

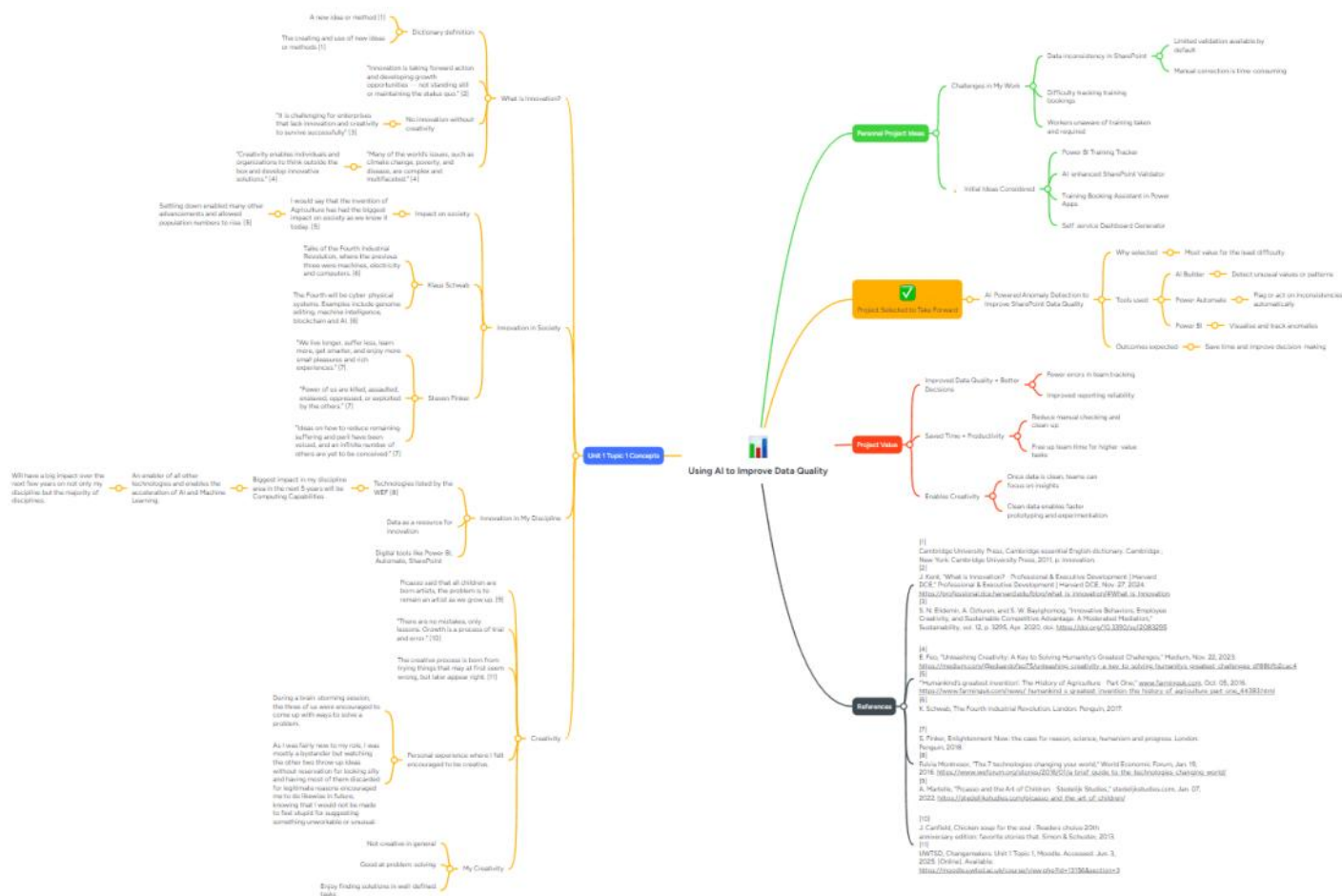
AI-Powered Anomaly Detection to Improve SharePoint Data Quality

Paul Griffiths

2410226

Computing (Data and Information Systems)

Ideation & Mind Map



Clearer version of this Mindmap can be accessed at the link below.

