in a variety of under-resourced communities (NCES, 2021). While access still must be improved, the landscape has changed dramatically in recent years, since many of these learning frameworks were drafted.

Data Ethics: only a subset of existing frameworks, including discrete data science guidelines, cover the ethical implications of data use in society. Moreover, coverage is relatively limited to a narrow focus on data privacy (K-12 CS, GAISE II, Dana Center). While an important ethical principle, other values-based issues frequently emerge in the analysis, utilization, or communication of data, including but not limited to: systemic bias in datasets, system bias in data-fed algorithms ("garbage in, garbage out,"), factual authority of data, visual manipulation of data, and monetization of data. The Dana Center course framework goes furthest in this regard, emphasizing students "should identify bias and sources of bias in data, and describe how bias in data impacts people and society." Curriculum providers have also emphasized the need for a "fourth ingredient" of civic responsibility and ethics to be included in the definition of data science (Schanzer et al, 2022). Explicit attention is likely needed in future learning guidelines.

Relevant Pacing: many educators interviewed for this paper noted the increased speed, quantity, and variety of information that students (particularly adolescents) consume on a daily basis. In simple terms, students are outpacing their teachers – both in technology uptake and the speed by which they process new information (or data). Student attention has many willing vendors: social media, e-sports, mobile games, podcasts, digital newsletters, all running on information cycles that are now much shorter than 24 hours. Students also have a plethora of online learning resources and supplements to learn material –- whether in data science, mathematics, or any other subject. This explosion of resources may fundamentally shift the role of the teacher, and the strategies required to engage students in any material regardless of subject. Building relevance, autonomy, and recency into data science and datasets will be critical to meeting students where they are today. New approaches to professional development will also be critical for educator toolkits to navigate this new world with confidence.

State Standards

Many states and districts around the country are working to create explicit learning opportunities in data science and data literacy. These "discrete" efforts exist in a wider context of subject-specific learning frameworks that recommend several foundational data concepts and building blocks (as outlined above), all with varying degrees of focus on data. However, both stakeholder interviews and research literature caution that data-related learning is often "left-out" of the school day due to a variety of factors – including perceptions of student assessments and evaluation, teacher confidence and training, and the order it appears in syllabi or standards. Even harder to measure through mapping standards is the degree of computer-based technology integration and training in the classroom.

To sort through this landscape, this section (*State Standards)* outlines both the *Context* for data education and *Discrete Data Science* learning expectations that have been articulated at the state-level to teach students about data. The following section (*Implementation Models*) gives further context to the discrete data science efforts (depicted in Figure 2), accompanied by a number of case studies to elicit the critical yet nuanced differences that may comprise a holistic and modern *data science* education.

In total, the number of discrete *data science* programs articulated in state policy (either implemented or planned) span 14 states, with efforts including:

- 7 state-wide standards or learning guidance (5 mathematics, 2 CTE)
- 8 state-wide data science course programs (5 mathematics, 3 computer science)
- 4 micro-credential programs (2 mathematics, 2 CTE)

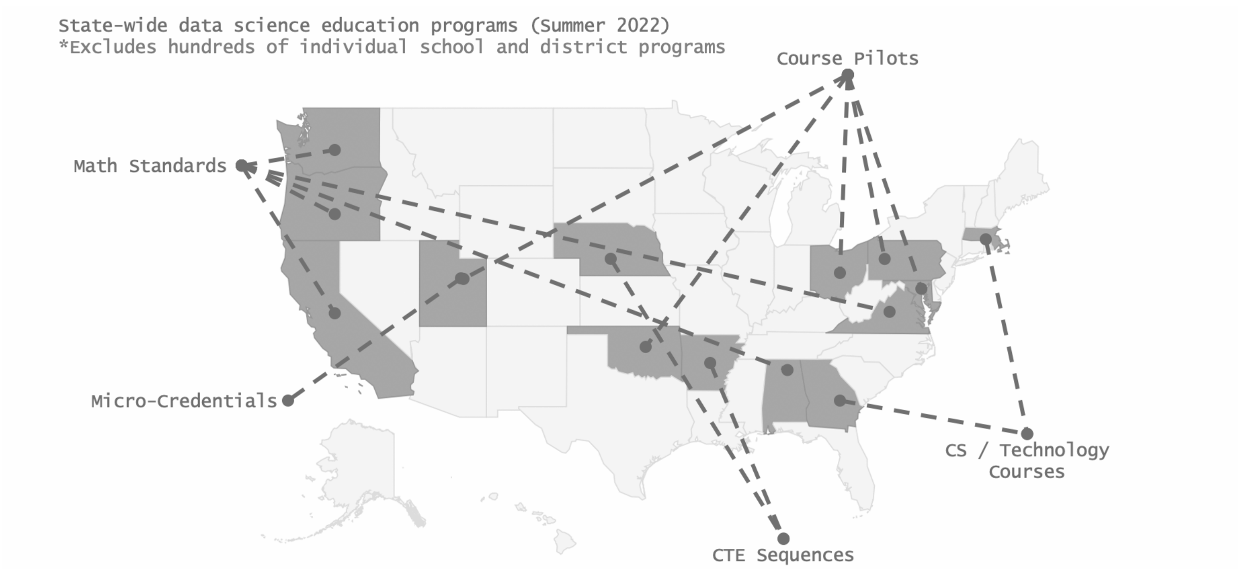

Figure 2: State-wide map on data science education programs (author data)

Math standards & requirements

Context: All 50 states have mathematics standards related to data in at least one grade-level band. 34 states have either implemented Common Core mathematics standards as originally written or with minor modifications – all which incorporate Measurement & Data in K-5, Statistics & Probability in 6-8, and then a discrete course in high school. Of the remaining states, 10 give explicit guidance for statistics in high school (1 in the context of an Advanced Math Modeling course), while 5 do not. However, across the board, standards related to data, statistics, or probability most frequently appear at the end of the school year – especially for state standards that rely exclusively on Geometry and Algebra 2 to convey a unit of Statistics at the very end. The state of New York even makes this explicit through their “Plus” standards, communicating to educators that statistics or data units in Algebra 1, Geometry, and Algebra 2 are optional and only expected for advanced students (New York State Department of Education, 2019). Additionally, only a handful of states explicitly recommend or mention utilizing computer-based software. Given current technology, this finding should be a significant red flag.

Discrete Data Science: The majority of data science standards have been implemented through mathematics thus far, given nearly all high school data science courses in the

U.S. to date are considered a mathematics course. 5 states have codified or recommended data science standards in their math frameworks, including California, Virginia, Oregon, Washington, and Alabama (*partial integration of "big data" and data wrangling skills). Moreover, Virginia is one example of a state that has drafted detailed learning expectations for high school data science from scratch – opting to design their own program instead of a pre-written curriculum. Data science topics implemented through math standards include data visualization, data collection and analysis, data wrangling and management, drawing conclusions from data, statistics and probability, and a computer-based technology component. These early examples showcase how data science learning can be incorporated into the K-12 mathematics infrastructure.

Science standards & requirements

Context: While data science programs have generally been implemented under the umbrella of mathematics, states have also articulated general data practices through science standards. Most of the data standards incorporated within science standards are contained in Practice 4 of the NGSS, "Analyzing and interpreting data," or are closely based off of this framework. Through Practice 4 of NGSS students are expected to:

1. expand their capabilities to use a range of tools for tabulation, graphical representation, visualization, and statistical analysis
2. improve their abilities to interpret data by identifying significant features and patterns, use mathematics to represent relationships between variables, and take into account sources of error
3. use digital tools to analyze and interpret data

20 states have adopted NGSS, 24 states have standards based off of NGSS, and 6 states have neither.

CS standards & requirements

Context: Many elements of data science education appear in state computer science standards. Data topics contained in computer science standards in various states include modeling and simulation, data collection and analysis, using data to support claims, and prediction with large data. However, similar to mathematics sequences, data analysis components are typically an individual thread or small cluster of learning expectations among many – creating risk that data topics may not be emphasized in practice.

As of 2021, 39 states have adopted the CSTA K-12 Computer Science Framework standards that include components for data literacy and analysis (Code.org 2021).

Discrete Data Science: embedded within computer science standards guidance, 3 states have either created guidance or list approved options for an explicit data science course, including Arkansas (*blended with CTE), Massachusetts, and Georgia. While difficult to generalize, courses articulated in computer science include topics relevant to technical management and operating on data, including: data types, software logic, data security, and algorithms. A smaller number of CS courses have incorporated both multivariable data analysis and data communication / visualization, along with other career-focused skills that relate to the use of data. Other examples include courses dedicated to databases and database management (e.g. Indiana's Computer Science III: Databases framework), though these courses exclude necessary focus on analysis and interpretation.

