Data Literacy for Educators
Professional Development

Dr Katie Williams

MAICD MAIS CISA

THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA

The University of Queensland (UQ) acknowledges the Traditional Owners and their custodianship of the lands on which we meet.

We pay our respects to their Ancestors and their descendants, who continue cultural and spiritual connections to Country.

We recognise their valuable contributions to Australian and global society.

Building strong relationships between Aboriginal, Torres Strait Islander and non-Indigenous peoples is important to The University of Queensland and its core business of education and research.

Through connections and partnerships, we build and expand the body of knowledge for creating change to advance ideas that benefit the world. In the spirit of reconciliation, we work together to develop and nurture these ideas.

Image: Digital reproduction of *A guidance through time* by Casey Coolwell and Kyra Mancktelow

01

Welcome & Icebreaker

Course Learning Objectives

- Understand the importance of data literacy in Education.

- Learn how to interpret and apply data for better decision making.

- Gain hands-on experience in working with classroom data.

Today's Journey….

Your Facilitator

With you today.

- Senior Lecturer, Business Information Systems (BIS), Business School

- Board Member, Chair / Independent Member, Audit & Risk Management Committees

- Former KPMG Partner specialising in Governance, Risk and Compliance

Introductions

- Name

- Role

- School / Organisation

02

What is Data?

What is Data?

There are many definitions of data, depending on the discipline, domain and context. The word 'data' and 'information' are often used interchangeably.

Data definitions:

- Facts, images or sounds that may or may not be pertinent or useful for a particular task (Alter, 2002).

- Facts or information, especially when examined and used to find out things, or to make decisions. (Hornby, 2015).

- Quantitative or qualitative; stored on analogue or digital media; collected through observation, as part of digital transactions or simply as by-products of digital lives; processed, structured, and analysed to be converted into information; about people, things and systems. (World Bank, 2021)

- Data are measurements or observations that are collected as a source of information. Data can exist in a various types and can be represented in different ways (Australian Bureau of Statistics, 2024).

- Indigenous data refers to information, in any format, that is about Indigenous Peoples, knowledge systems, customs, resources or territories or that impacts Indigenous lives at the collective and/or individual level (Rainie et al. 2019; Lovett et al. 2019).

Economically, data is an asset, yet legally and in accounting practice its no fully recognised.

***Datum* is the singular word in Latin and *data* is the plural form of *datum*.**

Images from Microsoft Windows Stock

Data Types

There are three general data types.

Unstructured: Data that is not pre-formed and exist in free-form formats such as text or numbers without any implemented ordering.

- Examples: social media posts, text documents, audio files

- **In Education**: student essays, classroom video recordings.

Semi-structured: Data that has some structure but does not conform to a strict schema.

- Examples: XML (eXtensible Mark-up Language), HTML (Hypertext Markup Language), JSON files.

- **In Education**: emails between staff and parents (metadata: sender, recipient, timestamp), but free-form message content.

Structured: Data that is organised in a predefined format, making it easier to read and understand by both humans and machines.

- Examples: credit card transactions, spreadsheets, relational databases

- **In Education:** student attendance records, standardised test scores (NAPLAN)

What this Looks Like in Practice

Using the same data set, three different data types look very different.

Unstructured Data

The university has 5600 students. John's ID number is 1, he is 18 years old and already holds a B.Sc. degree. David's ID number is 2, he is 31 years old and holds a Ph.D. degree. Robert's ID is number 3, he is 51 years old and also holds the same degree as David, a Ph.D. degree.

Semi-structured Data

```
<University>
<Student ID="1">
<Name>John</Name>
<Age>18</Age>
<Degree>B.Sc.</Degree>
</Student>
<Student ID="2">
<Name>David</Name>
<Age>31</Age>
<Degree>Ph.D. </Degree>
</Student>
...
</University>
```

Structured Data

ID	Name	Age	Degree
1	John	18	B.Sc.
2	David	31	Ph.D.
3	Robert	51	Ph.D.
4	Rick	26	M.Sc.
5	Michael	19	B.Sc.

Adapted from by J. Cardoso, in A.F. Salam and J.R. Stevens (Eds.). (2007). *"Developing Dynamic Packaging Applications Using Semantic Web-Based Integration".*
Semantic Web Technologies and E-Business: Toward the Integrated Virtual Organization and Business Process Automation (p.11. Idea Group Publishing.

03

Data Literacy

What is Data Literacy?

Definitions are important.

- Being able to read, work with, analyse, and communicate with data (Bland, 2025).

- The knowledge, skills and dispositions to use data ethically for decisions about learning and wellbeing (Queensland Government, 2025). 🔗

- **Data literacy for teaching** is the ability to transform information into actionable instructional knowledge and practices by collecting, analysing, and interpreting all types of data (assessment, school climate, behavioural, snapshot, longitudinal, moment-to-moment, etc.).

  - It combines an understanding of data with standards, disciplinary knowledge and practices, curricular knowledge, pedagogical content knowledge, and an understanding of how children learn (Gummer & Mandinach, 2015, p. 2).

Why is *Data Literacy* Important?

Why are we talking about?

- Foundational skill for navigating modern life.

- A skill set that educators must acquire and nurture throughout their careers (Mandinach & Gummer, 2016).

- Data skills support core curriculum outcomes (numeracy, digital literacy, critical thinking), as well as future workforce needs.

- Increasing skills in data literacy enables evidence-based decisions.

  - Improved student performance.

  - Better school outcomes.

- Supports equity for diverse learners by enabling educators to identify learning gaps across socio-cultural groups through monitoring performance trends by student demographic (Datnow & Hubbard, 2016).

- Elevates awareness of ethical and responsible data use (privacy, transparency and fairness in the handling of data).

Conceptual Framework for Data Literacy

Integrating content and pedagogy with data literacy (Gummer & Mandinach, 2015).

Three primary interacting domains:

1) **Discipline specific content knowledge**

2) **Data use for teaching**

3) **Pedagogical content knowledge**

Additional aspects include the ethical and responsible use of data, and considerations of the privacy of students (Gummer & Mandinach, 2015).

The domain of data use for teaching includes the components of the Inquiry process.

Beneath the inquiry cycle are subcomponents of the inquiry process, each comprising specific knowledge and skills.

Figure 1. Organization of data literacy conceptual framework
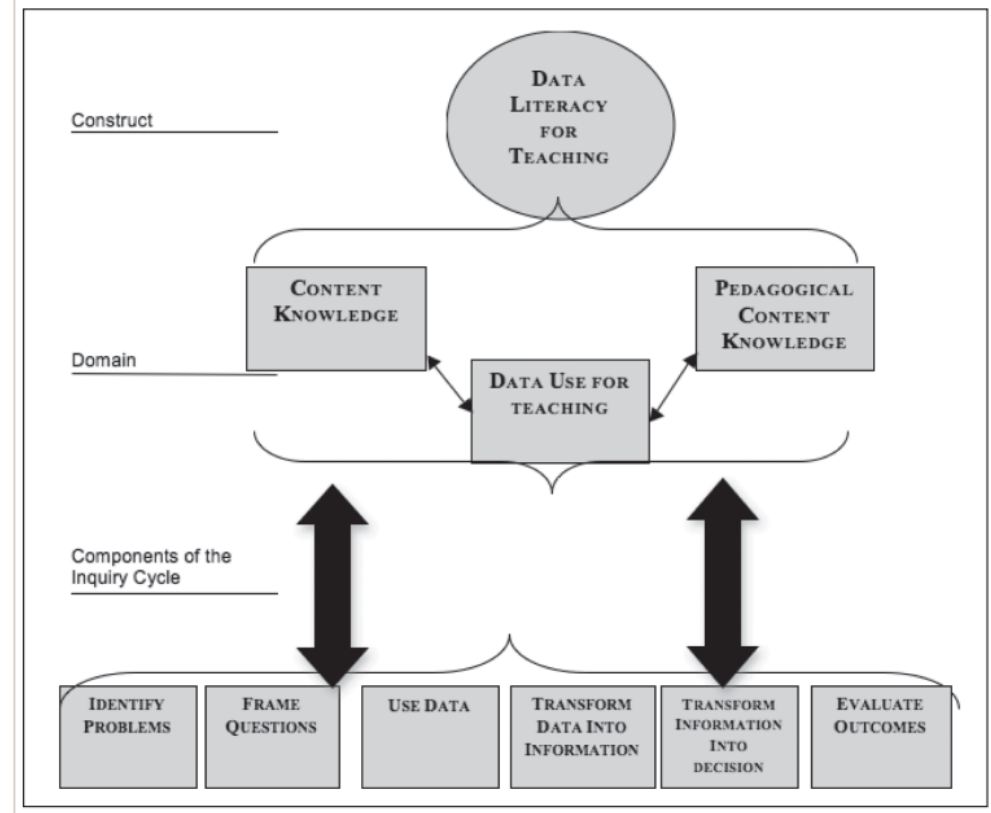
Moving Beyond Content and Pedagogy (Mandinach & Gummer, 2016).

Revised conceptual framework integrates elements of Schulman's (1987) knowledge base for teaching.

- Emphasis on elements that had clear connection to data use:

  - **Knowledge of learners and their characteristics** (connects with the inclusion of multiple forms of data beyond student performance).

  - **Knowledge of educational contexts** (classroom, school, community, cultures).

