Department of Education

GOVERNMENT OF WESTERN AUSTRALIA

CURRICULUM RESOURCE MODULE

Big data

YEAR 9

Acknowledgements

The STEM Learning Project respectfully acknowledges the Traditional Custodians of the lands upon which our students and teachers live, learn and educate.

The STEM Learning Project is funded by the Western Australian Department of Education (the Department) and implemented by a consortium in STEM education comprising the Educational Computing Association of WA, the Mathematical Association of WA, the Science Teachers Association of WA and Scitech. We acknowledge and thank the teachers and schools who are the co-creators of these resources.

Copyright and intellectual property

Table of contents

The STEM Learning Project

The aim of the STEM Learning Project is to generate students' interest, enjoyment and engagement with STEM (Science, Technology, Engineering and Mathematics) and to encourage their ongoing participation in STEM both at school and in subsequent careers. The curriculum resources will support teachers to implement and extend the Western Australian Curriculum across Kindergarten to Year 12 and develop the general capabilities.

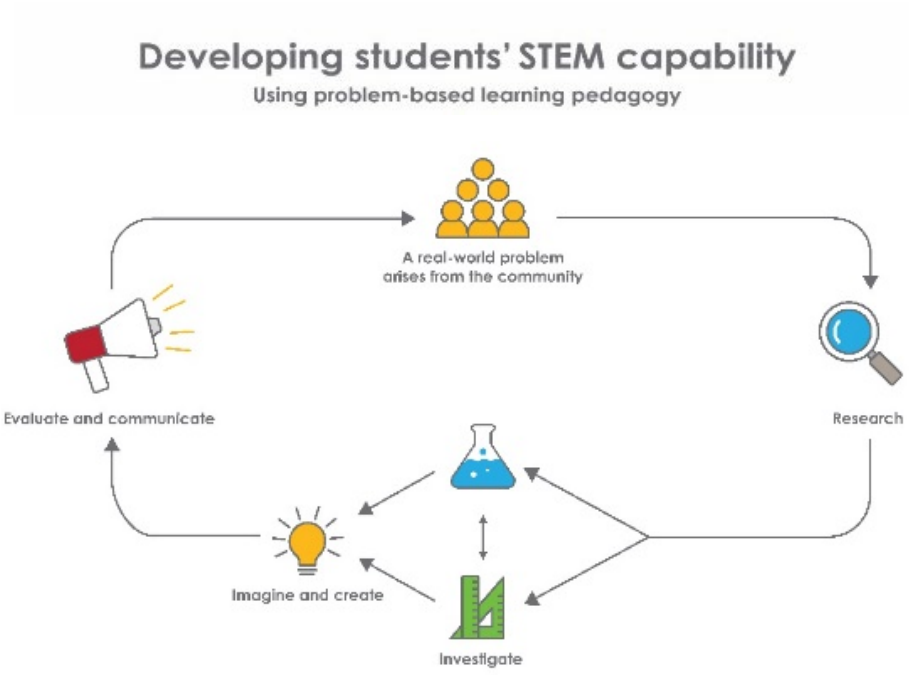
Why STEM?

A quality STEM education will develop the knowledge and intellectual skills to drive the innovation required to address global economic, social and environmental challenges.

STEM capability is the key to navigating the employment landscape changed by globalisation and digital disruption. Routine manual and cognitive jobs are in decline whilst non-routine cognitive jobs are growing strongly in Australia. Seventy-five per cent of the jobs in the emerging economy will require critical and creative thinking and problem-solving, supported by skills of collaboration, teamwork and literacy in mathematics, science and technology. This is what we call STEM capability. The vision is to respond to the challenges of today and tomorrow by preparing students for a world that requires multidisciplinary STEM thinking and capability.

The approach

STEM capabilities are developed when students are challenged to solve open-ended, real-world problems that engage students in the processes of the STEM disciplines.



STEM Consortium

Year 9 – Big data

Overview

The International Data Corporation (IDC) predicts that by 2025 worldwide data will grow by 61% to 175 zetabytes. This datasphere is described as the collective sum of the world's data, which is growing at an exponential rate. For example, every two days humans are generating as much data as was generated from the dawn of humanity through the year 2003.

Data age 2025 whitepaper (IDC, 2018) *www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf*,

While analysing sets of data is nothing new, the evolution of a new scientific discipline called *big data analytics* (often abbreviated to big data) has emerged due to the sheer volume of data now being generated (much of it through digital devices) and the sophisticated capabilities of Artificial Intelligence (AI).

The key distinguishing feature of big data is that it involves data sets that are too large or complex to analyse by traditional methods and can only be analysed via a computer.

In short, big data is all about computational analysis giving us the ability to see the entire planet for the first time as one giant interconnected set of data. This allows us to understand and potentially solve problems on a global scale such as homelessness, traffic, school results, food supply, the economy, education, medicine and climate change.

What is the context?

In this digitally connected world, digital interactions and daily activities like searching online, using debit or credit cards, tracking a workout, buying takeaway or using commercial transport all leave behind a trail of data. This data is continuously collected and stored.

Businesses, governments, scientists and many of the apps and services we use every day already use big data analytics to interpret this pool of data, which can have an exponential creation growth rate.

For example, in the health industry, data collected by smartphone health apps use big data analytics to help find links between lifestyle and illness. Similarly, rail services use ticket data to map travel patterns and improve services, while Census data is used to plan for new schools, shopping centres and hospitals based on population demographics.

Many questions are raised with the creation and usage of data such as: So what? Why is this data important? How can this data be used effectively?

The influx of data through social networks, business and government websites, and the increasing number of websites and tools for global monitoring in different disciplines reflects the growing interest in data-based decision making.

Understanding the nature of the captured data and anticipating the types of innovations or new knowledge that can be built upon this data is increasing the role of the data scientist.

In this module, students consider their contribution to big data through their daily activities and online interactions and gain insight into big data analysis. They take on the role of the data scientist to use big data to identify real-world problems and design possible solutions.

The United Nations has developed 17 *Sustainable Development Goals* (United Nations, n.d.) *https://sustainabledevelopment.un.org/* that are designed to address some of the biggest issues facing the planet.

Each of the sustainable development goals has several targets that are designed to be met by 2030. These targets are based around key problems and issues that have been identified using data analysis of many factors including mortality rates, income, poverty and education. These goals provide an authentic context in which students can not only explore the concept of big data but also examine authentic examples of how data is being used to help generate solutions. For example, many of these goals are being addressed through organisational initiatives (eg *UN Global Pulse - Big data for development and humanitarian action* (UN Global Pulse, 2018) *www.unglobalpulse.org/* and the *Data collaboratives explorer* (GovLab, n.d.) *https://datacollaboratives.org/explorer.html*)and company projects designed to solve these problems using big data analysis (eg *The plume air cloud* (CORDIS, 2020) *https://cordis.europa.eu/project/id/778501*, *Driving directions, traffic reports & carpool rideshares by Waze* (Waze Mobile, 2020) *www.waze.com/*).

Students in this module will identify a real-world problem and determine how and what big data can be used to address, mitigate or solve the problem.

What is the problem?

How can big data analysis be used to help solve global problems?

How does this module support integration of the STEM disciplines?

Science

Biological Science knowledge is developed as students consider how ecosystems consist of communities of interdependent organisms as they examine factors that affect population sizes such as seasonal changes (ACSSU176), and use scientific understanding to investigate how such models can be used to predict changes in populations due to environmental changes (ACSHE157). Students examine how values and needs influence scientific research, and investigate how scientific and technological advances can be applied as they create and analyse large data sets to identify problems and develop solutions in influencing new trends and innovations (ACSHE228).

Students use internet research to identify problems and formulate questions that can be investigated scientifically (ACSIS164). They evaluate information from secondary sources as part of the research process, they revise and refine research questions to target specific information and data collection. Students develop ideas from their own or others' investigations and experiences to investigate further.

Students collect reliable data for modelling and simulations and use digital technology to investigate situations and events (ACSIS165). They analyse patterns and trends in data, including describing relationships between variables and identifying inconsistencies, and use spreadsheets to present data in tables and graphical forms and to carry out mathematical analyses on data (ACSIS169). Students also design and construct appropriate graphs to represent data and analyse graphs for trends and patterns.

Students use knowledge of scientific concepts to draw conclusions that are consistent with the evidence and can suggest more than one possible explanation of the data presented (ACSIS170). Students also evaluate conclusions and identify alternative explanations that are also consistent with the evidence (ACSIS171), and critically analyse the validity of information in primary and secondary sources and evaluate the approaches used to solve problems (ACSIS172).

Students communicate scientific ideas and information for a particular purpose, they construct evidence-based arguments and use appropriate scientific language, conventions and representations (ACSIS174) when presenting results in oral presentations and in contributing to group discussions.

Technology

As students determine how big data analysis can be used to help solve problems identified by the United Nations, they develop techniques for acquiring, quantitative and qualitative data from a range of sources (ACTDIP036) to determine what may be considered a significant problem to research. They analyse and visualise data to create information around their chosen problem (ACTDIP037). Students create and use interactive solutions for sharing ideas and information online, considering safety, social contexts and legal responsibilities (ACTDIP043).

Students identify and define the needs of a stakeholder (WATPPS54) as they investigate a selection of resources to develop solution ideas, identifying and considering constraints (WATPPS55) and apply design thinking, creativity and enterprise skills (WATPPS56). Students design solutions, assessing alternative designs against given criteria (WATPPS57), evaluate design processes and solutions against student-developed criteria (WATPPS59), and work independently and collaboratively using digital technology and an iterative and collaborative approach (WATPPS60).

The *Design process guide* has also been included as a resource to assist teachers in understanding the complete design process as developed in the Technologies curriculum.

Mathematics

Mathematic understandings are developed in *Activity 1* and used throughout the other activities. Students investigate reports of surveys for information on how data were obtained to estimate population means and medians (ACMSP227). They identify everyday questions and issues involving at least one numerical and at least one categorical variable, collect data directly from secondary sources (ACMSP228), and compare data displays using mean and median to describe and interpret numerical data sets (ACMSP283).

General capabilities

There are opportunities for the development of general capabilities and cross-curriculum priorities as students engage with *Big data*. In this module, students:

- Develop problem-solving skills as they research the problem and its context (*Activity 1*); investigate parameters impacting on the problem (*Activity 2*); imagine and develop solutions (*Activity 3*); and evaluate and communicate their solutions to an audience (*Activity 4*).
- Utilise creative thinking as they generate possible design solutions; and critical thinking, numeracy skills and ethical understanding as they choose between alternative approaches to solving the problem of using big data analysis to help solve problems identified by the United Nations in its Sustainable Development Goals.
- Utilise personal and social capability as they develop socially cohesive and effective working teams; collaborate in generating solutions; adopt group roles; and reflect on their group work capabilities through self and peer evaluation.
- Utilise a range of literacies as they collate records of work completed in a journal; work with data and data representations; and represent and communicate their solutions to an audience using digital technologies in *Activity 4*.
- Develop information and communication technology (ICT) capability while using big data analysis and using spreadsheet formulas and graphing functions.

istockphoto.com

What are the pedagogical principles of the STEM learning modules?

The STEM Learning Project modules develop STEM capabilities by challenging students to solve real-world problems set in authentic contexts. The problems engage students in the STEM disciplines and provide opportunities for developing higher order thinking and reasoning, and the general capabilities of creativity, critical thinking, communication and collaboration.

The design of the modules is based on four pedagogical principles:

- Problem-based learning
  This is an underlying part of all modules with every module based around solving an initial problem. It is supported through a four-phase instructional model: research the problem and its context; investigate the parameters impacting on the problem; design and develop solutions to the problem; and evaluate and communicate solutions to an authentic audience.
- Developing higher order thinking
  Opportunities are created for higher order thinking and reasoning through questioning and discourse that elicits students' thinking, prompts and scaffolds explanations, and requires students to justify their claims. Opportunities for making reasoning visible through discourse are highlighted in the modules with the icon shown here.

- Collaborative learning
  This provides opportunities for students to develop teamwork and leadership skills, challenge each other's ideas, and co-construct explanations and solutions. Information that can support teachers with aspects of collaborative learning is included in the resource sheets.
- Reflective practice
  Recording observations, ideas and one's reflections on the learning experiences in some form of journal fosters deeper engagement and metacognitive awareness of what is being learnt. Information that can support teachers with journaling is included in the resource sheets.

These pedagogical principles can be explored further in the STEM Learning Project online professional learning modules located in Connect Resources.

Activity sequence and purpose

Activity 1

RESEARCH

Big data analytics

Students develop an understanding of the nature and application of data as they research the concept of big data analytics. They create and interrogate large data sets from authoritative sources and analyse the data that are collected.

Activity 2

INVESTIGATE

Investigating data sets

Students locate and investigate data from a range of datasets. They explore the various ways data can be represented and determine which representations will best communicate their solution.

Activity 3

IMAGINE & CREATE

Big data solutions

Students apply understandings about big data as they identify a problem and develop data sets to support their design decisions.

Students describe the key features of and develop implementation strategies for their innovation.

Activity 4

EVALUATE & COMMUNICATE

Showcase

Groups present their innovation at a showcase event, highlighting the importance of big data analysis in informing their innovation.

Background

Expected learning	Students will be able to: 1. Examine and explain factors that affect population sizes. 2. Analyse information and data to construct evidence-based conclusions. 3. Analyse, process and identify patterns and trends in data. 4. Interpret and analyse data using scientific and mathematical principles. 5. Define 'big data' and identify some of the major sources from which it is derived. 6. Understand how big data is being used by organisations such as the United Nations to identify and solve real-world problems. 7. Interrogate datasets from a range of sources and use data analysis to form and support conclusions. 8. Enter data into a spreadsheet, generate data visualisations to compare and analyse the data. 9. Based on data analysis, design and create a product that will address an identified issue or need. 10. Evaluate innovations that have been designed using the design process and which is supported by data analysis. 11. Work collaboratively to plan, develop and communicate ideas and information.
Vocabulary	This module uses subject-specific terminology, some of which is shown in *Teacher resource sheet 1.5: Vocabulary*. The list contains terms that need to be understood, either before the module commences or developed as they are used.
Timing	There is no prescribed duration for this module. The module is designed to be flexible enough for teachers to adapt. Activities do not equate to lessons; one activity may require more than one lesson to implement.
Consumable materials	There are no consumable resources required for this module.

Safety notes	Potential hazards specific to this module include but are not limited to: • Possible exposure to cyber bullying, privacy violations and uninvited solicitations when using the internet.
Enterprise skills	The *Big data* module focuses on developing outcomes related to the general capabilities and enterprise skills. Enterprise skills include problem solving, communication skills, digital literacy, teamwork, financial literacy, creativity, critical thinking and presentation skills. Further background is available from the *Foundation for Young Australians.* An example is *The New Basics: Big data reveals the skills young people need for the New Work Order* (Foundation for Young Australians, 2016) *www.fya.org.au/wp-content/uploads/2016/04/The-New-Basics_Web_Final.pdf*
Assessment	The STEM modules have been developed to provide students with learning experiences to solve authentic real-world problems using science, technology, engineering and mathematics capabilities. While working through the module, the following assessment opportunities will arise: • Anecdotal notes of observations as students work collaboratively through the activities • Data analysis through spreadsheet use • Reflective and content focused journal entries • Design thinking through a multimedia presentation • Reflections, peer and self-evaluations and justification of understandings when students present their learning in *Activity 4*. *Links to the Western Australian Curriculum* indicates the expected learning students will engage in as they work through the module. Evidence of learning from journaling, presentations and anecdotal notes from this module can contribute towards the larger body of evidence gathered throughout a teaching period and can be used to make on-balance judgements about the quality of learning demonstrated by

	the students in the science, technologies and mathematics learning areas. Students can further develop the general capabilities of Information and communication technology (ICT) capability, Critical and creative thinking and Personal and social capability. Continuums for these are included in the *General capabilities continuums* but are not intended to be for assessment purposes.

Activity 1: Big data analytics

Activity focus 	Students develop an understanding of the nature and application of data as they research the concept of big data analytics. They create and interrogate large data sets from authoritative sources and analyse the data that are collected.
Background information	**What is 'big data'?** Big data is a simple term used to describe the emergence of incredibly powerful ways to gather and analyse digital information using multiple computers working together. Large, complex and layered groups of data can be analysed to identify patterns and trends. Through the use of data from multiple sources, we can gain deeper and more accurate information that can be used to our benefit. The creation and accumulation of data are growing at an exponential rate. According to *IBM Big data analytics* (IBM, 2019) https://www.ibm.com/au-en/analytics/hadoop/big-data-analytics, each day about 30 billion emails are sent, more than 230 million Tweets are made, 45 billion Facebook messages are sent and 3.5 billion Google searches are made. Added to that is the predicted 6.1 billion smartphone users sending, interacting, storing and retrieving data at an unprecedented rate. While this data alone may have little value in and of itself, when combined with other data it can provide valuable information about trends, patterns and relationships which would have otherwise not been identified. **Where does big data come from?** In such a digitally connected world, any kind of digital interaction will create a trail of data that is collected and stored. Big data also comes from non-human sources, such as weather data and financial market information. Data about people's everyday life is constantly being created, from the apps on their phones, the time their train arrives (and when it leaves), how much they spend, and even what they spend on it. All this information (and much more) is collected and added to the ever-growing pool of big data. This data can then be analysed and interpreted

	and used for multiple purposes to gain greater understanding of what people do and when they do it.
Instructional procedures	It is recommended that students work in the same small groups of four for all activities. Mixed ability groups encourage peer tutoring and collaboration in problem-solving. Collaboration is an important STEM capability. Throughout this module, students will be required to keep track of their progress and reflect on their experience. Therefore, it is necessary to consider how this will be carried out. See *Student journal* for more information. The journal could be handwritten or electronic using a program such as *Sway*, *Microsoft Word*, *OneNote* or Weebly. *Student activity sheet 1.0: Journal checklist* also supports students' journaling while working through the module.
Expected learning	Students will be able to: 1. Identify and explain the concept of 'big data' and distinguish between data and information (Mathematics, Science and Technologies). 2. Perform mathematical operations on rainfall and kangaroo population data sets to calculate values for the mean, median and range (Mathematics). 3. Investigate how data can be used to predict changes in populations due to environmental changes (Science). 4. Extract climate data from the Bureau of Meteorology, store it in a format that is useful for analysis, and visualise the data to create information about rainfall trends (Mathematics, Science and Technologies). 5. Link big data analysis to problem identification and solution-finding and identify where it can be used to assist humanity (Mathematics and Technologies). 6. Identify and critically analyse patterns and trends in rainfall, population and big data sets, and use knowledge of scientific concepts to draw conclusions that are consistent with the evidence (Mathematics and Science).
Equipment required	**For the class:** Device for projection and internet connection

	For the students: Digital device with internet access, and access to a spreadsheet app *Student activity sheet 1.6: Annual rainfall comparisons* *Student activity sheet 1.8: Kangaroo populations and rainfall comparisons* *Student activity sheet 1.10: Gapminder data 'deep dive'* *Student resource sheet 1.13: Data sets from around the world* *Student resource sheet 1.14: Australian data sets* *Student resource sheet 1.15: United Nations Sustainable Development Goals data sets*
Preparation	Test links to videos and download online resources in case of issues connecting to the internet. Provide students with access to resource sheets. Organise students into small groups of four.
Activity parts	**Part 1: Data versus information** Begin by asking the question: *What is data?* Record and display student responses as a brainstorm. This could be done using sticky notes added to a question wall, on a whiteboard or using a digital platform such as *Padlet*. Ensure this information is available for review and reflection in later activities. Smartphone data Encourage students to think about how and what data is collected from various sources and devices both intentionally (eg formal data collection through completing forms and registrations etc.) and incidentally (eg informal data collection through parking meters, shopping transactions and location data). Working in their groups, students use the *placemat strategy* to brainstorm ideas about the ways a smartphone is used and what kind of data may be collected and shared.