College degree programs & requirements

Context: The goal of K-12 data science education is not necessarily to prepare students to become data scientists, but to better prepare students of all interests and backgrounds for a world surrounded by data. A scan conducted by the American Statistical Association indicated that over 138 college degree programs require some level of statistics and data science (ASA, 2017). For those interested in pursuing a focused education in data science at the college level, a 2018 NAS consensus study report indicated a varying level of college level offerings including majors, minors, certificates, and massive open online courses (MOOCs) (NAS, 2018). Most undergraduate college data science programs consist of a combination of statistics / mathematics, computer science, and a blend of interdisciplinary courses based on the degree program.

Discrete Data Science Requirements: For students interested in pursuing a formal degree and career in Data Science, students must complete a variety of advanced mathematics. Most undergraduate programs often require students to complete up to Calculus 3/Multivariable Calculus, and are also often required to take additional math courses such as Linear Algebra and/or a Statistics/theoretical probability course (author data). However, some institutions offer Data Science degree programs with no requirements for Calculus (California State University East Bay, 2022), opting to instead introduce brief conceptual knowledge through introductory statistics. Other institutions have created computational Calculus courses meant for students pursuing Data Science, with all assignments completed in R and using real data (University of Arkansas, 2020). Stakeholder interviews revealed other Data Science degree programs in some cases opted to require Calculus "not because we had to, but simply because of inertia," given the significant number of freshmen who already learn the subject. In short, data science degree programs require students to advance in many quantitative reasoning domains during college, with ongoing re-organization and modernization of how individual topics are introduced.

Models of Implementation

We conducted a field review of current and planned models for implementing data science into K-12 education in the United States, including stakeholder interviews with teachers, policymakers, and other educators; review of pilot programs or policy changes in 15 states; and aggregated data on over 150 districts and schools who have implemented data science learning opportunities. We include data science examples that are either 1) articulated explicitly in state or local policy or 2) explicit data science programs already implemented or in process of implementation this school year (courses, discrete semesters or units, after-school programs, competitions, etc.). We exclude a great number of ongoing research projects still in development, or non-discrete integrations. This review also excludes thousands of new data science programs starting just this school year, which continue to grow rapidly (e.g. YouCubed, 2022). ***Several case studies of each model can be found in the Appendix.***

In our review, we find data science opportunities most frequently articulated as 1) high school mathematics offerings – through additional student course pathways, modernization of existing courses (e.g. Algebra 2), or advanced electives, and 2) as Career & Technical Education (CTE) sequences. Data and Analysis units can also be found in many Computer Science classes, as articulated in the K-12 Computer Science Standards (K-12 Computer Science Framework, 2016). Surprisingly few explicit data science courses appear in CS extension courses or electives – possibly a reflection of the efforts to integrate computational thinking elsewhere in the curriculum (rather than within itself), as well as the evolution of data science as an interdisciplinary field, relying upon statistics and specific domain knowledge as equal components. We also surface explicit examples of integration in science or social studies; with the caveat that case studies submitted for this workshop will provide much richer examples across a wider number of subjects.

Many decision makers expressed a need to foster grade-appropriate data literacy beginning in elementary and middle school, as well as integration of data science techniques across all K-12 subject areas – not as an optional pathway, but as universal content necessary for all learners. Given the interdisciplinary nature of data science, discrete subjects in K-12 present a challenge for creating authentic learning experiences. However, many stakeholders highlighted how this blending already occurs in the K-5 curriculum (given the diverse responsibilities of an elementary teacher), and then begins to unravel to distinct silos in secondary education. While several examples blending these disciplinary perspectives can be found in higher-education (e.g. Vance et al., 2022), this review only surfaced a handful of explicit programs integrating data science across subjects with authenticity (i.e. including meaningful technology integration and quantitative or digital information analysis). We also do not claim this review's sample to be comprehensive; however, challenges in finding these examples implemented in schools would suggest insufficient frequency or scale of these otherwise critical efforts.

Mathematics Pathways

Adding an additional course pathway in high school mathematics is the most frequent approach for incorporation of data science in K-12 education to date. This model has two primary approaches: 1) a 2+1 model, wherein students complete two years of combined coursework in Algebra, Geometry, and Statistics / Data, and can then pursue additional coursework in a variety of mathematics disciplines; or 2) an additional option to the Algebra 2 → Precalculus pathway, following Geometry. In both cases, students may have the opportunity to enroll in courses such as Data Science, Honors Data Science, AP Statistics, AP Computer Science courses, and even a Data Science II course – which may impart methods in multivariable regression, machine-learning, and artificial intelligence. Through these models, students are given additional choices in mathematics course offerings that can more closely align to their interests or career goals (see fig. 3). *See case studies on Oregon, Utah, Ohio, and San Diego Unified School District.*

A key advantage of these models is that they offer an additional option to meet high school graduation requirements. Algebra 2 courses in particular can be a significant graduation barrier (Burdman, 2015), in some cases even regardless of explicit policy (Stoker et al., 2018), with some suggesting low career relevance and narrow procedural focus as potential foils (Hacker, 2016). An evaluation of a pathways model in Texas college and universities found both an increase in passage rates and assessment scores (Schudde & Yonah Meiselman, 2021). Challenges of this model include 1) ensuring students can later "switch" to a post-secondary pathway that requires Trigonometry or Calculus if they choose 2) trade-offs in enrollment for data science course pathways and 3) stakeholder perception issues (parents, counselors) that new pathways are equivalently rigorous to existing pathways. Recent developments in college admissions standards may mitigate both these concerns (University of California Schools, 2021; Harvard University, 2022), which emphasize the importance of broad preparation in mathematics before college. Additionally, stakeholder interviews revealed introductory college mathematics is itself changing at many institutions nationally, enabling students to pursue a variety of STEM careers with more targeted quantitative skill sets (e.g. Math Q at Harvard, UCLA LS30, PRIME at CUNY), rather than a procedural preparation in topics like Calculus.

Modernized Algebra 2 (or other courses)

The modernized Algebra 2 method entails two steps: 1) re-organizing the contents of a traditional Algebra 2 course to focus on depth for key topics needed in later courses and 2) adding topics relevant to data science, which provide a context for existing Algebra 2 concepts to be modeled and connected to other disciplines or data. This method's aim is not only to increase data literacy taught in the high school setting, but also to update otherwise procedure-heavy topics less relevant in the computer age. Of the many positive impacts this method may generate, the greatest may be for the Algebraic content itself: boosting focus, depth, and relevance for critical mathematical topics may be helpful to students who find Algebra 2 to be a barrier to graduation (Hacker, 2016). The detail-oriented process to revise existing courses may prove challenging for a state-wide effort, though early case studies have shown promise for collaboration across sectors (districts, state boards, industry, higher-education) to ensure diverse student and societal needs are accounted.While Algebra 2 is the most frequent example to date of this approach, pilots have also expanded to Algebra 1 and Geometry, and the IDSSP framework focuses on updates to Pre-Calculus – suggesting many high school mathematics courses may benefit from similar work. Multiple states and school districts are currently piloting this approach. *See case studies on Washington state and Khan Lab School (associated with Khan Academy).*

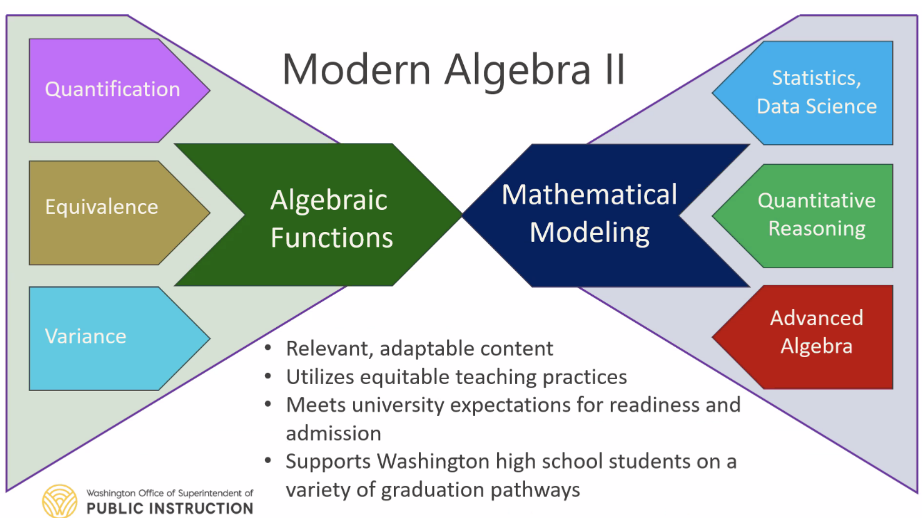

Figure 4: Washington's new vision for an Algebra 2 course (Washington Office of Superintendent of Public Instruction, n.d.).

