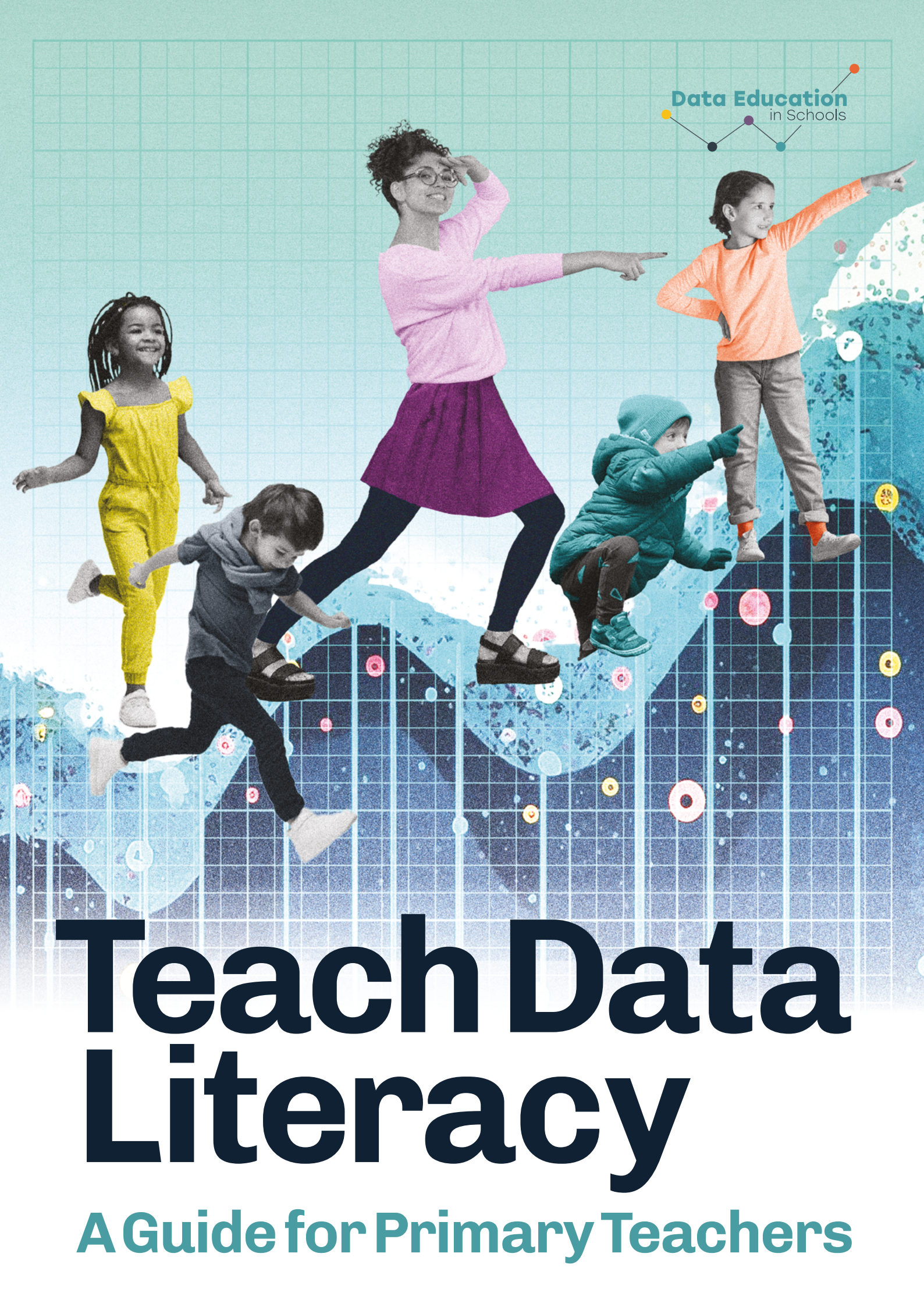


Teach Data Literacy

A Guide for Primary Teachers

Contents

The Data Education in Schools project is part of the Edinburgh and South East Scotland City Region Deal Data Skills Programme, funded by the Scottish and UK Governments. The Data Skills Programme brings together industry, universities, colleges, schools and others to provide routes into data or digital careers.

This handbook is part of the Data Education in Schools' commitment to developing an interdisciplinary data education curriculum for Scotland and a set of engaging real world data science teaching materials for primary and secondary school teachers. By Judy Robertson, Holly Linklater, Kate Farrell, Jasmeen Kanwal, Serdar Abaci, Claire Sowton and the Data Education in Schools team. Graphic design by Elspeth Maxwell. Edinburgh, 2023.

www.dataschools.education

Teach Data Literacy

A Guide for Primary Teachers

Data literacy is the set of skills and concepts which people need to understand, interpret and make decisions based on the data they encounter in the world around them.

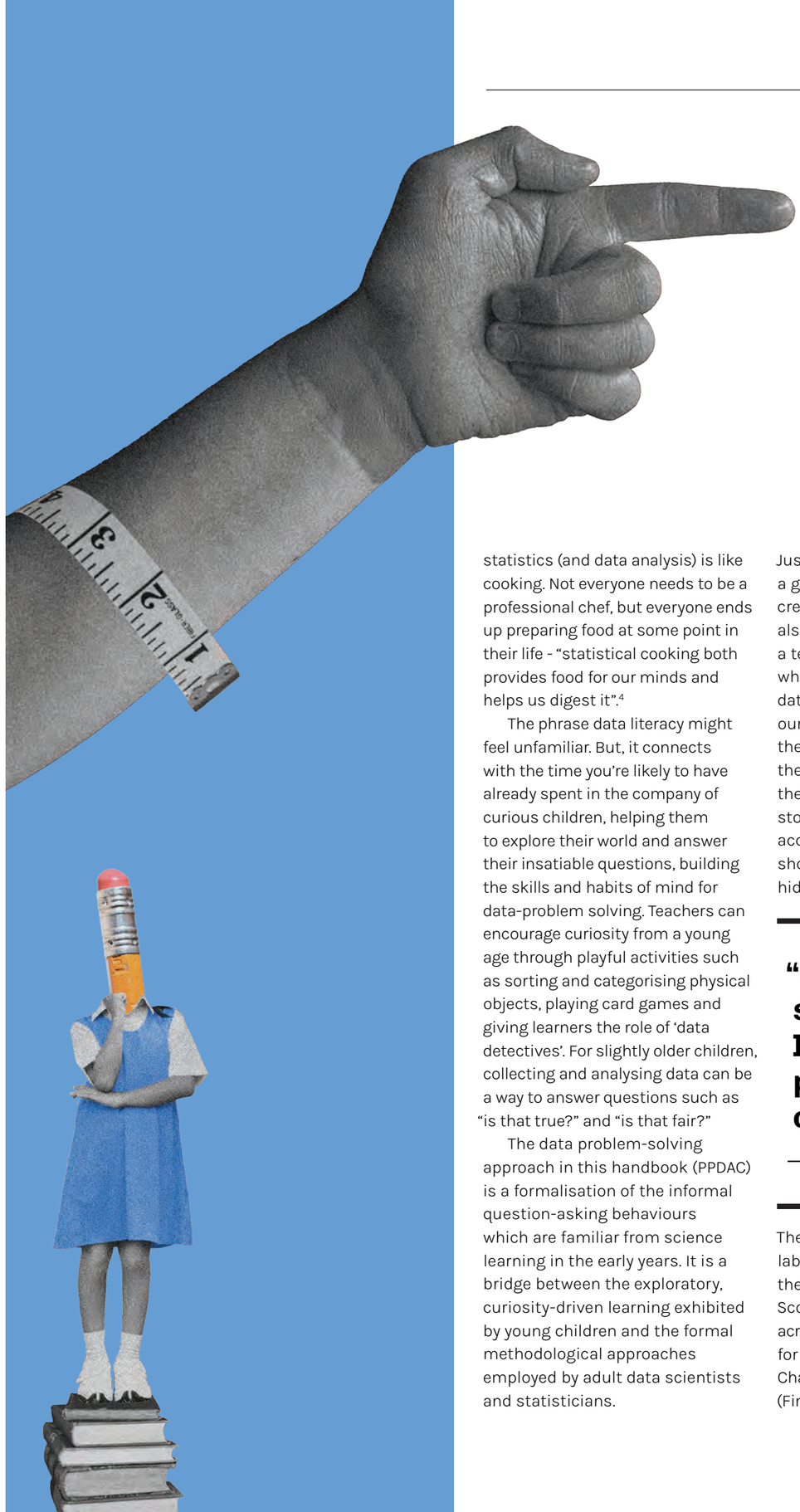
The skill of critically analysing data to solve problems has always been necessary, but with advances in digital technologies which collect vast amounts of data, the skill has become simultaneously more important and more complex.

In addition to the traditional "number crunching" which we might associate with handling data, learners should be aware of the ethical and societal dimensions of how data is used by technology companies. They should know how data is generated by smart phones and the websites they use, why companies collect and make money from it, and how it impacts their privacy. In an online world full of [mis]information, they need the data skills to tell the difference between evidence and false claims. And they should have opportunities to use data to understand and address challenges in their local community.

Data literacy is considered as one of the core foundations for the OECD 2030 Learning Compass, alongside digital literacy, physical and mental health, and social and emotional skills, and is recognised as "essential for thriving in the 21st century".[1]

Data literacy "*...is 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... [which] include the ability to select, clean, analyse, visualise, critique and interpret data, as well as to communicate stories from data and use data as part of a design process*".[2]

There are many exciting and well-paid career opportunities available to young people who can analyse data fluently – as data scientists, digital marketers, data journalists, machine learning engineers, sports performance coaches and many more futuristic-sounding roles.[3] Even for young people who don't go on to use data in their jobs, data literacy is an important aspect of being an educated citizen in a digital world. The statistics educator Chris Wild uses the metaphor that



> **Data literacy is a key skill for the modern world.**
>
> *– David Spiegelhalter*[7]

statistics (and data analysis) is like cooking. Not everyone needs to be a professional chef, but everyone ends up preparing food at some point in their life - "statistical cooking both provides food for our minds and helps us digest it".[4]

The phrase data literacy might feel unfamiliar. But, it connects with the time you're likely to have already spent in the company of curious children, helping them to explore their world and answer their insatiable questions, building the skills and habits of mind for data-problem solving. Teachers can encourage curiosity from a young age through playful activities such as sorting and categorising physical objects, playing card games and giving learners the role of 'data detectives'. For slightly older children, collecting and analysing data can be a way to answer questions such as "is that true?" and "is that fair?"

The data problem-solving approach in this handbook (PPDAC) is a formalisation of the informal question-asking behaviours which are familiar from science learning in the early years. It is a bridge between the exploratory, curiosity-driven learning exhibited by young children and the formal methodological approaches employed by adult data scientists and statisticians.

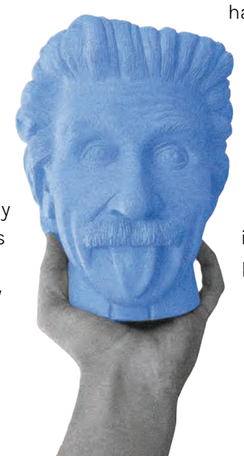
Just as curiosity is key to being a good data problem-solver, creativity and communication are also needed. "Data storytelling" is a technique used by designers who create beautiful displays of data to catch our eyes and change our minds. It's the art of distilling the most important points from the numbers and communicating them clearly. Of course, data storytelling should honour accuracy and truth, and learners should be able to spot "fake news" hidden within a fancy infographic.

> **"I have no special talents. I am only passionately curious."**
>
> *— Albert Einstein*[5]

There is no official area or topic labelled "data literacy" within the Curriculum for Excellence in Scotland, but it is embedded firmly across the curriculum in outcomes for Maths (Data and Analysis, Chance and Uncertainty), Literacy (Finding and Using Information), Technologies (Computing Science and Digital Literacy) and Social Studies (People in Society, Economy and Business).[6]

The purpose of this handbook is to draw together aspects of data literacy to make it easier for teachers to understand the data problem-solving cycle, and how learners' skills and understanding of key concepts progress. There is a collection of resources from around the world appropriate for different stages of learning in the Resources and Activities section, as well as examples of the prompts and conversation starters that teachers might use to guide and reinforce the data literacy ideas.