  - **Knowledge of educational ends, purposes, and values** (awareness of broader educational goals, why we teach).

  - **Curriculum knowledge** and **General pedagogical knowledge.**

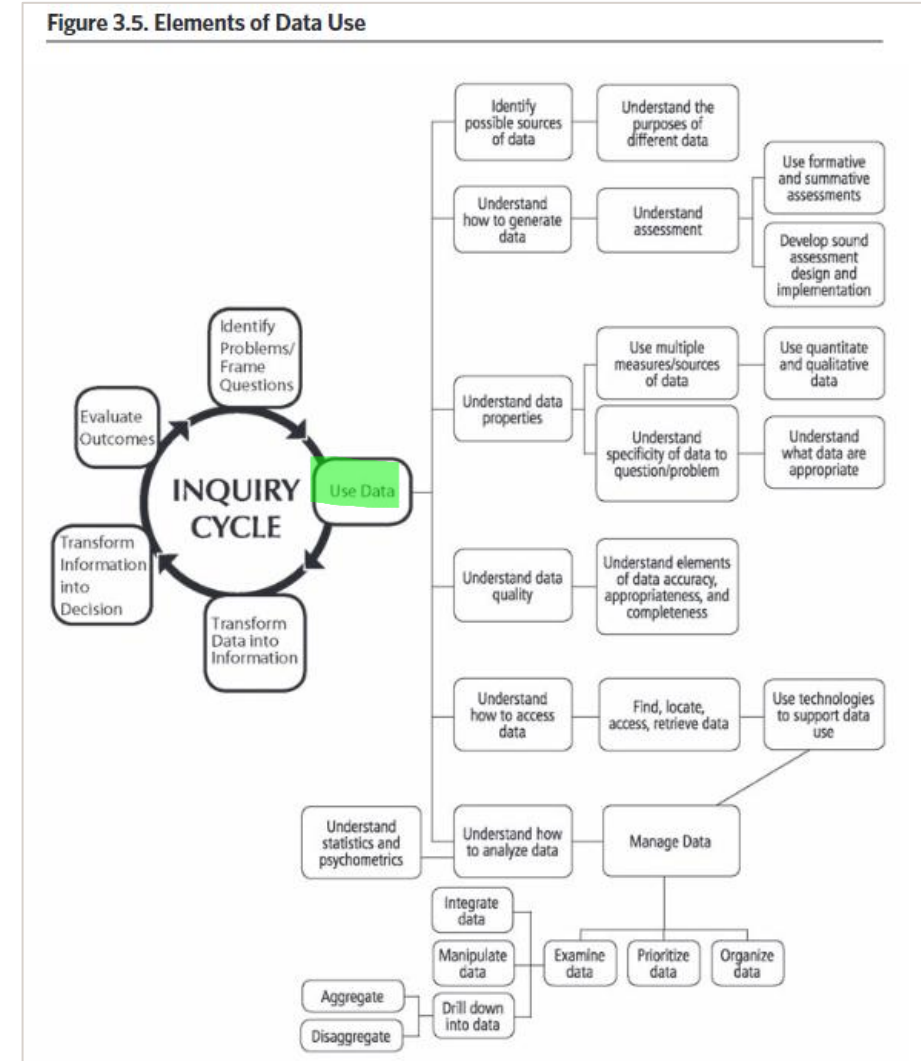**Figure 3.3. Revised Conceptual Framework for Data Literacy for Teachers**
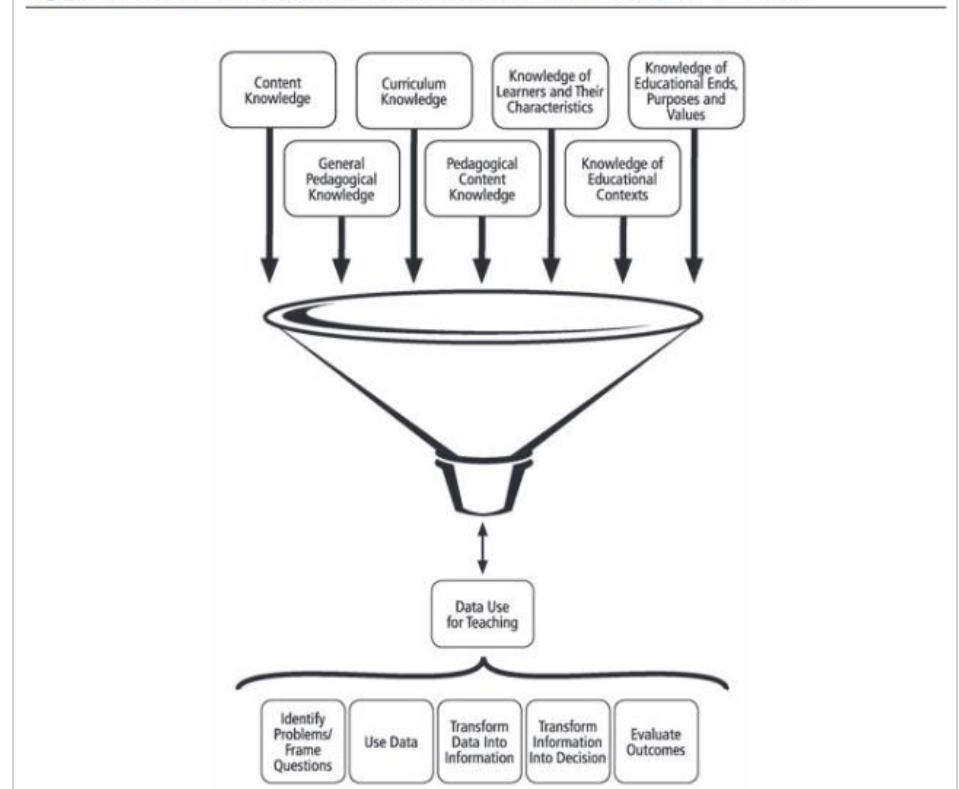
MANDINACH, E. B. & GUMMER, E. S. 2016. *Data literacy for educators : making it count in teacher preparation and practice, New York, N.Y, Teachers College Press. Page 47*

Domain of Data Use for Teaching

Revised conceptual framework (Mandinach & Gummer, 2016).

- The **Inquiry Cycle** is a process of moving from the identification of a teaching problem to refining that into a question (or set of questions) that can be examined empirically.

- As the teaching problem / issue is defined, teachers must identify and use data that might illuminate and refine what the problems are.

- The cycle includes the need to interpret and draw insights from the data to transform the data into information.

- Teachers then must transform information into decisions to identify and implement a potential solution to the problem or issue.

- The final component of the cycle is to evaluate the outcomes of the inquiry cycle work when the solution is implemented.
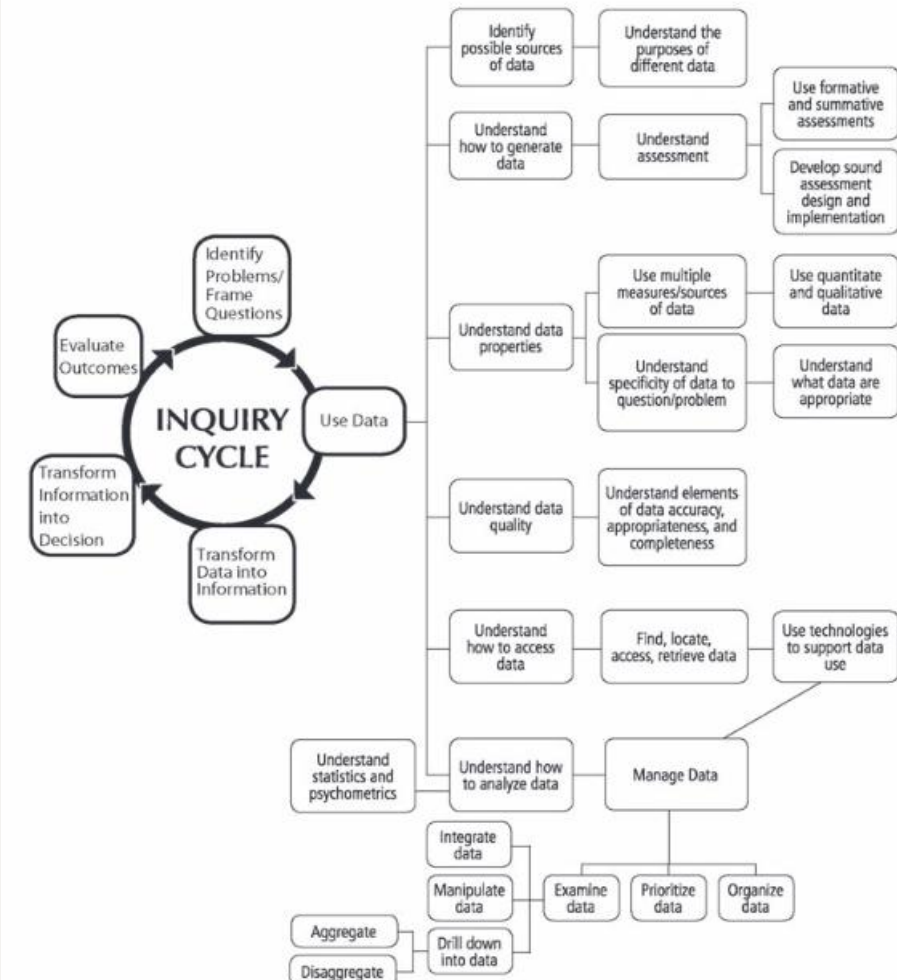