Advanced Electives

The elective model allows students to take an optional data science course in high school. An elective data science course gives students the chance to explore the field without risking the pressure of a graduation requirement. An acute benefit to this model is ease of implementation, as there is no need for a change in standards, student pathways, or any cross-institution alignment. Additionally, data science coursework may give students an opportunity to demonstrate independent challenge and rigor to colleges and universities, making this an advantageous course for students interested in selective colleges (NACAC 2019). While these courses may incorporate mathematical concepts and prepare students for many aspects of college-level math, not all necessarily count toward required math credits for graduation, likely a headwind for enrollment. Additionally, multiple stakeholder interviews warned how an optional elective model will lead to lower enrollment by definition, as opposed to other models of implementation. While a productive step for building curricula, educator, and technology capacity, this model is less likely to build universal data literacy in the long-term without complementary K-10 integration. *See case studies on Virginia, North Carolina School of Science and Mathematics, and district-level implementation of the Introduction to Data Science Program.*

Career & Technical Education Sequence

The Career & Technical Education (CTE) approach gives students the opportunity to personalize their education based on career interests and unique learning needs – launching students into a formal data science career pathway or data-focused careers in a variety of sectors. Through CTE, students have a wide variety of course sequences to choose from that can lead them to their specific educational goal. While this is not exclusive to data science, a small number of states have created a data science option or include data science courses as a part of available sequences. According to U.S. Department of Education Data, students who focused on CTE courses while in high school had higher median annual earnings than students who did not focus on CTE (DOE, 2019). Career and technical education may serve as an especially natural fit for data science programs over time, given the tight connection between industry and academia in the associated professional data science field. *See case studies on Arkansas and Nebraska.*

Both CTE programs reviewed for this paper referenced challenges in scaling professional development, preventing state-wide adoption. Nebraska saw especially strong interest in their first year; however, this interest was stymied by the small number of schools that currently offer courses or have capacity to teach introductory statistics or computer science, let alone data science. Arkansas has similarly seen a small number of schools offer all three courses, relying on higher-education partners to cover strong enrollment demand, due to teacher capacity constraints. To help alleviate these barriers, the Arkansas Department of Education has integrated data science into their Year-1 credential program for computer science educators, and is currently developing an intermediate Year-2 training program focused exclusively on data science. The University of Arkansas is also leading a "teach the teacher" model, with expanding summer cohorts passing down training from the prior summer. Greater investment in professional training at scale, flexibility in alternative licenses, technology support, and expanded dual-credit offerings with higher-education may be needed to achieve scale and equitable access in other states.

Integration into science, social science, and other subjects

While learning from data is a major component of both science and the social sciences, stakeholder interviews suggest this content is too often "brushed over" at the high school level in particular. For example, only 50% of U.S. high school graduates can currently interpret a scatter plot in a science context (compared to 84% of college graduates) (Goo, 2014).
Greater focus in science and social science can re-emphasize data concepts already necessary for existing standards, and incorporate both quantitative reasoning and technology tools to explore subject matter at greater depth. A new study conducted at North Carolina State University suggests that computational learning methods and data analysis can help students problem-solve in social studies classes, and provide a common language for social analysis in classroom discussions (McGlinn Manfra et al., 2021). Stakeholder interviews suggest that social science and science teachers are already often burdened with teaching their students to understand graphical representations and statistics, absorbing time for other required material. If an increased focus on data science were to be articulated in these subject areas, or if these topics were covered more thoroughly in other areas, more time may be unlocked to cover subject-specific material. In other words, adding more time for data science and data literacy may save time in the curriculum on-net. *See integration case studies in KIPP NYC schools and New Jersey schools.*

Data is already part of science education standards (e.g. NGSS) and social science assessments (e.g. NAEP History and Civics assessments; several AP courses) in many ways (NGSS, 2013, 1; NAEP, 2022). A review of middle and secondary science curricula identified a plethora of technology and data tools currently used in science classrooms nationally, including probeware, wearable sensors, log-based data, and simulations (Lee & Wilkerson, 2018). However, the emphasis on classroom data collection and "sanitized" experiences may leave students unprepared for larger, complex, and "messy" data found in modern digital mediums or in automation contexts (i.e. machine-learning). Moreover, educator's comfort with retrieving and preparing public datasets may control the feasibility of creating authentic experiences for students, with a need to balance minimizing technology barriers while also ensuring skill-building in "data wrangling" (ibid).

K-8 Integration

The K-8 Integration model incorporates basic data literacy concepts into early education. While most data science programs currently appear at the high school level, it may be critical to build basic data acumen into math education starting at an early age (Engel, 2017). With today's pervasive digital sphere, younger students are exposed to data at younger ages over time. Students should gain accurate impressions of data as organized quantitative information (rather than just their cell service "data," for example), and be able to parse quality data analysis from misleading data (Wineburg et al. 2016; Zucker, Noyce, & McCullough, 2020).

Early programs are seeking to strengthen data literacy across the curriculum in K-8. Early resources have been piloted for the K-5 level, with the incorporation of bite-size lessons such as "data talks", a collection of short classroom discussions to help students develop data literacy (YouCubed, 2022). Other exemplars demonstrate innovative ways in which younger students can interact with data through cross-subject collaborations. *See case study on Data Literacy through the Arts (DLTA).* Project-based learning, semester-long experiences, and data modeling with basic math can be incorporated into the 6-8 level. Very basic data modeling and collection is already part of K-5 standards (Common Core State Standards, 2022), and could be expanded or supplemented with classroom-specific technology tools or spreadsheets. Moreover, existing research in informal statistical inference provides a promising framework for building early data intuition and acumen with K-5 students – including the relationship between data and probability (i.e. Beysian priors and updates), identifying data used in evidence-based claims, and the distinction between sample-based claims and generalizations (Makar & Rubin, 2009).

Both student and educator confidence will be critical for this new content, especially in younger grades. Students may be intimidated by technical or abstract technology. Stakeholder interviews revealed many K-8 teachers in particular feel unsuited or unprepared to teach data science as well. Teachers may need to navigate gray space with students, and may not always be a subject matter expert in ever-changing technology. Creating new norms for classroom environments may be critical for adapting to faster-changing technology. Moreover, interviewees highlighted how few educators benefited from a strong data or statistics preparation in their own education. This same issue may explain, in part, why the statistics and data literacy components of existing math curriculums are generally cut first. While promising research and individual curricular resources have been generated for K-8 students, additional work is needed to fully take advantage of these potential opportunities, and build student confidence during a formative age (Mason et al., 2020), especially given compounding effects of math anxiety for young students (Ramirez et al., 2013), and likely similar challenges for their teachers across subjects.

Out-of-School Programs

A small but growing number of out-of-school programs have been developed around the country to give students exposure to working with data in real and authentic contexts. These include team-based models (Young Data Scientists League in WA, grades 6-12), instructor-facilitated modules (Data Clubs in MA, grades 6-8) (NetApp Data Explorers, grades 6-8), or several regional and national open-registration data challenges (e.g. ASA Fall Data Challenge, grades 9-12). All of these programs universally rely on engaging students in 1) real-world authentic data and 2) "big" or complex data which cannot be read or manipulated manually. Both YDSL and the ASA programs also explicitly reward or encourage students to find and bring in external data found online to enhance their analysis, demonstrating consistent choices to emphasize "messy" or "imperfect" data found through digital sources - mimicking encounters students may experience later in life.

Competition models in particular can demonstrate the "low ceiling, high floor" potential for data science education, with students employing advanced mathematics and statistical tests (multi-variable linear regression, hypothesis testing, etc.), as well as data communication and collection acumen (complex data visualization, specific policy recommendations with mathematical modeling) and light programming through data-focused software (R, Python, etc.). High school students have also employed unsupervised machine-learning techniques and created A.I. programs leveraging large data, completing "the type of work that people with Master's degrees are performing." Other programs working with younger students (Data Clubs, Data Explorers) focus on novel approaches to impart the foundational concepts of data: distributions, correlation, and understanding the provenance of data ("who, what, where, how, why"). As data science is articulated from high school through kindergarten, out-of-school programs will be critical for innovation that can later support the traditional classroom, and can provide enrichment experiences that comprise a full ecosystem of student learning.