Ask students the following questions to help direct their thinking:

- How is the data collected (intentionally and/or unintentionally)? Why do you think this?
- What kind of data is collected?
- Is the data qualitative or quantitative?

Quantitative data are measures of values or counts and are expressed as numbers (eg how many; how much; or how often). Qualitative data are measures of 'types' and may be represented by a name, symbol, or a number code (eg what type).

- Who collects this data?
- What do you think this data is used for?
- What could this data reveal about their actions, preferences and behaviours?

After initial responses are recorded, students access digital resources such as *How Google Collects Data about you on the Internet*, *Data vs Information*, *Reinventing society in the wake of big data*, *The Human Face of Big Data*, *Target knows when you're pregnant* and *What is Big Data?* (see *Digital resources*).

Discuss the differences between data and information. Record student responses as a brainstorm using the following prompt questions as headings:

- What is the difference between data and information?
- How can both data and information be represented, and in what formats can they be found?
- Who collects data and information?
- Who uses the data and information?
- What data may be generated by what students post or view on social media?
- What information may this reveal to others about them?
- To whom may they be revealing this information?

Encourage students to also consider if any data sets could be combined to generate information that would otherwise be difficult to obtain. For example, ask:

- Could there be any links between the weather data (eg temperature) and the volume of cold drinks purchased from vending machines?

School excursion scenario

Provide students with a scenario where an outside company wants to pay for a Year 9 class to go on an excursion. This will involve both an activity and class lunch at a 5-star restaurant in Perth.

Use a *think, pair, share strategy* to determine what data would need to be collected by the company to assist in the planning. Encourage the class to participate in the 'share' stage, and record class responses using the following prompt questions as headings:

- How will the data be collected?
- What suitable survey questions should be asked? How will these allow useful data to be collected?
- What actual data and information does the company need?
- Does the company need access to both the data and the information? Why or why not?
- Who will collect and store this data/information?
- How will the data be analysed?
- What would be the most appropriate way to communicate the results?

Ask students to reflect upon their understandings of data and information and summarise them in their journal.

Part 2: Comparisons of annual rainfall

Weather forecasting requires vast amounts of data to be collected and included in the modelling. Data that is often incorporated include temperature, air pressure, wind speed and direction, humidity, amount of precipitation, cloud coverage and type of cloud. These and other data are collected by weather observers, radiosondes (balloons), radar and satellite imagery.

In Australia, the Bureau of Meteorology collects, analyses and stores weather and climate data. This data is important for weather and long-range weather forecasting, and to identify long-term trends, which are usually referred to as climate.

Students read through *Student activity sheet 1.6: Annual rainfall comparisons*. Working in pairs, students use a spreadsheet to compare annual rainfall data for three regions in Western Australia. Students collect secondary data from the Bureau of Meteorology and estimate the mean and median rainfalls for three weather stations within each of these regions.

Discuss the answers to this activity as a class and ask students to reflect on the following, encouraging them to use relevant language:

- What rainfall value is the most appropriate to use when describing the annual rainfall for the regions?
- What did you notice about your calculated values for the mean annual rainfall when compared to the 'summary statistics for all years' values?
- What did you notice about the rainfalls for the weather stations within similar districts? What information could this provide about their locations?

Teacher resource sheet 1.7: Exemplar spreadsheet for annual rainfall comparisons provides a sample that contains the appropriate data with mean and median rainfalls calculated.

Part 3: Kangaroo populations

It is estimated that there are up to 50 million kangaroos in Australia, although unlike humans, the exact number is difficult to determine. Estimates of populations throughout Australia are based on ground and aerial surveys where commercial harvesting occurs. Where individual states allow commercial harvesting, management plans require population figures to be collected. Since the whole of the country is not surveyed, the actual national population would be significantly higher if estimates for areas not surveyed were included. In Western Australia, populations of

Red and Western Grey kangaroos are estimated for commercial harvesting purposes.

Students read through *Student activity sheet 1.8: Kangaroo populations and rainfall comparisons*. Working in their groups, students use the sample *Student kangaroo population* spreadsheet to compare annual Western Grey and Red kangaroo populations and rainfall data for districts in Western Australia. Students use secondary Bureau of Meteorology climate data and kangaroo population estimates to start to construct their own big data set for analysis.

Students analyse the data to determine if rainfall influences kangaroo populations in Western Australia. Using data from 2001–2018, students calculate the annual rainfall within a rainfall district, estimate the mean and median rainfalls for three rainfall districts within the kangaroo monitoring regions, and compare the rainfall trends to that of the populations for each of the two species.

Discuss students' answers as a class and ask them to reflect on the following questions. Encourage students to use relevant language and to use their data analysis to support their responses:

- What trends did you observe when you compared the population and rainfall data?
- Were the trends consistent across both species of kangaroo?
- Were the observed trends consistent across the class? What may have been analysed differently by different groups?
- What effect might differences in wet seasons have on the data being analysed?
- What additional data or information could be useful to determine the effect of rainfall on kangaroo populations in Western Australia?
- Is rainfall the only influence on kangaroo populations?
- What else could impact kangaroo populations?
- What other data sets would be useful to compare the kangaroo population?
- What could have influenced the data you have worked with?

Teacher resource sheet 1.9: Exemplar spreadsheet for kangaroo populations and rainfall comparisons and the sample *Exemplar teacher kangaroo* spreadsheet provides the appropriate data with annual total, mean and median rainfalls calculated. Visual representations of the data are also shown to allow patterns and relationships between the data and population information to be identified.

Part 4: Data 'deep dive'

Students explore the *Gapminder* website to examine large data sets. They closely examine the tools and data menu categories used by *Gapminder* and find data that is of interest to them.

Using the class projection device, demonstrate how to access the 'tools' and 'data' dropdown menus. Using a 'Bubble' data representation, choose 'Life expectancy' for the vertical axis and 'Income' for the horizontal axis, and then 'play' the data to visualise how it changes with time. Select 'Australia' and then 'play' again to create the trace on the visualisation.

Working in their cooperative group, students individually record their findings on the *Student activity sheet 1.10: Gapminder data 'deep dive'*. This allows them to examine the data visualisations and data sources, help them organise their findings, and start to consider who might use this information.

Encourage students to explore data visualisation tools (Bubbles, Income, Maps, Trends, Ranks, Ages, Spreadsheet) to best show the data, and to use the 'Data' menu to learn more about the 'indicators' and where the data is sourced.

Encourage students to use a *jigsaw strategy* to summarise information from the resources and share their findings.

Ask students the following questions to help direct their thinking:

- What types of data did you find?
- What are the main categories for the various data sets?
- How is the data represented?
- What type of data representation do you find most useful?

- What unexpected indicators (variables) did you find?
- What were some interesting facts that you found?
- Did you find any 'odd' information or trends?
- Where does *Gapminder* source their data?

This activity could be extended by introducing students to other large data sets (eg *Data.gov*, the *Australian Bureau of Statistics*, *The World Factbook*) to compare how and what data is collected. See *Digital resources* for examples.

Part 5: What is big data and how is it being used

Big data analysis can provide a vast amount of otherwise unbelievable information, and big data is often described using characteristics such as volume, variety, velocity, variability and veracity (see *Teacher resource sheet 1.5: Vocabulary*).

Discuss the definitions and ask students to reflect on why these might be important to consider when looking to use big data. Use *Teacher resource sheet 1.11: Big data fact sheet 1 – What is big data?* for prompts to further discussion such as What's a byte?, What's creating the deluge of information? and Big data facts and stats.

Provide students with access to *Student resource sheet 1.13: Data sets from around the world* and *Student resource sheet 1.14: Australian data sets*. Students work in pairs and need to access at least three data sets from each of the International and Australian contexts (at least six in total).

Ask students the following questions to help direct their thinking and have them write their answers directly on the resource or into their journal to allow participation in a class discussion on their findings.

- What are the main categories for the data sets?
- How is the data represented?
- Who is most likely to use this data set and what information may be gained?

Discuss the answers to these questions for the list of data sets in the resources. This will allow students to learn about the data sets they have yet to explore.

Provide students with an opportunity to further explore their choice of big data sets from the resources. Students need to find at least two interesting outcomes in the big data sets they have interrogated. This will allow students to feel that they can explore data on a topic of interest to them. Once complete, students use a modified *think, pair, share strategy* to 'share' these interesting outcomes to other pairs or with the whole class.

After having examined a range of data sets, students explore the various applications of big data and how they are applied to a range of disciplines including health, town planning, media, insurance etc. Provide students with examples of where big data has been used to affect change. *Student resource sheet 1.12: Big data fact sheet 2 – Applications of big data, now and in the future* can be used to stimulate thinking and provide examples of big data applications. Ask students to consider the following guiding questions when considering these examples:

- What type of data is needed?
- Is this data already readily accessible? If not, where could this data be sourced?
- How could additional big data be used to provide information or insights into these applications?

To conclude this activity, use the above guiding questions to discuss the data types needed for some of the example applications.

The interrogation of big data has also been designed to introduce students incidentally to the United Nations Sustainable Development Goals, which they will use in subsequent activities.

Part 6: The Sustainable Development Goals

Introduce students to the *United Nations Sustainable Development Goals Project* by showing the video *Big Data for Sustainable Development – UN Global Pulse Project*.

In pairs, students explore the range of resources on the United Nations Sustainable Development Goals including the United Nations Sustainable Development Goals Project, *Big Data for Sustainable Development* and *About the Sustainable Development Goals*. Students should access each of the goals and consider how big data can be used

to find solutions to the goal targets. Students should also access the data infographics for each of the goals.

Students can further their understanding of the United Nations Sustainable Development Goals by interrogating some of the links on *Student resource sheet 1.15: United Nations Sustainable Development Goals data sets*.

Encourage students to consider the following questions when accessing these data set links:

- What Sustainable Development Goal does the data set allow new solutions to be found for?
- How is the data represented?
- What information does the data provide?
- What was the new solution that came out of the use of this big data?

Provide time for student reflection and journaling of key information and ideas from this activity. Also, it may be useful to have a class discussion about which Sustainable Development Goals the students feel are most easily supported by the big data they have accessed.

Resource sheets

Student journal

Teacher resource sheet 1.2: Cooperative learning - Jigsaw

Teacher resource sheet 1.3: Cooperative learning – Placemat

Teacher resource sheet 1.4: Cooperative learning – Think, Pair, Share

Teacher resource sheet 1.5: Vocabulary

Student activity sheet 1.6: Annual rainfall comparisons

Teacher resource sheet 1.7: Exemplar spreadsheet for annual rainfall comparisons

Student activity sheet 1.8: Kangaroo populations and rainfall comparisons

Student kangaroo population spreadsheet

Teacher resource sheet 1.9: Exemplar spreadsheet for kangaroo populations and rainfall comparisons

Teacher exemplar kangaroo spreadsheet

Student activity sheet 1.10: Gapminder data 'deep dive'

	Teacher resource sheet 1.11: Big data fact sheet 1 – What is big data? *Teacher resource sheet 1.12: Big data fact sheet 2 – Applications of big data, now and in the future* *Student resource sheet 1.13: Data sets from around the world* *Student resource sheet 1.14: Australian data sets* *Student resource sheet 1.15: United Nations Sustainable Development Goals data sets*
Digital resources	How Google Collects Data About You and the Internet (Pingdom AB, 2018) *royal.pingdom.com/how-google-collects-data-about-you-and-the-internet/*
	Data vs Information (EDUCBA, 2019) educba.com/data-vs-information/
	Reinventing society in the wake of big data (Edge Foundation, 2012) *edge.org/conversation/alex_sandy_pentland-reinventing-society-in-the-wake-of-big-data*
	The Human face of Big Data (PBS, 2016) *youtu.be/kAZ8lK224Kw*
	Target knows when you're pregnant (WCPO 9, 2012) *youtube.com/watch?v=XH1wQEgROg4*
	What is Big Data? (World Economic Forum, 2016) *youtube.com/watch?v=eVSfJhssXUA*
	Computing basic statistics (Khan Academy, n.d.) *khanacademy.org/computing/ap-computer-science-principles/data-analysis-101/data-tools/a/learning-from-data-sets*
	Finding patterns in data sets (Khan Academy, n.d.) *khanacademy.org/computing/ap-computer-science-principles/data-analysis-101/data-tools/a/finding-patterns-in-data-sets*

Mean, median, & mode example (Khan Academy, n.d.) *khanacademy.org/math/cc-sixth-grade-math/cc-6th-data-statistics/mean-and-median/v/mean-median-and-mode*
Climate Data Online (Australian Government Bureau of Meteorology, 2020) *bom.gov.au/climate/data/*
Gapminder (Gapminder Foundation, n.d.) *gapminder.org/*
Australian Bureau of Statistics (Commonwealth of Australia, n.d.) *abs.gov.au/*
Data.gov (U.S. Government, 2020) *data.gov/*
The World Factbook (Central Intelligence Agency, n.d.) *cia.gov/library/publications/the-world-factbook/*
Doing Business Data (The World Bank Group, 2020) *doingbusiness.org/en/data*
The United Nations Sustainable Development Goals Project (UNSDGP, n.d.) *unsdgproject.com/*
Big Data for Sustainable Development - UN Global Pulse projects (UN Global Pulse, 2016) *youtube.com/watch?v=v-zGHqMyd7o&feature=youtu.be*
Big Data for Sustainable Development (United Nations, 2017) *un.org/en/sections/issues-depth/big-data-sustainable-development/index.html*
About the Sustainable Development Goals (United Nations, n.d.) *un.org/sustainabledevelopment/sustainable-development-goals/*

Activity 2: Investigating data sets

Activity focus 	Students locate and investigate data from a range of datasets. They explore the various ways data can be represented and determine which representations will best communicate their solution.
Background information	The data sets provided in this activity allow students to analyse authentic and often live data sets as they develop their understanding of a range of mathematical concepts, especially around probability and statistics.
Instructional procedures	Establish a balance between the available information; too much overwhelms and misdirects students from their task, while too little does not provide the necessary scope for them to access relevant data sets to analyse. Teachers should familiarise themselves with the resources in the appendices and decide on the scope of the tasks they will set for students. If specific mathematical concepts need to be taught, then some data sets will be more useful and relevant than others. The ability level of students may also dictate the complexity of the data sets chosen. When students are undertaking data analysis on large data sets in Part 2 of this activity, opportunities arise for students to analyse data from the different science understanding content strands. When using the *Gapminder* website, students could analyse data from the Biological Sciences (examining factors that affect population sizes), Earth and Space Sciences (occurrence of earthquakes and volcanic activity) or the Chemical Sciences (describing the effect of the products of combustion on the environment).
Expected learning	Students will be able to: 1. Interpret and compare data sets that reflect the United Nations Sustainable Development Goals (Mathematics and Technologies). 2. Perform mathematical operations on data sets to visualise data and create information (Mathematics and Technologies).

	3. Investigate and critically analyse a range of big data sets for patterns and trends (Mathematics, and Technologies). 4. Identify alternative explanations that are also consistent with the observed evidence (Mathematics and Science).
Equipment required	**For the class:** Device for projection and internet connection Access to online articles and video clips in the *Digital resources* **For the students**: Digital device with internet access, and access to a spreadsheet app *Student resource sheet 1.13: Data sets from around the world*
Preparation	Ensure relevant data sets are available to students, either through the resource material provided or as a set of online links they can follow. Review all of the data sets and establish what mathematical and technology concepts the students need to apply to their chosen data sets. Conducting a modelled lesson around one of the sets could be useful.
Activity parts	**Part 1: Finding patterns in data sets** Discuss the definition of correlation with the class. Select a series of *spurious correlations* examples (see *Digital resources*) and show them one by one on the class projection device or interactive whiteboard. Using the *think, pair, share strategy*, students consider what each example shows. Ask students the following questions to help direct their thinking: • What is the graph showing? • How is the data related? • How likely is it that one causes the other? Students watch the video on *Correlation and causality*, access the *Gapminder – Correlations, Anomalies, and Trends*

examples and consider the importance of causality when comparing data sets (see *Digital resources*).

Students watch the *Gapminder CO2 emissions and child mortality* Bubble visualisation (see *Digital resources*). Ask students to conduct a think, pair, share and use the following questions to help elicit student thinking:

- Does this show a spurious correlation?
- In this example, does correlation imply causation? Why?
- Is child mortality greater for lower CO2 emissions?
- What else needs to be considered when looking at this visualisation?

Discuss the importance of interrogating exactly what the data is showing.

In pairs, students explore *Gapminder* and find two patterns in the data or visualisations that could indicate an abnormal trend or spurious correlation. Students report their findings to the class.

Part 2: Data analysis

Students perform data analysis on large data sets.

As a class, view the *Life expectancy and income* visualisation from the *Gapminder* website which shows the life expectancy and income change over time (see *Digital resources*). Students should make notes on their observations. Hold a class discussion to explore students' observations.

Working individually or in pairs, students collect life expectancy and income raw data (csv or xlsx) by selecting these indicators in the Data menu.

Students choose three countries, each from a different continent, and create at least four charts or graphs based on this data. Students should include at least three facts about high vs low life expectancy and income (things that stand out for them) for the countries they are comparing.

Students should also include at least three facts about the economy for these selected countries that they find from another source (see *Student resource sheet 1.13: Data sets from around the world*). These graphs and facts will be

displayed as an infographic on a single sheet, either printed or digitally. How students present will depend on whether they worked alone or in groups. Students share their findings with another partner (or pair).

In their groups, students use either *Gapminder* or other appropriate big data sets (see *Student resource sheet 1.13: Data sets from around the world*) to investigate different indicators for three countries of their choice. Students collect raw data (csv or xls), create a raw data spreadsheet, and produce four charts or graphs based on this data. Students can choose to use more than two data sets and need to include at least three facts based on the environment or populations of these chosen countries. These facts can be from other sources but need to back up or reinforce the raw data the group is interrogating.