Figure 3.4. Domain of Data Use for Teaching
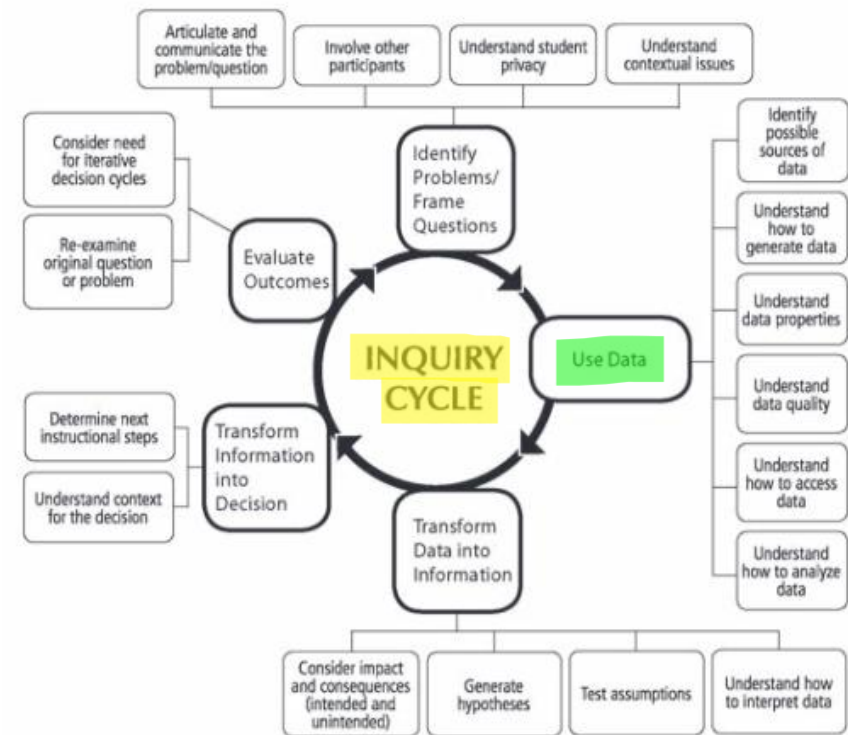
MANDINACH, E. B. & GUMMER, E. S. 2016. *Data literacy for educators : making it count in teacher preparation and practice, New York, N.Y, Teachers College Press. Page 48*

Elements of Data Use

Revised conceptual framework (Mandinach & Gummer, 2016).

- This figure displays the multiple subcomponents, associated elements and sub-elements that define the complex component of **Use Data.**

- To identify possible sources of data, there is a need to know that different data have different purposes; what data is needed to answer the question?

- Not all necessary and useful data will be immediately available. There is a need to understand assessment and identify and develop fair, reliable, and valid assessments.

- There is a need to understand data properties to determine what is appropriate and valid data; use of both quantitative and qualitative data to move beyond a focus on numerical student assessment results.

Figure 3.5. Elements of Data Use

MANDINACH, E. B. & GUMMER, E. S. 2016. *Data literacy for educators : making it count in teacher preparation and practice, New York, N.Y, Teachers College Press. Page 51*

Elements of Data Use

Revised conceptual framework (Mandinach & Gummer, 2016).

- Understanding data quality involves different levels of knowledge and skills to determine data accuracy, appropriateness, and completeness.

- Teachers must understand how to access and retrieve data, often using disparate systems.

- Understanding how to **analyse data** is key to effective data use.

- Teachers must know how to manage data, especially as data sources and systems become more complex.

- Key activities focus on organising and preparing data for analysis, as well as examining the data.

Figure 3.5. Elements of Data Use

MANDINACH, E. B. & GUMMER, E. S. 2016. *Data literacy for educators : making it count in teacher preparation and practice, New York, N.Y, Teachers College Press. Page 51*

Choosing an Appropriate Visual Display

A logical taxonomy of charts and what purposes they support.

- At a high level, the taxonomy divides the purpose into four different purposes: comparisons, compositions, relationships, and distributions.

- It divides the branches into subcategories based on the number of variables involved and time dependency of the visualisation.

- There are many choices in building visualisations.

- A general rule of thumb is to choose the simplest chart, as users will likely have a hard time following it.

Sharda, R., Delen, D., Turban, E., & King, D. (2018). Business intelligence, analytics, and data science (4th ed.). Pearson Education, Limited.
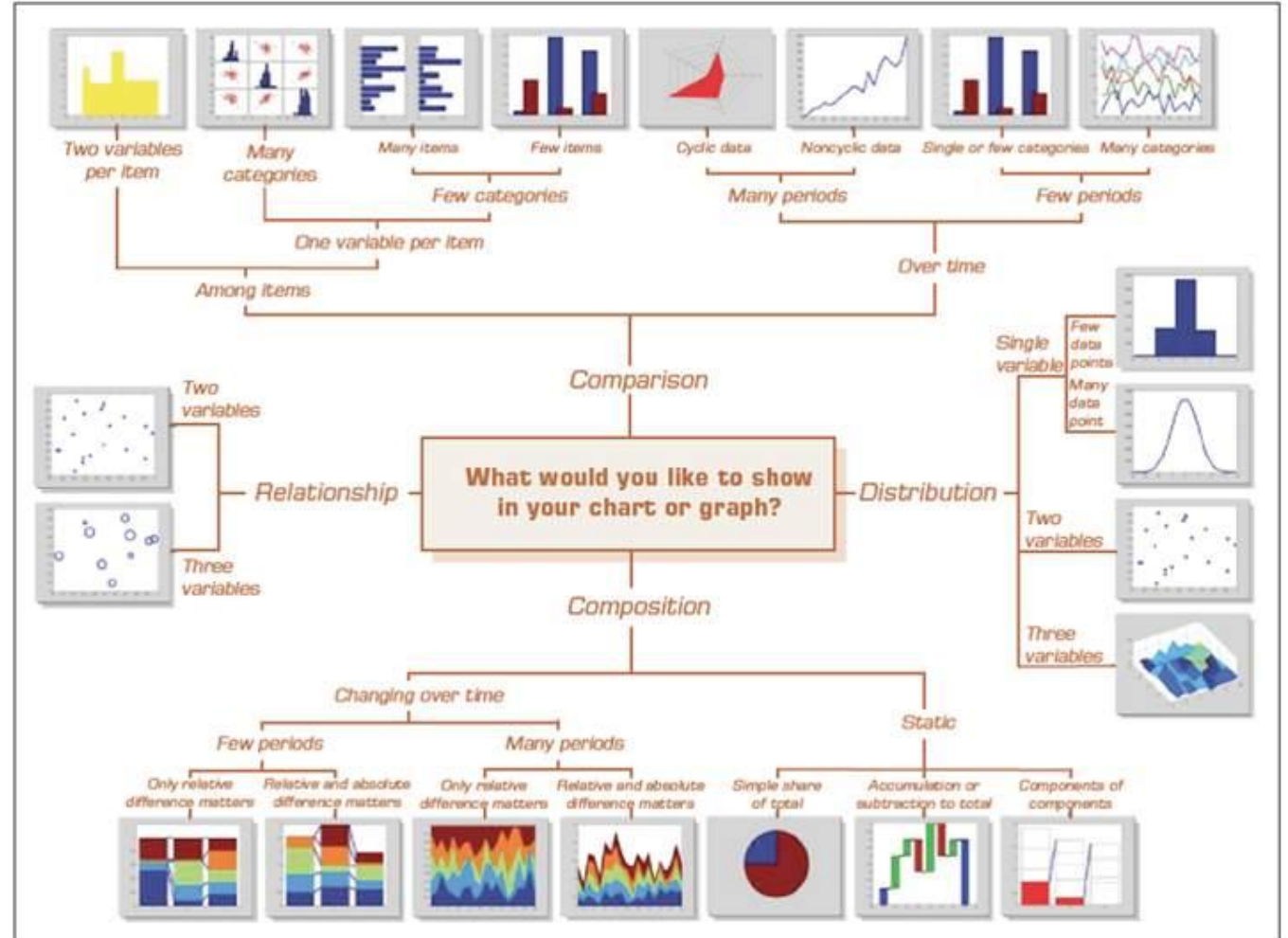
04

Data Quality

The Quality of Data

Data quality is the degree to which a particular data set meets a users' requirements. Common dimensions of data quality include:

Completeness

- All required data is present, with no missing records or fields.
- **Example**: A student's profile in the school database includes name, date of birth, and parent contact details. Medical information is not captured.

Accuracy

- The degree to which data correctly reflects real-world object or event.
- **Example**: A student's grade is recorded as an A in the system, but the marked paper shows a C.

Timeliness

- The extent to which data is up to date and available when needed.
- **Example**: Attendance data is updated weekly instead of daily.

Validity

- The degree to which data conforms to defined format, rules and business requirements.
- **Example:** A student's date of birth is entered as 00/13/2026 (invalid).

Reliability

- The consistency of the data across sources and time.
- **Example:** NAPLAN results reported in the school dashboard match those in the national database.

