INTEGRATING THE CURRICULUM: MATHEMATICS, STATISTICS, DATA LITERACY, AND DATA SCIENCE

Gail Burrill
Michigan State University
burrill@msu.edu

Many argue that all students should be prepared for a data driven world, implying statistical/data literacies should be part of their school experience. Across countries, the content in curricular documents at the secondary level related to these topics varies greatly from very little to some data analysis to simulation-based inference. In many countries, statistics is an elective course or added onto the mathematics curriculum. In the former case, students not choosing those courses are not prepared to make sense of a world of data and in the latter, statistical/data content is at the end of a mathematics course and often omitted. This paper argues for integrating mathematics and statistics through data and addresses two research questions: 1) to what extent and how are statistical ideas involving data envisioned and enacted in typical secondary school curricula across different countries and 2) in what ways can data-driven activities be integrated into the secondary school mathematics curriculum.

INTRODUCTION

The importance of preparing students to navigate a world driven by data has been recognized by many researchers (Andre & Lavicza, 2019, Gal & Gigner, 2021; Gould, 2021; Engel, 2017). There have been increasing calls to include data science in the high school curricula (Bargagliotti et al., 2020; Gould, 2023; International Data Science at Schools Project (IDSSP), 2019; Sukol, 2024). Statistics has been included in standards/curriculum frameworks by countries such as Brazil (Ministério da Educação, 2018) and Australia (Australian Curriculum). In some countries such as the United States, however, the result is often the addition of statistical content to the mathematics curriculum as the final unit in a course, which is often omitted. In other countries such as the Philippines, content at the upper secondary level related to these topics is an elective course, leaving students not choosing those courses unprepared for a data-driven world.

From another perspective, many students see traditional mathematics as a collection of rules and procedures unrelated to their world (Li & Schoenfeld, 2019; Matthews, 2019). Research has suggested that students are motivated to learn when they engage with relevant topics using real data (Czocher et al., 2021; Dunn & Marshman, 2019; Neumann et al., 2013; Stinson & Wager, 2012). This paper, aligned with that research, argues that statistical and mathematical concepts can be integrated through data-rich contexts. The elements of mathematics, statistics, data literacy, and data science involved when engaging in data driven contexts are taken as those in Figure 1, where content objectives specific to each domain overlap and support each other. The figure is informed by Lee et al. (2022), Smith et al, (2023) and Sukol (2024), but data literacy is presented at the heart of the other three domains because of its key role in both statistics and data science.

Note that Gould (2021) and others argued that little mathematics is necessary for data science, and some argue that statistics and mathematics should be separate entities. While this position might be optimal, a full treatment of each content area would demand more time and a deeper look into each domain. The reality is that secondary school schedules are limited by the number of subjects typically required of students, and yet, it is important that all students have exposure to a basic set of ideas related to each domain, regardless of their educational trajectory. Gould (2021) pointed out that the enormous changes caused by the growing importance of data in all aspects of cultural, social, economic, and scientific life mean there is a need to re-think secondary-level statistics education. Given these constraints, integrating the content through data seems like a reasonable path to follow. Schoenfeld and Daro, (2024) aligned with this stance stating, "Modeling with functions (i.e., algebra) is at the heart of data science, and the calculus pathway needs more work with data and computation" (p. 1298-1299). Thus, this paper addresses two research questions: 1) to what extent and how are statistical ideas involving data envisioned and enacted in typical secondary school curricula and 2) in what ways can data driven activities be integrated into the mathematics curriculum.

In: Kaplan, J. & Luebke, K. (Ed.) (2024). *Connecting data and people for inclusive statistics and data science education. Proceedings of the Roundtable conference of the International Association for Statistics Education (IASE), July 2024, Auckland, New Zealand.*

