Mathematical and Data Literacy

Competencies and curriculum implications at the intersection of mathematics, data science, statistics and computing

Cathy Smith, Vinay Kathotia, Robert Ward-Penny, Oli Howson and Michel Wermelinger.
April 2023

The Open University

Contents

1.   Introduction

This report contributes to the Mathematical Futures Programme aim of building a new vision of mathematics education that anticipates and supports the role of mathematics for individuals, economies and society, strengthening diversity and reducing inequity. Previous work on this project had identified three areas with the most potential and need: (1) the integration of data science into the mathematics curriculum, (2) opportunities for enhancing programming and computational thinking in mathematics, and (3) a coherent and cross phase approach to the use of digital technology in mathematics.

The Royal Society has commissioned this report to investigate the intersection of mathematics, data science, statistics and computing, thereby addressing areas (1) and, to some extent, (2) above. A combination of desk research and expert interviews was used to synthesise sector-specific documents that set out visions for competencies needed by future school leavers and reported curriculum initiatives relevant to those goals. This was used to create a body of knowledge at the intersection of mathematics, data science, statistics and computing, listing top-level competencies that young adults should have when they leave school or college, and to consider the implications for various models of curriculum change.

2.   Summary

Part One of this report establishes a framework of competencies for activities at the intersection of mathematics, statistics, data science and computing. It provisionally terms this area of activity 'mathematical and data literacy'. It was based on a survey of curriculum documents and research papers, supplemented by expert interviews, within the four disciplines of mathematics, statistics, data science and computing education.

Two principles have guided the identification of mathematical and data literacy competencies:

- Competencies included in the intersection should, as far as possible, align with consensus views in the literature, increasing the potential for adoption.
- The overall set of competencies should be minimal and coherent, providing a basis for future curriculum development work.

Recommendations for curriculum mapping

- For the purposes of future-facing school curricula, statistics should be disaggregated from mathematics and, instead, data science and statistics can be treated as cognate. Data science, with its contextual relevance and ethical aspects, provides the opportunity for an engaging, data-driven education that includes all the statistics needed by school leavers.

- In defining the intersections of mathematics, statistics/data science and computing, a helpful focus is on mathematics competencies most closely associated with data-driven questions, whether from pure or applied mathematics, rather than on finding connections to all of mathematics.

- While fluent use of technology is clearly related to computing (which includes programming), their changing relationship and different take-up in schools makes it helpful to consider the two separately.

- Knowledge and decision-making within a context, and with relevant tools, are so important to statistics/data science that they cannot be omitted from the intersection. We conclude that the intersection of mathematics, statistics/data science and computing is concerned with the study of variation and generalisability, with precision and context and technology and ethics.

Intersectional Competencies
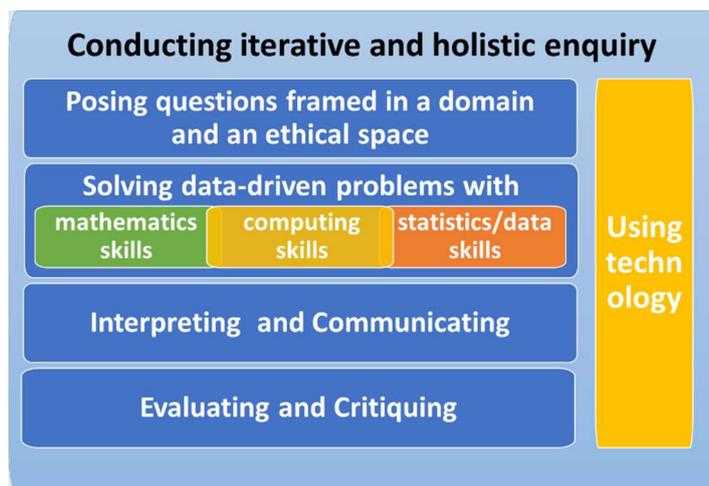
Figure 1 Model for intersectional competencies

The work on consensus areas produced a framework shown in Figure 1 that aligns mathematics, statistics/data science and computing around the overarching activity of

posing and answering questions. This is based on six high-level competencies common across the reference literature. They are related but different, and each is necessary for the resulting activity to be considered as characteristic of the intersection of mathematics, statistics data science and computing. See 5.1 for details.

Within this framework, problem solving can draw on 'toolkits' in mathematics, computing and statistics/ data science, not just a collection of techniques, but the ways of thinking and reasoning, and the language, concepts and tools inherent to those disciplines.

- The mathematical toolkit involves: Using quantities and methods; Analysing covariation; Reasoning mathematically; Using mathematical representations; Using mathematical aids and tools.
- The statistics/data science toolkit involves: Data stewardship; Handling data; Data representation; Statistical thinking and methods; Probabilistic reasoning; Using computational aids and tools.
- The computing toolkit involves: Computational thinking (including developing algorithms; Using algorithms; Programming; Representing and manipulating data; Safe use of technology; Using a range of aids and tools including emerging technologies.

For detailed competency descriptions, see 5.2 – 5.4.

Competencies for school leavers

A repeated call in workplaces is for 'T-shaped' knowledge - broad across many areas with deep knowledge in one area. The six high-level mathematical and data literacy competencies, and their toolkits, can be further elaborated as competencies for school leavers. The competencies elaborating skills 'for all' and 'for many' provide the broad knowledge and activities that allow students to engage fully and critically in society, and to move into entry-level roles that involve decision-making based on quantitative data. Skills 'for specialists' allow students to develop depth at school. Together they could provide the basis of a new course of study, or show how students studying e.g. computing, could use this specialist knowledge to extend their mathematical and data literacy.

See Section 6 for competencies for all, for many and for specialists.

The feasibility of curriculum models for embedding mathematical and data literacy

Part Two reports an analysis of curriculum models for achieving the competency goals set out in Section 6. Ten case studies were identified that are compatible with these aims (see Section 8, which may be treated as an appendix). These were compared with nine

curriculum models, ranging from 'No change', through a series of changes and additions to the current curriculum and assessment at KS 3, KS 4 and KS 5, to 'A new baccalaureate system'.

Evidence from the case studies suggests:

- There are few precedents for teaching interdisciplinary content within computing and small numbers studying at GCSE reduce its reach.

- The opposite holds: there are several successful trials of teaching computing through other subjects, including some of developing teachers in other subjects to teach mathematical and statistical skills. Gains are noted mainly in social science and science subjects, with little evidence for arts, reducing the reach in a free choice system.

- Curricula that have attempted to raise the profile of what was previously considered a subset of mathematics (for example promoting statistics in New Zealand or modelling in Ontario) do so both by creating separate strands that carry teaching and assessment requirements and by providing trialled teaching materials for teachers to adapt.

- Curricula such as the International Baccalaureate, New Zealand and Ontario that provide multiple pathways for 16-18 study have good take up of pathways related to mathematical and data literacy as well as those leading to STEM careers. All these systems have a minimum requirement for studying mathematics at KS 5.

- Projects and curricula that have successfully embedded aspects of mathematical and data literacy have all included substantive project work. This may be non-assessed, assessed by teachers, also externally moderated or even externally marked. Extra teacher workload associated with assessment is noted, but also recognition of its value.

- Case studies speak to the potential for teaching 11-14 year olds to pose and answer questions based on data relevant to them. They also emphasise the need for teaching.   Stand-alone courses that include a data skills project would need curriculum time and teacher support.

- Integrative, focused and ongoing professional development is a common factor in making effective change. Similarly; an approach involving multiple stakeholders and organisations supports curriculum design and delivery.

-

- Three features of current assessment structures recur as barriers for creating mathematical and data literacy courses for all:
  - the complicated relationship between English teachers, policy makers, inspectors and multiple exam boards;
  - the stifling of any inter-disciplinary teaching activity by high-stakes GCSE assessment at 16;
  - unconstrained subject choice at 14 and 16 creates subject silos and prevents coordinated approaches to planning interdisciplinary study.

Part One

3.   Competencies

For this report we take competencies to be enactments of activities appropriate within the emerging domain of mathematical and data literacy. Taken as a whole, a mathematics competency framework responds to the question 'What does it take for a learner to become a *doer* or *user* of mathematics?', rather than the question 'What does it take to become a knower?' [1]. Individual competenc*ies* add specificity to a general notion of mathematical competence. Thus: "a mathematical competency is someone's insightful readiness to act appropriately in response to a specific sort of mathematical challenge in given situations" [2]. While there is widespread agreement that knowing and doing are interdependent [3], a curricular focus on competencies emphasises a sociocultural understanding that a main purpose of education is for students to play an active role in society.

Interest in mathematics competencies developed in the early 2000s, stimulated by curriculum reform in the US (National Research Council, 2001) and the growth of comparative curriculum/performance studies such as TIMSS 1995 [4]. One reason given for working with competencies is that they allow a focus on processes independently of

---

[1] Mogens Niss et al., 'Conceptualisation of the Role of Competencies, Knowing and Knowledge in Mathematics Education Research', in *Proceedings of the 13th International Congress on Mathematical Education*, ed. Gabriele Kaiser, ICME-13 Monographs (Cham: Springer International Publishing, 2017), 235–48, https://doi.org/10.1007/978-3-319-62597-3_15.

[2] Mogens Niss and Tomas Højgaard, 'Mathematical Competencies Revisited', *Educational Studies in Mathematics* 102, no. 1 (1 September 2019): 14, https://doi.org/10.1007/s10649-019-09903-9.

[3] A. Schoenfeld, 'Learning to Think Mathematically: Problem Solving, Metacognition and Sense Making in Mathematics', in *Handbook of Research on Mathematics Teaching and Learning*, ed. D. Grouws (New York: Macmillan, 1992), 335–70; Bethany Rittle-Johnson, Michael Schneider, and Jon R. Star, 'Not a One-Way Street: Bidirectional Relations Between Procedural and Conceptual Knowledge of Mathematics', *Educational Psychology Review* 27, no. 4 (December 2015): 587–97, https://doi.org/10.1007/s10648-015-9302-x.

[4] Jeremy Kilpatrick, 'Competency Frameworks in Mathematics Education', in *Encyclopedia of Mathematics Education*, ed. Stephen Lerman (Cham: Springer International Publishing, 2020), 110–13, https://doi.org/10.1007/978-3-030-15789-0_27.

subject content areas. This facilitates tracing similar mathematical activities at different levels of schooling, where content statements may appear very different, and thus supports coherent curriculum progressions. It also facilitates comparisons across the curricula of different educational systems. In curricula where content areas and competencies are distinguished, their relationship may be represented via dimensions in an array (Germany, PISA) or facets of a polygon (e.g. the well-known Singapore pentagon).

Between them, content statements and cognitive competencies still only describe part of what is intended as educational outcomes. Some frameworks also include affective statements such as resilience and confidence, and attitudinal statements such as understanding that data shapes our world. This wider view responds to the international movement around 21st century skills for lifelong learning [5].

Competencies are used with many purposes. They are used *normatively* to design the aims, structure, organisation and assessment of curricula; *descriptively* by teachers and researchers to characterize mathematics teaching and learning; *diagnostically* and *metacognitively* in interactions with students [2]. A useful distinction is offered between internal and external use [6]. Competencies for *internal* use support communication between practitioners within and about a discipline. Mathematics competencies for *external* use function as a communication with parents, government, society; they position and justify the role of mathematics in education. Both types of communication are necessary but they may require different emphases. As a relevant example, for pure mathematicians, 'mathematical modelling' may appear as just one of several competencies that characterise mathematics but it is central to applied mathematics and usually emphasised in external communications about why everyone should study mathematics.

Finally, competencies are written at different granularities, ranging from titular noun phrases (e.g. China's core competencies of intuitive imagination, data analysis [7]) to statements of actions with specific qualities, contexts and tools. The elaboration of high-

[5] Malcolm Tight, 'Twenty-First Century Skills: Meaning, Usage and Value', *European Journal of Higher Education* 11, no. 2 (3 April 2021): 160–74, https://doi.org/10.1080/21568235.2020.1835517.

[6] Jens Hojgaard Jensen and Uffe Thomas Jankvist, 'Disciplinary Competence Descriptions for External Use', *Nordic Studies in Mathematics Education* 23, no. 2 (2018): 3–24.

[7] Jing Cheng, Jiansheng Bao, and Dianzhou Zhang, 'From "Two Basics", to "Four Basics" to "Core Mathematics Competencies" in Mainland China', in *Beyond Shanghai and PISA: Cognitive and Non-Cognitive Competencies of Chinese Students in Mathematics*, ed. Binyan Xu, Yan Zhu, and Xiaoli Lu, Research in Mathematics Education (Cham: Springer International Publishing, 2021), 1–13, https://doi.org/10.1007/978-3-030-68157-9_1.

level competencies into, for example, 'specific expectations' with teaching examples, tips, prompts, and sample tasks [8] is a necessary part of curriculum development and – crucially - continues for years as teachers, task designers and quality controllers negotiate what is considered educational achievement in teaching and assessment.

4.   Approach

The purpose of Part One of this report is to survey existing vision literature in mathematics, data science, statistics and computing (MDSC) in order to create a body of knowledge listing the top-level intersectional data/ statistical/ problem-solving/ computing competencies that young adults should have when they leave school or college to prepare them for employment and life.

The evidence base for this synthesis was provided by a literature review and a series of interviews with experts associated with curriculum initiatives in MDSC (see Appendix A), each strand establishing credentials and high-level interpretations for the other.

Sources were purposively sampled based on a combination of:

- Recommendations from experts in MDSC education of papers considered to be significant scholarship that aligns with consensus views.

- Sources derived from research databases, using the *Academic Search Complete*, *British Education Index*, *Education Abstracts*, *Education Research Complete*, and *ERIC* databases, for the period 01/01/2018 to 02/02/2023, applying the search terms:

    i.   "mathematics education" AND "computational thinking" AND curriculum (29 results, which led to 5 relevant papers)
    ii.   "mathematics education" AND (computing or "computer science") AND (statistics or "data science") (85 results, which led to 4 further relevant papers)

- Following up references from key papers and authors.

The question that originally framed our reading was broad: "What does the existing vision literature in MDSC have to say about competencies for citizens and the proposed school curriculum models?" This developed into:

---

[8] Ontario Ministry of Education, *The Ontario Curriculum, Grade 9: Mathematics (2021)*, 2021, https://www.dcp.edu.gov.on.ca/en/curriculum/secondary-mathematics/courses/mth1w.

- What does the existing vision literature in MDSC indicate are areas of consensus *within* and *between* the relevant disciplines?

- What are examples of stable, well-documented and high-performing curricula or curriculum initiatives considered to address these intersections of MDSC?

- What are the implications for competencies for citizens and the proposed school curriculum models? (cf. Part Two)

4.1   Reference points

Expert interviews and an initial review of the literature enabled us to identify examples of stable, well-documented and high-performing curricula or curriculum initiatives. We used these as reference points, comparing them within each discipline to establish what was agreed to be both in scope and possible (a sense of central tendency and spread, perhaps!). We compared them also between disciplines, now focusing on what was common and what was compatible.

In Mathematics the three reference frameworks chosen are: the Danish KOM competency framework (2003), the Ontario Curriculum (2007, revised 2020), the OECD PISA mathematical literacy framework (2003, revised 2012, 2021). The Danish competencies of mathematical learning (KOM) project was chosen because (together with the US NCTM Standards) it is regarded as one of the first, most influential, competency frameworks. It is stable, with subsequent educational initiatives (to reduce inequality, for example) focusing on whole-school enactment and upskilling teachers [9]. KOM's attention is squarely on the discipline of mathematics. In contrast, the 2005 Ontario curriculum exemplifies curricula organised around mathematics as a tool for everyday life [10], corresponding to an aim of Mathematical Futures. Ontario's 2020-21 reforms have since incorporated programming into elementary mathematics. Canadian students out-perform English students on PISA tests, and Ontario is close to the Canadian average [11]. In particular, they are strong on

---

[9] OECD, *Education Policy Outlook 2019: Working Together to Help Students Achieve Their Potential*, Education Policy Outlook (OECD, 2019), https://doi.org/10.1787/2b8ad56e-en.
[10] Cathy Smith and Candia Morgan, 'Curricular Orientations to Real-World Contexts in Mathematics', *The Curriculum Journal* 27, no. 1 (2 January 2016): 24–45, https://doi.org/10.1080/09585176.2016.1139498.
[11] Kathryn O'Grady et al., 'Measuring up: Canadian Results of the OECD PISA 2018 Study' (Council of Ministers of Education, Canada, 2019).

interpreting mathematics in context [12]. Finally, the racy definitions of mathematical literacy are internationally influential and mobilised for large-scale computer-based assessment.

In Statistics, reference frameworks are: the New Zealand Mathematics and Statistics school curriculum [13] (2007, with updates in subsequent years, and being revised 2023), and the GAISE (2007, updated 2020) framework [14]. New Zealand is well regarded internationally for its statistics education [15]. It gives statistics equal footing alongside mathematics with its three inter-related strands of Statistical investigation, Statistical literacy and Probability being developed from the start of primary education through to Year 13. A high proportion of students continue to study statistics beyond compulsory schooling: the New Zealand take-up rate for advanced mathematics or statistics is three times that of England [16]. The *Pre-K–12 Guidelines for Assessment and Instruction in Statistics Education II* (GAISE II) framework is an internationally developed framework that bridges statistics and data science, that officially underpins the statistics education standards of the US National Council of Teachers of Mathematics (NCTM) and is endorsed by the American Statistical Association. GAISE II has recommendations and exemplar teaching and assessment activities that support the development of statistical reasoning and statistical literacy across three levels.

In curricular terms, **data science** is the newest arrival on the block (post-2010). Moreover, data science has not been an academic innovation but is an instance where the education sector is playing catch up with a rapidly expanding industry and responding to growing demand/clamour for data science courses at tertiary level [17], and recognising its value for

[12] Pierre Brochu et al., *Measuring up, Canadian Results of the OECD PISA Study: The Performance of Canada's Youth in Mathematics, Reading and Science: 2012, First Results for Canadians Aged 15* (Toronto: CMEC, 2014), http://site.ebrary.com/lib/celtitles/docDetail.action?docID=10831057.
[13] https://nzcurriculum.tki.org.nz/The-New-Zealand-Curriculum/Mathematics-and-statistics OR https://new.censusatschool.org.nz/new-zealand-statistics-curriculum/
[14] A. Bargagliotti, C. Franklin, P. Arnold, R. Gould, S. Johnson, L. Perez, and D. Spangler, Pre-K-12 Guidelines for assessment and instruction in statistics education (GAISE) report II, American Statistical Association, Alexandria, VA, 2020. www.amstat.org/asa/files/pdfs/GAISE/GAISEIIPreK-12_Full.pdf
[15] Sharleen Forbes, 'The Coming of Age of Statistics Education in New Zealand, and Its Influence Internationally', *Journal of Statistics Education* 22, no. 2 (July 2014): 2, https://doi.org/10.1080/10691898.2014.11889699.
[16] Jim Ridgway, 'Implications of the Data Revolution for Statistics Education: The Data Revolution and Statistics Education', *International Statistical Review* 84, no. 3 (December 2016): 528–49, https://doi.org/10.1111/insr.12110.
[17] David Donoho, '50 Years of Data Science', *Journal of Computational and Graphical Statistics* 26, no. 4 (2 October 2017): 745–66, https://doi.org/10.1080/10618600.2017.1384734.

school education [16]. School-based courses are less developed [18] and have greater challenges. Given the industry-driven emergence of data science, our reference points for data science started with the Standards of the UK's Association for Data Science Professionals [19], contrasting them with GAISE II and three research-informed data science frameworks developed for use in schools.

The *International Data Science in Schools Project* (IDSSP) was set up in 2018 by statisticians and computer scientists from across the world. Envisaged school-based trials were disrupted by the Covid-19 pandemic but the IDSSP frameworks [20] (for curriculum and teacher professional development) have been widely endorsed and influential.

The UCLA and Los Angeles Unified School District's *Introduction to Data Science* (IDS)[21] is a data-driven course for high school that brings together statistical and computational thinking, civic engagement and mobile technology to support 'students to think critically about and with data' [22]. First implemented in 2014, the course has engaged over 42,000 students in a total of 151 high schools across 74 districts in the USA.

Project Data Science and Big Data at School (ProDaBi) [23], developed by researchers at Paderborn University, Germany has been developing and trialling courses since 2018. The courses span all of secondary school (Grades 5 through 12 in Germany) and the project includes development of teaching materials, professional development courses for teachers and related evaluation and research.

Finally, for **computing**, the reference frameworks are the 2013 national curriculum in England [24], the GCSE, A level and league-table approved computing qualifications in England and Wales, the US Common Core Standards for Computer Science [25] (and the US Advanced Placement Program (AP) Computer Science courses [26]), and the Singapore

---

[18] Neville Davies and Neil Sheldon, 'Teaching Statistics and Data Science in England's Schools', *Teaching Statistics* 43 (2 June 2021): S52–70, https://doi.org/10.1111/test.12276.

[19] AfDSP, 'Data Science Standards' (Alliance for Data Science Professionals, 2022), https://alliancefordatascienceprofessionals.co.uk//standards.

[20] http://www.idssp.org/files/IDSSP_Frameworks_1.0.pdf

[21] https://www.idsucla.org/

[22] https://www.idsucla.org/other-mobilize-curricula

[23] https://www.prodabi.de/en/

[24] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/239067/SECONDARY_national_curriculum_-_Computing.pdf

[25] https://www.codelicious.com/k-12-computer-science-standards

[26] https://apcentral.collegeboard.org/courses/ap-computer-science-a and https://apcentral.collegeboard.org/courses/ap-computer-science-principles/course

curriculum based around an O level and A level in Computing. As we considered digital technology skills alongside computing, we also looked at Dr Erika Kispeter's *Digital Skills and Inclusion Research Working Group Evidence Brief*[27] which looked at 'What digital skills do adults need to succeed in the workplace now and in the next 10 years?'

4.2 Examining the Intersections

The Royal Society project brief has the title "Intersection of mathematics and computing, statistics, and data science". Given the different granularities at which one can list competencies, and the variety of subject definitions in play, the intersection could range from the entirety of these subject areas to the empty set. This is not simply a semantic issue or a territorial dispute but is crucial for how we delineate the subject matter for this report. Our survey of the literature revealed three areas on which opinions can clearly differ but on which we needed to take a position. In this section we examine each in turn:

1. The relationship between statistics and data science.

2. The curriculum balance of pure and applied mathematics.

3. The role of computing and programming in mathematical and data literacy.

4.2.1   The relationship between statistics and data science

Analysing and making sense of data, and consequently data literacy, are central elements of both data science and statistics. There are contrasting views as to whether data science is the same as statistics.

Tukey (1962) [29] and Owen (2015) [28], would say that the distinction is fundamental, cultural and much needed. Owen draws attention to the extra importance in data science of software engineering, with teams bringing together diverse skills working within live data, at scale and in real time. While the rise of computational power, the opening up of vast fields of data and proliferation of data types are at the heart of the emergence of data science, that does not encompass the whole story. Tukey argued that data analysis should be viewed as a new science rather than an application of mathematical statistics: data science "must seek for scope and usefulness rather than security" and "must use

---

[27] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/807831/What_digital_skills_do_adults_need_to_succeed_in_the_workplace_now_and_in_the_next_10_years_.pdf

[28] https://medium.com/@srowen/what-50-years-of-data-science-leaves-out-2366c9b61d3d

mathematical argument and mathematical results as bases for judgment rather than as bases for proof or stamps of validity" [29].

From this perspective, mathematics and statistics are implicated in data science (providing the representations and tools for data manipulation and analysis) but equally important are domain-specific knowledge, computing and the use of technology, and aspects of the data such as its nature, size, provenance, accuracy, reliability, ethical use, and the needs of the industry/clients posing the problems.

Such arguments for the distinctiveness of data science tend to treat statistics as part of mathematics. This perspective is common in school: in many curricula, including the English one, probability and statistics are strands of mathematical content knowledge. Similarly, the OECD's PISA 2021 framework states that mathematical reasoning includes "both mathematical (deductive) and statistical (inductive) type reasoning". [30]

Nevertheless, this narrower view of school statistics as a subset of mathematics does not represent how it has been defined by later statisticians, or in Higher Education, as "the only discipline whose primary activity is finding meaning in data". [31] This wider view does include those valued elements of data provenance, accuracy, usage, etc that a purely mathematical lens might abstract away. Moreover, the omission of these in the enacted school curriculum is identified as a barrier to good teaching of statistics. [32] **Thus, for the purposes of this future-facing paper on school curricula, we have preferred to treat data science and statistics as cognate.** At school level, what the disciplines have in common far outweighs the differences that exist in approach and content. Those divergences which are attended to by school-level study are often in the finer detail met only by specialists; for example, hypothesis testing based on a probability appears in the 2017 A level Mathematics

---

[29] John W. Tukey, 'The Future of Data Analysis', *The Annals of Mathematical Statistics* 33, no. 1 (1962): 6.

[30] OECD, 'PISA 2021 Mathematics Framework (Second Draft)', 2018, 3, https://www.oecd.org/pisa/sitedocument/PISA-2021-mathematics-framework.pdf.

[31] Robert Gould, 'Toward Data-Scientific Thinking', *Teaching Statistics* 43, no. S1 (July 2021), https://doi.org/10.1111/test.12267.

[32] ACME and the Royal Statistical Society (RSS), 'Embedding Statistics at A Level: A Report on Statistical Requirements and Assessment across A Level Courses in Biology, Business, Chemistry, Geography, Psychology and Sociology', 2015, http://www.acme-uk.org/media/32719/embeddingstatistics-at-a-level.pd; Neville Davies and Neil Sheldon, 'Teaching Statistics and Data Science in England's Schools', *Teaching Statistics* 43, no. S1 (2021): S52–70, https://doi.org/10.1111/test.12276.

curriculum, whereas a data science approach might focus instead on using data to determine prediction intervals.

This approach of considering statistics/data science together is supported by the literature about innovations in statistics and data science education, which provide similar and certainly compatible messages [33]. Data science aligns with statistics in its central concern with gathering, processing and analysing data to generate insights and predictions in complex situations. In the reference curricula for data science this translates to a central role for the statistical investigation cycle, and the use of statistical representations and tools for analysis, communication and evaluation. Much of this may also be automated via technology, resulting in a greater overlap between data science and computing than *was* the case with previous school statistics but entirely compatible with its actual use beyond school. The consensus amongst school-based statistics/data science educators is that **data science, with its contextual relevance and ethical aspects, provides the opportunity for an engaging, data-driven education that includes all the statistics needed by school leavers.**

4.2.2   The curriculum balance of pure and applied mathematics

The importance of extra-mathematical contexts in statistics/data science raises questions about how such contexts feature in school mathematics curricula, and whether there are opportunities for data-driven mathematics outside what is normally deemed to be 'statistics' or 'data handling' content. A curriculum that aims to promote the applications of statistics via data science should pay equal attention to applications of mathematics.