<https://mm.tt/map/3726405536?t=xkVXynTYwz>

Aim & Objectives and Research Area

Project Aim and Objectives

Aim:

To design and evaluate an AI-powered anomaly detection solution using Power Platform tools that improves the quality of data in SharePoint Lists used for reporting.

Objectives:

1. Investigate common data integrity and inconsistency issues in SharePoint Lists that negatively impact reporting accuracy and decision-making
2. Explore the use of Microsoft's AI Builder Prediction Model to identify and classify anomalous patterns within business data
3. Design and build a Power Automate flow to automatically flag or act upon anomalies in SharePoint data entries detected in real-time
4. Use Power BI to develop an interactive dashboard that visualises anomalies and highlights trends over time
5. Evaluate the effectiveness of the solution through collecting and analysing user feedback

The overall goal is to deliver a low-code, intelligent system that enhances data reliability in organisational reporting while reducing the manual effort required to identify data quality issues.

Research Philosophy

The most suitable research philosophy for this project is pragmatism as it prioritises real-world solutions over abstract ideas. Pragmatism is especially suited to projects like this one, where the priority is to create a solution that solves a genuine business problem - in this case, poor data quality in SharePoint Lists used for reporting. [1]

At the same time, my personal preference leans toward positivism, which views reality as objective and measurable. From this viewpoint, data errors and anomalies can be defined, counted and corrected. Cleaner data can be directly correlated with improved business intelligence outputs. [1]

I will therefore take a blended approach, combining these philosophies:

- Pragmatism allows me flexibility in method selection. I can combine:
 - Quantitative data such as anomaly frequency and error rates
 - Qualitative feedback such as user experience feedback
- Positivist tendencies will guide my data analysis - assuming that clean data leads to objectively better outcomes.

By combining both philosophies, the hope is to balance data with how people actually experience the system, making sure it's not just effective on paper but also useful in real life.

Search Tools and Strategies

To build a solid foundation for the project, I will carry out a structured literature and technical review using both academic and practical

sources. After some study [2], I have decided that for searching I will use:

- **Google Scholar** - for discovering academic papers [3]
- **Microsoft Learn** - for up-to-date platform-specific documentation [4]
- **ResearchGate** - to follow researchers working in related fields [5]
- **IEEE Xplore** - for articles on anomaly detection and AI [6]
- **ACM Digital Library** - for research on low-code systems and applied AI [7]

To ensure relevant and focused search results, I will use a combination of Boolean logic, keywords and filters to narrow down my search.

Examples include:

- "AI anomaly detection" AND SharePoint AND "Power Platform"
- "Low-code data validation" OR "automated data quality"
- "Business intelligence" AND "Power BI" AND accuracy NOT finance
- "AI Builder" AND "Power Automate" AND prediction
- SharePoint AND "data integrity" OR "reporting accuracy"

Each source will be assessed for relevance using the CRAAP test [8]:

- **Currency** - Look for the most recent (past 5 years) sources due to fast-moving tech
- **Relevance** - Prioritise studies and articles that focus on practical application in business contexts
- **Authority** - Prefer peer-reviewed journals, Microsoft docs or recognised experts
- **Accuracy** - Cross-check data with trusted platforms
- **Purpose** - Avoid commercial content or overly theoretical work without practical examples

This methodical approach will ensure that the research is thorough, up-to-date and satisfies both academic standards and technical needs.

References

[1]

D. Jansen, "Research Philosophy & Paradigms: Positivism, Interpretivism & Pragmatism," *Grad Coach*, Jun. 15, 2023.

<https://gradcoach.com/research-philosophy> (accessed Jun. 17, 2025).

[2]

"List of Top 10 Search Engines for Research Papers [2024] [Updated]," *Medium*, Jan. 22, 2024. <https://medium.com/@scholars-edge/list-of-top-10-search-engines-for-research-papers-2024-updated-752634a7adb6> (accessed Jun. 17, 2025).