Data science out-of-school programs, like their in-school counterparts, also borrow heavily from project-based learning literature, and have demonstrated how increasing student relevance can drive deeper learning engagement. Some programs have moved even further to engage industry mentors and students directly in program design and implementation – in these cases, the teacher is no longer alone in content delivery, but part of a multi-generational "learning team." The move toward relevance is also reflected in the diversity and recency of topics: climate change, deep space exploration, TikTok trends, COVID cases, and art or dance trends. Early research evaluating student dispositions confirms the case for relevance: middle school students show strong prior interest even before engaging in these programs (PERG, 2020). Moreover, after-school programs show promise to generate meaningful increases in intrinsic interest, competencies, and values toward data and its societal importance (ibid). Some programs have also witnessed early success in student persistence, observing students will "dig deeper" and employ increasingly complex techniques to answer lingering questions (Higgins et al., 2021) – increasing their engagement with mathematical concepts and content.

Analysis and Discussion

Our review identified a wide diversity of K-12 data science education programs across school subjects, content progressions, and settings – true to the interdisciplinary nature of the professional and academic field. We identified several examples where data science is expanding foundational skills and approachable on-ramps for modern STEM + C education, and also serving as particularly advanced project-based STEM capstones in high school – especially in mathematics. Data science experiences also appear in other school subjects (social studies, humanities, and even the arts) and younger grades (K-8), though these programs have yet to be articulated into policy or discrete district programs – creating an opportunity for additional guidance work. Emergent themes from stakeholder interviews and program review include:

Deep Student Engagement: educators consistently report that courses and curricula in data science have been uniquely engaging for students. Many teachers expressed that students enjoyed the courses with surprise, and in some cases decided to pursue data science degrees in higher education (interview, 2022), while others have asked for opportunities and internships in the field (interview, 2022). Mathematics educators in particular were shocked by the engagement: "I have never gotten those questions before in nearly 20 years of teaching,: with others reporting "this was the first mathematics course [my students] enjoyed" (interview, 2022). Several educators also observed a plethora of students engaging beyond the required material or assignment, including those who previously expressed little interest or engagement in other mathematics courses. "Usually math is the deterrent for STEM pathways for my students; in this class however, they seem to be enjoying the material despite their past experiences" (interview 2022). We heard many similar stories from teachers in California, Ohio, Connecticut, and Idaho. Additional research is needed to understand the effects of different data science instruction strategies, through the rich project-based learning literature may provide some guidance (Krajcik and Blumenfeld 2006). Research referenced in this report suggests that the combination of content relatability, student autonomy, and inquiry-driven learning inherent to data science may partially explain the strong student excitement (Matuk et al., 2021; Heinzman, 2022; Rubin et al., 2021). Topics and datasets relevant to adolescents with recency will be critical for ensuring these opportunities are realized - especially given the speed at which online information, social media hashtags, and other “data” now travel.

Professional Development is Key: Many teachers reported not feeling comfortable teaching a topic that is new to both their students and themselves. While teachers have the option to teach statistics and probability in Algebra 2, our conversations revealed that in many cases, teachers choose not to cover these topics because of lack of time or confidence. Teachers also grappled with the challenge of helping their students learn new technology, including coding. While many existing curriculum programs provide teacher training as part of their offerings, most only teach educators the basics of programming in time-constrained PD programs. Teachers mentioned troubleshooting technology with their students during the school-year to be particularly time-intensive. To help mitigate technology barriers, many curriculum providers have integrated any programming into browser-based tools, so that students only must manipulate a few lines of code, and can avoid hassling with direct desktop downloads; others offer office hours where teachers can ask questions directly to developers for additional assistance. Notably, teachers have concurrently reported significant enjoyment in teaching the course after completing initial professional development: “their assessments were fun to grade.” (interview, 2022); others reported increased confidence in their profession: data science is a “signal it’s innovative, responding to what society is looking for right now” (interview, 2022). Many teachers (especially in math) have also been using their data science course to collaborate with other teachers in their school. We found this to be consistently true across interviews, with project collaborations cited with science, psychology, journalism, and AP Research teachers.

Risk to Becoming an “Alternative:” a paradox has emerged in the early success of data science education; the often shocking degree of student enthusiasm, including from students who have struggled with mathematics in the past, is creating confusion among some parents, teachers, and even students themselves. Because of the enjoyment, some students and parents have come to perceive data science to be the “easier” course option compared to other courses like Calculus. Teachers consistently told us that these perceptions were far from reality. Educators have tried to communicate that “it’s not less challenging, it’s just different, and we think you might like this challenge” (interview, 2022). As detailed throughout this landscape review, students are consistently engaging in advanced concepts. One current data science teacher told us that “watering down content is not equity; we believe all students can learn challenging content if it’s taught well.” One contributing factor may be that parents and students believe completing Calculus in high school is critical for higher-education, factoring into their admissions decision (Gewatz, 2020). In an interview with EdWeek, Steven Strogatz, professor of applied mathematics at Cornell, said “once you start talking about alternate pathways to Calculus, it’s just code for a rigorous pathway and a weak one” (ibid). Teachers consistently expressed that *all* students should take data science courses, including math-inclined students, because the content is stimulating and relevant to the real world, rather than a less “rigorous” pathway.

Summary

In our national review, we find that data science education is being implemented in many states across the country in different capacities and approaches. State standards on science, computer science, and mathematics education help frame many of the courses that we explored, while also providing an existing scaffold on which to further innovate. Moreover, this review excludes many ongoing efforts at hundreds of individual schools and districts that are only in the design stages of new programs, mirroring the significant momentum for data science undergraduate programs in recent years in higher-education. Many curriculum providers, researchers, and industry partners are also working to create additional data science-specific content. Most importantly, our interviews revealed that students are engaging in new data science course offerings to a shocking degree, and have expressed the desire to further their studies of quantitative analysis and STEM. Additional investment in professional development, policy guidance, course recognition and credit, and assessment tools to build seamless learning pathways will likely be needed to capture these promising opportunities.

Appendix I - Framework Commonalities

Baseline Commonalities Across K-12 Subjects

*Excludes "discrete" data science frameworks, focusing only on existing U.S. guidelines
**Cites earliest grade-level appearance of concept in a given framework

1. Data collection (NGSS, K-12) (CCSS Math K-5, via measurement) (CS, K-2, tech / digital tools focus) (GAISE II, A)
2. Data tables (NGSS, 3-5) (CCSS, HS) (CS, 6-8, tables for data storage) (GAISE II, B)
3. Basic variance (mean / median / mode) (6-8 NGSS) (CCSS HS) (GAISE II, A)
4. Variables in data (GAISE II, A) (NGSS, P3, 3-5) (CCSS, HS)
5. Statistical or inference-based questions (CS, K-2) (CCSS, HS) (NGSS, P4, 3-5) (GAISE II, A)
6. Data visualizations (dot plot, histogram, box plots) (NGSS, 3-5) (CCSS HS) (CS, K-2) (GAISE II, A, as variability)
7. Outlier data (CCSS Grade 8, HS) (GAISE II, C)
8. Data types (qualitative, quantitative, etc.) (NGSS, K-2) (CCSS, 8) (CS 6-8, as characters, noms, bits) (GAISE II, A)
9. Correlation vs. causation (NGSS, P4, 6-8) (CCSS, HS) (CS, 9-12, in prediction) (GAISE II, C)
10. Function / Model Fit (NGSS, P4, 9-12) (CCSS HS, emphasizes linear, quadratic, and exponential models; residuals) (CS, 6-8, focus on tech fit)
11. Slope / intercept / correlation coefficient (NGSS, P4, 9-12) (CCSS, HS) (GAISE II, C)
12. Sample selection, sample vs. population (NGSS, P4, 9-12) (CCSS, HS) (GAISE II, B)
13. Bayesian probability; updates to priors (NGSS, P4, 9-12) (GAISE II) (CS K-12, 6-8)
14. Units, ratios, percents, compound units, basic algebra (NGSS, P5, 9-12) (CCSS, throughout)
15. Analysis types: sample surveys vs. experiments vs. observational studies (CCSS HS) (GAISE, C) (U.T. Dana Center)
16. Sampling: random sampling, sample vs. population (CCSS 6-8) (CS K-12 6-8, simulation) (GAISE, B)
17. Randomization (CCSS, HS) (GAISE II, B)
18. Compare two things with data:
    a. NGSS: two alternate solutions (P4, 3-5)
    b. CCSS HS: two treatments & significance (or two distributions) (6-8)
    c. GAISE II: two groups & association between two variables (A)
19. Data collection from modern tech devices (NGSS, devices only) (CS, devices & online data) (GAISE II, devices & online data)
20. Data cleaning (GAISE II, A) (CS, 6-8)
21. Limitations of data (NGSS, P4, 6-8) (GAISE II, B) (CS, 3-5)

