A vision for meaningful **data education** across the four nations of the UK

A report prepared for the Joint Mathematical Council of the UK

Laurie Jacques and Marie Joubert

About this report

In this report, commissioned and funded by the Joint Mathematical Council of the United Kingdom, we set out to propose a set of **data skills** for learners across the four nations of the United Kingdom at three milestones during their statutory school age and to identify the kind of system and classroom level resources necessary to provide a **data education** such that they achieve them. We draw from examples of **data skills** in the workplace and evidence from selected jurisdictions who are also responding to the increased demand for **data skills** from their citizens in a 21st century world. This report extends the work on a similar theme produced for the Royal Society's Mathematical Futures Programme (Smith, Kathotia, Ward-Penny, Howson, & Wermelinger, 2023) which proposed a set of competencies for mathematical and data literacy for school leavers. In this report we also specify **data skills** for learners at the end of primary and lower secondary. We contribute to the on-going discussions at the Royal Society on this topic (The Royal Society, 2024) by making recommendations for achieving a meaningful **data education** across the four nations of the United Kingdom.

Acknowledgements

We wish to thank academic colleagues both nationally and internationally and across all phases of education and in each of the four nations, for their willingness to discuss **data education** with us during the writing of this report. We would also like to thank our contacts who provided such insightful examples of **data skills** in the workplace and were happy for us to use them in this report. We would like to thank the Joint Mathematical Council for entrusting us with this work.

Dr Laurie Jacques and Dr Marie Joubert, October, 2024

Contents

About this report 2

Section 1: Introduction 4

Section 2: **Data skills** for the workplace 6

Section 3: Current practices in **data education** 13

Section 4: **Data education** in the four nations of the UK 19

Section 5: A vision for meaningful **data education** across the four nations of the United Kingdom 24

Summary of Recommendations 30

References 33

Appendices 36

Section 1: Introduction

What do we mean by **data skills** and **data education**?

This report is concerned with the **data education** that all learners[1] should experience and the **data skills** they should acquire in school between the ages of 4 and 16 in whichever of the four nations of the United Kingdon (UK) they are educated. This report follows on from the recent publication by The Royal Society setting out a case for **Mathematical and data education** (MDE) (The Royal Society, 2024).

There is a multitude of terms used in educational discourses in reference to **data education** and skills. For example, 'data literacy' (Wolff, Gooch, Cavero Montaner, Rashid, & Kortuem, 2016), 'statistical literacy' (Aziz & Rosli, 2021; McMaster, 2022) and 'data acumen' (National Academies of Sciences, 2018) are used. For the purpose of this report, we adopt the term **data skills** to describe the competencies learners need to achieve 'data literacy' as defined by Wolff et al. (2016). We choose this because it captures the complexity of working with data:

> *the ability to ask and answer real-world questions from large and small data sets through an inquiry process, with consideration of ethical use of data. It is based on core practical and creative skills, with the ability to extend knowledge of specialist data handling skills according to goals. These include the abilities to select, clean, analyse, visualise, critique and interpret data, as well as to communicate stories from data and to use data as part of a design process.(p. 23)*

We use **data education** to mean the necessary conditions both at system- and classroom-level that learners need in order to develop *foundational*[2] **data skills**. i.e. these are the **data skills** needed by *all* learners and upon which those needed for further *advanced* study and employment are built.

Why is **data education** important for citizens of the UK in the 21st century?

We live in an increasingly data driven world. The UK's citizens are both *sources* of data arising from our digital footprints, generated from our day-to-day lives, as well as *consumers* of data. We have (and give) access to a mass of data about our own and others' activities, visualised in varied ways that enable us to set personal goals, monitor our health, make decisions and satiate our curiosity. For instance, the exercise-tracking app, STRAVA[3] has a freely available global heatmap that indicates popular exercise routes which can be filtered by locality and exercise type. Whilst it might be interesting

---

[1] We use *learners* as a collective term to refer to pupils in early years settings or primary schools and students in secondary mainstream schools.

[2] Foundational mathematics establishes essential capabilities for further mathematical study and for learning more generally (The Royal Society, 2024, p. 31)

[3] https://www.strava.com/onboarding

to see this data, we should be aware of how individuals and organisations make use of our data and be assured that this data is collected and used in ethical ways (Wolff et al., 2016). This requires knowledge of *data stewardship*.

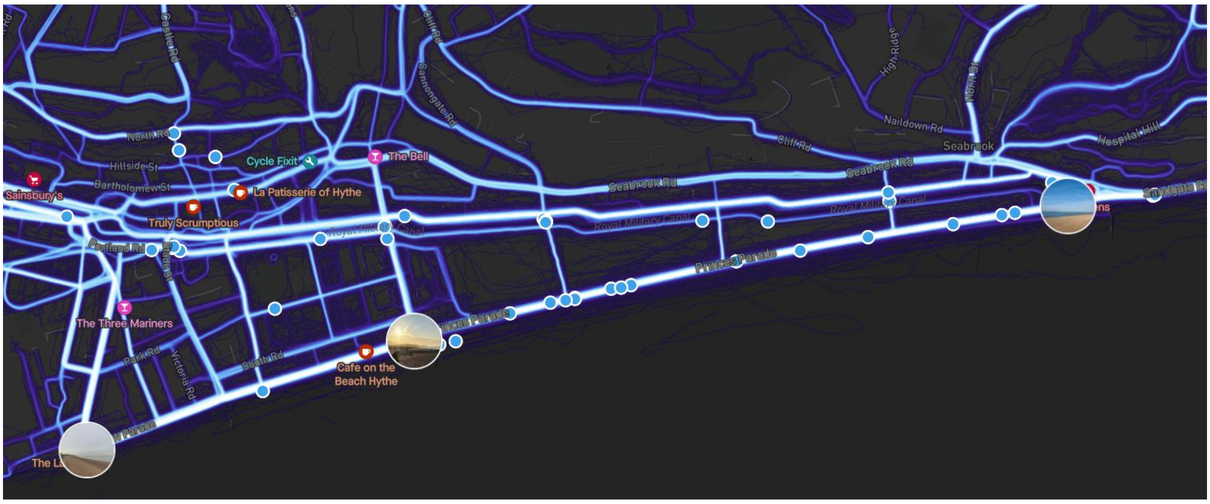

Figure 1 STRAVA heat map of running routes in Hythe, Kent, lighter routes indicate popular routes. (accessed 6/9/24)

In 2007 The Statistics and Registration Service Act established the UK Statistics Authority. This is an independent body at arm's length from government, which reports directly to the Westminster parliament, the Scottish Parliament, the Welsh Parliament and the Northern Ireland Assembly with the statutory objective of "promoting and safeguarding the production and publication of official statistics that serve the public good" and providing oversight of the Office for National Statistics and the Office for Statistical Regulation. They intervened during the recent UK election, reminding politicians to be transparent about any claims they were making during their campaigns. Its five year long strategy document (UK Statistics Authority, 2020) sets out a vision for the UK statistics system from 2020-2025. However, their strategy document does not include engagement with **data education**. There are clearly important relationships to be developed there.

As technology advances to enable data visualisations to communicate information from public and private sources in more engaging ways, so too must we be capable of engaging with these visualisations such as the heatmap above (Figure 1)(Royal Statistical Society, 2024). Furthermore, UK citizens of the 21st Century need to be able to critically engage with data and ask questions of information presented to them. For instance, by asking *how* the data was collected, *when* it was collected, whether there might be any biases to be aware of or how likely this data is to be representative? This kind of thinking is modelled well on BBC Radio 4's regular series, *More or Less*.[4]

How well do learners in the UK understand data?

In PISA 2022, for each of the four nations of the UK *uncertainty and data* was the most successfully answered subdomain in the mathematics assessment, maintaining a

[4] https://www.bbc.co.uk/programmes/b006qshd

trend over several cycles of PISA for the last 12 years. Whilst scores for Wales and Scotland were not significantly different from the OECD average, both Northern Ireland and England's average scores were significantly higher with only the top six East Asian education systems (Singapore, Macao, Taiwan, Hong Kong, Japan and South Korea) scoring a significantly higher average score than England (Ingram, Stiff, Cadwallader, Lee, & Kayton, 2023a, 2023b, 2023c; Scottish Government, 2023).

Despite these apparent strengths, our UK learners' capabilities do not appear to transfer successfully to the growing needs of UK society. There is increasing demand for specialist **data skills** at all competency levels across all industrial sectors and in all regions of the four nations of the UK (Fearns, Harriss, & Lally, 2023; The Royal Society, 2019). However, preparing and training learners to meet this demand is lagging significantly behind. Based on data of graduates from UK universities between 2017 and 2018, the estimated supply of data scientists from UK universities is unlikely to be more than 10 000 per year while up to 178 000 roles which require specialist skills will need to be filled in the coming years (Department for Digital Culture Media and Sport, 2021). Furthermore, of 9.2 million job adverts analysed in 2019, just over one in ten (996 000) required some form of data expertise such as basic IT skills, with over a third of these based in London (The Royal Society, 2019).

This report builds on substantial thinking that has taken place in the UK over last five years, beginning with a discussion held at The Royal Society, (The Royal Society & STEM Learning, 2019) by stakeholders with an interest in **data education** across the four nations of the UK (see Appendix 2) we make reference to these contributions without unnecessary elaboration or critique and blend them with some international and further thinking to consider the following research questions:

1. What are the **data skills** that learners should have developed by a) by the end of primary school, b) the age of 14, c) the end of formal education in the UK?
2. What is the **evidence base** on effective practices **in developing data skills** in mathematics (and elsewhere in the curriculum)? Are these small-scale insights and/or examples of effective system-wide **data skills** development?
3. What **resources** do we need to achieve meaningful improvement **in data education** across the four nations?

At the end of each section we summarise key points and highlight in bold, key messages that later form the recommendations at the end of the report.

Section 2: **Data skills** for the workplace

In this section we consider **data skills** that all learners may need for employment. We begin by considering the general workplace (rather than specialised data analyst or data scientist roles), and presenting vignettes of how data is used in different professions. We then identify the **data skills** that these occupations demand. These vignettes were gathered from people known to the authors via email communications and from a private online community. For the purposes of this report, we asked our contacts to describe how they process data in their work.

Air BnB Owner

I use Excel to input incoming payments, expenses (outgoings), tax payable (assumed 20%) and calculate net profit, month by month, line by line. This is simple book-keeping.