[3]

Google Scholar, 2025. <https://scholar.google.com/>

[4]

Microsoft, "Training," *learn.microsoft.com*, 2024. <https://learn.microsoft.com/en-us/training/>

[5]

ResearchGate, 2025. <https://www.researchgate.net/search>

[6]

IEEE, "IEEE Xplore Digital Library," *ieee.org*, 2025. <https://ieeexplore.ieee.org/Xplore/home.jsp>

[7]

Acm, "ACM Digital Library," *Acm.org*, 2025. <https://dl.acm.org/>

Research Strategies

Validity, Reliability, Generalisability, Ethics

Validity in research refers to how accurately a method measures what it is intended to measure. [1] For this project, validity will be ensured by testing the anomaly detection solution on real-world data. This will involve cross-checking results through manual inspection. Any flagged entries will be reviewed by team members familiar with the data.

Reliability refers to consistency of results. [2] This will be assessed by running the anomaly detection software multiple times on static datasets and over time on live datasets. A consistent identification of anomalies in static data will show good reliability. To improve reliability, the detection logic will be rule-based, with parameters that can be reviewed and adjusted as needed.

Generalisability refers to whether any findings can apply to contexts beyond the study. [3] While this anomaly detection will first be applied to a single SharePoint List, the approach is being developed with scalability in mind. The components used (AI Builder, Power Automate, Power BI) are standardised and can be applied to other datasets with only minor changes.

Ethics refers to a set of moral principles and guidelines that govern how researchers should conduct their work. [4] This project will only use data from internal SharePoint Lists with authorised access. No personal or sensitive data will be included in the analysis. Also, all feedback from stakeholders will be anonymous and participation in surveys or interviews will be voluntary.

Data Collection Methods & Pilot

To gather evidence for evaluating the solution's performance and value, a combination of primary data and qualitative feedback will be used. [5]

Primary data will include:

- Logs of anomalies collected before and after implementation
- Metrics such as number of anomalies and error reduction percentage
- Stakeholder feedback through short survey or interview to measure perceived usefulness, accuracy and potential time-saving

A pilot version of the solution will be implemented using a SharePoint List used by a small team. The pilot will:

- Apply AI Builder to detect outliers or inconsistent entries
- Use Power Automate to highlight anomalies through notifications
- Log the outcomes for analysis

Pilot feedback will be collected after stakeholders have used the tool for at least one week. The original group of users will be 5 team members. They will be asked whether it improved their workflow and reduced time spent manually correcting errors. This feedback will help improve the solution.

Quantitative & Qualitative Analysis

This project will use both quantitative and qualitative data to assess outcomes.

Quantitative analysis will include:

- Descriptive statistics: Number and type of anomalies before and after implementation, false positive rates, frequency of data errors
- Inferential statistics: Comparison of average anomaly rate before and after implementation
- Visual analysis: Power BI will display these metrics in appropriate charts to enable identification of trends and patterns over time

These quantitative measures will be complemented by qualitative ones.

Qualitative analysis will include:

- Responses to open-ended survey questions or interviews. These will be coded and thematic analysis applied [6]

- Emerging themes may include “time-saving” and “data integrity” and will be coded and grouped using Microsoft Excel
 - Insights from these themes will be used to guide improvements or further development down the line
- Combining quantitative metrics and qualitative feedback ensures that there is a clearer picture - not just how the system performs on paper, but how helpful it actually feels to the people using it. [7]

References

[1]

C. M. Patino and J. C. Ferreira, “Internal and External validity: Can You Apply Research Study Results to Your patients?,” *Jornal Brasileiro De Pneumologia*, vol. 44, no. 3, p. 183, 2018, doi: <https://doi.org/10.1590/S1806-37562018000000164>.

[2]

P. C. Price, “Reliability and Validity of Measurement: Research Methods in Psychology,” *Opentextbc.ca*, 2018. <https://opentextbc.ca/researchmethods/chapter/reliability-and-validity-of-measurement/> (accessed Jul. 01, 2025).