The current mathematics specification of the English national curriculum is largely context-free. Subject content is listed with the overarching statement that, through the mathematics content, pupils should be taught to develop fluency, reason mathematically and solve problems. The value of applied mathematics is recognised, since some problems must be set in financial contexts and others may involve "model[ling] situations mathematically [34]. However there is no explicit statement of the relationship between pure and applied mathematics or of the range of situations in which mathematics should/could

---

[33] As a sign of merging messages, in 2021, the American Statistical Association's *Journal of Statistics Education* evolved to become the *Journal of Statistics and Data Science Education*.
[34] Department for Education, *Mathematics Programmes of Study: Key Stage 4. National Curriculum in England*, 2014, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/331882/KS4_maths_PoS_FINAL_170714.pdf.

be applied. The lack of an explicit statement has implications for what is assessed. Comparison with the reference curricula shows that other choices can be made.

Out of the three reference curricula for mathematics, the high-level Danish KoM framework chooses to distinguish problems arising within mathematics itself and those which answer questions relating to an external, often real-world domain, naming these 'mathematical problem-handling' and 'mathematical modelling' (see Figure 1). The contrast and balance between them has been retained and clarified over time [35] suggesting that the explicit distinction is deemed valuable. Problem-handling includes devising and implementing strategies for problems that are already formulated mathematically but are non-routine to the problem solver. Modelling has at its heart mathematical actions that take account of the purposes, data, facts, features and properties of an extra-mathematical domain. It also gains extra emphasis as a way for students to understand that an essential feature of mathematics as a discipline is how it is applied within other fields of practice.
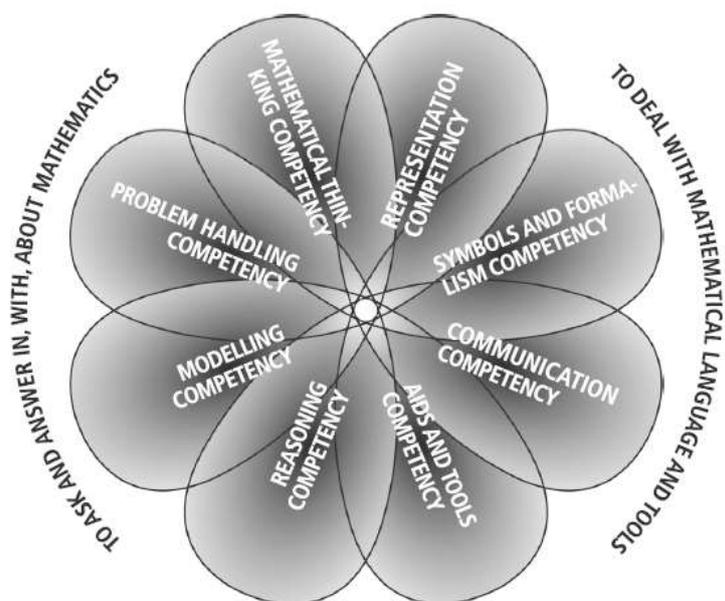
Figure 2 A visual representation of the eight KOM mathematical competencies [35]

In contrast, the Ontario curriculum "balances" fundamental concepts and skills in mathematics with opportunities to apply them and requires students at all grades "to make connections between their lived experiences, mathematical concepts, other subject areas, and situations outside of school." (The Ontario Curriculum, Grades 1-8: Mathematics, 2020, p65). Grade 11 course titles such as *Mathematics for Work and Everyday Life*, or *Functions and Applications* are a good representation of this more integrated approach. Questions

---

[35] Niss and Højgaard, 'Mathematical Competencies Revisited'.

and data pertaining to real-world situations are pervasive as a purpose and a motivation for learning mathematics and, largely, also for developing new mathematics. In this curriculum, data largely serves the purpose of illustrating applications of mathematics. A less realistic example is using Pythagoras's theorem to calculate ramp heights (where practitioners would use a table, an app, or CAD software); a more realistic one is using exponential graphs to predict growth.

Similarly, the OECD's PISA framework has moved towards integrating pure and applied mathematics. The Formulate-Employ-Interpret modelling cycle has been central to its definition of mathematical literacy since 2003. In 2022, however, it added mathematical reasoning within that cycle and argued that solving problems within mathematics can be considered as applying mathematics in a scientific domain, in this case that of mathematics itself. As in Ontario, the questions use real-world contexts but are not necessarily driven by data.

The similarity of approach between the Ontario curriculum and PISA may account for the Ontarian students' relative success in the international assessment. In our opinion it achieves simplicity by ignoring some distinctions that are relevant to this study of the intersection of MDSC. Specifically, it does not provide a means of examining the different roles played by numerical examples or data in mathematics, for example distinguishing between the nature and provenance of numbers generated by a geometric sequence and those obtained by measurement or sampling. **Our resolution in this report is to focus on identifying mathematics competencies most closely associated with data-driven questions, whether from pure or applied mathematics.** We have also thought it most helpful to consider a smaller set of close associations, i.e. those competencies that *must* be included in a data-driven strand or version of the mathematics curriculum. While mathematics for its own sake can be inspiring and important, it is already widely enacted and laid out in the English curriculum (e.g. as symbolic fluency, circle theorems).

4.2.3 The role of computing and programming in mathematical and data literacy

The role of methods - computations/calculations/algorithms - in how we transform quantitative inputs into outputs (and similarly in how we process data in data science) is integral to mathematical and statistical problem solving. Commenting on the role of computation in determining what we teach, Cobb (2007) observes that "historically, we have always tended to underestimate the extent to which what we are able to do shapes

what we think we ought to do." [36] This, coupled with modern computing having evolved as a branch of applied mathematics, makes it difficult to disaggregate/delineate the intersection of computing with mathematics.

Moreover, this enactment of computation/algorithmics can play out at different levels. In mathematics, students can be taught to follow algorithms, as in the widespread 'I do, we do, you do' teaching routine and/or be tasked to choose and adapt them. In computing, students could write and implement their own programs, work with user-friendly formats that permit some adaptation (e.g. using Python via a menu-based Jupyter Notebook) [37], or the computation/algorithm can be packaged and treated as a 'black box', whose outputs are more relevant than understanding its workings. This variation (in how we compute) can be represented as a spectrum of computing, with programming at one end and use of pre-packaged technology at the other end.

Policy and pedagogic issues add further tensions to the spectrum discussed above. In 2013, the English national curriculum shifted from 'ICT' to 'computing' (including using two or more programming languages). The curriculum is specified in detail for Key Stage (KS) 3, which may in some schools be only 2 years, and is not assessed. Some pupils will be taught either in a cross-curricular fashion or within a potentially optional subject up until the end of KS 4. A subset of these may choose to study computing up until the end of KS 5. The changes from ICT to Computing are still becoming established and provision in schools can be highly variable.

Our current readings and consultations have produced no consensus as to how much programming is needed for school leavers, where the balance should lie between text-based and visual/block-based programming, and whether programming should be developed within computing or in context. **Our resolution is that, for schools, we will consider the use of technology and computing (which includes programming) as separate but related**.

---

[36] George W Cobb, 'The Introductory Statistics Course: A Ptolemaic Curriculum?', *Technology Innovations in Statistics Education* 1, no. 1 (12 October 2007): 10, https://doi.org/10.5070/T511000028.

[37] R. Biehler and Y. Fleischer, Introducing students to machine learning with decision trees using CODAP and Jupyter Notebooks, Teaching Statistics 43 (2021), S133–S142. https://doi.org/10.1111/test.12279

4.3 Terminology

We end this section with some notes on terminology in this report.

Schools and colleges: in 2021, 38% of 16-17 year olds studied full-time in colleges, 38% in state-funded schools and 6% in independent schools (with 16% not in education) [38]. It would be appropriate to refer to the competencies needed by school and college leavers. Nevertheless, for brevity, we use 'school' in its wider sense to refer to both.

Mathematical and data literacy: for the reasons given above, our use of 'mathematics' does not include 'statistics'. We have adopted the terminology 'mathematical and data literacy' as a pragmatic label with antecedents in mathematics, statistics and computing. In our proposed framework, we will consider separately the problem solving toolkits provided by mathematics, statistics/data science and computing.

Computing: *Computing and Computer Science* are often used interchangeably in higher education and industry. Within the UK – and specifically English and Welsh - systems, Computer Science is generally used to refer to the problem solving, computational thinking, programming-based part of the subject area. [39] In this report, computing is used as an umbrella term to include Computer Science, Information Technology (IT) and Digital Literacy. [40]

Disciplines and domains: we have referred to mathematics, statistics, data science and computing as 'disciplines', given that they have their own academic and professional standards. We note that professional data scientists argue that their work is truly inter-

---

[38] Department for Education, 'Participation in Education, Training and Employment Age 16 to 18, Calendar Year 2021', 2021, https://explore-education-statistics.service.gov.uk/find-statistics/participation-in-education-and-training-and-employment.

[39] https://www.computingatschool.org.uk/resource-library/2014/september/computing-in-the-national-curriculum-a-guide-for-secondary-teachers

[40] After the reboot: computing education in UK schools, 2017, Royal Society. https://royalsociety.org/topics-policy/projects/computing-education/

disciplinary. We follow them by referring to knowledge of the context where analytics will be applied as 'domain knowledge' [41].

Variation, covariation and variability: Statistics is sometimes described as the study of variation [42]. 'Variation' is also used widely within mathematics education: to name a learning theory (variation theory [43]), to describe a phenomenon underpinning functions (covariant relationships [44]) or as a big idea related to abstraction and generalisability (invariance and change [45]). Despite these many meanings variation is a useful umbrella term.

In statistics, 'variability' which occurs due to chance, to errors or to sampling, depending on context [46]. The mathematical version of variation is 'co-variation', that is studying the possible relationships between variables that change together. Given a set of numbers alone, without context or purpose, either approach can be useful and appropriate.

Extending this, we argue that the intersection of mathematics, statistics/data science and computing is concerned with the study of variation and generalisability, with precision and context and technology <u>and</u> ethics.

5.   Competencies at the intersection

We consider mathematical and data literacy to be defined by common ways of working within mathematics, statistics/data science and computing, where we attend not only to

---

[41] Institute of Analytics, 'IoA Analyst Competency Framework', 2022, https://ioaglobal.org/wp-content/uploads/2022/06/The_IoA_Analyst_Competency_Framework.pdf; AfDSP, 'Data Science Standards'.

[42] e.g. 'Understanding variation as the heart of statistics.' OECD, 'PISA 2021 Mathematics Framework (Second Draft)', p15.

[43] Angelika Kullberg, Ulla Runesson Kempe, and Ference Marton, 'What Is Made Possible to Learn When Using the Variation Theory of Learning in Teaching Mathematics?', *ZDM* 49, no. 4 (1 August 2017): 559–69, https://doi.org/10.1007/s11858-017-0858-4.

[44] Anne Watson, *Key Ideas in Teaching Mathematics: Research-Based Guidance for Ages 9-19* (Oxford, United Kingdom: Oxford University Press, 2013).

[45] Ilya Sinitsky and Bat-Sheva Ilany, *Change and Invariance: A Textbook on Algebraic Insight into Numbers and Shapes*, 2016, https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1429092.

[46] ACME and the Royal Statistical Society (RSS), 'Embedding Statistics at A Level: A Report on Statistical Requirements and Assessment across A Level Courses in Biology, Business, Chemistry, Geography, Psychology and Sociology'.

the interplay at the intersections but also to what are necessary complementary aspects for future school leavers. The competencies that follow are organised with this in mind, including to allow consideration of different curricular structures. Our Intersectional competency framework is a simplification, although not so simple as to obscure real differences of perspective. We have sought to integrate where sensible but also to show where differences and possibilities exist. Thus we have distinguished between mathematical, statistical and computational approaches, or 'toolkits' for solving problems. It is by no means the only possible way of slicing this curricular space, but we suggest it is one that should be intelligible and agreeable to the main stakeholders. Deciding on a curriculum framework and then developing it – a process of years – will involve re-organisation as well as detailing expectations at the intersections of content, competencies, dispositions and experiences.

We propose an alignment of mathematics, statistics/data science and computing around the overarching activity of posing and answering questions. This activity allows for flexibility in what one asks questions about, the types of information you draw on, the tools and methods you use and the oversight you bring; choices of which contribute to leaning the enquiry closer either to what has been called mathematics, or data science or computing. Nevertheless, the activity remains recognisably the same in each. **The six high-level competencies shown in Figure 3 are common across the literature reviewed. They are related but different, and each is necessary for the resulting activity to be considered as characteristic of the intersection of mathematics, statistics data science and computing.**
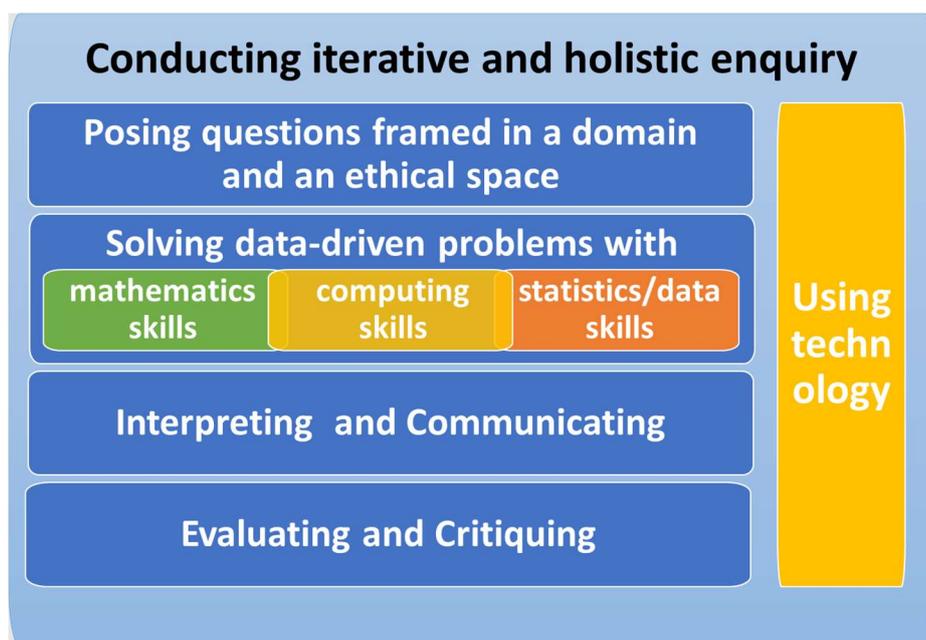
Figure 3 Mathematical, computational and data literacies as posing and answering questions

We continue this section by briefly introducing the six intersectional competencies encapsulated by this framework for mathematical, computational and data literacy. In 5.1 we show how this framework is positioned in relation to the reference frameworks in mathematics and statistics/data science, which serves to explain the nature of the intersectional competencies further. In 5.2 - 5.4 we unpack the competencies in mathematics, computing and statistics/data science needed to solve data-driven questions. In 6 we then elaborate how the competencies could be achieved at different levels to support school-leavers' choice of work and study focus.

Six intersectional competencies

Iterative and holistic enquiry: conduct investigations using the full mathematical/statistical/computational enquiry cycle in situations enabling access to multivariate data and technological tools.

Posing questions: pose and understand questions framed in (extra-mathematical) domain knowledge and an ethical space that anticipate the properties of variation, complexity or uncertainty inherent in data and the analytic methods available.

Solving data-driven problems: devise and implement strategies that may be non-routine to the problem solver, drawing on a repertoire or 'toolkit' of appropriate concepts, analytic methods, reasoning, representations and aids. There are differences between what is foregrounded in a toolkit that forms a basis for data and statistical skills, one that forms a basis for mathematics, and one for computing, which we elaborate in sections 5.2, 5.3 and 5.4 below.

Interpreting and communicating: report and justify findings to others using language, displays and measures appropriate to the domain and the audience.

Evaluating and critiquing: compare different measures, methods, interpretations and modes of communication; discuss their own and others' choices and the implications for ethics and for confidence in the results. Reflect on and make judgements about the appropriateness of their own and others' approaches and conclusions.

Using technology: as technologies enter classrooms, it will often be appropriate to use technology-supported methods as part of problem-solving, interpreting, communicating and critiquing. Students should have opportunities to engage in guided enquiry with access to the internet and computer-based representational tools.

5.1   Positioning the competencies amongst other mathematics and statistics/data science frameworks

Why posing and answering questions?

The framework that we propose is based overall on posing and answering questions. This is not currently foregrounded in the English mathematics curriculum but is a high-level priority in all three mathematics and data skills/statistics reference curricula. 'Posing and answering questions' is one of the two overarching competences offered by the Danish KoM framework, with the other being 'the ability to deal with mathematical language and tools'.

Since 2003, PISA's approach to mathematical literacy has similarly started with a cycle based on formulating models or problems in a context, employing mathematical methods, interpreting and evaluating the results. The 2022 PISA framework retains this cycle and adds a new emphasis that these processes are underpinned by mathematical reasoning, described as "The ability to reason logically and present arguments in honest and convincing ways"[47].

The Ontario curriculum is organised differently and focuses on principles for effective learning. Throughout, it balances understanding *and* application of concepts and skills. For example, the following premise appears in the Grade 1-8 curriculum:

> "It is based on the belief that all students learn mathematics most effectively when they develop a solid understanding of the fundamental concepts and skills in mathematics and are given opportunities to apply these concepts and skills as they solve increasingly complex tasks and investigate mathematical ideas, applications, and situations in everyday contexts." [48]

By grades 11 and 12, the curriculum "embeds the learning of mathematics in the solving of problems based on real-life situations" [49]. Answering questions derived from extra-mathematical domains is thus central to this curricular approach. Posing such questions is

---

[47] https://pisa2022-maths.oecd.org/#Mathematical-Reasoning

[48] Ontario Ministry of Education, *The Ontario Curriculum, Grades 1-8: Mathematics (2020)*, 2020, 65, https://www.dcp.edu.gov.on.ca/en/curriculum/elementary-mathematics.

[49] Ontario Ministry of Education, *The Ontario Curriculum, Grades 11 and 12: Mathematics (2007)*, 2007, 4, https://www.edu.gov.on.ca/eng/curriculum/secondary/math1112currb.pdf.

described as an ability that students develop over time (ibid., p.6) and one that "develops their math agency" (Grades 1-8, p.75).

The computing literature suggests that students should show creativity in creating and combining artefacts to solve a problem with goals specified by themselves and others [50]. Developing a computer program involves iterative cycles of posing requirements, designing, coding, testing, debugging and feeding back to modify/extend requirements. Computer science curricula such as the one in England may not centre-stage this explicitly in terms of 'posing and answering data-driven questions' but even in the England curriculum it is implicit in the very first expectation of the Key Stage 3 content: "design, use and evaluate computational abstractions that model the state and behaviour of real-world problems and physical systems" [51]. The first of six practices in the US Advanced Placement Program (AP) *Computer Science Principles* course, is Computational Solution Design: Design and evaluate computational solutions for a purpose. It also has Creative Development as its first foundational 'Big idea', which involves "a formal iterative design process or a less rigid process of experimentation. While using either approach, developers will encounter phases of investigating and reflecting, designing, prototyping, and testing." [52]

All the reference statistics/data science curricula include posing and answering questions as a framing concept. There are a number of representations for the statistical or data enquiry cycle. We share two here:
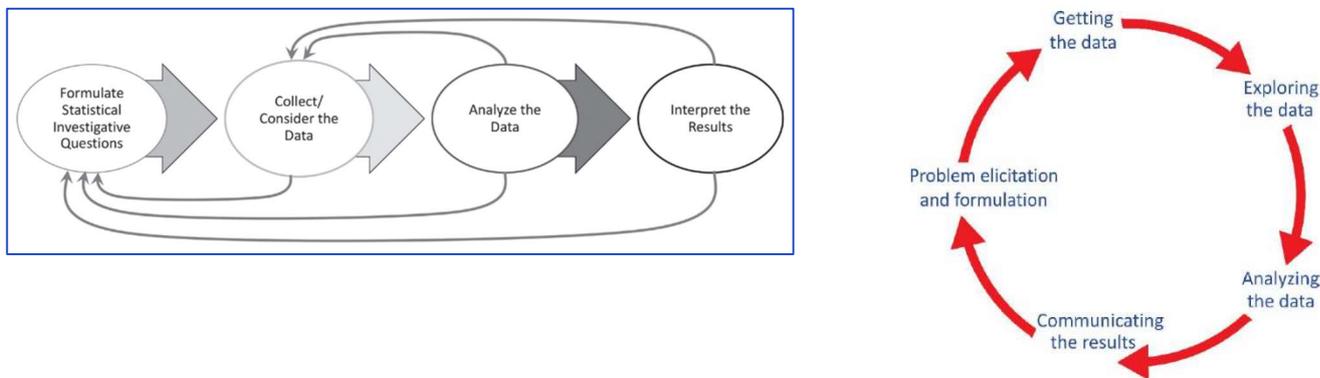
Figure 4 GAISSE II data cycle [14] (left) and IDSSP data cycle [20] (right)

[50] Shuchi Grover and Roy Pea, 'Computational Thinking in K–12: A Review of the State of the Field', *Educational Researcher* 42, no. 1 (1 January 2013): 38–43, https://doi.org/10.3102/0013189X12463051.
[51] DfE, Computing programmes of study: key stages 3 and 4, 2013.
[52] AP Computer Science Principles Course and Exam Description, 2020, p. 17. https://apcentral.collegeboard.org/courses/ap-computer-science-principles/course

Moreover, it is from this perspective that including a competency *Conducting iterative and holistic enquiry* becomes necessary. Not only is each step of the cycle a competency but being able to plan for and enact the cycle is an essential and overarching competency. Iterations are not simply repeats of the full cycle but there may be multiple overlapping sub-cycles, as in the GAISE II diagram.

As a contrast, we note that Grillenberger and Romeike, working solely on the intersection of computing and data science, developed a more minimal competency framework that focuses solely on analysis and evaluation. However they comment that they are unusual in specifically excluding the aspect '*problem – ask questions from data*' which they regard as "being out-of-scope of the model, but without neglecting its importance" [53]. Their high-level processes, such as cleansing and optimizing, fit rather into our statistics/data science or computational toolkits. This is not surprising since they are integrating only two disciplines.

It has been argued within mathematics education (e.g. PISA 2021 [54]) that it is often not necessary to engage in every part of the modelling cycle, especially in the context of an assessment. For statistics/data science, however, the anticipation and refinement of questions, data representations, analytic methods and convincing arguments for end users are unavoidable [55]. Recent mathematics education research also stresses the role of anticipation in modelling, so that "modellers must be aware of the specific role of sub-competencies and how they interrelate in parallel and in sequence. [...] This means that modellers must be able to both enact modelling subcompetencies and synthesise these into a global competency." [56]

The centrality of domain knowledge and ethics
Two aspects that substantially distinguish data science from traditional mathematics are the integral roles of (a) domain-specific knowledge and (b) the ethics, provenance and stewardship of data. Accounts of the modelling cycle (e.g. in PISA 2021) can make a virtue of

[53] Andreas Grillenberger and Ralf Romeike. 2018. Developing a Theoretically Founded Data Literacy Competency Model. In Proceedings of the 13th Workshop in Primary and Secondary Computing Education (WiPSCE '18), October 4–6, 2018, Potsdam, Germany. ACM, 7. https://doi.org/10.1145/3265757.3265766

[54] OECD, 'PISA 2021 Mathematics Framework (Second Draft)'.

[55] Anna Bargagliotti et al., 'Pre-K-12 Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report II' (American Statistical Association and National Council of Teachers of Mathematics., 2020), https://hdsr.mitpress.mit.edu/pub/cqncbp3l.

[56] Vince Geiger et al., 'Developing a Task Design and Implementation Framework for Fostering Mathematical Modelling Competencies', *Educational Studies in Mathematics* 109, no. 2 (February 2022): 317, https://doi.org/10.1007/s10649-021-10039-y.

abstraction, referring to context only at the beginning or end of a process, and with no acknowledgement that domain knowledge will be required. Similarly, in GCSE mathematics, the 'Interpret' part of a data handling question will typically have 1 mark available, and students are routinely advised to ignore it.

In statistics/ data science, data come with a context and the frameworks take an opposing position: "In mathematics, context obscures structure. In data analysis, context provides meaning" [57]. The IDSSP project adds the arguments that school-level data science is not only concerned with teaching analytic techniques but teaching how data is used in society, *and* that this is motivating to young people: "At all times, the focus will be on questions, problems and data that are meaningful to their lives and attendant social and ethical issues that arise in acquiring and working with data" [58].

Situating problems in a non-mathematical context, and engaging with the potential messiness of that context, is therefore essential for a curriculum that addresses data science. Indeed, there is considerable evidence from mathematics education that teachers too readily assume that students can make connections between domain and mathematical knowledge. Teaching contextualised problem solving, for example in apprenticeships or Core Maths, requires time spent discussing domain-specific issues, variables and representations [59]. An antecedent lies in the Royal Society's recommendations for embedding mathematics in T-levels through "the use of mathematical models that provide mathematical insight into real (and complex) situations and problems that are simplified by making a number of assumptions." [60]

---

[57] Cobb & Moore, 1997, p. 803 cited in Bargagliotti et al., 'Introducing GAISE II', 7.

[58] International Data Science in Schools Project (IDSSP) Curriculum Team, 'Curriculum Frameworks for Introductory Data Science', 2019, 11, https://www.idssp.org/pages/framework.html. www.idssp.org/files/IDSSP_Frameworks_1.0.pdf

[59] Geoff Wake, 'Making Sense of and with Mathematics: The Interface between Academic Mathematics and Mathematics in Practice', *Educational Studies in Mathematics* 86, no. 2 (June 2014): 271–90, https://doi.org/10.1007/s10649-014-9540-8; Jennie Golding, Cathy Smith, and Mick J Blaylock, 'Learning to Teach Contextualized Problem-Solving in a Non-Calculus Mathematics Pathway', *Teaching Mathematics and Its Applications: An International Journal of the IMA* 37, no. 2 (4 June 2018): 69–83, https://doi.org/10.1093/teamat/hry009.

[60] Royal Society's Advisory Committee on Mathematics Education's (ACME) Post-16 Contact Group, 'Mathematics for the T Level Qualifications: A Rationale for General Mathematical Competences (GMCs)', 2019, https://royalsociety.org/~/media/policy/topics/education-skills/Maths/Mathematics%20for%20the%20T%20Level%20Qualifications%20-%20a%20rationale%20for%20GMCs.pdf?la=en-GB.

The ethical space of the domain is a necessary part of the domain knowledge. Indeed, ethics is the one compulsory area required by Data Science professionals, including knowledge of ethical issues relating to the domain and to the data. Undergraduate frameworks agree that: "Academic institutions should ensure that ethics is woven into the data science curriculum from the beginning and throughout." [61] It is also an area of consensus in the literature on digital literacy: "The digital world also expects its citizens to use the technology and to respond to the ethical issues of using technology appropriately."[62] Interviewees strongly recommended extending this to school curricula. One noted some children can and do learn to code by studying alone, but they do not learn ethics or the need for it.