I also account for numbers of nights booked per month vs number of nights it was open for bookings; I work out a percentage of my occupancy rates (knowing this number helps with fulfilling a tax requirement for furnished holiday lets and business rates). This has helped me to make decisions. E.g. Looking back over two years' occupancy rates we now close completely in January as it's just not worth it.

Tax advisor

I take raw data supplied by clients in the form of numbers on documents. I have to find the right numbers on the documents, often inputting them into an Excel spreadsheet to collate them into groups of same type and add them up for a yearly total. I then present those numbers in an understandable format to help the client know how much money they made from their business. I then take those numbers to disclose that information in a prescribed format to the government and calculate how much tax that person should pay.

School data manager

I collate data into datasets in Excel spreadsheets which are heavily formula based using Power BI semantic models and some structured query language (SQL). I interpret the data and communicate it to middle and senior leaders in school. I also do it at a regional level for other academies and head office but the datasets we use there are pre-made so the data is all interpreted and communicated centrally.

Social media analyst

For my role I need to measure for each post, impressions, engagements and clicks. To do this I need to understand reporting in percentages vs in number terms and understand that number and percentage are different e.g. if you get 1k views but you have 10k followers then 1k views isn't that great. But, if you get 280 views but you only have 350 followers, then obviously that's much better. So, I need to be able to see things in the context of the whole. I need to have an eye for spotting outliers (i.e. if a post gets a high number of clicks) and then dive into why that specific post did well. I need to be able to look at the averages vs totals, and understand that the more posts there are, the lower the averages could be. E.g. if you have 1 post which gets 5k impressions, then your average number of impressions will be 5k. But if you have 10 posts which each get 2k impressions, then the average will be 2k. But the 2k average is better than the 5k average. I also need to explore patterns over time – trends - that help me or my client to make informed decisions. I need to present the data to my client in a way that provides clear and actionable strategic adjustments.

Senior Environmental acoustic consultant

I work in renewable development planning. We measure baseline noise data and analyse it in terms of mean/mode/median to obtain the representative level for different time periods. We collate source noise data which may be in a sounds pressure level and convert it to sound power level. We compare the data collected with limits/guidance. This is easy maths (subtraction) but presenting data in a table is important as this is how decisions are made. We use graphs in for a time history and % of periods with each level to show the mode. We sometimes calculate the noise levels across a grid of points and use contour maps to present these.

Senior Technical Specialist

I work in HR but my background is more in financial analysis and financial policy. My role involves providing the powers-that-be with insights about pay and reward and using them to shape pay and reward policies. The data manipulation itself isn't that complex - I use Excel mostly (5k-ish people) and my Excel skills are better than most but not exceptional. Where I add value is in understanding what the data means, how we got there, what that means going forward, what the impact might be on how we structure pay etc. and in presenting it in a way that tells a story and gets senior management to understand the issues quickly and comprehensively. I'd also add that if I could do what I do and develop my skills in data analytics packages it would be beneficial too.

Change Manager

I identify, measure and report project benefits to stakeholders. This involves gathering data from business teams, tracking it and putting it into shiny graphs for senior leaders. I do deep dives into our staff survey results and EDI reports to see trends and put interventions in place. E.g. I cut the data down by different characteristics to see what support we can give specific teams/groups for digital skills. I also do comms so I take data and create stories from it. All of this is in Excel. I'm totally self-taught. And I've probably forgotten more than I know because I use it a lot less than I used to (when I was a business analyst and project manager).

Head of Public Sector Transport

I need statisticians who know more than R. Many statisticians only know R and SQL and it is stifling the analysis and the visualisation that could be done. I think there is a massive skill gap - there's a bunch of people who know how to use R who are unable to evaluate if the outputs are right or turn the outputs into a story (i.e. the 'what is really going on here?') and then turn that story into visuals that non-statisticians actually understand.

The way we analyse data is changing. We need people who know how to write prompts (Prompt Engineers) to interrogate the data via a large language model (LLM) through asking the right questions and phrasing them in the right way. We need people who understand the underlying data in order to determine if the model is hallucinating. LLMs can use many forms of data. I am concerned about keeping data in spreadsheets which are prone to error and a security risk.

Weather forecaster

I developed a weather forecasting website. I started it because the people tasked with improving the data display did not listen to the end users (the forecasters) so I created what I wanted. The data is generated from outputs from different weather forecasting models run with different configurations. To work with the data, you need to know about meteorology (i.e. the physics of what's happening in the atmosphere and why) and a bit about the models and how they are run. As a forecaster I add value to the data because I know the strengths and weaknesses of different models and I can incorporate local effects that the models can't capture. People are working on producing entirely data-driven weather forecasts but there are many issues with this!

Communications consultancy: account director

As an account director, I need a range of **data skills**. In essence, it's pretty important to know some basic formulas and not be scared off by lots of numbers! I need to be able to produce a scope/fee proposal, which breaks down everything we're proposing by activity (drafting a press release, say), and by grade (each of which costs a different amount per hour). I need to be able to formulate the spreadsheet so it's clear and accurate for the client. I develop programme/activity trackers which to break down activities and give them timelines and owners. I also provide forecasts as to how much each project under my remit will earn that month, based on the fee proposal explained above. Each month has 'actuals' (tracking actual time logged), 'target' (which I need to adjust according to most recent developments' and 'WIP' (which tracks all unbilled time).

I feed into a pipeline which is basically a forward look at all projects we're bidding for and set out how much we think we'll earn for each. Then, I need to put a % of how likely we are to win the work. Using that, I need to provide a 'weighted' pipeline. So, if Project A is £10k/month for the next 12 months, and we're 60% likely to win it, the weighted pipeline will show £6k/month. I visualise stats for client/internal reports and analyse the results from public polling, for example, and draw conclusions about what that means for our clients.

Data skills in the workplace

Some of our contributors have developed the necessary **data skills** to fulfil the demands of their role whilst on-the-job. Others will have experienced professional or academic qualifications in order to be appointed to their role. Despite this, it is clear that there are a core set of skills being used to a greater or lesser degree.

Collecting and preparing data

The roles above require that data be gathered from already existing sources or generated as part of the role (e.g. senior environmental acoustic consultant). Data is then collated, cleaned and stored securely. To do this one needs to know and understand what will happen to the data in the next stage of the process. This might involve inputting data from raw sources (e.g. tax advisor), segment or cut the data by different characteristics (e.g. Head of Digital Engagement), transform it (e.g. senior environmental acoustic consultant) or combine it with other related data (e.g. school data manager, weather forecaster). Once this preparation process is complete, the next stage of working with the data can begin.

Analysing and interpreting data

The contributors above all had a role in preparing the data and then also analysing and interpreting it. The examples highlight different purposes for analysing and interpreting the data. For example, using the data to measure or define key performance indicators (e.g. school data manager, social media analyst), analyse trends and make decisions (e.g. senior technical specialist, Air BnB owner), solve problems (e.g. Head of Digital Engagement and Change) or predict future outcomes (e.g. weather forecaster).

The practical **data skills** required at this stage include drawing on mathematical skills (e.g. understanding percentages, rates and averages), using digital tools such as a spreadsheet (e.g. formulating formulae to perform calculations) and, for some roles, making use of more complex functionalities such as using Power BI semantic models or running SQLs in Excel. The head of public transport flagged the importance of knowing the correct tool with which to interrogate data and knowing the tool's limitations. Knowing how to optimise a tool is also an important skill, e.g. prompt engineers knowing how to write prompts for LLMs to get the best outputs from a query.

The head of public sector transport also indicated how having the necessary technical skills is not always enough. There are also more nuanced **data skills** required such as drawing on knowledge of the context in which the data is situated (e.g. senior, technical specialist, social media analyst) or understanding how the data had been generated in the first instance in a way that AI may not be able to (e.g. weather forecaster). This knowledge enables the data to be critically analysed and sense-tested such as recognising when to explore outliers or looking for trends over time (e.g. social media analyst), having a sense of the validity of the results from analysis and the approach used and comparing results with other indicators/ guidance (e.g. senior environmental acoustic consultant) or related data (school data manager).

For some of these roles, the consumers of the analysis of these data were not necessarily our contributors and therefore others need to have the skills to present and communicate data that has already been analysed.

Story-telling with data

When data needs to be communicated to a wider audience there is a need to present the data in such a way to meet the needs of that audience and to serve the purpose for which the data was analysed in the first instance (see *Analysing and interpreting data*). Several of our contributors referred to 'telling a story' (e.g. senior environmental acoustic consultant, senior technical specialist and head of public sector transport) as the goal of presenting the results of analyses of data. Having the skills to analyse the data may not necessarily mean having the skills to communicate the analysis of the data to others, particularly if the data has been analysed by a specialist and the audience are non-specialists (see weather forecaster and head of public sector transport). Therefore, there is a need to be able to construct a story in the data through representations and visualisations and have the skills to use the tools that generate them to tell the story clearly (e.g. senior technical specialist, head of public sector transport, senior environmental acoustic consultant).

The **data skills** described above provide some insights into what learners need to be prepared for when they leave school and move into certain types of employment and the examples clearly offer concrete examples of the problem solving cycle (Wolff et al., 2016), referred to as the PPDAC cycle, as shown in Figure 2.

Figure 2 The problem-solving cycle (PPDAC: problem, plan, data, analysis and conclusion) (Wolff et al., 2016)

The whole cycle should form the basis of the **data education** that learners all experience from early years to school leaving age. In the next section we discuss what we know about how **data education** is being implemented around the world.

Summary: **data skills** for the workplace

Data skills are necessary for making sense of our complex and multi-disciplinary world. Working with data has the potential to engage learners in meaningful mathematical activity. To meet the future needs of UK citizens and the demands for workforce skills, **any reviews of national curricula in each of the four nations should increase the prominence of *data education* at all phases and be expanded to include data skills that go beyond simple statistical manipulation**.

Being able to construct and interpret varied data visualisations are increasingly important **data skills**. Whilst bar charts, pie charts and box plots are often used to communicate data in the media and are simple for learners to construct by hand (for pencil and paper examinations) or interpret, learners need to develop the skills, including using digital technologies to construct visualisations and make sense of and question the varied representations of data that they encounter every day across different media and their workplace; a view also expressed by the Royal Statistical Society (2024). **Data skills need to include constructing and interpreting varied visualisations beyond the more traditional graphical representations**.

When collecting, analysing, interpreting and communicating with data in the real world (in contrast to the school world) it is done with purpose, for instance our workplace vignettes illustrated how the whole of the problem-solving cycle was followed, driven by relevant and purposeful contexts. **Data skills need to be developed for the whole of the problem-solving cycle starting with relevant and purposeful enquiry and working with increasingly larger data sets**.