[3]

R. Bermingham, “Interpreting research evidence,” *UK Parliament*, Sep. 2020, Accessed: Jul. 01, 2025. [Online]. Available: <https://post.parliament.uk/interpreting-research-evidence/>

[4]

D. Resnik, “What is ethics in research & why is it important?,” *National Institute of Environmental Health Sciences*, Dec. 23, 2020. <https://www.niehs.nih.gov/research/resources/bioethics/whatis> (accessed Jul. 01, 2025).

[5]

A. Ahmed, L. Pereira, and K. Jane, “Mixed Methods Research: Combining Both Qualitative and Quantitative Approaches,” *ResearchGate*, 2024. https://www.researchgate.net/publication/384402328_Mixed_Methods_Research_Combining_both_qualitative_and_quantitative_approaches (accessed Jul. 01, 2025).

[6]

B. Williams, “Market Research Sustainability: 5 Eco-Friendly Practices - Insight7 - AI Tool For Interview Analysis & Market Research,” *Insight7 - AI Tool for Interview Analysis & Market Research*, Jul. 2024, doi: <https://doi.org/10962523357/hYhvCNfjiqkZEN2pq-so>.

[7]

J. Brereton, “Qualitative vs Quantitative Metrics: A Comprehensive Comparison,” www.launchnotes.com, Sep. 01, 2023. <https://www.launchnotes.com/blog/qualitative-vs-quantitative-metrics-a-comprehensive-comparison> (accessed Jul. 01, 2025).

Appendix

Unit 1 Topic 1

According to the Cambridge Dictionary, “innovation is a new idea or method, or the use of new ideas and methods.”

5 things that were not around 10 years ago:

- Chat GPT
- Remote Working being normalised
- Google Assistant
- Google Wallet
- Food Delivery Apps

These have all changed my life for the better in terms of making day-to-day things a little easier but none of them have made a major difference.

I would say that the invention of Agriculture has had the biggest impact on society as we know it today. This is because without agriculture, humans would still be spread out in tribes, unable to settle down in one place. Settling down enabled many other advancements but also allowed population numbers to rise.

In general, innovation has changed things for the better. Deaths from natural disasters has halved over the last century while the percentage of people living in extreme poverty has also halved. People are living longer and have more time and money to enjoy their lives.

Collective wellbeing should be our goal and we can use science and reason to achieve it.

Stephen Pinker wrote “We live longer, suffer less, learn more, get smarter, and enjoy more small pleasures and rich experiences. Fewer of us are killed, assaulted, enslaved, oppressed, or exploited by the others. Much suffering remains, and tremendous peril. But ideas on how to reduce them have been voiced, and an infinite number of others are yet to be conceived.”

I still believe that Agriculture has had the biggest impact on the development of society. This has had many benefits – civilisation as we know it depends upon this. However, it can be argued that it has forced humans to live in a way that we are not suited to and that we were happier before.

Klaus Schwab talks of the Fourth Industrial Revolution, where the previous three were machines, electricity and computers. The Fourth will be cyber-physical systems. Examples include genome editing, machine intelligence, blockchain and AI.

Of the technologies listed by the WEF, I think the one that will have the biggest impact in my discipline area in the next 5 years will be Computing Capabilities. This is because not only is an enabler of all other technologies but it also enables the acceleration of AI and Machine Learning which I think will have a big impact over the next few years on not only my discipline but the majority of disciplines.

Nikolai Kondratieff studied the impact of innovation on society and produced a chart referred to as Kondratieff Waves or Cycles. Certain Innovations cause spikes of prosperity in society before returning to normal.

Creativity is important as it is innovation, not replication, that will give us a competitive edge. There is no innovation without creativity. Creativity is essential for solving the social and environmental issues in the world around us which will only get worse as it increases in complexity.

We have an innate ability to be creative, but we lose that as we get older. Picasso said that all children are born artists, the problem is to remain an artist as we grow up. We need to be aware that it is wrong to believe that mistakes are the worst things you can do. The creative process is born from trying things that may at first seem wrong, but later appear right.

I do not consider myself to be a creative person under most circumstances as I struggle to think outside the box. However, I think I am actually quite good at thinking up solutions to clearly defined problems.