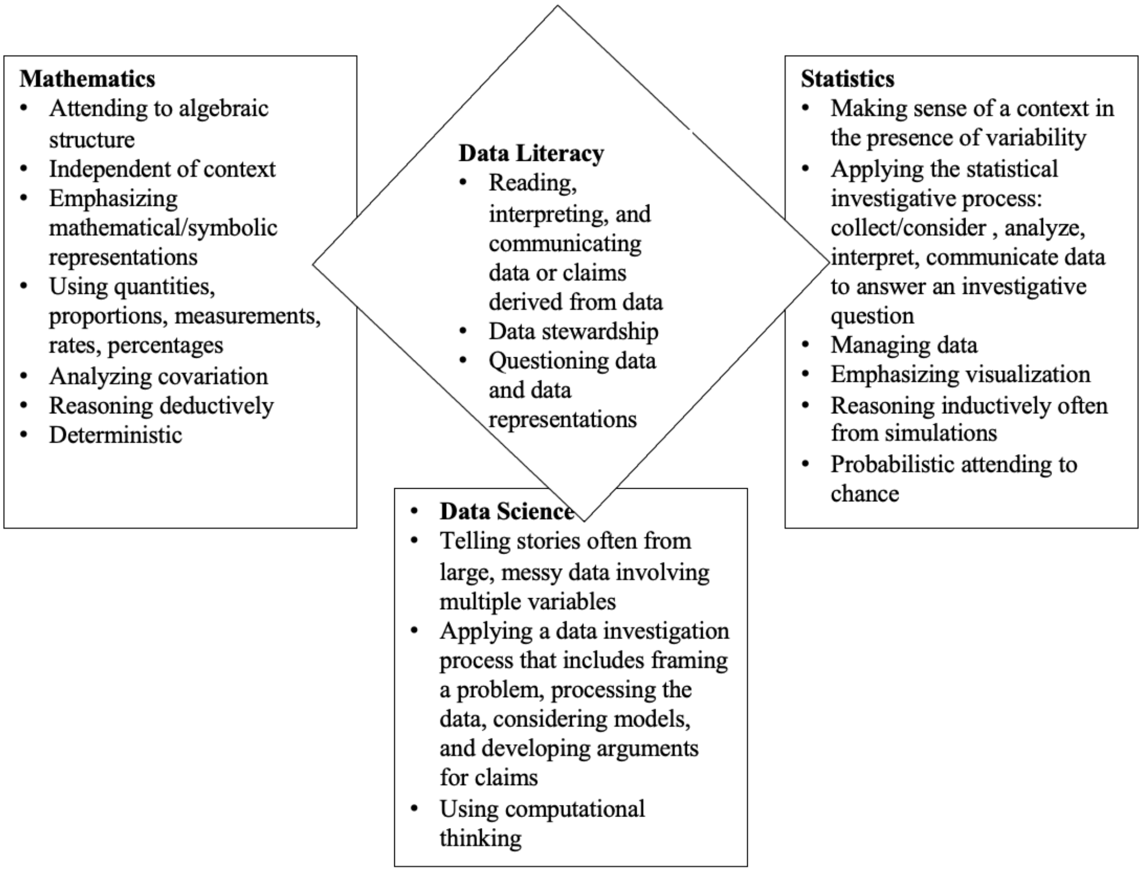

Figure 1. Content Relative to Solving Data Driven Problems: Mathematics, Statistics, Data Literacy, and Data Science.

METHOD

Educators from 14 countries responded to the question, "Describe the status of statistics/data science education in your country." The countries, Australia, Brazil, Canada (three provinces), Columbia, England, Finland, Germany, Japan, Korea, New Zealand, Spain (Catalonia), South Africa, Turkey, United States, were chosen to represent different areas of the world with different educational systems (Burrill, 2023). Curricular frameworks/standards from these countries were analyzed as well. In addition, articles about India, China, Ireland, and Bangeladesh related to the question were examined. The data were analyzed adapting a descriptive classification approach (Parsons, 1996). A first pass through the data, consistent with an open coding technique drawn from grounded theory (Strauss & Corbin, 1998), looked for similarities and meaningful differences (Rosch, 1978), such as the presence/absence of statements related to the use of simulation or counting techniques. The results, described below, were organized into four categories based on the degree of similarities (Burrill, 2023).

Sources for ways to integrate the content in a secondary mathematics/data curriculum, included curricula designed specifically to focus on data and written in English such as Data Driven Mathematics (https://www.amstat.org/education/k-12-educators), Introduction to Data Science (IDS; Gould et al., 2018), websites devoted to statistics activities such as education.ti.com and https://codap.concord.org, and research articles related to the implementation of data driven activities. Relevant research articles were identified by a non-exhaustive examination of abstracts from publications within the last five years in *Statistics Teacher*, *Journal of Statistics and Data Science Education, Teaching Statistics,* and proceedings of recent International Association for Teaching Statistics conferences. If the abstracts described a data rich activity, the article was reviewed for possible connections to mathematics. Several of these connections are described in the results section.