Data skills need to include awareness and understanding of ethical acquisition and processing of data (data stewardship). Whilst not mentioned explicitly by our contributors the processing of data with ethical consideration is never more important and is not explicit in the planning and data collection stages of the problem-solving cycle. As mentioned in section 1, UK citizens are both sources of and consumers of data. Learners need to be aware of the ethical issues associated with leaving digital footprints such as how this data can be used by others who have access to this footprint (e.g. for marketing) and whether this data has been accessed with permission.

Section 3: Current practices in **data education**

Data education around the world

In this section we discuss approaches to **data education** taken in four countries that are generally seen as successful or interesting in their approach. We include the US and New Zealand as examples of forward-thinking initiatives that were also identified by Royal Society's Mathematical Futures horizon scan of international mathematics education policy (Adams & Boylan, 2023) and the Royal Statistical Society (Royal Statistical Society, 2024) for raising the profile of **data education**.

Finland

Finland is known for its progressive education system, and it integrates data literacy and statistical thinking across subjects from an early age. Statistics and data science are not standalone subjects but are embedded in mathematics, science, and even humanities with students learning to collect, interpret, and analyse data as early as primary school. There is, further, an emphasis on programming across the curriculum (Kadijevich, Stephens, & Rafiepour, 2023) with clear implications for applications in statistics and data science.

Phenomenon-based learning (Schaffar & Wolff, 2024) is a cornerstone of Finnish education, where students work on real-world problems that can require statistical analysis and data modelling. For instance, students might work on projects analysing environmental data, which helps them develop **data skills** in a context that is relevant and meaningful to them.

Singapore

Problem solving is at the heart of Singapore's school mathematics curriculum, which balances mastery with foundation basic skills and applying higher-order thinking skills for problem solving (Kaur, 2019). Throughout school, learners experience interdisciplinary project work, real-word applications of mathematics and modelling tasks (Dawn, 2011). The Ministry of Education has introduced computational thinking and data literacy from a young age, using real-world problems to teach these concepts (Seow, Looi, How, Wadhwa, & Wu, 2019). The curriculum suggests that students in the fifth year of primary should have opportunities to, for example, work in groups to collect data to solve a problem, use a spreadsheet to construct a graph and critically discuss examples of representations of data from newspapers[5]. In secondary they are expected to experience a cycle of formulating a problem, making sense of data, including real data presented in graphs and tables, applying appropriate concepts and skills to solve the problem and interpreting the mathematical solution to the problem[6].

[5] https://www.moe.gov.sg/-/media/files/primary/mathematics_syllabus_primary_1_to_6.pdf
[6] https://www.moe.gov.sg/-/media/files/secondary/syllabuses/maths/2020-express_na-maths_syllabuses.pdf

United States

In the United States, the Guidelines for Assessment and Instruction in Statistics Education: A Pre-K–12 Curriculum Framework (GAISE I), published in 2005 and revised in 2007 [7] addressed the perceived growing need for improved statistics education. It promoted increased statistical literacy and active learning and provided a framework for instruction. In 2020, an updated version of guidelines, GAISE II, was published[8]. It aimed to build on the core principles of GAISE I but to expand these to take into account the changes in the nature and availability of data. The report explains that it highlights, for example, asking questions all the way through the statistical problem-solving process, careful planning of data collection to help answer investigative questions and the importance of using technology. GAISE II recommends that, in whatever assessment is used for statistics, students should need to use reasoning at all stages of the statistical cycle and that, if technology is used in the assessment, this should align with the technology used in the classroom.

The teaching approaches recommended in the GAISE frameworks represent a shift away from teaching-as-usual in US classrooms; it is important that teachers engage in professional development to prepare them to use these approaches and to provide them with an understanding of foundational topics (Franklin, 2013, 2014).

The American Statistical Association(ASA)[9] is dedicated to and involved in enhancing statistics education at all levels across the US, including the production of a K-12 Curriculum (Bargagliotti et al., 2020). In 2016, Christine Franklin was appointed as the inaugural ASA Statistical Ambassador[10] advocating for statistical education and raising its profile amongst stakeholders and policy makers.

The recommendations for the statistics curriculum published by the Royal Statistical Society (Royal Statistical Society, 2024) were informed by GAISE II.

New Zealand

According to Forbes (2014), "*New Zealand has been leading the world in terms of the data handling, and in more recent years, data visualisation approach in its school statistics curriculum*." The approach they have adopted appears to have been shaped by the involvement of professional statisticians and led to renaming the mathematics curriculum to 'mathematics and statistics' as early as 2007 (Callingham & Watson, 2024).

New Zealand is in the process of a 'curriculum refresh', "*in response to gathering pressures to be more deliberate in how we educate our young people for the complex and uncertain futures that await us all*" (Chamberlain et al., 2021, p. 3). In the draft

[7]https://www.amstat.org/asa/files/pdfs/GAISE/GAISEPreK-12_Full.pdf
[8] https://www.amstat.org/asa/files/pdfs/GAISE/GAISEIIPreK-12_Full.pdf
[9] https://www.amstat.org/education/k-12-educators
[10] https://www.amstat.org/education/asa-k-12-statistical-ambassador

curriculum for mathematics and statistics[11], at all levels the statistics statements reflect the steps of the PPDAC cycle and one about statistical literacy. For the latter, at Year 4, students "*use the statistical processes to investigate, interpret, critique, and check the claims made about data presented in tables, pictographs, bar graphs, line graphs, and pie charts*" and at Year 8 they "*examine the data-collection methods, data visualisations, and findings of others' statistical investigations to see if their claims are reasonable*." There is a clear indication on the Curriculum refresh page[12] that teachers will need support in teaching the new curriculum.

The above examples highlight possibilities that countries are exploring to implement **data education** at system and classroom level. At system level, countries and states are developing curricula where **data skills** feature prominently or form part of **interdisciplinary or project-based learning**. In some cases, this is response to the OECD's thinking about curricular design for the 21st Century (OECD, 2020a, 2020b, 2020c; Taguma, Makowiecki, & Gabriel, 2023). There are examples where **data skills**, although part of the mathematics curriculum, are also **integrated into other curriculum content** and sometimes in unexpected places, for instance, Japan incorporates data literacy mostly into first language content (Figure 3).

Not only have **data skills** have become a curriculum priority, but the above country examples also illustrate components that support **data education** at classroom level such as supporting teachers to use **innovative teaching practices,** making use of **digital technologies** and investing in **educational initiatives** to support teachers and learners e.g. websites, professional learning opportunities and curriculum resources. What they do not tell us, however, is how well these initiatives work in practice and therefore we need more research about this.

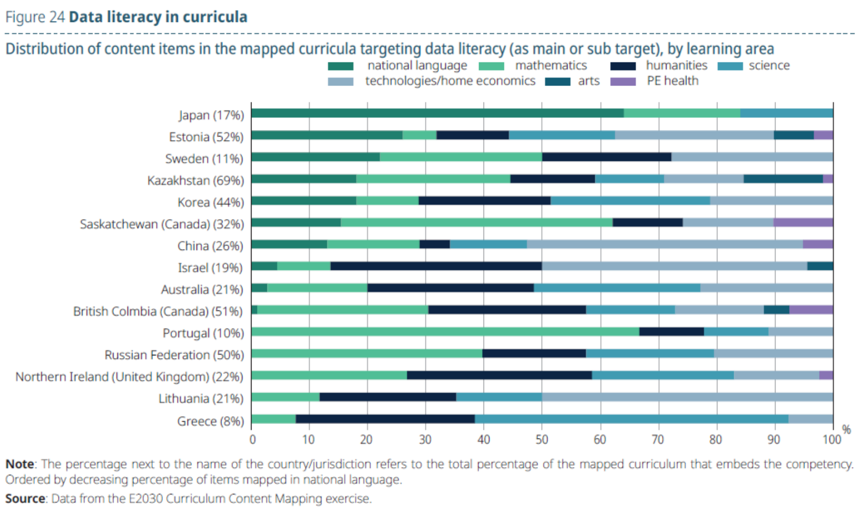

Figure 3 Data literacy as part of an interdisciplinary curriculum, (OECD, 2020b, p. 59)

[11] https://files-au-prod.cms.commerce.dynamics.com/cms/api/qwxsnqcpfm/binary/ML70rC

[12] https://newzealandcurriculum.tahurangi.education.govt.nz/new-zealand-curriculum-online/new-zealand-curriculum/new-zealand-curriculum-refresh/563714466

Resources to support **data education**

A high-quality **data education** requires high-quality resources. This was discussed at a workshop hosted in the US in 2022 by the Board on Science Education at the National Academies of Sciences, Engineering, and Medicine. We identify some key messages from the proceedings (National Academies of Sciences, Engineering and Medicine, 2022) along with further examples presented at the at the 15th International Congress of Mathematics Education survey report on statistics and data science education as a vehicle for empowering citizens. We identified three types of resource. For each type we provide examples from around the world.

Cross disciplinary projects

The Thinking with Data project (TWD) (Vahey et al., 2012). The TWD project in the US designed and evaluated a set of four 2-week, integrated modules for cross-disciplinary implementation in 7th grade social studies, mathematics, science, and English language arts by providing a set of teaching modules. The project measured the impact using a quasi-experimental design and showed that learners were able to engage in more sophisticated data literacy activities as a result of participating in the TWD intervention when compared to the control group.

Data skills in Geography (Royal Geographical Society, 2019). This project in the UK was funded by the Nuffield Foundation[13] with the aim to "*upskill teachers of today and enhance the understanding of teachers of tomorrow*" (p. 5) The project provided an integrated approach across Higher Education and schools and included sharing good practice and expertise, producing teaching materials for GCSE and A level with accompanying teacher professional development and collaborations with initial teacher education departments and other learned societies. The project reached over one thousand teachers who reported that their confidence in the use of **data skills** had increased.

Digital Technological tools

Writing data stories project (Wilkerson, Finzer, Erickson, & Hernandez, 2021). This project in the US seeks to introduce students to data storytelling. As part of this project, the developers have designed a plug-in for the freely available data visualisation platform CODAP[14] (Common Online Data Analysis Platform) with which learners can integrate data gathered from their own enquiry, trace their analysis process and reflect and share their work with others. The resource is currently being piloted with 10-14 year olds.