Data-literate children and young people have the skills to fluently collect and use data to understand the world more deeply and take informed action to improve their lives and community. This handbook is intended to support teachers in weaving data literacy into the curriculum in thought-provoking and meaningful ways. •



Data Literacy Concepts

> *"The biggest needs for every statistical cook, from the person who simply heats up TV dinners to the Michelin-star chef, are needs for conceptual understanding."* [8]

Learners should explore three main statistical concepts at primary school: data types, variability and probability.

Data literacy overlaps with statistical literacy, and they share important concepts which will be familiar from the data and analysis, chance and uncertainty strands of the maths curriculum. There are three main statistical concepts which learners should informally explore at primary school: data types, variability and probability.

Knowing about the "big ideas" in statistics and how conceptual understanding progresses is useful for teachers when planning for and supporting learning (diagnosing learners' difficulties). It isn't necessary for learners to use statistical vocabulary.

Similarly, supporting the ability to think about data and draw valid conclusions is more valuable than rehearsing mechanical calculations for statistical formulae or tests. It makes sense to use software tools for calculations and drawing graphs to give the learners more time to think about what the results mean, particularly when working with larger datasets. •

Data Types & Variables

> *A dataset is a collection of measurements or records about a group of related items or people. A weather dataset might consist of a collection of temperature recordings from cities in the UK, and an opinion poll is a dataset of a group of people's views of the same topic.*

In statistics, a variable[9] is a characteristic of interest that you measure, record, and analyse for all of the items in a dataset. A variable is an attribute that describes a person, place, thing, or idea. The value of the variable will be different for different items/records in the dataset which you measure – it will vary.

There are two main types of variables which learners will encounter in primary schools: numerical and categorical. Numerical variables record amounts, quantities, or measurements in the form of numbers, while categorical variables record categories or groups. Examples of numerical variables include height, income, age and temperature. Teachers can draw more experienced learners' attention to the fact that numerical variables can have different ranges of expected values. For example, temperature could be negative when it is cold outside, but height would not be zero or less.

Categorical variables are used to record ideas or characteristics which can be categorised but don't have an obvious numerical value e.g. colour, political party, or gender. It is possible to count how many items in the dataset belong in each category but it doesn't make sense to do arithmetic on the categories. For example, I can count that there are 5 people with brown eyes, 3 people with green eyes and 2 people with blue eyes. I can even say that a total of 5 people have green eyes or blue eyes. But if I tried to add up the categories themselves (e.g. blue+green), it would have no meaning.

Developing an understanding of different types of data and how they can be measured by taking part in hands-on data-problem-solving projects will lay the foundation for learners to move on to more formal statistical analysis in later stages of secondary school. •

There are two main types of variables which learners will encounter in primary schools: numerical and categorical.

Variability

Variability is a key concept for statistical or data literacy. Statistics is the business of "understanding, explaining, and quantifying variability in the data within the given context".[10]

Variability is about recognising that even though not everything is the same, it is still possible to find patterns in the variation which help us to understand the world better. Not everyone in my class has the same favourite sweetie, but if I do a survey of their preferences, I can say something useful about which are the most popular sweets within this particular group of people (useful when planning a party).

Because not every baby is born with the same height and weight, a dataset of heights and weights of a large number of babies is useful when midwives are checking that any individual baby is growing in a healthy way (this might be familiar to parents in Scotland from the little red books with growth charts issued by the NHS).

Variability will occur in any dataset; the purpose of statistics is to try to find out how much of that variability can be explained by factors which we know about and can measure, how much could be caused by unknown factors and how much variance occurs just by random chance.

At first, it's useful to focus learners' attention on describing the variability of a single variable within just one group. For example, "This sunflower is taller than me. This sunflower is shorter than me. This sunflower is the same height as me." Then: "the sunflowers in this flowerbed are all between 50cm and 200cm tall, and the most common height is 150cm". This is a numerical variable of the sunflower height, measured in centimetres.

Once learners master describing the variability for one variable, they could move on the compare the variability between two (or more) groups. They could compare the variability of heights of sunflowers which were grown in sunshine or in the shade and make comparisons. Younger learners might begin by talking about heights comparatively without measuring anything; they may be able to compare variability before they learn how to measure: "The sunflowers in the sunny flower bed were bigger than the ones in the shade. The tallest of the plants in the sunny flowerbed was the same size as our teacher, but the tallest plant in the shady bed was only as tall as Sam who is the smallest in our class".

They could then progress to using measurements and writing the results in a more formal way: "The sunflowers in the sunny flowerbed tended to grow taller than the ones in the shady flowerbed. I noticed that the maximum height of the shady sunflowers was 140cm and the minimum height for this group was 30cm. This compares to the tallest sunflower in the sunny flowerbed which was 200cm and the smallest which was 50cm. The most common

"Younger learners might begin by talking about heights comparatively... they may be able to compare variability *before* they learn how to measure.

> **Describing variability in a dataset is often a clear way to communicate statistical findings to help people make decisions.**

height was 100cm for the shady sunflowers, but it was 150cm for sunny ones. This makes me think that sunshine is important for growing tall sunflowers."

Another way of extending the concept of variability is to measure the same variable on multiple occasions, to compare how it changes over time. With the sunflower example, this could be recording the height of the sunflowers in a flowerbed at the end of every week and spotting trends in how much they had grown: "I noticed that every sunflower plant grew at least 2 cm this week. The biggest change was 8cm but it was most common for them to grow about 3 cm. The plants are growing more quickly now because last week I noticed that only three of the plants grew as much as 2cm" (or, if measuring with Duplo, "I noticed that every sunflower grew more than one brick taller this week.")

More experienced learners can explore the associations between two variables. If variables A and B are associated then as variable A changes, so does variable B. Height and weight are often associated: the taller a dinosaur was, the heavier it would tend to be (and small dinosaurs tended to be lighter). For some variables, the association is negative, so a higher value of variable A tends to correspond to a lower value of variable B (e.g. the number of fillings you have may be negatively associated with the number of toothpaste tubes you get through a year).

Describing variability in a dataset is often a clear way to communicate statistical findings to help people make decisions. Even for researchers and statisticians, calculating and communicating clear descriptive statistics such as average and variance is a good starting point before conducting more complex statistical analyses. •

Probability and Chance

Curriculum for Excellence suggests that in primary school, learners will be describing or comparing datasets, or investigating the relationship between variables.

At later stages of their learning (senior phase of secondary school or above) they may begin to use data to formally predict outcomes. This requires an understanding of probability, the foundations of which can be laid in primary school. Learners can informally discuss how likely or unlikely events might be using everyday language. ("Sunflowers planted in the sunny bed will probably grow taller than sunflowers in the shady bed.")

The concept of randomness is important – learners can explore this in practical ways in simulations with coins or dice. This can lead to the idea of using random samples to avoid bias in experiments or surveys. For example, if you wanted to find out about the travel habits of the people who attend your school, but you didn't have time to ask everyone, you could sample the school population by asking twenty people to complete a survey. If you asked only the people who got off the school bus in the morning, the sample would be biased towards those who travel to school by bus. It might be better to randomly select people to complete the survey, for example by choosing names out of a hat or asking every 10th pupil entering the school door. This might give a more accurate summary of travel habits, as long as you didn't pick the moment when everyone coming off the bus was walking through the school door. (The progression of probability concepts in the Scottish Curriculum for Excellence can be found in the Chance and Uncertainty topic within Numeracy.) •

> **The concept of randomness is important – learners can explore this in practical ways in simulations with coins or dice.**

The Data Problem-solving Cycle

Data can be used to help us make better decisions. The underlying goal of this curriculum is to teach learners how to use data to solve real-world problems.

Figure 1 *(opposite page)* shows the data problem-solving cycle, also known as PPDAC[11] (**P**roblem, **P**lan, **D**ata, **A**nalysis and **C**onclusions).

The teacher can choose which stage of the problem-solving cycle they will focus on for each lesson – it isn't necessary to pack the full problem-solving cycle into one lesson. Sometimes it might be appropriate to spend time developing analysis skills (such as creating graphs) using a pre-prepared data set. Other lessons might focus more deeply on interpreting and critiquing conclusions about data made by others.

Although this handbook describes the PPDAC stages in sequence (Problem, Plan, Data, Analysis and Conclusions), problem-solving is often messier than that. For example, when exploring a data set – such as the temperature of the classroom over the course of »

PPDAC Stage	Skills Needed
Problem	• Discover a claim which needs to be investigated (e.g. news story) – wonder whether it can be true • Ask an interesting question
Plan	• Predict an answer • Find a data set and evaluate data quality • Design how to collect data
Data	• Collect a data set • Organise and structure data • Prepare data • Check that data is correct
Analysis	• Explore the data set • Visualise the data in tables and graphs • Analyse the data using simple descriptive statistics
Conclusions	• Answer the question • Explain what the data reveals • Decide whether a claim makes sense based on results • Communicate results to others • Make decisions and recommendations for change based on the data • Act on data to change own behaviour

Above: Table 1. Skills needed for the data problem-solving cycle
Opposite page: Figure 1. The data problem-solving cycle

a week – in the Analysis phase, the learner might notice an unusual data point which would cause them to return to the Data phase to check that the value was entered correctly. In the Conclusions phase, learners may start to think of many more questions that they didn't consider at the outset in the Problem phase and may go around the cycle one or more times. The skills required in each stage of PPDAC are shown in Table 1.

Children and young people will particularly benefit from project-based learning in which they take ownership of the whole problem-solving cycle to pursue a personal interest. However, learners who are still developing self-directed learning skills, and the meta-cognitive abilities to plan, monitor and evaluate their progress in a task will require support to manage this. •

> **Teachers can choose which stage of the problem-solving cycle they will focus on for each lesson. It isn't necessary to pack the full problem-solving cycle into one lesson.**

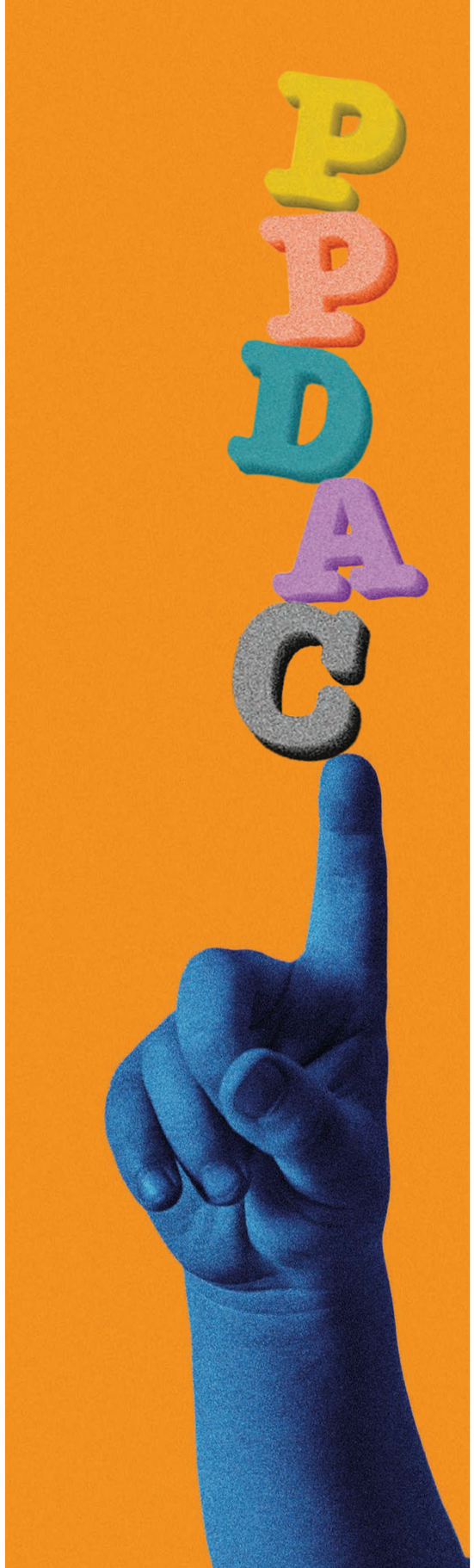
Real-world Problem-solving Contexts

The purpose of teaching data literacy is to develop skills so that learners can use data to understand the world and use it to inform their decision-making in everyday problems.