RESULTS

Status of statistics/data in secondary curricula

The analysis with respect to the research question related to how statistical ideas involving data are envisioned in typical secondary school curricula resulted in four categories, for which the content in category *n+1* subsumes the content in category *n*: a focus on

1) creating and using basic graphical representations (e.g., bar graphs, circle graphs) and elementary probability formulas such as those for finding the probability of intersections, unions, conditional probability.
2) introductory data analysis, often restricted to contrived, univariate data with typical graphical representations (box plots, dot plots, histograms), and the computation of statistical summary measures or formulaic probability.
3) introductory inference, in which the ideas in the treatment of statistics are primarily mathematical often with a relatively heavy emphasis on probability.
4) using data from real contexts with simulation-based procedures as a precursor for formal inference.

Most of the countries were in categories 1 and 2, with Germany, the United Kingdom, Finland, Korea, and the Canadian provinces of Nova Scotia and British Columbia in category 3. Only New Zealand and some states in the United States were in category 4.

Descriptions of how statistical ideas involving data were being enacted in typical secondary school curricula showed a gap between the vision and the reality for almost every country except New Zealand. The reasons for this difference consistently cited by those surveyed pointed to the lack of time in the curriculum for statistics (e.g., South Africa, Catalonia in Spain) and the reluctance on the part of teachers to engage with statistical content (Japan and the United States). In addition, in almost all the countries, assessments established barriers such as those in the UK where only a small percent of the marks at A level are on statistics or in the US and Australia where assessments primarily focus on other mathematical content and items that supposedly assess statistics often are not about statistical understanding but rather about performing mathematical operations. New Zealand stands out because the vision of what statistics should look like in K-12 schools and what takes place in practice seemed to be aligned (Pfannkuch & Wild, 2013). In general, implementation of statistics seemed weak; statements such as the following from the UK were typical: "Many national reports have been published about teaching statistics with few of their recommendations being implemented" (Davies & Sheldon, 2021, p. 552).

Together, the findings suggest that while ambitious goals for working with data are often in place, the reality is the traditional mathematics sequence crowds these out of many enacted curricula. One possible solution is to rethink what we teach in secondary mathematics and how we teach it. The increasing power of technology to support the learning of mathematics suggests that much of the content we now teach can be done effectively by mathematical action technology- technology designed to do mathematics. (Dick & Hollebrands, 2011). Using the technology to do the mathematical work can reduce the emphasis on topics such as factoring and rational polynomials creating room to incorporate data driven activities in the curriculum (Burrill, 2024). The next section highlights some ways to make the integration of statistics, data, and mathematics possible.

Integrating mathematics, statistics, and data

In the 1990s, the US National Science Foundation funded a project called Data Driven Mathematics, which produced 11 modules covering content related to beginning algebra through matrices. (Note the author was Project Director.) The focus of these modules, currently available on the American Statistical Association website (https://www.amstat.org/education/k-12-educators), was to develop traditional mathematical concepts through data driven activities. Variables and equations are introduced by formulas for rating/ranking in multiple contexts such as sports teams or movies, z-scores for comparison such as scores on college entrance tests, and Spearman's rank correlation to analyze preferences. Inequalities are developed by comparing the prices for groceries at two different food chains, and the mathematics of linear equations through analyzing car prices. Absolute value is central in mean absolute deviation and fitting linear models. The content includes percentages, ratios, rates of change, absolute and relative frequencies, statistical summary measures, and graphical representations.

The study of quadratic functions leads to minimizing the sum of the deviations for a linear model and the least squares regression line; logarithms and exponential functions are developed through examples such as federal debt, population, and number of kinds of stamps issued by the US post office, and matrices are used to create multiple regressions in contexts such as ratings and tree ages.