I had an experience where I felt encouraged to be creative. During a brain-storming session in work, the three of us were encouraged to come up with ways to solve a problem that we had, no matter how outside the box it was. As I was fairly new to my role, I was mostly a bystander but watching the other two throw up ideas without reservation for looking silly and having most of them discarded for legitimate reasons encouraged me to do likewise in future, knowing that I would not be made to feel stupid for suggesting something unworkable or unusual.

Creativity is important in my subject area as we regularly have situations where all we have is a goal and how we get there is up to us. For most tasks, there are many possible ways to go so having some creativity to explore all options for benefits and drawbacks is vital.



✓ **You have completed Section 3!**

After completing this section, you should

- Think of yourself as someone capable of being creative (if you don't already)
- Identify personal passions as sources for idea generation.

Back

Continue →

Unit 1 Topic 2

How do YOU solve a problem?

The importance of creativity

The last time I felt I was being creative was writing for this assignment. I feel at my most creative when putting thoughts into words in a way that other people can understand. It makes me feel like I'm creating order out of chaos, clarity out of obfuscation. It feels satisfying.

Being creative is good for your mental wellbeing

In a study by BBC Arts and UCL, 50,000 people participated in a creativity test. The research showed that there are three main ways we use creativity as coping mechanisms to control our emotions:

- Distraction – using creativity to avoid stress
- Contemplation – using creativity to allow us to reassess problems and make plans
- Self-development – to face challenges by building up self-esteem and confidence

Nothing beats taking part in live creative activities that involve face-to-face social interaction.

Regardless of skill level

New pursuits is more beneficial than doing one for more than ten years as activities can become less effective at regulating negative emotions

When we face hardships, creative activities are particularly beneficial for our emotions

Lower psychological wellbeing significantly impacts on student learning and experience while good wellbeing can lead to better learning and performance

Research suggests adults who make time for play have greater productivity, more creativity and greater health and happiness

This is why learning games are becoming a popular learning strategy

Creative confidence is having the ability to think freely, without fear of shame, or fear of judgement, or any fear.

According to the founders of Ideo, "creativity is something you practice, not just a talent you're born with."

There are four fears that hold most of us back from being creative:

- Fear of the messy unknown
- Fear of being judged
- Fear of the first step
- Fear of losing control

Recognising how your philosophy impacts your investigation

Research philosophy is your belief about the ways in which you will be collecting, analysing and using data in your project

Philosophy of research is split into three broad forms:

- Ontology – the study of being – researcher assumes the knowledge on what exists has already been created, but then asks what the world must be like with this assumption. Split into two branches – objective (reality is controlling the person) and subjective (person controls reality)
- Epistemology – the study of knowledge – split into three – foundationalism (all knowledge is based on a few statements that cannot be doubted), coherentism (all knowledge is true so long as it does not contradict itself) and pragmatism (if it works, it's true and will get you by for the time being)
- Axiology – the researcher is trying to find what is valuable and why – split into two – ethics (moral principles and dilemmas) and aesthetics (sensory emotional value in science, critical reflection on art, culture and nature in social science and branches of art in art)

Each approach has pros and cons

Positivism is led by science. Assumption that there is only one truth that can be gained by observing, collecting data and interpreting strictly

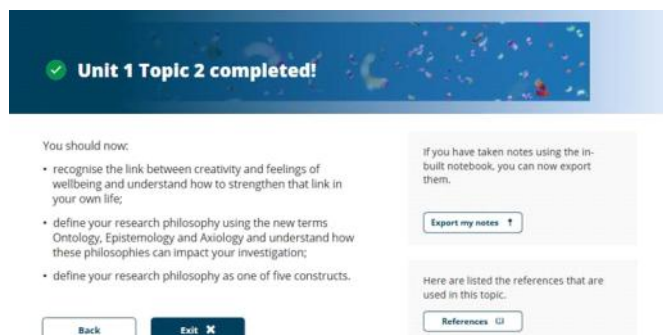
Critical realism is similar but does not require the restrictions of science to be proven.