Uniqueness

- The degree to which each record is distinct, with no duplication.
- **Example:** A student is entered twice in the system with different first (i.e. Jackie Smith and Jacquelin Smith).

The Data Dictionary

Why do data dictionaries matter?

- A **data dictionary** is a centralised repository that defines and standardises the meaning, format, and usage of data elements within a system, or an organisation.

- Data dictionaries can be used to ensure that all users and stakeholders interpret data terms consistently – this supports both data quality, data governance, and communication.

- Within this course, let's consider a data dictionary as a decision enabler:

  - Avoids disputes over '*whose numbers are right*'?

  - Measures and KPIs are consistently defined.

  - Facilitates trust in dashboards and reports.

Exploratory Data Analysis

One way of understanding the quality of your data is to use Exploratory Data Analysis (EDA).

The goal of EDA is to identify quality issues and understand what data is, rather than to discover patterns and correlations in data.

A basic form of EDA involves opening data in a text editor and looking at individual records and values. Statistical methods could then be used to explore issues more systematically.

Through EDA data quality issues can be identified, such as:

- Duplicates

- Missing or incomplete fields

- Inconsistent formats or measurement units

- Inaccurate data

It is common to face data quality and inconsistency challenges **when combining multiple datasets** that have different schemas or are populated with different rules.

Example of Data Quality Issues

The demonstration uses a real dataset with errors to showcase typical data quality problems.

Accuracy

Uniqueness

Accuracy

CaseNumber	Sentence	Registration	Category	firstName	middleName	lastName
006A00005	Guilty	10/31/2011	Baltimore	Eugene	Mark	Smith
009G14291	Guilty		Incarcerated	Craig		Williams
011D04321	Guilty	09/15/2011	Out-lining Jurisdiction	Dontay		Purnell
012CO1639	guilty	11/26/2012	Baltimore	Howard	Nicholas	Motley
044J12261	Guilty	6/01/11	Out-lining Jurisdiction	Frank	Joseph	Blank
044J12261	Guilty		Incarcerated	Kenneth	Wayne	Locus
1.08E+224	Guilty		Incarcerated	Michael		Mott
01K04873	60	10/24/2012	Baltimore	Joseph	Lee	Griffin
	Guilty		Incarcerated	Markie		Cole
	Guilty	1/02/13	Baltim	Kali	Alexander	Moulton

Completeness

Validity

Sadiq, S. and Indulska, M (2017). Open data: Quality over quantity, International Journal of Information Management, 37(3), p. 151

Activity

Fairplay Vouchers

The Fairplay Program, open to Queensland children aged between 5-17 years of age provides a $150 voucher for children to use towards their sport and active recreation membership.

Tasks

1) Download file "Voucher_Data_extract.xls" and explore the dataset. Alternatively, review a hardcopy of the data.

2) Review the supporting Data Dictionary for the data set.

3) Identify **10 data quality issues**. Highlight these issues using

Time Allocation

• Identify data quality issues = 20 minutes.

• Class debrief = 5 minutes

Entire open data set available at www.data.qld.gov.au (41700 rows).

Data Challenge

Student Academic Achievement – Open Data

Tasks

1) Download file "Academic Achievement Data.xlsx".

2) Use the following questions to help explore the dataset.

Time Allocation

- Explore data set and dictionary = 5 minutes

- Identify at least 2 data insights = 25 minutes

- Create simple visuals to communicate these insights = 10 minutes

- Debrief = 5-10 minutes

1. How many students are in each year level?

2. What is the most common achievement result?

3. How many different languages (MLOTE) are represented in the data set?

4. What proportion of students in Year 4 achieve a grade level of B?

5. How does achievement distribution change as students progress from Year 1 to Year 10?

6. In which year levels are the highest and lowest achievement results concentrated?

Morning
Tea

Some Insights

C

Year	Total
1	96
2	8
3	17
4	52
5	27
6	6
7	9
8	8
9	7
10	6
11	8
12	1
PY	63
Grand Total	308

Year 4 Achievement Results

Result: C, B, D, A, N
No. of Students: 0, 5, 10, 15, 20, 25

DISTRIBUTION OF LANGUAGES

English, 89.61%
Pidgin, nfd, 8.44%
Welsh, 0.97%
Irish, 0.97%

100.00%, 80.00%, 60.00%, 40.00%, 20.00%, 0.00%
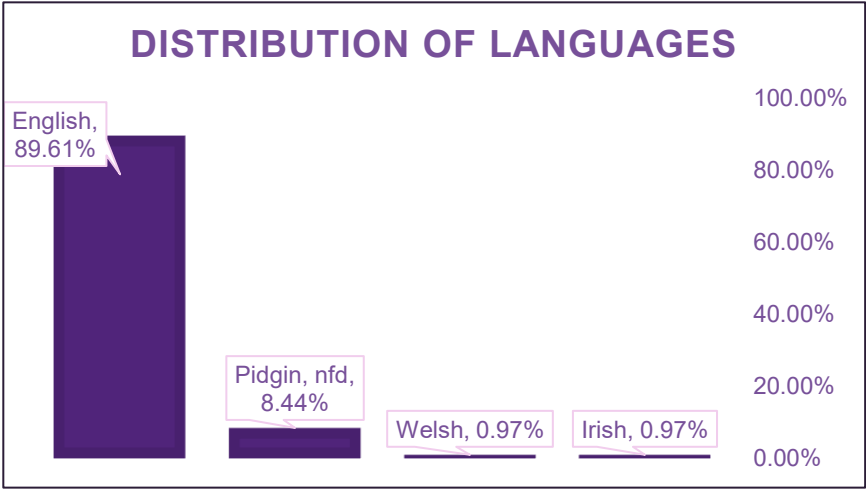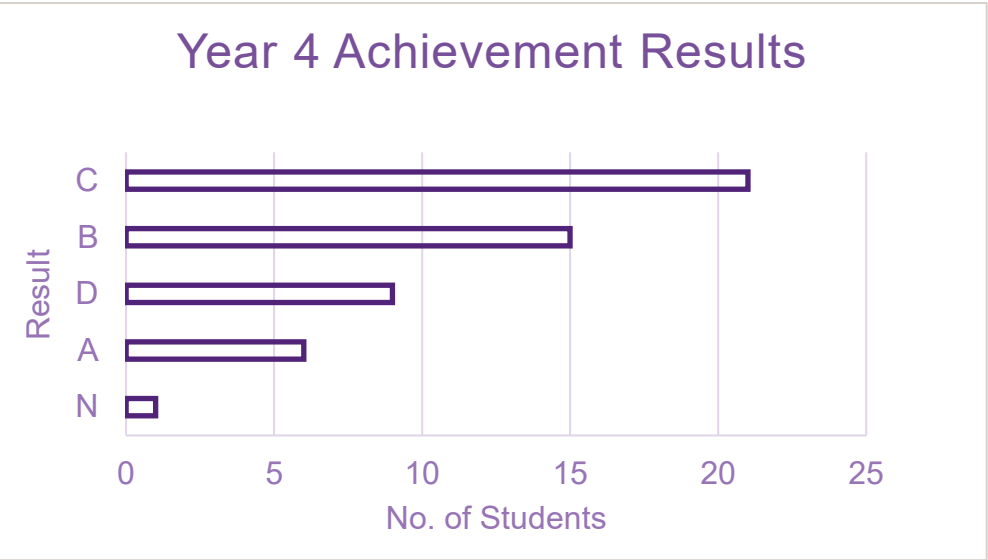
05

Data Tools

Which Data Tools Do You Use to Inform Your Teaching / Practice?
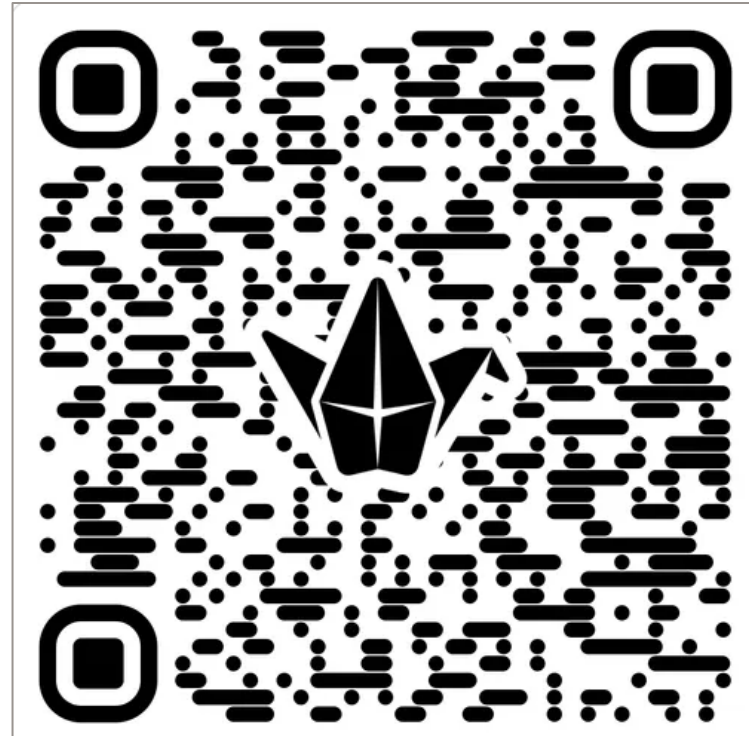
[Padlet](#)

Google Forms
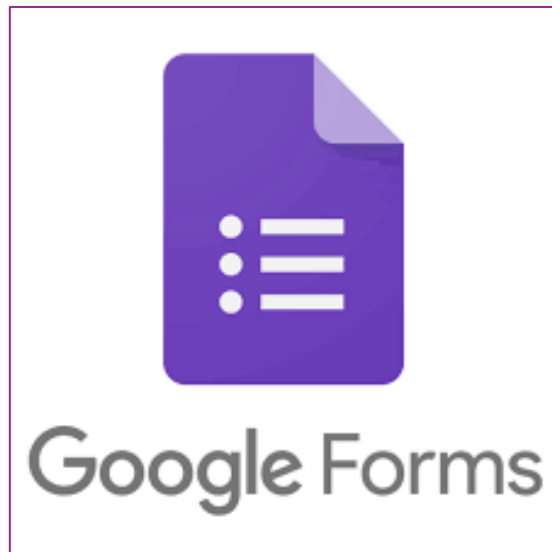
- Google Forms can be used to collect, organise and analyse data.