Students will present this information to another group in the next part of this activity. Allow students to choose the format they want to use for their presentation.

Direct student thinking by using the following prompts:

- What is the data telling us?
- What question do our findings answer?
- What are the key findings we discovered?
- How can we best present this information?
- What other data could support these findings?

Part 3: Explanation of data

In their groups, students present their data analysis findings from *Part 2* to another group. Groups discuss the data they have identified and analysed, explain 'what this data is telling them', and identify any additional data that could be useful to interrogate further based on relevant environmental or population data facts they found.

While the first group is presenting, the listening group may make notes for follow up questions. After the presentation, the two groups should discuss the findings. A spokesperson from each group will summarise their partner groups' findings for the class.

Students should consider the following questions to help direct their discussion:

- What question did these findings answer?

- Where and how could this information be used by others?
- What other data could be interesting to analyse to support these findings?
- What other methods could be used to present these findings?
- Was there any data that did not support these findings?

Hold a class discussion to share students' findings. One member from each group summarises what their partner group had presented and what the data had told them about their chosen countries.

Part 4: Applications of big data sets and collecting data

Using the applications of big data case studies (see *Student resource sheet 1.12: Big data fact sheet 2 – Applications of big data, now and in the future*), students explore ways big data sets could be applied to solve problems.

Students deconstruct each example in their group by considering the following questions:

- Who is this application most appropriate for?
- What is the problem type (environmental, social, economic, global)?
- What is the impact of the problem?
- What big data would need to be collected and analysed?
- Where could this data be sourced from?
- What type of solution is it to the problem (is it a product, initiative, innovation or data solution)?

Students use these understandings to assign each of the examples to the most appropriate United Nations Sustainable Development Goal (see *About the Sustainable Development Goals* in *Digital resources*).

Discuss each groups' findings as a class and record the Sustainable Development Goals that each group assigned each example to allow students to revisit this at a later stage.

The structure of this task will help form the basis for groups to develop their solutions to address one of the United Nations Sustainable Development Goals.

Resource sheets	*Teacher resource sheet 1.12: Big data fact sheet 2 – Applications of big data, now and in the future* *Student resource sheet 1.13: Data sets from around the world*
Digital resources	Finding patterns in data sets (Khan Academy, n.d.) *khanacademy.org/computing/ap-computer-science-principles/data-analysis-101/data-tools/a/finding-patterns-in-data-sets*
	Our World in Data (Global Change Data Lab, n.d.) *ourworldindata.org/*
	The era of big data (Khan Academy, n.d.) *khanacademy.org/computing/ap-computer-science-principles/data-analysis-101/big-data/a/what-is-big-data*
	Spurious correlations (Tylervigen, n.d.) *tylervigen.com/spurious-correlations*
	Correlation and causality (Khan Academy, 2011) *youtube.com/watch?v=ROpbdO-gRUo&feature=youtu.be*
	Gapminder – Correlations, Anomalies, and Trends (ARCGIS, n.d.) *arcgis.com/apps/MapSeries/index.html?appid=6688d48ebab54f2dbb51b830427b8527*
	Gapminder CO2 emissions and child mortality (Gapminder Foundation, n.d.) *gapminder.org/tools/#$state$time$value=2014;&marker$axis_x$which=child_mortality_0_5_year_olds_dying_per_1000_born&domainMin:null&domainMax:null&zoomedMin:null&zoomedMax:null&spaceRef:null;&axis_y$which=co2_emissions_tonnes_per_person&domainMin:null&domainMax:null&zoomedMin:null&zoomedMax:null&scaleType=genericLog&spaceRef:null;;;&chart-type=bubbles*
	Life expectancy and income (Gapminder Foundation, n.d.) *gapminder.org/tools/#$chart-type=bubbles*
	Gapminder (Gapminder Foundation, n.d.) *gapminder.org/*

About the Sustainable Development Goals (United Nations, n.d.)
un.org/sustainabledevelopment/sustainable-development-goals/

Activity 3: Big data solutions

Activity focus 	Students apply understandings about big data as they identify a problem and develop data sets to support their design decisions. Students describe the key features of and develop implementation strategies for their innovation.
Background information	After selecting one of the United Nations Sustainable Development Goals, and an associated target, students identify and analyse relevant data sets to determine the nature and characteristics of their chosen problem. Students use a design process to develop an innovation (which could be an initiative, product, or data solution) that is designed to help meet their chosen target. Further information about the design process can be found in the *Design process guide*. Students will present their ideas using a TED talk type presentation where they must 'pitch' their innovation to their audience. Initiatives around the world are using big data to help solve environmental, social and economic problems. Many of them are linked to and focus on addressing one or more of the 17 United Nations Sustainable Development Goals. Examples of these initiatives include: UN Global Pulse Library (UN Global Pulse, 2020) *www.unglobalpulse.org/library/?_types=report* Data for Climate Action (Data for Climate Action, n.d.) http://dataforclimateaction.org/ Global Partnership for Sustainable Development Data (Global Partnership for Sustainable Development Data, n.d.) *www.data4sdgs.org/* GSMA Launches Big Data for Social Good Initiative – Newsroom (GSM Association, 2020) *www.gsma.com/newsroom/press-release/gsma-launches-big-data-for-social-good/*

Expected learning	Students will be able to: 1. Use a design process and identify data sets and information to support the development of an innovation that addresses a United Nations Sustainable Development Goal target (Mathematics and Technologies). 2. Work collaboratively to collect, analyse, visualise and identify trends in data and to present their innovation idea to an audience (Mathematics, Science and Technologies). 3. Analyse and evaluate data on the needs of contemporary society and provide feedback to peers on an innovation idea (Mathematics, Science and Technologies).
Equipment required	**For the class:** Device for projection and internet connection Access to online articles and video clips in *Digital resources*
	For the students: Digital device with an internet connection Access to a spreadsheet app *Student resource sheet 1.13: Data sets from around the world* *Student resource sheet 1.14: Australian data sets* *Student resource sheet 1.15: United Nations Sustainable Development Goals data sets* *Student activity sheet 3.1: Project plan* *Student activity sheet 3.2: Data innovation one-pager*
Preparation	Ensure students have access to resource sheets and digital resources. Consider appropriate formats for a digital portfolio to allow students to collaboratively track their group's progress.
Activity parts	**Part 1: Deciding on a Sustainable Development Goal target** In groups, students discuss what they learnt about big data and how it can be used to inform the creation of innovations, products and solutions that address real-world problems. Through this process, students select one of the

United Nations Sustainable Development Goals, and an associated target, to which the group will propose a solution.

To aid in this discussion, pose the following questions to guide student thinking:

- Which United Nations Sustainable Development Goal target will your group research?
- Why is this target important to your group?
- Where will you access the most appropriate and relevant data sets about this target?
- What type of information will you analyse?
- How will you generate or collect the data you require to inform your product or initiative?

Introduce the following project description/design brief to the class:

Project description and design brief

After exploring a range of United Nations Sustainability Development Goal related datasets and examples of existing projects, your group will use extensive analysis of relevant data to gain a greater understanding of the target that has been chosen. Your group will develop a detailed plan that outlines the problem that you have identified, and which relates to your chosen United Nations Sustainable Development Goal target.

With this United Nations Sustainable Development Goal target as their focus, students commence interrogating big data sets to help understand the scope of the problem. Students should consider addressing the target at a local, regional, or global level. This will help them choose the most appropriate big data sets to explore.

Remind students that the data set resources will provide sufficient information for them to acquire suitable data sets from (see *Student resource sheet 1.13: Data sets from around the world*, *Student resource sheet 1.14: Australian data sets* and *Student resource sheet 1.15: United Nations Sustainable Development Goals data sets*).

Once students have sourced and started to analyse data related to their chosen goal target, they should define the problem. A well-defined problem will assist when

developing ideas for their innovation and help ensure a suitable solution is proposed.

Part 2: Ideation

Once students have identified the United Nations Sustainable Development Goal target they are going to address, they begin to generate ideas for a solution. This solution is based on the needs identified through the analysis of big data.

Students use the knowledge and skills they have acquired in the previous activities to develop an innovation to help meet their chosen United Nations Sustainable Development Goal target. This innovation that the group proposes could be an initiative, a product or a data solution.

As students develop their innovation, they consider strategies and identify additional data that may be required to further inform their innovation (eg weather data, traffic data, deaths due to malnutrition etc).

Students explain how the data has informed their innovation and justify why they are developing their innovation in the form they have chosen.

To assist with ideation, students should use processes with which they are familiar. Such processes could include brainstorming, mind mapping, sketching, or drawing diagrams and plans. Further information can be found in the *Design process guide*.

Encourage groups to use *Student activity sheet 3.1: Project plan* to organise their information.

Once the groups have agreed on an innovation and supporting data, prompt student thinking by asking:

- What type of innovation is your solution?
- How are you going to promote your innovation?
- How will you know if your innovation has been successful?
- What resources will you use to create your innovation?

Part 3: Production

Depending on the innovation being proposed, each group must identify the stages they will follow in the development of their innovation. Further information on how students can use computer-aided design drawing can be found in the *Drawing in the design process* resource.

Ask each member of the group to complete *Student activity sheet 3.2: Data innovation one-pager* and the group to complete *Student activity sheet 3.1: Project plan* together. To achieve this, student groups need to discuss their innovation (solution) and must consider such things as:

- What supporting data do you have on your United Nations Sustainable Development Goal target?
- What is the anticipated goal or outcome?
- What resources will be required to develop the innovation – physical materials, technologies, building, human resources, data?
- What costs will be incurred? Who will cover these costs?
- What is the proposed timeframe for the development of the innovation prototype/representation?
- How will the innovation be promoted?
- Who will the innovation assist the most?
- How you will know if the innovation is a success?

This information should be considered and included on both *Student activity sheet 3.1: Project plan* and *Student activity sheet 3.2: Data innovation one-pager* which will be used in *Part 4*. Each member of the group will require a copy of both activity sheets and needs to understand what the innovation will be, and how this is supported by the big data the group has analysed.

Part 4: Evaluation

Using the jigsaw strategy (see *Teacher resource sheet 1.2: Cooperative learning – Jigsaw*) one member from each group will join a master group (so that each group has a representative from every other group). Students take turns

	to present their *Student activity sheet 3.2: Data innovation one-pager* to the other members of the master group. Encourage students to ask questions of the presenter after the 'pitch' is given and to provide feedback on the innovation. Once all presentations are completed, group members return to their original groups and, using the placemat strategy (see *Teacher resource sheet 1.3: Cooperative learning – Placemat*), record the feedback they had received about their group's innovation. Students discuss the feedback and consider how their innovation could be refined. New ideas should be recorded on the placemat. Using this feedback, groups refine their *Student activity sheet 3.2: Data innovation one-pager*. Encourage students to consider if their innovation is sufficiently supported by the big data analysis, and make modifications to the *Student activity sheet 3.1: Project plan* as necessary.
Resource sheets	*Design process guide* *Drawing in the design process* *Teacher resource sheet 1.2: Cooperative learning – Jigsaw* *Teacher resource sheet 1.3: Cooperative learning – Placemat* *Student resource sheet 1.13: Data sets from around the world* *Student resource sheet 1.14: Australian data sets* *Student resource sheet 1.15: United Nations Sustainable Development Goals data sets* *Student activity sheet 3.1: Project plan* *Student activity sheet 3.2: Data innovation one-pager*

Activity 4: Showcase

Activity focus 	Groups present their innovation at a showcase event, highlighting the importance of big data analysis in informing their innovation.
Background information	Students prepare and deliver a presentation detailing how they used big data to inform an innovative solution to a world problem, as identified in the *United Nations Sustainable Development Goals*. *TED: Ideas worth spreading* (TED Conferences LLC, n.d.) www.ted.com/ is a non-profit organisation devoted to spreading ideas, usually in the form of short, powerful talks (18 minutes or less) from innovators and leaders in fields such as science, medicine, business and global issues. This platform has become a key mechanism for stimulating new thinking, ideas and innovations around topics that are impacting us globally. *TED-Ed* (TED Conferences LLC, n.d.) https://ed.ted.com/, based on the same premise, has evolved as a platform for school-aged children (8-18 years old) to give 'voice' to their innovative and creative ideas. The TED-Ed Student Talks program encourages students to develop 'TED-like' talks around their chosen topic and provides a scaffolded structure to help them develop their presentation. These presentations can then be live and/or recorded and uploaded to the school website or school *YouTube* channel. Teachers can register their class to actively participate in the program or use the materials in a more localised way such as presenting student 'TED talks' at a school assembly or parent night. If the decision is made to fully participate in the program, then some student presentations could be submitted to TED-Ed to be included in their online showcase of student talks. Please note, registration is free and necessary to access all materials. If an alternative presentation format is desired, the *TED-Ed Lessons* (TED Conferences LLC, n.d.) https://ed.ted.com/lessons?direction=desc&sort=featured-

	position could be used to stimulate ideas for presentations. Students can choose from a range of multimedia tools including *PowerPoint, Sway, Keynote* or *iMovie*. *Teacher resource sheet 4.1: Student evaluation* has been included to provide scaffolding to assist with the possible assessment of the module, as well as provide scaffolding to assist students in their reflections. During the presentation process, it is suggested that students complete *Student activity sheet 4.2: Peer evaluation* and/or *Student activity sheet 4.3: Self-evaluation*.
Instructional procedures	Students maintain the same groups. To scaffold cooperative group work, each member of the group could have a role and responsibility. See *Teacher resource sheet 1.1: Cooperative learning – Roles*. When preparing for the showcase, all students contribute to deciding on content, preparing the media presentation and delivering the presentation; although one student may have overall responsibility for managing the project. When students are preparing the presentation script, timing and public speaking skills need to be discussed as well as content for the slides (ie slides should not be text-heavy). Direct students to resources that scaffold the process for creating a TED-like talk (see *Digital resources*). After the student presentations, groups provide peer group feedback. Two stars and a wish, plus, minus and what's next?, and warm and cool feedback are *Strategies to enhance peer feedback* (Education Services Australia, n.d.) *www.assessmentforlearning.edu.au/professional_learning/peer_feedback/peer_strategies_enhance.html#1* that could be used. Modelling of this strategy will be needed.
Expected learning	Students will be able to: 1. Explain and evaluate how big data is used to solve real-world problems as identified by the United Nations Sustainable Development Goals (Mathematics, Science and Technologies). 2. Use digital technology to organise and present information about their innovation and the design process (Mathematics and Technologies). 3. Work collaboratively to design, develop and present a TED type talk to communicate the importance of big

	data analysis in informing their innovation (Mathematics and Technologies).
Equipment required	**For the students**: A digital device with an internet connection Access to a spreadsheet app *Student activity sheet 4.2: Peer evaluation* *Student activity sheet 4.3: Self-evaluation* *Student activity sheet 1.0: Journal checklist*
Activity parts	**Part 1: Showcase preparation** Explain to students they are going to create a presentation along the lines of a TED Talk to promote their innovation, which is designed to address a real-world problem as identified through the analysis of big data sets around the United Nations Sustainable Development Goals. To give students an idea of the format of a TED Talk, remind them of TED talks they have watched as part of this module. Also show the *Start a TED-Ed Club today* video where students present their ideas using the TED Talk format (see *Digital resources*). Ask students to consider the following questions as they prepare the structure of their presentation: • What makes for an engaging TED Talk? • How can you best promote your innovation? • What key messages are to be communicated, and who is the intended audience? • How has big data analysis been used to help ideate your solution to one of the United Nations Sustainable Development Goal targets? • How are you going to structure your presentation and distribute the tasks within your group? Discuss with the class how they will be able to evaluate whether or not their TED Talk presentations are successful, and how they will highlight the importance of big data analysis in informing their innovation.

Part 2: Showcase presentation

Students create the script, identify the visual assets they will use, and select the 'voice' of the group to perform the talk.

Students work in groups to prepare their presentations. Provide a combination of class and homework time for development, review and refining of the presentations.

Students will need help developing the skills needed for promoting their innovation/ product to a wider audience.

Part 3: Presentation delivery

The student representative for each group delivers their TED-like talk to the audience highlighting how data helped inform their choice of problem and strategy for finding a solution. There is an expectation that after the TED-talk the audience can ask probing or clarifying questions to find out more about the innovation. All members of the group are expected to be able to answer questions such as:

- What types of data did you analyse to help you to understand the problem?
- What data sets did you combine to identify patterns and relationships to inform the creation of your innovation?
- How will you know if your innovation is successful?

Part 4: Peer group feedback

After the presentation, time should be allocated for students to give feedback. Groups take turns to provide written feedback to all the other groups. Questions that could be used to help structure feedback could include:

- How has big data analysis been used to help ideate the solution to the United Nations Sustainable Development Goal target?
- What improvements to the data analysis could you suggest to better support the innovation?
- How could the presentation be improved?

Part 5: Reflection and journaling

The teacher may use this opportunity to complete the *Teacher resource sheet 4.1: Student evaluation*.

The teacher may hold individual student discussions, or have students use their journal to give and receive feedback using *Teacher resource sheet 4.3: 3 – 2 – 1 Reflection*.

Debrief the module with the class and provide time for students to complete reflections and document their thoughts on the following questions in their journal:

- What are the most important things you learnt about the use of big data analysis to help identify patterns in data and using these to ideate solutions to problems?
- How well did the data support you in ideating your innovation to the United Nations Sustainable Development Goal target?
- What additional data would help support your solution to the target goal? Why?
- How might the STEM skills you have developed in this module better prepare you for a successful future?
- What are three things I learnt? What are two things I found interesting? What is one thing I found difficult?

Allow time for students to individually consider peer group feedback. Encourage students to incorporate insights from this peer group feedback and to reflect on their learning journey in their journals or digital portfolio, specifically for the use of big data analysis and on their presentation.

Students should complete and ensure that all relevant activities on *Student activity sheet 1.0: Journal checklist* are included in their journal. Advise students if they need to submit their journal for feedback and assessment.

Students complete *Student activity sheet 4.2: Peer evaluation* and *Student activity sheet 4.3: Self-evaluation*.