Topics, appropriate for introductory algebra, could be extended by including formulas for odds ratio (Eisenhauer, 2022), risk (Herd Immunity), and financial literacy, such as calculating the inflation rate using the Consumer Price Index. Higgins et al. (2021) described an activity that connects data to mathematics in relation to deer ticks and Lyme disease. The datasets by states progress from information about the presence of deer ticks, rates of infection, and location; to infection rates by year; to a complex set of attributes that could hypothetically be related to the spread of deer ticks and Lyme disease (percent forest cover, average summer high temperature, average summer low temperature, average winter low temperature, summer moisture level, and winter precipitation). The investigation involves understanding rates, relative and absolute frequencies, proportions, percentages, and statistical concepts such as statistical summary measures, and correlation and regression.

Many of the relevant articles involved regression. A building height activity (https://codap.concord.org) leads students to linear equations and interpreting rate of change and y-intercept. An investigation of the wage gap using data on gender, occupation, education, race, and location can engage students in working with fractions, ratios, and proportions, exploring linear relationships and inequalities, solving systems of equations, and piecewise functions as well as statistical summary measures, regression and residuals, and confidence intervals for slope (Burrill et al., 2023). Lock (2022) described a sampling process connecting car prices to positive and negative rates of change and to exponential and quadratic regression. Whitaker (2021) explored polynomial and power regression relating the length and weight of fish. The next section describes three data-based examples highlighting the mathematical and statistical concepts that can emerge in analyzing the data.

EXAMPLES

Example 1: Modeling Quarterback Passing Ratings (education.ti.com)

While this activity describes what is known as US football, the approach and the mathematics is relevant for many sports including hockey, soccer, baseball, and basketball. In the author's experience most students have been engaged by the activity, and an extensive understanding of the sport itself is useful, but not necessary, for the analysis. The data are statistics on 58 National Football League (NFL) quarterbacks recognized as very good over the years to determine who could be the "best" passing quarterback. The data, from football.db.com as of the 2022 season, consist of the number of passing attempts, completions, touchdowns, interceptions, and the yards gained (Table 1).

Table 1. A Sample of Quarterback Passing Data

	Touchdowns	Completions	Yards gained	Attempts	Interceptions
Tom Brady	649	7753	89214	10551	253
Otto Graham	174	1464	23584	2626	135
Patrick Mahomes	219	2386	28484	3590	63

As students explored the data, they recognized the need to account for the difference in years of experience (e.g., Brady who played for 23 seasons and Mahomes for seven). Typically, they either sought out additional data on seasons played or divided each of other four variables by the number of attempts. Accounting for the differences in both magnitude and units can be done in several ways, each involving slightly different mathematics. Proportions can be used to compare each player to the best in every category. For example, Otto Graham had the most yards gained per attempt at 8.98. Mahomes, at 7.934 yards per attempt, has achieved 0.8835 as many yards per attempt as Graham; note the units are no longer relevant. Z-scores can also address magnitude and units. The mean number of yards gained per attempt is 7.35 yards, and the standard deviation is 0.42 yards. Graham's z-score is 3.88 and Mahomes 2.63. A third strategy is to use the formula, normalized x = (x-x minimum)/range of x, which gives a normalized x for Graham of 0.707 and for Mahomes, 0.401. Any of these approaches will transform the data into unitless values, allowing students to create an equation that combines the variables, weighing some variables more than others depending on their insights and perspective. For example, Player rating = 10 TD/Att +5Comp/Att -20Int/Att + Yd/Att. The NFL actually has a formula for ranking quarterbacks on passing (see Modeling QB Passing Rates), which can be analyzed and compared to students' ideas.

Other contexts involving similar mathematical and statistical content include ranking regions of a country according to climate risk (see Climate education.ti.com for such a ranking of US states) or investigating countries in terms of their well-being (Well Being OECD; Ubilla & Gorgorio, 2020)

Example 2: Modeling CO_2 Levels (education.ti.com)

Rising temperatures are of increasing concern for citizens and governments around the world. One of the key contributors to climate change is the level of carbon dioxide (CO_2) in the atmosphere. CO_2 emissions, measured in parts per million (ppm), are largely caused by burning fossil fuels like coal, oil and natural gas that keep heat, which would normally disappear into space, trapped on Earth. In this activity students used data on CO_2 levels collected from the Mauna Loa Observatory in Hawaii from 1968-2023. The yearly mean levels seem approximately linear (Figure 2), but the residuals show a pattern indicating a linear model is not appropriate. An exponential model, while a bit better fit still has a pattern in the residuals. The monthly mean CO_2 levels, however, show a cyclic pattern (Figure 3), and a sinusoidal regression fits the data fairly well.