This evidence informs our second competency *Posing questions framed in a domain and an ethical space.* It is the latter part of that competency that extends what has previously been attempted in mathematics. It is perhaps worth noting again that these competencies anticipate each other: their order is not fixed. Enquiry may not always start with question formulation. This is particularly relevant for data science in industry, where available data may be the starting point for a cycle. In education, students may use given data as a way of starting a discussion of domain issues, and then pose their own questions.

Two other competencies, *Interpreting and Communicating* and *Evaluating and critiquing*, appear universally, more or less in this form, across the literature about data-driven mathematics, statistics/data science and computing [63]. As with problem posing, there is an added dimension to the traditional mathematics approach because data science requires analysts to report and justify findings with a sense of audience and using language, displays and measures appropriate to the domain [64], and not only to mathematics. Critique extends to comparing algorithms, interpretations, modes of communication and discussing the implications for ethics and for confidence in the results.

---

[61] National Academies of Sciences, Engineering, and Medicine, *Data Science for Undergraduates: Opportunities and Options* (Washington, D.C.: National Academies Press, 2018), 3, https://doi.org/10.17226/25104.

[62] Meize Guo and Anne Ottenbreit-Leftwich, 'Exploring the K-12 Computer Science Curriculum Standards in the U.S', in *Proceedings of the 15th Workshop on Primary and Secondary Computing Education* (WiPSCE '20: Workshop in Primary and Secondary Computing Education, Virtual Event Germany: ACM, 2020), 1, https://doi.org/10.1145/3421590.3421594.

[63] The 2007 England curriculum, the International Baccalaureate (IB) MYP and Diploma Programme, and the US Common Core are examples of curricula that treat these as key competencies. This is also true for the statistics and data science curricula and frameworks we have referenced (New Zealand, California, GAISE II, IDSSP, IDS, ProDaBi).

[64] https://alliancefordatascienceprofessionals.co.uk/documents/AfDSP_Standards_June22.pdf

Solving data-driven problems with mathematical, statistics/data science and computing toolkits

A central competency in this framework is *Solving data-driven problems*. At a high level, the constituent competencies of problem-solving (decomposition, pattern recognition, abstraction, pattern generalisation) are common to mathematical or statistical thinking. With the addition of algorithm design, they also constitute computational thinking. [65] However, staying with the surface similarity of these competencies would mask how they play out quite differently in practice. Here they are represented as distinctive aspects of problem solving; any could be foregrounded depending on the requirements of the problem and the skills of the solver. A conscious decision was made to keep the problem-solving (and analysis) toolkits of mathematics, computing and statistics/ data science separate. By toolkit, we don't just mean a collection of techniques, but the ways of thinking and reasoning, and the language, concepts and tools inherent to those disciplines. We exemplify these in Sections 6, 7 and 8, but at this stage we would like to draw attention to the following contrasts.

Prior to the arrival of modern computing, mathematical methods were our main tool for grappling with complexity, by 'finding the simple in the complex' [66]. Computing on the other hand enables us to hold, process and infer from large and complex data sets. This sets up a fundamental difference in mathematical and statistical problem solving. Mathematics flattens the data, abstracting away the context to hone in on structural meaning. In statistics/data science, context and data lead. Mathematical reasoning seeks to justify and prove, whereas statistical reasoning supports inferences and makes predictions.

An illustration of this is curve-fitting. In mathematics, curves are fitted to data to explain and classify the co-variation of target variables. The purpose is to find and use a functional relationship. Students may ignore other variation (seen as 'error') or seek to minimise it by optimising the fitting based on relevant criteria (e.g., least squares). In statistics and data science, variation is the subject of study – not to be minimised but to be anticipated, accounted for, analysed and interpreted.

---

[65] https://www.computingatschool.org.uk/resource-library/2014/september/computing-in-the-national-curriculum-a-guide-for-secondary-teachers

[66] Jacob Schwartz in his essay on Computer Science in the book Discrete Thoughts: Essays on Mathematics, Science and Philosophy https://doi.org/10.1007/978-0-8176-4775-9

Using technology

Our three reference mathematics frameworks were chosen for stability, and it is not therefore surprising that they are not built around using digital technologies. In the KOM framework, digital tools appear within the second overall competence of "Handling mathematical constructs, language and tools". The Ontario curriculum has recently introduced 'coding' as a strand of algebra for grades 1-8, involving writing, executing, reading and altering programmes of increasing sophistication and efficiency. [48] PISA is developing computer-based assessment. The global direction of travel in mathematics is for wider inclusion of technology in both teaching and assessment, an issue that is being studied in a complementary report for the Royal Society. Moreover, education researchers increasingly argue that technologies do not only amplify mathematical or statistical thinking but change it [67]. Technologies used in school-level data science curricula include mobile phones, data cards, spreadsheets, menu-driven user-friendly software environments and programming in Python or R. All agree that engaging with technology is not only a tool but a purpose of teaching data science.

There is less consensus around programming. In the Los Angeles based Introduction to Data Science (IDS) project, which sees computational aspects as an essential component of data science, programming (e.g. via RStudio) is used to merge computational and statistical thinking [31] . Programming is a way of executing and communicating reproducible strategies, and code serves as a notation system and as a modelling tool. Other experts believe there is a middle-ground of building on unplugged approaches (at younger ages) and then using user-friendly interfaces such as Jupyter Notebooks that support tool-use without programming [68]. Others, such as the New Zealand curriculum, consider that it is more important for students to use easily accessible visualisation technologies as a basis for inferential reasoning.

Given the diversity of viewpoints, and resource in schools, we have decided to frame the intersectional competence *Using technology* at a high-level, reflecting the consensus that students should have opportunities to engage in guided enquiry with access to the internet

---

[67] Gould, 'Toward data-scientific thinking'; E. Geraniou and M. Misfeldt, 'The Mathematical Competencies Framework and Digital Technologies', in U. T. Jankvist, E. Geraniou (eds.), Mathematical Competencies in the Digital Era, (Cham: Springer International Publishing, 2022), pp. 39–60.

[68] Rolf Biehler and Yannik Fleischer, 'Introducing Students to Machine Learning with Decision Trees Using CODAP and Jupyter Notebooks', *Teaching Statistics* 43, no. S1 (July 2021), https://doi.org/10.1111/test.12279.

and computer-based representational tools. This competency informs all stages of posing and answering problems.

5.2   Unpacking the mathematics toolkit

In this section we do not attempt to identify or characterise the whole of mathematics; there are plenty of curricula and frameworks that provide ways of doing this. Instead we have compared the three reference frameworks with the intersectional competencies outlined above to identify their connections. What follows are six high-level areas of distinctively mathematical activity that are most clearly involved when solving data-driven problems with technology in the (late) secondary curriculum.

Solving data-driven problems within mathematics involves:

Using quantities and methods: estimating and measuring, choosing appropriate number operations (e.g. division) and representations (e.g. rates, percentages), understanding the effect of choices and errors.

Analysing covariation: finding and describing patterns and relationships between variables. Proportional reasoning. Choose and apply appropriate graphical, spatial and symbolic methods to determine/investigate values taken by one variable given knowledge about related variable(s), including if appropriate the rate of change.

Reasoning mathematically: producing and/or critiquing extended arguments that use mathematical methods and inferences to justify and/or prove conclusions. Use logical comparators within methods, arguments and discussion.

Using representations: choosing and, importantly, comparing, translating and combining different mathematical representations (e.g. 2-way tables, schematics, graphs, symbolic forms) generated with technology where possible.

Using aids and tools: choosing and using appropriate manipulatives, visual aids and technologies in mathematical work, understanding their affordances, limitations and connections.

These were checked for compatibility with the ten General Mathematical Competencies (GMCs) identified for vocational T-levels[69]. The first two, *Using quantities and methods* and *Analysing covariation,* are derived most directly from the focus in the Ontario curriculum on understanding how numerical data and mathematical models are combined to explain variation in the world. The special nature of *Reasoning mathematically* is highlighted in both PISA 2022 and the KoM framework (and is not explicit in T-level GMCs). The final two competencies describe aims for how students will use and think with representations and tools which are inherently mathematical. One surprise in this listing is that we have not explicitly included spatial thinking. That is solely because of the focus on school-leavers and data-driven approaches. Spatial thinking is important for its own sake. It may also be a context from which to derive measurements and data. More significantly, the awareness given by spatial thinking (what the Chinese curriculum calls 'intuitive imagination' [70]) pervades these competencies for school leavers. It underpins notions that are central to computational thinking such as finding structure, decomposition. It also aids understanding of visual representations such as 2-way tables and graphs.

---

[69] Measuring with precision; Estimating, calculating and error spotting; Working with proportion, Using rules and formulae; Processing data; Understanding data and risk; Interpreting and representing with mathematical diagrams; Communicating using mathematics; Costing a project; Optimising work processes.

[70] Yanjun Hong and Uffe Thomas Jankvist, 'Comparing Two Mathematical Competencies Frameworks from China and Denmark', *Asian Journal for Mathematics Education* 1, no. 4 (2022): 494–514, https://doi.org/10.1177/27527263221144193.

5.3 Unpacking the data skills toolkit

Solving problems within statistics/data science involves:

Data stewardship: attend to privacy, security, and ethical aspects such as findability, accessibility, interoperability, and reusability and how these interact. While ethical considerations are fundamental to posing and answering questions within the domain, these are the ethical aspects that relate specifically to holding and processing data.

Handling data: gathering, organising, cleaning, displaying data, purposefully and with respect to its use within a legal and ethical framework.

Data representation: choosing and, importantly, comparing, connecting and combining different data representations, generated with technology where possible.

Statistical thinking and methods: anticipating, recognising and accounting for variety and variability (including error or uncertainty) in data and how they shape our analyses and predictions; engaging in exploratory data analysis: finding, describing and analysing patterns, relationships, trends and variation using data visualisations and shape representations; using appropriate measures of central tendency, spread and variation; choosing and applying statistical models to support making inferences and predictions, and testing hypotheses.

Probabilistic reasoning: as appropriate, choosing and applying a probabilistic model (and often proportional reasoning), using randomisation and simulations.

Using aids and tools: as relevant, using a range of computational tools, including calculators, spreadsheets, mathematical and statistical software, data analysis packages, visualisation tools, machine learning algorithms and programming.

5.4 Unpacking the computing toolkit

Solving data-driven problems with and within computing involves:

Computational thinking (including developing algorithms): for a given problem, carrying out abstraction, decomposing it into simpler sub-problems, recognising and generalising patterns, and developing relevant algorithms (without using a programming language).

Using algorithms: being able to follow and implement a range of common algorithms, for example for searching and sorting data, and recognising how they could apply to a given problem.

Programming: creating programs to solve problems using text-based or visual/block-based programming environments.

Representing and manipulating data: choosing relevant representations and using these to harness, transform and create data.

Safe use of technology: have a good understanding of how to use technology safely, an awareness of threats and the protection options that are available.

Using aids and tools: using a range of hardware, software, sensors and networks, design and computational tools, including calculators, spreadsheets, databases, component-based software packages, and new and emerging technologies.

6.   Citizen categories and school leavers

The Royal Society proposed an outline of five citizen categories, and invited reconsideration of these in relation to specifying the needs of school leavers. Curricula with differentiated pathways usually describe these pathways in terms of immediate goals for work or further study. Original categories 1 and 2 describe mathematical needs necessary for citizenship, without including any further workplace requirements. Categories 4 and 5 are graduate STEM professions while category 3 covered a wide range in between.

Original RS categories	1 Basic functional numeracy for personal decision making.	2 Mathematical literacy to understand and critique numerical claims	3 Traditionally non-quantitative jobs/professions, but requiring mathematical expertise: e.g., lawyers, journalists, civil servants, politicians, health practitioners, administrators, teaching practitioners (not mathematics), technical and vocational professions – including technicians.	4 Jobs/professions where mathematical competences are a core component, e.g. analysts, engineers, financial professionals, scientists, social scientists.	5 Mathematical sciences, e.g. mathematicians, mathematics teachers.

Table 1 Original Royal Society citizen categories

The original framing of category 1 appears to arise from a combination of three aspects: mathematics in this category is applied to everyday contexts where *individuals* typically make decisions; it serves to *construct* a solution and the data involved appears *well-determined*. Activities in category 2 involve interactions with, and for, *other people*, perhaps in the workplace or society; they are framed as *responding* to claims that may involve *uncertainty*.

The constructive/responsive aspect to this distinction is one that is valued in mathematics and data sciences competency frameworks, for example the KOM framework describes constructive/receptive facets of each competency. This duality seems worth keeping. However, there are reasons to question the other two aspects. Few personal decisions of any significance are based on neat and certain data. They are instead likely to involve finding and interpreting data sources that need to be understood and critiqued. In addition, a 2018 UK review of adults' digital skills concludes that, in a society with widespread digital access to information, there are no clear distinctions between basic skills and workplace

needs and that basic digital and information skills act as a gateway to employment[71]. Our suggestion is that categories 1 and 2 be combined as the mathematical and data literacy that contributes to active citizenship and employment **for all** school leavers.

In contrast, category 3 seems wide. What Kispeter calls 'higher-level digital skills', and others have called 'techno-mathematical literacies' (TMLs), are typically learnt within the specialised setting of the workplace, incorporating the vital aspect of domain knowledge. TMLs do not necessarily require new mathematical knowledge but, instead, knowledge of the relevant data and models accessible through technology that support working practices. Examples might be teachers' use of pupil attainment packages to predict performance, or electricians' entering variables on online platforms that guide the process of issuing safety certificates. Kispeter's review concludes that such workplaces require 'T-shaped' knowledge - broad across many areas with deep knowledge in one area. It seems appropriate then to create competencies **for many** students, who should use data skills in a range of contexts with some involving multiple variables and responsibility to others.

In such settings a distinction has been made between practitioners and advanced practitioners. Practitioners are those who make decisions using models and quantitative data that are provided for them (including when those go wrong [72]), whereas advanced practitioners have extra responsibility (and skills) for choosing and critiquing those models and choices of data. This distinction appears not only in the digital skills literature, but also in employment-based data skill literature and in the statistics literature about knowledge of policy makers. However, these all argue that there is transition between the two levels – practitioners can develop to use advanced knowledge; experts will need to be embedded in practice. [73] For this reason, we have not indicated extra intersectional competencies for this category beyond those specified 'for many'. We also note findings from ACME's

---

[71] Erika Kispeter, 'What Digital Skills Do Adults Need to Succeed in the Workplace Now and in the next 10 Years?', Report (London: Department of Culture, Media and Sport, June 2018), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/807831/What_digital_skills_do_adults_need_to_succeed_in_the_workplace_now_and_in_the_next_10_years_.pdf.

[72] Nathalie J. van der Wal, Arthur Bakker, and Paul Drijvers, 'Which Techno-Mathematical Literacies Are Essential for Future Engineers?', *International Journal of Science and Mathematics Education* 15, no. 1 (1 May 2017): 87–104, https://doi.org/10.1007/s10763-017-9810-x; Celia Hoyles, ed., *Improving Mathematics at Work: The Need for Techno-Mathematical Literacies*, 1st ed (New York, NY: Routledge, 2010).

[73] https://ctale.org/2020/10/06/what-skills-and-tools-do-policy-makers-use-to-analyse-data/; https://wol.iza.org/articles/big-data-in-economics/long

Mathematical Needs report [74] that interviewees felt the need to have studied mathematics to a level higher than that habitually used at work. The 16-19 curriculum could include suitable courses that develop these same competencies over a wider range of contexts and/or with wider mathematical or statistical toolkits.

Finally, we blend categories 4 and 5 that need competencies **for specialists**. Category 5 differed in nature to the others since it requires in-depth knowledge about mathematics as a discipline with valued ways of reasoning (such as deductive proof) and sociocultural practices. From a Bernsteinian perspective, mathematicians, sociologists of mathematics, curriculum writers and mathematics teachers are the archetypal insiders and maintainers of boundaries in this practice. For graduates the 4/5 distinction matters as mathematical research contributes strongly to the UK's economic and academic progress. [75] However, at school-leaver level, in terms of mathematical and data literacy, the needs of future mathematicians are similar to the needs of future physicists/engineers etc in RS category 4.

The competencies that we have developed are thus proposed at three levels: for all, for many and for specialists. They correspond to categories as shown in Table 1.

Mathematical and data literacy for all:		Data skills for many (plus wider contexts or tools)		For specialists (and mathematics specialists)	
Mathematical and data literacy to understand and respond critically to numerical information and claims and to use these for decision-making e.g. to buy insurance, run a hair salon.		Workplace roles that involve decision-making based on quantitative data and models that are now accessible through technology e.g. classroom teacher, health practitioners, electricians, some journalists, civil servants, politicians, artists	Workplace roles that involve responsibility for choosing or critiquing quantitative data and models: e.g., lawyers, journalists, civil servants, politicians, lead teachers.	Workplace roles where data science competencies (developed in parallel with e.g. computing, engineering, mathematics, science) allow the creation of analytic models e.g. analysts, engineers, financial professionals, scientists, social scientists.	Those who need to understand mathematics as a discipline, e.g. mathematicians, mathematics teachers and educators.
RS 1	RS 2	RS 3		RS 4	RS 5

Table 2 Reworked categories for school leavers: competencies for all, many and specialists

---

[74] ACME Mathematical Needs report (2011)

[75] Philip Bond, 'The Era Of Mathematics: An Independent Review of Knowledge Exchange in the Mathematical Sciences.' (UKRI, 2018), https://www.ukri.org/publications/the-era-of-mathematics/.

In terms of the current English curriculum structure, the 'For all' level would correspond to what must be taught in KS 3 and 4, thus Foundation level mathematics GCSE and KS 4 computing (although our competencies start from the different goal of promoting mathematical and data literacy). At the 'For specialists' level, current students choose amongst A levels in Mathematics, Further Mathematics, Computer Science (A level Statistics is negligeable, and there is no Data Science A level). There is no single set of qualifications corresponding to the 'For many' level; currently some of the cohort study Higher Level GCSE, some study Core Maths, some study mathematics, statistics and computing as part of T levels and vocational qualifications.

6.1 Intersectional competencies for school leavers

The following statements give more detailed competency statements for all, most and specialist school leavers in the six intersectional competencies. They were synthesised from the vision papers, report trials and curricula identified as reference points in 4.1, combining their reported goals with our knowledge of young people's educational progress at 18.

	All students can:	Most students can:	Specialists can:
Conducting iterative and holistic enquiry	Create and conduct investigations using the full mathematical/statistical/computational enquiry cycle in domains related to health, personal finances, civics, including using data that has been gathered from familiar digital interactions (e.g. Youtube views), attending to the context and variability in the data.	ALSO use the full enquiry cycle in situations requiring detailed knowledge from another school subject.	ALSO use the enquiry cycle with attention to sophisticated tools and formal reasoning appropriate to mathematical modelling/ statistical reasoning/ computing.
Posing questions	Pose and relate to questions framed in domain knowledge; identify ethical implications of different ways of formulating questions; reformulate a question to allow use of a given mathematical/statistical/computational construct or tool.	ALSO formulate, relate to and connect a sequence of questions: the mathematical and domain questions used to frame an investigation, those used to collect data, and those used to guide analysis and interpretation.	ALSO formulate questions using specialist language, tools and reasoning of mathematics/ statistics/ computing. For example, in statistics, posing inferential investigative questions regarding causality and prediction for surveys, observational studies and experiments.
Solving problems (with mathematics, computing and data toolkits)	Explore and use patterns of association between variables, attending to quantitative, algebraic, visual and spatial features; plan approaches using appropriate models, tools and representations; propose solutions and make predictions.	ALSO formally model relationships among multiple variables and account for variation from proposed models or between models; adapt approaches to take account of data complexity and computational efficiency.	ALSO create analytic methods to investigate associations; estimate and quantify errors and develop confidence/reliability measures for solutions and predictions.

The Open University

Interpreting and communicating	Use analytic outcomes as evidence to answer questions; identify implications in the originating domain (e.g. to inform decision making). Interpret information in the media involving the use of data. Use language, digital artefacts and visualisations to communicate answers and how the answers were arrived at, including a statement of uncertainty and plausibility when needed.	ALSO interpret comparisons or combinations of analytic outcomes (e.g. interpreting using measures of central tendency _and_ spread, or value _and_ direction of change) and provide comprehensive answers; draw inferences; compare interpretations; make sense of others' code or programs and annotate and present programs so that others can follow; create and reflect on different ways of communicating answers and conclusions for relevant audiences.	ALSO use evidence from analyses to interpret, classify and communicate errors, bias or uncertainty and discuss their implications; produce well-documented programs that are accessible and re-usable; communicate results and formal mathematical/statistical reasoning through reports and presentations using multiple applications and a variety of formats.
Evaluating and critiquing	Compare effects of different conditions; reflect on and make judgements about the appropriateness of their own and others' conclusions; consider possible sources of confusion, error and bias (e.g. choice of training data) and take these into account when evaluating their own and others' reasoning and communication; appreciate there are multiple stakeholders and viewpoints; discuss implications for ethics and for confidence in the results.	ALSO compare the effect of systematically varying methods and conditions (including of implementing different common algorithms) within the given domain; comment on the validity of generalisations based on inputs and choices; evaluate the impact of outliers, missing values, and the errors in modelled solutions and predictions; compare the outcomes of machine learning algorithms for a classification or prediction problem.	ALSO distinguish between different levels of rigour in evidence and arguments; evaluate the efficiency and applicability of algorithms and programs; justify or prove conjectures and conclusions using formal mathematical/statistical/computational reasoning and methods (e.g. for training and testing machine learning algorithms).
Use technology	Use data derived from online platforms. Experience working with data collected from personal devices. Use technologies to implement routines and create text, graphs, charts and visualisations.	ALSO: use varied sources of data (e.g. sensors, large datasets); use current technologies (e.g. spreadsheet, graph plotter, Orange) to implement and adapt existing routines.	ALSO: use specialised languages and current technologies to create representations and algorithms that support computation, modelling, prediction, data analysis and display.

Table 3 Intersectional competencies for school leavers: for all, many and specialists

6.2 Mathematical toolkit competencies for school leavers

These competency statements expand on the mathematical toolkit needed for solving data-driven problems. They are synthesised from goals set in the vision papers, report trials and mathematics curricula identified as reference points in 4.1 and our knowledge of mathematics pedagogy. They are not comprehensive in that they do not outline all the mathematics needed by school leavers. Instead they outline a subset that would be needed for a coherent course at the intersection of MSDC.

	All students can:	Most students can:	Mathematics specialists can:
Using quantities and methods	Specify and relate to familiar quantities (time, money) with a contextual meaning, appreciate role of units, choose appropriate number operations (e.g. division) and representations (e.g. rates, percentages), use number sense and fluent estimation to assess reasonableness of claims; compare effects of choices and errors.	ALSO construct formulae for multi-stage operations; use exponential expressions (e.g. for compound interest); use representations of number to show appropriate accuracy. Distinguish between absolute and relative change (e.g. percentage vs percentage points). Consider the effect of systematically changing a parameter.	ALSO: use and manipulate symbolic expressions for multi-stage operations with fluency.
Analysing covariation	Find patterns in data or in familiar contexts, and thus Identify and describe proportional, linear and non-linear relationships. Read graphs and tables to determine values taken by one variable given knowledge about related variable(s). Interpret gradients of linear functions qualitatively (0, +, -) as indicating rate of change.	ALSO Identify and describe quadratic and exponential relationships in familiar contexts. Use a graph plotter to construct graphs from equations or data. Choose and apply appropriate graphical, spatial and symbolic methods to investigate values taken by one variable given knowledge about related variable(s).	ALSO Identify and describe periodic relationships. Use methods of differential calculus. Investigate values taken by one variable given knowledge about related variable(s) and the rate of change. Interpret area under a speed-time graph.
Reasoning mathematically	Explain how to arrive at an answer, working forwards or working backwards. Identify and use logical language involving comparators, conditions and negations e.g. 'If, Then, Else', AND, OR, NOT. Determine whether statements are always, sometimes or never true.	ALSO; use general properties in informal deductive arguments; use rules purposefully to create algebraic equivalences, or to construct and adapt algorithms and sequences of instructions.	ALSO: Produce and/or critique extended arguments that use mathematical methods and inferences to justify and/or prove conclusions.
Using representations	Use and relate to tables, 2-way tables, graphs and other schematics using length to represent quantity, generated with technology where possible.	ALSO: choose, compare, combine and translate between mathematical representations, including algebraic expressions and schematics using area, volume or colour to represent quantity.	ALSO: use function notation. Compare and combine representations of related variables (e.g. use the graph of a function to sketch the graph of its derivative).
Using aids and tools	Use hand-held and online calculators; use tools to organise data and thinking (e.g. writing frames, grids, algebra tiles, spreadsheets).	ALSO: use visual aids and technology that support graph plotting, scale drawing, sorting, searching and repeated calculations.	ALSO: Understand the affordances, limitations and connections of manipulatives, visual aids and technologies.

Table 4 Mathematical toolkit competencies for school leavers: for all, many and specialists

6.3 Statistics/data science toolkit competencies for school leavers

These competency statements expand on the statistics and data science toolkit needed for solving data-driven problems. They were synthesised from the curricular levels in the GAISE II statistics and data science standards and in the New Zealand statistics curriculum, with some input from other reference data science sources mentioned in 4.1.

	All students can:	Some students can:	Specialists can:
Data stewardship	recognise the purposes for which data are collected and processed, and attend to privacy, security, and ethical concerns in a responsible manner.	ALSO: ensure data that may contain sensitive information are handled appropriately; appreciate how to manage data so that it is accessible, usable, re-usable and well-documented.	ALSO: enact practices that enhance reproducibility and ensure ethical use of data; appreciate the ethical consequences of their experiments and analyses; design or adapt data processing systems to address the privacy and security of data.
Handling data	gather, organise, clean, store, display and interrogate data purposefully (primary and secondary data); distinguish between types of variables (e.g., categorical or quantitative) and understand what characteristic they represent; engage with data arising from familiar sources (e.g. mobile phones) and from multiple sources.	ALSO: use appropriate data handling methods depending on the type of data or setting (e.g., a survey or experiment); devise appropriate data collection plans; have some experience of working with large data sets and know how to address missing data.	ALSO: select data analysis techniques appropriate for the type of data they have; connect data from different sources, including for large data sets.
Data representation	choose and, importantly, compare, connect and combine different data representations, generated with appropriate technology; use appropriate representations, (e.g. proportions for categorical data, two-way tables, tree diagrams and confusion matrices); represent the variability in data using appropriate visualisations.	ALSO: work with multi-dimensional interactive visualisations and graphics.	ALSO: design and code visualisations for large data sets (e.g. hexagonal bin plots).
Statistical thinking and methods	anticipate, recognise and account for variety and variability in data (including via error or chance) and appreciate how they shape analyses and predictions; engage in exploratory data analysis: find, describe and analyse patterns, relationships and trends; look for association between two variables; use appropriate measures of central tendency, spread and variation; describe key features of distributions such as mean, median, range, symmetry; use scatter plots to explore association between variables; compare conditional proportions across categorical variables; understand that sample data can be used to answer questions about a population.	ALSO: identify appropriate ways to summarise quantitative or categorical data using tables, displays, and summary statistics, including using standard deviation as a measure of variability; summarise and describe relationships among multiple variables; recognise the uncertainty caused by sample to sample variability; use measures of correlation; choose and apply appropriate statistical models and methods to support making inferences and predictions (e.g. least-squares regression line); compare results for different conditions in an experiment.	ALSO: summarise their data using graphical displays and numerical summaries; understand the constraints on generalising and on predictions based on attributes of the data; understand the issues of bias and confounding variables in observational studies; use multivariate thinking to understand how variables impact one another; understand the power and limitations of conclusions based on samples and experiments; be able to quantify the related uncertainty using margins of error; describe associations between two categorical variables using measures such as difference in proportions and relative risk; implement bootstrapping where relevant.