Learners should work with data sets gathered in the real world where possible. Initially, they are likely to be gathered within the class. Learning is most powerful when children and young people work with data which is personally meaningful, rather than focussing solely on manufactured examples. Learners should work with small data sets (relating to their class or schools) in the early stages. As they progress, they can use their skills to explore the larger publicly accessible data sets which reflect interesting aspects of the wider world, which has the advantage of fostering their curiosity and deepening their knowledge of the world in which they live. For example: how temperature/rainfall is changing across Scotland.[12] •

Learners should have opportunities to:

- ✓ ***Collect new data*** through surveys in the school or local community (e.g. surveys of how pupils travel to school or the amount of single use plastic used at lunchtime)
- ✓ ***Process and explore data*** gathered through sensors in the local environment (e.g. rain gauges or thermometers in the playground)
- ✓ ***Find, explore and analyse data sets*** relating to their personal interests (e.g. movies or trading cards)
- ✓ ***Explore and analyse existing public data sets*** published by governments, international organisations and researchers (e.g. World Health Organisation or Gapminder); (Level 2)
- ✓ ***Contribute to open-source data sets*** or take part in citizen science projects to analyse data (e.g. Zooniverse).
- ✓ ***Share datasets with other schools,*** community groups and scientific teams to be part of a wider research project (e.g. gathering and sharing data from national projects to document insect, flowers or birds across the country).



The Impact of Data on Personal Lives and Society

Data has always been used to make decisions in matters which affect our lives – for example, census data is used to allocate funding to local authorities.

> **It is important that children and young people are aware of their rights concerning personal data privacy and that they are included in discussions about how our society makes decisions.**

However, technology has changed and increased the usage of data in decision-making: decisions about which songs or books you might like to listen to or buy, whether to grant mortgage applications, short-listing for jobs, medical diagnoses, and crime prevention are made based on automatic data analysis by computer algorithms.

This raises complex ethical issues about privacy and fairness. It is important that children and young people are aware of their rights concerning personal data privacy and that they are included in discussions about how our society makes decisions. Ethics underlies the data problem-solving cycle: ethical awareness is required at each stage of the problem-solving cycle because unfairness can emerge unintentionally as a result of decisions about which people are represented in the data set, how data is collected, and how the results are analysed and shown. •

Learners should:

- ✓ Be aware of their rights and responsibilities about ***personal data privacy*** – The 5Rights campaign is a good place to start.[13]
- ✓ Understand how companies and other organisations gather information on individuals, and what ***decisions are made using this data.***
- ✓ Evaluate how conclusions that are drawn from data may be flawed if the ***data collection or analysis methods are biased.***
- ✓ Know how to avoid or minimise unfairness and ***bias in their own data problem-solving.***

"If we want to make the world add up, we need to ask questions; open-minded, genuine questions. Once we start asking them, we may find it is delightfully difficult to stop."

— Tim Harford[14]

Problem

At the first stage of the cycle (Problem), learners follow their curiosity to find a problem they would like to investigate. This could come from personal interest or be prompted by reading or watching something in the media.

Thinking of questions to answer is a highly creative activity. Some learners are particularly curious by nature and regularly seek answers to questions such as "How many…?"; "What is…?"; or "Why does…?". The "problem" might result in learners seeking to answer a question to find a fairer outcome for a particular group. Or to investigate whether a claim is wrong because it is a topic which affects their lives (such as examining the evidence on the impact of screen time on health) or simply because they enjoy arguing! Teachers should encourage learners to be sceptical and question whether claims they encounter are true, and think about the context in which they are made. •

Skill/Task	Classroom support	Progression
Ask an interesting question	With support, formulate a specific question which can be answered using data within a problem-solving approach.	Teacher will model this in Early and First levels, reducing scaffolding by Level 2. Questions which describe variables (such as simple frequency counts) are suitable for Early and Level 1. As learners get more experienced, they will pose questions which compare two groups, compare a group over time and find associations between variables.
Identify a claim which needs to be investigated (e.g. news story) – wonder whether it can be true…	Have the opportunity to learn in an information-rich environment (containing books, posters, reputable websites and documentaries) which encourages them to take an interest in the world around them and motivates them to find new problems to solve or new questions to answer. Have opportunities to read and watch claims based on data, e.g. from politicians, charities, companies and other organisations which may have an agenda.	All primary learners can benefit from this. See the resources section *(page 40)* for picture book suggestions for younger learners. As learners progress to Level 2, they will begin to examine claims online and in the media from politicians, charities etc. Fact checking websites can be useful here (e.g. fullfact.org).

Table 2. Skills progression for the Problem Stage of PPDAC

Asking good questions

Learners will often need help from their teacher to turn general curiosity into specific investigative questions which can be answered by collecting data within the time scale of the project, using the resources available.

It's important to strike a balance between inviting learners to pose questions that they find interesting, and guiding them towards selecting questions which will be feasible to answer.

Teachers can model how to frame good investigative questions with their learners. One way is to identify questions which will be hard to investigate is to consider how many variables they involve. It's quite common to brainstorm questions which contain many variables because often there are many possible causes for an effect. Learners might correctly suggest, for example, that the number of birds spotted in the playground could depend on how many trees are nearby, whether there are bird feeders and how many cats live in the area. But it would be difficult for primary learners to investigate data about all of these variables at one time. If the question involves more than two variables, it is usually hard to answer it satisfactorily by looking at just tables and graphs.

The teacher could guide learners towards picking one variable to focus on at a time, e.g. "How many birds can we count in the playground per day when we have a bird feeder in the playground, compared to the number of birds per day when there is no bird feeder?".

A good way to get learners started with posing questions for their investigations is to give them the prompts "I notice…" and "I wonder…" after watching a video or reading an article about a topic. "I notice" prompts encourage the learners to state facts they have already spotted in the data. "I wonder" prompts can encourage the learners to speculate about possible trends in the data which require more investigation. The wonderings can be descriptive "I wonder what TV shows people in primary 6 like to watch?" or comparative "I wonder if S1 students typically send more texts than P7 students? How many more?"

It is worth noting that "how much…?" (estimation) questions are more likely to give informative answers than "is there a difference…?" (dichotomous) questions. For example, it is not enough to know whether there is a difference between the grams of sugar eaten by teachers and pupils at the Sports Relief Day bake sale. The revealing information is how much of a difference there is in the sugar guzzling. If the teachers eat on average 500g and the pupils eat 501g then the difference probably doesn't matter very much: both groups ate more than is good for them in a single sitting. Furthermore, is it possible to measure the amount of sugar everyone consumes in cakes accurately enough to believe there is really a difference of 1g? On the other hand, if we work out that on average that teachers eat roughly double the amount of sugar compared to children during the bake sale (400g compared to 200g) we might want to restrict teachers to oatcakes only next year. •

> **Estimation questions are more likely to give informative answers than dichotomous questions.**

As a teacher, you can ask yourself:

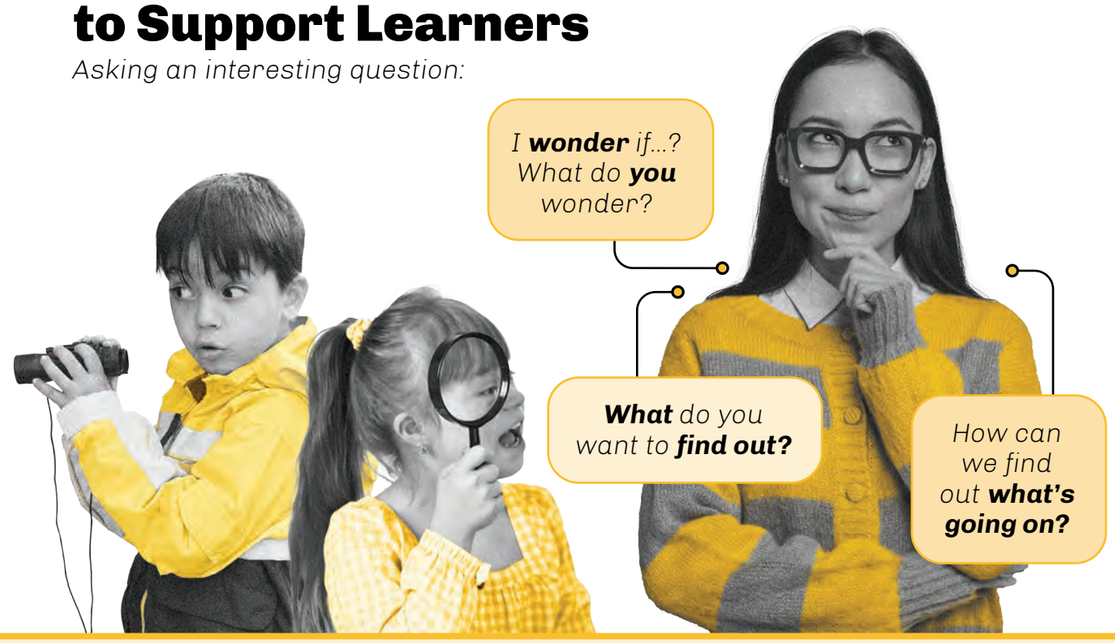
- ✓ Is the question worth investigating and ***interesting?***
- ✓ Is the ***intent*** of the question clear: will the learners be asked to describe, compare or find an association?
- ✓ Is it possible to answer the question using data that the ***learners*** can feasibly collect?
- ✓ Is the ***variable*** of interest clear?
- ✓ Is the population of interest ***clear*** and specific (e.g. this class/ all classes in the year in our school/birds in our playground)?

— Arnold and Franklin.[24]

Prompts for Teachers

to Support Learners

Asking an interesting question:



Looking at a claim that needs to be investigated:

A SUPER GUIDE TO...

ASKING GOOD QUESTIONS!

Data Education in Schools

Is it an **interesting** question?

Is it **worth** answering?

USING DATA TO CHANGE THE WORLD!

WHAT ARE YOU MEASURING?

This is your **variable.**

HOW WILL YOU ANSWER YOUR QUESTION?

How high? How many? How often? Most? Least?

Most popular? Most frequent?

IS IT POSSIBLE TO COLLECT THIS DATA?

How high are the **clouds** above the school...?

WHICH GROUPS WILL YOU SURVEY?

This is your **population.**

HOW LONG WILL YOU GATHER DATA?

Hourly? Daily? Weekly? For a term... for a year?

WWW.DATASCHOOLS.EDUCATION

Plan

At the planning stage of the cycle, the learner decides how to answer the question(s) they have developed in the previous stage. The teacher can support the learners in planning fresh data collection.

Groups of learners could generate lists of possible data which they could collect, and with support, identify which types of data will be feasible and most useful to collect. For learners at earlier levels, the teacher has an important role in guiding the class to make ethical decisions for projects involving human participants, relying on their professional judgement about topics which might be private or sensitive for participants, and ensuring that their time is not wasted.

More confident and experienced learners can move towards using existing datasets from reputable sources, as well as requiring less support from the teacher in collecting their own datasets. For example, learners may choose to gather sensor data (e.g. temperature, atmospheric pressure, CO2 levels) using technology sensors, and so will need to decide which to use and where are the most appropriate places to put them. Teachers might assist by demonstrating how to set up the sensors and illustrating what can be measured. In the planning phase of the cycle, it is good practice to decide when data collection will start and stop, for example specifying an end date or setting the number of participants in advance and stopping when that number is reached. This helps to make sure that the project will be possible within a certain time-scale, and develops good research habits.

More experienced learners will benefit from working with public datasets when studying real-world problems. The first step would be to identify what existing data sources there are which could be helpful – which organisation keep data on this? It is helpful to start by checking reputable sources. For example, does the government or local authority gather and releases data sets on this topic? If so, is the data recent and freely accessible? Do sports teams have statistics on their website? Can you find tables of information on box office takings to settle the endless debate about 'what's the best Marvel/Star Wars film?'.