Appendix II - Case Studies by Model

Mathematics Pathways Case Studies

Oregon recently implemented a 2+1 Math Pathways model after updating the 2021 Oregon Mathematics Standards. All students take two core math courses after eighth grade—one credit of algebra, ½ credit of geometry, and ½ credit of data science/statistics—followed by a third-credit course option geared towards students' different postsecondary plans (Oregon Department of Education 2022). Students can choose the Calculus Pathway, Data Science Pathway, or Quantitative Reasoning Pathway (ibid). The Oregon Department of Education (ODE) anticipates the standards being implemented by many schools in the 2023-2024 school year (ibid). The state has awarded $2 million in grants to support implementation of the new pathway options this coming school year, which will include supporting districts to pilot curricula in support of data science pathway options. ODE is also working with their state universities to align college entry requirements with the new K-12 guidance, to ensure students may pursue a variety of careers, including in STEM, with any chosen pathway (ibid).

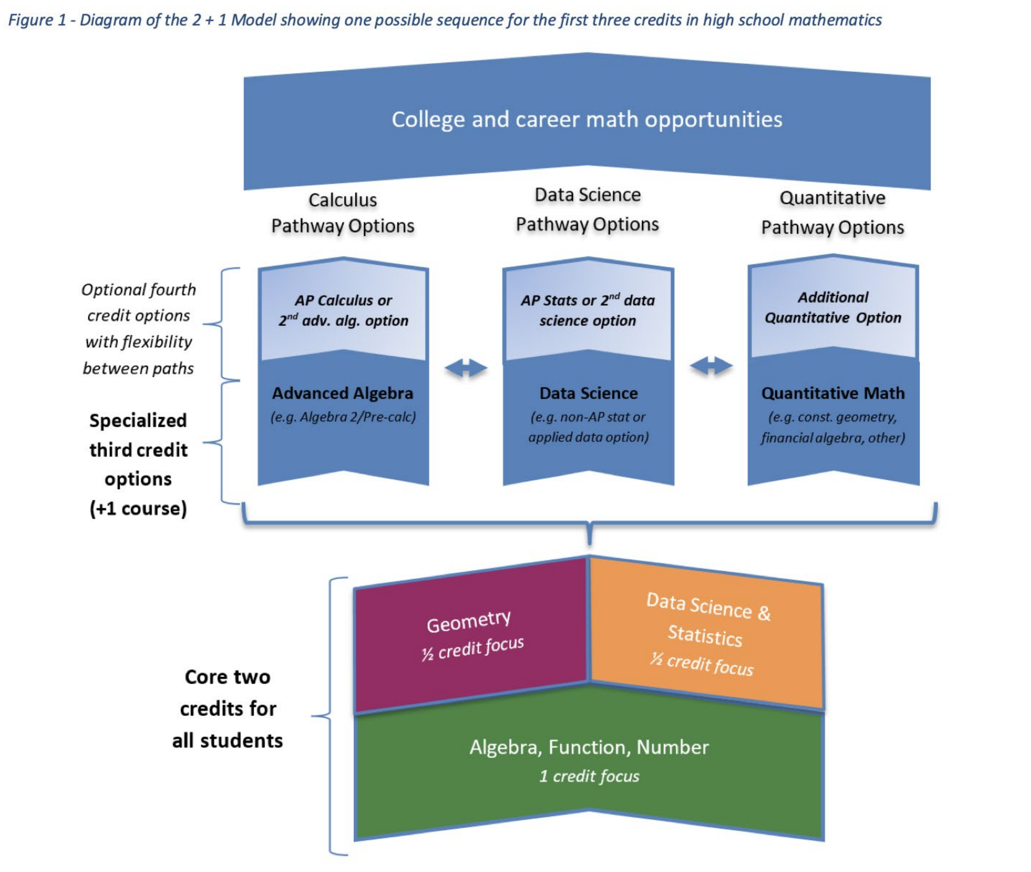

Figure 3: Oregon 2+1 Model (Oregon Department of Education, 2022)

Utah will be offering a data science course for the 2023-2024 school year as a third year mathematics option (Lindsey Henderson, Utah State Board of Education, 2022). The program garnered significant support from industry leaders in Utah, including the "Silicon Slopes," a growing technology community in Utah, due to the increasing importance of data literacy in 21st century jobs in both Utah and beyond (ibid). The pilot will draw on resources from multiple data science curriculum programs aligned to Utah standards, which will streamline the integration process (ibid). In advance of the program, Utah will also help facilitate professional development cohorts (summer 2023) in collaboration with curriculum providers for teachers for all interested districts. This builds on a recently developed micro-credential in data science education, offering stackable training opportunities state-wide to support lesson plans or bite-size integration of data science content into existing courses (ibid). Beyond high school, Utah also hopes to foster data literacy in K-8 mathematics in future programming (ibid).

San Diego Unified School District created a data science course pathway with the CourseKata curriculum (SDUSD, 2022). As a result, there are now two pathways to graduation in SDUSD. First, students must take Integrated Math 1 and 2 as part of the San Diego Enhanced Mathematics initiative (ibid). After completion of these requirements, SDUSD students can either take Integrated Math 3 then go on to Calculus, or take Data Science 1,2, followed by Data Science 3,4/Honors if they intend to pursue further study (ibid). In Data Science 3,4, students engage in several advanced data science topics including machine-learning, k-clustering, and multiple-linear regression. The University of California recently updated their A-G subject requirements so that both data science courses fulfill the "C" Mathematics requirement, meaning that they sufficiently prepare students for freshman year university-level math courses (University of California, 2022). This provides an opportunity to high schools across the state to expand their rigorous math course offerings while still preparing students for higher education. The Data Science 1,2 year-long course is approved as a "C" course by the University of California A-G requirements, and validates the Algebra II requirement. SDUSD also partnered with the CourseKata authors and other researchers to infuse a greater focus on data in the Integrated Math sequence (ibid).

Ohio created a new pilot program—currently running in 13 school districts—that offers five different math pathways after Algebra 1 and Geometry: Statistics and Probability, Data Science Foundations, Discrete Math and Computer Science, Quantitative Reasoning, or Algebra 2 (Ohio Department of Education, 2022). The Data Science Foundations course is intended for students who need a third or fourth mathematics credit (ibid). The program was created with the intention of modernizing their math offerings to ensure their students are better prepared for careers in the 21st century (ibid). Teachers participating in the pilot program have reported that students who they believed would not otherwise succeed in mathematics have successfully completed these other course options: "it saved them from not graduating high school" (interview, 2022). One teacher reported that her class of 24 students collaborated with an AP Research Capstone course, aiding with data analytics that the AP students did not have the statistical background to do. Several educators shared that students became more excited to pursue quantitative disciplines as a result of the course, including in Statistics, Data Science, and other STEM fields. In addition to relevance, survey work analyzing student attitudes suggested student autonomy in data science coursework may be an additional explanatory factor for the strong engagement witnessed thus far (Heinzman, 2022).

Modernized Algebra 2 Case Studies

Washington's state education agency is working to modernize Algebra 2 into a course that is aligned to a wider variety of students' interests and careers (Washington Office of Superintendent of Public Instruction, 2022). In existing state policy, students are required to complete Algebra 1 and Geometry (or corresponding Integrated Mathematics courses), with a third credit "based on student's interest and their High School & Beyond Plan" – naturally enabling the aforementioned Mathematics Pathways (ibid). Yet to both support extensive enrollment and meet existing local higher-education requirements, Washington opted to instead modernize their existing Algebra II course, sharpening focus on more critical Algebra II content in the first semester (sequences, functions & transformations, polynomials & rational functions, rational exponents, and exponential functions), and opening the second semester for mathematics supporting mathematical modeling and data science (including regression analysis, correlation vs. causation, machine-learning and predictive modeling, linear programming & matrices, and financial modeling). The course is running for the first time this academic year in nine districts across the state (ibid). Washington organized a development team to produce and open-source eleven modules with flexibility of implementation in the second semester (see fig. 4) to aid in content delivery. To equip educators for the new course, the state is also running a Summer Institute in advance of the course pilot, and will then be facilitating ongoing support during the school year through in-person and virtual PD cohorts - needed to meet the many rural locations in the state with fewer local resources.