Tinkerplots (Saldanha & Hatfield, 2021) is a dynamic digital data exploration environment that can be used for the full PPDAC cycle. The design research project

---

[13] https://www.nuffieldfoundation.org/project/data-skills-in-geography
[14] https://codap.concord.org/

(Saldanha & Hatfield, 2021) involving *Tinkerplots*[15] aims to develop understandings of distribution and box plots with 7th grade learners in Canada. They used pre- and post-tests to demonstrate that the learners enriched their imagery of box plots. Tools within the software such as *dividers* and *percentile hats* contributed to learning, together with 7th graders reflections on using these tools that served to orientate their attention to data quarters' location and spread.

Curriculum resources

Data education in Schools[16] is part of the Edinburgh and Southeast Scotland City Region Deal Data skills Programme, funded by the Scottish and UK Government. The programme brings together a range of stakeholders to provide expertise and support for teachers and learners in the form of professional learning workshops, data literacy courses for primary and secondary teachers and a suite of **data education** resources[17] for primary and secondary classrooms.

CensusAtSchool – TataurangaKiTeKura[18] is a national data project accompanied by curriculum resources that gathers data biennially from learners in New Zealand aged 3-13 and makes the data available for learners to conduct authentic enquiry with a large dataset. The project also includes curriculum resources organised by year group and different topics to support teachers to develop local enquiry projects.

Summary: Current practices in **data education**

We set out to identify effective practices in data education, however there was no evidence of system level effective practices and limited evidence of smaller scale practices. We identified those which we considered to have potential.

Domain specific competencies are **data skills** that are used and applied in cross-disciplinary spaces (The Royal Society, 2024, p. 22), the national curriculum for each of the four nations should **ensure that foundational data skills are specified across all subjects, for all phases from the beginning of primary using consistent terminology and defined coherently for each phase of study.**

If data skills are to be integrated into other curriculum content, this will demand that all teachers across all subjects provide a meaningful data education, including the effective use of digital technologies. However, it is well recognised that most teachers, including many secondary mathematics specialist teachers, do not have the confidence or knowledge to teach data skills effectively. This implies the **need for significant professional learning to broaden teachers' own knowledge of this growing discipline and to build their confidence in teaching approaches including their use of a range of appropriate digital technologies for working with data. By**

[15] https://www.tinkerplots.com/
[16] https://dataschools.education/about-us/
[17] https://dataschools.education/data-education-resources/
[18] https://new.censusatschool.org.nz/

bringing together teachers across disciplines, and across phases, there is the potential to develop a coherent view of how cross-disciplinary data skills progress over time.

In section 2 we suggest the prominence of **data education** in each of the four nations should be raised. To do this, those most passionate will be best placed to promote this priority. This could be achieved by **appointing a single *Ambassador for Data Education* across all four nations of the UK, such as in the US, to advocate for high-quality, meaningful data education** and with a particular focus on all phases of education from EYs to school leaving age and beyond. The *Ambassador* would provide a conduit between member bodies of JMC, policy makers, other MDE stakeholders, including teachers, school leaders and the UK Statistical Authority. Their role would be to influence and support policy-level and other educational initiatives addressing the improvement and sustainability of meaningful **data education** across the UK and to stimulate the production of **data education** curriculum resources, including signposting appropriate digital technologies.

There is a plethora of curriculum materials available online to support teachers to teach **data skills** in engaging and meaningful ways for all phases of education. These require careful curation particularly to ensure that they are relevant to teachers in each of the four nations of the UK. For some resources, such as specialist data analysis and data visualisation digital tools (such as CODAP or Tinkerplots), teachers will require additional support to use them effectively in the classroom, so teachers need to know where to access high-quality professional development. What is required is the **creation of a single online curated repository of quality-assured and freely available data education curriculum materials and signposts to high-quality professional learning opportunities**.

In our research literature searches we noted that the main focus for researching **data skills** and **data education** involved teaching and learning in secondary and post-16 education and often conducted in jurisdictions outside of the UK. The current literature lacks relevance to the earlier years of **data education** and to the contexts of the four nations of the UK**. Further UK-centric research is needed to understand how learners develop data skills in national contexts in early years, primary and secondary settings and develop effective teaching practices.**

Data education is a growing world-wide priority. Policy makers need to find system-level mechanisms to address this. There is limited evidence about those mechanisms and what works, and what does not work, at scale. **Further research is needed that focuses on understanding and addressing barriers to successful data education policy implementation in the context of the UK** and thus contributing to the growing global debate about how to accommodate this priority.

Section 4: **Data education** in the four nations of the UK

In this section we contextualise **data education** within the education systems of the four nations by drawing considering curriculum and assessment. We then compare specifically the current **data skills** expectations for learners at the end of three stages in their **data education**. We draw on the comparative analysis conducted by Forsythe and Smith (2020).

Curriculum

England

- Most recent curriculum was introduced in 2014 and is currently under review[19].
- Specifies content for learners from age 5 to 16.
- Data handing/ statistics has reduced in content in successive curriculum iterations since 1990 (Boylan & Adams, 2023).
- **Data skills** are referred to as *statistics* in the programmes of study.

Scotland

- The most recent curriculum was introduced in 2010, known as The Curriculum for Excellence (CfE) with a fresh narrative introduced in 2019.
- There are four capacities at the curriculum's centre: successful learning, confident individuals, responsible citizens and effective contributors.
- The curriculum specifies a substantial number of benchmarks up to S3 (around 14 years of age).

Wales

- The Curriculum for Wales was introduced in 2022.
- **Data skills** are integrated across various areas of learning and experience.
- Skills are emphasised alongside/as well as the content.
- **Data skills** are called data in the specification in the curriculum.

Northern Ireland

- The most recent curriculum was introduced in 2007 and is currently under review[20].
- Mathematics and numeracy form part of the curriculum with *data handling* as sub-domains.
- There is requirement to develop cross-curricular skills of communication, using mathematics and ICT and thinking skills and personal capabilities (managing information; thinking, problem-solving and decision-making; being creative; working with others; and self-management.

---

[19] https://www.gov.uk/government/news/government-launches-curriculum-and-assessment-review
[20] https://www.education-ni.gov.uk/news/review-northern-ireland-curriculum-announced

Assessment

England

- Mathematics Teacher assessment at the end of KS1, with optional end of KS1 tests and statutory end of KS2 tests.
- Almost all learners sit GCSE Mathematics Foundation and Higher Tier options (a Level 2 qualification).
- Optional additional GCSEs in statistics.
- Beyond 16, learners can take Core Maths qualifications, designed for students who want to continue to study mathematics but have chosen not to study A-level mathematics. Core maths helps students develop their quantitative and problem-solving skills.

Scotland

- Teacher assessments in numeracy informed by Scottish National Standardised Assessment in Primary 1, Primary 4 and Primary 7 (end of primary) and Secondary 3.
- Personalised online tests since 2019.
- At 16, most students gain National 4s (internally assessed) or National 5s (externally).
- Range of other qualifications includes National 4 and 5s in Numeracy which include outcomes for Interpret graphical data and situations involving probability to solve straightforward, real-life problems involving money/time/measurement.

Wales

- Annual numeracy test performance from Year 2 and teacher assessment of Mathematics National Curriculum levels at ages 7 and 11.
- Online adaptive tests started 2019.
- Numeracy tests and Mathematics levels are reported at 14.
- From 2016, there are two GCSEs, *Mathematics* and *Mathematics Numeracy*, with three tiers. At 16, most students take both, from 2025 there will be a combined double award GCSE.
- Numeracy GCSE includes content for statistics and some probability. Mathematics GCSE includes all Numeracy content plus additional probability.

Northern Ireland

- At end of Key Stages 1 & 2, there are standardised and moderated teacher assessments in Using Mathematics.
- At KS4 the study of mathematics and numeracy is not compulsory but in practice it is one of the main areas of learning.
- Nearly all schools use the Council for the Curriculum, Examinations & Assessment (CCEA's) GCSEs in Mathematics and Statistics.

Data skills expected in the curricula of the four nations of the UK

Programmes of study, assessment frameworks and examination specifications from the four nations define learner expectations at key points in learners' education. We used these centrally produced documents to extract the current expectations relating to **data skills** at the end primary, at the end of lower secondary and at the end of upper secondary when learners would be examined (see Appendix 2). We did this to compare expectations across the four nations and to gain a sense of the current landscape of **data education** across the UK.

At the end of primary, expectations appeared to be similar across all nations in terms of representing and interpreting data. However, in using the end of KS2 test framework[21] for England, the statements only represented what could reliably be assessed in test conditions, therefore there was no mention of all stages of the PPDAC cycle. Notably England had no statements about probability. Both Northern Ireland and Scotland included expectations for learners to use digital tools such as spreadsheets, databases and technologies to display data simply and clearly.

At the end of lower secondary, expectations varied in detail and how they progressed from the **data skills** expected of respective primary learners. Compared to the other three nations, Scotland's expectations build substantially and coherently on the expectations for primary whereas there was hardly any progression in Northern Ireland for the **data skills** described in the levels of progress (LOP) to LOP5. Like England, Northern Ireland expected learners to construct representations, building on interpreting them in primary. Northern Ireland and Scotland continued to refer to the use of technologies with which to work with data.

At school leaving age detail in the national examination specifications are largely the same for Northern Ireland, England and Wales (the latter two sharing the same specification approach) [22]. However, in Scotland, the National 5 specification only expects students to compare data sets by performing simple descriptive statistics and simple regression. Since the specifications are for an examination, it is unsurprising that the expectations require learners to demonstrate statistical knowledge and skills that can be reliably assessed under exam conditions e.g. performing statistical processes rather than the full PPDAC cycle.

Overall, the four nations' expectations vary most in the earlier phases of education but converge by the time learners reach the school leaving age. There are several observations of note. Firstly, that there is little attention to the purpose for collecting and analysing data and as such there is no sense of relevant contexts within which working with data should arise. Secondly, the expectations focus on statistics (and

---

[21] https://assets.publishing.service.gov.uk/media/5a822bbded915d74e62362a2/2016_KS2_Mathematics_framework_PDFA_V5.pdf

[22] We used the specifications for GCSE foundation tier which are relevant to all learners.

probability) and ignore the broader skills associated with data literacy, critical thinking and awareness of ethical data use (data stewardship).

Whilst there were some examples of digital technologies in some of the specifications for younger learners these disappeared by the time learner reach 16, since the current assessment arrangements do not accommodate digital data manipulation.