Probabilistic reasoning	use natural and expected frequencies to estimate and communicate probabilities and risks; interpret contextual statements to determine appropriate denominators for probabilistic scenarios; choose and apply a probabilistic model (and often proportional reasoning) where relevant, in particular for risk and statistical literacy problems.	ALSO: use two-way tables, tree diagrams and Venn diagrams for calculating probabilities including conditional probabilities; use simulations to investigate associations between two variables and to compare groups; understand the role of random selection and random assignment in experiments and its implications for generalisability.	ALSO: understand the role of randomness and the effect of sample size on the variability of estimates; understand how and why bootstrapping works; use simulations to determine approximate sampling distributions; calculate and interpret expected values and standard deviations of discrete random variables; use distributions such as the Poisson, binomial, and normal.
Using aids and tools	use a range of relevant computational tools, including calculators, spreadsheets, mathematical and statistical software, data analysis packages, visualisation tools, and machine learning algorithms.	ALSO: use appropriate technology to implement machine learning prediction or classification models.	ALSO: use technology to decompose and filter data sets and transform variables, including smoothing for time series data; use technology and/or program to compare and evaluate machine learning algorithms and consider their impact.

Table 5 Statistics / Data science toolkit competencies for school leavers: for all, many and specialists

6.4 Computational toolkit competencies for school leavers

These competency statements expand on the computational toolkit needed for solving data-driven problems. The competency statements below were synthesised from the national curricula of England and Wales (subject content documents and related guidance), individual examination board specifications as well as input from the Singapore and US K-12 curricula, which can be mapped to the competencies below. Only competencies which have some level of intersection with mathematics and data science have been included.

	All students can:	**Some students can**:	**Specialists can**:
Computational thinking	look at a problem and carry out basic abstraction – removing unneeded detail; decompose a problem by breaking it down into sub-problems that are easier to solve; write a set of instructions (an algorithm) to solve the particular problem they are faced with.	ALSO: compare a problem to others faced previously; decompose a problem by breaking it down into sub-problems using formal techniques such as top down design and with the concept of a subroutine in mind so that the sub-problems are easier to solve; write an algorithm to solve a particular problem in general terms; identify some areas of repetition or patterns within the problem and identify the potential to reuse parts of the algorithm.	ALSO: look at a complex problem and carry out a range of abstraction techniques appropriate to the problem at hand; decompose a problem using formal techniques, with the concepts of subroutines and modules in mind and with a view to re-composition; when writing algorithms consider developing an efficient solution in terms of run-time and data storage efficiency; identify patterns within the problem and make efficient and effective use of repetition, subroutines and pre-written solutions.
Using algorithms	follow a range of common algorithms for searching and sorting data and see their applicability to real-life situations; apply comparators when searching and sorting data.	ALSO: follow, discuss and implement a range of common algorithms for searching and sorting data and choose between them based on their efficiency and applicability to a given problem.	ALSO: follow, discuss and implement a range of algorithms for traversing graphs and domain specific problems and identify their use in modelling real-life situations.

Programming	create simple programs to solve carefully curated problems; have experience of at least two programming languages, one of which should be text based; use input, output, manipulate variables and use simple iteration and selection.	ALSO: create programs to solve well-articulated problems within a given domain; use at least three languages, including a text-based imperative language and a text-based declarative language such as SQL; use 1-D and 2-D lists or arrays, nested selection and a range of iterative techniques and have had some exposure to subroutines.	ALSO: create programs to solve a range of novel problems within a new domain; use an imperative language that offers exposure to object-orientated programming techniques and have also had some exposure to other programming paradigms such as functional programming (such as Haskell) or logical programming (such as Prolog); use pre-developed code and subroutines using varying parameters and return methodologies, and are comfortable with modular programming and the use of classes, objects, records and similar structures.
Representing and manipulating data	convert between different numerical bases and have an understanding that all data can be represented in binary regardless of what it is; choose suitable variables for a given situation or problem and manipulate individual variables to achieve a desired effect; work with data generated by themselves and their peers.	ALSO: perform basic binary manipulations (add, shift); manipulate individual and list variables to achieve a desired effect; perform simple tasks upon data held within a relational database; work with data from multiple sources and a range of data-types.	ALSO: perform binary manipulations (add, shift, subtract, multiply, mask); manipulate multi-dimensional list variables to achieve a desired effect; perform relatively complex tasks upon data held within a relational database; work with more specialised and unfamiliar data-types, including live and automated data collection, e.g. remote or embedded sensors and log data.
Safe use of technology	use standard technologies safely, for example choosing safe passwords, use of anti-virus software, and an awareness of relevant legal structures.	ALSO: recognise different types of malware and other threats; know the protection options that are available and have an awareness of the legal, social, ethical and moral guidance under which they should act.	ALSO: act based on a thorough understanding of different types of malware and other threats and the protection options that are available; discuss and debate the legal, social, ethical and moral guidance under which they should act as well as the relevant professional guidelines.
Using aids and tools	use a range of hardware, software, and networks, design and computational tools, including calculators, spreadsheets, software packages, and programming; engage with novel and emerging technologies such as robotics, Internet of Things, Artificial Intelligence and Virtual Reality as users.	ALSO: create and combine artefacts from two or more applications; use databases and component-based software packages; work on projects that may include some aspects of physical computing, data collection or logging with novel and emerging technologies.	ALSO: work with novel and emerging technologies, including the opportunity to develop and create robotics, Internet of Things, Artificial Intelligence, Virtual Reality, or other emerging technology artefacts.

Table 6 Computational toolkit competencies for school leavers: for all, many and specialists

Part Two

The second part of this report is a feasibility study on a set of curriculum models that could be the basis of reform in the English national system of teaching and assessment, delivering on the educational aims proposed in Part One. In short, these are that all students have opportunities to develop- and are assessed on - mathematical and data literacy; some also have opportunities – and are assessed on – using mathematical and data skills in context; finally, others are prepared to become STEM and data science specialists. We have used the literature form Part One and the curriculum models proposed to identify case studies that are relevant to achieving these aims. Section 7 outlines the curriculum models; Section 8 contains the 10 case studies; Section 9 provides an overview of these in the light of feasibility criteria and may be read without Section 8.

7.   Curriculum models

The set of curriculum models considered is informed by the indicative list provided by the Royal Society (Appendix B). Table 7 shows it fleshed out and organised by where change is envisaged. After the 'No change' model 1, the top of the table (models 2- 4c) situates change primarily in the Key Stage (KS) 4 curriculum, affecting GCSEs or equivalent qualifications, but with the current system of specialised A levels. This would have implications for KS 1, KS 2 and KS 3. The lower part of the table (models 5-8) envisages major change all the way to KS 5. This could have implications for the whole future landscape of Level 2 and 3 qualifications [76]. Amongst these, the new T levels (vocational assessments) already have some integration of mathematics within domain knowledge; BTECs (alternative academic assessments) are more varied. The academic pathway of GCSE resit, Core Maths and A levels could be most affected structurally.

---

[76] Department for Education, 'Guide to the Post-16 Qualifications Landscape at Level 3 and below for 2025 and Beyond' (Department for Education, 2023).

		Curriculum models	RS #	Most relevant examples
Change within KS 3 & 4 but not KS 5	**1**	No change. Current structure in England	1	
	2	Current Structure for KS 3 and KS 4 with content added to Computing	2b	ProDaBi project [77] (KS 3 onwards, Germany)
	3	Current Structure for KS 3 and KS 4 with content added across subjects	3 (GCSE)	Urban Data School project (KS 2/3, England)[78]
	4a	Current Structure for KS 3 and KS 4 with content added to a Mathematics course	2a	
	4b	Additional content at KS 3 and 4 creates two parallel or sequential Mathematics GCSE courses	4 (GCSE)	New Zealand curriculum has parallel mathematics and statistics pillars. MEI proposal has Essential Maths GCSE followed by Maths [79]
Change in KS 5 drives change in KS 3 & 4	**5**	Current KS 5 structure with the addition of credit-bearing project-based work e.g. EPQ.	5	MEI data science courses, IB project. Scotland NPAs
	6	Current KS 5 Structure with content added across subjects	3 (post-16)	Nuffield and Score projects adding quantitative content to A levels [80]. Geography in Core maths.
	7	Current KS 5 Structure with additional content creating choice of mathematics pathways (examined as A/AS level or Core Maths)	4 (post-16)	IB has choice of course length and analysis/applications. Ontario is pathway-based..
	8	A new baccalaureate system at KS 5 and below (requiring study in core subjects)	6	IB (highly-constrained); NZ and Ontario curricula (less constrained).

Table 7 Curriculum models

[77] https://www.prodabi.de/en/

[78] Annika Wolff, Michel Wermelinger, and Marian Petre, 'Exploring Design Principles for Data Literacy Activities to Support Children's Inquiries from Complex Data', *International Journal of Human-Computer Studies* 129 (September 2019): 41–54.

[79] MEI, 'Curriculum and Assessment in Mathematics in England at KS4 and KS5 - Position Paper' (Mathematics in Education and Industry, 2022).

[80] SCORE, 'Mathematics within A Level Science 2010 Examinations' (SCORE (Science Community Representing Education), 2012),

8. Case Studies

The case studies start with larger, well-established national curricula and end with smaller-scale innovative projects trialled in England. An overview and discussion are presented in section 9. The cases included are:

8.1   Ontario curriculum pathways

Ontario offers an example of a curriculum with multiple pathways for older students (our Model 7), and one where real-world applications are notably pervasive [81]. The 2005 mathematics curriculum for elementary grades 1-8 is undergoing a rolling transition from 2020 [82]. Ontario also "de-streamed" grade 9 in 2021, following other states in a drive to reduce entrenched inequities worsened by early pathway choices. The grade 10-12 curriculum remains in effect [83].

Curriculum change

The 2005 Ontario curriculum (influenced by US Standards) heavily emphasises mathematical processes (problem solving; reasoning and proving; reflecting; connecting; communicating; representing; selecting tools and strategies) in the familiar content areas of number, algebra, measurement, geometry and statistics. It also stresses use of

---

[81] Smith and Morgan, 'Curricular Orientations to Real-World Contexts in Mathematics'.

[82] Ontario Ministry of Education, *The Ontario Curriculum, Grades 1-8: Mathematics (2020)*, 1–8.

[83] 'Course Descriptions and Prerequisites', accessed 23 March 2023, https://www.dcp.edu.gov.on.ca/en/course-descriptions-and-prerequisites/mathematics.

mathematics: "Mathematical knowledge becomes meaningful and powerful in application. This curriculum embeds the learning of mathematics in the solving of problems based on real-life situations." [84]

Recent changes to the grade 1-9 curriculum have retained the seven processes and much of the content [85]. New strands have been added for financial literacy, responding to a concern about basic skills [86], and for social-emotional learning (SEL) skills while using mathematical processes.   Measures and geometry have been combined in to one strand of Spatial Sense, with attention to how this underpins mathematical reasoning in other areas [87]. Coding and mathematical modelling have been added into the algebra strand. While applications of mathematics were already present, the new curriculum means that students must be taught to model "messy problems". [88] This identification of strand content matters as they define areas that must be assessed.

All assessment is by teachers: it has a common structure in all subjects but can be tailored to local circumstances. [89] Teachers report on student learning at 4 levels in in four main areas: knowledge and understanding, thinking, communication, and application. For secondary students, 70% of the assessment occurs during the course and 30% at the end. Provincial assessments at grade 3, 6 and 9 inform the Ministry about performance trends.

Implementing the curriculum

Curriculum development in Ontario is seen as having decades of change and consolidation, with the recent change responding to the availability of technology. [90] Teacher professional development has the highest priority in good mathematics teaching and embedding change. Ontario's consistent approach to improving mathematics performance is by increasing subject-specific CPD, funding release time for teachers to

---

[84] Ontario Ministry of Education, *The Ontario Curriculum, Grades 11 and 12: Mathematics (2007)*, 4.

[85] Christine Suurtamm, 'Elementary Mathematics Curriculum, Teaching, And Learning: The Past Ten Years', *Ontario Mathematics Gazette* 60, no. 3 (2022): 19–24.

[86] Egan J. Chernoff, 'The Canadian Math Wars', *Canadian Journal of Science, Mathematics and Technology Education* 19, no. 1 (1 March 2019): 73–76, https://doi.org/10.1007/s42330-018-0037-9.

[87] *Research to Inform Elementary Mathematics Curriculum Revision*, 2020, https://www.youtube.com/watch?v=AmZtIpAHZs4.

[88] Suurtamm, 'Elementary Mathematics Curriculum, Teaching, And Learning: The Past Ten Years'.

[89] Jill Lazarus, 'Secondary Curriculum Over the Last Decade', *Gazette - Ontario Association for Mathematics* 60, no. 3 (2022): 52–54.

[90] Lazarus.

work collaboratively, in particular to low-performing schools. [91] The Ministry works within an education ecosystem, partnering with practitioners and researchers to develop professional learning and classroom-ready resources before implementing a new curriculum. [92] Assessment resources are an important part of this change.

Overlapping Pathways from 15

As in other Canadian states, mathematics is effectively compulsory until age 17 (grade 11). Students need at least 3 mathematics credits to graduate high school, with one from grade 11 or 12.

Ten mathematics courses are specified for grades 10 – 12, in three overlapping pathways, most with online teaching versions to allow individual study. Aims are described in terms of preparing for school leavers' next destinations: the workplace (referred to as W below), college or university. Within college preparation, a distinction is made between technology-related programs (CT) and others (C). Within university preparation, the distinction is between degree majors needing mainly statistical data management such as social sciences (UDM); those needing knowledge of functions such as business, health science (UF) and those also needing calculus such as STEM or economics (US).

Each course is described in terms of strands, corresponding broadly to topic areas, plus overall and specific expectations that outline activities and provide examples. While there is overlap between course strands, the only overarching framework is the processes. We have therefore synthesised the course expectations to highlight five competencies that show how Ontario brings data literacy and real-world applications into mathematics for the range of school leavers.

Making financial decisions: <u>all</u> students solve problems using compound and simple interest, more or less formally linked to sequences. Substantial parts of the W and C courses (only) involve sourcing and dealing with data representations (e.g. payroll/tax statements, airline websites) to make budgeting decisions as citizens and consumers.

Investigating and handling data; by 15 all students have solved problems involving 1- and 2-variable data gathered collectively from experiments (e.g. with sensors) or online sources. W courses focus on constructing and critiquing arguments with categorical data.

---

[91] 'Ontario Dedicating $60 Million for Renewed Math Strategy', news.ontario.ca, accessed 23 March 2023, https://news.ontario.ca/en/release/36364/ontario-dedicating-60-million-for-renewed-math-strategy; Haizhen Mou and Michael M. Atkinson, 'Want to Improve Math Scores? An Empirical Assessment of Recent Policy Interventions in Canada', *Canadian Public Policy / Analyse de Politiques* 46, no. 1 (2020): 107–24.

[92] Suurtamm, 'Elementary Mathematics Curriculum, Teaching, And Learning: The Past Ten Years'.

C and CT courses extend this to dealing with 2-variable data in context; plus measures of correlation, with equations of lines of best fit for CT; relative measures e.g. indices for C courses. University courses vary considerably. UF and UC courses fit functions to data without theorizing related statistical ideas. The UDM course includes making informed decisions about sampling, questioning data validity, drawing conclusions from presented data, all brought together in a substantive, assessed culminating investigation.

Dealing with uncertainty: all students have previously calculated and represented theoretical probabilities. W courses revisit representations that guide personal decisions (e.g. health risks); C course includes interpreting media use of statistics, questioning data provenance and exploring the effects of changing assumptions. UDM course is the only one to include combinatorics and probability distributions for discrete and continuous variables. CT, UF and US courses only implicitly deal with uncertainty, by replacing data with models.

Modelling with quantitative data: all students have extensive experience of solving problems involving linear relationships or graphs of non-linear relationships. W courses focus on situations that involve proportional and non-proportional reasoning including the effect of errors. Other students model data with functions, taking graphical, algebraic and computational approaches to solving contextual and mathematical problems, supported by technology. UDM stops this in year 12, while C, CT, UF and US courses increase the range of functions. UF includes an understanding of rate of change, and US applies this to optimization problems.

Solving spatial problems: all students have applied measurement concepts and skills to solve spatial problems. UDM students do no further work. W courses focus on area and perimeter in context and consider errors. C courses include optimisation problems, sine and cosine rule. CT courses add solving problems using trigonometry in 3-D and circle properties. UF introduces radian measure. US course introduces and applies vector operations and applications (e.g. in 3-D space).

Use coding to solve problems: rolling out from 2020 onwards, all grade 1-9 students will create computational representations by writing, or reading and altering, code and executing it in order to inform and communicate decisions. At higher grades this involves code to analyse numerical data. Choices about using plugged/unplugged approaches or which technologies are left to the school boards.

Benefits

- Throughout the curriculum, new concepts are introduced alongside their applications in context.
- Opportunities to use technology are frequently integral to the expectations, e.g. coding and investigating varying parameters of graphs
- The pathways in year 10-12 ensure all students have studied situations involving linear and exponential functional relationships (whether numerically, graphically or algebraically).
- They also allow specialisms in data management and calculus to prepare for further study.
- There are online versions for the range of grade 10-12 courses allowing more accessibility. Checking a sample of these suggests they do cover the whole curriculum including assessing applications.[93]
- The relative stability of the curriculum means CPD has been focused on improving pedagogy and low-achieving schools.

Barriers

- Needed to assure parents and community that fluent number skills were considered important.

Adaptations

- There was a need for extensive CPD for elementary teachers around coding in the new curriculum.
- Have created a digital curriculum tool so that resources and sample tasks for teachers are all accessed in one place.[94]

---

[93] https://ontariomath.weebly.com/grade-11---mbf3c.html

[94] Christine Suurtamm et al., 'Assessment in Mathematics Education', in *Assessment in Mathematics Education: Large-Scale Assessment and Classroom Assessment*, ed. Christine Suurtamm et al., ICME-13 Topical Surveys (Cham: Springer International Publishing, 2016), 1–38, https://doi.org/10.1007/978-3-319-32394-7_1.

8.2 'Mathematics and Statistics' in Aotearoa New Zealand

Aotearoa New Zealand has a well-established 'Mathematics and Statistics' curriculum which has been informed by research and is seen as progressive by the statistics education community. The national curriculum is notable for taking a strong stance by positioning statistics alongside mathematics:

> Mathematics is the exploration and use of patterns and relationships in quantities, space, and time. Statistics is the exploration and use of patterns and relationships in data. These two disciplines are related but different ways of thinking and of solving problems.[95]

The recent history and practice of statistics teaching in New Zealand illustrates the benefits afforded by creating parallel 'pillars' for mathematics and statistics within the overall subject area up to age 15 (as in Model 4b). A third 'applied mathematics' pillar is under discussion for university preparation, with debate about whether/how its content can combine applied mathematics (e.g. modelling) and data science[96].

The National Curriculum in New Zealand

Schooling is compulsory for children in New Zealand between the ages of 6 and 16 (Years 1 to 11) and then optional until 18 (Year 13). Most students leave school around the age of 17 [97], but opportunities within employment and education can influence students to stay on.

The national curriculum is intended to be broad and coherent. Mathematics and Statistics are compulsory areas of study until the end of Year 10. In Years 11, 12 and 13 student typically work towards one of the three Levels of National Certificate of Educational Achievement, New Zealand's senior school qualification. For each NCEA Level students need to earn 80 credits, meeting standards and being graded through a combination of internal and external assessments. As part of this, some students choose to specialise in Mathematics or Statistics, but all students have to meet a mandatory numeracy requirement of 10 credits in order to attain the NCEA Level 1 award (as in Models 7 and 8).

The latest version of the national curriculum was released in 2007, but there have been refinements since. The New Zealand government has invested extensively and collaboratively in curriculum development, for example by getting curriculum advisors and

---

[95] https://nzcurriculum.tki.org.nz/The-New-Zealand-Curriculum/Mathematics-and-statistics
[96] https://new.censusatschool.org.nz/wp-content/uploads/2021/12/NZSA_Statistics%20and%20data%20science%20at%20CL7_8.pdf
[97] https://www.education.govt.nz/our-work/our-role-and-our-people/education-in-nz/

teacher participants to codevelop materials for both teaching and assessment. When Statistics was introduced in its own right, there was a recognition of a "practice-facing imperative"[98], the importance of establishing teacher confidence, and the need to align assessment practices with the direction of curricular change. As an example, all educators were provided with sample tasks for internal assessments before the curriculum was implemented.

Statistical Investigation and Statistical Literacy

The New Zealand statistics curriculum is divided into three strands: statistical investigation, statistical literacy and probability[99]. The first two of these are most relevant to this discussion.

From the start of the curriculum, *Statistical investigation* involves conducting investigations using the 'statistical enquiry cycle'. One Year 1 project[100] starts with students going through their lost property box, wondering what questions they could ask and answer. They sort and classify the items, before organising them into a physical bar graph, then moving to (and reflecting on) an iconic bar graph. Older primary school students engage with the data provided by Census at School[101], a database both of and for students from Year 5 upwards. As students move into secondary school, they work with more sophistication, considering sources of variation, cleaning data sets and making inferences. There is an emphasis on doing real investigation with real data. At the top end, this strand involves content such as stratified sampling, box plots, linear regression and informal confidence intervals. There is a focus on simulation-based inference, so that students gain experience with sampling and randomisation, as well as an informal appreciation of distributions. Correspondingly, some formal statistics is postponed until university.

Statistical literacy encourages students to evaluate both the veracity of statistical conclusions and the effectiveness or appropriateness of statistical methods. For younger primary school students this might simply involve recognising when a pie chart is more appropriate than a bar graph, but by the start of secondary school there is a move to thinking beyond the data. A 2021 position statement[102] exemplified this by discussing the

---

[98] http://www.tlri.org.nz/sites/default/files/projects/TLRI_Project_Plus_Web_0.pdf (page 30)
[99] https://new.censusatschool.org.nz/new-zealand-statistics-curriculum/
[100] https://sdse.online/lessons/SDSE19-006/
[101] https://new.censusatschool.org.nz/explore/
[102] https://new.censusatschool.org.nz/wp-content/uploads/2021/12/NZSA_Statistics%20and%20data%20science%20at%20CL7_8.pdf (see page 13)

use of 'worry questions', such as "Where does the data come from?", "Is it reasonable to generalise from this data?" and "Could there be alternative explanations for the findings?" By the end of secondary school, a student might have worked to evaluate statistical reports in the media, or had experience of critiquing causal-relationship claims. There is a definite awareness that statistical literacy is a component of citizenship.

Technology

Technology features throughout these strands to aid both statistical work and understanding. Students working in upper primary/lower secondary classes might use Microsoft Excel or the free online data analysis platform, CODAP[103] to analyse the Census at School data. One widely used tool in secondary schools is iNZight[104], a free piece of software developed using R that allows students quickly to interrogate, summarise and graphically present data sets. One unusual way in which the curriculum takes advantage of technology occurs at the top end of the curriculum, where students can use bootstrapping to estimate a population parameter in a sampling situation. This leverages the ability of the computer to produce large numbers of samples quickly to arrive at a confidence interval empirically.

Assessment

The current assessment processes for the NCEA offer schools a lot of choice, but in brief mathematics and statistics are both assessed through a mixture of internal and external assessment. This allows educators to integrate experience with real data sets and technology into the assessment process. Although it can have its own difficulties, internally assessed work is considered important for bringing elements of validity beyond what would be afforded by a simple written exam. The various ways into formal assessment can create complications however, and there is a move to simplify practice at the time of writing, while retaining both internal and external elements[105].

Where Does Data Science Fit?

The New Zealand Statistical Association Education Committee released a statement in 2021, giving their perspective on statistics and data science education at senior secondary

---

[103] https://codap.concord.org/

[104] https://inzight.nz/

[105] Royal Society Te Apārangi Expert Panel. (2021). *Pāngarau Mathematics and Tauanga Statistics in Aotearoa New Zealand. Advice on refreshing the English-medium Mathematics and Statistics learning area of the New Zealand Curriculum*. Royal Society Te Apārangi.

level[106]. They noted that, whilst the present curriculum already touches on some aspects of data science, proper 'data scientific thinking' would integrate more computational aspects, as well as a greater awareness of data ethics, responsibilities, and ownership. Some of the societal implications of modern technology are covered elsewhere in the New Zealand curriculum under 'Digital Technologies', an area which is compulsory up to year 10, but this course does not touch on topics such as machine learning and the use of modelling approaches involving testing and training data.

At the time of writing there is some discussion as to whether the curriculum could integrate a third strand of mathematics provision, 'applied mathematics', and how data science could feature - perhaps heavily - within this. The same statement argues that a specific approach to data science would allow for targeted pedagogy (and learning experiences) that combined statistical, computational, and societal aspects of the field. For example, a student could be challenged to develop an interactive dashboard which summarised and reported on dynamic time series data, identifying trends and making predictions.

Benefits

- Distinguishing mathematics and statistics throughout the curriculum has afforded more space for development in the pedagogy and practice of statistics teaching.

- The curriculum structure gives time and attention to statistical literacy as well as statistical calculations and methods, and there is deliberate interest in promoting critical citizenship.

- Teachers and students have access to free, statistical software, fit for purpose.

- Conversations with New Zealand based researchers have suggested to us that the present curriculum supports higher level study for many, preparing students well for a popular first year statistics course at a well-known university.

Barriers

- Mathematics teachers can lack a strong statistical background; this and other factors can lead to a gap between the intended and enacted curriculum.
- The added distinction between the two areas of study contributes to some tensions between mathematics and statistics.
- The emphasis on statistics and various pathways available to students can hinder a student's ability to specialise in higher education; some universities have needed to

---

[106] https://new.censusatschool.org.nz/wp-content/uploads/2021/12/NZSA_Statistics%20and%20data%20science%20at%20CL7_8.pdf

provide first year calculus courses to support choice of STEM majors. There is a need to consider carefully pathways from high school to university, and how and when to teach the mathematics that is required to support further study in both mathematics and statistics.

- Common concerns about teaching statistics and general education continue under this model, such as how to strike the balance between context and content when conducting investigations, and how to promote achievement across diverse groups of students. This last point is of particular interest to New Zealand educators as they work to decolonise the curriculum, support students from indigenous backgrounds and enable all students to bring their worldviews to their education.