- In the classroom, typical applications include collecting assignment data, creating quizzes, and gathering feedback.

- Offers various question types, customisation options, and real-time data analysis capabilities.

- Data collected through Google Forms can be exported for further analysis, aiding in data-driven decision-making.

Open your Google Form

Start with responses.

Begin by selecting the Google Form which contains responses.

- Note: if starting with a new form, you must disseminate the form link and allow sufficient time for users to complete.

- Ensure the form is properly configured to collect the required data and respondent information effectively.

- Verify that the form's access permissions are appropriately set to gather submissions from the target audience.
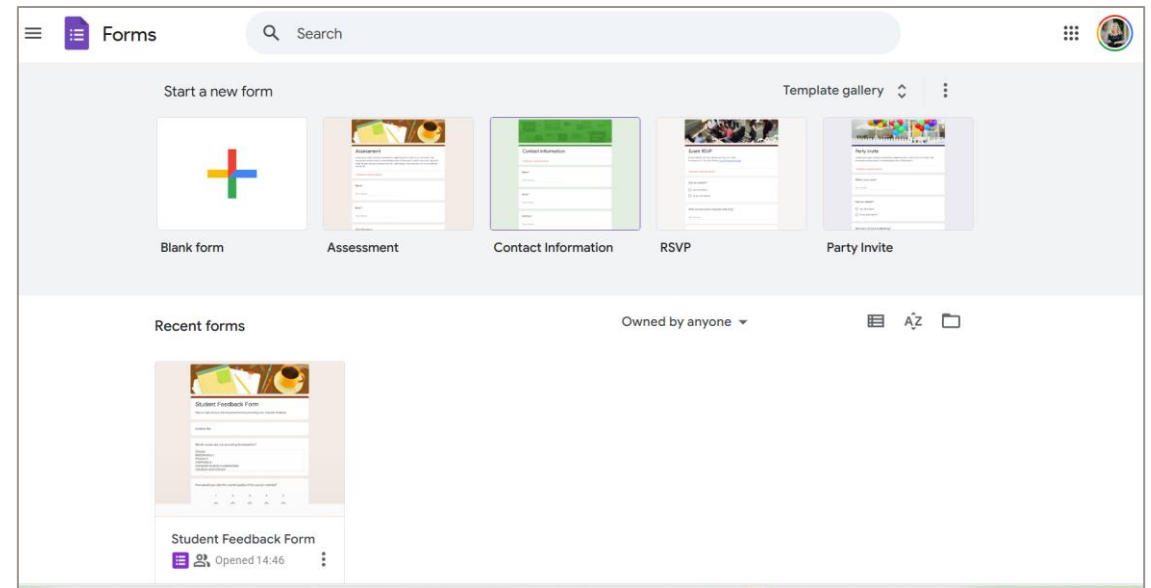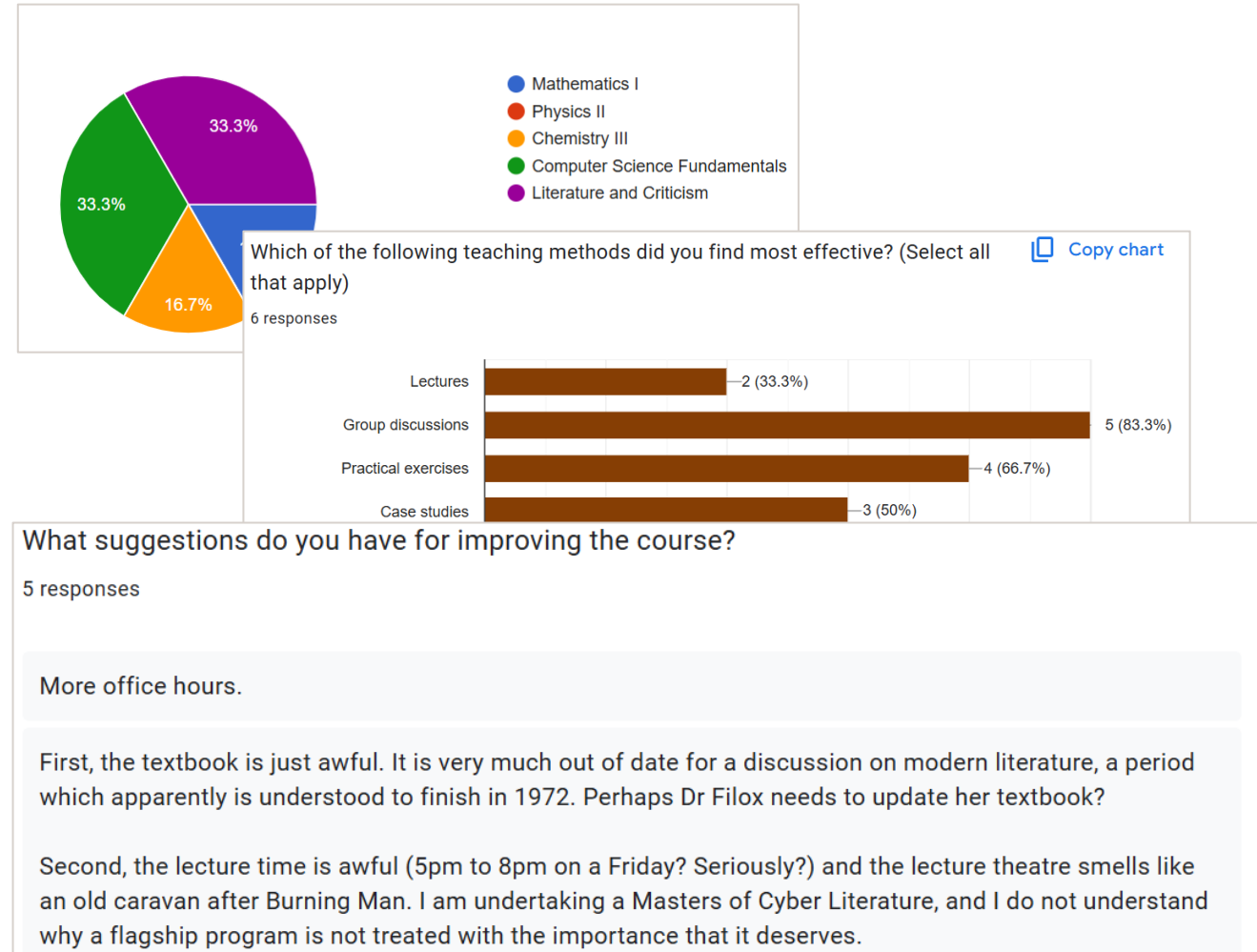
Access the Responses

Navigating Data

- Within the created Google Form, locate and click on the "Responses" tab – this serves as the central hub where you'll find consolidated response data, ready for analysis.

- Familiarise yourself with the layout of the "Responses" tab, which provides multiple options for data exploration.

- *Note: the tab dynamically updates, reflecting responses as they are submitted, offering a real-time view.*
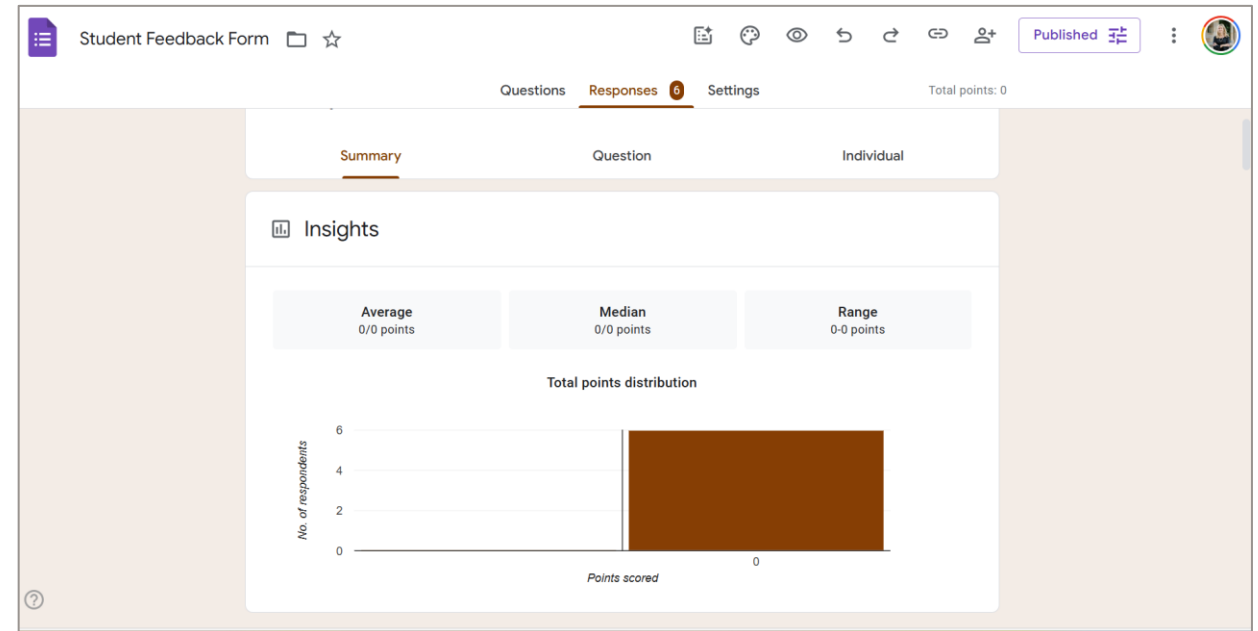
Explore Summary & Individual Responses