Finally, it is striking how what is prioritised in the national assessments bears only minor relationships to the skills and experiences that our workplace vignettes contributors described, making full use of the PPDAC cycle. The current method of assessment cannot assess learners' capabilities to identify a problem, construct a question and plan how to answer it by collecting and preparing data, analysing and interpreting data and then using data and visualisations to tell stories in the way our workplace vignette contributors so clearly described.

It is well documented that high stakes assessment influences the intended curriculum; as Hugh Burkhardt said, *WYTIWYG* (what you test is what you get). What UK nations currently *get* from testing young people on the **data skills** content of the national examinations at school leaving age is not what is *needed* when they move on to further study or take up their first job. What these learners *do* need is a set of **data skills** to enable them to work on and solve **purposeful** problems in **relevant** contexts. This need is also recognised in a set of recommendations for a statistical curriculum (Royal Statistical Society, 2024) and we illustrate how this can be addressed in section 5.

Summary: **Data education** in the UK

Whilst there are already national examinations for statistics across the four nations for school leaving age, these are currently optional for learners. When national assessment systems are reviewed, foundational **data skills** should form a key component of the assessed course of study for all learners as they reach school leaving age. Future reforms of **GCSE Mathematics or Scottish National 3/4/5 examinations should include a compulsory component for *all* learners involving foundational data skills** to ensure that all learners are taught the necessary **data skills** for when they are employed or prepare them for further study, for instance for Core Maths.

How **data skills** are assessed inevitably drives how they will be taught. Given that the majority of data in the workplace is generated and processed in a digital space as shown in the workplace vignettes above, it seems pragmatic to teach **data skills** using digital technologies, wherever possible. Digital tools are available as shown in section 3. The most recent PISA assessments used a digital platform to assess mathematical literacy including for the sub-domain uncertainty and data. The items in this assessment enabled the learners to interactively explore data in a digital environment. Whilst previous attempts to use course work as an assessment method have broadly been unsuccessful because of marginalising groups of students, inconsistent marking and teacher workload, the use of digital environments to simulate a data-led project during school time seems a reasonable solution. Furthermore, generative AI is developing rapidly and could be utilised to assess learners' work in simulated 'project' spaces. This would relieve teachers or external agencies of the responsibility of marking and optimise objective assessment. **Digital assessments of data skills projects should be trialled for respective school leaver examinations across the four nations of the UK. Data skills** at KS3 in the low stakes assessments proposed by The Royal Society (2024, p. 83) could also utilise such a digital data project environment.

Data education is a priority for governments of all four nations and the UK Statistics Authority has a potential part to play in this. When its Five Year Strategy (UK Statistics Authority, 2020) comes to an end in 2025 it could be valuable to **include an objective for it to engage with data education, to assist in emphasising the importance of data education across the four nations**.

Our review of literatures relevant to **data education** in the UK has not found any substantial evidence of current practices in schools. To gain a better understanding of **data education** in the four nations of the UK it might be useful to c**onduct state-of-the-nation reviews by getting directly into schools, observing data education in practice and discussing with teachers and school leaders how data education is incorporated into school activities.**

Section 5: A vision for meaningful **data education** across the four nations of the United Kingdom

Data skills at key points in learners' **data education**

In this section we identify an aspirational and coherent progression of **data skills** defined at three points in learners' **data education** from the end of primary school through to school leaving age (pp 26-28). For comparison, Appendix 1, summarises the existing expectations for **data skills** as described in the mathematics curricula for each of the four nations at the end of primary and lower secondary and school leaving age.

From their evaluation of several exemplar **data skills** curricula from the US (Bargagliotti et al., 2020) and New Zealand[23], Smith et al. (2023), proposed a set of competencies for school leavers. We draw on their framework to propose an additional set of **data skills** for the end of primary and lower secondary to accompany the original set. The Royal Society report for Mathematical and Data Education (The Royal Society, 2024) does not make such detailed recommendations, therefore we offer this set of **data skills** to add to the wider discussion about MDE. We felt that the 'intersectional competencies' *interpreting and communication*, *evaluating and critiquing* and *use of technology* proposed by Smith et al. (2023) were adequately subsumed into the proposed **data skills** statements, therefore our proposal includes **data skills** only for *conducting iterative and holistic enquiry*; *posing questions* and *solving problems* and the sub-domains for the *statistics/ data science* toolkit (renamed here as the **data skills** toolkit).

Our **data skills** for the end of primary and end of lower secondary emphasise **relevance** and **purpose** throughout (see those proposed by The Royal Society (2024, p. 32), with **relevance** gradually expanding out from learners immediate experience as they grow. We borrow this approach from New Zealand[24]. For instance, in primary, the problems learners work on should be relevant to their immediate locality but as they grow older, they will work on problems in their wider locality, or of national and global **relevance**. We also propose that the scale of a data set increases over the period of a learner's **data education**, for instance in primary, learners might collect only a small sample of data, whereas in later phases they will have developed skills to handle larger sets of data that are better handled with digital tools. In England, in contrast to the devolved nations, the 2014 National Curriculum no longer included the beginnings of probability in the programmes of study for KS1 and KS2. However, probability is experienced first-hand by the youngest children, and as we believe that any **data skills** developed should

[23] https://nzcurriculum.tki.org.nz/The-New-Zealand-Curriculum/Mathematics-and-statistics/Achievement-objectives#collapsible1
[24] Private conversation with a New Zealand academic colleague

build on learners' prior experiences, and we argue that the beginnings of probability should be reinstated for primary school learners in England.

The **data skills** we propose are detailed and aspirational, for which we make no apology. We acknowledge the concern that the current curricula of the four nations is overcrowded (The Royal Society, 2024, p. 61) and argue that to accommodate our proposed set of **data skills**, there might need to be compromises in other areas of mathematics *and* other subject curricula. However, working with data provides a meaningful context within which to use and apply skills learned in other mathematical domains, such as arithmetic and measurement, thus *capitalising on connected content* (The Royal Society, 2024, p. 61). We therefore argue that there is potential to leverage the prioritisation of **data education** by simultaneously improving the profile of mathematical and statistical problem-solving, application and reasoning.

Data education and Early Years

Our brief was, in part, to include **data skills** at the end of primary. Experience of all elements of the PPDAC cycle should begin in the early years. Young learners are naturally curious about their environment. They gather information and make sense of it by playing with, exploring and sorting collections of objects and or repeating events to establish patterns or solve problems. These experiences are the beginnings of statistical inference. As for all learners, data should be situated in the learners' natural environment. Synthetic experiences such as gathering information about pets, eye colour etc. which do not begin with **purpose** (i.e. to assist in solving a problem or making a decision), will be a meaningless exercise and potentially lay the foundation for negative attitudes towards working with data. On the other hand, exploring a wide range of visualisations of information in 3D and 2D provokes spatial reasoning. Stories can form a meaningful context for gathering data e.g., organising and representing the number of animals that join the old lady in *A Squash and a Squeeze*[25]. Our youngest learners can also experience the beginnings of data stewardship through discussing what it means to get a fair vote when choosing to do something in class. We do not make any specific recommendations for **data skills** in the early years, but we do advocate **purposeful** and **relevant** experiences of working with data throughout the early years just as for other phases of education.

[25] https://www.juliadonaldson.co.uk/books/picture-books/a-squash-and-a-squeeze/

		End of primary	End of lower secondary	School leaving age
Intersectional competency	Conducting iterative and holistic enquiry	Create and conduct investigations for different **purposes** (e.g. solving a problem, making a decision, observing changes, better understanding of an issue/ risk) using the statistical enquiry cycle to understand issues **relevant** to their immediate locality and context, using primary data they have collected.	Create and conduct observational and experimental investigations for different **purposes** (e.g. solving a problem, making a decision, observing changes, better understanding of an issue/ risk) using the statistical enquiry cycle to understand and propose solutions to issues **relevant** to their immediate and wider locality and context, combining primary and secondary data.	Create and conduct for different **purposes** 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.
	Posing questions	Pose questions for different **purposes** framed in **relevant** local/ familiar contexts including recognising contexts that might be inappropriate about which to ask questions.	Pose and relate to observational and experimental questions framed in **relevant** local/ familiar contexts; recognise and address ethical and unethical implications when formulating different questions. Observe, compare and describe different outputs from similarly worded prompts in large language model AI systems.	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.
	Solving data-driven problems (**data skills**)	Explore and compare single variable primary data to solve problems or make decisions **relevant** to their immediate locality and familiar contexts.	Explore univariate and simple relationships in bivariate observational and experimental data gathered for purposes **relevant** to their immediate and wider locality/ unfamiliar contexts.	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.

Data skills[26] toolkit	Data stewardship	Describe simple methods of processing data securely e.g. choosing and setting a secure password. Understand that apps on digital devices can leave data trails. Understand that people have rights about how their data is used. E.g. 5Rights campaign.	Recognise that some data might be sensitive and may need to be processed differently from other data. Understand different ways to protect data including personal data on digital devices. Know that ethical data collection involves seeking consent, transparency and limiting harm to those from whom data is collected.	Recognise the **purposes** for which data are collected and processed, and attend to privacy, security, and ethical concerns in a responsible manner.
	Handling Data	Describe the difference between primary and secondary data. Gather, clean, organise and store discrete and continuous primary data of **relevance** to local/ familiar issues and contexts of a single variable using a variety of appropriate methods. (e.g. simple paper and/or digital surveys/ forms; simple data bases)	Gather, organise, clean, and store discrete, continuous and grouped data (primary and secondary) of a single variable to understand issues **relevant** to their immediate and wider locality and context, using appropriate methods including digital technologies. Critically consider how secondary data has been gathered, including potential biases.	Gather, organise, clean, store, display and interrogate data **purpose**fully (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.
	Data representation	Recognise and describe how some data can only take certain values (discrete, e.g. by counting) and some data can take any value (continuous, e.g. by measuring). Understand that data visualisations tell stories about data gathered for a particular **purpose**. Interpret and compare different representations for discrete and continuous data to understand issues **relevant** to their immediate locality. Represent discrete and continuous data gathered for **purposes relevant** to their immediate locality and context using appropriate visualisations generated by hand and with appropriate technology.	Identify whether data is discrete, continuous or grouped/ungrouped and differentiate between primary and secondary data. Interpret visualised data by identifying stories in the data and critically compare and evaluate the different ways visualisations communicate data to understand issues **relevant** to their immediate and wider locality and contexts and according to the **purpose** for which the data were collected. Critically consider how secondary data has been represented, including looking for the possibility of miscommunicating or misrepresenting data. Choose and generate visualisations of data gathered to understand issues **relevant** to their immediate and wider locality and context using appropriate technology according to the **purpose** for which the data were collected.	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.