Adaptations

- Curriculum development was informed by both teacher feedback and educational research, both at the outset and in review.
- Teacher development continues to be recognised as important.
- Significant effort has also been put into task design and the production and sharing of resources including those for assessment. It was recognised that teachers would want 'tried and tested' resources from the start.
- New Zealand is a relatively small country, where the state closely informs curriculum development and educational agencies tend to be centralised. Whilst a lot may be gleaned from New Zealand's approach, some of the initiatives and positions taken by New Zealand may be more challenging to enact in the United Kingdom, where there are multiple awarding organisations and interests; at the very least the larger number of stakeholders may reinforce or multiply barriers.
- In PISA 2018 New Zealand scored 494 in mathematics, slightly higher than the average of 489 in OECD countries[107]. However, there has been some decline in performance since New Zealand has been involved in PISA, and there is also gender disparity and meaningful achievement differences between cultural and socioeconomic groups.

---

[107] https://www.oecd.org/pisa/publications/PISA2018_CN_NZL.pdf

8.3 'Analysis' and 'Applications' pathways in the International Baccalaureate

The International Baccalaureate (IB) is a globally recognised provider of international education. The IB takes a broad and balanced approach to curriculum design, requiring students to work both within and across a range of curriculum areas, and placing a high value on 'learner profile' capacities such as communication, risk-taking and reflection. Although the IB has extended its provision since its 1960s origin, introducing Middle Years and Primary Years Programmes in the 1990s, it is best known for its Diploma Programme (IBDP), a two-year pre-university course of study and qualification presently offered worldwide in over 5600 schools in 159 countries.

A student undertaking the standard IBDP must choose one subject to study from each of six subject groups: language and literacy, language acquisition, individuals and societies, sciences, mathematics, and the arts. Typically, a student will study three subjects at higher level (HL), each of which corresponds to 240 hours of study, and the other three at standard level (SL), which each require 150 hours. To qualify for the diploma each student must also fulfil the requirements of three core elements: studies in the theory of knowledge, a 'creativity, activity, service' project, and an independent, self-directed piece of writing called the 'extended essay' (EE). The IB promotes continuity between all these components.

Developing the IBDP Curriculum for Mathematics

The IBDP curriculum for each subject is reviewed in seven-year cycles, with curriculum change informed in part by the changing needs and demands of higher education and employment. The mathematics curriculum was last revised in 2019 (for first assessment in 2021) and instituted some major changes.

In the previous iteration, students could fulfil their mathematics requirement by choosing one of four options: HL (with or without Further mathematics HL), SL and Mathematical Studies SL. Mathematical Studies was presented as a course that was appropriate for students with a wide variety of mathematical abilities. It emphasised the application of mathematics, and contained a large number of statistical techniques. Whilst the IB internally considered Mathematical Studies to be an SL course which met diploma requirements, it was looked on less favourably by some university admissions teams, and in practice it was ultimately perceived as an easier option by many schools and students.

In the new iteration, there is a clearer separation of the domain of the courses and the level of difficulty. Students now choose from two strands of mathematics, known as Analysis and

Approaches (AA) and Applications and Interpretation (AI), each of which is offered at either SL or HL. All four courses contain common Level 3 topics such as introductory calculus, sequences, and graphs of functions, ensuring a firm base for university-level study, but the strands later diverge to reflect differences in students' interests and career aspirations.

The AA strand serves as a pre-university route for mathematically intensive disciplines, and as such can be thought of as more of a standard 'pure mathematics' strand, including (at HL) topics such as proof by induction and Maclaurin series. In contrast, the AI strand focuses more on statistical literacy and mathematical modelling skills and incorporates 'applied mathematics' topics such as algorithms on graphs and Voronoi diagrams. A student might favour AI if they were planning on work or further study in a field such as business, or psychology, where quantitative reasoning and statistical/technological fluency were more relevant than deductive and algebraic proof. The AI course contains much more statistics than the AA one; although an AA student will meet linear regression and the normal distribution, an AI student will gain additional experience with, for example, hypothesis testing, t-tests and (at HL) confidence intervals. Technology permeates teaching and assessment in both strands, but especially so for AI students.

The SL courses maintain the same foci as their HL counterparts but have less content. A student might choose AA SL if they enjoyed the more abstract, logical and problem-solving aspects of mathematics but did not want to take it at HL because of their other interests or career plans. Topics within AA SL include the laws of logarithms and how to solve quadratic equations in sin x, cos x and tan x. A student might choose AI SL to see how mathematics could be used in various contexts. Topics within AI SL include amortization and annuity, fitting mathematical models to data, and chi squared tests.

When it was first introduced, the AI course was heavily promoted as it was a significant change to what had come before. Although the AI HL course is academically demanding and suits those going into a range of university courses, there was a danger that stakeholders would see AI as the 'easy' option. The IB assuaged some concerns by researching and sharing how university admissions teams would perceive this new option. The new AI course has proved popular. In May 2022, out of 104,000 candidates, 56% studied the Analysis and Approaches course (19% took HL and 37% SL) and 44% the Applications and Interpretations course (7% took HL and 37% SL).[108] Overall, the uptake of HL has increased to

---

[108] https://www.ibo.org/globalassets/new-structure/about-the-ib/pdfs/final-statistical-bulletin-dp-cp-may-2022.pdf

26% from 15% in 2019. Teachers' mindset regarding the new courses continues to shift as the AI strand becomes more established, and the IB continue to refine their presentation.

Assessment and Practical Issues

The assessment of these four courses involves externally assessed written examination papers. For AA, one of the papers is non-calculator, but AI students use calculators as needed. There is also an independent project which is internally assessed (but centrally moderated) and worth 20% of the final grade. This is intended as a chance for students to choose and develop a piece of mathematics that is significant and of interest to them. Some students find that the mathematics they need stretches slightly beyond the course, or across strands (for instance, an AA student might find that their data does not have a linear association and learn about Spearman's rank correlation coefficient, which is only in the AI strand.) The resulting projects are very diverse, but a fair number of students use some element of statistical modelling within their projects and get to deploy important data science skills such as sorting and cleaning data in practice. The assessments are valid but the IB recognises that they incur significant workload for teachers. Even though they allocate 30-hours of curriculum time, plus guidance and toolkits for developing research skills, the production of the projects can be challenging.

HL candidates in both strands have a third written examination paper which is one hour long and consists of two questions. This is intended to test sustained reasoning. Although not mandated, generally, it is common for one of the two questions on the AI HL paper to have a statistics focus, allowing for the testing of more involved statistical work under examination conditions. However, this does not encompass the skill of working with large data sets, which the England example shows can be problematic to integrate into externally assessed examinations.

Since 2012, the IB also offers an alternative career-related programme[109], where students take subjects from a minimum of two DP areas and carry out a career related study with an external provider. It is interesting that there has been significant uptake of the AI strand of mathematics by students on this programme, even at the challenging HL level.

What Can We Learn from The Changes to the IB?

The recent changes to the IBDP mathematics courses demonstrate that it is possible to arrange choice pathways for pre-university students (our model 7) within a baccalaureate

---

[109] https://www.ibo.org/programmes/career-related-programme/what-is-the-cp/key-facts-about-the-cp/

The Open University

structure (model 8). The four combinations of AA/AI and SL/HL go a long way towards disentangling the level of difficulty from the focus of the curriculum, and this has undoubtedly given students more options and broadened the appeal of HL mathematics. A student who is good at mathematics but wants to study psychology or business, for example, can now study a course which is more relevant to their goals but still appropriately challenging. There was some concern amongst different groups of stakeholders when these courses were first introduced which had to be managed.

The assessment model of the IBDP highlights both the potential and problems of extended project work (as in model 5). The IA can result in authentic work of a very high level, and many students successfully work with statistics for this element of the course. However, this can place a large demand on teachers' time and attention. Mathematical writing, and the integration of technology into mathematical work involves new skills, and the assessment rubric must recognise and balance the mathematics with how it is communicated and structured. On a related note, when students choose their topic areas for the core extended essay, mathematics is one of the least popular topic areas, partly because schools do not share a consensus about what constitutes critical thinking with and about mathematics.

What Might the Future Hold for Data Science within the IB?

The closest relation to data science within the IBDP at the time of writing is mathematics, particularly the AI strand, as this contains the statistical elements. Further, all students can take a data science approach to their individual projects within the mathematics component. The IB is already undertaking its review for the 2026 mathematics curriculum, and one of the aspects it is looking at is how data science and modelling are handled. A relevant question is whether mathematical modelling (with an extra-mathematical component) may in future be considered separately from problem solving. Another is whether elements of data science might be integrated elsewhere, for instance in the 'digital society' course that was introduced in 2022. The ability to place at least some elements of data science outside of mathematics is an advantage of a diploma structure and, with coordination, has the benefit of reinforcing to students how uncertainty is a real-world feature. Although a diploma programme does provide common structure to students' learning, in practice it is still difficult to write good examination questions that involve concepts from other subjects. These are usually very carefully scaffolded.

The IB is also currently studying how data science competencies track through their programmes for primary and middle year students, using a continuum moving from data literacy to data handling to data science.

Benefits

- The provision of four mathematics courses, arranged across two parallel strands and matched at two levels of difficulty, disentangles the choice of content from the level of challenge.
- The new suite of courses is more flexible, so that the mathematical study better aligns with students' future plans (both in terms of education and careers), as well as supporting students entering IB study from different mathematical backgrounds.
- The inclusion of an internally assessed project (for all students) and sustained reasoning paper (HL only) offers valid opportunities for students to apply statistical and data science skills, and new avenues for assessment.
- Technology is recognised and integrated throughout the courses of study.
- The mathematics curriculum is regularly revisited in light of changing international expectations.

Barriers

- The initial introduction of the courses met with some concern and confusion from various stakeholders, including universities, teachers, parents and students.
- The independent project is introduced, managed and assessed internally by schools, and this can have a impact on teacher workload.
- Whilst large data sets may feature in individual students' projects, they do not feature in the wider curriculum or assessment yet.
- All changes and developments in assessment and curriculum have to sit within and match the wider IB qualifications and framework.
- A limitation: the IB cohort largely consists of those preparing for university.

Adaptations

- The IB conducts significant liaison with stakeholders prior and during the teaching of the new courses.
- The IB has significant investment in ongoing teacher professional development, running courses and conferences worldwide, as well as providing substantial online content. Each school also has a designated 'IB coordinator' who not only has administrative responsibility but is tasked with pedagogical leadership and support.

8.4 ProDaBi – Project Data Science and Big Data in German schools

Project Data Science and Big Data at School (ProDaBi)[110] is a project that has been run since 2018 by computer science education and mathematics education researchers at Paderborn University. The project is funded by the foundation Deutsche Telekom Stiftung.

ProDaBi has been developing and trialling data science education materials (and some associated professional development) intended to raise students' awareness of data as a model of the world, data processing, exploration of data, and to teach them how to use data-driven machine learning (ML) for making predictions and decisions. The project started with a year-long trial with Grade 12 (Year 13) students in two schools in Paderborn, with three-hour weekly sessions, run over seven months in the school year 2018/2019 [111]. In subsequent years materials have been developed and trialled lower down the age range, first with teaching modules for lower secondary aged students, and then into the top end of primary school. Lessons have been typically taught by computer science teachers.

ProDaBi offers an example of how a computing-focus approach to data science is presently being enacted (our Model 2 for adding content to computing). It also embodies the paradigm that machine learning is a fundamental component of data science which can be communicated to school-aged students.

Predictive modelling as a fundamental idea in data science education

A starting premise for the ProDaBi project is that the industrial processes of data science are distinct from the stages of the statistical modelling cycle in many important ways. Significant amounts of data may already exist which have not necessarily been collected to a plan; the domain/business context is a crucial starting point; data preparation and cleaning are separate, time-intensive steps; and training and testing a model with data is a distinct and important phase of the data science cycle [112]. Instead of reaching a statistically rigorous conclusion, now the goal often is to develop, validate and deploy a model which can be used for context-informed classification or prediction. This in turn raises new and important issues of social responsibility and ethics.

ProDaBi developers list three aims for their activities – demystifying the data-based machine learning (ML) model building process, investigating ML by developing their own

---

[110] https://www.prodabi.de/

[111] https://publikationen.bibliothek.kit.edu/1000127944/97090875

[112] Please see slides and video at https://iase-web.org/Webinars.php?p=230307_2000

models, and reflecting on responsible use of ML. One reason that ML is seen as integral is that data sets (and real-world contexts) typically involve multiple variables and tracking the resulting multitude of correlations by hand is simply not possible. Students could be introduced to simple classification models as an example of machine learning, for example deciding whether an image is of a cat or a dog, or determining whether a lizard has come from a natural or disturbed habitat based on attributes including its mass, length of limbs, tail length, head depth and toe-pad width [113].

Pedagogical and computational models

The ProDaBi project uses the notion of *decision trees* to develop and test its predictive algorithms, where at each stage of a multi-stage process one uses a particular attribute (and a threshold) to categorise a candidate. Decision trees serve as a transparent model for students, helpful for understanding the 'machine room' of artificial intelligence (AI), compared to, say, artificial neural networks (ANN). Decision trees provide an opportunity for teachers and students to discuss potential advantages and disadvantages of automation, and the developing role of humans in AI [114].

Students meet decision trees for the first time at the top end of primary school, where they work with physical cards that name an object and list some of its attributes. This 'unplugged' approach doesn't introduce them to the mathematics or computation of data-based ML but the students get to explore the heuristics and properties of the process, recognising that different algorithms can result in different outcomes and having an opportunity to compare and discuss the accuracy and biases of their choices.

The resources provided for early secondary school students involve the students using larger data sets with semi-automatic support, specifically the computing environment Common Online Data Analysis Platform (CODAP) [115] which facilitates 'drag and drop' for articulating commands. At the top end of secondary school students can use Jupyter notebooks (cell-based environments that support creation or adaptation of Python code) or Python programming to access and work with large data sets, with Python libraries to create decision trees and related visualisations. Students are also tasked to grapple with

---

[113] Example from GAISE II report, p. 97 https://www.amstat.org/asa/files/pdfs/GAISE/GAISEIIPreK-12_Full.pdf

[114] Slide 28 at https://iase-web.org/Webinars.php?p=230307_2000

[115] https://codap.concord.org/

The Open University

assessing rules using misclassification rates, issues of overfitting data, false positives and false negatives, growing and pruning decision trees, and evaluating their algorithms [116].

Meaningful contexts to engage students

As well as giving students experience of machine learning processes, the ProDaBi resources demonstrate the importance of context in Data Science and are based around recognisable scenarios from the real world.

At the top end of primary school ProDaBi uses the example of choosing/recommending food items (based on nutritional information). Students are given 55 data cards and green and red paper clips to label the cards. Two example cards are given below:

Figure 5 Data cards – Nutritional facts of food items (Slide 56, at https://iase-web.org/Webinars.php?p=230307_2000)

In this activity, students get to manually create a decision tree as a rule system for classifying food items, test their tree with test data, compare different trees/approaches and explore how a machine learning method might create a decision tree.

[116] Biehler, R., & Fleischer, Y. (2021). Introducing students to machine learning with decision trees using CODAP and Jupyter Notebooks. Teaching Statistics, 43(S1), S133–S142. https://doi.org/10.1111/test.12279
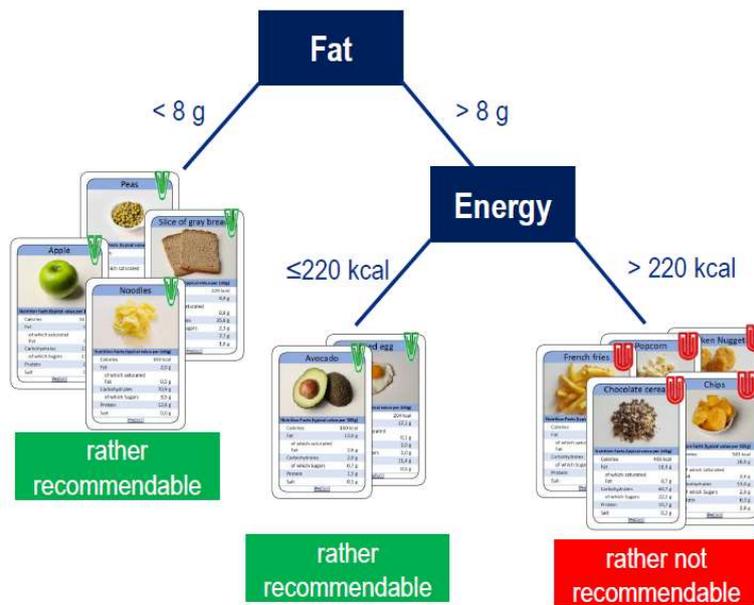
Figure 6 A decision tree for food items (Slide 64, at https://iase-web.org/Webinars.php?p=230307_2000)

One of the contexts in the materials for older students is based around social media platforms' use of online gaming behaviours to decide which advertisement is shown to a player, with data fields including the type of device someone uses to play games, the online platforms they use, the types of game, how often they plan, and so on. An authentic data set to support this resource was created by the ProDaBi developers working with secondary schoolchildren.

Figure 7 A student's solution creating a decision tree with CODAP (Biehler & Fleischer, Teaching Statistics. 2021;43:S140 [37])

Students can use the Jupyter Notebooks to focus on one menu-based action at a time or move on to use more versatile code-based actions; they can also use the notebooks to add explanations and narrative.

Benefits

- The unplugged/embodied, menu-driven and Python code-based implementations allow different age groups to explore ML.
- Relevant and authentic contexts support student engagement and sense-making.
- Students get to appreciate the role and need for computational tools, automation and ML.
- Students see and grapple with the ethical implications of ML.

Barriers

- The mathematical demand of some of the constructs (confusion matrices / two-way tables) relating to the decision trees used in some of the resources can be problematic – students need to be able to use these.
- Teaching was done by computing teachers and there are related teacher supply issues.
- Some teachers may have difficulties filling in student gaps in mathematical knowledge.

Adaptations

- Team teaching could be considered to draw in complementary knowledge across disciplines.
- The developers are experimenting with different sized data sets to draw out different aspects of ML.

8.5 Scotland's post-15 National Progression Awards in Data Science

Scotland's "Curriculum for Excellence" [117] lays out a programme for students from ages 3 to 18, consisting of a broad general education followed by a senior phase from the ages of 15 to 18, during which students can begin to specialise and attain formal qualifications.

Scotland is notable for having recently added a set of three national qualifications to the list of options: National Progression Awards (NPA) in Data Science [118]. Although these have had limited uptake to date, their construction and outworking speaks to how an additional distinct qualification in data science can exist alongside other qualifications in mathematics. (This approach supplements rather than replaces existing qualification so is similar to Model 5)

The Data Science NPA

There are three levels of NPA in Data Science, labelled 4, 5 and 6 after the matching levels of award in the Scottish Credit and Qualifications Framework (SCQF) [119]. The qualifications are intended to provide students with a foundation of knowledge in data science, developing both applicable skills and an "awareness of the societal aspects of this important, emerging technology" [120].

The three Data Science NPA sit against a background of related national mathematics qualifications, such as the Higher 'Applications of Mathematics' (also SCQF 6) which includes statistical content such as regression and hypothesis testing. However, the NPA have a specific focus on data science and have been designed with flexibility in mind; students can work towards these qualifications in school or in college, such that they might also be appropriate for adults wishing to retrain.

The level 4 NPA consists of two units: data citizenship and data science. These cover basic statistical techniques and data visualisations, but also require students to reflect upon how

---

[117] https://education.gov.scot/education-scotland/scottish-education-system/policy-for-scottish-education/policy-drivers/cfe-building-from-the-statement-appendix-incl-btc1-5/what-is-curriculum-for-excellence
[118] https://www.sqa.org.uk/sqa/91458.html
[119] These are comparable with levels 1, 2 and 3 in the UK's Regulated Qualifications Framework.
[120] https://www.sqa.org.uk/sqa/91458.html

the ways in which data is collected, processed, and presented can have consequences and make a difference, both to individuals and in society.

The level 5 and level 6 awards each contain more advanced versions of these core units as well as a third module from a choice of: data security, computer programming, machine learning, a data science project, data science statistics, and (at level 6 only) an additional statistics option[121]. Within these modules students might engage with primary data from environmental sensors, see how machine learning uses testing and training datasets, or develop their programming skills. The choice of modules again reflects the flexibility of the qualification; it is envisioned that the overall award could be taught across departments, with teachers from mathematics, computer science or geography teaching towards their own strengths and interests. Similarly, whilst students can develop programming skills in a language such as Python if they choose, each NPA can be attained using only Microsoft Excel or Google Sheets.

At the time of writing there has been limited but developing take up of the qualifications; an estimated 1000 students have studied the NPA since it was introduced in 2020.

Delivery Models, Assessment and Teacher Support

The adaptability of the Data Science NPA has been promoted as a strength. The educator's guide[122] is written with an awareness of the fact that many teachers have to work with classes where students are working at different levels of the NPA, or even on different qualifications. Even where students are all studying together, they might encounter the Level 4 NPA as an intense enrichment experience within their mathematics lessons, or tackle all three levels over multiple years with a view to progressing to progressing ultimately to a degree in Data Science. There is even overlap with other qualifications; for instance, the data security module also counts towards the NPA in Cybersecurity.

Schools can set their own assessments for their students, although these need to be checked against unit specifications and approved before use. Alternatively, the core units can be assessed through a combination of computer-based tests and centrally provided projects which are marked by teachers then submitted online.

The courses are strongly supported by work carried out by Data Education in Schools [123], a group of researchers, educators and consultants based out of Edinburgh University. They

---

[121] https://dataschools.education/wp-content/uploads/2020/11/educators-guide_Oct2020.pdf
[122] https://dataschools.education/wp-content/uploads/2020/11/educators-guide_Oct2020.pdf
[123] https://dataschools.education/

provide an extensive array of materials [124] developed with the data science organisation Effini [125], and specifically tailored to the NPA.

Future Developments

Although the NPA is only of direct relevance to senior secondary school students, there is growing interest in developing data science content and data literacy throughout the curriculum stages. For example, Data Education in Schools also provides resources for younger students. These link to the curricula of multiple subject areas, touching on many recognisable contexts, and integrating technology at various points.

Benefits

- Having a separate qualification allows a broad range of data science content and skills to be covered.
- The flexibility built into the qualifications allows students, schools and further education colleges to engage with data science in a number of ways, promoting access and ultimately supporting equity.
- The modular approach to the qualification encourages participation of staff from different curriculum areas, hence using the cross-curricular nature of data science as a strength.

Barriers

- The qualification exists in a crowded field.
- It can be challenging to balance the relative difficulty of various modules.

Adaptations

- To increase the relevance of the qualification it has been designed and promoted so that it sits comfortably across the school/further education boundary.
- Teachers are supported with some high-quality resources and there are opportunities for professional development (although it is recognised that more are needed.)

---

[124] https://dataschools.education/npa-course-materials/
[125] https://effini.com/

8.6 MEI supplementary Data Science Courses

MEI is a charity organisation concerned with advancing mathematics education. They see data science as an important interdisciplinary field which is gradually filtering down into schools. Since 2020 they have supported schools by delivering online data science courses aimed at students in Years 12 and 13, designed to sit alongside either A level Mathematics or Core Mathematics. Such courses could be the basis of credit-bearing extra qualifications (our Model 5)

The MEI data science courses

MEI presently offer two online courses. The shorter course is a free, open access self-study programme designed as an introduction to data science [126]. Over six lessons, students cover topics including cleaning data, data visualisation, and a brief introduction to machine learning. The course also involves work with the A level Mathematics large data sets, and teachers can recommend it to their students as a way of increasing their understanding of working with data.

The longer course is a term-long taught programme. MEI runs each presentation of the course with limited spaces; teachers apply on behalf of their students, so one cohort might involve about 150 students from 30 different schools. Each week students take part in a one-hour live session and support this with independent study. This course is formally assessed through a combination of a practical task and a one hour written assessment, each worth 50%. Although schools must administer the assessment, MEI set and mark the work, and award certificates. There are three levels of award: pass, merit and distinction. Students who submit the assessments but do not achieve at least a pass in both are sent a certificate of attendance.

The vast majority of the students involved with the longer course are A level students, but the mathematics involved is accessible to both A level and Core Mathematics students, covering topics such as linear functions, scatterplots, correlation and regression. There has been a conscious decision to circumvent probability within the course and to emulate the processes of data science; the goal is to create predictive models rather than explanatory ones, processes such as linear regression are presented as 'black boxes' enacted by code, and relative judgements are based on metrics, not probabilities.

Both courses involve some element of programming. During development, MEI considered various tools but ultimately chose to use the freely available, web based Kaggle platform

---

[126] https://mei.org.uk/introduction-to-data-science/

which allows students to combine text and coding in Python (which is already taught in many schools as a foundation language and is widely used in industry). However, MEI is clear that the data science course is not a programming course. Students mainly import libraries for data analysis, visualisation, and machine learning, and alter routines using the commands in those libraries. By providing students with sample blocks of code to copy, adapt and build on, the courses intend to recognise the relevance of programming to data science and support transferrable skills without assuming too much at the point of entry.

Data science sensibilities

The courses were developed in collaboration with computing educators and influenced by the International Data Science in Schools Project (IDSSP) [127]. The courses have also been reviewed by data scientists in industry and academia. Their content and delivery both reflect certain key data science sensibilities.

The course positions technology as central to data science. The course developers believe that data at scale needs technology, and that meaningful, authentic work can rarely be done with pencil and paper. They also recognise the import of representations in data science, for example when students are exploring distributions, or presenting their findings. Students need access to a range of visualisations, including some dynamic ones, and the affordances of these visualisations are unavailable without technology.

Both courses also evince the importance of context, or domain knowledge in the practice of data science. The contexts covered within the courses are varied and have been picked so as to be accessible and of interest to 17-year-olds. They include hotel cancellations, house prices, engine sizes and emissions, climate data and basketball players. Students access these contexts via historical, open-source data sets that are closed to revision, and which may require some data cleaning. Context is also used explicitly to inform method; for instance, students might be required to make a sensible context-based decision as to whether an outlier constitutes an error or a relevant piece of data.

The courses touch on some aspects of ethical issues surrounding data. There is no discussion about data stewardship or privacy, possibly due in part to the time available, but there is exploration of the implications of using models in context, and how bias (conceptualised as a mismatch of the deployment of the model to the data) can be introduced to a model through the choice of training data.

---

[127] http://www.idssp.org/

The short course integrates insight into career opportunities in data science, including video interviews with working data scientists.

Machine Learning, AI and the relationship Between mathematics, computing and data science

Both courses integrate aspects of machine learning and AI. A pyramid shared with students in the short course starts with data sources and pre-processing at the base, moves up through data exploration and analysis, then goes onto machine learning before finally reaching decision making. Practical experience with training and testing data also features in some of the tasks; a large data set might be split, so that 80% can be used to build a model whilst the remaining 20% is held back to test it.

However, a representative of MEI made it clear to us in conversation that they saw data science as part of the mathematics and statistics curriculum. They offered three main arguments for this: that mathematical and statistical analysis and skills form the foundation of data science and artificial intelligence; that data science is pedagogically informed by statistics education; and that whilst data analytics is grounded in computing, data science places a higher import on context than computing has traditionally done. Notwithstanding, they also recognised that computing had an important part to play, and that teachers moving from mathematics might be uncomfortable with some aspects of computational thinking, including the acceptance of 'black box' procedures and the readiness to engage with complicated, multivariate contexts.