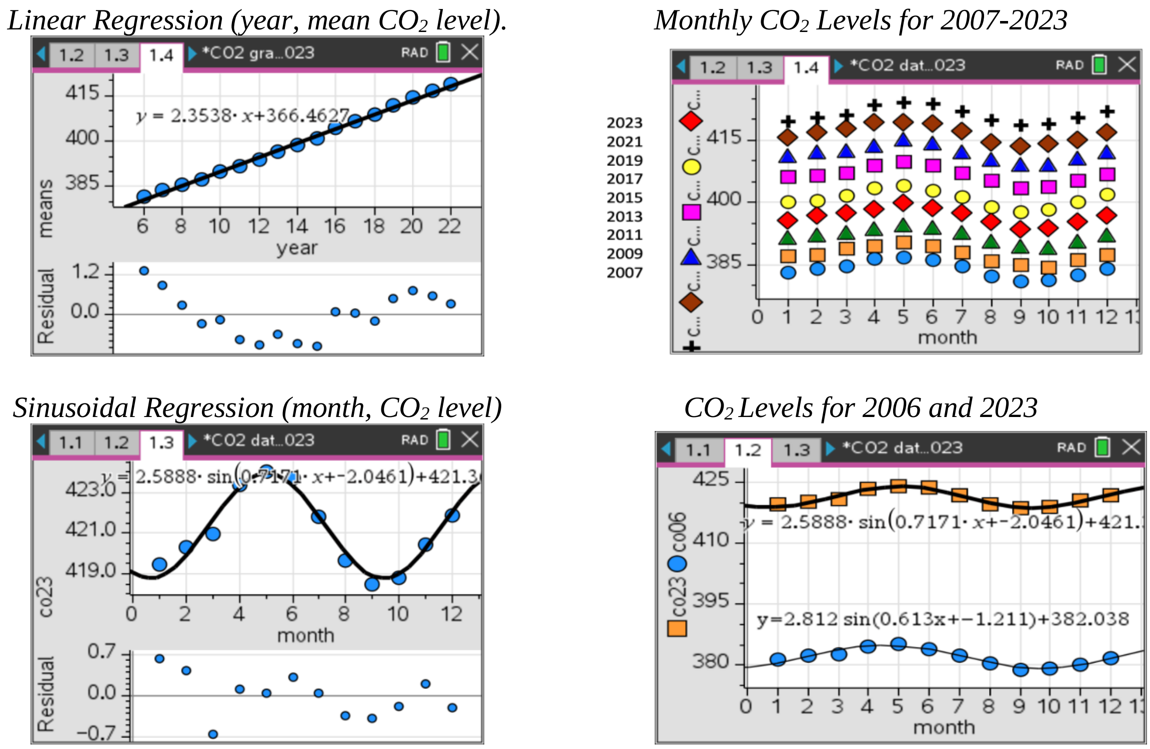

Figure 2. Representations of CO_2 data

Students can analyze the period (0.7171/6.2831, approximately 11 months) and the amplitude, which seems to be decreasing over time, providing an opportunity to explore why this might be. (Warming temperatures enable a longer time for trees and plants to grow and thus absorb CO_2 from the atmosphere.) The constant, the starting amount of CO_2 in the atmosphere for the time interval, is clearly increasing. Because these two patterns of change are evident in the data, one reflecting cyclic variation across the seasons of the year, and another reflecting an overall general increase in mean CO_2 levels for each year, a challenge for students is to create a model involving both changes. The actual model used by climatologists is $f(x) = 257 + e^{0.016(x-1958)} - 2.84\sin\big(2\pi 1.005(x + 23.339)\big) + 256.63$ (Krinos & Maurais, 2019). Investigating how to use this model for predicting (assuming the trend continues) when the CO_2 level will be greater than 600 ppm, the danger level for humans, can involve solving complicated equations.