If the learner finds a data set by searching the internet, they may need support to evaluate the quality of the data source by considering the source of the data. Could the organisation which collected the data have an agenda or motive which may cause them to present a misleading sample of the data? Was the data gathered from a representative sample of people who are relevant to the question identified in the "problem" phase of the cycle? »

Skill/Task	Classroom support	Progression
Find a data set and evaluate data quality	Check existing reputable data sets to decide whether there is a high-quality data set which could be used to answer the question. Search for other public data and evaluate the data quality in terms the agenda of the individual or organisation which collected the data.	Working with existing public datasets requires a firm foundation which comes from working with smaller datasets collected in the class or school. So these skills may start to develop towards the end of Level 2.

Table 3. Skills progression for the Plan Stage of PPDAC (continued overleaf)

"In the rush to get data and start analysis, attention to design is often glossed over."

— David Spiegelhalter [15]

Skill/Task	Classroom support	Progression
Design how to collect data	• Decide and justify when it is necessary to collect new data, and if so, plan exactly what data is needed. • Decide and justify how and by whom the data will be collected (such as collected by humans using a survey, or collected by sensors automatically, or both). • Decide and justify from whom data will be collected (if applicable), taking into account ethical principles of not asking questions which invade privacy or take up too much time of the participants. • Decide and justify when and where the data collection will take place and when data collection will stop.	At Early and First level, the teacher models how to make these planning decisions, before gradually removing the scaffolding to enable Level 2 students to work through these decisions.
Predict an answer	Predict possible and likely answers to the question.	Teachers can guide learners through thinking about the logically possible answers to a question, and as learners progress, guide them to thinking about the likelihood of different answers (with justification). It may be helpful to use vocabulary about probability from Level 1 (see MNU 1-22a and MNU 2-22a).

Table 3. Skills progression for the Plan Stage of PPDAC

Part of the Planning stage involves predicting what the possible answers logically or plausibly could be, and which is most likely given what we already know. This is an informal version of hypothesis testing which will be familiar from science lessons. Thinking this through in advance is often useful in deciding which analysis method should be used. For example, imagine a project in which learners investigate whether indoor air quality in a classroom affects concentration. It is known that when CO2 levels reach a threshold (over 1500 ppm), people lose focus and their attention suffers.[16]

Learners could first discuss which classrooms feel stuffy, or times of day when they feel sleepy in particular rooms. This discussion could lead to a prediction about the room or time of day in which they expect to find the highest levels of CO2, and why. For example, they might predict that the classroom which has a broken window catch might have the highest CO2 levels because it is not possible to let in fresh air. They can also predict a sensible minimum and maximum range for the data value – it won't be possible to measure lower than 0 ppm (unless the sensor is broken) and it is unlikely to reach much more than 5000ppm because that would break health and safety in the workplace laws.[17] If they put a CO2 sensor in each classroom, they would have an objective measure of how much CO2 was present in the room, and how this changed over time. At the analysis stage, they could check graphs of the data for spikes in CO2 in the rooms and times that they predicted. •

Prompts for Teachers

to Support Learners

"The core advantage of data is that it tells you something about the world that you didn't know before."

— Hilary Mason

Data

> *Having planned for data collection, the next stage in the cycle is to gather data. If the learners are gathering a fresh data set then they need to decide in advance how to structure or organise the data so that they can use it in the Analysis stage.*

Learners in the early stages may collect toys or other objects before turning them into data displays. The next stage is to use symbols such as picture graphs or tally marks to represent data items, and then on to recording numbers in a table by hand. Once learners are more secure in their understanding of data collection and data types, then can move on to using digital tools such as spreadsheets. »

Skill/Task	Classroom support	Progression
Collect a data set	• Collect data locally within the class or school. • Extend data collection to the community. • Set up digital tools to assist in data collection where appropriate (e.g. online survey tools or sensors).	Early level learners benefit from gathering physical objects which they can then categorise and sort (e.g. leaves, toys). This progresses to collecting data locally including measuring within the school environment or designing class surveys. Once learners have gained experience, they can extend their data collection to wider groups.
Prepare data	Clean and prepare data for analysis or processing.	For learners in the early stages this may involve collecting or sorting physical cards to make it easier to create a chart. For more advanced learners, this might mean tidying up survey entries or removing blank entries.

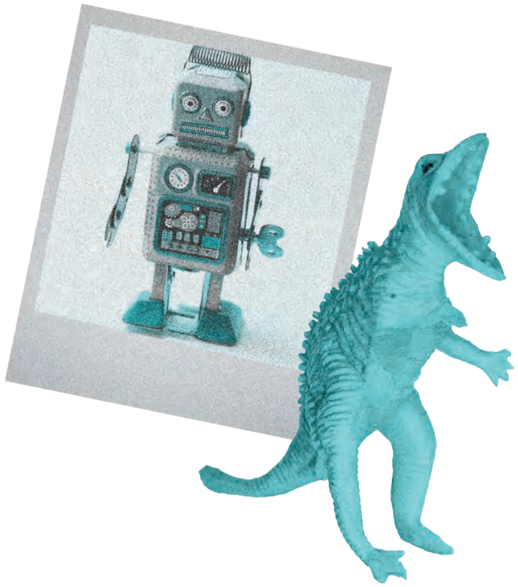
Table 4. Skills progression for the Data Stage of PPDAC (continued overleaf)

Skill/Task	Classroom support	Progression
Organise and structure data	• Categorise and sort data. • Organise and structure data digitally.	In the early years, teachers can help learners with structuring information through hands-on grouping and sorting objects such as toys according to colour, shape or size. As they progress, learners will learn about two types of data: numerical and categorical, which in turn enables them store data in a structured way in digital tools such as spreadsheets.
Check that data is correct	Check data for accuracy.	This skill is important in the case when learners are typing in written survey answers, or checking that participants in an online survey have filled in the questions correctly (Level 2). If an answer looks unusual, this can lead to a discussion about whether it should be included or not, and why. Learners could be encouraged to think about what they can do to avoid errors/improve accuracy -- e.g. two learners independently type in the answers and cross-check for errors.

Table 4. Skills progression for the Data Stage of PPDAC

If the learners are using a publicly available dataset, then the task is to download it and store it on a local computer for analysis. It is possible to explore some datasets interactively online without requiring downloads (e.g. Gapminder[18] and CODAP[19]).

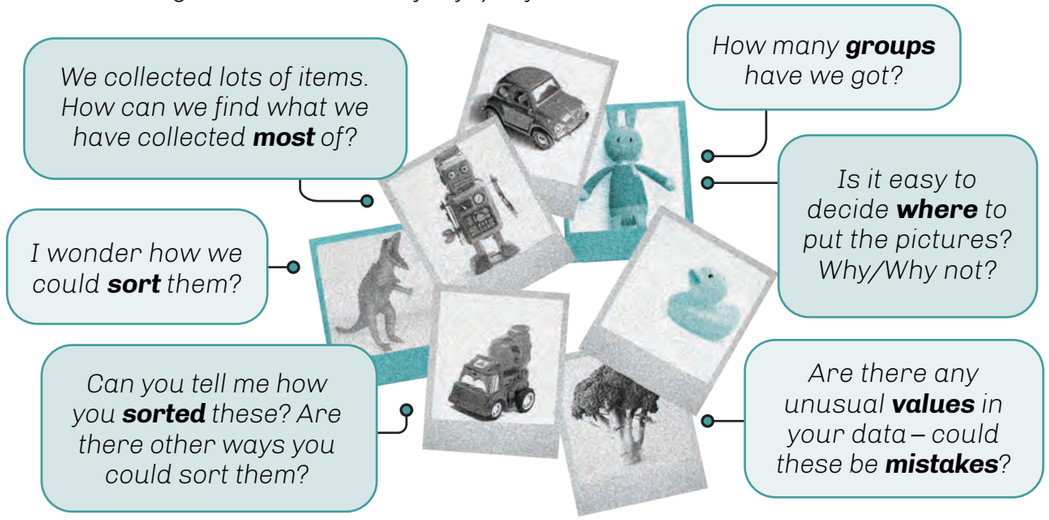
If the learners have collected the data themselves, they should check it for accuracy to make sure that they have not introduced copying errors or typos which might cause misleading results later. •



Prompts for Teachers

to Support Learners

When working with collections of toys/objects:



When using a pre-existing dataset:

"Good statistics are not smoke and mirrors; in fact, they help us see more clearly."

— Tim Harford [20]

Analysis

> *Children and young people will need to learn how to make sense of the data in the context of the problem they are trying to solve. The first step in data analysis is to get to know the data, often by "eye-balling" it in a graph or table.*

For primary school learners, this means knowing how to read and reason about tables, charts, graphs and infographics. Children encounter a wide range of engaging, interactive and informative techniques for showing data in the media and on the internet.

These representations are often considerably more complex and informative than traditional pie or bar charts, and they will need support as they learn to understand them. They will also be learning how to create tables, charts and diagrams either by hand (particularly in the early levels) or by using software packages such as Excel and Codap. •

Skill / Task	Classroom support	Progression
Explore the data set	Read tables, charts, graphs or infographics of data.	At an earlier stage, beginners may find it helpful to construct simple stacks of real objects or create pictographs as a way to learn how to read similar displays made by others. When exploring a dataset it can be useful to switch between different sorts of graphs and tables, as different displays make different facts more obvious. See pages 32-33 for a guide to graphs in increasing order of complexity.

Table 5. Skills progression for the Analysis Stage of PPDAC (continued overleaf)

Skill/Task	Classroom support	Progression
Visualise the data in tables and graphs	Create tables, charts, graphs or infographics of data.	Early level learners will create graphs and displays with objects (or people!), and this will gradually progress to displays using symbols such as pictographs, dot plots, tables with tally marks and bar charts (see pages 32-33). As learners begin to work with larger datasets, they will find it helpful to use digital tools for making graphs including spreadsheets and CODAP. It's better to spend time thinking about what the data means than laboriously drawing a chart by hand. More experienced learners will begin to appropriately select a type of graph which displays the dataset in a way which is easy to read.
Analyse the data using simple descriptive statistics	• Describe a dataset - highest and lowest values, the difference between the smallest and largest, most commonly occurring/typical value, least and most common category. • Compare the data between two groups - consider whether there is a difference between these groups. If so, how much of a difference is there between these two groups and does it matter? • Compare the data between the same group at different time points – are there changes in the variable of interest at different time points? If so, how much of a change and does it matter? • Explore the association between two variables – if the value of one variable increases or decreases, what happens to the value of the other?	Learners will begin by describing the frequency of a single variable in a dataset, e.g. counting how many red toys or blue toys they found in the class for categorical variables. As they advance, they can compare to see how one variable differs across two groups; e.g. do boys in P6 tend to be taller than girls in P6, and if so by how much? At a similar level of difficulty, learners could compare one group over time, e.g. how tall were the children in P5 and how tall are the same children now? How much have these children tended to grow in a year? Learners could then explore the relationships or associations between two variables, by looking at displays such as scatter plots. They should use these graphs to help them notice what happens to one variable as the other increases or decreases. It is helpful for them to consider whether there really is an association between the variables, or whether it could just be chance. Teachers could also prompt them to consider whether one variable causes the other to change or not.