Resource sheets

Student activity sheet 1.0: Journal checklist

Teacher resource sheet 1.1: Cooperative learning – Roles

Teacher resource sheet 4.1: Student evaluation

Student activity sheet 4.2: Peer evaluation

	Student activity sheet 4.3: Self-evaluation *Teacher resource sheet 4.3: 3 – 2 – 1 Reflection*
Digital resources	Bring TED-Ed Student Talks to Your School (TED Conferences LLC, n.d.) *ed.ted.com/student_talks*
	TED: Ideas worth spreading (TED Conferences LLC, n.d.) *ted.com/*
	TED-Ed (TED Conferences LLC, n.d.) *ed.ted.com/*
	TED-Ed Lessons (TED Conferences LLC, n.d.) *ed.ted.com/lessons?direction=desc&sort=featured-position*
	Start a TED-Ed Club today (TED-Ed, 2014) *youtube.com/watch?v=Qj_OzmP8iWg&feature=emb_logo*
	What students taught TED-Ed this year (TED-Ed Student Talks, 2018) *youtube.com/watch?v=ly7PSsgh_QM*
	TED-Ed Student Talks Information Packet (TED-Ed, 2019) *teded.tedcdn.com/club-resources/ted-ed-club-info-packet-2019-02.pdf*
	Kids coached to pitch world-changing ideas (Phys.org, 2014) *phys.org/news/2014-01-kids-pitch-world-changing-ideas.html*
	How to Create Your TED Talk: An 8-Step Process (Speak Up for Success, 2017) *speakupforsuccess.com/create-a-ted-talk/*

Appendix 1: Links to the Western Australian Curriculum:

The *Big data* module provides opportunities for developing students' knowledge and understandings in science, technologies and mathematics. The tables below show how this module aligns to the content of the Western Australian Curriculum and can be used by teachers for planning and monitoring.

BIG DATA Links to the Western Australian Curriculum	ACTIVITY			
	1	**2**	**3**	**4**
SCIENCE				
SCIENCE UNDERSTANDING				
Biological sciences: Ecosystems consist of communities of interdependent organisms and abiotic components of the environment; matter and energy flow through these systems (ACSSU176)	●			
SCIENCE AS A HUMAN ENDEAVOUR				
Nature and development of science: Scientific understanding, including models and theories, is contestable and is refined over time through a process of review by the scientific community (ACSHE157)	●	●		
Nature and development of science: Values and needs of contemporary society can influence the focus of scientific research (ACSHE228)		●	●	●
SCIENCE INQUIRY SKILLS				
Questioning and predicting: Formulate questions or hypotheses that can be investigated scientifically (ACSIS164)		●	●	
Planning and conducting: Plan, select and use appropriate investigation types, including field work and laboratory experimentation, to collect reliable data; assess risk and address ethical issues associated with these methods (ACSIS165)	●	●	●	
Processing and analysing data and information: Analyse patterns and trends in data, including describing relationships between variables and identifying inconsistencies (ACSIS169)	●	●	●	●
Processing and analysing data and information: Use knowledge of scientific concepts to draw conclusions that are consistent with evidence (ACSIS170)	●	●	●	

BIG DATA Links to the Western Australian Curriculum	ACTIVITY			
	1	2	3	4
SCIENCE				
SCIENCE INQUIRY SKILLS				
Evaluating: Evaluate conclusions, including identifying sources of uncertainty and possible alternative explanations, and describe specific ways to improve the quality of the data (ACSIS171)	●	●		
Evaluating: Critically analyse the validity of information in primary and secondary sources and evaluate the approaches used to solve problems (ACSIS172)	●	●	●	
Communicating: Communicate scientific ideas and information for a particular purpose, including constructing evidence-based arguments and using appropriate scientific language, conventions and representations (ACSIS174)	●	●		●
DIGITAL TECHNOLOGIES				
PROCESSES AND PRODUCTION SKILLS				
Collecting, managing and analysing data: Explore techniques for acquiring, storing and validating quantitative and qualitative data (ACTDIP036)	●	●		
Collecting, managing and analysing data: Analyse and visualise data to create information and address complex problems (ACTDIP037)	●	●	●	●
Digital implementation: Create and use interactive solutions for sharing ideas and information online, taking into account social contexts (ACTDIP043)		●	●	
CREATING SOLUTIONS BY				
Investigating and defining: Identify and define the needs of a stakeholder, to create a brief, for a solution (WATPPS54)			●	●
Investigating and defining: Investigate a selection of components/resources to develop solution ideas, identifying and considering constraints (WATPPS55)			●	●
Designing: Apply design thinking, creativity and enterprise skills (WATPPS56)			●	●

BIG DATA Links to the Western Australian Curriculum	ACTIVITY			
	1	2	3	4
DIGITAL TECHNOLOGIES				
CREATING SOLUTIONS BY				
Designing: Design solutions assessing alternative designs against given criteria, using appropriate technical terms and technology (WATPPS57)			●	●
Evaluating: Evaluate design processes and solutions against student developed criteria (WATPPS59)				●
Collaborating and managing: Work independently, and collaboratively to manage projects, using digital technology and an iterative and collaborative approach. Considers time, cost, risk and safety (WATPPS60)			●	●
MATHEMATICS				
STATISTICS AND PROBABILITY				
Chance: Investigate reports of surveys in digital media and elsewhere for information on how data were obtained to estimate population means and medians (ACMSP227)	●			
Data representation and interpretation: Identify everyday questions and issues involving at least one numerical and at least one categorical variable, and collect data directly and from secondary sources (ACMSP228)	●	●	●	
Data representation and interpretation: Compare data displays using mean, median and range to describe and interpret numerical data sets in terms of location (centre) and spread (ACMSP283)	●			

Further information about assessment and reporting in the Western Australian Curriculum can be found at *k10outline.scsa.wa.edu.au/home*

Appendix 1B: Mathematics proficiency strands

Key ideas

In Mathematics, the key ideas are the proficiency strands of understanding, fluency, problem-solving and reasoning. The proficiency strands describe the actions in which students can engage when learning and using the content. While not all proficiency strands apply to every content description, they indicate the breadth of mathematical actions that teachers can emphasise.

Understanding

Students build a robust knowledge of adaptable and transferable mathematical concepts. They make connections between related concepts and progressively apply the familiar to develop new ideas. They develop an understanding of the relationship between the 'why' and the 'how' of mathematics. Students build understanding when they connect related ideas, when they represent concepts in different ways, when they identify commonalities and differences between aspects of content, when they describe their thinking mathematically and when they interpret mathematical information.

Fluency

Students develop skills in choosing appropriate procedures; carrying out procedures flexibly, accurately, efficiently and appropriately; and recalling factual knowledge and concepts readily. Students are fluent when they calculate answers efficiently, when they recognise robust ways of answering questions, when they choose appropriate methods and approximations, when they recall definitions and regularly use facts, and when they can manipulate expressions and equations to find solutions.

Problem-solving

Students develop the ability to make choices, interpret, formulate, model and investigate problem situations, and communicate solutions effectively. Students formulate and solve problems when they use mathematics to represent unfamiliar or meaningful situations, when they design investigations and plan their approaches, when they apply their existing strategies to seek solutions, and when they verify that their answers are reasonable.

Reasoning

Students develop an increasingly sophisticated capacity for logical thought and actions, such as analysing, proving, evaluating, explaining, inferring, justifying and generalising. Students are reasoning mathematically when they explain their thinking, when they deduce and justify strategies used and conclusions reached, when they adapt the known to the unknown, when they transfer learning from one context to another, when they prove that something is true or false, and when they compare and contrast related ideas and explain their choices.

Source: ACARA – *www.australiancurriculum.edu.au/f-10-curriculum/mathematics/key-ideas/?searchTerm=key+ideas#dimension-content*

Appendix 2: General capabilities continuums

The general capabilities continuums shown here are designed to enable teachers to understand the progression students should make with reference to each of the elements. There is no intention for them to be used for assessment.

Information and communication technology (ICT) capability learning continuum

Sub-element	Typically by the end of Year 6	Typically by the end of Year 8	Typically by the end of Year 10
Create with ICT **Generate ideas, plans and processes**	use ICT effectively to record ideas, represent thinking and plan solutions	use appropriate ICT to collaboratively generate ideas and develop plans	select and use ICT to articulate ideas and concepts, and plan the development of complex solutions
Create with ICT **Generate solutions to challenges and learning area tasks**	independently or collaboratively create and modify digital solutions, creative outputs or data representation/transformation for particular audiences and purposes	design and modify simple digital solutions, or multimodal creative outputs or data transformations for particular audiences and purposes following recognised conventions	design, modify and manage complex digital solutions, or multimodal creative outputs or data transformations for a range of audiences and purposes
Communicating with ICT **Collaborate, share and exchange**	select and use appropriate ICT tools safely to share and exchange information and to safely collaborate with others	select and use appropriate ICT tools safely to lead groups in sharing and exchanging information, and taking part in online projects or active collaborations with appropriate global audiences	select and use a range of ICT tools efficiently and safely to share and exchange information, and to collaboratively and purposefully construct knowledge

Critical and creative thinking learning continuum

Sub-element	Typically by the end of Year 6	Typically by the end of Year 8	Typically by the end of Year 10
Inquiring – identifying, exploring and organising information and ideas **Organise and process information**	analyse, condense and combine relevant information from multiple sources	critically analyse information and evidence according to criteria such as validity and relevance	critically analyse independently sourced information to determine bias and reliability
Generating ideas, possibilities and actions **Imagine possibilities and connect ideas**	combine ideas in a variety of ways and from a range of sources to create new possibilities	draw parallels between known and new ideas to create new ways of achieving goals	create and connect complex ideas using imagery, analogies and symbolism
Generating ideas, possibilities and actions **Seek solutions and put ideas into action**	assess and test options to identify the most effective solution and to put ideas into action	predict possibilities, and identify and test consequences when seeking solutions and putting ideas into action	assess risks and explain contingencies, taking account of a range of perspectives, when seeking solutions and putting complex ideas into action
Reflecting on thinking and processes **Transfer knowledge into new contexts**	apply knowledge gained from one context to another unrelated context and identify new meaning	justify reasons for decisions when transferring information to similar and different contexts	identify, plan and justify the transfer of knowledge to new contexts

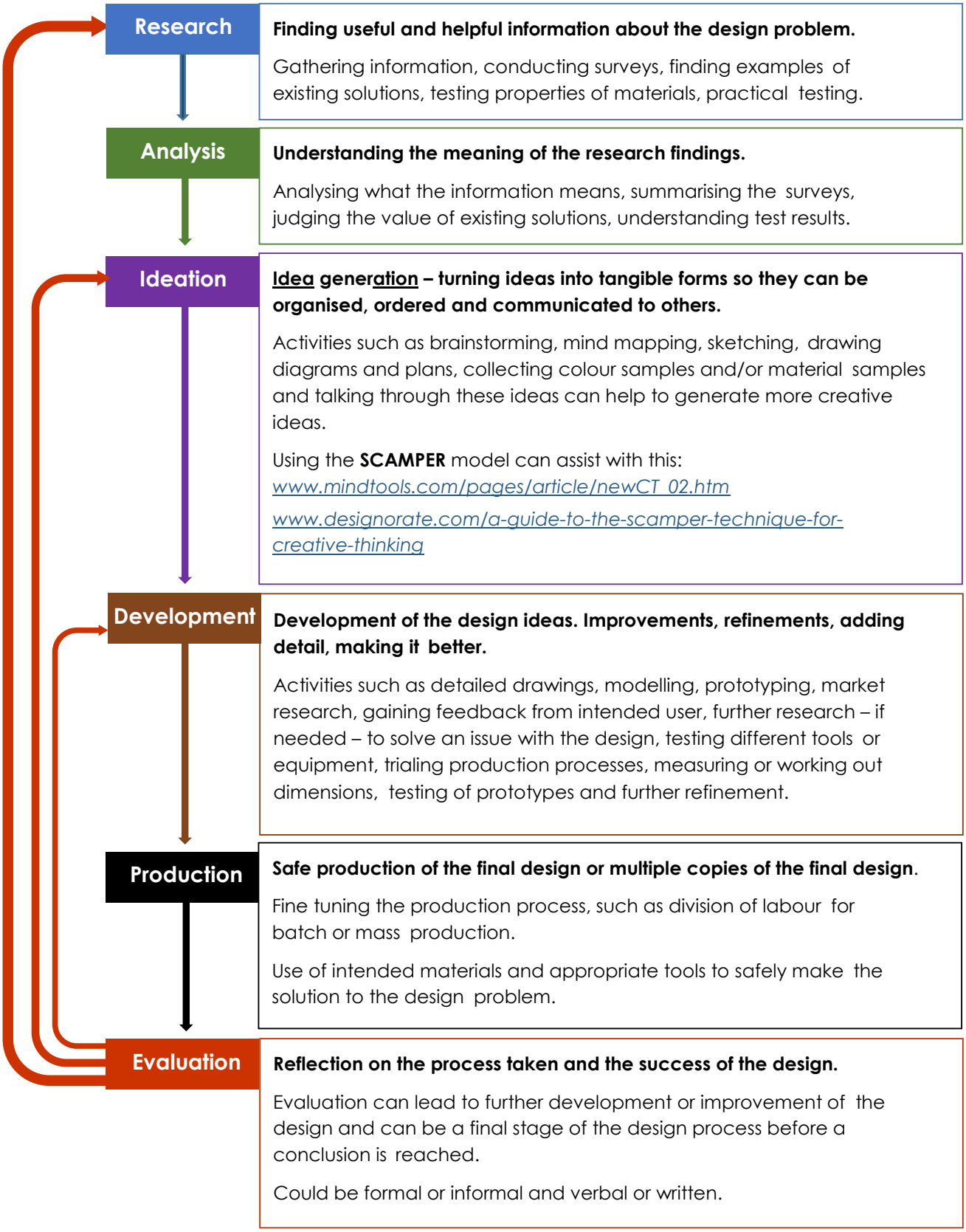
Personal and social capability learning continuum

Sub-element	Typically by the end of Year 6	Typically by the end of Year 8	Typically by the end of Year 10
Social management **Work collaboratively**	contribute to groups and teams, suggesting improvements in methods used for group investigations and projects	assess the extent to which individual roles and responsibilities enhance group cohesion and the achievement of personal and group objectives	critique their ability to devise and enact strategies for working in diverse teams, drawing on the skills and contributions of team members to complete complex tasks
Social management **Negotiate and resolve conflict**	identify causes and effects of conflict, and practise different strategies to diffuse or resolve conflict situations	assess the appropriateness of various conflict resolution strategies in a range of social and work-related situations	generate, apply and evaluate strategies such as active listening, mediation and negotiation to prevent and resolve interpersonal problems and conflicts
Social management **Develop leadership skills**	initiate or help to organise group activities that address a common need	plan school and community projects, applying effective problem-solving and team-building strategies, and making the most of available resources to achieve goals	propose, implement and monitor strategies to address needs prioritised at local, national, regional and global levels, and communicate these widely

Further information about general capabilities is available at:

k10outline.scsa.wa.edu.au/home/p-10-curriculum/general-capabilities-over/general-capabilities-overview/general-capabilities-in-the-australian-curriculum

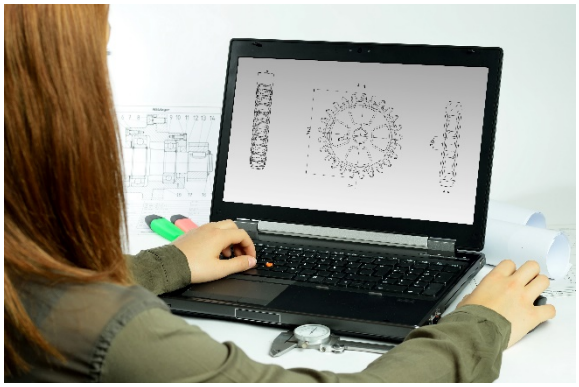
Appendix 3: Design process guide



Appendix 3B: Drawing in the design process

Incorporating the design process into the STEM modules will often result in the need for students to draw plans of their designs. This can be done at a simple level using hand-drawn sketches or at a more technical level using computer-aided design (CAD).

By developing skills using industry-standard software, students may be well-placed to explore future career pathways.



istockphoto.com.au

There are several CAD software options, two free examples are detailed below. *Autodesk* is a third package that is also free for educational use.

Tinkercad

- Format: Web-based app requiring internet access via a browser
- Purpose: A simple, online 3D design and 3D printing app
- Home: *www.tinkercad.com*
- Blog: *blog.tinkercad.com*
- Tutorials: *www.tinkercad.com/learn*
- Feature: Connects to 3D printing and laser cutting.

SketchUp

- Format: Can be downloaded and installed on devices, or used in a browser
- Purpose: Enables students to draw in 3D
- Home: *www.sketchup.com* 'Products' '*SketchUp for Schools*'
- Help centre: *help.sketchup.com/en*
- Blog: *blog.sketchup.com*
- Tutorials: *www.youtube.com/user/SketchUpVideo*. From beginner tool tips to intermediate and advanced modelling techniques, the video tutorials help to build *SketchUp* skills.

Appendix 4: Student journal

istockphoto.com.au

When students reflect on learning and analyse their ideas and feelings, they self-evaluate, thereby improving their metacognitive skills.

These modules encourage students to self-reflect and record the stages of their learning in a journal. This journal may take the form of a written journal, a portfolio or a digital portfolio.

Using digital portfolios can help develop students' Information and Communication Technology (ICT) capability.

Reflective practice and recording can be supported in classrooms by creating opportunities for students to think about and record their learning through notes, drawings or pictures. Teachers should encourage students to revisit earlier journal entries to help them observe the progress of their thoughts and understanding.

Journals are a useful tool that gives teachers additional insight into how students value their own learning and progress, as well as demonstrating their individual achievements.

The following links provide background information and useful apps for journaling.

Reflective journal (University of Technology Sydney)
www.uts.edu.au/sites/default/files/reflective_journal.pdf

Reflective writing (University of New South Wales Sydney)
student.unsw.edu.au/reflective-writing

Balancing the two faces of ePortfolios (Helen Barrett, 2009)
electronicportfolios.org/balance/Balancing.jpg

Digital portfolios for students (Cool tools for school)
cooltoolsforschool.wordpress.com/digital-student-portfolios

Kidblog – digital portfolios and blogging
kidblog.org/home

Evernote (a digital portfolio app)
evernote.com

Weebly for education (a drag and drop website builder)
education.weebly.com

Connect – the Department of Education's integrated, online environment
connect.det.wa.edu.au

Appendix 5: Student activity sheet 1.0: Journal checklist

As an ongoing part of this module, you have been keeping a journal of your work.

Before submitting your journal to your teacher please ensure you have included the following information.

- Tick each box once complete and included.
- Write N/A for items that were not required in this module.

istockphoto.com.au

Your name and group member's names or photographs	
An explanation of the problem you are solving	
Your notes from *Activity 1*	
Your notes from *Activity 2*	
Your notes from *Activity 3*	
Your notes from *Activity 4*	
Student activity sheet 1.6: Annual rainfall comparisons	
Student activity sheet 1.8: Kangaroo populations and rainfall comparisons	
Student activity sheet 1.10: Gapminder data 'deep dive'	
Student activity sheet 3.1: Project plan	
Student activity sheet 3.2: Data innovation one-pager	
Student activity sheet 4.2: Peer evaluation	
Student activity sheet 4.3: Self-evaluation	

Student activity sheet 1.0: Journal checklist	

Appendix 6: Teacher resource sheet 1.1: Cooperative learning – Roles

Cooperative learning frameworks create opportunities for groups of students to work together, generally to a single purpose.