Data related to contexts such as the number of births per month or per day, the lunar cycle, body temperature, or tides will also engage students in working with similar mathematical and statistical concepts.

Example 3: Roller coasters

Roller coasters first appeared in Russia in the 17th century (Wikopedia, https://en.wikipedia.org/wiki/Roller_coaster), and since then, thousands of roller coasters have been constructed across the world. The site https://rcdb.com contains a historical data base of roller coasters where individual roller coasters can be accessed randomly and that, depending on availability, may include length, height, drop, speed, inversions, vertical angle, duration, location, and whether it is made of wood or steel. Because the data are generated at random, students can investigate how sample size can affect outcomes (the mean height of a random sample of 10 roller coasters was 68.5 ft, while the mean height of a sample of 50 was 105.4 ft.). Comparing different random samples of size 10 and size 50 illustrates the variability in small samples, how distributions stabilize as sample size increases, and how samples of size 50 have approximately the same center and variability. Students can find confidence intervals for the population means and analyze whether the differences in mean speeds or heights by region of the world is likely due to chance. The mathematics involves different regression models and using logarithms to straighten the data to find a better model. For example, most of the residuals for the linear model to predict height from length are below the regression line (Figure 3), indicating the line most often predicts a height greater than the observed height, which suggests the assumptions of the model should be questioned and not used for predications. Transforming the data using logarithms to find a better fit, then manipulating the equations allows students to make predications, e.g., to estimate the height of a 7.5 ft. roller coaster, students would have to solve the equation $ln(height) = 0.7237\ ln(7.5)\ -1.253$.

Roller Coasters (length, height) *Roller Coasters (ln length, ln height)*

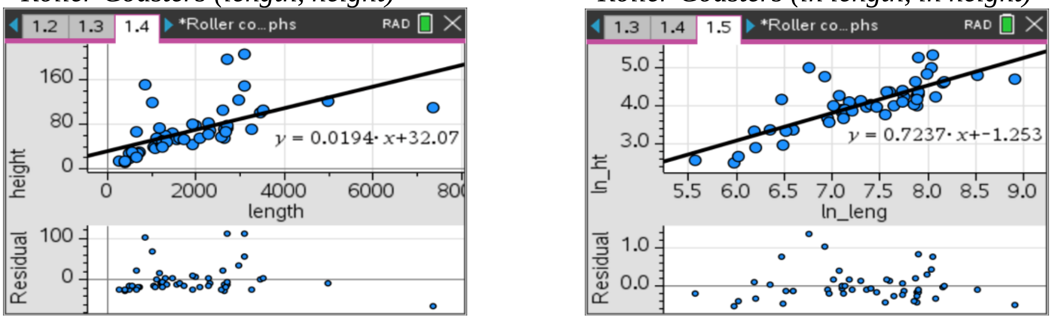

Figure 3. Representations of Roller Coaster data

CONCLUSION

The discussion above highlights the importance of preparing all students for a world driven by data, provides some evidence that suggests despite intentions articulated in curriculum documents, the reality is engaging with real data is not part of what students do in secondary schools in many countries. This paper argues for the integration of concepts from mathematics, statistics, data literacy, and data science to give all students the necessary foundation to live in this data driven world and to give them the tools they will need to move forward in any pathway. The examples illustrate ways in which this integration could be done, and a variety of resources are suggested that could be used by those deliberately seeking ways to make the connections between mathematics and data.

The challenges are many: helping students (and teachers) understand that, while theoretical mathematics is algorithmic and certain given a set of assumptions and definitions, statistics deals with data and uncertainty, but it is possible for the two to complement each other (Scheaffer, 2006); supporting teachers who often lack preparation in statistics and are reluctant to teach in anything other than a tightly controlled context; overcoming the inertia of tradition, which can value processes and procedures that are no longer relevant; ensuring that policies and high stakes assessments align with an integrated curriculum. Weiland and Engledowl (2022, p. 2) considered these challenges and others that exist in the United States education system and offer a "set of recommendations for building capacity to develop the data literacy of all students through the teaching of data science and statistics concepts and practices in the K–12 mathematics curriculum" (p. ??), recommendations that may be useful for other countries interested in this direction. The bottom line is without change, we will continue to prepare students for yesterday's world and not for tomorrow.