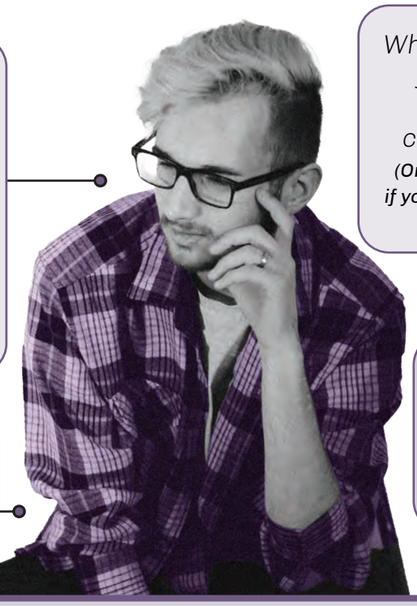
Table 5. Skills progression for the Analysis Stage of PPDAC

Prompts for Teachers

to Support Learners

*What is the smallest or biggest value that this **variable** takes in your dataset? What's the difference between the smallest and largest value?*
(Or spread/variance if your learners have covered these concepts)

*What is a **typical value** for this variable? The most/least common category?*
(Or mean, median or mode, if your learners have covered these concepts)



*What kind of **graph** or **table** will help you to understand the data most clearly?*

*Is the data **arranged** in the best way to help you answer this question?*

When looking at a graph:

*Do you see any **clusters?***

*What do you **notice,** what do you wonder?*

*Can you explain the **pattern?***

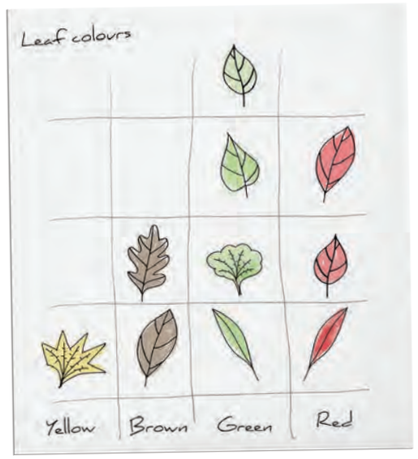
*How does this group **compare** to that group?*

*What **shapes** do you see here and what do you think they **mean?***

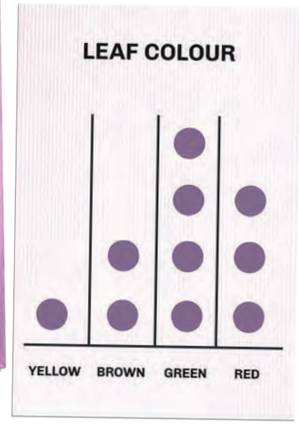
Getting to Grips with Graphs

Teaching how to represent frequency counts in graphs of increasing difficulty. Start with categorising real objects, then move on to showing the objects as symbols.*

1. Categorised objects



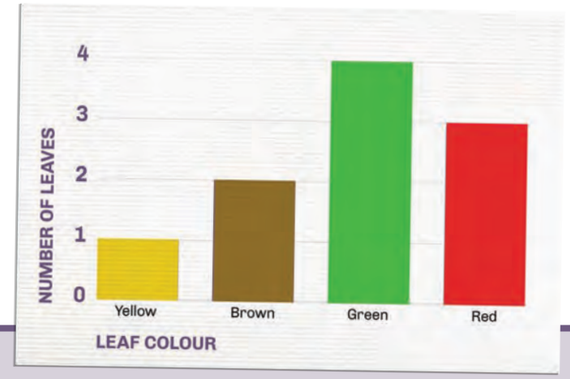
2. Picture graph



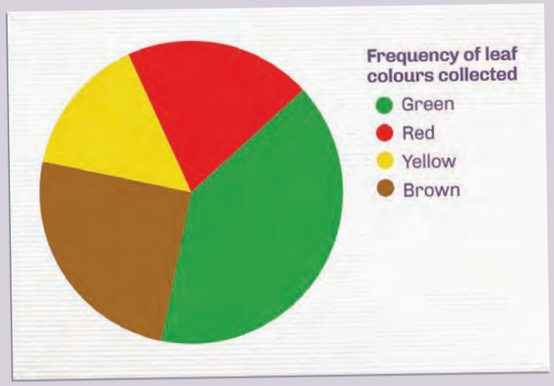
3. Tally marks/dot plots

Leaf colour	Frequency
Yellow	1
Brown	2
Green	4
Red	3

4. Frequency chart



5. Bar chart



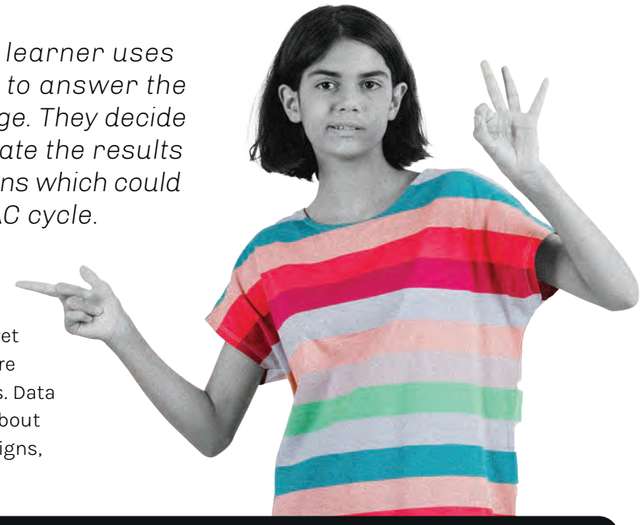
6. Pie chart

(once proportional reasoning is secure)

* *Note: other examples will be more appropriate for learners with colour blindness, e.g. categorisation by shape.*

Conclusions

> *In the Conclusions stage, the learner uses the results from the analysis to answer the question from the Problem stage. They decide how to explain and communicate the results to others and form new questions which could be answered in the next PPDAC cycle.*



The Conclusions stage arguably involves the most important skills for data literacy because most people are required to interpret claims about data made by other people more often than they conduct analyses themselves. Data interpretation skills enable critical thinking about claims in newspaper articles, election campaigns, adverts and in the workplace. »

Skill/Task	Classroom support	Progression
Answer the question	Look back at the question from the first stage in the PPDAC cycle (Problem) and present the answer clearly and concisely.	In the early years, answers might best be presented verbally during a discussion, perhaps with reference to a graph or table. The teacher can model the language for clearly answering questions during the discussion. Level 1 learners can be supported with fill-in-the-blank templates which help them to summarise the key information, whereas Level 2 learners will begin to use the skill of writing clear summaries in their own words.
Explain what the data reveals	Clearly summarise the relevant data.	This skill is the ability to clearly and succinctly summarise the "story" or main findings in a graph or table. It relies on the analysis skill to identify the key points of interest in a graph/table, and then accurately explain them. This can be practiced verbally, or modelled in template form. Exercises such as writing the caption for a graph may be useful in honing the skill.

Table 6. Skills progression for the Conclusions Stage of PPDAC (continued overleaf)

"The numbers have no way of speaking for themselves. We speak for them. We imbue them with meaning."

– Nate Silver [21]

Skill/Task	Classroom support	Progression
Interpret what output from analysis means	Given numerical results in tables graphs, decide what this means for the problem under consideration, and what the implications of these results might be.	This skill requires the learner, with the support of the teacher, to connect their results back to the original investigative question which they were trying to answer. As the learners progress, they can be supported to think about wider implications of the results, including actions which could be taken (see below).
Decide whether a claim makes sense based on results	If the PPDAC cycle started from a prompt in an news story, on a website or in a video, look back at the article or re-watch the clip and consider whether the claims in the article are likely to be true, given the results of the investigation.	This skill is likely to be used by learners at Level 2 in cases where they have explored existing datasets as part of fact checking. The teacher can model how to check the language used in the claim against the findings in the data e.g. "The article says that most people believe X but from the data you looked at, it seems that 55% believe it. Do you think this is fair?" Look for words like "most", "many", "always", "never" in the claims.
Decide, recommend and communicate analysis to others	Based on interpretation of the findings from the previous stages of the PPDAC cycle, consider whether any action should be taken to change the situation, and if so which action. Communicate data-informed recommendations clearly and persuasively.	Investigations in the school or local area can lead to the learners realizing that a local situation could be improved. Teachers can support them to consider what action could be taken, and how this could be arranged. For example, if the results suggest that the school library layout could usefully be changed, who should they approach with their layout suggestions? Teachers can help the learners to decide on appropriate ways to communicate their findings and suggestions to those in power e.g. a presentation to the headteacher or a letter to an MP.
Act on data to change own behaviour	In projects where the data is personal (such as fitness), consider what the data shows about personal habits and what changes (if any) might be beneficial.	Learners at Level 2 may be working on datasets about their own habits (see the Dear Data or DataFit resources), in which case the teacher has a role in helping them to decide whether and how they could change their health and wellbeing habits in a realistic, healthy way (e.g. walking to school each day or drinking a glass of water every lunchtime).

Skill/Task	Classroom support	Progression
Form new questions	At the end of the first cycle of PPDAC, consider whether more investigation is needed, such as asking additional or more in-depth questions	When analysing data it is common to notice and wonder about more questions, which can lead to another cycle of PPDAC. The teacher can prompt learners at all stages to list questions which they still want to know the answers to after they have finished the conclusions phase. Regardless of whether there is time to actually go through another full cycle, this can be used to reflect on whether the Plan stage of PPDAC could be improved for next time (e.g. we could have phrased the question differently), or perhaps another dataset would be needed to answer the new questions.

Table 6. Skills progression for the Conclusions Stage of PPDAC

In some cases, the data may have been collected to cast light on an issue within the local community. Imagine that a group of learners has collected data on the frequency of visits to the local park by various groups of people. They could use the data to recommend how the local authority should decide to spend money refurbishing the park to meet the needs of the different users. The skills of clearly communicating results, storytelling with data and persuasion are required here.

If the learner has conducted a study about a topic within their daily life – such as personal health or fitness – in the conclusions stage they can use the data to make decisions about whether and how they should change their behaviour to meet their goal. For example, if they were using a sleep tracker to see if they get enough sleep per night for their age, they could compare a visualisation of their data to a recommended figure for young people of their age. With support they could arrive at a plan to improve their sleeping patterns by changing their bedtime habits.

Conducting a PPDAC project often raises as many new questions as it answers - and this is a good thing! At the final stage of the first cycle, learners can reflect on what they have learned to decide what new or refined questions they should ask if they or other researchers were to go through another cycle in the data problem solving process. •

> **"We should ask who is missing from the data we're being shown... and whether our conclusions might be different if they were included.**
>
> — *Tim Harford*[22]

Prompts for Teachers

to Support Learners

Think back to the questions you wanted to answer in the Problem stage. What are your ***answers?***

Which of your ***results*** *did you find most* ***surprising?***

Should we ***trust*** *these results? Why/Why not?*

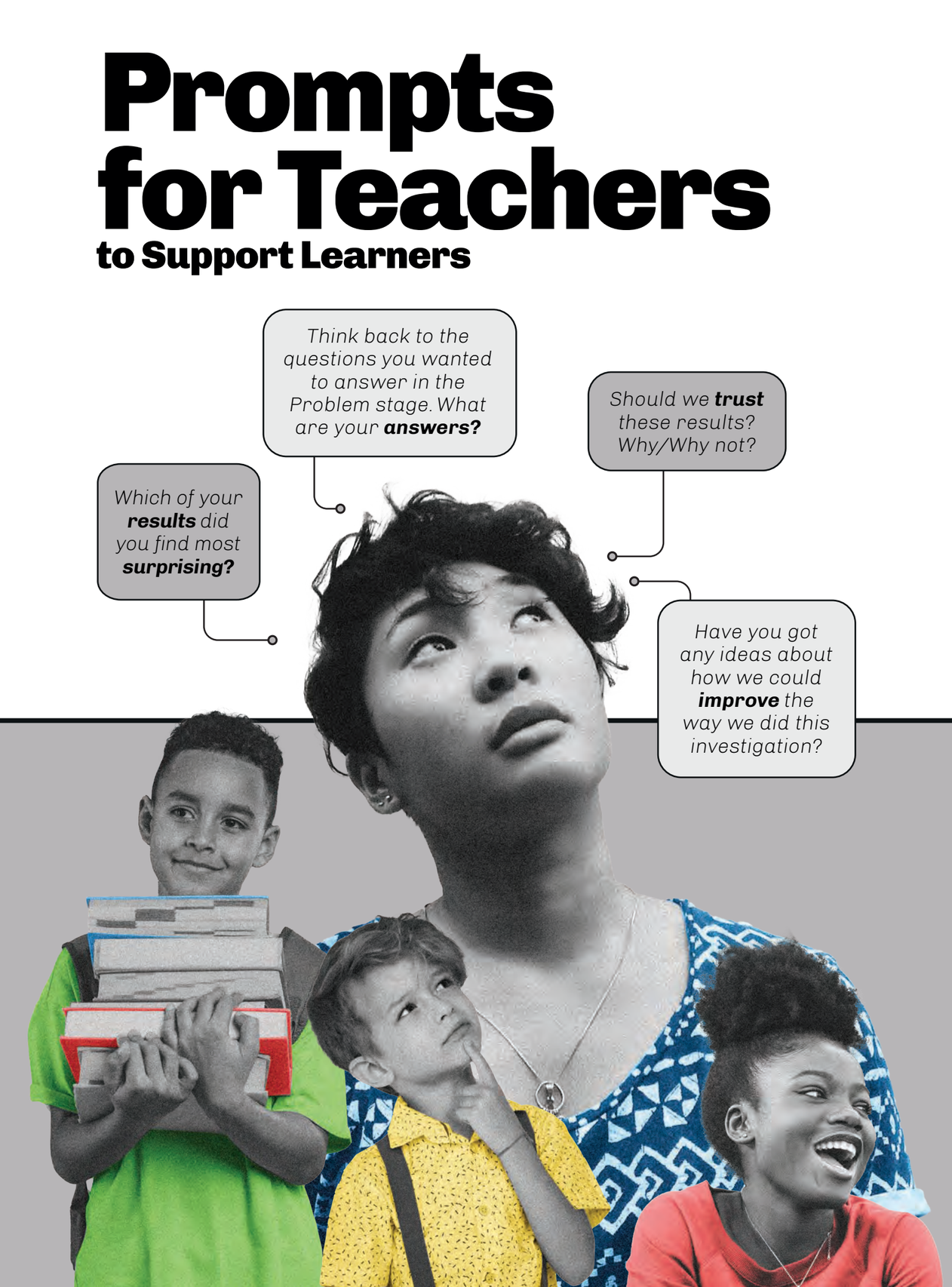
Have you got any ideas about how we could ***improve*** *the way we did this investigation?*

Would the ***results*** *of your investigation be the same if you had worded the questions in a* ***different*** *way?*

How does what you found out relate to our ***community?***

Do you have any ***more questions*** *you'd like to answer now?*

What ***action*** *do you think we should take to change things based on your analysis? How could we* ***persuade others to act?***



The following pages show a collection of data literacy activities to inspire you with ideas for lessons which your class will enjoy.