[26] In Smith et al. (2023) this is referred to as the statistical and data science toolkit

	Statistical thinking and methods	Understand the difference between mean, mode and median and know how to find them from data gathered to understand issues **relevant** to their immediate locality and context. Choose when mode, median or mean is the more appropriate measure according to the **purpose** for which the data were collected. Identify calculations required to compare data of relevance to issues in their immediate locality and context. (e.g. one-step and two-step sum and difference problems).	Interpret and use measures of central tendency (mean, mode, median) and spread (range, consideration of outliers) to make sense of data **relevant** to their immediate and wider locality and context based on the **purpose** for which the data were collected.	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.
	Probabilistic reasoning	Describe events that are impossible, unlikely, likely or certain. Distinguish between fair, unfair and random events. Make predictions from data analysis gathered for a **purpose relevant** to a local/ familiar issues or contexts. Describe whether and give reasons why data analysed for a **purpose relevant** to a local/ familiar issues or context was as predicted. Describe how an event may relate to other events.	Record, describe and analyse the frequency of outcomes of simple probability experiments involving randomness, fairness, equally and unequally likely outcomes, using appropriate language and the 0-1 probability scale. Understand why the probabilities of all possible outcomes sum to 1. Generate theoretical sample spaces for single and combined events with equally likely, mutually exclusive outcomes and use these to calculate theoretical probabilities. Use probabilistic reasoning to make sense of outliers or other unusual/ unexpected features of data **relevant** to their immediate and wider locality and context and based on the **purpose** for which the data were collected.	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.

	Using aids and tools	Use calculators to perform calculations with small sets of data gathered for a **purpose relevant** to their immediate locality/ context. Use simple paper based and digital forms to gather data for a **purpose relevant** to their immediate locality/ context.	Use digital tools (e.g. online forms, data loggers) to gather data for a **purpose relevant** to their immediate and wider locality and context. Use spreadsheets and data bases to store, organise and generate visualisations of data. Write simple formulae in spreadsheets to perform calculations involving data gathered for a **purpose relevant** to their immediate or wider locality/ context, involving more than one cell.	Use a range of relevant computational tools, including calculators, spreadsheets, mathematical and statistical software, data analysis packages, visualisation tools, and machine learning algorithms.

Summary of Recommendations

Throughout this report we have identified issues that need addressing according to our three research questions (see p.6). We summarise these recommendations according to the questions identifying those potentially responsible for actioning them.

*1 What are the **data skills** that learners should have developed by a) by the end of primary school, b) the age of 14, c) the end of formal education in the UK?*

In section 2 we identified the **data skills** for different occupations and in section 4, the different curriculum contexts of the four nations. We make the following recommendation:

Recommendation 1a (Policy makers)

Reviews of the national curricula in each of the four nations should increase the prominence of ***data education*** at all phases and be expanded to include **data skills** that go beyond simple statistical manipulation. These skills should:

- include constructing and interpreting varied visualisations beyond the more traditional graphical representations;
- be developed by experiencing the whole of the PPDAC problem-solving cycle starting with **purposeful** and **relevant** enquiry and working with increasingly larger data sets during learners' **data education**;
- include awareness and understanding of ethical acquisition and processing of data (*data stewardship*).

In Section 4, we described the different assessment contexts for **data education** in each of the four nations. We make the following recommendations:

Recommendation 1b (Policy makers)

Reform **GCSE Mathematics or Scottish National 3/4/5 examinations to include a compulsory component for *all* learners involving foundational data skills** to ensure that all learners are taught the necessary **data skills** for when they are employed or to prepare them for further study, for instance for Core Maths.

Recommendation 1c (Policy makers)

Digital assessments of **data skills** projects should be trialled for school leaver examinations across the four nations of the UK.

In Section 5, we proposed a set of **data skills** for the end of primary, lower secondary, adding to those proposed by Smith et al. (2023) for school leavers.

Recommendation 1d (Policy makers, JMC and other MDE stakeholders)

The proposed **data skills** should be used to stimulate discussion about what skills are included in national curricula in each of the four nations of the UK to develop meaningful **data education**.

*2 What is the evidence base on effective practices in developing **data skills** in mathematics (and elsewhere in the curriculum)? Are these small-scale insights and/or examples of effective system-wide **data skills** development?*

In Section 3, we considered current practices in **data education**. We make the following recommendations:

Recommendation 2a (Policy makers)

Ensure that foundational **data skills** are specified coherently across all subject curricula, for all phases from the beginning of primary education, using consistent terminology for each phase of study.

Recommendation 2b (Policy makers)

All teachers of all subjects should have the opportunity to participate in professional learning to broaden teachers' own knowledge of this growing discipline and to build their confidence in teaching approaches including their use of a range of appropriate digital technologies for working with data. By bringing together teachers across disciplines, and across phases, there is the potential to develop a coherent view of how cross-disciplinary **data skills** progress over time.

Further research

We recommend further research in these two areas:

Recommendation 2c

Understanding how learners develop **data skills** in national contexts in early years, primary and secondary settings and develop effective teaching practices to further develop their skills.

Recommendation 2d

Understanding barriers and possible solutions to successful **data education** policy implementation in the context the UK.

*3 What **resources** do we need to achieve meaningful improvement **in data education** across the four nations?*

In Section 3, we identified resources to achieve meaningful improvement in **data education**. We make the following recommendations:

Recommendation 3a (JMC member bodies and other MDE stakeholders)

Identify and appoint a single *Ambassador for Data Education* across all four nations of the UK, such as in the US, to advocate for high-quality, meaningful data education.

Recommendation 3b (JMC member bodies and other MDE stakeholders)

Create a single online repository for curated quality-assured and freely available **data education** curriculum materials and signposts to high-quality professional learning opportunities.

In Section 1, we referred to the role of the UK Statistics Authority and its five-year strategy plan that end in 2025. We make the following recommendation.

Recommendation 3c (UK Statistics Authority with JMC member bodies and other MDE stakeholders)

The UK Statistics Authority include an objective for engaging with stakeholders in **data education** in the next five-year strategy plan, to assist in emphasising the importance of **data education** across the four nations.

Further research

We recommend further research in this area:

Recommendation 3d (Policy makers, JMC member bodies and other MDE stakeholders)

Conduct state-of-the-nation reviews, by getting directly into schools, observing **data education** in practice and discussing with learners, teachers and school leaders how **data education** is incorporated into school activities.

References

Adams, G., & Boylan, M. (2023). *Landscaping Mathematics Education Policy: Horizon scanning of international policy initiatives*. Retrieved from London:

Aziz, A., & Rosli, R. (2021). *A systematic literature review on developing students' statistical literacy skills.* Paper presented at the Journal of Physics: Conference Series.

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D. A. (2020). Pre-K–12 guidelines for assessment and instruction in statistics education II (GAISE II): A framework for statistics and data science education. *American Statistical Associatio*.

Boylan, M., & Adams, G. (2023). *Landscaping Mathematics Education Policy: Landscaping national mathematics education policy*. Retrieved from https://royalsociety.org/-/media/policy/projects/maths-futures/landscaping-national-mathematics-education-policy.pdf

Callingham, R., & Watson, J. (2024). Statistics education research at the school level in Australia and New Zealand: A 30-year journey. *Mathematics Education Research Journal, 36*(Suppl 1), 91-122.

Chamberlain, M., Darr, C., Hipkins, R., McKinley, S., Murphy, H., & Sinnema, C. (2021). New Zealand curriculum refresh: progressions approach. *NZCER. https://www. nzcer. org. nz/research/publications/new-zealand-curriculum-refreshprogressions-approach*.

Crisan, C., Geraniou, E., & Hodgen, J. (2023). Educational technologies in Mathematics education. The Royal Society. https://royalsociety.org/-/media/policy/projects/maths-futures/educational-technology-mathematics-education.pdf

Dawn, N. K. E. (2011). Towards mathematical literacy in the 21st century: Perspectives from Singapore. *Southeast Asian Mathematics Education Journal, 1*(1), 19-29.

Department for Digital Culture Media and Sport. (2021). *Quantifying the UK data skills gap : full report : published 18 May 2021*. Retrieved from https://www.gov.uk/government/publications/quantifying-the-uk-data-skills-gap/quantifying-the-uk-data-skills-gap-full-report

Fearns, J., Harriss, L., & Lally, C. (2023). *Data science skills in the UK workforce*. Retrieved from https://post.parliament.uk/research-briefings/post-pn-0697/

Forbes, S. (2014). The Coming of Age of Statistics Education in New Zealand, and its Influence Internationally. *Journal of Statistics Education, 22*(2), null-null. doi:10.1080/10691898.2014.11889699

Forsythe, S., & Smith, C. (2020). *Mathematics Education 5-18 in the Four Nations: A comparative analysis*. Retrieved from

Franklin, C. (2013). Common core state standards and the future of teacher preparation in statistics. *The Mathematics Educator, 22*(2).

Franklin, C. (2014). *The statistical education of teachers (SET): An American Statistical Association policy document*. Paper presented at the Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics.