Understanding Trends

- Review the automatically generated response summary for initial overview insights which includes visual representations such as charts and graphs to highlight response patterns.

- Identify frequently selected choices or response trends to quickly grasp dominant opinions from your survey data.

- Analyse the summary visualizations to reveal overall sentiment from survey respondents about the questions asked.

- Switch to the 'Individual' view to examine each respondent's answers for a more granular level of analysis.

Pie chart with legend:
- Mathematics I
- Physics II
- Chemistry III
- Computer Science Fundamentals
- Literature and Criticism

33.3%, 33.3%, 16.7%, 33.3%

Which of the following teaching methods did you find most effective? (Select all that apply)   Copy chart
6 responses

- Lectures — 2 (33.3%)
- Group discussions — 5 (83.3%)
- Practical exercises — 4 (66.7%)
- Case studies — 3 (50%)

What suggestions do you have for improving the course?
5 responses

More office hours.

First, the textbook is just awful. It is very much out of date for a discussion on modern literature, a period which apparently is understood to finish in 1972. Perhaps Dr Filox needs to update her textbook?

Second, the lecture time is awful (5pm to 8pm on a Friday? Seriously?) and the lecture theatre smells like an old caravan after Burning Man. I am undertaking a Masters of Cyber Literature, and I do not understand why a flagship program is not treated with the importance that it deserves.

Data Tools

- NotebookLM can be used to help analyse a variety of data types, uploaded by the user.

- The AI works within the context of your own content. It can provide educators with tools to explore a vast amount of information, take notes, ask questions, and organise information efficiently.

- NotebookLM uses Google's Gemini model to generate flashcards, quizzes, and reports for enhanced learning.

- Linkage with Google Docs, allowing users to upload documents and synthesise information effectively and quickly.

Getting Started with Notebook LM

- Navigate to the NotebookLM website and log in using your Google account.

- You will then be prompted to create your first notebook – simply click on "Create".

- 5 different options for how to upload documents.
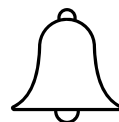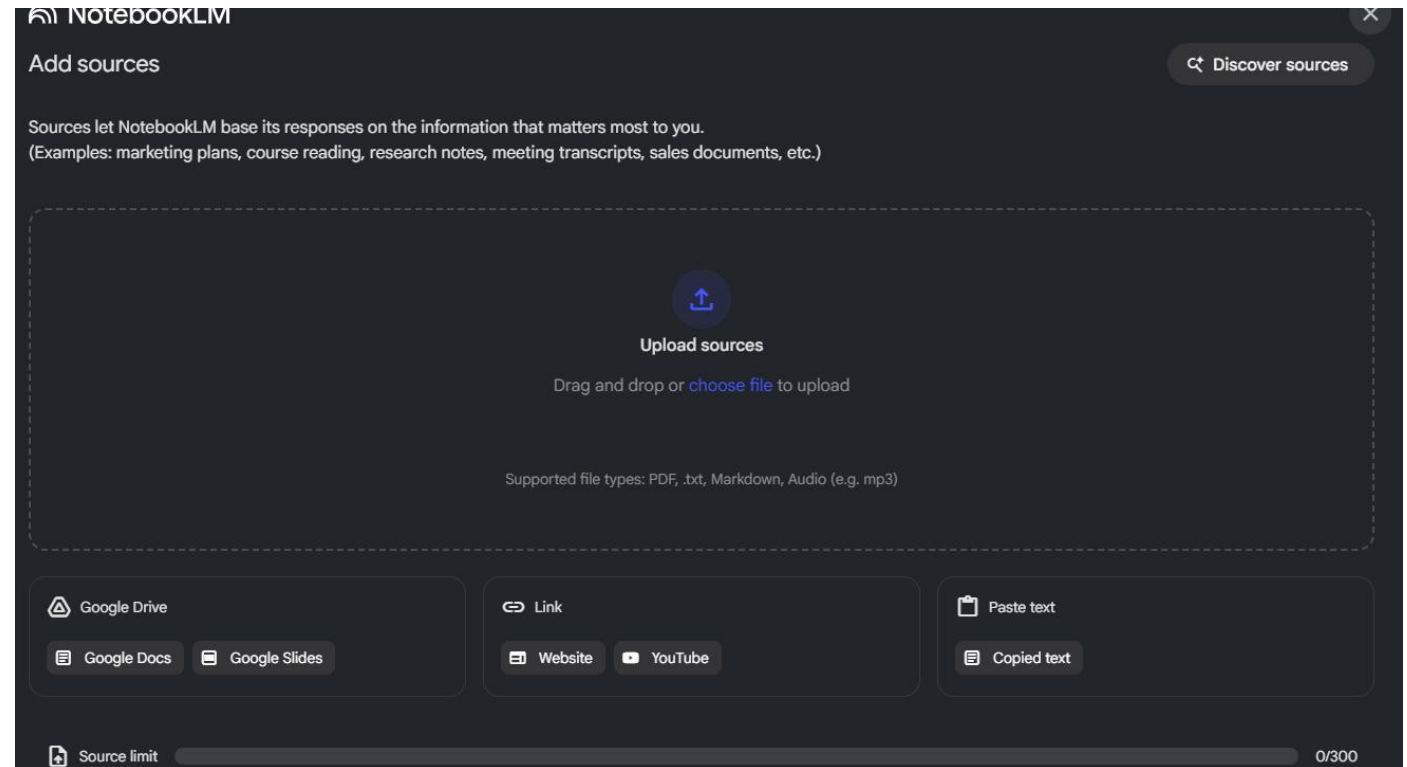
AI Insights NotebookLM for School Data Analysis

The NotebookLM Dashboard

The Notebook Guide is a dashboard which contains:
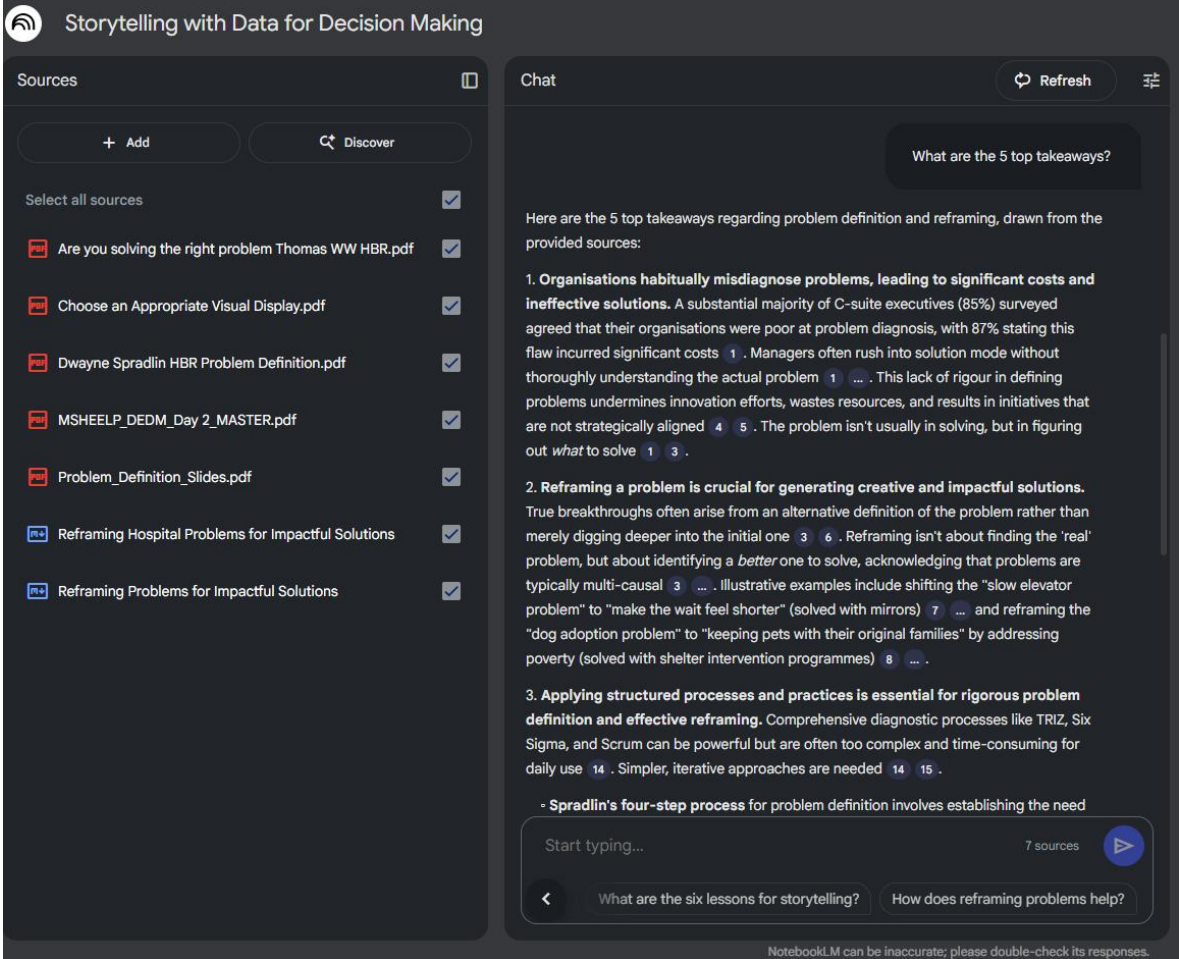
- An automatically generated summary of your uploaded documents