As well as having the potential to increase learning for all students involved, using these frameworks can help students develop personal and social capability.

When students are working in groups, positive interdependence can be fostered by assigning roles to group members.

istockphoto.com.au

These roles could include:

- Working roles such as Reader, Writer, Summariser, Time-keeper.
- Social roles such as Encourager, Observer, Noise monitor, Energiser.

Further to this, specific roles can be delineated for specific activities that the group is completing. It can help students if some background to the purpose of group roles is made clear to them before they start, but at no time should the roles get in the way of the learning. Teachers should decide when or where roles are appropriate to given tasks.

istockphoto.com.au

Appendix 7: Teacher resource sheet 1.2: Cooperative learning – Jigsaw

Cooperative learning frameworks create opportunities for groups of students to work together, generally for a single purpose.

As well as having the potential to increase learning for all students involved, using these frameworks can help students develop personal and social capability.

The jigsaw strategy typically has each member of the group becoming an 'expert' on one or two aspects of a topic or question being investigated. Students start in their cooperative groups, then break away to form 'expert' groups to investigate and learn about a specific aspect of a topic. After developing a sound level of understanding, the students return to their cooperative groups and teach each other what they have learnt.

Within each expert group, issues such as how to teach the information to their group members are considered.

<table>
<tr><td>Step 1</td><td>Cooperative groups
(of four students)</td><td colspan="2">1 2 3 4</td><td colspan="2">1 2 3 4</td></tr>
<tr><td>Step 2</td><td>Expert groups
(size equal to the number of groups)</td><td>1 1</td><td>2 2</td><td>3 3</td><td>4 4</td></tr>
<tr><td>Step 3</td><td>Cooperative groups
(of four students)</td><td colspan="2">1 2 3 4</td><td colspan="2">1 2 3 4</td></tr>
</table>

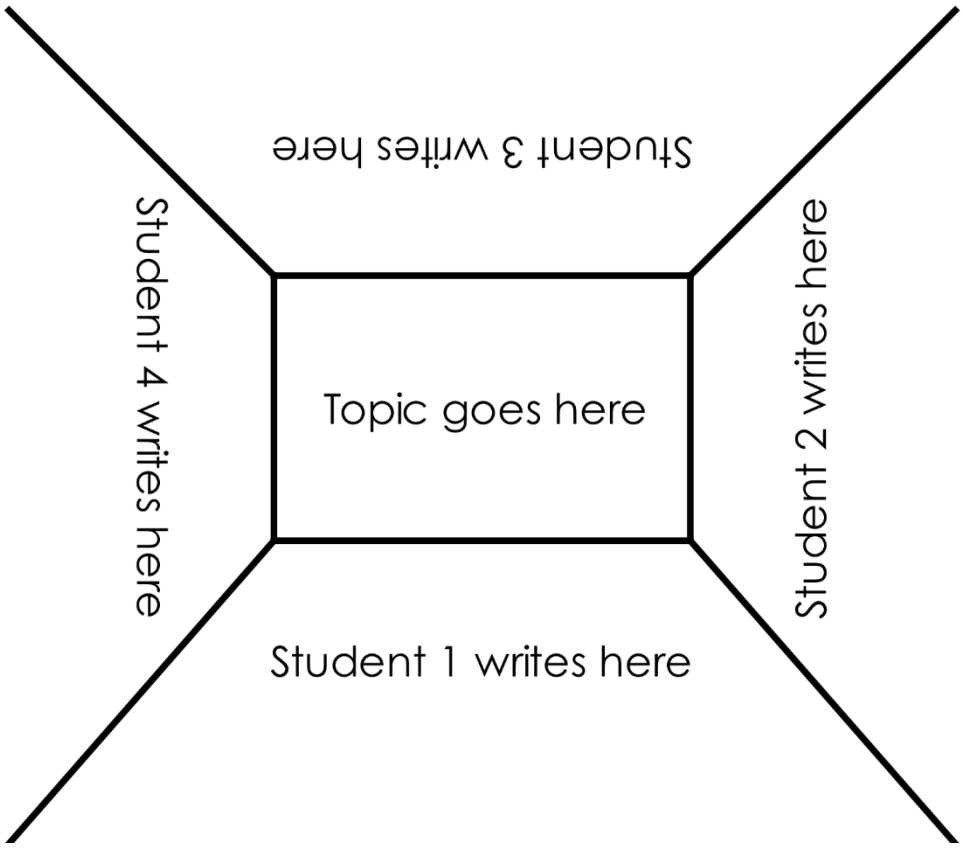
Appendix 8: Teacher resource sheet 1.3: Cooperative learning – Placemat

istockphoto.com.au

Cooperative learning frameworks create opportunities for groups of students to work together, generally for a single purpose.

As well as having the potential to increase learning for all students involved, using these frameworks can help students develop personal and social capability.

The placemat strategy involves students working collaboratively to record prior knowledge about a common topic and brainstorm ideas. It also allows teachers to readily see the contribution of each student. The diagram below shows a typical placemat template.



STEM Consortium

Appendix 9: Teacher resource sheet 1.4: Cooperative learning – Think, Pair, Share

Cooperative learning frameworks create opportunities for groups of students to work together, generally to a single purpose.

istockphoto.com.au

As well as having the potential to increase learning for all students involved, using these frameworks can help students develop personal and social capability.

In the 'think' stage, each student thinks silently about a question asked by the teacher.

In the 'pair' stage, students discuss their thoughts and answers to the question in pairs.

In the 'share' stage, the students share their answer, their partner's answer or what they decided together. This sharing may be with other pairs or with the whole class. It is important also to let students 'pass'. This is a key element of making the strategy safe for students.

Think–pair–share increases student participation and provides an environment for higher levels of thinking and questioning.

istockphoto.com.au

Appendix 10: Teacher resource sheet 1.5: Vocabulary

Below is a list of vocabulary appropriate for the teaching of *Big data*.

Technological vocabulary	**Definition**
Behavioural data	Behavioural data refers to information produced as a result of actions, typically commercial behaviour using a range of devices connected to the Internet, such as a PC, tablet, or smartphone. Behavioural data tracks the sites visited, the apps downloaded, or the games played.
Big data	Big Data is a term for large, complex and layered groups of data that can be analysed to spot patterns and trends. Using data from multiple sources we can gain deeper and more accurate information that we can use to our benefit. Big Data analysis requires multiple computers to work together as the amount of data collected is so large.
Comma-separated values	A common way to organize, store and read tabular data (numbers, text) in which each data record (or field) is separated by commas. Abbreviated as CSV.
Data	Data is raw, unorganized facts that need to be processed. Data can be something simple and seemingly random and useless until it is organsied. When data is processed, organised, structured or presented in a given context, so as to make it useful, it is called information.
Data analytics	A process of examining data to discover patterns, correlations and trends in order to find useful information, prompt conclusions, and enhance and fine tune decision making
Data cleaning	Removing non-relevant data from datasets. Could include duplicate results, unwanted data and extraneous, superfluous information.
Dirty data	Dirty data, also known as rogue data, is inaccurate, incomplete or inconsistent *data*, especially in a computer system or *database*. Dirty data can contain such mistakes as spelling or punctuation errors, incorrect data associated with a field, incomplete or outdated data, or even data that has been duplicated in the database. They can be cleaned through a process known as *data cleaning*.

Exponential growth	When a growth rate becomes even more rapid in proportion to the growing total number or size.
Hacker	Someone who uses computers and the Internet to access computers and servers without permission.
Information	Data formatted in an organised way which can be used for processing.
Mean	The **statistical mean** refers to the **mean** or average that is used to derive the central tendency of the data in question. It is determined by adding all the data points in a population and then dividing the total by the number of points. The resulting number is known as the **mean** or the average.
Median	The **median** is a simple measure of central tendency. To find the **median**, we arrange the observations in order from smallest to largest value. If there is an odd number of observations, the **median** is the middle value. If there is an even number of observations, the **median** is the average of the two middle values.
Moore's Law	The principle that the speed and capability of computers can be expected to double every two years, as a result of increases in the number of transistors a microchip can contain.
Qualitative data	Qualitative data can be observed and recorded and is non-numeric in nature. This type of data is collected through methods of observations, one-to-one interviews, focus groups, surveys etc.
Quantitative data	Quantitative data is data expressing a certain quantity, amount or range (how many, how much, how often etc.)
Trendlines	Are best fit linear regression lines.
Variability	Inconsistency of the data set can hamper processes to handle and manage it.
Variety	The type and nature of the data. This helps people who analyze it to effectively use the resulting insight. Big data draws from text, images, audio, video; plus it completes missing pieces through *data fusion*.
Velocity	The speed at which the data is generated and processed to meet the demands and challenges that lie in the path of growth and development. Big data is often available in real-time. Compared to *small data*, big data are produced more continually. Two kinds of velocity related to big data are the

	frequency of generation and the frequency of handling, recording, and publishing.
Veracity	The quality of captured data can vary greatly, affecting accurate analysis.
Volume	The quantity of generated and stored data. The size of the data determines the value and potential insight, and whether it can be considered big data or not.

Appendix 11: Student activity sheet 1.6: Annual rainfall comparisons

Introduction

The Bureau of Meteorology collects, analyses and stores weather and climate data. The amount of rainfall is one type of data that is collected at weather stations; with weather stations located throughout Australia allowing the amount of rainfall in individual locations and regions to be compared. Students collect secondary data from the Bureau of Meteorology to develop spreadsheets to analyse and compare average annual rainfall data for regions. These regions are commonly referred to as rainfall districts and are used to group multiple weather station sites that have relatively similar rainfall climates. Within Western Australia, three weather stations within three rainfall districts will be compared in this activity.

The problem that is being explored is: *Which of the three rainfall districts has the highest annual rainfall?*

Task

In this activity, students collect and analyse secondary data from the Bureau of Meteorology using a spreadsheet software program. By calculating the mean, median and range of rainfalls for three weather stations within each of three rainfall districts, annual rainfalls can be compared. Rainfall from the years 2015-2019 will be used.

Summary of steps:

Monthly rainfall data from nine weather stations will be analysed. A summary of the steps to be undertaken includes:

- collecting monthly weather station rainfall data from the Bureau of Meteorology website
- calculating rainfall mean, median and range for the years 2015-2019
- comparing the rainfall for three rainfall districts.

Instructions for accessing the rainfall data

1. Collect rainfall data by going to *www.bom.gov.au/climate/data/*.
2. In box 1 select *Data about* **Rainfall** from the drop-down list and select the *Type of data* to be **Monthly** Observations. **Monthly rainfall** data will now be *Selected*.
3. In box 3, *Get the data* by entering the *Station number* from the following table. Selecting *Get Data* will then open the *Monthly rainfall data* in a new window. Repeat for all listed weather stations.

Instructions for creating the spreadsheet

4. Open a new spreadsheet.

5. For the Central Coast District, select the Wanneroo (9105) weather station data window, highlight and copy the monthly rainfall data for the years 2015-2019, then paste into your spreadsheet.
6. Repeat for the Swanbourne (9215) and Perth Airport (9021) weather stations within this Central Coast District. Copy all data from the same district into the same sheet.
7. Repeat Steps 5 and 6 for the South Coast and West Gascoyne District weather stations.

Central Coast District (Perth Metropolitan)		South Coast District		West Gascoyne District	
Wanneroo	9105	Albany	9500	Edaggee	6020
Swanbourne	9215	Myrup	9584	Steep Point	6102
Perth Airport	9021	Northcliffe	9590	Shark Bay Airport	6105

Individual weather stations data

8. For each weather station, sum the monthly data for each year to get the *Annual Rainfall*. Confirm these values with those on the Bureau of Meteorology Monthly rainfall data pages.
9. Calculate the *5-year Mean Annual Rainfall* and *5-year Median Annual Rainfall* values for each weather station by calculating the mean and median values of the *Annual Rainfall* values over the 5 years (2015-2019). Compare these values to the Summary statistics for all year's data. Comment on your findings.
10. For each weather station, calculate the *Mean Rainfall* and *Median Rainfall* for each month over the 5 years (2015-2019) by using the individual monthly data for each month. Compare to the *Summary statistics for all year's* data.

District data

11. Calculate the *District Monthly Mean Rainfall* for each of the individual years 2015-2019 by finding the mean rainfall for each month for the three weather stations within the district. Repeat for all three districts.
12. Calculate the *5-year District Monthly Mean Rainfall* values by finding the mean of the *District Monthly Mean Rainfall* values for the years 2015-2019. Repeat for all three districts.
13. Calculate the *District Monthly Mean Annual Rainfall* for each year by summing the *District Monthly Mean Rainfall values*.
14. Calculate the *5-year District Mean Annual Rainfall* by finding the mean of the *District Monthly Mean Annual Rainfall* values for the years 2015-2019. Repeat this to calculate the *5-year District Median Annual Rainfall* and *5-year District Annual Rainfall Range* values. Repeat for all three districts.
15. Once all values have been calculated, create a new sheet that contains only District data. Copy and paste the *5-year Monthly Mean Rainfall* data, the *District Monthly Mean Annual Rainfall* data for each of the years 2015-2019,

and the *5-year District Annual Rainfall Mean, Median* and *Range* data for each of the three districts. Display the District data in tables for comparison. Sample tables are as follows:

5y Monthly Mean Rainfall Data	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
Central Coast													
South Coast													
West Gascoyne													

	District Monthly Mean Annual Rainfall					5y District Annual Rainfall		
	2015	2016	2017	2018	2019	Mean	Median	Range
Central Coast								
South Coast								
West Gascoyne								

Analysing your data

Represent the information in the above tables in graphs.

. Use these graphs and tables above look for trends.

Questions to answer

1. By comparing the weather stations within a district, what information might the rainfall data provide about the individual weather station locations?

_____

_____

_____

_____

_____

2. Which year had the highest annual rainfall? Does each weather station within a district show the same trend? Comment.

_____

_____

_____

_____

_____

3. Which district has the highest annual rainfall? What evidence supports this?

4. By comparing the 5-year individual weather station data, was the rainfall between 2015-2019 higher or lower than the summary statistics for all year's data? Is this data sufficient evidence to indicate if the annual rainfall is increasing or decreasing? Comment.

5. The Bureau of Meteorology sometimes combine smaller rainfall district segments into a single rainfall district. Would it be suitable to have the Central Coast and South Coast districts be combined into one larger rainfall district? What additional information or data might you need before making this decision?

6. If the Central Coast, South Coast and West Gascoyne rainfall districts were combined into a much larger district, what impact might this have on the mean annual rainfall data? Would this still be a suitable representation of the rainfall data across the larger district? Use evidence from your data analysis to support your answer.

Appendix 12: Teacher resource sheet 1.7: Exemplar spreadsheet for annual rainfall comparisons

Introduction

The following sample spreadsheet is intended for use by the teacher only to provide an example of what might be generated by students in Activity 1 using *Student activity sheet 1.6: Annual rainfall comparisons*.

Exemplar spreadsheet

App 12: Teacher exemplar rainfall

Explanation of the spreadsheet

The exemplar spreadsheet contains the following:

- weather station data sheets for the Central Coast, South Coast and West Gascoyne rainfall districts
- sample calculation sheets for the Central Coast, South Coast and West Gascoyne rainfall districts
- blank District calculation sheets (locked and unlocked)
- district summary sheet.

Contained within these sheets are highlighted cells. Yellow highlighting indicates values that can be compared across weather stations and Bureau of Meteorology summary statistics data. Green highlighting indicates values that can be compared across the rainfall districts.

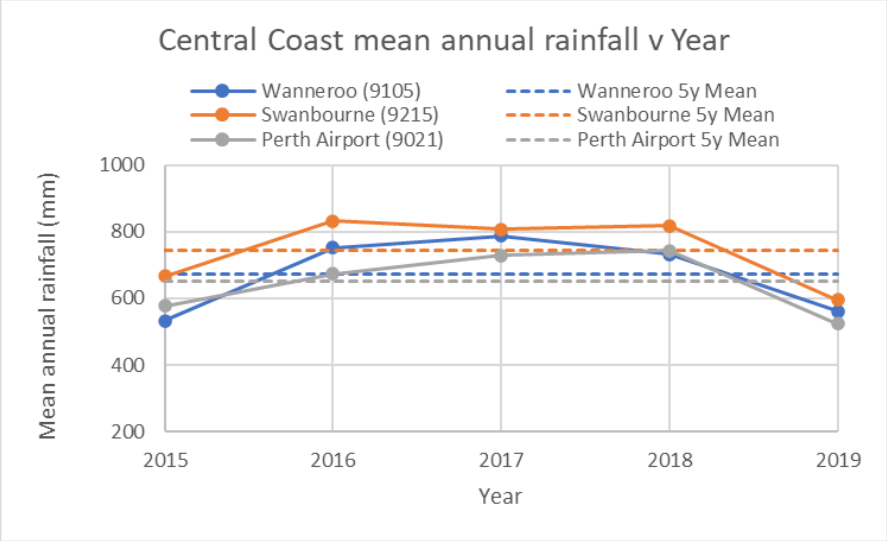
It may be useful to view the formulae used in the exemplar spreadsheet. Clicking on cells will display the formula in the formula bar. Bureau of Meteorology monthly rainfall data contained within this spreadsheet is locked, with data source references included.

The following weather stations and rainfall districts have been used in this exemplar, however data for other weather stations and districts is available from the Bureau of Meteorology climate data website.

Central Coast District (Perth Metropolitan)		South Coast District		West Gascoyne District	
Wanneroo	9105	Albany	9500	Edaggee	6020
Swanbourne	9215	Myrup	9584	Steep Point	6102
Perth Airport	9021	Northcliffe	9590	Shark Bay Airport	6105

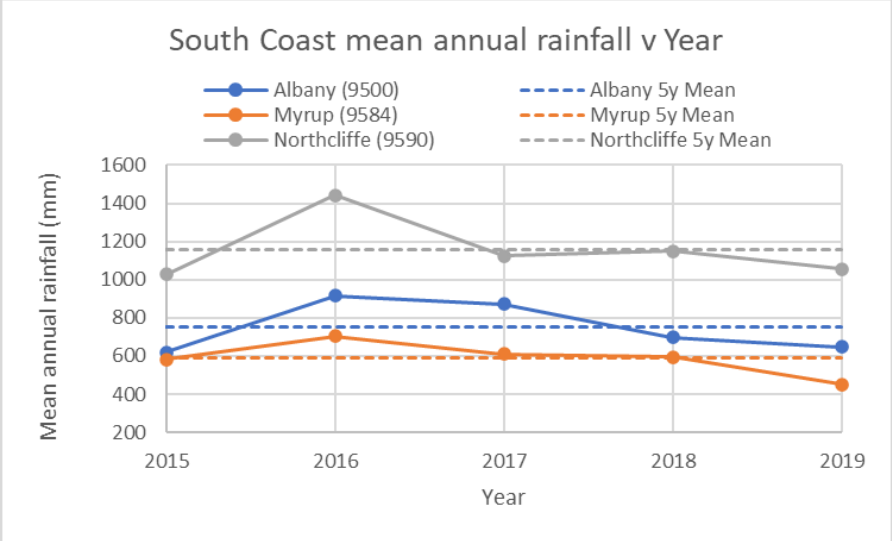
The following pages contain calculated data for weather stations within districts and when comparing districts, and graphical representations that may be useful when having students answering the questions.