Washington consulted both higher-education and local industry (including construction, agriculture and the U.S. Military), many of which expressed talent needs for these changes. Completion of the course will now align with the highest enlistment requirement for the Navy – which includes eligibility for working with nuclear submarines, as one example of a STEM career enabled by modernized content. Arlene Crum, OSPI's Director of Mathematics, emphasized that there is now a need to both "work forward and backward" from the course when considering future course revisions, given many students will "be doing math they may have never heard of before" in the second semester. Washington will also soon move to support districts in their implementation of additional optional fourth credit mathematics courses, including in data science.

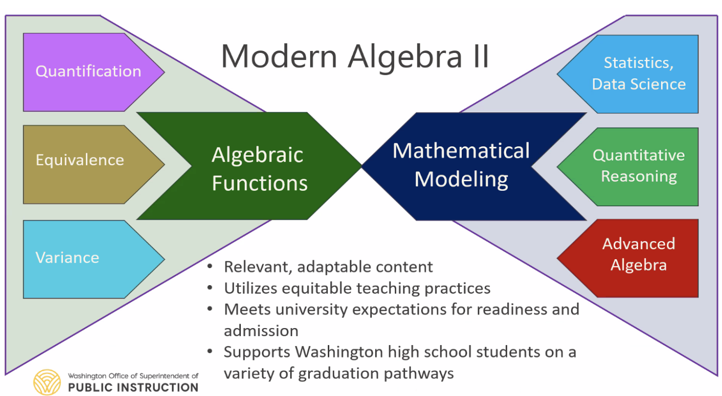

Figure 4: Washington's new vision for an Algebra 2 course (Washington Office of Superintendent of Public Instruction, n.d.).

Khan Lab School (associated with Khan Academy) in Mountain View, CA is currently piloting an iteration of the modernized Algebra 2, re-organizing the traditional "geometry sandwich" (Algebra 1 -> Geometry -> Algebra 2). In their revised sequence, KLS compresses Algebra 1 and Geometry into a single-year course, and then adds a full-year course dedicated to data science, statistics, and modeling (K., 2019). Students then complete a "full" Algebra 2 course before moving onto several opportunities in advanced mathematics courses (Calculus, Multivariable Calculus, Linear Algebra, Advanced Data Science, etc.), including the option for dual-enrollment in local college courses. In essence, KLS created a "data sandwich," rearranging the traditional mathematics sequence with modern data science content in the middle. This course articulation ensures all students learn data science content, and do so early in high school, while also preserving a path to high school Calculus for advanced students. In addition to their revised Algebra / Geometry sequence and their first Data Science & Statistics course, KLS will also be adding a second "Advanced Data Science" course this school year, entirely project-based. Topics will include polynomial regression, function building in programming tools, data wrangling, geospatial data, web-scraping, and other methods to be determined by student interest.

Advanced Electives Case Studies

The North Carolina School of Science and Mathematics—a specialty high school for juniors and seniors interested in STEM—adopted the University of California, Berkeley's popular undergraduate data science course (Data 8: Foundations of Data Science) to high school students. The curriculum, available for free online at data8.org, is intended to build inferential thinking, computational thinking, and real-world skill transfer for students (University of California, Berkeley, 2022). NCSSM originally adopted this semester-long course as an elective that students could choose to take in their junior or senior year. The pilot course was so successful that they decided to make it a requirement for all juniors at their new Morganton campus beginning this school year. Additionally, NCSSM will now offer an advanced data science course for students who enjoyed the foundational material. Taylor Gibson, Dean of Mathematics at NCSSM, highlighted that professional development for educators was the primary challenge in implementation, which both a summer "crash course" and on-going check-ins during the year have helped to overcome. Students immediately gravitated to the course and the technology integration. Given the significant interest, NCSSM is also now creating both a Data Science club, and a Data Science Expo to seed a state-wide opportunity for students to present their work.

Virginia is currently piloting a data science course, with the state-wide launch slated for this school year. State data science learning standards, the *Data Science Standards of Learning* (SOL) and *Data Science Standards of Learning Curriculum Framework*, were approved by the State Board of Education on April 22, 2022 to help guide implementation (Virginia Department of Education, 2022). The first data science courses would be offered in the 2022-2023 school year either as a semester or year-long course that counts toward students' mathematics graduation requirements, for students in or above 10th grade. Many students plan to enroll in the course for an *Advanced Studies Diploma*, enabling them to pursue a diverse mathematics option in the fourth year, including students who would not have otherwise taken mathematics senior year or would have earned a Standard Diploma with three rather than four math credits (ibid). The course will also be unique in exposing students to multiple software tools, including CODAP, spreadsheets (e.g. Microsoft Excel, Google Sheets), and Python. During the 2022-2023 school year pilot of Data Science, individual districts are facing early challenges with meeting student data privacy requirements using free web tools, web-filters and online activity regulations. Additional investments in this area will be needed to ensure authentic experiences for students that also respect reasonable cybersafety standards.

Virginia is also making significant and intentional investment in professional development: VDOE developed detailed unit-guides for each component of the course (with over 400 individual guides), worked with Virginia ASCD and the Rural Math Innovation Network to create two microcredentials for teachers who wish to teach Data Science: “How to Teach Inquiry-Based Methodology” (through lens of DS) and “Community Connections” (using the lens of DS to collect data and identify problems of practice), and will be building additional resources online into a online learning management system (LMS) for educators to access year-round (ibid). Pre/post survey analysis has shown significant gains in technology confidence after early PD interventions.

Introduction to Data Science (IDS) curriculum has experienced particularly wide adoption since its creation in 2014, funded by the National Science Foundation (NSF). Following pilots in Los Angeles Unified School District (LAUSD), the course quickly grew across Southern California (with 3,200 students across 51 schools enrolled in the 2020-21 school year). Today, in addition to a few aforementioned states, many high schools across the United States have piloted IDS curriculum, in addition to Australia and soon in Mexico. We spoke with teachers in Idaho, Connecticut, and Ohio who have experienced success with teaching the course, particularly in engaging students. Teachers have also expressed a need for an “Advanced” or an “Honors” version of the course, with many students advancing through the content quicker than expected, in addition to more dynamic dataset and topic options. Most importantly, the initial public investment in research & development for IDS – one of the earliest high school data science courses – paved the way for additional year-long curricula (e.g. CourseKata, Youcubed, Bootstrap, adoption of college-level courses, etc.) and countless modules and lesson plans from other providers. The IDS program is just one example of how public investment in new or unproven areas can catalyze follow-on development. Public investment to cover persistent gaps in data science education – including in K-8, assessment tools, and for subjects beyond math or science – may generate similar systemic effects.

Career & Technical Education Case Studies

Arkansas created a Data Science CTE pathway in 2021. The three-year program consists of three data science courses, with the option of including computer science and college-level classes in the third year (Arkansas Department of Education, 2021). The majority of third-year classes are taught by higher-education partners via dual-enrollment programs. The program also boasts eleven industry certifications that students can earn, including a Data Science Certification and various programming language certifications. The sequence always allows students to qualify directly for second- or third-year courses, given a significant number of Arkansas students who utilize online training programs in data science and computer programming in their own time. The pathway was designed in collaboration with the University of Arkansas and local industry representatives (e.g. Walmart, Tyson Foods, JB Hunt). Industry representatives were particularly invested in equipping students with the "ability to select the right model" over formulaic or procedural fluency, as well as tangible training in modern software including R, Python, and Tableau, rather than older programming languages (e.g. C++). Program designers also focused carefully to differentiate between computer science and data science thinking, emphasizing "programming as a means to an end, critical thinking, and recommendation-focused analysis" for data science curricula. To help create alignment to post-secondary mathematics education, the University of Arkansas also created a specialized Calculus course (Math 2594), which teaches students both multi-variable Calculus and Linear Algebra in the context of computational programs, with all student assignments completed in R.