REFERENCES

Andre, M., & Lavicza, Z. (2019). Technology changing statistics education: Defining possibilities, opportunities and obligations. *Electronic Journal of Mathematics and Technology, 13*(3), 253–264.

Australian Curriculum, Mathematics. https://v9.australiancurriculum.edu.au/f-10-curriculum/learning-areas/mathematics/foundation-year_year-1_year-2_year-3_year-4_year-5_year-6_year-7_year-8_year-9_year-10?view=quick&detailed-content-descriptions=0&hide-ccp=0&hide-gc=0&side-by-side=1&strands-start-index=0&subjects-start-index=8

Bangladesh National Curriculum Framework. (2021). National Curriculum and Textbook Board, Bangladesh.https://file-chittagong.portal.gov.bd/uploads/3501485c-b357-4cd0-8624-66e835222e4f/65b/8da/4cd/65b8da4cdd0d9185517429.pdf

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D. (2020). Pre-K-12 Guidelines for assessment and instruction in statistics education (GAISE) report II, American Statistical Association. https://www.amstat.org/asa/files/pdfs/GAISE/GAISEIIPreK-12_Full.pdf

Burrill, G. (2023). An international look at the status of statistical education. In Burrill, G., Reston, E., & Souza, L. (Eds.), *Research on reasoning with data and statistical thinking: International perspectives* (pp. 11-16). Springer.

Burrill, G. (2024). Preparing students for tomorrow: The role of interactive dynamic technology. Paper presented in *ICME-15 Topic Study Group 3.8, The role and the use of technology in the teaching and learning of mathematics at upper secondary and tertiary level*. Sydney, Australia.

Burrill, G., Funderburk, J., Byer, B., & Gorsuch, R. (2023). Using technology to explore the wage gap. *Mathematics Teacher: Learning and Teaching Mathematics PreK-12*. *116*(5), 378-386. https://doi.org/10.5951/MTLT.2022.0348

Czocher, J., Melhuish, K., & Kandasamy, S. (2019). Building mathematics self-efficacy of STEM undergraduates through mathematical modelling. *International Journal of Mathematical Education in Science and Technology, 51*. 1-28. https://doi.org/10.1080/0020739X.2019.1634223.

Davis, N., & Sheldon, N. (2021). Teaching statistics and data science in England's schools. *Teaching Statistics*, *43*(51), 52-70. https://doi.org/10.1111/test.12276

Dick, T., & Hollebrands, K. (2011). *Focus in high school mathematics: Technology to support reasoning and sense making*. National Council of Teachers of Mathematics.

Dunn, P., & Marshman, M. (2020). Teaching mathematical modelling: A framework to support teachers' choice of resources. *Teaching Mathematics and its Applications: An International Journal of the IMA, 39*, 127-144. https://doi.org/10.1093/teamat/hrz008.

Engel, J. (2017). Statistical literacy for active citizenship: A call for data science education. *Statistics Education Research Journal 16*(1): 44-49. https://doi.org/10.52041/serj.v16i1.213

Eisenhauer, J. (2022). Probabilities, odds, and vigorish. *Teaching Statistics, 44*(3), 126-132. https://doi.org/10.111/test.12318

Gal, I., & Geiger, V. (2022). Welcome to the era of vague news: A study of the demands of statistical and mathematical products in the COVID-19 pandemic media. *Educational Studies in Mathematics, 111*(1), 5–28. https://doi.org/10.1007/s10649-022-10151- 7.

Gould, R. (2021). Towards data scientific thinking, *Teaching Statistics 43*(11), 11–22. https://doi.org/10.1111/test.12267

Gould, R. (2023). All students need to learn data science. *Edsource*. https://edsource.org/2023/all-students-need-to-learn-data-science/703117

Gould, R., Moncado-Machado, S., Molyneux, J., Johnson, T., & Trusela, L. Mobilize: A data science curriculum for 16-year-old students. (2018). In M. A. Sorto, A. White, & L. Guyot (Eds.). *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10).* International Statistical Institute. https://iase-web.org/documents/papers/rt2016/Gould.pdf

International Data Science in Schools Project (IDSSP). (2019). Curriculum frameworks for introductory data science, http://idssp.org/files/IDSSP_Frameworks_1.0.pdf.