Central Coast	District Monthly Mean rainfall						WS#9105		WS#9215		WS#9021	
	2015	2016	2017	2018	2019	5y Mn	5y Mean	5y Median	5y Mean	5y Median	5y Mean	5y Median
Jan	0.5	15.3	43.6	112.8	6.4	35.7	34.7	8.5	40.2	15.2	32.3	23.4
Feb	23.8	5.5	114.4	0.4	0.4	28.9	29.7	14.0	33.0	1.8	24.0	0.8
Mar	19.7	26.0	25.8	2.8	5.5	16.0	16.2	15.8	17.9	21.4	13.9	19.8
Apr	55.2	77.2	0.1	24.0	35.8	38.5	39.3	37.1	43.4	35.2	32.7	35.2
May	89.1	111.4	64.7	64.8	15.3	69.1	69.3	66.1	74.6	71.0	63.2	68.8
Jun	91.3	113.1	55.6	150.8	193.5	120.9	117.2	109.0	136.4	144.0	109.1	100.2
Jul	108.5	142.8	180.0	166.2	115.0	142.5	138.1	141.2	151.1	157.8	138.3	129.4
Aug	95.8	137.8	152.0	165.8	119.9	134.3	132.7	142.7	134.6	139.0	135.5	131.6
Sep	42.0	69.1	89.3	37.4	29.2	53.4	51.8	44.0	56.2	37.2	52.1	46.8
Oct	27.7	38.3	26.8	34.5	25.0	30.5	28.0	28.0	34.8	32.8	28.5	26.0
Nov	20.6	8.7	1.9	3.7	12.4	9.5	6.3	6.7	13.2	6.6	8.9	5.6
Dec	18.3	7.7	20.7	1.5	1.7	10.0	10.0	8.4	8.4	6.0	11.4	8.6
Total	**592.6**	**752.8**	**774.9**	**764.7**	**560.3**	**689.1**	**673.3**	**733.0**	**743.8**	**807.6**	**650.1**	**674.4**



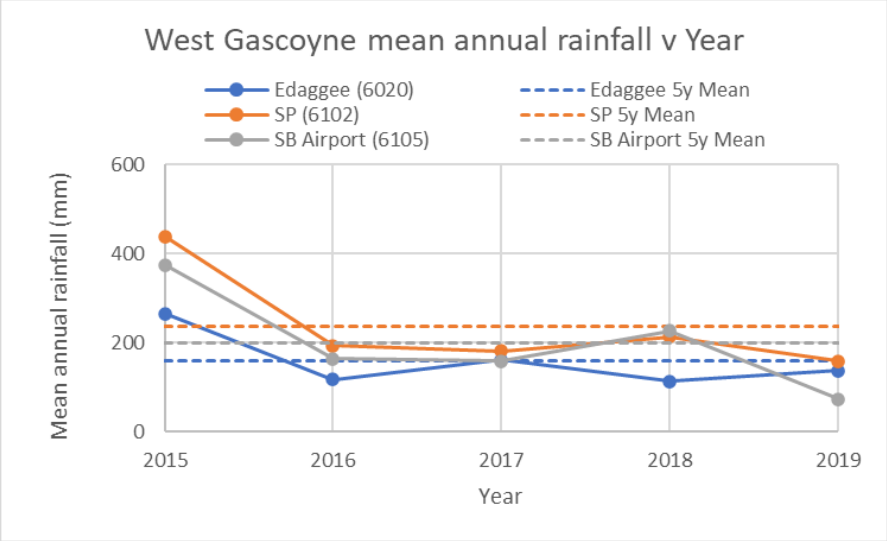
STEM Consortium

South Coast	District Monthly Mean rainfall						WS#9500		WS#9584		WS#9590	
	2015	2016	2017	2018	2019	5y Mn	5y Mean	5y Median	5y Mean	5y Median	5y Mean	5y Median
Jan	4.1	88.8	21.5	18.7	13.1	29.2	22.6	14.6	28.3	8.4	36.8	22.0
Feb	12.6	21.4	71.9	39.6	2.1	29.5	20.5	20.9	49.7	15.0	18.4	13.2
Mar	37.2	39.0	51.4	18.8	33.4	35.9	37.3	42.1	29.1	23.5	41.5	39.6
Apr	74.3	70.1	24.5	39.0	52.9	52.1	65.8	54.2	32.4	39.9	58.2	67.1
May	84.7	126.1	86.8	59.3	49.1	81.2	58.6	57.7	57.5	66.7	127.5	122.8
Jun	99.3	128.0	50.2	95.9	153.5	105.4	85.0	83.2	57.1	61.5	174.0	181.7
Jul	111.4	142.8	176.4	166.9	116.0	142.7	132.9	135.8	63.6	54.4	231.6	249.8
Aug	131.4	153.3	135.0	179.6	130.3	145.9	128.4	119.1	110.5	108.6	199.0	190.2
Sep	67.7	112.5	120.1	60.3	61.2	84.3	82.8	65.3	59.9	67.2	110.3	109.8
Oct	52.1	80.3	54.8	67.7	57.6	62.5	58.7	59.5	48.6	42.4	80.2	73.9
Nov	22.3	17.4	20.1	36.1	38.2	26.8	25.7	22.6	26.2	28.0	28.5	24.4
Dec	47.1	40.9	56.3	32.1	11.2	37.5	32.3	34.1	26.0	18.8	54.2	53.0
Total	**744.1**	**1020.5**	**869.0**	**813.9**	**718.5**	**833.2**	**750.6**	**696.6**	**588.9**	**595.9**	**1160.1**	**1123.8**



STEM Consortium

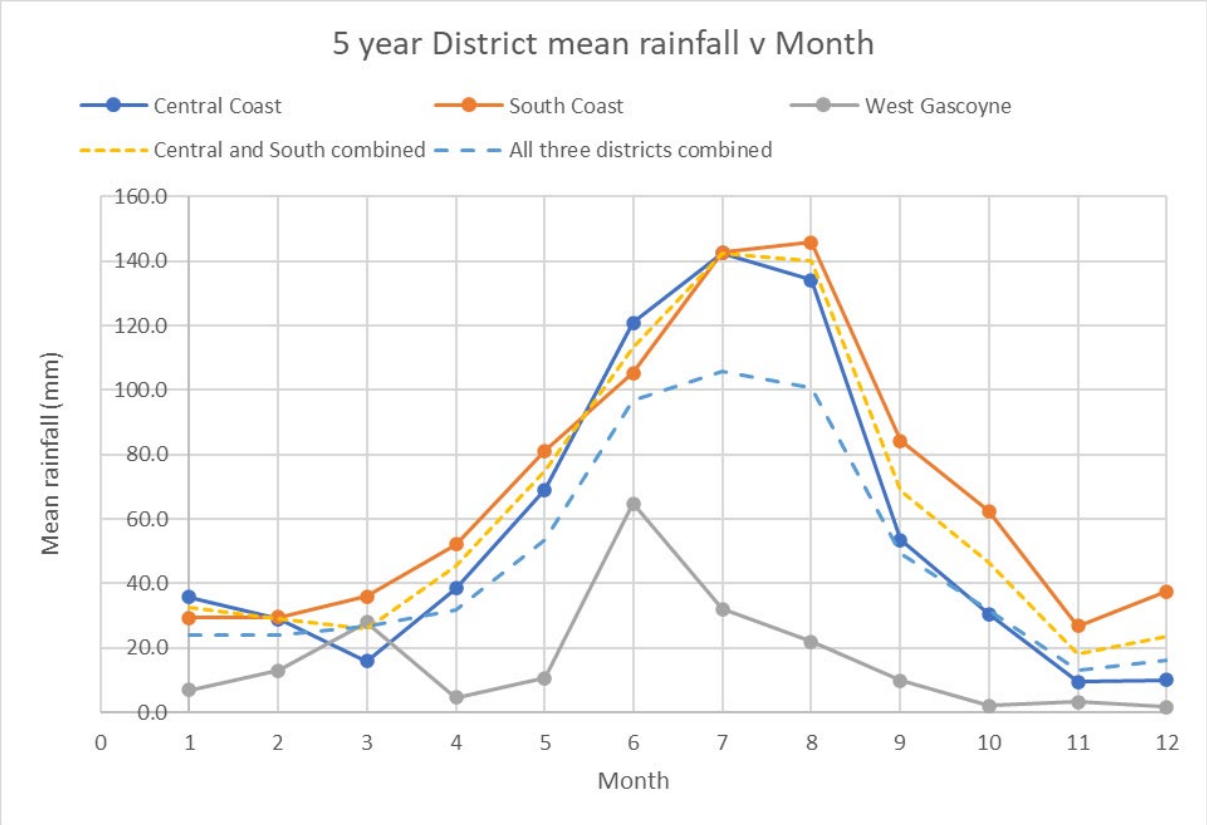
West Gascoyne	District Monthly Mean rainfall						WS#6020		WS#6102		WS#6105	
	2015	2016	2017	2018	2019	5y Mn	5y Mean	5y Median	5y Mean	5y Median	5y Mean	5y Median
Jan	0.1	4.1	28.2	0.7	1.1	6.8	5.8	2.3	9.2	0.4	5.5	2.0
Feb	10.5	0.0	52.4	1.5	0.0	12.9	20.7	0.5	4.1	0.4	13.9	3.6
Mar	137.6	1.1	0.0	0.9	0.1	27.9	28.0	0.2	19.2	0.2	36.6	0.2
Apr	2.4	4.6	0.0	0.1	16.2	4.7	8.7	0.4	3.3	3.0	2.0	2.4
May	15.9	13.2	2.8	21.3	0.1	10.6	5.4	3.4	16.1	18.1	10.4	7.8
Jun	103.6	43.1	5.6	89.1	83.1	64.9	44.8	53.4	95.1	107.4	54.8	44.4
Jul	42.6	61.8	13.7	30.7	11.8	32.1	18.8	13.0	42.0	42.0	35.6	37.0
Aug	32.1	23.0	22.9	22.9	8.7	21.9	12.1	11.8	28.8	25.8	24.8	30.0
Sep	2.7	5.5	33.0	7.1	0.9	9.9	6.0	3.0	14.2	9.2	9.3	2.6
Oct	0.1	0.5	1.4	7.5	0.5	2.0	1.8	1.0	1.5	1.2	2.7	0.6
Nov	12.1	1.4	2.2	0.0	0.1	3.2	3.7	1.0	2.4	1.0	3.4	2.2
Dec	0.1	0.0	5.1	2.5	0.1	1.6	3.1	0.0	1.3	0.4	0.3	0.0
Total	**359.8**	**158.2**	**167.3**	**184.3**	**122.8**	**198.5**	**158.8**	**136.8**	**237.3**	**193.8**	**199.4**	**163.4**



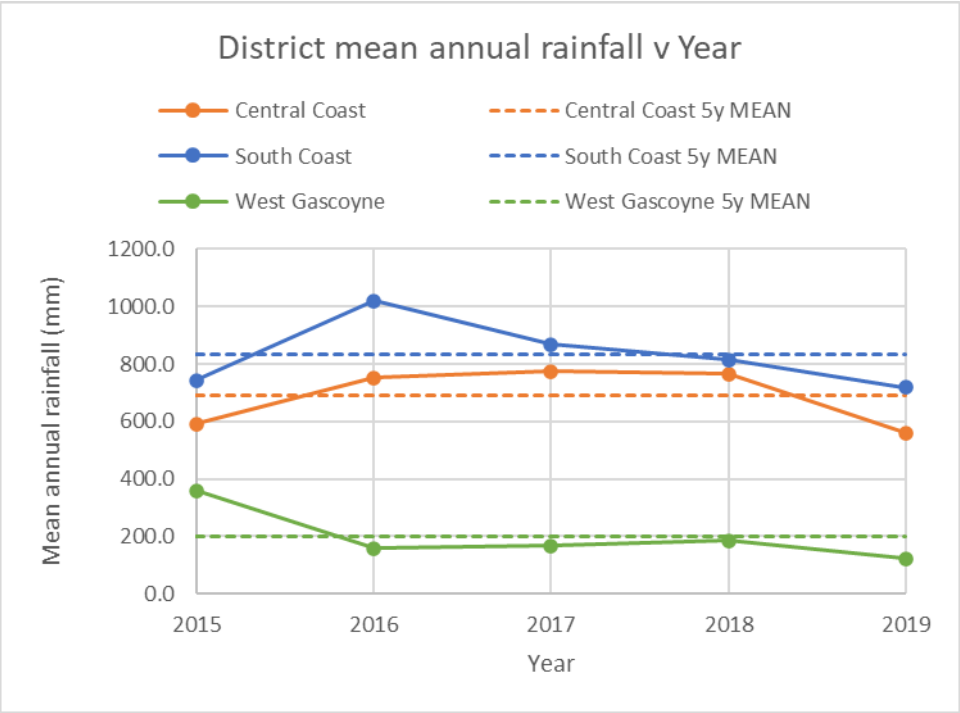
STEM Consortium

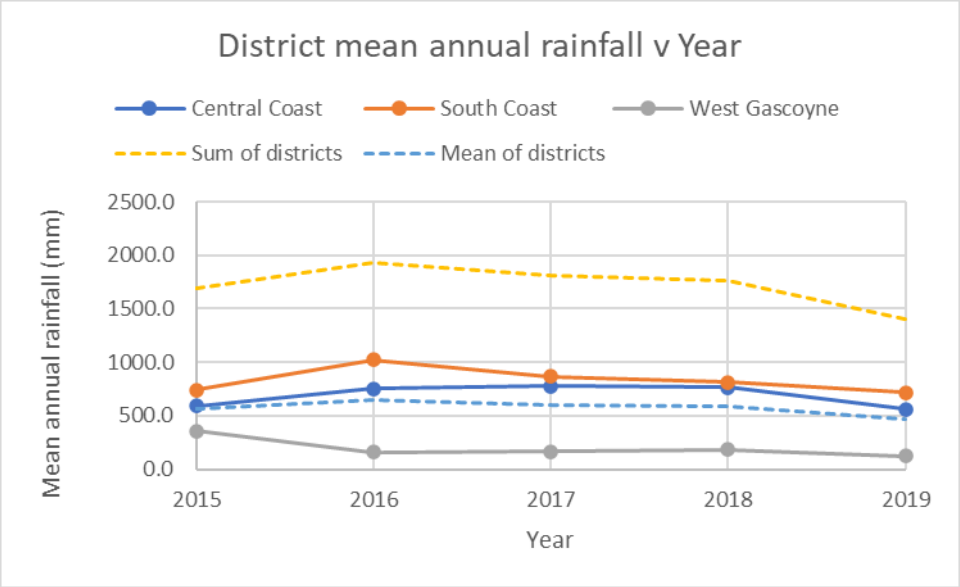
5y Monthly Mean Rainfall Data	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
Central Coast	35.7	28.9	16.0	38.5	69.1	120.9	142.5	134.3	53.4	30.5	9.5	10.0	689.1
South Coast	29.2	29.5	35.9	52.1	81.2	105.4	142.7	145.9	84.3	62.5	26.8	37.5	833.2
West Gascoyne	6.8	12.9	27.9	4.7	10.6	64.9	32.1	21.9	9.9	2.0	3.2	1.6	198.5

	District Monthly Mean Annual Rainfall					5y District Annual Rainfall		
	2015	2016	2017	2018	2019	Mean	Median	Range
Central Coast	592.6	752.8	774.9	764.7	560.3	689.1	752.8	214.7
South Coast	744.1	1020.5	869.0	813.9	718.5	833.2	813.9	302.0
West Gascoyne	359.8	158.2	167.3	184.3	122.8	198.5	167.3	237.0



STEM Consortium





STEM Consortium

Bureau of Meteorology climate data from the following sources has been:

- Central Coast District (Perth Metropolitan)
    - Wanneroo (9105)
      bom.gov.au/jsp/ncc/cdio/weatherData/av?p_nccObsCode=139&p_display_type=dataFile&p_startYear=&p_c=&p_stn_num=009105
    - Swanbourne (9215)
      bom.gov.au/jsp/ncc/cdio/weatherData/av?p_nccObsCode=139&p_display_type=dataFile&p_startYear=&p_c=&p_stn_num=009215

    - Perth Airport (9021)

      bom.gov.au/jsp/ncc/cdio/weatherData/av?p_nccObsCode=139&p_display_type=dataFile&p_startYear=&p_c=&p_stn_num=009021

- South Coast District
    - Albany (9500)

      bom.gov.au/jsp/ncc/cdio/weatherData/av?p_nccObsCode=139&p_display_type=dataFile&p_startYear=&p_c=&p_stn_num=009500
    - Myrup (9584)

      bom.gov.au/jsp/ncc/cdio/weatherData/av?p_nccObsCode=139&p_display_type=dataFile&p_startYear=&p_c=&p_stn_num=009584
    - Northcliffe (9590)

      bom.gov.au/jsp/ncc/cdio/weatherData/av?p_nccObsCode=139&p_display_type=dataFile&p_startYear=&p_c=&p_stn_num=009590
- West Gascoyne District
    - Edaggee (6020)

      bom.gov.au/jsp/ncc/cdio/weatherData/av?p_nccObsCode=139&p_display_type=dataFile&p_startYear=&p_c=&p_stn_num=006020
    - Steep Point (6102)

      bom.gov.au/jsp/ncc/cdio/weatherData/av?p_nccObsCode=139&p_display_type=dataFile&p_startYear=&p_c=&p_stn_num=006102
    - Shark Bay Airport (6105)

      bom.gov.au/jsp/ncc/cdio/weatherData/av?p_nccObsCode=139&p_display_type=dataFile&p_startYear=&p_c=&p_stn_num=006105

The following links can be used as a review of features of *Excel* used to create these spreadsheets if required:

- SUM function (Microsoft)
  support.office.com/en-us/article/sum-function-043e1c7d-7726-4e80-8f32-07b23e057f89?ui=en-US&rs=en-AU&ad=AU
- AVERAGE function (Microsoft)

  support.office.com/en-us/article/AVERAGE-function-047BAC88-D466-426C-A32B-8F33EB960CF6
- MEDIAN function (Microsoft)

  support.office.com/en-us/article/MEDIAN-function-D0916313-4753-414C-8537-CE85BDD967D2
- MIN function (Microsoft)

  support.office.com/en-us/article/MIN-function-61635D12-920F-4CE2-A70F-96F202DCC152

- MAX function (Microsoft)

  support.office.com/en-us/article/MAX-function-E0012414-9AC8-4B34-9A47-73E662C08098

- TRANSPOSE function (Microsoft)

  support.office.com/en-us/article/TRANSPOSE-function-ED039415-ED8A-4A81-93E9-4B6DFAC76027

- Absolute cell references (Excel Trick)
  www.exceltrick.com/formulas_macros/excel-relative-and-absolute-references/

Appendix 13: Student activity sheet 1.8: Kangaroo populations and rainfall comparisons

Introduction

The problem that is being explored is:

Does annual rainfall influence kangaroo populations in Western Australia?