Nebraska is implementing a similar program, with initial demand from the business community raised in 2016. Industry representatives were particularly looking for graduates who could understand, manipulate, and visualize data for external audiences, with a heavy emphasis on accurate and effective communication. Version 1 of the program will run from the 20-21 to 22-23 school years; Version 2 (already released) will be implemented thereafter (Nebraska Department of Education, 2022). The program combines perspectives from database management, statistics, and even graphic design. Students gain an introduction to advanced techniques (K-clustering, unsupervised vs. supervised machine-learning) by the end of the sequence. The program also includes an exploration of data privacy, ethics, and data collection practices, in addition to classroom discussions on the implications of Artificial Intelligence. Finally, students gain concrete technology skills needed for real-world analysis, including merging, uploading and management, and transformations (e.g. numeric to categorical data) in the very beginning, enabling hands-on exploration throughout the rest of the sequence. The three-course CTE sequence ends in a credential in Microsoft Excel. As of the 2021-22 school year, 48 schools out of 244 in the state are offering the three-course sequence (ibid).

Subject Integration Case Studies

KIPP NYC worked to develop a curriculum where 5th grade students work with data in their social studies and science classes as part of their efforts to incorporate computational thinking in other subjects, in partnership with Bootstrap at Brown University. Featured in a February 2020 EdWeek issue, KIPP's program engages students in a history class to "develop theories about what led to the downfall of Mayan civilization, using data to investigate various factors that could have played a role, including their diets, their human-sacrifice practices, and deforestation," according to an interview with KIPP-New York's science director Chéla S. Wallace (Gewertz, 2020). Additional units will enable students to analyze and compare "American Food: Past and Present," leveraging data to compare nutrition from and diets between present-day and native american cultures (KIPP STEM, 2022). Students will learn about the history of agricultural industrialization, the construction of America's continental supply chain, and present-day health challenges. KIPP NYC plans to expand the pilot in both directions (K-4, and 6-8) in future years.

New Jersey area public schools have partnered with PD providers to integrate data literacy learning into existing science courses and units, with the earliest professional development workshops beginning in 2017 (Dataspire, 2022). Rather than focusing on discrete courses, several schools in the state have chosen to provide curriculum-agnostic professional training to science and social studies teachers. Multiple institutions also focused on targeted strategies to impart NGSS data analysis practices through dynamic technology tools and real-world data. In working with teachers, professional development leaders have also intentionally avoided vocabulary in the field ("data science") ("data literacy"), etc., instead describing discrete student learning outcomes in greater detail to reduce intimidation or educator anxiety for imparting these topics. Training has also focused on a variety of accessible technology tools (Google Sheets, CODAP, Tuva Labs) to ease introductory barriers for educators.

K-8 Integration Case Studies

Data Literacy through the Arts (DLTA), a partnership between New York University Steinhardt, Fordham University, and Education Development Center (EDC) created a series of lesson plans intended to blend mathematics, arts, and computational thinking through the lens of data (DLTA, 2022). Since the 2020-2021 school year, students across multiple New York City and Boston public schools have engaged in integrated project-based experiences centered on data collection, analysis, and communication. Students co-enrolled in mathematics and arts classes build complementary skills, and then work across both courses on associated projects. Project examples have included data collection on mental health (with image vs. numeric data), communicating data through dance (e.g. TikTok) and through storytelling (e.g. comics). One learning module engaged students in a full cycle of data inquiry, including 1) data mining that identified indicators on neighborhood infrastructure (with public access geospatial data), 2) collecting data on missing indicators not available online (with photography), and then 3) communicating the information to others through storytelling (with poetry about their neighborhoods). Early research has indicated cross-disciplinary and project-based data science projects have the potential to drive deeper engagement with the material, increasing student autonomy and connections to real-world applications (Matuk et al., 2021).

Resources

American Statistical Association. 2017. “College Majors Requiring Statistics.” amstat.org.

Arkansas Department of Education. 2021. “STEM Programs of Study.” March 2021. https://dcte.ade.arkansas.gov/docs/Resources//STEM%20POS%20Pathways%20with%20Certifications%20March%202021.pdf

Battal A, Afacan Adanır G, Gülbahar Y. Computer Science Unplugged: A Systematic Literature Review. Journal of Educational Technology Systems. 2021;50(1):24-47. doi:10.1177/00472395211018801

Bootstrap. 2022. “Bootstrap.” 2022. https://bootstrapworld.org/.

Burr, Wesley. 2019. “International Data Science in Schools Project.” International Data Science in Schools Project. 2019. http://www.idssp.org/.

California State University East Bay. 2022. “Data Science Minor.” Department of Statistics and Biostatistics. Accessed August 2022.

Burdman, Pamela. 2015. “Degrees of Freedom: Diversifying Math Requirements for College Readiness and Graduation.” PACE. April 2015. https://files.eric.ed.gov/fulltext/ED564291.pdf.

Crawford, Deborah and Tina Mazzacane. 2022. Interview by Zarek Drozda, Davis Johnstone, and Brooke Van Horne. August 5, 2022.

Code.org. 2021. “2021 State of CS Report.” Code.org. (n.d.). Retrieved August 25, 2022, from https://advocacy.code.org/stateofcs

Common Core Standards. 2022. “Statistics & probability.” Statistics & Probability | Common Core State Standards Initiative. Retrieved from http://www.corestandards.org/Math/Content/SP/

Concord Consortium. 2022. “Narrative Modeling with StoryQ.” Concord Consortium. 2022. https://concord.org/our-work/research-projects/storyq/.

CourseKata. 2022. “CourseKata.” 2022. https://coursekata.org/. Data-8.Github.io. 2022. Data-8.Github.io. 2022. http://data8.org/.

Dataspire, 2022. “School / District Partnerships.” Accessed August 2022. https://dataspire.org/school-partnerships

Data Literacy Through the Arts (DLTA). Accessed August 2022. https://sites.google.com/nyu.edu/data-literacy-through-the-arts/projects?authuser=0

Donoho, David. 2017. “50 Years of Data Science.” *Journal of Computational and Graphical Statistics, 26* (4), 745-766. https://doi.org/10.1080/10618600.2017.1384734

Engel, J. 2017. “Statistical literacy for active citizenship: A call for data science education.” *Statistics Education Research Journal, 16* (1), 44-49.

Gewertz, Catherine. 2020. “Teaching Students to Wrangle ‘Big Data.’” Education Week. February 5, 2020. https://www.edweek.org/teaching-learning/teaching-students-to-wrangle-big-data/2020/02.

Goo, Sara Kehaulani. “The Art and Science of the Scatterplot.” *Pew Research Center*. 16 September 2015. https://www.pewresearch.org/fact-tank/2015/09/16/the-art-and-science-of-the-scatterplot/

Hacker, A. 2018. “The math myth: And other stem delusions.” The New Press.

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

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.

Indiana Department of Education. 2018. "Academic Standards Course Framework: Computer Science III: Databases." March 2018. https://www.in.gov/doe/files/cf-bmite-computer-sci3-data-3.7.18.pdf

K-12 Computer Science Framework. 2016. Retrieved from http://www.k12cs.org.

KIPP: STEM, "Data Science in Social Studies," 5th Grade Pilot. Accessed August 2022. https://www.kippstem.org/data-science-in-social-studies

Klein, David. 2022. "A Quarter Century of US 'Math Wars' and Political Partisanship." California State University, Northridge. 2022. http://www.csun.edu/~vcmth00m/bshm.html.

Krajcik, Joseph S., and Phyllis C. Blumenfeld. 2005. "Project-Based Learning." Chapter. In *The Cambridge Handbook of the Learning Sciences*, edited by R. Keith Sawyer, 317–34. Cambridge Handbooks in Psychology. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511816833.020.

K., Charlie. 2019. "Jay B.'s 'Merge and Purge' Project" The KLHS Khranicle. 2019. https://www.khranicle.com/jay-b-merge-and-purge-project.

LaMar, T. & Boaler, J. 2021. "The importance and emergence of K-12 Data Science." kappanonline.org. Retrieved August 25, 2022, from https://kappanonline.org/math-importance-emergence-k12-data-science-lamar-boaler/

Lee, Andrew. 2020. "Common Core State Standards: What You Need to Know."
Understood.org.
October 22, 2020. https://www.understood.org/en/articles/common-core-state-standards-what-you-need-to-know.

Lee, V. R., & Wilkerson, M. (2018). Data use by middle and secondary students in the digital age: A status report and future prospects. Commissioned Paper for the National Academies of Sciences, Engineering, and Medicine, Board on Science Education, Committee on Science Investigations and Engineering Design for Grades 6-12. Washington, D.C.