Interpretivism assumes that access to reality is only through social constructions. You cannot use science to understand people

Postmodernism establishes established thinking by exposing ideas embedded in the way we use language to explain the reality in which we exist

Pragmatism only considers concepts relevant if they support an action

I would say that I feel most drawn to Positivism as I work with data a lot and like to think that there is only one truth that the data leads towards. With enough data, anything can be understood



The screenshot shows a digital interface with a blue header bar at the top containing a green checkmark icon and the text "Unit 1 Topic 2 completed!". Below the header, the page is divided into two main columns. The left column, titled "You should now:", contains a bulleted list of three tasks: recognizing the link between creativity and feelings of wellbeing, defining research philosophy using new terms, and defining research philosophy as one of five constructs. At the bottom of this column are "Back" and "Exit" buttons. The right column contains two sections: "If you have taken notes using the in-built notebook, you can now export them." with an "Export my notes" button, and "Here are listed the references that are used in this topic." with a "References" button.

Unit 1 Topic 3

Known Knowns

Selecting an appropriate database

Four categories, known knowns, known unknowns, unknown knowns and unknown unknowns

Unknown unknown is something we are neither aware of nor completely understand

In research, it is prudent to be aware of what is known and what is not known

This can be explored using library databases

A database is any digital resource of information

Most published work uses between 4 and 8 databases

The CRAAP test evaluates the credibility and relevance of information found:

- Currency – is it recent & has it been superseded?
- Relevance
- Authority – is it credible?
- Accuracy
- Purpose – what is it written for?

Search engines have no quality control and is disorganised

Google scholar can be used

Using Boolean search strategies for identifying research

Keywords are important words that can be used to target our search

Speech marks indicate multiple words are one search term

Asterisk indicates the search should be for all forms of the word

Booleans (And, Not, Or) can be used to refine the search

Resources found can improve future searches

Analysis paralysis can be dispelled by using a concept map

A concept map is a chart that is a visual approach to organisation, like a mind map

A concept map starts with more discipline from the top of the page and works down

It introduces structure and form

Done to expose concepts and propositions

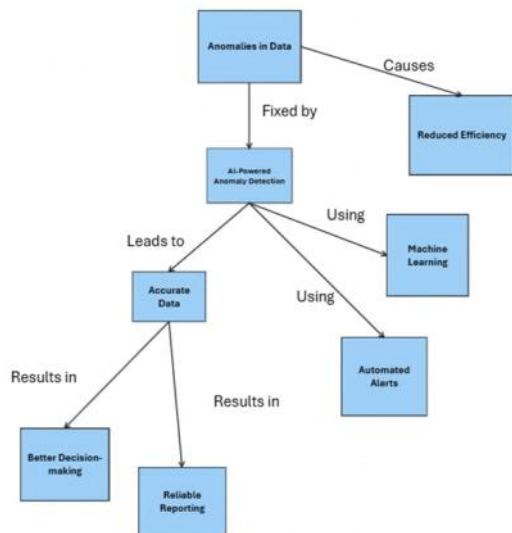
A proposition is an academic term used to describe how two concepts are joined

Can use Microsoft Whiteboard, PowerPoint, Word, Google Docs, Slides, etc

As simple a diagram as possible that organises thoughts

One page and three units max

Concepts in shapes, relationships are arrows and text on arrows explains the relationship



A focus question to outline what is going to be in or out of scope for my research will be ""How can AI-driven anomaly detection improve the accuracy and reliability of data in SharePoint Lists used for reporting?"

Unit 4 Topic 1 completed!

You should now:

- Be able to imply a numerical precision to arguments in a casual way that brings a lack of precision to our work.
- Be able to analyse data using different examples of graphs and charts and present data using a variety of methods.
- Have a deeper understanding of creating charts that communicate in the best way possible.

Back

Exit

If you have taken notes using the in-built notebook, you can now export them.

Export my notes

Here are listed the references that are used in this topic.

References



You have completed Section 3!