Comparisons between the l Western Grey and Red kangaroo populations estimated annually and the average annual rainfall data for regions within Western Australia will need to be developed. Students construct a data set for analysis using climate data from the Bureau of Meteorology and kangaroo population estimates. The following spreadsheet provides data for kangaroo population and monthly district rainfall averages for the years 2001-2018. This data is locked within the individual sheets.

Exemplar spreadsheet

App 13: Student Kangaroo Pop.

Explanation of the spreadsheet

The exemplar spreadsheet contains the following:

- Bureau of Meteorology average monthly rainfall values for the 16 rainfall districts in Western Australia for the years 2001-2018
- annual population estimates of Western Grey and Red kangaroo population.

Task

In this activity you will compare and analyse data and information from numerous sources to explore possible relationships between rainfall and size of kangaroo populations. Using the spreadsheets created in *Student activity sheet 1.6: Annual rainfall comparisons* as a guide, information from other sources will guide your analysis and allow you to visualise data and analyse patterns and trends to draw conclusions that are consistent with the evidence.

Summary of steps:

A summary of the steps to be undertaken includes:

- identifying kangaroo management areas within Western Australia
- identifying kangaroo species distribution within Western Australia
- comparing kangaroo management areas to the rainfall districts
- selecting suitable rainfall districts for each of the kangaroo species
- calculating annual rainfall for the years 2001-2018
- comparing mean annual rainfall across chosen rainfall districts to the kangaroo species population
- visualising and analysing trends to draw conclusions

Instructions for identifying kangaroo management areas and species distribution

1. Identify the West Australian kangaroo management monitoring areas in Figure 1 (pg. 4) and Figure 2.2 (pg. 5) at:
    - *https://www.dpaw.wa.gov.au/images/documents/plants-animals/animals/kangaroos/Management%20plan%20for%20the%20commercial%20harvest%20of%20kangaroos%20in%20Western%20Australia%202019-2023.pdf*
2. Western Red and Grey kangaroo distributions in Western Australia can be found in Figure 3 (pg. 1) and Figure 4 (pg. 15):
    - *https://www.dpaw.wa.gov.au/images/documents/plants-animals/animals/kangaroos/commercial_harvest_of_kangaroos_management_plan.pdf*
    - *dpaw.wa.gov.au/images/documents/plants-animals/animals/kangaroos/fauna_note_29_western_grey_kangaroo.pdf*

    Compare the Western Australian kangaroo management monitoring areas with the kangaroo distributions for each of the two species.
3. Australia is divided up into rainfall districts. Information on these districts is available at:
    - *bom.gov.au/climate/cdo/about/rain-districts.shtml*
    - *bom.gov.au/climate/cdo/metadata/maps/raindist.pdf*
4. Using the rainfall district, kangaroo management monitoring area and kangaroo distribution information, identify the rainfall districts where the kangaroo populations are monitored. Identify the rainfall districts for both the Western Grey and Red kangaroo species.
5. For each of the Western Grey and Red kangaroo species, choose three rainfall districts that you will analyse to investigate the relationship between the annual rainfall ion the kangaroo population estimates.

Student kangaroo population spreadsheet data

The *Student kangaroo population* spreadsheet contains data from the following sources:

- Kangaroo population data
    - *environment.gov.au/biodiversity/wildlife-trade/natives/wild-harvest/kangaroo-wallaby-statistics/kangaroo-population#2010*
    - *environment.gov.au/biodiversity/wildlife-trade/natives*
    - *environment.gov.au/system/files/pages/ee20f301-6c6c-44e4-aa24-62a32d412de5/files/kangaroo-statistics-states-2018.xlsx*
- District monthly average rainfall data for the years 2001-2018
    - *bom.gov.au/anon/home/ncc/www/awa/rainfall/district_averages/monthly/*

Student kangaroo population spreadsheet use

Rainfall district data

6. Open the *Student kangaroo population* spreadsheet. Calculate the annual total rainfall for all districts for the years 2001-2018.
7. Copy the annual total rainfall data for the years 2001-2018 into the student analysis sheet.
8. Identify and highlight the rainfall districts that you have chosen to analyse for the Western Grey kangaroo species. These will be your Western Grey kangaroo rainfall districts.
9. For the years 2001-2018, calculate the mean and median annual rainfall for your Western Grey kangaroo rainfall districts.
10. Identify and highlight, in a different colour, the rainfall districts that you have chosen to analyse population numbers of Red kangaroo species. These will be your Red kangaroo rainfall districts.
11. For the years 2001-2018, calculate the mean and median annual rainfall for your Red kangaroo rainfall districts.
12. Using your Western Grey kangaroo rainfall districts, combine the data to calculate the mean and median rainfalls for each of the individual years 2001 to 2018.
13. Create a suitable visualisation of the Western Grey kangaroo combined district annual rainfall to show the rainfall trend over the years from 2001 to 2018.
14. Using your Red kangaroo rainfall districts, combine the data to calculate the mean and median rainfalls for each of the individual years 2001 to 2018.
15. Create a suitable visualisation of the Red kangaroo combined district annual rainfall to show the rainfall trend over the years from 2001 to 2018.

Kangaroo population estimate data

16. Using the data in the kangaroo population sheet, create a suitable visualisation of the Western Grey, Red and total kangaroo population estimates for every year from 2001 to 2018.

Analysing your data

Using your kangaroo population estimates and combined district annual rainfall visualisations, and any additional visualisations or graphs that make your data easier to analyse, look for trends. Consider what effect higher or lower than average rainfalls may have on kangaroo populations and compare the traces individually for the Western Grey and Red kangaroo species. When comparing these traces, look for trends over the whole range of years from 2001 to 2018, and consider making comparisons over 5-year blocks. Especially take note of peaks and troughs, in what years these occur, and consider what your data might be showing around the years of 2006 and 2012.

Discuss your findings within your group, and then complete the questions.

Questions to answer

1. Identify three interesting observations from your data analysis for each of the kangaroo species and explain what additional data may have been useful to analyse to provide additional insight or information.

2. Based on your data analysis, does annual rainfall influence the Western Grey kangaroo population in Western Australia? Comment.

3. Based on your data analysis, does annual rainfall influence the Red kangaroo population in Western Australia? Comment.

4. When population surveys are incomplete or not conducted every year, kangaroo population estimates are scaled according to trends in regional rainfall (Department of Biodiversity, Conservation and Attractions, 2018). Why might rainfall be a good indicator for populations of kangaroos?

5. To gain a greater understanding on this rainfall influence on kangaroo populations, what additional information may you need to analyse?

Reference:

2019 Commercial Kangaroo Harvest Quota Submission for Western Australia (Department of Biodiversity, Conservation and Attractions, 2018)

dpaw.wa.gov.au/images/documents/plants-animals/animals/kangaroos/2019%20Commercial%20kangaroo%20harvest%20quota%20submission%20for%20Western%20Australia.pdf

Appendix 14: Teacher resource sheet 1.9: Exemplar spreadsheet for kangaroo populations and rainfall comparisons

Introduction

The following sample spreadsheet is intended for use by the teacher only to provide an example of what might be generated by students in Activity 1 using *Student activity sheet 1.8: Kangaroo populations and rainfall comparisons*.

Exemplar spreadsheet

App 14: Exemplar Teacher kangaroo

Explanation of the spreadsheet

The exemplar spreadsheet contains the following:

- Bureau of Meteorology average monthly rainfall values for the 16 rainfall districts in Western Australia for the years 2001-2018
- annual Western Grey and Red kangaroo population estimates
- sample analysis and visualisation of data.

It may be useful to view the formulas used in the exemplar spreadsheet. Clicking on cells will display their formula in the formula bar. Data for kangaroo population estimates and monthly district rainfall averages for each year are in individual sheets. This secondary source data is locked to prevent accidental changes being made, and each sheet contains the reference for the data source. Both locked and unlocked sample analysis have been included.

Data acquisition process

The kangaroo population and district monthly average rainfall data used in the *Exemplar Teacher kangaroo population* spreadsheet has been acquired and processed through the following process:

- Extraction of the annual Western Australian Western Grey and Red kangaroo population estimates and placing into the spreadsheet. Data acquired from:
    - *environment.gov.au/biodiversity/wildlife-trade/natives/wild-harvest/kangaroo-wallaby-statistics/kangaroo-population#2010*
    - *environment.gov.au/biodiversity/wildlife-trade/natives*
    - *environment.gov.au/system/files/pages/ee20f301-6c6c-44e4-aa24-62a32d412de5/files/kangaroo-statistics-states-2018.xlsx*
- Extraction of the district monthly average rainfall data for the years 2001-2018 and placing into the spreadsheet. Data acquired from:
    - *bom.gov.au/anon/home/ncc/www/awa/rainfall/district_averages/monthly/*

Comparisons of kangaroo distributions and rainfall districts

The kangaroo population estimates are acquired through surveys of monitoring regions. It is advised that students compare the Western Grey and Red kangaroo species distributions separately. When determining the rainfall districts of interest for each species, students should first carefully compare the kangaroo species distribution to the kangaroo management areas. The kangaroo management areas in Western Australia do not cover all rainfall districts, and therefore not all rainfall districts will be suitable for analysis. The following is a suggestion for possible rainfall district data sets to compare to population estimates for each of the species of kangaroo:

- Western Grey Kangaroo
    - 8. North Coast
    - 9. Central Coast
    - 9A. South Coast
    - 10. North Central
    - 10A. South Central
    - 11. Eucla
    - 12. Southeast
- Red Kangaroo
    - 4. De Grey
    - 5. Fortescue
    - 6. West Gascoyne
    - 7. East Gascoyne
    - 7A. Murchison
    - 11. Eucla
    - 12. Southeast

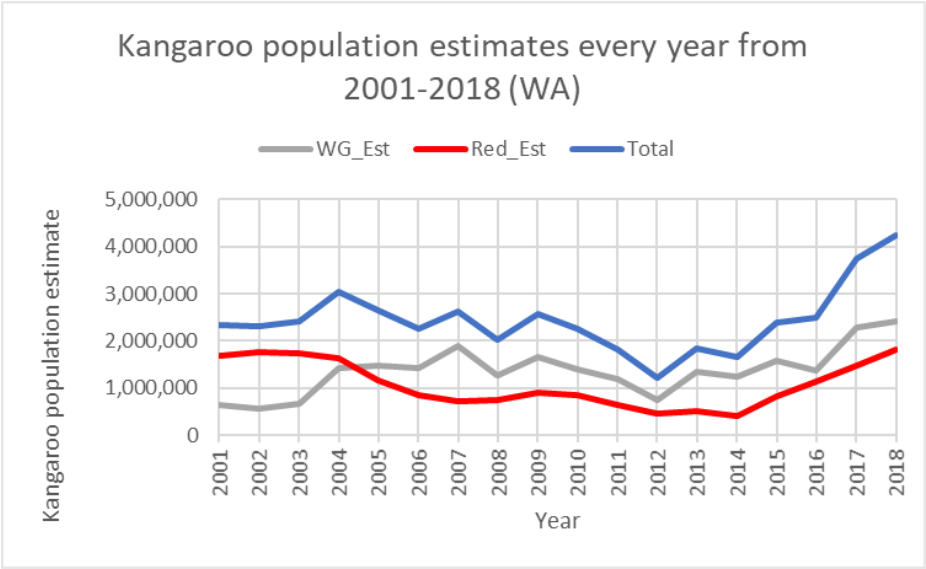
Visualisations have been included as an example of how students may display their data to allow trend analysis. Western Australian kangaroo population estimates from 2001-2018 is included, together with sample rainfall district combinations that may be suitable when comparing the kangaroo population estimate data for each of the Western Grey and Red kangaroo species.

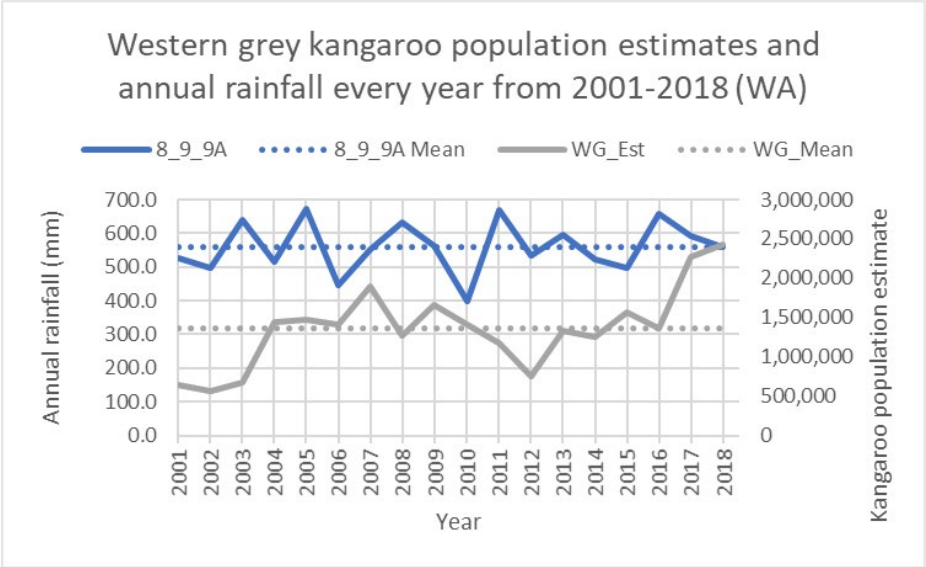
Students should observe that by combining neighbouring annual rainfall district data, it may be easier to observe trends. They should also be able to observe that there are trends that could be explained further by increasing the data set, and by bringing in additional information.

When analysing this data, students should be encouraged to consider aspects of the data that may influence the results. Such aspects may include the differences in wet seasons and the comparison of annual rainfall rather than monthly data, the breeding cycle of kangaroos, the sustainable harvesting of kangaroos, the times of the year when kangaroo populations are surveyed and the lack of publicly available district information for kangaroo populations. Despite this, trends can be observed that indicate a delay of one to two years in the population data

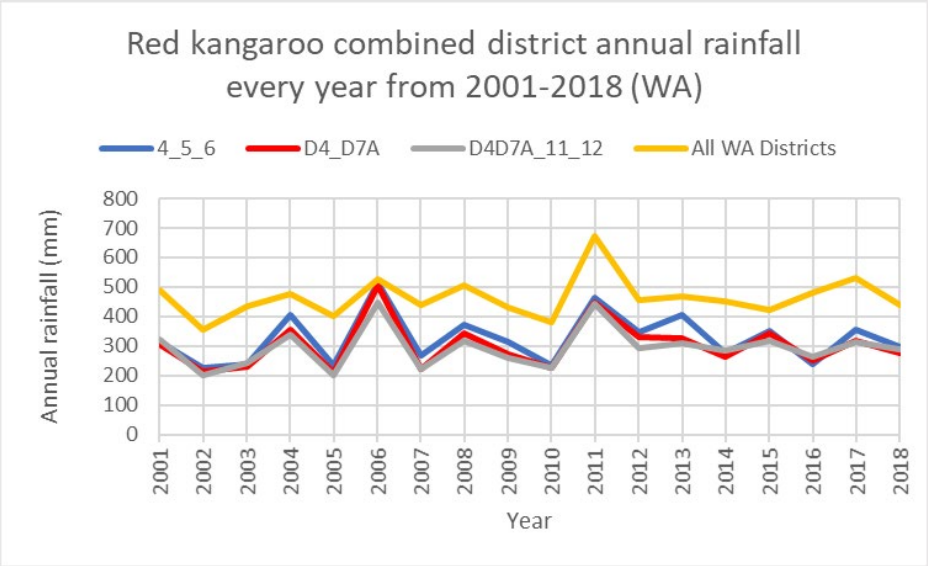
when compared to peak and trough annual rainfall events. This is most evident in the Western Grey kangaroo population estimates. To assist with observing trends, mean kangaroo population estimates and mean annual rainfall traces may be useful to include. Red kangaroo population data trends are less conclusive, but this is also impacted by the larger districts over which the red kangaroo surveys are conducted, and the lower population numbers.

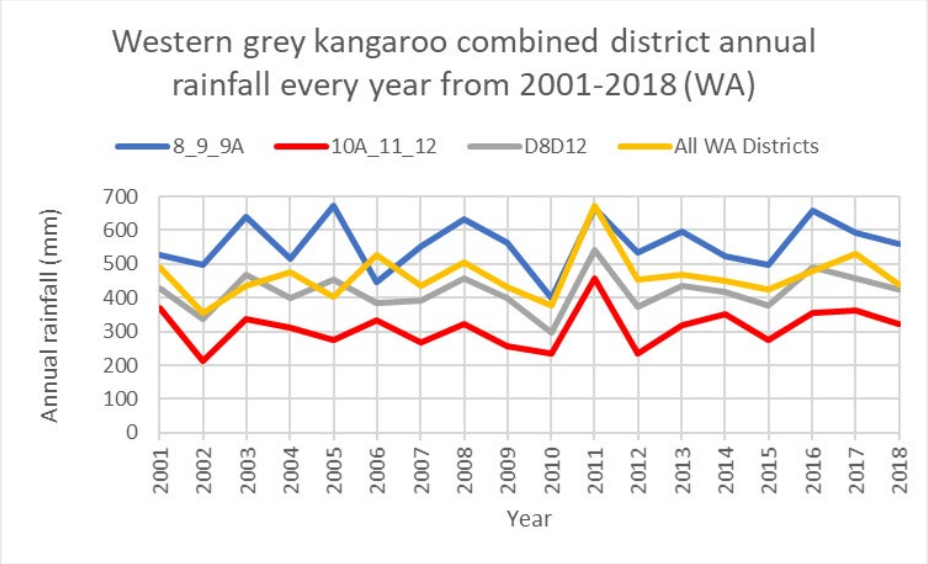
Encourage students to consider what additional data could be included to conclusively determine the annual rainfall influence on kangaroo populations, and how limited data sets may make making such conclusions more problematic. In spite of this, rainfall data is used for sustainable kangaroo harvesting population estimates when insufficient surveying has been conducted.