Resources & Activities

> *We recognise that teaching a new area of the curriculum presents new challenges for teachers – it places demands on an already full timetable, as well as you learning what to teach about a distinct set of knowledge and skills. But, we hope that by now we have convinced you of the importance of integrating teaching and learning about data literacy into your practice.*

We encourage you to begin by thinking about what you already know about the learners in your class and what might capture their curiosity. They are already creating and consuming data through their daily lives and it can be fun to explore with them what they already notice and understand about the world around them. For example, are they aware of thermostats and smart meters?

It's not necessary to initially try to diagnose where you expect the learners to be with data literacy – pick an open-ended activity on a topic your learners will like (e.g. you could choose from our Data Explorer Cards),[23] listen to the learners' discussions and then talk to them about their learning. After you have explored some of the resources with your learners, you will become more confident in assessing which learners need support or challenge in which distinct aspects of data literacy. At this point the PPDAC and skills labelling of the resources might be helpful. The activities are also grouped into curriculum levels for convenience.

The thinking skills of data literacy are "for" everyone, but different people will have different interests and should have the opportunity to develop their data skills by pursuing topics close to their hearts. It's important to avoid inadvertently reinforcing stereotypes about which types of people are "good at" science or technology-related topics. Through engaging with learners' interests and curiosity, we can avoid marginalising learners.

Data literacy activities often give scope for creative and authentic assessment approaches, particularly if the learners are keen to share their findings with others, and request that action should be taken. For example, recommendations for parents in the form of an infographic could be evidence of skills from the Conclusions stage of PPDAC, and photos of sorted toys could be evidence of skills from the Data stage.

Many of the activities shared in this section are drawn from the New Zealand maths activities and NRich maths (England). Where the curriculum levels are different in the original source we have mapped them to the Scottish levels guided by expectations for learning outcomes in Curriculum for Excellence. However, you may find that the children and young people you are working with have needs and capacities out-with what was imagined for the hypothetical learner. As you become more familiar with the curricula of data literacy you will be able to use your knowledge and understanding of the young people you teach to choose and adapt the best resources for you.

You'll notice that there is often an overlap between the different PPDAC stages which an activity might cover. For example, the questions considered in the Conclusions section rely on Analysis skills, and Problem or Plan activities often naturally lead into the Data stage of the cycle. •

We share these resources and activities with the hope that you find them useful starting points when planning for learning.

Early Level

Resources that support teaching foundational skills young learners need to work with data.

> *These typically involve working with real objects and pictures, with the aim of transforming these into more abstract graphs and charts later.*

Socks

This idea, from the NRich maths project, suggests resources in a nursery/P1 classroom to encourage open ended maths learning. The data literacy aspects relate to categorising and sorting information. In the early stages, learners can categorise everyday objects such as socks by colour or size. Hanging socks on a washing line in order of size introduces the concepts of ordering and sorting which become important later. The prompts "How can we remember which sock goes where? (When sorting or ordering)" and "Can we sort them in a different way?" are particularly useful.

SKILLS & CONCEPTS
Categorising and ordering objects.

NRich Maths
https://edin.ac/3iJs0za

Collecting

Another categorising and sorting activity from NRich, which could be combined with health and well-being and natural science topics. Challenging the learners to consider alternative ways of sorting the collection introduces awareness of a powerful computational thinking concept which is relevant to real-world data sets. It is helpful for the learners to hypothetically consider which other objects would belong in a collection. Useful data literacy prompts from the activity are "Can you find another way to sort your collection? What if you sorted them into the divided tray? Is that tray big enough? (Can you find one that is?)" and "Is there something else you can think of that could belong here?"

SKILLS & CONCEPTS
Categorising and ordering objects.

NRich Maths
https://edin.ac/3wcAtJy

I like trucks

An introductory level series of lessons about gathering data on learners' favourite toys (not just trucks!). It involves sorting and counting photographs. It goes through the PPDAC cycle with support from the teacher.

SKILLS & CONCEPTS
Pose investigative questions with support from the teacher; with the teacher, decide on how to collect the data to answer the investigative question; sort objects into categories for display; make a display of the data collected (pictograph); make statements about data displays.

NZ Maths
https://edin.ac/3Wi9Ey7

All about us

Introduction to data literacy unit of work, using pictographs of favourite activities and categorisation. Sessions 1 and 2 are most relevant to data literacy. Session 1 focuses on working with real-world objects, while session 2 moves on to creating simple charts by arranging cut-outs of pictures.

SKILLS & CONCEPTS
Conduct investigations using the statistical enquiry cycle: posing and answering questions; gathering, sorting and counting, and displaying category data; discussing the results; create picture graphs about category data and discuss patterns in the data.

NZ Maths
https://edin.ac/3XQswW9

The Garden

The learners are given drawings of insects found in a garden and are asked to make a data display with categories. This leads on to the idea of pictogram. Could be linked to learning for sustainability.

SKILLS & CONCEPTS
Categorisation, sorting, counting and data display.

NZ Maths
https://edin.ac/3XCDMVV

First Level

Projects for the whole PPDAC cycle

Healthy Hands

A five-lesson sequence with health and well-being linked to hand-washing. There are extensive teacher notes explaining how to move from a picture representation to tally marks and then bar charts. They compare two variables: hand-washing and absence rates visually and through verbal discussion. There are a couple of books about germs mentioned in the resources which are available from online bookshops (schools may have their own similar books).

SKILLS & CONCEPTS

- Identify a health issue for investigation
- Pose an investigation question with support from the teacher
- Participate in developing a plan to collect data
- Gather, sort, count & display category data
- Make a pictograph and identify its key features
- Discuss results of an investigation
- Make 'I wonder' statements and pose new investigative questions
- Draw a picture or diagram to show the investigation process

NZ maths
https://edin.ac/3ZIOnRb

Parties and favourites

A sequence of lessons to model the PPDAC cycle. Loosely based on a birthday party theme. It has extensive teachers' notes. Suitable for teachers who are trying PPDAC for the first time, or who want to use this as a model for an investigation for their own class.

SKILLS & CONCEPTS

Modelling the full PPDAC cycle. Learners count, compare, organise, analyse, display and interpret data and at the same time, apply early additive strategies for combining numbers.

NZ Maths
https://edin.ac/3ZHm45F

Getting to School

This is the familiar activity of collecting and displaying data on how people travel to school, but with useful notes for teachers about how to support data literacy learning through the PPDAC cycle.

SKILLS & CONCEPTS

Problem: posing a question to investigate
Plan: planning an investigation
Data: collecting and recording whole number and category data
Analysis: sorting data, including frequency plots
Conclusion: making summary statements from data displays

NZ Maths
https://edin.ac/3QOOc2v

Problem

Ladybird count

The learners are shown pictures of cards with ladybirds on them, and encouraged to think about what investigative questions using 'What do you notice?' and 'What questions would you like to ask?' prompts. Suitable for younger learners, but could be adapted for older learners by including different graphs.

SKILLS & CONCEPTS

Ask an interesting question; create a data display.

NRich Maths
https://edin.ac/3QKaL8N

Plan

Open or Closed

A quick worksheet-style exercise about the difference between open and closed survey questions.

SKILLS & CONCEPTS

Conduct investigations using the statistical enquiry cycle: posing and answering questions; gathering, sorting, and displaying category and whole-number data; communicating findings based on the data.

NZ Maths
https://edin.ac/3GW9QxA

Data

Make a Guess Who Game

Learners will probably be used to playing the *Guess Who* game - in this activity, they get to make their own version of the game. In doing so, they learn about categorical variables (for example hair colour, eye colour of the game characters).

SKILLS & CONCEPTS

- Exploring variability in categorical variables
- Structuring information

Data Education in Schools
https://edin.ac/3IRlNai

Data

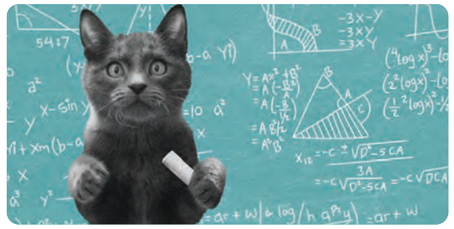
Make a Top Trumps Game

Use the familiar Top Trump cards to teach learners the concept of numerical variables (such as length, or weights of every animal in a collection). Learners also design their own card set.

SKILLS & CONCEPTS

- Exploring variability in numerical variables
- Structuring information

Data Education in Schools
https://edin.ac/3ISZj8V



Data

Cat Maths

A Teacher's Guide with a series of games, lessons and activities centred around a set of cards about cats. It teaches categorical variables (of course!). There are free printable sets of some of the cat cards on the website.

SKILLS & CONCEPTS

- Sorting information in categories
- Structuring information
- Handing categorical information

Creative Maths
https://edin.ac/3ITqwbx

Data

Data Cards

Learners gather information about simple attributes of classmates (e.g. eye colour) using paper data cards. They then organise them in piles. There are links to the Computing Science strand of the Technologies curriculum in Scotland, with structuring information.

SKILLS & CONCEPTS

- Pose investigative questions
- Write data collection or survey questions to support collecting information for investigation
- Make sensible statements about the information and be able to back up their statements with appropriate displays
- Display information to answer investigative questions
- Collect information and sort into categories
- Answer investigative questions by sorting, organising and arranging information

NZ Maths
https://edin.ac/3XikKEM

Analysis

Data Sweets

A simple way of introducing data analysis and simple visualisation.

SKILLS & CONCEPTS

Analysing data by creating frequency tables with tally marks and numbers and creating bar charts.

Data Education in Schools
https://edin.ac/3XGLgYi

Analysis

Match-ups

A set of activities where the learners investigate favourite colours, learn to interpret bar graphs and choose true or false statements about other people's interpretations of the graphs.

SKILLS & CONCEPTS

- Statistical literacy
- Interpreting graphs

NZ Maths
https://edin.ac/3XqTPXu

Conclusions

Tony's Spelling Troubles

The task is to interpret a bar chart which has a graph of spelling practice and compare it to a table of spelling scores. Learners have to give advice to the fictional learner about how to adjust his homework habits. It's quite a nice example of reasoning about cause and effect using data but without formal statistical calculations. It could be related to the Curriculum for Excellence four capacities.