Krinos, A., & Maurais, A., (2019). Parameter and uncertainty estimation for a model of atmospheric CO2 observations. *SIAM Undergraduate Research Online 12*, 305-321. https://www.siam.org/media/fl0l1vto/s01753pdf.pdf

Lee, H., Mojica, G., Thrasher, E., & Baumgartner, P. (2022). Investigating data like a data scientist: Key practices and processes. *Statistics Education Research Journal*, *21*(2), https://doi.org/10.52041/serj.v21i2.4

Li, Y., & Schoenfeld, A. (2019). Problematizing teaching and learning mathematics as "given" in STEM education. *International Journal of STEM Education*, *6*(44). https://doi.org/10.1186/s40594-019-0197-9

Lock, R. (2022). Some favorite car activities. In S. A. Peters, L. Zapata-Cardona, F. Bonafini, & A. Fan (Eds.), *Bridging the gap: Empowering & educating today's learners in statistics. Proceedings of the 11th International Conference on Teaching Statistics (ICOTS11*). International Statistical Institute. http://iase-web.org/icots/11/proceedings/pdfs/ICOTS11_167_LOCK.pdf?1669865530

Matthews, J. (2017). When am I going to use this in the real world? Cognitive flexibility and urban adolescents' negotiation of the value of mathematics. *Journal of Educational Psychology*, *110*(5). https://doi.org/10.1037/ edu0000242

Ministério da Educação. Base Nacional Comum Curricular – BNCC. Brasília: MEC/SEF.

Neumann, D., Hood, M., & Neumann, M. (2013). Using real-life data when teaching statistics: Student perceptions of this strategy in an introductory statistics course. *Statistics Education Research Journal*, *12*(2), 59-70. https://doi.org/10.52041/serj.v12i2.304

Parsons J. (1996). An information model based on classification theory. *Management Science* *42*(10) 1437–1453. http://www.jstor.org/stable/2634376

Pfannkuch, M., Wild, C., & Parsonage, R. (2012). A conceptual pathway to confidence intervals. *ZDM – The International Journal on Mathematics Education*, *44*(7), 899-911. https://doi.org/10.1007/s11858-012-0446-6.

Higgins, T., Rubin A., Mokros, J., Sagrans, J., & Ren-Mitchell, A. (2021). When 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)*. International Association for Statistical Education. https://doi.org/10.52041/iase.fjrqt

Rosch, R., (1978). Principles of classification. In E. Rosch, & B. Lloyd (Eds.), *Cognition and categorization*, (pp. 27-48). Lawrence Erlbaum.

Scheaffer, R. (2006). Statistics and mathematics: On making a happy marriage. In G. Burrill (Ed.), *Thinking and reasoning with data and chance, 68th NCTM Yearbook*, (pp. 309–321). National Council of Teachers of Mathematics.

Schoenfeld, A., & Daro, P. (2024). More math: Less war. *Science*, *383*(6689), 1297-1299. https://www.researchgate.net/publication/379152401_More_math_less_math_war

Smith, C., Kathotia, V., Ward-Penny, R., Howson, O., & Wermelinger, M. (2023). *Data literacy: Competencies and curricular implications at the intersection of mathematics, data science, statistics and computing*. The Open University.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Sage Publications, Inc.

Stinson, D., & Wager, A. (Eds.). (2012). *Teaching mathematics for social justice: Conversations with educators.* National Council of Teachers of Mathematics

Sukol, S., (2024). *Beyond borders 2024: Primary and secondary education around the world*. Data Science 4 Everyone.

Ubilla, F., & Gorgorió, N. (2020). From a source of real data to a brief news report: Introducing first-year preservice teachers to the basic cycle of learning from data. *Teaching Statistics*, *43*(1), 110-123. https://oi.org/10.1111/test.12246

Weiland, T., & Engledowl, C. (2022). Transforming curriculum and building capacity in K–12 data science education. *Harvard Data Science Review*, *4*(4). https://doi.org/10.1162/99608f92.7fea779a

Whitaker, D. (2021). Nonlinear modeling: Something fishy. *Statistics Teacher*. https://www.statisticsteacher.org/2021/03/29/nonlinear-modeling-something-fishy