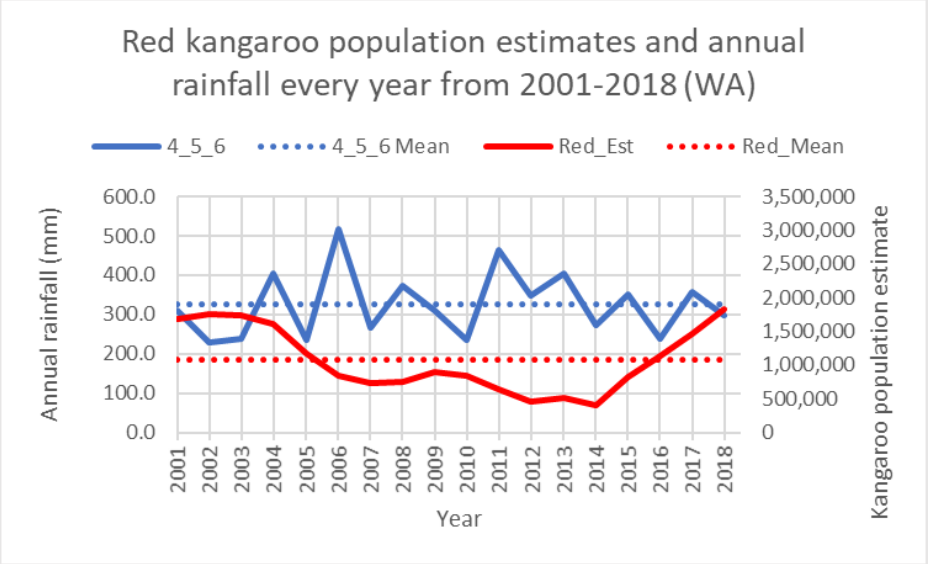




STEM Consortium







STEM Consortium

The following links can be used as a review of features of *Excel* used to create these spreadsheets if required:

- SUM function (Microsoft)
  support.office.com/en-us/article/sum-function-043e1c7d-7726-4e80-8f32-07b23e057f89?ui=en-US&rs=en-AU&ad=AU
- AVERAGE function (Microsoft)

  support.office.com/en-us/article/AVERAGE-function-047BAC88-D466-426C-A32B-8F33EB960CF6
- MEDIAN function (Microsoft)

  https://support.office.com/en-us/article/MEDIAN-function-D0916313-4753-414C-8537-CE85BDD967D2
- MIN function (Microsoft)

  support.office.com/en-us/article/MIN-function-61635D12-920F-4CE2-A70F-96F202DCC152
- MAX function (Microsoft)

  support.office.com/en-us/article/MAX-function-E0012414-9AC8-4B34-9A47-73E662C08098
- TRANSPOSE function (Microsoft)

  support.office.com/en-us/article/TRANSPOSE-function-ED039415-ED8A-4A81-93E9-4B6DFAC76027
- Absolute cell references (Excel Trick)
  www.exceltrick.com/formulas_macros/excel-relative-and-absolute-references/

Appendix 15: Student activity sheet 1.10: Gapminder data 'deep dive'

Resource ('Tool' or 'Data' menu, 'Tool' type – Bubbles, Trend)	**Type of data representation** (eg spreadsheet, graphs)	**Sources of big data** (eg Source of the 'indicators' for the axes)	**Interesting observations** (eg What do you notice in the data)	**Who might use this information** (eg Police, local council, store owners)

Appendix 16: Teacher resource sheet 1.11: Big data fact sheet 1 – What is big data?

What is big data?

"It is the accumulation and analysis of information. Lots of information. Oceans of information. Every time someone clicks on something at Amazon, it's recorded and another drop is added to the ocean. Every time a scanner beeps at the supermarket checkout. Every time a home electricity meter reports a reading. Every time a parcel passes a FedEx checkpoint. Every time a customs officer checks a passport, every time someone posts to Facebook, every time someone does a Google search — the ocean swells."

—Dan Gardner, 'An Ocean of Data' (excerpted from *The Human Face of Big Data*)

Big data is shaping businesses, shifting markets, and transforming our world. During the first day of a baby's life, the amount of data generated by humanity is equivalent to 70 times the information contained in the Library of Congress.

Big data is a simple term used to describe the emergence of incredibly powerful ways to gather and analyse digital information to gain new insights about nearly every aspect of our world and lives.

It is the ability to extract meaning: to sort through huge masses of numbers and find the hidden patterns, the unexpected correlations and the surprising connections.

What's creating the deluge of information?

Each of us now leaves a trail of digital exhaust, an infinite stream of phone records, texts, browser histories, GPS data, and other information that will live on forever. This data comes from:

- Sensors on mobile phones
- Online shopping
- Medical imaging
- GPS-enabled cameras and smartphones
- Satellites
- Smart electrical grids
- Video surveillance
- Social media: Facebook, Twitter, Google
- Digital music
- Digital photographs
- Large databases in the banking sector
- Gene sequencing.

You can think about big data as the process of helping the planet grow a nervous system, one in which we are just another human type of sensor.

How to measure big data and what's a byte?

(A byte reminder – each one is 1,000 times larger than the one before it!)

- A ***byte*** = one character, or a grain of sand
- A ***kilobyte*** = a sentence, or a couple of pinches of sand
- A ***megabyte*** = a 20-slide *PowerPoint* show, a small book, or a tablespoon of sand
- A ***gigabyte*** = 9 metres of books on a shelf, or a shoebox full of sand
- A ***terabyte*** = 300 hours of good-quality video, a tenth of the Library of Congress, or a playground sandbox
- A ***petabyte*** = 350,000 digital pictures, or a 1.6km-long stretch of beach
- An ***exabyte*** = half the information generated worldwide in 1999, or a beach stretching from Perth to Karratha
- A ***zettabyte*** = unimaginable, or a beach as big as all the coastlines in the world

Big data facts and stats

- The world's data doubles every two years.
- In 2011, humans created 1.8 zettabytes of data. This is equivalent to 200 billion high-definition movies that are at least 120-minutes long. It would take one person 47 million years to watch all those movies.
- *YouTube* users upload 48 hours of new video every minute of the day.
- 20 petabytes of data are processed daily (that's 20 billion megabytes).
- There are one billion Tweets created every 72 hours.
- 30 billion pieces of content are shared on *Facebook* monthly.
- A personal computer holds about 500 gigs, so it would require about 20 billion PCs to store all of the world's data.
- $600 buys a disk drive that can store all of the world's music.
- Lady Gaga has 28 million *Twitter* followers. Barack Obama has 18 million.

Appendix 17: Student resource sheet 1.12: Big data fact sheet 2 – Applications of big data, now and in the future

The following are examples of how big data is currently being applied to a range of contexts. Where else could big data be used to solve a problem, mitigate risk or create an opportunity?

Big data in health

The level of data generated within healthcare systems is extensive. Big data analytics has helped healthcare improve by providing personalised medicine and prescriptive analytics, clinical risk intervention and predictive analytics, waste and care variability reduction, automated external and internal reporting of patient data, standardised medical terms and patient registries and fragmented point solutions.

With the added adoption of mHealth (mobile health), eHealth (information and communications technologies in health care) and wearable technologies the volume of data will continue to increase. This includes electronic health record data, imaging data, patient-generated data, sensor data, and other forms of difficult to process data. There is now an even greater need for such environments to pay greater attention to data and information quality.

Big data in health research is particularly promising in terms of exploratory biomedical research, as data-driven analysis can move forward more quickly than hypothesis-driven research. Then, trends seen in data analysis can be tested in traditional, hypothesis-driven follow-up biological research and eventually clinical research.

Examples:

- Data scientists and engineers compile all the data from wearable fitness trackers to continually revise and revamp their products and their interfaces to be better, plus all that collected health information generates huge databases for analysis by the healthcare industry.
- Ingestible Sensors: Tiny vitamin sized pills the size of a grain of sand containing integrated circuits and a battery that transmits heart rate, respiration, body posture, even sleeping patterns as they travel through your body.

Big data in media

The industry appears to be moving away from the traditional approach of using specific media environments such as newspapers, magazines, or television shows and instead taps into consumers with technologies that reach targeted people at optimal times in optimal locations. For example, publishing environments are increasingly tailoring messages (advertisements) and content (articles) to appeal to

consumers that have been exclusively gleaned through various *data-mining* activities.

Examples:

- A great amount of data is also generated from the media and social media, and the entertainment industry uses what it learns from that data to customize its products. For example, Netflix discovered a correlation between the colour of a TV show title's cover art and customer response. The media mogul also uses its collected data on viewer habits and preferences—all those clicks and choices you make when you're watching (or stopped watching) shows and deciding on your "watch list"—to tailor its productions, including its hit series House of Cards.
- Finding the next Lady Gaga: A company called Next Big Sound analyses social media to predict which emerging musicians are likely to write the next big hits.
- "Relying on social networks and analytics, companies are gathering volumes of data from the web to help musicians and music companies better understand their audiences by tracking hundreds of thousands of artists' new fans, plays, views, and comments across social media, and them correlating them with events such a song releases or shows, a company called Next Big Sound can identify a band's popularity on the web before record sales rise. Next Big Sound's technology singled out Alabama Shakes, headed by lead singer Brittany Howard, as one of the 15 fastest accelerating artists across the Internet, even before the band's shows sold out and before they appeared on the Late Show with David letterman.

Michael Malone (excerpted from *The Human Face of Big Data*)

Big data in insurance

Health insurance providers are collecting data on social "determinants of health" such as food and TV consumption, marital status, clothing size and purchasing habits, from which they make predictions on health costs, to spot health issues in their clients. It is controversial whether these predictions are currently being used for pricing.

Example:

"Instead of 'find my iPhone', some auto insurance companies are offering a service that may enable parents to 'find my teenager'. Progressive Insurance, for example, offers the Snapshot, a tracking device that reports on a car's location, acceleration, braking, and distance travelled. Owners who install the device can get a 10 to 15 per cent discount on their policy. Privacy activists, however, fear the technology is ripe for abuse".

Michael Malone (excerpted from *The Human Face of Big Data*).

Big data in the Internet of Things (IoT)

Big data and the IoT work in conjunction. Data extracted from IoT devices provides a mapping of device interconnectivity. Such mappings have been used by the media industry, companies and governments to more accurately target their audience and increase media efficiency. IoT is also increasingly adopted as a means of gathering sensory data, and this sensory data has been used in medical, manufacturing and transportation contexts.

Kevin Ashton, digital innovation expert who is credited with coining the term, defines the Internet of Things in this quote: "If we had computers that knew everything there was to know about things—using data they gathered without any help from us—we would be able to track and count everything, and greatly reduce waste, loss, and cost. We would know when things needed replacing, repairing or recalling, and whether they were fresh or past their best."

Other examples:

- Food Fight: When nine-year-old Martha Payne of Scotland photographed and ranked her school lunches according to quality, quantity, nutritional value and pieces of hair found on her "food-o-meter," the local government council tried to ban her website. Martha's millions of supporters took to Twitter and Facebook forcing the council to reverse its decision. Martha used her newfound fame to raise $185,000 from her followers to provide meals to children in 16 countries.
- Sensing Earthquakes: In Taiwan, students at the Lanyang Girls High School are among the thousands of people in 67 countries participating in the crowdsourced Quake-Catcher Network, which uses the accelerometers in everyday laptops to detect and provide warnings of earthquakes.
- Daily Emissions: The railroad company Union Pacific uses alert systems to measure daily emissions using data comprised of 20 million pattern matches. That's a big dataset!
- Transportation: Rio de Janeiro's government uses big data to improve its regional transportation, natural disaster relief, and population migration.
- Dairy Industry: Even non-high-tech industries are collecting data now. For example, engineers who work for the dairy industry compile data from genomics and productivity statistics to improve the industry's methods and production in coordination with feedback from sensors on each animal.
- Data scientists and engineers compile all the data from wearable fitness trackers to continually revise and revamp their products and their interfaces to be better, plus all the collected health information generates huge databases for analysis by the healthcare industry.

Appendix 18: Student resource sheet 1.13: Data sets from around the world

International context			
The World Factbook (CIA)	https://www.cia.gov/library/publications/resources/the-world-factbook/index.html	World Weather Information Service	worldweather.wmo.int/en/home.html
World Health Organisation	www.who.int/en/	World Weather Online	www.worldweatheronline.com/lang/en-au/
Landmatrix	landmatrix.org/	World Wide Science	worldwidescience.org/
The United Nations Human Development Reports	hdr.undp.org/en/data	Nature International journal of Science	www.nature.com/
The World Economic Forum	www.weforum.org/events/world-economic-forum-on-africa-2019#	UN Sustainable Development Goals	sustainabledevelopment.un.org/
The World Bank	www.doingbusiness.org/en/data	UN Environment	www.unenvironment.org/
OECD Aid Database	www.oecd.org/dac/financing-sustainable-development/development-finance-data/	International Institute for Sustainable Development	www.iisd.org/
UN Data	data.un.org/	National Geographic	www.nationalgeographic.com/environment/

Appendix 19: Student resource sheet 1.14: Australian data sets

Australian context			
Data.gov.au	data.gov.au/	CSIRO	www.csiro.au/
Australian Bureau of Statistics	www.abs.gov.au/	Climate Change in Australia	www.climatechangeinaustralia.gov.au/en/
Australian Data Archive	ada.edu.au/	Australian Geoscience Information Network	www.geoscience.gov.au/
Open Data Department of Prime Minister and Cabinet	www.pmc.gov.au/public-data/open-data	Australian Bureau of Agricultural and Resource Economics and Sciences	www.agriculture.gov.au/abares
The Office of the Australian Information Commissioner	www.oaic.gov.au/	Australian Institute of Marine Science	www.aims.gov.au/
Geoscience Australia	www.ga.gov.au/	Australian Government Science Resources	www.industry.gov.au/about-us/what-we-do/science-in-our-department
GovPond	www.govpond.org/index.php	Sustainability: Cool Australia	www.coolaustralia.org/sustainability/
Bureau of Meteorology	http://www.bom.gov.au/	Sustainability Australia	http://www.sustainabilityindex.com.au/

Appendix 20: Student resource sheet 1.15: United Nations Sustainable Development Goals data sets

United Nations Sustainable Development Goals data sets			
UN Global Pulse Initiative	www.unglobalpulse.org/library/?_types=report	Global Partnerships for Sustainable Development Data	www.data4sdgs.org/
Big Data for Social Good	www.gsma.com/newsroom/press-release/gsma-launches-big-data-for-social-good/	SDG Data Hub	unstats-undesa.opendata.arcgis.com/
The UN World Data Forum	unstats.un.org/unsd/undataforum/index.html	World Bank Innovation for Big Data	www.youtube.com/watch?v=KtTuA-PLrAg
Data for Climate Action	dataforclimateaction.org/	SDG Indicators	unstats.un.org/sdgs/indicators/database/
Data Collaboratives	datacollaboratives.org/explorer.html	Big Data Project Inventory	unstats.un.org/bigdata/inventory/ msutoday.msu.edu/feature/2019/using-big-data-to-solve-the-worlds-biggest-challenges/
A World That Counts	www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf	One World Data	ourworldindata.org/

Appendix 21: Student activity sheet 3.1: Project plan

What is your chosen United Nations Sustainable Development Goal?

_____

_____

What target(s) are you focusing on for your project /initiative?

_____

_____

What data sources are you using to inform your ideas? Give websites and data types (eg diseases, immunisations, salary, food production, locations, etc)

_____

_____

What have you found out about your chosen goal based on the analysis of your chosen data? How did you come to that conclusion?

_____

_____

Describe your initiative / project and what you hope to achieve?

_____

_____

Appendix 22: Student activity sheet 3.2: Data innovation one-pager

<table>
<tr><td colspan="2">Title of innovation:</td></tr>
<tr><td>UN Sustainable Development Goal

Which goal(s) and target(s) are you hoping to address through your innovation.</td><td rowspan="3">Draw a picture here of your innovation or a mind map or flow chart of how it works.

The visual should be an image or graphic that you think helps illuminate your topic, the innovation itself, the data it uses, or your specific example. It does not need to be anything you created by hand, but it can be. A good visual would add depth or additional information to your written text. If necessary, one of the sections should explain the purpose of the visual: what it is depicting? Make sure you also cite the source of the image.</td></tr>
<tr><td>Innovation purpose

What is the intended goal or purpose of this innovation? Why was it created in the first place? What need or problem led to it being created? Assume that the audience doesn't know much about the topic.</td></tr>
<tr><td>Innovation function

How does the innovation work? What data does this innovation use, produce or consume? Where does the data come from? How is it collected? Who is collecting it? If possible, how and where is it stored? Do they provide public access to this data?</td></tr>
<tr><td>Beneficial effect on society, economy, or culture

Think big picture. What group of people will benefit from this innovation? How specifically will they benefit? State clearly if you think this benefit is on society, economy of culture.</td><td>Sources

List all websites that you used to find your information. Include the URL of each.</td></tr>
</table>

Remove the italics and write your own text. Remember it should only be ONE PAGE.

Appendix 23: Teacher resource sheet 4.1: Student evaluation

Key: **1. Sometimes (S)** **2. Consistently (C)** **3. Independently and consistently (I)**	**Student name**												
Remains focused on tasks presented													
Completes set tasks to best of their ability													
Works independently without disrupting others													
Manages time effectively													
Cooperates effectively within the group													
Contributes to group discussions													
Shows respect and consideration for others													
Uses appropriate conflict resolution skills													
Comes to class prepared for activities													
Actively seeks and uses feedback													

Appendix 24: Student activity sheet 4.2: Peer evaluation

	Always	Usually	Sometimes	Rarely
Remains focused on tasks presented				
Completes set tasks to best of their ability				
Works independently without disrupting others				
Uses time well				
Cooperates effectively within the group				
Contributes to group discussions				
Shows respect and consideration for others				
Uses appropriate conflict resolution skills				
Comes to class prepared for activities				
Actively seeks and uses feedback				

Comments:

_____

_____

_____

Appendix 25: Student activity sheet 4.3: Self-evaluation

	Always	Usually	Sometimes	Rarely
Remains focused on tasks presented				
Completes set tasks to best of their ability				
Works independently without disrupting others				
Uses time well				
Cooperates effectively within the group				
Contributes to group discussions				
Shows respect and consideration for others				
Uses appropriate conflict resolution skills				
Comes to class prepared for activities				
Actively seeks and uses feedback				

Comments:

_____

_____

_____

Appendix 26: Teacher resource sheet 4.4: 3 – 2 – 1 Reflection

3 – 2 – 1 Reflection			
Name	**3 things I learnt**	**2 things I found interesting**	**1 thing I found difficult**

Notes