Ingram, J., Stiff, J., Cadwallader, S., Lee, G., & Kayton, H. (2023a). *PISA 2022: National report for England* (1838705228). Retrieved from https://www.gov.uk/government/publications/pisa-2022-national-report-for-england

Ingram, J., Stiff, J., Cadwallader, S., Lee, G., & Kayton, H. (2023b). *PISA 2022: National report for Northern Ireland*. Retrieved from https://www.education-ni.gov.uk/publications/pisa-2022-national-report-northern-ireland

Ingram, J., Stiff, J., Cadwallader, S., Lee, G., & Kayton, H. (2023c). *PISA 2022: National report for Wales*. Retrieved from https://www.gov.wales/achievement-15-year-olds-program-international-student-assessment-pisa-national-report-2022

Kadijevich, D. M., Stephens, M., & Rafiepour, A. (2023). Emergence of computational/algorithmic thinking and its impact on the mathematics curriculum. In *Mathematics Curriculum*

Reforms Around the World: The 24th ICMI Study (pp. 375-388): Springer International Publishing Cham.
Kaur, B. (2019). *Evolution of Singapore's school mathematics curriculum*: Springer.
McMaster, J. (2022). *Statistical Literacy*. Retrieved from https://osr.statisticsauthority.gov.uk/wp-content/uploads/2023/02/OSR_Statistical_Literacy_Research_Report.pdf
National Academies of Sciences. (2018). *Data science for undergraduates: Opportunities and options*: National Academies Press.
National Academies of Sciences, Engineering and Medicine. (2022). *Foundations of Data Science for Students in Grades K-12: Proceedings of a Workshop*.
Noyes, A., Clark-Wilson, A., Hodgen, J., & Button, T. (2023). Mathematics education and digital technology. https://discovery.ucl.ac.uk/id/eprint/10176790/1/JMC%20Digitech%20Report%20July%202023.pdf
OECD. (2020a). *Curriculum (re)design A series of thematic reports from the OECD Education 2030 project OVERVIEW*. Retrieved from
OECD. (2020b). *Curriculum Overload*.
OECD. (2020c). *What Students Learn Matters Towards a 21st Century Curriculum*: OECD Publishing.
Pittard, V. (2018). The integration of data science in the primary and secondary curriculum. *Final Report to the Royal Society Advisory Committee on Mathematics Education*. https://royalsociety.org/news-resources/publications/2018/integration-data-science-in-primary-secondary-curriculum/
Royal Geographical Society. (2019). *Data Skills in Geography Project Review*. Retrieved from https://www.rgs.org/schools/projects-and-partnerships/data-skills-in-geography/
Royal Statistical Society. (2024). *Key recommendations for the statistics curriculum in the UK*. Retrieved from https://rss.org.uk/RSS/media/File-library/Policy/2024/Cover_Recommendations-for-statistics-curriculum-in-the-UK.pdf
Saldanha, L., & Hatfield, N. (2021). Students Conceptualizing the Box Plot as a Tool for Structuring Quantitative Data: a Design Experiment Using TinkerPlots. *Canadian journal of science, mathematics and technology education, 21*(4), 758-782. doi:10.1007/s42330-021-00184-0
Schaffar, B., & Wolff, L.-A. (2024). Phenomenon-based learning in Finland: a critical overview of its historical and philosophical roots *Cogent Education, 11*(1), 2309733. doi:10.1080/2331186X.2024.2309733
Scottish Government. (2023). *Programme for International Student Assessment (PISA2022): Highlights from Scotland's Results*. Retrieved from https://www.gov.scot/publications/programme-international-student-assessment-pisa-2022-highlights-scotlands-results/documents/
Seow, P., Looi, C.-K., How, M.-L., Wadhwa, B., & Wu, L.-K. (2019). Educational policy and implementation of computational thinking and programming: Case study of Singapore. *Computational thinking education*, 345-361.
Smith, C., & Forsythe, S. (2020). Mathematics Education 5-18 In The Four Nations: A comparative analysis. https://oro.open.ac.uk/73134/
Smith, C., Kathotia, V., Ward-Penny, R., Howson, O., & Wermelinger, M. (2023). *Mathematical and Data Literacy: Competencies and curriculum implications at the intersection of mathematics, data science, statistics and computing*. Retrieved from https://royalsociety.org/~/media/policy/projects/maths-futures/intersection-mathematics-data-statistics-computing.pdf
Taguma, M., Makowiecki, K., & Gabriel, F. (2023). OECD Learning Compass 2030: Implications for Mathematics Curricula. In *Mathematics Curriculum Reforms Around the World: The 24th ICMI Study* (pp. 479-509): Springer International Publishing Cham.

Royal Geographical Society. (2019). *Data Skills in Geography Project Review*. Retrieved from https://www.rgs.org/schools/projects-and-partnerships/data-skills-in-geography/
The Royal Society. (2019). *Dynamics of data science skills*. Retrieved from https://royalsociety.org/news-resources/projects/dynamics-of-data-science/
The Royal Society (2023). A new approach to mathematics and data education: A discussion paper from the Mathematical Futures Board of The Royal Society's Advisory Committee on Mathematics Education. https://royalsociety.org/-/media/policy/projects/maths-futures/new-approach-to-mathematics-and-data-education.pdf
The Royal Society. (2024). *A new approach to mathematical and data education*. Retrieved from https://royalsociety.org/-/media/policy/projects/maths-futures/mathematical-and-data-education-policy-report.pdf
The Royal Society & STEM Learning. (2019). *Data skills for all: Note of discussions at a Royal Society and STEM Learning workshop on data science skills*. Retrieved from https://royalsociety.org/news-resources/projects/dynamics-of-data-science/
UK Statistics Authority. (2020). *Statistics for the public good: UK Statistics Authority Five Year Strategy 2020-2025*. Retrieved from https://uksa.statisticsauthority.gov.uk/statistics-for-the-public-good/
Vahey, P., Rafanan, K., Swan, K., Kratcoski, A., Stanford, T., & Patton, C. (2012). Thinking with Data: A Cross-Disciplinary Approach to Teaching Data Literacy and Proportionality. *Educational studies in mathematics, 81*. doi:10.1007/s10649-012-9392-z
Wilkerson, M., Finzer, W., Erickson, T., & Hernandez, D. (2021). Reflective Data Storytelling for Youth: The CODAP Story Builder.
Wolff, A., Gooch, D., Cavero Montaner, J. J., Rashid, U., & Kortuem, G. (2016). Creating an Understanding of Data Literacy for a Data-driven Society. *The journal of community informatics, 12*(3), 9-26. doi:10.15353/joci.v12i3.3275

Appendices

Appendix 1: Expectations for **data skills** at the end of primary and lower Secondary in each of the four nations

The tables on the next page draw from various statutory documents to summarise the current expectations for data skills in each of the four nations. These key documents are listed below along with details of learner ages.

England

End of primary age: 11
Key document: End of KS2 Assessment Framework[27]
Lower secondary age: 14
Key document: Mathematics programmes of study: key stage 3 [28]
End of secondary age: 16
Key document: Mathematics GCSE subject content and assessment objectives[29] (Foundation tier)

Northern Ireland

End of primary age: 11
Key document: Levels of progression for using mathematics primary levels[30] (level 4)
End of lower secondary age: 14
Key document: Levels of progression for using mathematics primary levels[30] (level 5)
End of secondary age: 16
Key document: CCEA GCSE Specification in Mathematics[31]

Scotland

End of primary age: 11/12
Key document: Benchmarks Numeracy and Mathematics[32] (second level)
End of lower secondary age: 13/14/15
Key document: Benchmarks Numeracy and Mathematics[33] (third and fourth level)
End of secondary age: 16/17/18
Key document: Core skills framework[34] (SCQF4 and SCQF5)

Wales

End of primary age: 11
Key document: Programme of Study for Mathematics[35](Year 6)
Lower secondary age: 14
Key document: Programme of Study for Mathematics: (Year 9)
End of secondary age: 16
Key document: WJEC GCSE in mathematics[36] (Foundation tier)

---

27 https://www.gov.uk/government/publications/key-stage-2-mathematics-test-framework
28 https://assets.publishing.service.gov.uk/media/5a7c1408e5274a1f5cc75a68/SECONDARY_national_curriculum_-_Mathematics.pdf
29 https://assets.publishing.service.gov.uk/media/5a7cb5b040f0b6629523b52c/GCSE_mathematics_subject_content_and_assessment_objectives.pdf
30 https://ccea.org.uk/learning-resources/levels-progression-using-mathematics
31 https://ccea.org.uk/downloads/docs/Specifications/GCSE/GCSE%20Mathematics%20%282017%29/GCSE%20Mathematics%20%282017%29-specification-Standard.pdf
32 https://education.gov.scot/media/s5edgtvx/numeracyandmathematicsbenchmarks.pdf
33 https://education.gov.scot/media/s5edgtvx/numeracyandmathematicsbenchmarks.pdf
34 https://www.sqa.org.uk/files_ccc/24CombinedCoreSkillsFrameworkV1.pdf
35 https://hwb.gov.wales/api/storage/e2aad4fc-faae-4054-a382-242d08ee2021/programme-of-study-for-mathematics.pdf
36 https://www.wjec.co.uk/media/u3rbzz1p/wjec-gcse-maths-spec-from-2015-r-e.pdf

	England	Northern Ireland	Scotland	Wales
End of primary expectations	interpret and present data using bar charts, pictograms and tables solve one-step and two-step questions [e.g. 'How many more?' and 'How many fewer?'] using information presented in scaled bar charts, pictograms and tables interpret and present discrete and continuous data using appropriate graphical methods, including bar charts and time graphs solve comparison, sum and difference problems using information presented in bar charts, pictograms, tables and other graphs complete, read and interpret information in tables, including timetables solve comparison, sum and difference problems using information presented in a line graph interpret and construct pie charts and line graphs and use these to solve problems calculate and interpret the mean as an average 2016 Key stage 2 mathematics test framework	collect and record relevant data for a given activity draw and label pictograms and bar charts read and interpret information from tables, pictograms, diagrams, lists, bar charts, simple pie charts and databases. collect, group, record and present data with given class intervals; present and interpret data using a range of graphs, tables, diagrams, spreadsheets and databases; understand and use the language of probability LOP Level 3 and 4	I have explored a variety of ways in which data is presented and can ask and answer questions about the information it contains. I have used a range of ways to collect information and can sort it in a logical, organised and imaginative way using my own and others' criteria. Using technology and other methods, I can display data simply, clearly and accurately by creating tables, charts and diagrams, using simple labelling and scale. I can use appropriate vocabulary to describe the likelihood of events occurring, using the knowledge and experiences of myself and others to guide me. Having discussed the variety of ways and range of media used to present data, I can interpret and draw conclusions from the information displayed, recognising that the presentation may be misleading. I have carried out investigations and surveys, devising and using a variety of methods to gather information and have worked with others to collate, organise and communicate the results in an appropriate way. I can display data in a clear way using a suitable scale, by choosing appropriately from an extended range of tables, charts, diagrams and graphs, making effective use of technology. I can conduct simple experiments involving chance and communicate my predictions and findings using the vocabulary of probability.	They collect discrete data, group data where appropriate, and use the mode and median as characteristics of a set of data. They draw and interpret frequency diagrams and construct and interpret simple line graphs. They understand and use simple vocabulary associated with probability. *(Level 4 statement)* represent data using: lists, tally charts, tables, diagrams and frequency tables bar charts, grouped data charts, line graphs and conversion graphs extract and interpret information from an increasing range of diagrams, timetables and graphs (including pie charts) use mean, median, mode and range to describe a data set use numbers to describe the likelihood of an event, e.g. a one-in-six chance v recognise that some events are equally likely identify the outcomes of simple events, e.g. flipping a coin, rolling a dice. *Y6 POS*