Makar, Katie & Rubin, Andee. 2009. "A Framework for Thinking About Informal Statistical Inference." *Statistics Education Research Journal, 8* (1), 82-105. May 2009. https://iase-web.org/documents/SERJ/SERJ8(1)_Makar_Rubin.pdf

Mason, Stacie L.; Rich, Peter J. 2020. "Development and Analysis of Elementary Student Coding Attitudes Survey." *Computers & Education*, Volume 153, August 2020. https://doi.org/10.1016/j.compedu.2020.103898.

Mazzacane, Tina. 2022. "Final Review of Data Science Standards of Learning and Data Science Standards of Learning Curriculum Framework." Virginia Board of Education Agenda Item G. April 21, 2022. https://www.doe.virginia.gov/boe/meetings/2022/04-apr/item-g.pdf

McGlinn Manfra, Meghan, Thomas C. Hammond & Robert M. Coven. 2022. "Assessing computational thinking in the social studies." *Theory & Research in Social Education*, 50:2, 255-296, https://doi.org/10.1080/00933104.2021.2003276

National Academy of Sciences, Engineering, and Medicine. 2018. "Envisioning the Data Science Discipline: The Undergraduate Perspective." NAS. https://doi.org/https://www.nationalacademies.org/our-work/envisioning-the-data-science-discipline-the-undergraduate-perspective

National Assessment of Educational Progress . (2022). "Assessments - civics: NAEP. Assessments - Civics" | NAEP. Retrieved from https://nces.ed.gov/nationsreportcard/civics/

National Association for College Admission Counseling. 2019. "The State of College Admission Report." NACAC.net. Retrieved from https://www.nacacnet.org/news--publications/publications/state-of-college-admission/

National Center for Education Research (NCER), 2021. "Catalyzing a New Field: Data Science Education in K-12." *Technical Working Group Summary.* October 26, 2021.

National Center for Education Statistics (NCES), 2021. "Students' Internet Access Before and During the Coronavirus Pandemic by Household Socioeconomic Status." September 20, 2021. https://nces.ed.gov/blogs/nces/post/students-internet-access-before-and-during-the-coronavirus-pandemic-by-household-socioeconomic-status.

National Council for the Social Studies. 2013. "C3 Framework." NCSS, Silver Spring, MD. (2013).

National Council of Teachers of Mathematics. 2018. "Catalyzing Change in High School Mathematics: Initiating Critical Conversations." April 25 2018. https://www.nctm.org/Store/Products/Catalyzing-Change-in-High-School-Mathematics-(Download)/

National Research Council. 2012. "A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas." Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Nebraska Department of Education, 2022. "Communication and Information Systems: Program of Study Areas 2023-2024." https://www.education.ne.gov/wp-content/uploads/2022/01/CIS-Final.pdf

New York State Department of Education. 2019. "New York State Next Generation Mathematics Learning Standards 2017." Updated 2019. http://www.nysed.gov/common/nysed/files/programs/curriculum-instruction/nys-next-generation-mathematics-p-12-standards.pdf

Ohio Department of Education. 2022. "High School Math Pathways." Accessed August 2022. https://education.ohio.gov/Topics/Learning-in-Ohio/Mathematics/Resources-for-Mathematics/Math-Pathways.

Oregon Department of Education. 2022. "Oregon Mathways Initiative High School Math Pathways-2+1 Model." March 2022. https://www.oregon.gov/ode/educator-resources/standards/mathematics/Documents/2plus1Model.pdf.

Oregon Department of Education. 2020. "Oregon math educator update - November 2020." ORED. https://content.govdelivery.com/accounts/ORED/bulletins/2ac7225

PERG: Program Evaluation and Research Group. 2020. “Data Dispositions Survey,” evaluating Data Clubs for Middle School Youth. https://www.terc.edu/dataclubs/wp-content/uploads/sites/10/2021/09/Fall-2020-Data-Dispostions-Results.pdf.

Ramirez, Gerardo; Gunderson, Elizabeth A.; Levine, Susan C.; Beilock, Sian L. 2013. “Math Anxiety, Working Memory, and Math Achievement in Early Elementary School.” *Journal of Cognition and Development (14)* 2, 187-202. 8 May 2013. https://doi.org/10.1080/15248372.2012.664593

Stanford Graduate School of Education. 2022. “Youcubed.” August 2022. https://www.youcubed.org/.

Schanzer; Emmanuel et al. 2022: Emmanuel Schanzer, Nancy Pfenning, Flannery Denny, Sam Dooman, Joe Gibbs Politz, Benjamin S. Lerner, Kathi Fisler, and Shriram Krishnamurthi. “Integrated Data Science for Secondary Schools: Design and Assessment of a Curriculum.” In Proceedings of the 53rd ACM Technical Symposium on Computer Science Education V. 1 (SIGCSE 2022). Association for Computing Machinery, New York, NY, USA, 22–28. https://doi.org/10.1145/3478431.3499311

Schudde, Lauren & Meiselman, Yonah. 2021. “Accelerating Success: A Multi-College Investigation of the Dana Center Mathways Project.” Policy Brief. February 2021. https://files.eric.ed.gov/fulltext/ED612363.pdf

Stoker, Ginger; Mellor, Lynn; Sullivan, Kate. 2018. “Trends in Algebra II: Completion and Failure Rates for Students Entering Texas Public High Schools.” Regional Educational Laboratory (REL) Southwest, Institute of Education Sciences. February 2018. https://files.eric.ed.gov/fulltext/ED580929.pdf.

The University of Texas Dana Center. 2022. “Launch Years.” Accessed August 2022. https://www.utdanacenter.org/our-work/k-12-education/launch-years.

The University of Texas Dana Center. 2021. “Data Science Course Framework.” May 2021. https://www.utdanacenter.org/sites/default/files/2021-05/data_science_course_framework_2021_final.pdf

University of Arkansas. 2022. “DASC 2594. Multivariable Math for Data Scientists.” Accessed August 2022. https://catalog.uark.edu/undergraduatecatalog/coursesofinstruction/dasc/

University of California. 2022. “Subject Area C: Mathematics.” University of California A-G Resource Guide. 2022. https://hs-articulation.ucop.edu/guide/a-g-subject-requirements/c-mathematics/.

University of California, Los Angeles. 2022 “Introduction to Data Science Curriculum V_5.0.” 2022. https://student.idsucla.org/overview/.

Utah State Board of Education. 2022. MIDAS Education Micro-Credentials. Accessed August 2022.

U.S. Department of Education. 2019. “Insights into how CTE can improve students' income after they graduate.” CTE Data Story. Retrieved from https://www2.ed.gov/datastory/cte/index.html

U.S. Department of Education. 2021. “Catalyzing a New Field: Data Science in K-12 Education Technical Working Group Summary” | October 26, 2020 Catalyzing a New Field: Data Science Education in K-12. Institution of Education Sciences. https://doi.org/https://ies.ed.gov/ncer/whatsnew/techworkinggroup/pdf/DataScienceTWG.pdf

Vance, Eric A.; Glimp, David R.; Pieplow, Nathan D.; Garrity, Jane M.; Melbourne, Brett A. 2022. “Integrating the Humanities into Data Science Education: Reimagining the

Introductory Data Science Course.” *Statistics Education Research Journal, 21* (2), Article 9. https://doi.org/10.52041/serj.v21i.

Virginia Board of Education. 2022. “Final Review of Data Science Standards of Learning and Data Science Standards of Learning Curriculum Framework.” https://doi.org/https://www.doe.virginia.gov/boe/meetings/2022/04-apr/item-g.pdf

Vance, E., Glimp, D., Pieplow, N., Garrity, J., & Melbourne, B. (2022). “Integrating the Humanities Into Data Science Education: Reimagining the Introductory Data Science Course.” International Association for Statistical Education.

Washington Office of Superintendent of Public Instruction (OPSI), 2022. “Aligning Mathematics Pathways with Student Career Goals” n.d. Accessed August 28, 2022. https://www.k12.wa.us/sites/default/files/public/legisgov/A5-Aligning-Mathematics-Pathways-with-Student-Career-Goals-ADA.pdf.

Wineburg, S., McGrew, S., Breakstone, J., & Ortega, T. 2016. “Evaluating information: The cornerstone of civic online reasoning.” Stanford Digital Repository.

YouCubed. 2020. “Data Talks.” https://www.youcubed.org/resources/data-talks/

YouCubed. 2022. Survey data on Youcubed High School Data Science Course: 1,830 teachers across 1,432 schools and districts serving 160,000 students, with 57% students of color.

Zucker, A., Noyce, P., & McCullough, A. 2020. “JUST SAY NO! Teaching students to resist scientific misinformation.” *The Science Teacher, 87* (5), 24-29.