It was also suggested that placing data science closer to mathematics and statistics than computing might steer around some of the significant gender gap present in A level computing; 37% of the students covered by the post-pilot survey [128] were female, compared to 15% of A level computing candidates in summer 2022 [129].

Benefits

- Positioning these courses as separate, supplementary study means that there are fewer demands and restrictions on the content.
- All the staff involved in delivering the course are specialists with good knowledge and understanding of data science as a discipline.

---

[128] https://mei.org.uk/app/uploads/2021/08/mei-ds-short-course-pilot-report.pdf
[129] https://www.jcq.org.uk/wp-content/uploads/2022/08/A-and-AS-Results-Summer-2022-Compared-to-2021.pdf

- Authentic, large-scale data is used within the courses, and the assessment for the taught course involves both a practical task and a written paper; the examination and certificate promote involvement and add cachet.
- Successful integration of user-friendly but commercial-grade technology

Barriers

- The courses have been popular, and the longer course has always had a large number of students applying, but it would be both practically and economically challenging to reproduce this collaborative approach at scale.
- The short lengths of the courses limit both the width and the depth of the content compared to other curriculum domains, and there is no provision at all for younger students.
- It can be hard to find good quality data sets in contexts which 17-year-olds will understand enough to appreciate the relevant attributes, whilst still be interesting enough for a student to be motivated genuinely to ask and answer a question.

Adaptations

- MEI have balanced or adapted the statistical and programming content so that the course is accessible to a wider range of students – in particular, the course is suitable for interested students studying either A level or Core Mathematics.
- The courses involve collaboration between MEI and schools, from the point of application to the administration of the examination.

8.7   Geography and Core Mathematics Project

The concepts and skills taught in school mathematics have relevance and application within other curricula. The 2017 Smith Report [130] for instance notes that mathematical and quantitative skills form part of the formal assessment of a third of non-mathematics A levels. This intersection is particularly evident through the teaching of Core Mathematics, with its emphasis on applied mathematical skills and authentic cross-curricular contexts.

The Geography and Core Maths Project [131] is a DfE funded partnership between the Advanced Mathematics Support Programme (AMSP) and the Royal Geographical Society (with the Institute of British Geographers) (RGS) intended to consolidate one such connection. Through a combination of continuing professional development (CPD) courses and online resources this initiative works to support the approximately 80% of A level Geography students who are not studying A level Mathematics [132], and to support their teachers in delivering quantitative and statistical content. It provides a single-subject precedent for our Model 6 (current KS 5 structure with content added across subjects) and speaks to how joined-up approaches inform future avenues for professional development.

Mathematical and statistical skills in geography

Throughout secondary school, geography lessons regularly incorporate aspects of mathematical and statistical reasoning. This can call on skills that are well covered by the mathematics curriculum, such as finding a percentage change or reading a compound bar chart, but it may also involve more domain specific work, for instance working with geospatial data or understanding a kite diagram. A level Geography also includes some inferential statistical methods which go beyond the content of GCSE Mathematics. Notably at A level, the geography programme of study includes a non-examined assessment (NEA) worth 20% of the grade, for which each student must conduct some fieldwork and write up a report, including sections of data processing and analysis.

In this way geography teachers are required to touch on a significant amount of mathematical and statistical content, some of which is likely to be new to their students, or even to them. This challenge is particularly pronounced for some when facilitating the NEA, as there is less support from textbooks and prepared examples. Projects often involve primary quantitative data drawn from observation or the use of tools such as anemometers and pH gauges, and in each case geography teachers must first steer their

---

[130] I'm assuming we already have this reference, but Smith (2017) para 113
[131] https://www.rgs.org/geomathsbriefing/
[132] https://www.rgs.org/geomaths/

students towards appropriate statistical methods for their chosen context, and then support them in reaching conditional or qualified conclusions.

The data-related elements of the present geography curricula are consonant with the mathematical and data literacy competencies presented elsewhere in this report. For instance, the fieldwork component of the A level course includes some discussion of data gathering and ethics, although there is limited discussion about cleaning data or data protection issues. Teachers can also bring large data sets into the geography classroom, for instance by accessing the NOMIS data on the labour market [133]. Data trends can be identified using online analysis and visualisation packages such as Gapminder [134] and students can work with location-linked data via representations such as heat maps and flow maps via software such as ArcGIS Online[135].

The RGS-AMSP Collaboration

Since 2019 AMSP, in liaison with the RGS, has delivered tailored CPD sessions for geography teachers working with mathematics. The broad goals of the project are to develop geography teachers' data skills and confidence, and to promote the parallel study of geography and core mathematics [136]. These sessions generated immediate interest, particularly as recent changes to geography specifications had increased the expectation for geography teachers to integrate data skills into their teaching. The CPD provision was adapted over the pandemic period, and at the time of writing AMSP are offering smaller sessions which focus on specific content or processes, such as chi-squared tests, Spearman's rank correlation coefficient, or how to engage with big data sets. These presently involve hundreds of teachers, with more teachers participating each year. Sessions are recorded for those who are unable to attend, with a certificate for those attending four sessions or more out of a set of six. AMSP also offers dedicated online support via the MEI Integral platform. Feedback from the sessions has been positive; comments from participants suggest that this programme has developed geography teachers' knowledge of statistics and how to teach it, as well as bolstering their confidence and willingness to integrate technology into their teaching.

---

[133] https://www.nomisweb.co.uk/

[134] https://www.gapminder.org/

[135] https://schools.esriuk.com/

[136] Similar provision is also offered for teachers of biology and psychology. Although this has also had some success, it is not detailed here for reasons of space, and because the respective subject organisations have been less closely involved.

Benefits

- For students, cross-curricular teaching of statistics can create learning situations where the questions and the answers are perceived as being meaningful; abstract teaching of statistics can lean on repeated calculation and ersatz contexts.

Barriers

- The developing prominence of statistical methods within A level Geography has moved some teachers outside of their comfort zones, leaving them feeling challenged and necessitating space for and provision of professional development. Whilst AMSP is addressing this need for some, it can be hard for teachers to access CPD.
- It can be challenging to assess meaningfully some statistical/data analysis skills through timed written assessment. Context-based learning of methods such as Spearman's rank correlation coefficient may end up being reduced to cloze-type completion questions.
- Cross-curricular delivery across siloed subject areas meets barriers due to small differences such as vocabulary, as well as larger difficulties due to specification changes. For instance, geography teachers have noted that students cannot now be assumed to be familiar with spreadsheet packages since Information and Communications Technology (ICT) has been superseded by Computing in many schools.

Adaptations

- For teachers, access to professional development which is jointly devised and then revised by both a subject association and a mathematics education organisation warrants relevant, considered content and sensitivity to the demands of both areas.
- A formal relationship between the RGS and AMSP has supported long-term provision which has developed in response to evolving teacher needs, combining live and asynchronous learning.
- Some awarding bodies (for example Pearson [137]) have provided additional documentation, outlining the nature of the curriculum overlap, differences in approach and vocabulary, and noting common misconceptions.

---

[137] https://qualifications.pearson.com/content/dam/pdf/A%20Level/Geography/2016/teaching-and-learning-materials/Maths_for_geographers_guide.pdf

8.8 Cross-curricular work in Civic Statistics

The term 'Civic Statistics' was coined by an international collaborative group (ProCivicStat) funded by Erasmus+. A product of the ProCivicStat project is a collection of teaching and learning materials designed to support statistics teaching that engages with current social issues. [138] The team behind ProCivicStat offer on their website that:

> We believe that students should see the usefulness of statistics in understanding evidence, and should engage with data about important social phenomena, to support their development as active and empowered citizens.
>
> Social phenomena are complex, and democracies need citizens who can explore, understand, and reason about information of a multivariate nature. Most statistics courses fail to teach these skills [139].

This position has been illuminated by the recent pandemic, where citizen engagement with statistics involved contested concepts and metrics, multivariate phenomena, and non-linear relationships [140]. The ProCivicStat authors go on to offer an eleven-facet framework for teaching civic statistics[141] which is designed to ensure fair coverage of engagement and action, knowledge, and enabling processes.

Civic Statistics in the Curriculum and the Importance of Citizenship

Although the ProCivicStat resources do not constitute a full tried-and-tested curriculum, and there is limited research on the efficacy of this approach, the decades of experience and thinking of the movement behind ProCivicStat offer some insight into how citizenship aspects of data science might be integrated into other subjects (Model 3). Together the resources are designed to promote civic engagement through statistics, although not necessarily in mathematics lessons. Materials cover a wide range of topics such as income inequality, migration, and ageing societies. Teacher guidance is provided in multiple languages and linked to outside sources of data and support materials. Many contexts sit

---

[138] These materials are hosted on the website of The International Statistics Literacy Project (ISLP) (which in turn is a subgroup of the International Association of Statistics Education (IASE), the education section of the International Statistics Institute (ISI)).

[139] https://www.iase-web.org/islp/pcs/

[140] J. Ridgway, Covid and data science: Understanding R0 could change your life, Teaching Statistics 43 (2021), S84–S92. https://doi.org/10.1111/test.12273

[141] https://iase-web.org/islp/pcs/documents/Conceptual_framework_brief.pdf

more comfortably in subjects such as geography or politics, where students might be better prepared to ask and investigate meaningful questions of interest to them.

One contemporary starting point for work in Civic Statistics which was shared with us was the OECD Better Life Index [142]. This website offers an easy-to-use graphical interface where students can adjust the relative importance of eleven indices and rank the quality of life in various countries, and illustrates well how teachers might adopt different but complementary approaches. A mathematics teacher might use this to support work on proportional reasoning and weighted averages, and a data science educator could get the students to download the dataset, edit it as necessary for specific purposes and then explore other representations. A critical citizenship perspective, which might be adopted by an economics teacher, would further bring in questions such as "what are the benefits and limitations of indicator systems such as this?" and "how can you measure democracy?" These questions are consonant with further data science sensibilities as students begin to interrogate the source and handling of data, as well as associated ethical issues.

The term 'civic statistics' might suggest that it should be taught within citizenship, a foundation subject at KS 3 and KS 4. However, an academic interviewed in associated with the project noted that the teaching of citizenship in UK schools is not always valued, even though it invokes a wide range of issues with substantial impact on students. They therefore felt that coordinating civic statistics across established curriculum areas such as English and Business Studies would be both more effective and more authentic than separating it as a component of taught citizenship. Teachers should be encouraged to challenge with their students the statements concerning their domain that they see in the media, perhaps by engaging with reports from fact-checking organisations such as FullFact [143] in lessons or, for example, demonstrating how choosing the position of the origin on time-series data for salaries or inflation can steer the implications of a graph.

Civic Statistics and Data Literacies

Whilst they are primarily built around the goal of getting students involved in social issues, the ProCivicStat resources regularly take an approach which resonates with many data science principles and practices. Context is paramount and data visualisation often forms an integral part of the work, which can be invaluable in helping students make sense of the data. Technology affords students the chance to work with real data sets, and the main PCS resource database is searchable not only by statistical topic and level of difficulty, but also

---

[142] https://www.oecdbetterlifeindex.org/
[143] https://fullfact.org/

by the applications or platforms which the students will be using. These include Gapminder, INZight, R and CODAP. Notwithstanding, civic statistics itself does not involve any compulsory use of explicit or original programming.

Future Developments in Civic Statistics

The nature and approaches of civic statistics continue to be discussed [144] against a background of pressing issues including climate change, diversity, and social equality. Developments in content include an increased awareness of machine learning processes: common techniques such as decision trees might be modelled with students through activity and analogy, and students of all ages can be expected to engage with and evaluate the output of AI, and to reflect on the role it might come to play in society. Advances in pedagogy include a consideration of how civic statistics might be assessed through project-work or other non-traditional means.

Benefits

- A civic statistics approach in principle offers an authentic enhancement of multiple curricula in a way that intersects and resonates with many aspects of data science.
- The broad and flexible scope of civic statistics could appeal to a lot of stakeholders and meaningfully contribute to progress with regard to some equality, diversity and inclusion issues in education.
- Civic statistics could be integrated into multiple curriculum models; for example, it could be delivered as part of a citizenship-facing component within a baccalaureate model or across multiple subjects in secondary school with the help of an in-school coordinator.
- Developments in technology and the interest associated with current affairs mean that there are opportunities for educators outside of mathematics and computing to engage with data science in a way that entails low buy in but high payoff.

Barriers

- Teachers would require professional development to learn how to think in terms of civic statistics, how to use new technologies, and (potentially) how to coordinate the delivery of civic statistics across and between subjects.

---

[144] Ridgway, J. (ed.) (2022) Statistics for Empowerment and Social Engagement: Teaching Civic Statistics to Develop Informed Citizens. Springer ISBN: 978-3-031-20748-8

- Formal assessment of civic statistics could be complicated by the need to balance validity and standardisation.
- Incautious development or presentation of civic statistics could make it appear as an 'easier' form of statistics, rather than a demanding approach that draws on knowledge and sensitivities from multiple disciplines.
- The contemporary focus of civic statistics would require resources to be constantly renewed, or for local parties to be trained and empowered to develop their own datasets and resources.

Adaptations

- Stakeholders could be invited in from multiple levels – not just from within the education sector but beyond, for instance official statistics providers and the media could contribute to training, the production, and the updating of resources.
- The website is designed to be user-friendly and accessible for educators from a wide variety of backgrounds and starting points, mapping content to statistical topics and tools. This could be developed even further so that subject specialists could more easily structure pathways for their students that attended to key skills and approaches from civic statistics.

8.9 Bootstrap integrating computing with other subjects

Bootstrap [145] is a US-based initiative, housed at Brown University, that provides free curricular materials and professional development for integrating Computer Science and Data Science into mathematics, science, business and social studies classes for students from US grades 5 to 12 (expected ages 10 to 19). The programme has expanded significantly since its origin: it has received multiple grants from charitable and corporate organisations, and is now well established in the US, such that in the academic year 2021/22, it reached nearly 30,000 students. Further, according to their website, "nearly 50% of our students are African-American or Latin(x), and nearly 45% are girls and young women." [146] The Bootstrap website foregrounds three principles: equity, scale and rigour. [147]

Bootstrap offers an example of how a programme of study can be developed explicitly around the intersection of computing and mathematics (and later computing and data science). It is also notable for promoting a model of provision which does not rely heavily on specialist computer science educators; according to their 2015 evaluation, approximately three quarters of the teachers did not have a computer science degree, and the same fraction did not have computer science as their main teaching assignment. [148]

Bootstrap:Algebra

Bootstrap:Algebra was originally devised as an after-school programme, but has since evolved into a nationally recognised curricular resource which is suitable for in-class use. The current version is promoted as a 25-hour course where students get to create a video game, whilst working with mathematical content and computer science concepts. [149] Game design is seen as an engaging starting point from which students can connect to key ideas from mathematics and programming.

The Bootstrap curriculum is underpinned by several pedagogic and design principles intended to support students in making connections and engaging with cross-curricular work. Notably, Bootstrap:Algebra uses the functional programming language Racket, which treats variables and functions in a similar way to mathematics. The intent is to smooth the

---

[145] This case study is based on content and research papers available via the Bootstrap website https://www.bootstrapworld.org/ . Many of the research papers, on the development and trial of various phases of Bootstrap have Kathi Fisler and Shriram Krishnamurthi, computer science professors at Brown University and co-founders of Bootstrap, as co-authors.

[146] https://www.bootstrapworld.org/index.shtml

[147] https://www.bootstrapworld.org/

[148] https://www.bootstrapworld.org/impact/EvalReportDec2015Medium.pdf (page 6)

[149] https://www.bootstrapworld.org/materials/algebra/Bootstrap-Algebra.pdf

connection between algebraic thinking and programming abstractions, anticipating that "functional languages reinforce – rather than undermine – the math that students already need to learn in school." [150] [151] Once students have built their initial understanding, they can move on to other languages in other Bootstrap modules (see below).

The Bootstrap curriculum is also designed to be flexible and helpful to teachers: students do not need to spend all 25 hours in a computer lab, the materials are designed in a modular way to aid organisation and delivery. Work is aligned to US State curricula and standards to encourage teacher buy-in, and professional development and the programme also runs workshops and professional development. This flexibility extends to assessment: using Bootstrap does not necessitate any particular model of assessment, allowing teachers to attend to the demands of their own context.

Expansion and Bootstrap: Data Science

The Bootstrap provision has been extended to include further modules which integrate computing into science lessons (Bootstrap:Physics) and social science lessons (Bootstrap:Data Science). There is also a fourth course, Bootstrap:Reactive, which focuses on developing more advanced programming skills.

The materials are presented in a highly flexible way, such that a teacher can use them to support anything from a one-week to a one-year course. [152]

Bootstrap:Data Science is a relatively new module which, like Bootstrap:Algebra, is designed for embedded or integrated delivery. The course authors have detailed the thinking behind their curriculum design in some detail [153], proposing that a "strong and equitable data science curriculum combines four ingredients". These are listed as: domains of study, statistics and mathematics, computing and civic responsibility. Instead of developing a computer game, students combine data science content (such as linear regression) with computing content (such as Boolean logic) to pose and answer questions; the course

---

[150] https://www.bootstrapworld.org/blog/software/FunctionalLanguages.shtml
[151] The effectiveness of this approach is examined in Wright, G., Rich, P. & Lee, R. (2013). The Influence of Teaching Programming on Learning Mathematics. In R. McBride & M. Searson (Eds.), *Proceedings of SITE 2013--Society for Information Technology & Teacher Education International Conference* (pp. 4612-4615). New Orleans, Louisiana, United States: Association for the Advancement of Computing in Education (AACE). Retrieved April 5, 2023 from https://www.learntechlib.org/primary/p/48851/.
[152] https://www.bootstrapworld.org/materials/data-science/
[153] https://cs.brown.edu/~sk/Publications/Papers/Published/spddplfk-integ-ds-desn-assm-curric/paper.pdf

materials provide 26 large data sets drawing on contexts such as politics, sport and the environment. Some aspects of data science, such as data cleaning and working with newer forms of data visualisation are not included in the course due to space constraints.

Equity, Access and Variability

The Bootstrap materials appear to be used widely and with thousands of students across the United States. It offers a proof of concept for integrating some elements of computer science into mathematics (and other such integrations) in a way that supports both domains. Further, the embedded design approach may be helping to promote equity and support access by bypassing bottleneck issues associated with a deficit of specialist computer science teachers. Nonetheless, it is relevant to note that the small sample sizes in the main evaluation document (where the participants could skew towards enthusiastic adopters) [154] make it hard to quantify some aspects of its success or reflect meaningfully on how well similar work might translate to the UK context.

Benefits

- The courses have been developed to appeal to and support teachers, particularly those with limited computer science experience. This has had benefits in terms of equity and access, and positively impacted the scalability of the programme.
- The Bootstrap materials seek to integrate knowledge and approaches from each involved discipline in an authentic way which has mutual benefit. This has extended to the combined pedagogy, for instance when error messages while programming are seen as opportunities for a student to test their understanding of abstract structure, rather than simply as mistakes or typos. [155]
- Meaningful and accessible large data sets connect statistics and programming with domains of interest and with civic responsibility.

Barriers

- The goal of establishing clearer links between subject areas could limit the included content from each discipline and impose design choices (for instance, some computer scientists might prefer programming languages which were more widely used over functional programming options).
- Some presentations of the courses rely on a community of volunteers (college students, specialist teachers, and software professionals) which cannot be assumed.

---

[154] https://www.bootstrapworld.org/impact/EvalReportDec2015Medium.pdf
[155] https://bootstrapworld.org/blog/software/Errors.shtml

- The flexibility built into the Bootstrap model might not transfer comfortably to a context with more constraints concerning curriculum and assessment.

Adaptations

- As more school districts and teachers are coming on board, Bootstrap can increase their pool of 'ambassadors' and professional development offerings.
- There is an increased focus on evaluation and iterative improvement of the materials.
- The Bootstrap team have been developing adaptations for students with visual impairments or other accessibility issues.

8.10 Urban Data School

Data literacy is often taught in schools using small, personally collected data which is easier for students to relate to. This does not give the pupils the experience of critically querying and deriving insights from larger, complex data that is collected at scale.

The Urban Data School study trialled a method for teaching from complex data, collected through a smart city project. It showed that existing data principles from the literature can be adapted to design data literacy activities that help pupils understand complex data collected by others and form interesting questions and hypotheses about it.

The study was carried out over two years in three primary and secondary schools in Milton Keynes, using datasets from the MK:smart project. The lesson plans were co-created with three teachers (maths, geography and triple science) and used with 67 students aged between 10–14 years, across 4 separate sessions, each of 2-3 lessons over a few weeks. It is a small-scale, longitudinal example of our Model 3, adding content across subjects).

Data literacy activity design principles

The study authors synthesised from the literature the following principles for designing data literacy activities.

P1 Inquiry principle: Lead the students first in a guided inquiry, followed by an open inquiry when they are more familiar with the data and the approach.

P2 Expansion principle: Start from a representative small part of the data set and expand out, rather than starting with the full, large data set and focusing in. This helps students orient themselves within the data, before navigating across it.

P3 Context principle: Teach in a context the student understands, using data from their own environment or somehow related to them.

P4 Foundational competences principle: Focus on developing foundational competencies rather than practical skills, for example how to ask 'good' scientific questions from data.

P5 STEAM principle: Take a STEAM approach by working collaboratively on creative activities alongside practical ones.

P6 Personal data collection principle: Students should engage with data they collected themselves. When analysing an external data set, they should be given additional activities

that help them understand what it is like to collect that type of data. This is to support them in contextualising and interpreting the external data.

The main change to existing literature is in P6: personal data collection can complement (rather than replace) interpretation of existing data.

Note that the principle '**P4 Foundational competences principle**: Focus on developing foundational competencies rather than practical skills' is not demoting the role of skills, but is intended to clarify that the skills build on and need the foundational competencies. All students should obtain the foundational competences of interpreting, critiquing and asking questions of data. This is especially important if the data set is not available, as in adverts or media reports. Data-literate citizens with a critical mind will more easily spot fake news, unwarranted conclusions and other ways of manipulating public opinion or prospective customers with data. The ability to ask relevant and feasible questions builds the foundation for the choice of the data cleaning, analysis and presentation methods and tools. Without the foundational competences, the practical skills (like applying statistical and visualisation methods) may be misguided and thus contribute to misinformation and erroneous conclusions.

Activities

The three main data literacy activities are described next. Their goal was to develop the students' skills for asking questions that they were personally interested in from the data. Therefore, all activities followed principles P1 and P4. Any additional principles followed are given in the individual activity descriptions. The first two activities were each done with and without computers, in different sessions.

Smart meter energy data: The students were shown anonymised data about whole house consumption, individual appliance consumption and generation of solar energy from ~70 houses in Milton Keynes (P3). Students were first given data about a single household over a single day and then for more households and over a time period (P2). Students were shown data on paper and using an interactive app.

Potential for solar energy production: The students were shown aerial photographs and an interactive Google-like aerial map of Milton Keynes, including of their school and neighbourhood (P3), with information about the potential annual energy yield if solar panels were installed on roofs, based on their direction, pitch and size.

Be a LiDAR device: This lesson followed the previous lesson to explain the LiDAR technology used to obtain the rooftop dataset. Students worked in groups to build plasticine houses

(P5) and measure their height for each square centimetre (P6). Students swapped their data to figure out the shape of the other group's house from the data alone.

The activities drew on mathematical and digital literacies to help pupils better understand an externally sourced complex dataset. Doing their own measurements and using interactive visualisations led pupils to more critical questions and insights, like whether the solar panel yield takes into account nearby obstacles (e.g. tall trees) and if smart metres are always measuring correctly. The design principles were effective in helping students make sense of the data and of the wider topic (energy consumption and generation).

Benefits

- Interactive data visualisations can support school children in asking valid questions from data.
- Snapshots of the data can help school children understand how to frame questions around the extended dataset, along dimensions such as space and time.
- The learning of data skills lends itself to cross-curricular learning and can begin with primary school students.
- When engaging with externally sourced data it can be useful to act in the role of a data collector to understand better where errors can creep into the data and to develop better data scepticism.

Barriers

- Even though the teachers co-created and co-delivered the lessons, they didn't feel confident in doing such activities by themselves after the project finished.
- Bespoke interactive visualisations may be needed, with simple user interfaces.

Adaptations

- The activities were refined and adapted by each teacher to their classroom and context.
- While many open data sets exist, they require curation for the local context of pupils.

9. Overview of case studies

Each of the case studies in Section 8 offers a particular perspective on how mathematical and data literacy education - or at least certain aspects of it - might be conceptualised, organised, and delivered. Although each account is necessarily a product of its own context, it offers an example of what is possible, and collectively the studies point towards several barriers and adaptations which are relevant to this analysis. The following discussion collects some comparisons and commonalities which speak to form and feasibility.

Through the discussion, we evaluate the effect of each curriculum model on the following criteria. Judgements are qualitative, based on our analysis of the case studies and the lessons learnt. The numerical scales are included to facilitate overall comparisons, and are not the result of any systematic quantitative process.

- Benefits for mathematical and data literacy: will the model allow the competencies to be achieved for all/many/specialist students, with the corresponding societal, personal and employment benefits? Scale +10 (allow) to -10 (prevent).

- Equitable access: will the model affect gaps in attainment or participation in STEM due to gender, social class or ethnicity? Scale +1 (reduce) to -1 (widen).

- Assessment: how extensive are the changes to assessment structure as compared to current provision? Scale 0 (none) to -2

- Teacher professional development: what extent of professional development is needed and how many teachers are affected? (It is relevant here that 6,500 secondary teachers taught some computing in 2021, compared to 16,500 for geography, 36,000 for mathematics, 50,000 for science and 222,000 primary and nursery teachers [156].) Scale 0 (none) to -5 (highest need)

- Teacher workforce: will this model result in secondary teachers teaching outside their specialist subject or need a non-trivial increase in the teaching workforce? Scale 0 (no effect) to -2

---

[156] DfE, School workforce in England 2021. https://explore-education-statistics.service.gov.uk/find-statistics/school-workforce-in-england#dataBlock-c26f37c7-2bca-4019-a55d-a26f58713754-tables

9.1 The case for change

The case studies repeatedly speak to the import and relevance of data-driven mathematics. Amongst other features, the advancing place of statistical/data knowledge and skills in the workforce is recognised within MEI's courses for students and the work with geospatial data in geography; the team behind Civic Statistics make a strong argument for data science as a critical component of contemporary citizenship; and the developing curricula and qualifications in Scotland, New Zealand and the IB offer examples of how statistics/ data science is already part of the international educational landscape. Together the cases make multiple, intersecting arguments for the deliberate inclusion of mathematical and data literacy in secondary education, and against keeping the current curriculum structure with no changes.