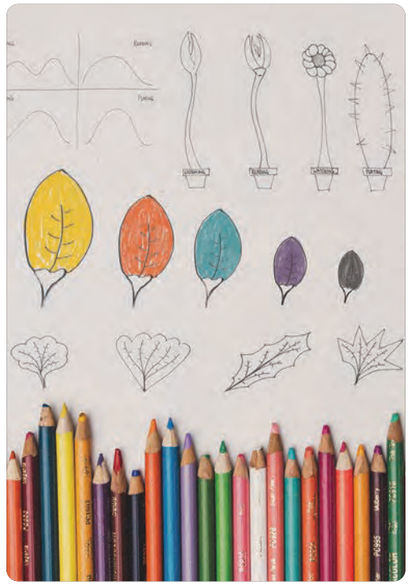
SKILLS & CONCEPTS

- Describe in their own words the situation represented in a bar chart and tally chart
- Use a bar chart and tally chart to make sensible predictions
- Interpret information and results in context

NZ Maths
https://edin.ac/3QM77uW

Second Level

Projects for the whole PPDAC cycle



Dear Data

This is an unusual and creative approach to teaching data literacy.

It focuses on the artistic side of hand-drawing visualisations of data which learners collect about their own lives.

The resource contains six lessons.

SKILLS & CONCEPTS

- How to ask questions to allow more depth to data collected
- Using a visualisation to present the data so that others can discover information
- How to uncover the story within the collected data

Data Education in Schools
https://edin.ac/3GReA7y

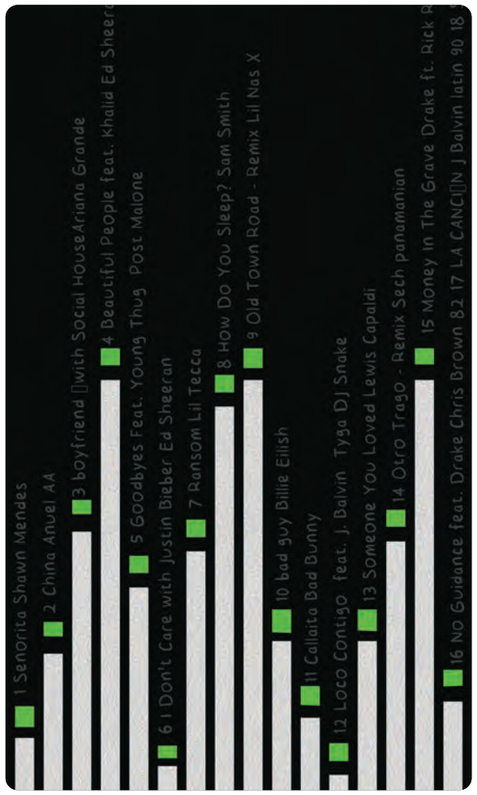
Climates, Microclimates & Temperature Data

This set of activities explores climates and microclimates and uses sensors for collecting data. It introduces the concept of independent and dependent variables in simple classroom experiments.

SKILLS & CONCEPTS

- Understand the differences between and be able to define dependant variable and independent variable
- Learn to measure temperature and recognise factors (dependant variables) that impact temperature

Data Education in Schools
https://edin.ac/3WiaGdt



Spotify Data Analysis

Spotify Data Analysis makes use of secondary data from Spotify. Learners will use data to identify patterns and relationships between different characteristics.

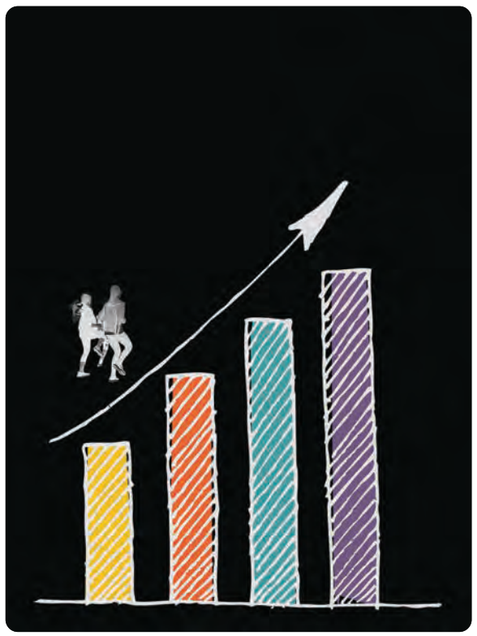
The activity will support learners in developing their ability to review and interpret a dataset. The activity starts by encouraging learners to think of questions that the dataset might answer and thus develop their "Problem" skills.

Once the dataset has been analysed learners will have the opportunity to reorganise and restructure data to help them answer their questions (draw conclusions). This activity can be used for Level 1 or Level 2, according to which datasets you use.

SKILLS & CONCEPTS

- Exploring a real-world dataset
- Creating graphs
- Drawing conclusions from data

Data Education in Schools
https://edin.ac/3CWnoHT



DataFit

DataFit is about helping learners to improve their understanding and confidence in health and well-being as well as data literacy, so that they feel empowered to use and act on their personal activity data.

In a series of four lessons, learners use step count data from devices such a Fitbits to monitor their physical activity.

SKILLS & CONCEPTS

- Devising and using a variety of methods to gather information
- Working with others to collate, organise and communicate the results
- Carrying out investigations
- Designing surveys

Data Education in Schools
https://edin.ac/3kpVKNh

Data Escape Rooms

Learners will become Trainee Agent of DATA to defend Scotland against the villainous high-tech VIKINGS.

They will solve puzzles in time to defeat the evil time-travelling baddies and foil their plans. The Agents will have to use graphs, charts, maps and logic to uncover the plots and crack the secret codes.

Along the way they are introduced to the basics of how the internet works with data centres and undersea fibre cables, how cryptocurrency works, and how to avoid phishing attempts by hackers.

SKILLS & CONCEPTS

- Interpreting data in tables, bar charts, line graphs and maps
- Solving logic puzzles and codes

Data Education in Schools
https://edin.ac/3D2b4pQ

Photo credit: Denise Bertacchi

Cotton Wool Catapults

Learners make cotton wool catapults and then graph their accuracy and distance.

They should decide on the best representation for their results. There is a link to the Technologies curriculum.

SKILLS & CONCEPTS

- Conduct investigations using the statistical enquiry cycle: posing and answering questions; gathering, sorting, and displaying category and whole-number data; communicating findings based on the data.
- Use of dot plots to show spread central tendency in data.

NZ Maths
https://edin.ac/3H9XuTD

Cool Colours

This hands-on science experiment about water temperature covers all stages of the PPDAC cycle. Learners plot three lines on a time series graph.

Good teachers' notes about connections to science.

SKILLS & CONCEPTS

- Statistical investigation
- Compare statements with the features of simple data displays from statistical investigations or probability activities undertaken by others
- Plot 3 series on a time series graph

NZ Maths
https://edin.ac/3JJSCoa

Data Explorer Cards

Our Data Explorer cards introduce children and young people to a data problem-solving approach (PPDAC).

The design builds on informal question-asking behaviours and supports learners' exploration of data in relation to the themes of *Lego! Birds! Lost Property! Plastics!* or a theme of their choosing.

Question prompts on the cards support learners to explore problems, plan, collect data, analyse and draw conclusions.

SKILLS & CONCEPTS

- Identifying and exploring questions about the world
- Developing and building skills and habits of mind for data-problem solving

Data Education in Schools
https://edin.ac/3ZOYnZf

Illustration by Jim Kettner and Kane Lynch

Youcubed K-12 Data Science comic

A comic book which introduces learners to key information about the subject of data science and how to look after their own personal data.

SKILLS & CONCEPTS

Data privacy

Youcubed

https://edin.ac/3DjLlcH

Dollar Street Investigation

Using the GapMinder Dollar Street photos, this Teacher's Guide suggests an investigation of how families live in different parts of the world on different income levels. It encourages learners to spend time exploring a rich real-world dataset. See Example 8 for Level B, p62-66.

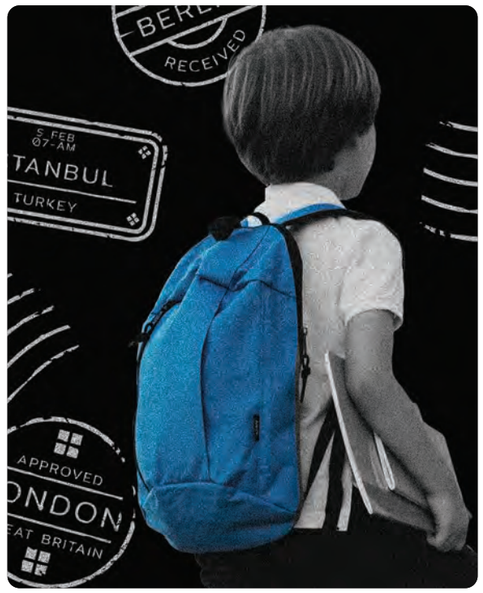
SKILLS & CONCEPTS

Consider how and why the data were collected, from whom the data were collected, how variables were measured, and what were the possible outcomes for the variables.

Consider the ethical implications of how data is collected.

National Council for Teachers of Mathematics (USA)

https://edin.ac/3XId5VY



The Boy at the Back of the Class

This activity, related to a book called *The Boy at The Back of The Class*, follows through the PPDAC cycle to help learners understand about where refugees are in the world, how many refugees there are, and how this compares to previous times.

This would be a good social studies/health and well-being activity.

SKILLS & CONCEPTS

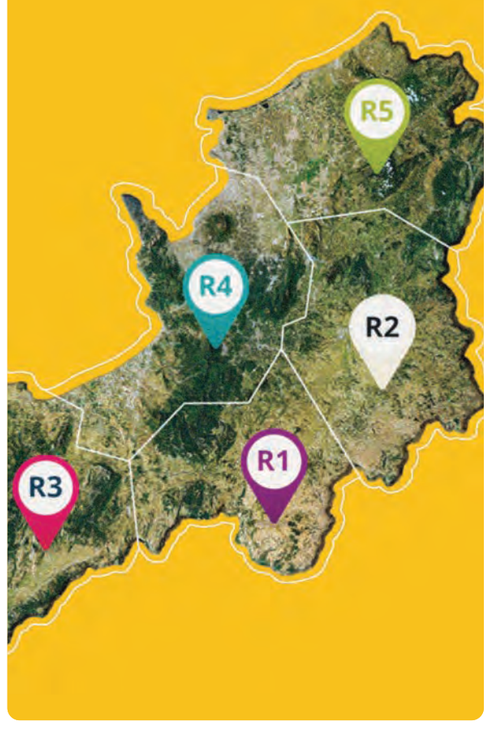
Carry out a data investigation of real-world data to understand our society.

Data Education in Schools

https://edin.ac/3kr9OpP

Problem

The Story of Statistopia



In this lesson there is a fictional island and learners work through series of data activities to better understand the fictional community and its people.

It includes lots of different ways of engaging with PPDAC from the problem stage to simple analysis tasks, to drawing conclusions from pre-existing data.

This would be suitable for upper primary learners who are ready to use large data sets with multiple variables.

SKILLS & CONCEPTS

- Identify different forms of chart and suggest an appropriate use
- Solve a range of graphical data problems
- Reflect on the importance of data in real-world decision-making

Office of National Statistics

https://edin.ac/3XA4I9v

Problem

Gapminder Cards: Data to Challenge our Assumptions

This activity encourages learners to explore the Gapminder dataset to find out up-to-date information about the world and challenge their assumptions. It is suitable for the Problem stage of PPDAC because exploring a data set often prompts curiosity and questioning – "I wonder if...". This can lead to planning to collect or find more data to answer the questions.

SKILLS & CONCEPTS

- Explore a real-world dataset
- Pose questions about the world based on looking at data
- Challenging assumptions

Data Education in Schools
https://edin.ac/3ZIA9zA

Plan

Books versus bean bags

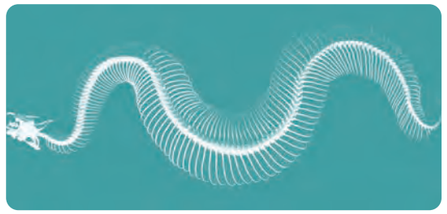
The learners are given the context of deciding whether the school library refurbishment should spend money on books or comfy furniture, and they have to plan how to collect data to find out.

The teachers' notes have useful questions for the learners, including how to spot bias in a sample.

SKILLS & CONCEPTS

Design an investigation that would lead to reliable results, using the enquiry cycle.