- Suggested questions you may want to ask that are tailored to your documents

- Buttons to create an audio overview, a video overview, mind maps and reports.
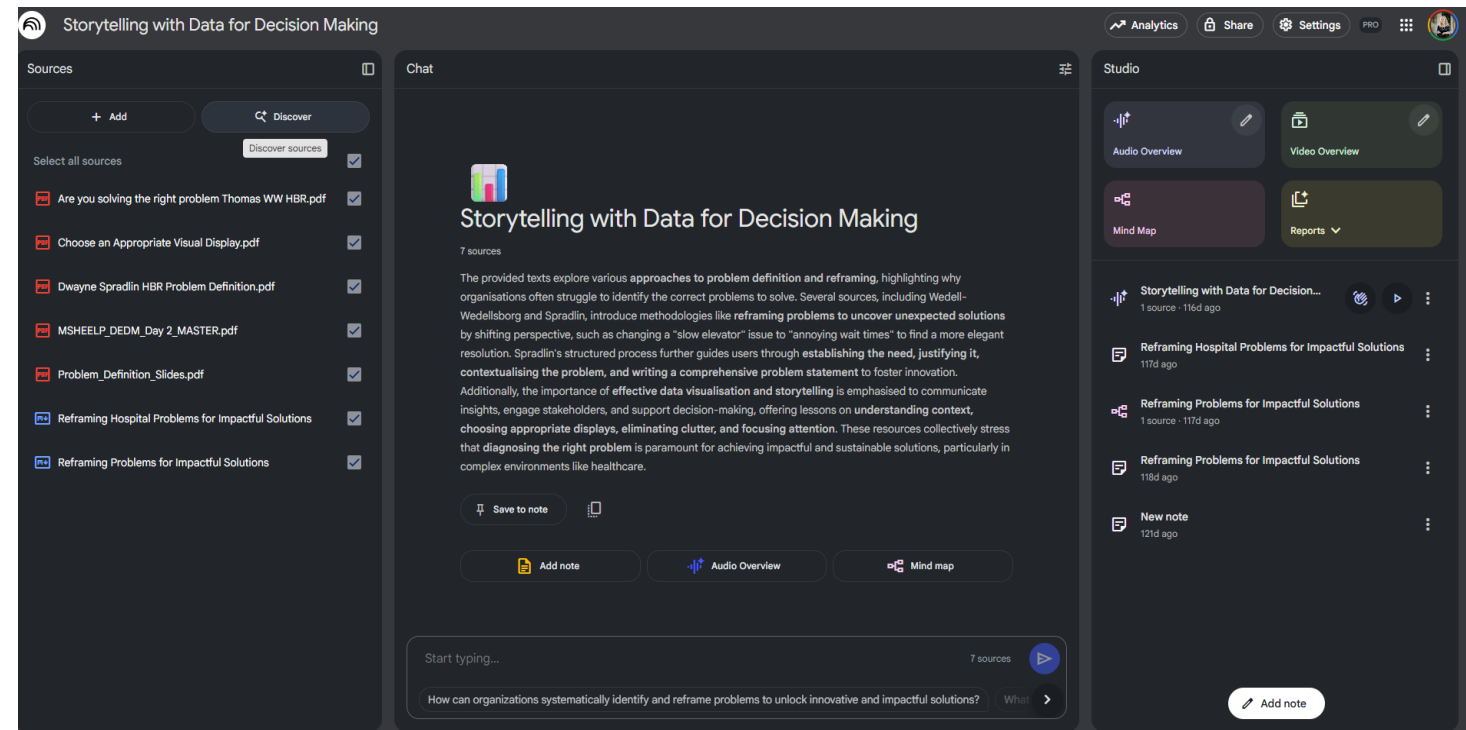
Asking Questions

NoteBookLM allows you to ask questions about the documents / sources you have uploaded.

- Click one of the "suggested questions" or start typing your own questions in the bottom of the page.

- In addition to answering the question, NotebookLM provides specific citations showing you where it got its answer from.

- These are shown as numbers in grey ovals next to the information. When you hover or click on the number, the specific location in the document appears.

Generating Ideas & Notes

Users can ask the AI to generate ideas in the chat page which can be saved as a 'Note'.

- There are two types of notes: written and AI generated. Written notes are notes you type in yourself. To create a new note, click "Add note" at the top of the notes page. This will generate a new, blank note for you.

- These functions enable users to quickly and easily consolidate thoughts into one organized area. *Note: any unsaved chat disappears when you leave the notebook.*
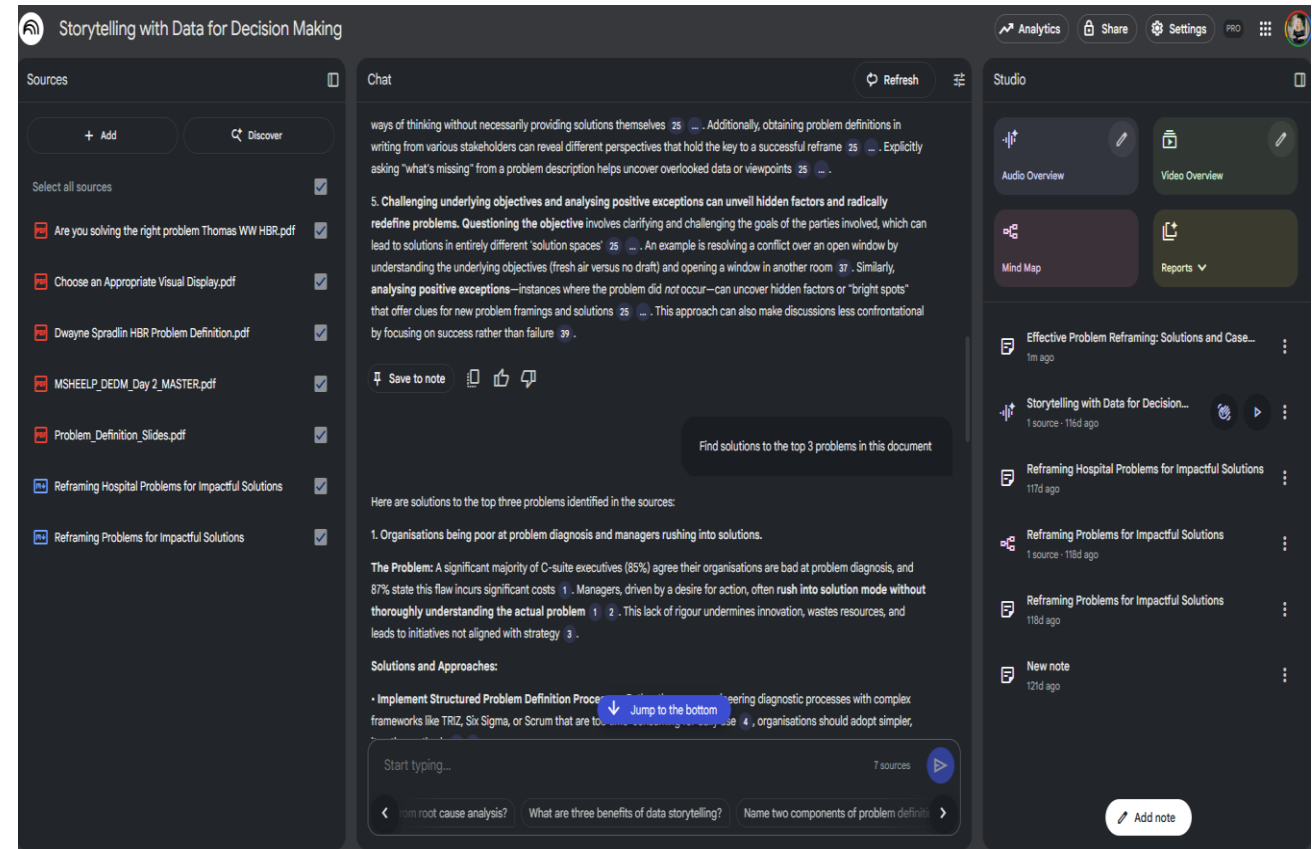
Podcast Generation with Notebook LM

Users can create customised podcasts with two AI voice discussing your data sources.