Failing to make the change is negative rather than neutral in its outcome because the demands of employers for T-shaped, broad and connected, knowledge [157] are already not met under the siloed nature of our current system. Moreover there are pedagogic implications in teaching mathematics content that is increasingly distant from its use. As technology increasingly disguises societal practices such as making change or using paper maps, these are not available as experiences which students can mathematise[158]. Instead, informal, out-of-school discussions become critical in making connections between classroom mathematics and its actual use. This has effects on equity, as it has been shown that such experiences are not uniformly distributed and affected by social class and linguistic capital [159].   In addition, teaching mathematics for meaning-making and for connections to realistic uses of mathematics supports the participation and engagement all students, but especially girls [160]. Finally, there is an ongoing PD cost for new

[157] Kispeter, 'What Digital Skills Do Adults Need to Succeed in the Workplace Now and in the next 10 Years?'

[158] Jenefer Golding, 'Mathematics Education and Financial Literacy: A Valued but Vulnerable Symbiosis', in *Proceedings of the British Society for Research into Learning Mathematics 40 (1) March 2020*, ed. Rachel Marks, 2020, http://www.bsrlm.org.uk/publications/proceedings-of-day-conference/ip40-1/.

[159] Robyn Zevenbergen, 'Mathematics, Social Class, and Linguistic Capital: An Analysis of Mathematics Classroom Interactions', in *Sociocultural Research on Mathematics Education*, ed. Atweh, Bill, Helen Forgasz, and Ben Nebres (Mahwah: Lawrence Erlbaum Associates, 2001), 201–15.

[160] Cathy Smith and Jennie Golding, 'Schools' Strategies for Promoting Girls' Participation in Mathematics', in *Proceedings Of the 42nd Conference of the International Group for the Psychology of Mathematics Education (Vol 4) Umeå, Sweden.*, ed. Ewa Bergqvist et al., 2018, 211–18; Cathy Smith, 'Gender and Participation in Mathematics and Further Mathematics A Levels: A Literature Review for the Further Mathematics Support Programme', 2014, https://core.ac.uk/download/pdf/79498409.pdf.

teachers who must understand, justify and motivate applying mathematics in inauthentic, outdated contexts in order to teach the curriculum.

Model 1 No change	
Benefits for mathematical and data literacy	-5
Equity	-1
Teacher Professional Development (Maths teachers only)	0
No effect on Assessment or Teacher Workforce	0

9.2 Promoting data skills through computing and across subjects

We have found no examples of data skills being added to computing curricula. Rather the opposite, since coding has recently been added to primary and early secondary mathematics curricula in Ontario and also in Sweden [161]. Several case studies take a proactive role in demonstrating the role of technology within mathematical and data literacy, including the Scottish NPA model and ProDaBi. Both projects aim for flexible models of delivery but have been crucially supported by computing teachers. As such they can inform Model 2, that is retaining the current structure for KS 3 and KS 4 with content added to Computing. ProDaBi was particularly notable for aiming to allow upper-primary aged students to explore and reflect upon some basic elements of machine learning through activity and analogy. Scotland's National Professional Awards in data science sit against a background of related academic and vocational qualifications for over-15s.

The fit with computer science does have problems. GCSE Computer Science is optional so adding to that content would not deliver the benefits for all, or even many. Through our reading and conversations, we encountered different opinions about whether students should be provided with software that automated key processes or (eventually) be required to develop their coding skills. Even without programming, any expansion of content in Computing as a foundation subject would require more time in KS 3 and, problematically, at KS 4 where the existing demands are already enacted inconsistently, in part because of

---

[161] C. Kilhamn et al., 'Variables in Early Algebra: Exploring Didactic Potentials in Programming Activities.', *ZDM - Mathematics Education* 54, no. 6 (November 2022): 1273–88; Susanne Kjällander et al., 'Elementary Students' First Approach to Computational Thinking and Programming.', *Education Sciences* 11, no. 2 (February 2021): 80.

the low profile of non-assessed subjects. Thus this model without any change to assessment and curriculum structure would not achieve the full aims.

The ProDaBi example suggests that computing teachers may well require less extensive PD than teachers of other subjects to start to teach data skills, and there are fewer of them. However there is already a shortage of qualified computing teachers, with only 53% of lessons being taught by a teacher with a relevant post A level qualification [149]; extending the subject time would further affect supply.

Teaching data science with a heavy emphasis on computing also invokes issues of equity and access. Not all students have equal access to technology or previous experience with technology, such that an overreliance on computing as an entry point could form a barrier to some students engaging with other, less technical, data skills. The MEI case study shows that positioning data science closer to mathematics and statistics than computing can reduce gender gaps in participation, with 37% of their students being female [162] compared to 15% of A level computing candidates in summer 2022 [163].

Model 2 Current Structure for KS 3 and KS 4 with content added to Computing	
Benefits for mathematical and data literacy	+4
Equity	-1
Teacher Professional Development (Computer teachers only)	-0.5
Assessment	0
Teacher Workforce	-2

The example of ProDaBi also touches on some wider issues. First, it suggests the value of teaching data science over time; the higher level ProDaBi project is quite ambitious, but it relies on previous exposure. A supporting comparison can be drawn with the New Zealand curriculum where (amongst other skills) students develop their statistical literacy from primary school, asking more sophisticated questions and gradually developing higher-order skills of critique and evaluation. Scotland is likewise working to develop the teaching of data science in primary school. Introducing and developing data science sensitivities over time advances both the potential depth and width of provision within later Key Stages.

---

[162] https://mei.org.uk/app/uploads/2021/08/mei-ds-short-course-pilot-report.pdf
[163] https://www.jcq.org.uk/wp-content/uploads/2022/08/A-and-AS-Results-Summer-2022-Compared-to-2021.pdf

The Bootstrap programme from the United States offers something of a counterpoint; through careful curriculum and resource design it aspires to empower non-specialists to teach computer science in conjunction with their own disciplines, including mathematics, physics and social science, thus working towards its goals of equity, scale and rigour. The provision is for c. 25-hour courses, with associated teacher PD, that can be dropped into existing curricula flexibly as suits teachers or school boards. This is enticing, and it speaks to the import and power of proper teacher support, but it is unclear how well this approach might generalise to more extensive slices of subject curricula, or to the contexts such as England where curriculum time and assessment are more standardised.

Mathematical and data literacy's central activity of posing and answering questions in context naturally suggests the potential for cross-curricular work (which is an advantage of the Intersectional framework proposed in Section 5 of this paper). The merits of this approach are recognised not only in ProDaBi, Scotland's NPA and Bootstrap case studies but also in Civic Statistics, Urban Data School and the PD collaboration between Geography and Core Maths. The approach laid out in the Civic Statistics case study foregrounds citizenship elements and data literacy with a global perspective. 'Drop-in' materials for teachers cover a wide range of topics such as income inequality, migration, and ageing societies that appeal to many stakeholders and allow informed discussion of equality, diversity and inclusion issues. The Urban Data School, aimed at primary/ lower secondary students, captured similar principles of engaging students through inquiry into data from their own environment and included experiences of collecting such data though smart city technologies.

While there is certainly an opportunist element of innovative education projects taking any available teaching space, this consensus suggests that adding content across subjects provides a viable way of achieving the benefits of mathematical and data literacy. The change also affects the teacher workforce equally, with no specific demands on shortage subject teachers. It is particularly welcomed as a way of introducing mathematical and data literacy to those with mathematics anxiety, and of broadening its appeal to girls through enhancing meaning-making [160]. Moreover, the example of Geography shows that A level teachers in England have, over time, and when incited by relevant drivers such as assessment, taken advantage of professional development opportunities to develop students' mathematical, statistical and technological skills that are authentically useful in their own subject.

KS 3 and 4 subjects (apart from mathematics) that were repeatedly suggested as particularly suitable for enhanced data skills were sciences, geography and religious education. At KS 5 this is broadened to include psychology, sociology and politics. We have

seen no trials of embedding in the arts or languages. This does expose a weakness of any model that embeds mathematical and data literacy across several subjects but retains current free subject choice at 14 or 16 (such as Model 6): there is no guarantee which combination of subjects individuals will study, and thus there would be no mechanism for ensuring all had met the desired competencies, reducing the equity advantage. Further investigation is needed to understand how/whether data literacy could be embedded in a wider range of subjects. In addition, changes to assessment would be needed, since the pressure of high-stakes examinations means that teachers would be likely to omit any non-assessed content in the subjects chosen.

Model 6 Current KS 5 Structure with content added across subjects	
Benefits for mathematical and data literacy	+7
Equity	+0.5
Teacher Professional Development (in Humanities, Science + ?)	−2.5
Assessment	−0.5
Teacher Workforce	0

Model 3, of enhancing some subjects at KS 3 and 4 (while leaving KS 5 unchanged) raises further issues of progression. The case studies suggest that compromises are inevitable – even desirable – in developing resources that promote competencies across domains. Students' experiences in Geography, for example, would not be the same as those in RE. A potentially uneven pattern of experience at KS 4 would not provide a reliable foundation for specialists to build their skills during KS 5. (A similar effect was noted in the Geography case study, where changes to KS 3 computing mean that GCSE teachers cannot rely on students' spreadsheet skills.) Reteaching could be needed, reducing curriculum time, or this could in effect push existing inequities arising from subject choice at 16 down to age 14.

Model 3 Current Structure for KS 3 and KS 4 with content added across subjects	
Benefits for mathematical and data literacy	+8
Equity	+0.5
Teacher Professional Development (in Humanities; Science + ?)	−3
Assessment	−0.5
Teacher Workforce	0

Two issues which came up in conversation with experts during the writing of the case studies was how teachers in other subjects might be supported in bringing data science into their lessons, and how cross-curricular approaches to mathematical and data literacy might be aligned when subjects are siloed. One suggestion was that schools might designate a 'data science coordinator', although it was also noted that previous efforts to have 'statistics coordinators' in UK schools had not been fruitful.

9.3 Introducing strands within mathematics

Several case studies site mathematical and data literacy within the mathematics curriculum. Among these, Ontario is unusual in that their approach to teaching or assessing data-driven problems is not marked or ring-fenced. In Ontario, most content strands are roughly balanced between expectations of developing a new mathematical skill and those of applying it in domain contexts, many of which are specified in the curriculum documents. Thus (although it extends to KS 5) it speaks to Model 4a of adding content to the compulsory mathematics curriculum at KS 3 and Foundation level GCSE. In England's current framework this is the only way of ensuring that students meet the competencies for all. With care, Higher level GCSE could be tailored to address the competencies for many. Without further change at KS 5, this model would not embed the specialist skills.

One (flawed) argument for Model 4a is that it involves minimal disruption: teaching with attention to data and applications is not precluded in the current curriculum, indeed is already some teachers' practice. Schools have moved towards spending two (rather than 3) years on KS 3, suggesting some space in that KS. However, this ignores the way that high-stakes assessment drives teachers to frame students' opportunities to learn [164]. GCSE Mathematics questions are set in a mathematical context [165] or in a simple context involving units specified in the curriculum (e.g. money, time, length, mass, litres, speed, density), and teaching is largely restricted to those. Embedding this extra content would involve teacher PD and increased curriculum time, with a subsequent effect on workforce capacity.

---

[164] ACME, 'Problem Solving in Mathematics: Realising the Vision through Better Assessment', 2016, https://royalsociety.org/~/media/policy/Publications/2016/problem-solving-in-mathematics-06-2016.pdf.

[165] Candia Morgan and Sarah Tang, 'To What Extent Are Students Expected to Participate in Specialised Mathematical Discourse? Change over Time in School Mathematics in England', *Research in Mathematics Education* 18, no. 2 (3 May 2016): 142–64, https://doi.org/10.1080/14794802.2016.1174145.

A further consideration from Ontario is that assessment of students is done by teachers, following a national reporting system. Assessment is able to follow what is taught, rather than drive it. 70% of the assessment must occur during the course and 30% at the end, drawing on teacher-designed tests and tasks. Provincial tests at grades 3, 6 and 9 separately monitor school performance against expectations. This system leaves a lot of responsibility to teachers in a very respected education ecosystem. The 10-year curriculum design cycle acts to influence the enacted curriculum through PD and through some internal structure; for example the 2020 move to designate modelling as a 'strand' in algebra is to ensure that it is assessed.

Model 4a Current Structure for KS 3 and KS 4 with content added to a Mathematics course	
Benefits for mathematical and data literacy	+5
Equity	0
Teacher Professional Development (in Maths)	-2
Assessment	-0.5
Teacher Workforce	-0.5

The New Zealand case study takes a stronger approach of ensuring attention to data literacy within mathematics by creating curriculum strands (or pillars). It draws a distinction between equally-weighted 'Mathematics' and 'Statistics' strands both of which must be taught and assessed throughout the primary and secondary curricula until the final, optional pre-university year of school when students may choose one. New Zealand is thus similar to Model 4b in which assessment is extended to two GCSEs (although NZ strands extend beyond KS 4). Previous arguments for a two-pillar or two-subject approach to teaching mathematics have often separated applications of mathematics from developing new mathematical tools (e.g. Use of Maths [166] or Essential Maths [167]). In the case of New Zealand, both strands involve applications to real-word contexts. There was considerable PD involved in educating mathematics teachers about pedagogy for statistics, and this was supported by extensive pre-trialling of curriculum materials and sample tasks. Much assessment is teacher-designed with some post-16 students able to

---

[166] Paul Hernandez-Martinez et al., 'Students' Views on Their Transition from School to College Mathematics: Rethinking "Transition" as an Issue of Identity', *Research in Mathematics Education* 13, no. 2 (2011): 119–30, https://doi.org/10.1080/14794802.2011.585824; Andy Noyes et al., 'Evaluating Mathematics Pathways' (Department of Education, 2011).
[167] MEI, 'Curriculum and Assessment in Mathematics in England at KS4 and KS5 - Position Paper'.

choose between studying 'standards' (of which three must involve numeracy) that have external assessments and those that are internally moderated.

This approach has been successful in increasing fidelity to the intended curriculum, and a recent expert panel review [168] has suggested it be retained, indicating its strength. There are concerns about inequitable attainment patterns in New Zealand, but the same review considers those to result from school differences and not to curriculum structure.

Again, without further change at KS 5, this model's benefits are limited since it would not embed the specialist skills.

Model 4b Additional content at KS 3 and 4 creates two parallel, or sequential, Mathematics GCSE courses	
Benefits for mathematical and data literacy	+8
Equity	+0.5
Teacher Professional Development (in Maths)	-2
Assessment	-1.5
Teacher Workforce	-1

9.4 Introducing pathways within mathematics

The approach taken at KS 5 in several of the case studies is to introduce pathways at 16-18. New Zealand and Ontario both set a minimum requirement of mathematical credits to obtains school graduation certificates and provide a range of courses to suit students' interests and needs. The International Baccalaureate (IB) requires students to choose between two pathways through mathematics, which can loosely be thought of as 'pure' and 'applied' mathematics routes. Their overall approach is seen as effective in providing benefits for all, through the minimum requirements, and in allowing specialisation.

In Ontario, the pathways are described in terms of preparing for school leavers' next destinations: the workplace, college or university. Within college preparation, a distinction is made between technology-related programmes and others. Within university preparation, the distinction is between degree majors needing mainly statistical data management

---

[168] Royal Society Te Apārangi Expert Panel, 'Pāngarau Mathematics and Tauanga Statistics in Aotearoa New Zealand. Advice on Refreshing the English-Medium Mathematics and Statistics Learning Area of the New Zealand Curriculum' (Royal Society Te Apārangi, 2021).

such as social sciences (which has been popular among girls [169]); those needing knowledge of functions such as business, health science and those also needing calculus such as STEM or economics. Online courses are provided to mitigate problems of accessing teaching. Pathways are overlapping, but it is recognised that early choice has equity issues and the entry point has recently been moved up by de-streaming grade 9.

In New Zealand, students have extensive choice in selecting standards to study, which may in practice be restricted by schools. The present curriculum supports higher level study, encouraging and preparing large numbers of students to choose a first-year statistics course at university. There is an accompanying concern that the emphasis on statistics may distract potential STEM undergraduates from specialising in calculus standards.

The two-pillar approach within IB mathematics has been well received by many and has turned out to be a useful way of grouping content for many diploma programme and career programme students. Students choose from two strands of mathematics, known as Analysis and Approaches (AA) and Applications and Interpretation (AI), each of which is offered at either SL or HL. There is core common content (including some calculus) ensuring that all four forms of IB Mathematics meet certain standards allowing university-level study. The AI versions of the curriculum could resonate with the Intersectional competencies suggested for all and for many school leavers (with little, however from the computational toolkits). All four courses include a non-examined but assessed project based on asking and answering questions.

These case studies show that pathways are an effective structure for ensuring an appropriate range of mathematical and data literacy competencies, and also that they have the flexibility to accommodate regular updates as mathematical needs evolve, speaking to Models 7 and 8.

For Model 7, adapting the current A/AS level and Core Maths structure, the issue recurs that there is no minimum mathematics requirement, and indeed there is currently little enthusiasm for compulsory Core Maths [170]. However, the case studies reveal no dissent once established as part of a baccalaureate system with minimum requirements for numeracy and literacy (as in the IB and as being developed pre-16 for the Welsh

---

[169] Jennifer Hall, 'Gender Issues in Mathematics: An Ontario Perspective', *Journal of Teaching and Learning* 8, no. 1 (5 January 2012), https://doi.org/10.22329/jtl.v8i1.3004.
[170] Andrew Noyes and Michael Adkins, 'Studying Advanced Mathematics in England: Findings from a Survey of Student Choices and Attitudes', *Research in Mathematics Education* 18, no. 3 (1 September 2016): 231–48, https://doi.org/10.1080/14794802.2016.1188139.

Baccalaureate [171]). The Core Maths structure does allow further specialisation that could include the statistics/data science and computation skills that are not currently included. In addition, increasing the numbers studying mathematics would have a large impact on teacher workforce. Relevantly, the RGS/Core Maths case study has started to develop non-mathematics teachers towards teaching current Core Maths although uptake is not known. The examination-based assessment remains problematic, given that all relevant case studies found the need for project-work to make valid assessments.

Model 7 Current KS 5 Structure with additional content creating choice of mathematics pathways (examined as A/AS level or Core Maths)	
Benefits for mathematical and data literacy	+7
Equity	0
Teacher Professional Development (Maths + others)	-2
Assessment	-0.5
Teacher Workforce	-0.5

Changing the curriculum so as to make not only a space, but an entitlement, for developing mathematical and data literacy is ambitious. Models 1-7 have all accepted the constraints of the current system where, after KS 3, students can choose not to study some subjects. The rarity of this has been commented on in mathematics [172], but applies even more to subjects where data-driven questions might be asked and answered.   This feature repeatedly means that England cannot follow other national curricula in planning minimum cross-curricular entitlements for all.

It is beyond the scope of this paper to investigate different baccalaureate options. Some, such as New Zealand, have an implicit baccalaureate where students are required to take a range of options with certain limiting requirements but a great deal of choice; others, such as the International Baccalaureate are more formal, based around a template where students choose within subject areas and meet some synoptic requirements. It is also worth recalling once more the new Curriculum for Wales where a coherent and broad approach to education is established throughout secondary school. Each of these options would afford the benefits of embedding mathematical and data literacy, either through a

---

[171] https://hwb.gov.wales/curriculum-for-wales

[172] Jeremy Hodgen, *Is the UK an Outlier?: An International Comparison of Upper Secondary Mathematics Education* (London: Nuffield Foundation, 2010).

coordinated, subject-based approach, or via a synoptic element. This is discussed further in the next section.

Model 8 A new baccalaureate system at KS 5 and below (requiring study in core subjects)	
Benefits for mathematical and data literacy	+10
Equity	+1
Teacher Professional Development	-2
Assessment	-2
Teacher Workforce	-1

These case studies indicate that any move to a pillars pathways system would need to be thought through carefully, balancing foci and attending judiciously to the needs of students, higher education, and employers. Any partition of mathematics will inevitably be imposed and imperfect, and it is entirely possible that incautious labelling may lead to content being neglected or even lost; the New Zealand case study reminds us for example that statistical modelling is distinct from, and does not supersede, mathematical modelling. The IB case further warns that when multiple versions of mathematics are on offer at the same level, it is natural for some stakeholders to perceive one as 'easier' than the other, and without handling and pre-emption this will affect provision and uptake.

9.5   Pitfalls and possibilities of mathematical and data literacy as a separate curriculum area

At first glance, two of the case studies successfully deliver data science content as an adjunct to other curricular provision. The MEI courses for A level and Core Mathematics students can be considered as enrichment, and the Scottish NPA in Data Science has been established independently of any other qualifications. Arguably, this approach has led to fewer restrictions on the content of these courses and awards them greater freedom to develop and evolve over time.

However, both cases implicate some significant feasibility issues surrounding scaling, equity and access. One of the strengths of the MEI courses is the presence of specialists with a deep knowledge of, and appreciation for data science; this would be hard to scale up at present. Equally, whilst the Scottish NPA offers a considered and detailed approach to teaching data science, it exists as one of many possible qualifications, and relies on the

presence of willing and capable staff for in-school delivery. If data science is positioned only as a separate and specialised discipline, it is likely that current workforce shortages and individuals' curriculum choices will mean that fewer students will have the opportunity to develop data science competencies.

A related but contrasting approach would be to introduce data science via a project-based component that has its own presence within the curriculum but is explicitly taught and positioned with reference to other curriculum areas. This model is not precisely exemplified within the case studies, but it can be illuminated by some comparisons with other examples of project-based qualification and interdisciplinary study.

One of the most common project-based qualifications in the UK at the moment is the Extended Project Qualification and a Level 3 data science project could exist in a similar way, requiring students to use their independent study and research skills to carry out some data-focused work for themselves. However, whilst such an approach might capture well the investigative nature of data science, an optional qualification would not reach all students, and it would rely on a level of knowledge of and experience with data science that Key Stage 5 students do not typically possess at present. There are also practical concerns surrounding independent work: MEI, for example MEI, note the difficulty for experienced *teachers* of finding large open access data sets suitable for 17-year olds.

An alternative approach would be to combine a project-based output with taught lessons, and to make it compulsory for all students. This is resonant with the delivery of Theory of Knowledge (TOK) within the International Baccalaureate diploma programme. Students have specific TOK lessons where they meet key ideas and are assessed via an exhibition and a short essay. There is also an expectation that subject teachers support the delivery of TOK by making explicit connections to ideas from the course. TOK plays a meaningful but bounded role within the wider diploma; compared to the main components it involves fewer curriculum hours and has fewer points attached to it. Teachers typically deliver TOK in addition to their regular subject specialism, reducing workload effects.

Model 5 Current KS 5 structure with the addition of credit-bearing project-based work.	
Benefits for mathematical and data literacy	+7.5
Equity	+0.5
Teacher Professional Development	-2
Assessment	-0.5
Teacher Workforce	-0.5

It is possible to map mathematical and data literacy into this model: for instance, students could meet new statistics/data science or computational content in two focused lessons each week, be supported in making connections with their other areas of study, and then complete an individual project which could play to their strengths and interests. This could be done discretely, or as a required component within a baccalaureate model, possibly even as a smaller component of a wider citizenship requirement. A flexible, considered and responsive form of this model could well serve many students on multiple trajectories, and the potential benefits for developing mathematical and data literacy are substantial. Still, quality delivery of such an approach would require teacher professional development, (just as TOK teachers benefit from focused training and support,) and the quality and depth of the output would be enhanced further if it was supported by the integration of data science content and sensitivities into earlier key stages.

9.6   The challenge and potential of assessment

One element of provision that runs through the case studies and cuts across the suggested models is assessment. Formal assessment practices not only codify expectations and standards, but also have the potential to shape teachers' impressions of what is truly important within subject areas, and to steer classroom pedagogy.

Two of the main issues to consider in the context of this report are what place teacher assessment might have, and the benefits and costs of project work as compared to examinations.

The case studies illustrate an extensive range of assessment models, from 100% teacher assessed work (Ontario) to a fully centralised assessment system, albeit on a small scale (MEI). A mixed approach seems to be particularly common (for instance New Zealand) with elements of support and moderation built in as considered necessary (as with the IB). It is of note that the Scottish NPA allows teachers to be involved in writing assessments as well as marking them; this presents an opportunity for teachers to attend to their contexts and perhaps address diversity and inclusion issues.

The majority of the cases discussed also appeared to endorse a balance between project work and written examinations; indeed, in the case of Ontario assessment could be fully project-based. It is not difficult to argue that students' facility with the linked, context-based and empirical tools of data science might be better tested through a report or project, rather than a formal written examination, and the work of the IB has the value and validity

of adopting this approach across mathematics. However, written examinations can reduce workload demands on teachers and offer some degree of standardisation, so a balance might seem optimal. Too many avenues for assessment can complicate matters, as seen in the case of New Zealand.

There were also differing approaches regarding the award and aggregation of marks; the MEI short course awards a certificate at either fail, pass, merit or distinction, whilst the IB IA mathematics project is marked out of 20 and then contributes 20% to the overall mark for the mathematics component. The appropriateness of different scoring systems depends on the model of provision they exist within; a pass/fail criterion might suit a civic statistics project instituted as a compulsory component of a baccalaureate style qualification, but it would not work well within the current A level system. Nonetheless, how assessment is graded can impact on its reliability, as well as influencing teacher workload.

9.8   The important role of integrative, focused and ongoing professional development

Another recurring aspect of the case studies was the importance of supporting teachers in delivering high quality content. The present shortage of teachers in both mathematics and computer science impacts heavily on the feasibility of many of the proposed options for curriculum development. Whilst approaches such as Ontario's online courses and initiatives such as the AMSP/RGS training go some way towards ameliorating these deficits, it is important to recognise that many teachers may not be mathematically confident – or may even be mathematically anxious – so the scope of cross-curricular training and delivery may be limited.

Teachers are at the pupil-facing side of any curriculum changes, and can be involved and supported at various stages of curriculum development. This was observed within the case studies, for example where:

- Teachers were involved in the creation and testing of resources before the statistics curriculum was introduced in New Zealand;
- Teachers can attend ongoing professional development and training to address concerns and shortfalls in their knowledge and previous experience, such as is provided by the IB or the AMSP/RGS collaboration;
- Teachers have access to online resource collections which exemplify and support good practice, such as is available for mathematics teaching in Ontario, or for the Scottish NPA.

The importance of timely, high-quality professional development and support for teachers was also a recurring theme in the accompanying conversations.

9.8   Feasibility overview of models

The preceding discussion has provided qualitative judgements on the feasibility of each model.   Figure facilitates an overall comparison, of the potential of each model for achieving the benefits of embedding mathematical and data literacy, together with the effects on equitable access, and the impact on teacher PD, assessment and teacher workforce capacity.