NZ Maths
https://edin.ac/3JetZ4L



Data

Dragonistics

This pack of Teachers' Guide, activities, videos and lesson plans centre around an attractive set of dragon trading cards. You can print a small set from the website, or borrow a class set from the Data Education in Schools team. It is a fun way to explore how information is structured, using categorical and numerical data.

SKILLS & CONCEPTS

- Structuring information
- Understanding sampling and variability in numerical and categorical data

Creative Maths
https://edin.ac/3GMoIhH

Data

Changing World

An exercise which explores ranking and rating data about learners' views on how technology has impacted people's lives. It's useful because it lays the foundations for designing surveys of people's attitudes and for correctly interpreting Likert scales (strongly disagree to strongly agree).

SKILLS & CONCEPTS

- Compare statements with the features of simple data displays from statistical investigations or probability activities undertaken by others
- Interpret tables showing rankings and ratings

NZ Maths
https://edin.ac/3GNsBmM

Data

Introduction to CODAP

Introduce your learners to the CODAP digital tool for analysing data. Try the Getting Started Tutorials 1 and 2: https://edin.ac/3iGrq0p

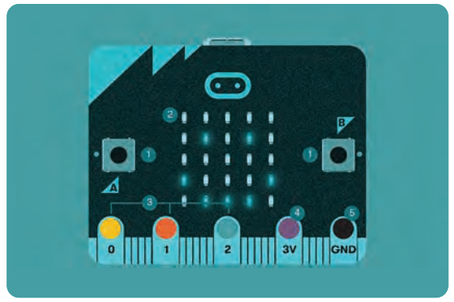
Then try these data treasure hunts, to check learners' data interpretation skills:

Mammals
https://edin.ac/3XlHT9b

Cats
https://edin.ac/3XGYiEO

SKILLS & CONCEPTS

Use a digital tool to manipulate datasets and create dot plots and bar charts as a way to understand data.



Data

Data Handling Unit

A series of five lessons using Microbits. Pupils learn about data through a variety of unplugged activities. They write and evaluate algorithms and programs using selection and repetition to use micro:bit as a temperature recorder, an automatic warning system and a digital assistant. This would be suitable for schools who have recently received their free Microbit devices and are looking for new projects to do with them.

SKILLS & CONCEPTS

- Understand, locate and use data
- Program sensors to collect data

Microbit Foundation
https://edin.ac/3w9ZrsO

Analysis

Deep sea diving

A short well structured activity to support learning how to interpret graphs which show ranges (in this case, the depths of sea creature habitats).

SKILLS AND CONCEPTS

- Conduct investigations using the statistical enquiry cycle
- Posing and answering questions
- Gathering, sorting, and displaying category and whole-number data
- Communicating findings based on the data
- Drawing and interpreting graphs which show ranges of values

NZ Maths
https://edin.ac/3IYLhmg

Analysis

Breaking down

Learners research how long different materials take to compost and chart them on a timeline. Open-ended activity with good teachers' notes and links to science and sustainability.

SKILLS AND CONCEPTS

- Represent data on a timeline
- Categorise data

NZ Maths
https://edin.ac/3iMD9ud

Conclusions

Gala Graphs

Learners interpret information from graphs which show frequencies of visitors at time points and then money at time points.

They write “stories” to explain what they see.

SKILLS & CONCEPTS

Interpret graphs; evaluate effectiveness of different displays in representing findings of a statistical investigation/probability activity undertaken by others.

NZ Maths
https://edin.ac/3XGHw8P

Conclusions

Too Much Telly

This exercise uses fictional data about TV watching/homework habits of children.

The learners interpret and draw conclusions from dot plots. There is an emphasis on using language to accurately describe results.

SKILLS & CONCEPTS

Compare statements with the features of simple data displays from statistical investigations or probability activities undertaken by others.

NZ Maths
https://edin.ac/3XHhmDd

Conclusions

Data Talks

Data Talks are a series of 10-15 minute class discussions to encourage learners and teachers to think about data together.

Each Data Talk comes with a single page containing a thought-provoking graph from a real research project, along with a set of questions to discuss together. There is a broad range of topics from Mindset to Melting Ice, so perhaps you can find a topic to match the interests of your class.

SKILLS & CONCEPTS

- Interpreting graphs
- Thinking of new questions to ask based on existing findings

Youcubed
youcubed.org/resource/data-talks/

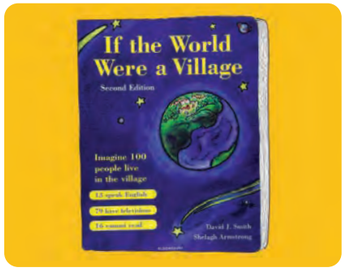
Data Education Resources

A library of resources created by the Data Education in Schools professional learning team. The library can be searched by curriculum level, area, topic & theme and by the Problem, Plan, Data, Analysis, Conclusion (PPDAC) framework stage.

Data Education in Schools
https://edin.ac/3Jc5fIo

Books for Learners

This list of books about data comes recommended by local primary school teachers who have enjoyed sharing them with their classes.



If the World Were a Village

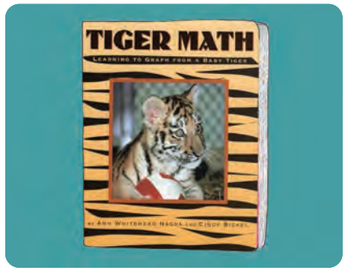
David J. Smith and *Shelagh Armstrong*

This picture book introduces children to concepts about variability in a population by illustrating the proportions of the sample by drawings of people in a village. For example, it shows how many people would have access to electricity, and how many people would speak different languages.

Beautiful News

David McCandless

This is a fascinating book full of visualisations which are both attractive and thought-provoking. It deliberately focuses on illustrating positive information about improvements in the world, particularly relating to green energy.



Tiger Math: *Learning to Graph from a Baby Tiger*

Ann Whitehead Nagda

Particularly suitable for children who love animals, this picture book explains how to interpret different sorts of graphs in the context of a true story of rearing a tiger cub at a zoo.

Observe, Collect, Draw! *A Visual Journal*

Giorgia Lupi and *Stefanie Posavec*

Written by two designers, this book is an invitation to sketch your own creative ways of representing data in your everyday life. It is a way to bring data topics to life for learners who love art. This would be useful to accompany the Dear Data resource.

Useful Datasets

Typical resources include:

Gapminder
www.gapminder.org
A collection of international datasets on health, education, poverty and other social issues.

Oceans of Data
https://edin.ac/3XtGRYF
Data sets and resources for teachers from the US organisation, Oceans of Data.

Data Science 4 Everyone
datascience4everyone.org/datasets
A collated list of resources and data sets for teachers from a US organisation which aims to promote data education in US schools.

Information is Beautiful
https://edin.ac/3R0BSMF
High-quality data visualisations which present facts about many aspects of society in a compelling way.

Scotland/UK-wide resources include:

Official UK government datasets
https://edin.ac/3iVNfci

Statistics about East Lothian
https://edin.ac/3H0Abun

Official Scottish Government statistics
https://edin.ac/3H1QNBI

Statistics about West Lothian
https://edin.ac/3QVKWTf

National Records of Scotland
https://edin.ac/3QXjZyq
Population data, available by geographical region.

Statistics about Scottish Borders
https://edin.ac/3JirEE4

Endnotes

1 OECD. (2019). *Well-being 2030 Action OECD Future of Education and Skills 2030: A Series of Concept Notes.* www.oecd.org.

2 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. P23

3 There are some great examples of job opportunities in data on our free poster at: https://dataschools.education/put-your-data-skills-to-work/

4 Wild, C. J. (2017). *Statistical Literacy as the Earth Moves.* Statistics Education Research Journal, 16(1), 31–37. P32

5 This quote is from a letter from Einstein to Carl Seelig, March 11, 1952, AEA 39-013: https://www.asl-associates.com/einsteinquotes.htm

6 You can find a complete list of data literacy related outcomes and expectations across the curriculum here: https://dataschools.education/resource/experiences-and-outcomes/

7 Spiegelhalter, D. J. *The Art of Statistics : Learning from Data* / David Spiegelhalter. UK: Pelican, an imprint of Penguin Books, 2019. Print. P18

8 Wild, C. J. (2017). *Statistical Literacy as the Earth Moves.* Statistics Education Research Journal, 16(1), 31–37. P32

9 Note that this is a slightly different meaning to “variable” in algebra or computer programming.

10 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.* https://www.amstat.org/education/guidelines-for-assessment-and-instruction-in-statistics-education-(gaise)-reports

11 Adapted from Wolff, A., Gooch, D., Montaner, C., Rashid, J. J., Kortuem, U., Wolff, A., … Kortuem, G. (2016). *Creating an Understanding of Data Literacy for a Data-driven Society.* The Journal of Community Informatics, 12(3), 9–26. Retrieved from www.ci-journal.net/index.php/ciej/article/view/1286.

12 You can find some examples of real-world datasets for classroom use here: https://dataschools.education/resource/seasonal-datasets/ and https://concord-consortium.github.io/codap-data/

13 https://5rightsfoundation.com/KnowYourRightsPoster.pdf

14 Harford, Tim. *How to Make the World Add up* / Tim Harford. The Bridge Street Press. 2020. Print. P294

15 Spiegelhalter, D. J. *The Art of Statistics : Learning from Data* / David Spiegelhalter. UK: Pelican, an imprint of Penguin Books, 2019. Print. P15

16 https://www.co2indicator.nl/documentatie/Ventilation-Rates-in-Schools-and-Pupils-Performance.pdf

17 http://www.iaquk.org.uk/ResourcesCO2.html

18 https://www.gapminder.org/tools/

19 https://codap.concord.org/

20 Harford, Tim. *How to Make the World Add up* / Tim Harford. The Bridge Street Press. 2020. Print. P10

21 Silver, Nate. *The Signal and the Noise : the Art and Science of Prediction* / Nate Silver. London: Penguin Books, 2013. Print. P9

22 Harford, Tim. *How to Make the World Add up* / Tim Harford. The Bridge Street Press. 2020. Print. P280

23 Available from: https://dataschools.education/data-education-resources/

24 Arnold, P., & Franklin, C. (2021). *What Makes a Good Statistical Question?* Journal of Statistics and Data Science Education, 29(1), 122–130. https://doi.org/10.1080/26939169.2021.1877582

Data Education in Schools

By Judy Robertson, Holly Linklater, Kate Farrell, Jasmeen Kanwal, Serdar Abaci, Claire Sowton and the Data Education in Schools team. Graphic design by Elspeth Maxwell using assets from UKblacktech.com, Freepik.com, and Unsplash.com.

Do you wince at the thought of data literacy, or sigh with relief that it is it now recognised as an essential part of all children and young people's learning?

Either way, this handbook is here to help.

Data literacy is the set of skills and concepts which people need to understand, interpret and make decisions based on the data they encounter in the world around them. There is no official area of the curriculum, or topic, labelled "data literacy" in the *Curriculum for Excellence* in Scotland. However it is relevant across the curriculum in outcomes for maths, literacy, technologies, and social studies.

This handbook sets out how, building on children and young people's curiosity about their world, teachers can enhance opportunities for all to build the skills and habits of mind relevant to data problem-solving.

- *Introduces and explains what is meant by 'data literacy' and how it relates to the school curriculum, including key terms and concepts*
- *Provides a clear framework for thinking about data problem-solving*
- *Reviews a wide range of resources for early, first and second level that can be used to support teaching and learning about data literacy*

THE UNIVERSITY of EDINBURGH

Data-Driven Innovation

Part of the Edinburgh & South East Scotland City Region Deal

UK Government

Scottish Government
Riaghaltas na h-Alba
gov.scot

CITY REGION DEAL
Edinburgh & South East Scotland

Data Skills Gateway

Data Education in Schools

The Data Education in Schools project is part of the Edinburgh and South East Scotland City Region Deal Data Skills Programme, funded by the Scottish and UK Governments. The Data Skills Programme brings together industry, universities, colleges, schools and others to provide routes into data or digital careers.

This handbook is part of the Data Education in Schools' commitment to developing an interdisciplinary data education curriculum for Scotland and a set of engaging real world data science teaching materials for primary and secondary school teachers.

www.dataschools.education