- As of September 2025, there are 4 different podcast formats.

- Users can choose multiple languages and one of three different podcast lengths.

- Additionally, customisation can be provided via prompting.

Lunch Break

Activity

Image licensed through Adobe Stock: 414170288
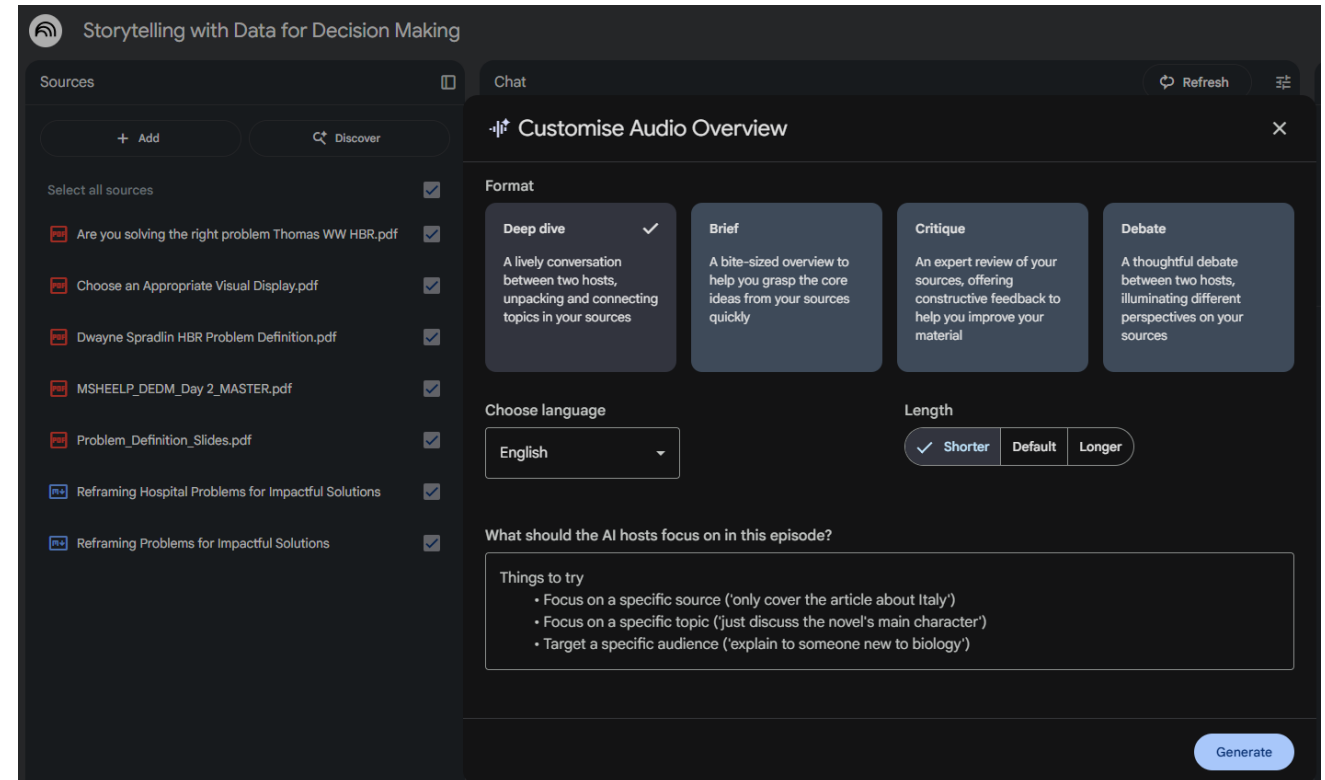
Hands-On Exploration

Work through the following tasks and explore NotebookLM's capabilities.

Tasks

1) Open NotebookLM and sign in with your Google Account.

2) Create a new notebook and title it "My Teaching Assistant".

3) Upload pre-prepared source materials.

- Using the documents as a source, ask NotebookLM to write a summary explaining the importance of **Science** within the Year 4 Curriculum.
- Review the response. What do you notice? Add a note capturing your initial observations.
- Discover / add 4 other relevant sources and upload these into your new notebook.
- Use the chat prompts to discover other insights across the sources.
- Use the Mindmap function across selected sources. Explore the connections.
- Explore other Notebook functions.

Time Allocation

- 30 mins

Hands-On Exploration

Work through a series of data quests to explore NotebookLM's capabilities.

Debrief

1) How can you see yourself using NotebookLM in your work?

2) What are the potential benefits and challenges of using an AI tool like this?

3) How could this tool help with data organisation and decision-making in your classroom?

Time Allocation

• Pair discussion: 10 minutes

• Class share: 10 minutes

Student Attendance and Performance

In this activity we will explore the relationship between student attendance and growth in reading from Years 3 to 5.

Activity Tasks

1) Download the dataset "Y5_Attendance_NAPLAN_Dataset.xlsx".

2) Complete the following

- Check the data dictionary

- Sort the dataset by *Attendance_%* (highest to lowest).

- Look at *Growth_Reading* alongside attendance.

- What do you observe?

3) Insert a new column: *Attendance_Band*, use these categories:

- Excellent = 95%+

- Good = 90-94%

- At risk = <90%

Fill the column by assigning each student to a band.

4) Calculate the average *Growth_reading* for each Attendance band. Use AVERAGEIF or create a PivotTable.

5) Explore the data, create simple charts to assist with analysis.

Reflection Questions

- What patterns do you see between attendance and growth?

- Is there a point (threshold) where low attendance leads to a clear drop in growth?

- Are there any exceptions?

- What other data points would be useful to include in this dataset, why?

Time Allocation

- 30 mins

Where to From Here?

A moment for reflection.

Course Learning Objectives

- **Understand the importance of data literacy in Education.**

- **Learn how to interpret and apply data for better decision making.**

- **Gain hands-on experience in working with classroom data.**

Start with a small project:

- Select a manageable data set. Use GoogleSheets to organise and visualise the data.

- Explore how GoogleForms could be used to support a deeper understanding of student learning preferences.

Collaborate with colleagues:

- Form a small group / community and discuss how you are using data in the classroom. Share insights, challenges, successes.

Try NotebookLM for deeper analysis:

- Use NoteBookLM to analyse school-related documents. Create notes and summaries within your notebook to help make connections between different data points. Identify patterns or outliers that can inform your practice.

Contact

Dr. Katie Williams

Senior Lecturer

Katie.Williams@uq.edu.au
+61 418 160 460

References

R1   Alter, S. (2002). Information systems : foundation of e-business (4th ed ed.). Prentice Hall. //doi.org/10.1201/b17511

R2   Hornby, A. S. (2015). Oxford advanced learner's dictionary of current English (9th ed ed.). Oxford University Press.

R3   World Bank. (2021). World Development Report 2021: Data for Better Lives. Washington, DC: World Bank.

R4   Gummer, E. S. & Mandinach, E. B. (2015). Building a Conceptual Framework for Data Literacy. *Teachers College record (1970), 117, 1-22.*

R5   ABS (2024, June 4). *Weighted average of 8 capital cities*. Monthly Consumer Price Index Indicator, April 2024 | Australian Bureau of Statistics (abs.gov.au)

R6   Cardoso, J. (2007) Developing Dynamic Packaging Applications Using Semantic Web-Based Integration", in A.F. Salam and J.R. Stevens (Eds.), Semantic Web Technologies and E-Business: Toward the Integrated Virtual Organization and Business Process Automation (p.11), 2007, Idea Group Publishing.

R7   Bland, S. (2025). Data Literacy Is an Essential Skill. Let's Teach It That Way. *EducationWeek*

R8   Mandinach, E. B. & Gummer, E. S. (2016). Every teacher should succeed with data literacy. *Phi Delta Kappan, 97, 43-46.*

R9   Datnow, A. & Hubbard, I. (2016). Teacher capacity for and beliefs about data-driven decision making: A literature review of international research. *Journal of educational change*, 17, 7-28

R10   Mandinach, E. B., & Gummer, E. S. (2016). Data literacy for educators : making it count in teacher preparation and practice (1st ed.). Teachers College Press.

R11   Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. Harvard educational review, 57, 1-22.

R12   Carroll, S. R., Garba, I., Figeroa-Rodriguez, O. L., Holbrook, J., Lovett, R., Materchera, S., Parsons, M., Raseroka, K., Rodriguez-Loneber, D., Rowe, R., Sara, R., Walker, J. D., Anderson, J. & Hudson, M. (2020). The CARE Principles for Indigenous Data Governance. Data Science Journal.

R13   Rainie, S. C., Kukutai, T., Walter, M., Figueroa-Rodríguez, O. L., Walker, J., & Axelsson, P. (2019). Indigenous data sovereignty. In T. Davies, S. Walker, M. Rubinstein, & F. Perini (Eds.), The state of open data: Histories and horizons (pp. 300-319). African Minds. https://doi.org/10.5281/zenodo.2677801

R14   Sharda, R., Delen, D., Turban, E., & King, D. (2018). Business intelligence, analytics, and data science (4th ed.). Pearson Education, Limited.