Figure 8 Feasibility for models
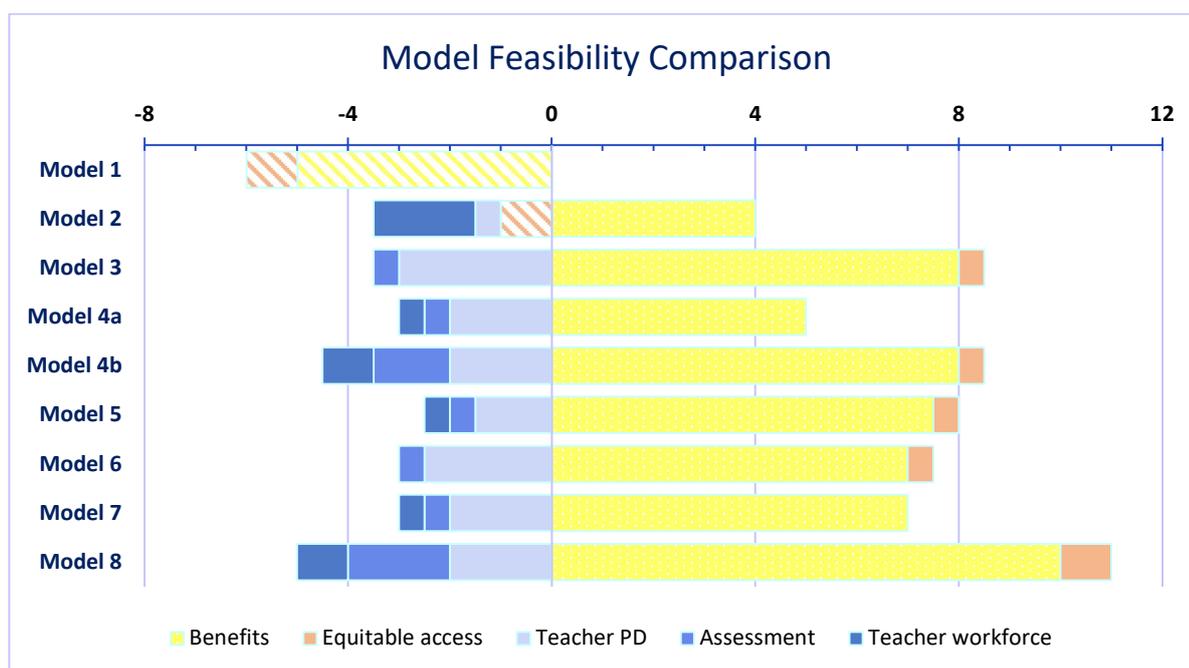
1	No change. Current structure in England	RS 1
2	Current Structure for KS 3 and KS 4 with content added to Computing	RS 2b
3	Current Structure for KS 3 and KS 4 with content added across subjects	RS 3 (GCSE)
4a	Current Structure for KS 3 and KS 4 with content added to Mathematics	RS 2a
4b	Additional content at KS 3 and 4 creates two parallel or sequential Mathematics GCSE courses	RS 4 (GCSE)
5	Current KS 5 structure with the addition of credit-bearing project-based work e.g. EPQ.	RS 5
6	Current KS 5 Structure with content added across subjects	RS 3 (post-16)
7	Current KS 5 Structure with additional content creating choice of mathematics pathways (examined as A/AS level or Core Maths)	RS 4 (post-16)
8	A new baccalaureate system at KS 5 and below (requiring study in core subjects)	RS 6

9.9 The value of a joined-up approach

One final recurring theme throughout the case studies and the attendant conversations was how an approach involving multiple stakeholders and organisations can support curriculum design and delivery. Sometimes connections were temporary, for instance when teachers were provided with information about how universities were responding to changes to the IB Mathematics curriculum, so that they could have informed discussions with students and parents. Other actions were more permanent, such as in the case of the NPA in Scotland which have been developed so that they can be delivered either by schools or further education providers.

The cases of MEI and AMSP/RGS demonstrate how subject associations can support teachers and enhance student outcomes; the involvement of RGS in particular has a history which builds on previous, successful collaborative work [173]. The cases of New Zealand and Ontario both suggest an educational ecosystem where policy makers, researchers and practitioners work together. It may be that the present UK climate, where awarding bodies compete over examination entries, is less inclined towards collaboration.

Whilst all curriculum areas are developing, this is especially true of data science; neither the pedagogy nor the content is settled. Cyclical models of curriculum development (such as the seven-year review of the IB) go some way towards coping with these changes, but a broad collaborative approach could also help keep data literacy materials up to date, for instance by inviting media organisations to help develop and update material on civic statistics. Finally, a joined-up approach seems apt: by engaging and involving stakeholders and interested parties from across education and beyond, it helps to frame the problem of mathematics, statistics/data science and computing education against its full domain, and to carry out important work in an ethical space.

---

[173] Nuffield Foundation, 'Mathematics in A Level Assessments' (Nuffield Foundation, 2012), https://www.nuffieldfoundation.org/wp-content/uploads/2012/04/Mathematics-in-A-level-assessments-A-report-on-the-mathematical-content-of-A-level-assessments-in-Business-Studies-Computing-Economics-Geography-Psychology-and-Sociology.pdf.

10.　References

ACME. 'Problem Solving in Mathematics: Realising the Vision through Better Assessment', 2016. https://royalsociety.org/~/media/policy/Publications/2016/problem-solving-in-mathematics-06-2016.pdf.

ACME and the Royal Statistical Society (RSS). 'Embedding Statistics at A Level: A Report on Statistical Requirements and Assessment across A Level Courses in Biology, Business, Chemistry, Geography, Psychology and Sociology', 2015. http://www.acme-uk.org/media/32719/embeddingstatistics-at-a-level.pd.

AfDSP. 'Data Science Standards'. Alliance for Data Science Professionals, 2022. https://alliancefordatascienceprofessionals.co.uk//standards.

Bargagliotti, Anna, Christine Franklin, Pip Arnold, Rob Gould, Sheri Johnson, Leticia Perez, and Denise Spangler. 'Pre-K-12 Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report II'. American Statistical Association and National Council of Teachers of Mathematics., 2020. https://hdsr.mitpress.mit.edu/pub/cqncbp3l.

Biehler, Rolf, and Yannik Fleischer. 'Introducing Students to Machine Learning with Decision Trees Using CODAP and Jupyter Notebooks'. *Teaching Statistics* 43, no. S1 (July 2021). https://doi.org/10.1111/test.12279.

Bond, Philip. 'The Era Of Mathematics: An Independent Review of Knowledge Exchange in the Mathematical Sciences.' UKRI, 2018. https://www.ukri.org/publications/the-era-of-mathematics/.

Brochu, Pierre, Maria Chuy, Marie-Anne Deussing, Koffi Philippe Houme, and Council of Ministers of Education, Canada (CMEC). *Measuring up, Canadian Results of the OECD PISA Study: The Performance of Canada's Youth in Mathematics, Reading and Science: 2012, First Results for Canadians Aged 15*. Toronto: CMEC, 2014. http://site.ebrary.com/lib/celtitles/docDetail.action?docID=10831057.

Cheng, Jing, Jiansheng Bao, and Dianzhou Zhang. 'From "Two Basics", to "Four Basics" to "Core Mathematics Competencies" in Mainland China'. In *Beyond Shanghai and PISA: Cognitive and Non-Cognitive Competencies of Chinese Students in Mathematics*, edited by Binyan Xu, Yan Zhu, and Xiaoli Lu, 1–13. Research in Mathematics Education. Cham: Springer International Publishing, 2021. https://doi.org/10.1007/978-3-030-68157-9_1.

Chernoff, Egan J. 'The Canadian Math Wars'. *Canadian Journal of Science, Mathematics and Technology Education* 19, no. 1 (1 March 2019): 73–76. https://doi.org/10.1007/s42330-018-0037-9.

Cobb, George W. 'The Introductory Statistics Course: A Ptolemaic Curriculum?' *Technology Innovations in Statistics Education* 1, no. 1 (12 October 2007). https://doi.org/10.5070/T511000028.

'Course Descriptions and Prerequisites'. Accessed 23 March 2023. https://www.dcp.edu.gov.on.ca/en/course-descriptions-and-prerequisites/mathematics.

Davies, Neville, and Neil Sheldon. 'Teaching Statistics and Data Science in England's Schools'. *Teaching Statistics* 43, no. S1 (2021): S52–70. https://doi.org/10.1111/test.12276.

———. 'Teaching Statistics and Data Science in England's Schools'. *Teaching Statistics* 43 (2 June 2021): S52–70. https://doi.org/10.1111/test.12276.

Department for Education. 'Guide to the Post-16 Qualifications Landscape at Level 3 and below for 2025 and Beyond'. Department for Education, 2023. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1128144/Guide_to_the_post-16_qualifications_landscape_at_level_3_and_below_for_2025_and_beyond.pdf.

———. *Mathematics Programmes of Study: Key Stage 4. National Curriculum in England*, 2014. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/331882/KS4_maths_PoS_FINAL_170714.pdf.

———. 'Participation in Education, Training and Employment Age 16 to 18, Calendar Year 2021', 2021. https://explore-education-statistics.service.gov.uk/find-statistics/participation-in-education-and-training-and-employment.

Donoho, David. '50 Years of Data Science'. *Journal of Computational and Graphical Statistics* 26, no. 4 (2 October 2017): 745–66. https://doi.org/10.1080/10618600.2017.1384734.

Forbes, Sharleen. 'The Coming of Age of Statistics Education in New Zealand, and Its Influence Internationally'. *Journal of Statistics Education* 22, no. 2 (July 2014): 2. https://doi.org/10.1080/10691898.2014.11889699.

Geiger, Vince, Peter Galbraith, Mogens Niss, and Catherine Delzoppo. 'Developing a Task Design and Implementation Framework for Fostering Mathematical Modelling Competencies'. *Educational Studies in Mathematics* 109, no. 2 (February 2022): 313–36. https://doi.org/10.1007/s10649-021-10039-y.

Geraniou, Eirini, and Morten Misfeldt. 'The Mathematical Competencies Framework and Digital Technologies'. In *Mathematical Competencies in the Digital Era*, edited by Uffe Thomas Jankvist and Eirini Geraniou, 39–60. Mathematics Education in the Digital Era. Cham: Springer International Publishing, 2022. https://doi.org/10.1007/978-3-031-10141-0_3.

Golding, Jenefer. 'Mathematics Education and Financial Literacy: A Valued but Vulnerable Symbiosis'. In *Proceedings of the British Society for Research into Learning Mathematics 40 (1) March 2020*, edited by Rachel Marks, 2020. http://www.bsrlm.org.uk/publications/proceedings-of-day-conference/ip40-1/.

Golding, Jennie, Cathy Smith, and Mick J Blaylock. 'Learning to Teach Contextualized Problem-Solving in a Non-Calculus Mathematics Pathway'. *Teaching Mathematics and Its Applications: An International Journal of the IMA* 37, no. 2 (4 June 2018): 69–83. https://doi.org/10.1093/teamat/hry009.

Gould, Robert. 'Toward Data-Scientific Thinking'. *Teaching Statistics* 43, no. S1 (July 2021). https://doi.org/10.1111/test.12267.

Grover, Shuchi, and Roy Pea. 'Computational Thinking in K–12: A Review of the State of the Field'. *Educational Researcher* 42, no. 1 (1 January 2013): 38–43. https://doi.org/10.3102/0013189X12463051.

Guo, Meize, and Anne Ottenbreit-Leftwich. 'Exploring the K-12 Computer Science Curriculum Standards in the U.S'. In *Proceedings of the 15th Workshop on Primary and Secondary Computing Education*, 1–6. Virtual Event Germany: ACM, 2020. https://doi.org/10.1145/3421590.3421594.

Hall, Jennifer. 'Gender Issues in Mathematics: An Ontario Perspective'. *Journal of Teaching and Learning* 8, no. 1 (5 January 2012). https://doi.org/10.22329/jtl.v8i1.3004.

Hernandez-Martinez, Paul, Julian Williams, Laura Black, Pauline Davis, Maria Pampaka, and Geoff Wake. 'Students' Views on Their Transition from School to College Mathematics: Rethinking "Transition" as an Issue of Identity'. *Research in Mathematics Education* 13, no. 2 (2011): 119–30. https://doi.org/10.1080/14794802.2011.585824.

Hodgen, Jeremy. *Is the UK an Outlier?: An International Comparison of Upper Secondary Mathematics Education*. London: Nuffield Foundation, 2010.

Hong, Yanjun, and Uffe Thomas Jankvist. 'Comparing Two Mathematical Competencies Frameworks from China and Denmark'. *Asian Journal for Mathematics Education* 1, no. 4 (2022): 494–514. https://doi.org/10.1177/27527263221144193.

Hoyles, Celia, ed. *Improving Mathematics at Work: The Need for Techno-Mathematical Literacies*. 1st ed. New York, NY: Routledge, 2010.

Institute of Analytics. 'IoA Analyst Competency Framework', 2022. https://ioaglobal.org/wp-content/uploads/2022/06/The_IoA_Analyst_Competency_Framework.pdf.

International Data Science in Schools Project (IDSSP) Curriculum Team. 'Curriculum Frameworks for Introductory Data Science', 2019. https://www.idssp.org/pages/framework.html.

Jensen, Jens Hojgaard, and Uffe Thomas Jankvist. 'Disciplinary Competence Descriptions for External Use'. *Nordic Studies in Mathematics Education* 23, no. 2 (2018): 3–24.

Kilhamn, C., K. Bråting, O. Helenius, and J. Mason. 'Variables in Early Algebra: Exploring Didactic Potentials in Programming Activities.' *ZDM - Mathematics Education* 54, no. 6 (November 2022): 1273–88.

Kilpatrick, Jeremy. 'Competency Frameworks in Mathematics Education'. In *Encyclopedia of Mathematics Education*, edited by Stephen Lerman, 110–13. Cham: Springer International Publishing, 2020. https://doi.org/10.1007/978-3-030-15789-0_27.

Kispeter, Erika. 'What Digital Skills Do Adults Need to Succeed in the Workplace Now and in the next 10 Years?' Report. London: Department of Culture, Media and Sport, June 2018. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/807831/What_digital_skills_do_adults_need_to_succeed_in_the_workplace_now_and_in_the_next_10_years_.pdf.

Kjällander, Susanne, Linda Mannila, Anna Åkerfeldt, Fredrik Heintz, João Piedade, and Nuno Dorotea. 'Elementary Students' First Approach to Computational Thinking and Programming.' *Education Sciences* 11, no. 2 (February 2021): 80.

Kullberg, Angelika, Ulla Runesson Kempe, and Ference Marton. 'What Is Made Possible to Learn When Using the Variation Theory of Learning in Teaching Mathematics?' *ZDM* 49, no. 4 (1 August 2017): 559–69. https://doi.org/10.1007/s11858-017-0858-4.

Lazarus, Jill. 'Secondary Curriculum Over the Last Decade'. *Gazette - Ontario Association for Mathematics* 60, no. 3 (2022): 52–54.

MEI. 'Curriculum and Assessment in Mathematics in England at KS4 and KS5 - Position Paper'. Mathematics in Education and Industry, 2022. https://mei.org.uk/curriculum-and-assessment-in-mathematics-in-england-at-ks4-and-ks5-position-paper/.

Morgan, Candia, and Sarah Tang. 'To What Extent Are Students Expected to Participate in Specialised Mathematical Discourse? Change over Time in School Mathematics in England'. *Research in Mathematics Education* 18, no. 2 (3 May 2016): 142–64. https://doi.org/10.1080/14794802.2016.1174145.

Mou, Haizhen, and Michael M. Atkinson. 'Want to Improve Math Scores? An Empirical Assessment of Recent Policy Interventions in Canada'. *Canadian Public Policy / Analyse de Politiques* 46, no. 1 (2020): 107–24.

National Academies of Sciences, Engineering, and Medicine. *Data Science for Undergraduates: Opportunities and Options*. Washington, D.C.: National Academies Press, 2018. https://doi.org/10.17226/25104.

news.ontario.ca. 'Ontario Dedicating $60 Million for Renewed Math Strategy'. Accessed 23 March 2023. https://news.ontario.ca/en/release/36364/ontario-dedicating-60-million-for-renewed-math-strategy.

1. Niss, M., R. Bruder, N. Planas, R. Turner, and J. A. Villa-Ochoa, 'Conceptualisation of the Role of Competencies, Knowing and Knowledge in Mathematics Education Research', in G. Kaiser (ed.), Proceedings of the 13th International Congress on Mathematical Education, (Cham: Springer International Publishing, 2017), pp. 235–48

Niss, Mogens, and Tomas Højgaard. 'Mathematical Competencies Revisited'. *Educational Studies in Mathematics* 102, no. 1 (1 September 2019): 9–28. https://doi.org/10.1007/s10649-019-09903-9.

Noyes, Andrew, and Michael Adkins. 'Studying Advanced Mathematics in England: Findings from a Survey of Student Choices and Attitudes'. *Research in Mathematics Education* 18, no. 3 (1 September 2016): 231–48. https://doi.org/10.1080/14794802.2016.1188139.

Noyes, Andy, P. Drake, G. Wake, and Murphy. 'Evaluating Mathematics Pathways'. Department of Education, 2011.

Nuffield Foundation. 'Mathematics in A Level Assessments'. Nuffield Foundation, 2012. https://www.nuffieldfoundation.org/wp-content/uploads/2012/04/Mathematics-in-A-level-assessments-A-report-on-the-mathematical-content-of-A-level-assessments-in-Business-Studies-Computing-Economics-Geography-Psychology-and-Sociology.pdf.

———. 'Mathematics in A Level Assessments'. Nuffield Foundation, 2012. https://www.nuffieldfoundation.org/wp-content/uploads/2012/04/Mathematics-in-A-level-assessments-A-report-on-the-mathematical-content-of-A-level-assessments-in-Business-Studies-Computing-Economics-Geography-Psychology-and-Sociology.pdf.

OECD. *Education Policy Outlook 2019: Working Together to Help Students Achieve Their Potential*. Education Policy Outlook. OECD, 2019. https://doi.org/10.1787/2b8ad56e-en.

———. 'PISA 2021 Mathematics Framework (Second Draft)', 2018.
https://www.oecd.org/pisa/sitedocument/PISA-2021-mathematics-framework.pdf.

O'Grady, Kathryn, Marie-Anne Deussing, Tanya Scerbina, Yitian Tao, Karen Fung, Vanja Elez, and Jeremy Monk. 'Measuring up: Canadian Results of the OECD PISA 2018 Study'. Council of Ministers of Education, Canada, 2019.

Ontario Ministry of Education. *The Ontario Curriculum, Grade 9: Mathematics (2021)*, 2021. https://www.dcp.edu.gov.on.ca/en/curriculum/secondary-mathematics/courses/mth1w.

———. *The Ontario Curriculum, Grades 1-8: Mathematics (2020)*, 2020. https://www.dcp.edu.gov.on.ca/en/curriculum/elementary-mathematics.

———. *The Ontario Curriculum, Grades 11 and 12: Mathematics (2007)*, 2007. https://www.edu.gov.on.ca/eng/curriculum/secondary/math1112currb.pdf.

Research to Inform Elementary Mathematics Curriculum Revision, 2020. https://www.youtube.com/watch?v=AmZtIpAHZs4.

Ridgway, Jim. 'Implications of the Data Revolution for Statistics Education: The Data Revolution and Statistics Education'. *International Statistical Review* 84, no. 3 (December 2016): 528–49. https://doi.org/10.1111/insr.12110.

Rittle-Johnson, Bethany, Michael Schneider, and Jon R. Star. 'Not a One-Way Street: Bidirectional Relations Between Procedural and Conceptual Knowledge of Mathematics'. *Educational Psychology Review* 27, no. 4 (December 2015): 587–97. https://doi.org/10.1007/s10648-015-9302-x.

Royal Society Te Apārangi Expert Panel. 'Pāngarau Mathematics and Tauanga Statistics in Aotearoa New Zealand. Advice on Refreshing the English-Medium Mathematics and Statistics Learning Area of the New Zealand Curriculum'. Royal Society Te Apārangi, 2021.

Royal Society's Advisory Committee on Mathematics Education's (ACME) Post-16 Contact Group. 'Mathematics for the T Level Qualifications: A Rationale for General Mathematical Competences (GMCs)', 2019. https://royalsociety.org/~/media/policy/topics/education-skills/Maths/Mathematics%20for%20the%20T%20Level%20Qualifications%20-%20a%20rationale%20for%20GMCs.pdf?la=en-GB.

Schoenfeld, A. 'Learning to Think Mathematically: Problem Solving, Metacognition and Sense Making in Mathematics'. In *Handbook of Research on Mathematics Teaching and Learning*, edited by D. Grouws, 335–70. New York: Macmillan, 1992.

SCORE. 'Mathematics within A Level Science 2010 Examinations'. SCORE (Science Community Representing Education), 2012. https://www.stem.org.uk/resources/elibrary/resource/25956/mathematics-within-level-science-2010-examinations#&gid=undefined&pid=1.

Sinitsky, Ilya, and Bat-Sheva Ilany. *Change and Invariance: A Textbook on Algebraic Insight into Numbers and Shapes*, 2016. https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1429092.

Smith, Cathy. 'Gender and Participation in Mathematics and Further Mathematics A Levels: A Literature Review for the Further Mathematics Support Programme', 2014. https://core.ac.uk/download/pdf/79498409.pdf.

Smith, Cathy, and Jennie Golding. 'Schools' Strategies for Promoting Girls' Participation in Mathematics'. In *Proceedings Of the 42nd Conference of the International Group for the Psychology of Mathematics Education (Vol 4) Umeå, Sweden.*, edited by Ewa Bergqvist, Marcus Österholm, Carina Granberg, and Lovisa Sumpter, 211–18, 2018.

Smith, Cathy, and Candia Morgan. 'Curricular Orientations to Real-World Contexts in Mathematics'. *The Curriculum Journal* 27, no. 1 (2 January 2016): 24–45. https://doi.org/10.1080/09585176.2016.1139498.

Suurtamm, Christine. 'Elementary Mathematics Curriculum, Teaching, And Learning: The Past Ten Years'. *Ontario Mathematics Gazette* 60, no. 3 (2022): 19–24.

Suurtamm, Christine, Denisse R. Thompson, Rae Young Kim, Leonora Diaz Moreno, Nathalie Sayac, Stanislaw Schukajlow, Edward Silver, Stefan Ufer, and Pauline Vos. 'Assessment in Mathematics Education'. In *Assessment in Mathematics Education: Large-Scale Assessment and Classroom Assessment*, edited by Christine Suurtamm, Denisse R. Thompson, Rae Young Kim, Leonora Diaz Moreno, Nathalie Sayac, Stanislaw Schukajlow, Edward Silver, Stefan Ufer, and Pauline Vos, 1–38. ICME-13 Topical Surveys. Cham: Springer International Publishing, 2016. https://doi.org/10.1007/978-3-319-32394-7_1.

Tight, Malcolm. 'Twenty-First Century Skills: Meaning, Usage and Value'. *European Journal of Higher Education* 11, no. 2 (3 April 2021): 160–74. https://doi.org/10.1080/21568235.2020.1835517.

Tukey, John W. 'The Future of Data Analysis'. *The Annals of Mathematical Statistics* 33, no. 1 (1962): 1–67.

Wake, Geoff. 'Making Sense of and with Mathematics: The Interface between Academic Mathematics and Mathematics in Practice'. *Educational Studies in Mathematics* 86, no. 2 (June 2014): 271–90. https://doi.org/10.1007/s10649-014-9540-8.

Wal, Nathalie J. van der, Arthur Bakker, and Paul Drijvers. 'Which Techno-Mathematical Literacies Are Essential for Future Engineers?' *International Journal of Science and Mathematics Education* 15, no. 1 (1 May 2017): 87–104. https://doi.org/10.1007/s10763-017-9810-x.

Watson, Anne. *Key Ideas in Teaching Mathematics: Research-Based Guidance for Ages 9-19*. Oxford, United Kingdom: Oxford University Press, 2013.

Wolff, Annika, Michel Wermelinger, and Marian Petre. 'Exploring Design Principles for Data Literacy Activities to Support Children's Inquiries from Complex Data'. *International Journal of Human-Computer Studies* 129 (September 2019): 41–54.

Zevenbergen, Robyn. 'Mathematics, Social Class, and Linguistic Capital: An Analysis of Mathematics Classroom Interactions'. In *Sociocultural Research on Mathematics Education*, edited by Atweh, Bill, Helen Forgasz, and Ben Nebres, 201–15. Mahwah: Lawrence Erlbaum Associates, 2001.

Appendix A   Expert interviews

We gratefully acknowledge the following experts who contributed to this report, providing knowledge of professional and education policy, curriculum initiatives, research and evaluations within their fields of mathematics, statistics, data science and computing.

Christian Bokhove, Professor in Mathematics Education, University of Southampton

Tom Button, Maths Technology Specialist, MEI

Sophie Carr, Vice President for Education and Statistical Literacy, Royal Statistical Society

Neville Davies, Emeritus Professor, University of Plymouth

Kate Farrell, Director of Curriculum Development, Data Education in Schools, Scotland

Anna Fergusson, Professional Teaching Fellow, Dept of Statistics, University of Auckland

Matt Forshaw, Senior Advisor for Skills, Alan Turing Institute

Eirini Geraniou, Associate Professor of Mathematics Education, UCL

Rachel Hilliam, Chair, Alliance for Data Science Professionals

Jenni Ingram, Associate Professor of Mathematics Education, University of Oxford

Rhys Jones, Associate Dean, Faculty of Health and Medical Sciences, University of Surrey

Christina Leslie, Computational Biologist, Sloan Kettering Institute, USA

Darren Macey, Framework Design, Cambridge Mathematics

Simon Pinfield, Manager: Resources, Projects and Partnerships, Royal Geographical Society

Tom Rainbow, Advanced Mathematics Support Programme

Jim Ridgway, Emeritus Professor in the School of Education, Durham University

Judy Robertson, Chair in Digital Learning, University of Edinburgh

Neil Sheldon, Chair, Teaching Statistics Trust

Deborah Sutch, Curriculum Manager, Mathematics, Diploma Programme, IBO

Christine Suurtamm, Professor of Mathematics Education, University of Ottawa

Clare Walsh, Head of Education, Institute of Analytics

Appendix B   Royal Society's indicative curriculum models

Model		Change in structure/content
1	Current structure with no change	• No change
2	Current structure with additional content added to the: a) mathematics curriculum b) computing curriculum	• Retain the siloed structure of mathematics and computer science but add new content to help students develop the relevant competences • Add problem-based/computer-based material to the computing curriculum
3	Current structure with additional content added across curricula	• Retain the siloed structure of mathematics and computer science but add new content to different subjects across the curriculum

4	Two pillars
	
	• Pillar 1 – Pure/techniques-based maths
	
	• Pillar 2 – Mathematical thinking/problem-solving
		• Two pillars which will need integration
		
		   ○ Pillar 1 – Pure/techniques-based maths
		
		   ○ Pillar 2 – Mathematical thinking/problem-solving/computing (similar to Core Maths model)
		
		• Focus on problem-solving, data science, stats, mathematical modelling and using computers as tools
		
		• Consider both GCSE and post-16

5	Current structure with the addition of project-based work, similar to the Extended Project Qualification or International Baccalaureate	• Shift to a broader curriculum with different types of assessment
6	A new structure based on a broader education system similar to the IB	• Re-design the GCSE/post-16 education system (e.g., RS work in broad and balanced)