	England	Northern Ireland	Scotland	Wales
End of lower secondary	describe, interpret and compare observed distributions of a single variable through: appropriate graphical representation involving discrete, continuous and grouped data; and appropriate measures of central tendency (mean, mode, median) and spread (range, consideration of outliers) construct and interpret appropriate tables, charts, and diagrams, including frequency tables, bar charts, pie charts, and pictograms for categorical data, and vertical line (or bar) charts for ungrouped and grouped numerical data describe simple mathematical relationships between two variables (bivariate data) in observational and experimental contexts and illustrate using scatter graphs. KS3 Mathematics POS	collect, organise, record and represent data design and use a data collection sheet construct, label and interpret a range of graphs, tables, diagrams, spreadsheets and databases understand, calculate and use mean and range place events in order of likelihood LOP Level 5	I can work collaboratively, making appropriate use of technology, to source information presented in a range of ways, interpret what it conveys and discuss whether I believe the information to be robust, vague or misleading. When analysing information or collecting data of my own, I can use my understanding of how bias may arise and how sample size can affect precision, to ensure that the data allows for fair conclusions to be drawn. I can display data in a clear way using a suitable scale, by choosing appropriately from an extended range of tables, charts, diagrams and graphs, making effective use of technology. I can find the probability of a simple event happening and explain why the consequences of the event, as well as its probability, should be considered when making choices. I can evaluate and interpret raw and graphical data using a variety of methods, comment on relationships I observe within the data and communicate my findings to others.	They use the mean of discrete data and compare two simple distributions. They interpret graphs, diagrams and pie charts. They use the probability scale from 0 to 1, and appreciate that different outcomes may result from repeating an experiment. *Level 5 statement* use the sum of all probabilities is 1 – simple cases, e.g. rolling a dice P (not 6) recognise that practice is different from theory and that repeated experiments may give different results understand that reliability/stability increases with a greater number of trials construct a sample space diagram and a two way table. Y9 Programme of study

End of lower secondary (cont.)			In order to compare numerical information in real-life contexts, I can find the mean, median, mode and range of sets of numbers, decide which type of average is most appropriate to use and discuss how using an alternative type of average could be misleading. I can select appropriately from a wide range of tables, charts, diagrams and graphs when displaying discrete, continuous or grouped data, clearly communicating the significant features of the data. By applying my understanding of probability, I can determine how many times I expect an event to occur, and use this information to make predictions, risk assessment, informed choices and decisions. Third and fourth level Numeracy and mathematics	

	England and Wales	Northern Ireland	Scotland
School leaver expectations	record describe and analyse the frequency of outcomes of probability experiments using tables and frequency trees apply ideas of randomness, fairness and equally likely events to calculate expected outcomes of multiple future experiments relate relative expected frequencies to theoretical probability, using appropriate language and the 0 - 1 probability scale apply the property that the probabilities of an exhaustive set of outcomes sum to one; apply the property that the probabilities of an exhaustive set of mutually exclusive events sum to one <u>understand that empirical unbiased samples tend towards theoretical probability distributions, with increasing sample size</u> enumerate sets and combinations of sets systematically, using tables, grids, Venn diagrams <u>and tree diagrams</u> construct theoretical possibility spaces for single and combined experiments with equally likely outcomes and use these to calculate theoretical probabilities <u>calculate the probability of independent and dependent combined events, including using tree diagrams and other representations, and know the underlying assumptions</u>	use 3 circle Venn diagrams to sort data estimate mean from a grouped frequency distribution identify the modal class and the median class from a grouped frequency distribution draw and/or use lines of best fit by eye, understanding what these lines represent draw conclusions from scatter diagrams use terms such as positive correlation, negative correlation and little or no correlation interpolate and extrapolate from data and know the dangers of doing so identify outliers; and appreciate that correlation does not imply causality. systematically list all outcomes for single events and for two successive events understand and use estimates or measures of probability from relative frequency compare experimental data and theoretical probabilities understand that increasing sample size generally leads to better estimates of probability. Taken from modular and completion test specifications.	Using Graphical Information SCQF 4 extract information from a straightforward table, graph, chart, or diagram use appropriate graphical forms to convey particular types of information communicate information in straightforward tables, graphs, charts, or diagrams Using Graphical Information SCQF 5 interpret information from a table, graph, chart, or diagram use an appropriate form of table, graph, chart, or diagram, to communicate information

	infer properties of populations or distributions from a sample, whilst knowing the limitations of sampling interpret and construct tables, charts and diagrams, including frequency tables, bar charts, pie charts and pictograms for categorical data, vertical line charts for ungrouped discrete numerical data, tables and line graphs for time series data and know their appropriate use interpret, analyse and compare the distributions of data sets from univariate empirical distributions through: appropriate graphical representation involving discrete, continuous and grouped data, apply statistics to describe a population use and interpret scatter graphs of bivariate data; recognise correlation and know that it does not indicate causation; draw estimated lines of best fit; make predictions; interpolate and extrapolate apparent trends whilst knowing the dangers of so doing	construct and interpret a wide range of graphs and diagrams including frequency tables and diagrams, pictograms, bar charts, pie charts, line graphs, frequency trees and flow charts, recognising that graphs may be misleading; examine data to find patterns and exceptions; compare distributions and make inferences; and plot and interpret scatter diagrams and recognise correlation.	

Appendix 2 Key reports relevant to **data education** in the United Kingdon

Title author	Year	Key Points in relation to **data education** in the UK
The Integration of Data Science in the Primary and Secondary Curriculum (ACME)	2018	At primary, core knowledge and skills are developed mainly through the statistics strand of mathematics at Key Stage 2 and algorithms and use of digital technology in computing at Key Stages 1 & 2. Disjuncture between mathematics in science and attainment targets. At secondary statistics content is carefully sequenced to support conceptual progression. The GCSE contains relevant content at both Foundation and Higher tiers. Key Stage 3 computing contains substantial content relating to data science. Secondary science emphasises the application of mathematics and the centrality of the classic data cycle. This includes content on appraising claims based on data. Key stage 3 and 4 geography strongly emphasises the collection and analysis of quantitative data
Data Skills for All (The Royal Society and STEM Learning)	2019	Note of discussions at a Royal Society and STEM Learning workshop on data science skills Proposes a framework for developing data skills.
The Dynamics of data science (The Royal Society)	2019	Foundational training in **data skills** begins at school, and there is an opportunity and a need to consult broadly on a future curriculum that addresses the breadth of **data skills** across mathematics and science, the arts and humanities. Students need to be well informed about the ever-widening range of opportunities working with data. New hybrid forms of education in data science at school age and later, such as apprenticeships, are becoming more prevalent.
Data skills in Geography Project Report (Royal Geographical Society)	2019	Example of a cross disciplinary professional development project for GCSE.
Mathematics Education 5- 18 in the four nations: A comparative analysis. (JMC)	2020	Brief document that compares the mathematics education systems of the four nations by comparing, curriculum, significant assessment from 5 to 16, mathematics pathways 16-18 and professional learning.
Mathematics education and digital technology (JMC)	2023	Future reforms of statutory curricula or non-statutory guidance should explicitly address the use of digital technology in the teaching of mathematics and data analysis across all years. Well-designed support for the implementation of reforms should also be provided. Appropriate high-stakes assessment of digital technology skills is needed to ensure that such skills are embedded in, and developed throughout, the learning process. National bodies should ensure that skills for working with data are better assessed, through ongoing practical work or high-stakes examinations in which students have access to appropriate digital technologies.

		All professional development for teachers of mathematics should embed appropriate uses of digital technologies such that they become normalised, rather in the same way that digital technologies should be embedded in all teaching and learning.
Educational Technologies in Mathematics Education (The Royal Society)	2023	Make mathematics-specific digital technologies (DT) easily available and accessible in mainstream education including data analysis software. Continue to research and develop the implementation and application of mathematics-specific DT in mathematics education including data analysis software and AI tools. All mathematics teachers are required to engage in a minimum level of DT related PD. High stakes assessments and assessments in school mathematics should change to integrate the use of DT for mathematical purposes.
Mathematical and Data Literacy Competencies and curriculum implications at the intersection of mathematics, data science, statistics and computing (The Royal Society)	2023	investigate the intersection of mathematics, data science, statistics and computing. Proposes 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.
Landscaping Mathematics Education Policy: Horizon scanning of international policy initiatives (The Royal Society)	2023	Provides an overview of recent developments in mathematical, statistical, and computational thinking and data literacy, informed by views of international mathematics education experts. The guiding aim was to use international comparisons and horizon scanning to identify high-performing practices, policy initiatives, and future directions of travel in mathematics education.
A new approach to mathematics and data education (The Royal Society)	2023	The purpose of this discussion paper is to air the underlying principles and possibilities, to gauge the appetite for change, and to explore how we could move from strategic vision to implementation with regards to Mathematical and data education (MDE)
Landscaping Mathematics Education Policy: Landscaping national mathematics education policy	2023	The review focuses on mathematics education policy and draws on broader education policy where this is relevant and/or impacts on mathematics education. The amount of probability and data handling has reduced in England National Curricula since 1990. The place of data handling and statistics education in the mathematics curriculum, together with concerns over teachers' knowledge and pedagogical skills to teach them, were issues at the start of the 1980s.
A New Approach to Mathematical and Data Education (The Royal Society)	2024	This report explores how the educational system can respond to the explosion in data availability and the way that the exploitation of data through technologies is pervading every

		aspect of life. The scope of mathematical education needs to change from 'mathematics' to what we have called mathematical and **data education** (MDE); a combination of mathematics, statistics and data science, underpinned by computational tools.
Key recommendations for the statistics curriculum in the UK (The Royal Statistical Society)	2024	A set of recommendations based on a comparison of the US and New Zealand and UK statistics curricula: A stronger emphasis on the use of information communication technologies A stronger emphasis needs to be placed on the use of real-world contexts Better connections need to be made with statistics and other disciplines Options for GCSE examination Clearer connections to the PPDAC cycle, starting at primary Using real world examples should be encouraged and adopted at the primary level and embedded and used throughout secondary level Carrying out calculations by hand should continue to be phased out, as schools in the UK continue to use technology More time needs to be focused on the visualisation of data in the teaching of statistics More opportunities need to be made available for young people to engage with the production of statistical reports, that include quantitative reasoning skills Teachers need more support and training to teach statistics successfully