Upon completion of this sub-section, you should be able to create your own concept map and come up with concepts that you can use for your literature search.

Back

Continue

Unit 2 Topic 2

Appendix Page 13

Surveys are a great way of collecting data. Can be Quantitative or Qualitative.

Individual Interviews can extract beliefs, values and attitudes.

Focus Groups involve around 7 people discussing a problem. This allows for a range of opinions.

Observation can be done qualitatively or quantitatively.

Secondary data can be used – from databases or websites.

The Hawthorne Effect shows that people perform better while under observation. This can bias the data collected.

For my project, the best way of collecting data would be to do observations on the SharePoint List to get Qualitative data about which errors are most likely to occur. This can then be paired with either surveys or individual interviews to assess the quality of the dataset before and after the solution is implemented. I do not believe that the Hawthorne Effect will be relevant here.

Quantitative data can be limited due to time and resources, limited results or limited replies. This needs to be accounted for.

Picking people for surveys has a number of approaches. However, as only a very limited number of people will be using the dataset, they can all be surveyed or interviewed.

Causal Relationship is an experiment – broadly manipulating a variable and measuring its effect.

For my research project, the data collection will be:

- Type – Causal Relationship
- Sample Size – All users of Dataset
- Method – Classification of errors followed by either surveying or interviewing users after solution has been in place for some time

Unit 2 Topic 2 completed!

You should now:

- Be familiar with different methods for collecting and analysing qualitative data including data sampling.
- Be able to consider how the acquisition of your data will impact your research.
- Understand your approach to collecting data for your research.

[Back](#) [Exit](#)

If you have taken notes using the in-built notebook, you can now export them.

[Export my notes](#)

Here are listed the references that are used in this topic.

[References](#)

Unit 4 Topic 1

Statistics are ways we can try to reveal patterns and trends.

When searching for “Quantitative Research of...” in Google, I get results for education, medicine, stem students and business.

With research statistics, the more accurate, the better.

There are two sorts of statistics – those used to describe our argument (Descriptive) and those used to make inferences from our argument (Inferential).

Descriptive – to turn data into a source of information to help describe the research topic. Can be numbers, charts, tables or graphs. Great way to take a snapshot of a data set and produce a summary that can be more easily consumed.

These statistics could include averages or frequencies.

Inferential – to help us conclude what the data is saying. Inference is necessary as it is usually not feasible to collect information from an entire sample.

Information gathered from these statistics can be strongest variable or proof of an opinion.

In my research, descriptive statistics could include a table showing types of error and their frequency in the dataset. An example question could be: Does error flagging reduce the error frequency and improve overall data quality?

With inferential samples, it should be random and unbiased.

Visualising data in charts is useful at helping us understand it. Different types of chart are useful for displaying different data.

To describe the correlation between academic success to student reading age results, I would use a scatter graph. This is because there are many different combinations possible, with academic success being continuous.

In my field of study, we use a lot of bar charts as it is useful at showing how many people have carried out specific training.

Aristotle said “The aim of art is to represent not the outward appearance of things, but their inward significance.”

I don’t agree with this – art does not have an aim. Artists create what they create for an infinite number of different reasons.

When creating charts, it is important to consider how the viewer will look at it.

Visual order of importance:

Size – Bigger catches the eye more and size hierarchy can show which order things should be read

Colour – Brighter colours catch the eye more

Contrast – High contrast catches the eye more

Proximity – The closer things are to each other, the more they seem related

Alignment – Important text should be more towards the left

When including charts, consider:

What are you trying to accomplish?

What story in your data do you want to share?

Are you trying to present too much?

Less is more

Unit 4 Topic 2

Qualitative research methods involve collecting and analysing data, normally non-numerical. Objective is often to understand stories, experiences and concepts of participants.

A researcher might use qualitative research methods as the topic of research is not numerical. Qualitative methods do not suit all research.

Limitations are that it is difficult to put the results into a tangible form.

Qualitative research follows four stages:

Data collection

Reading and rereading the data

Interpreting findings

Writing your report

I will organise my data in tables in the cloud.

A theme or pattern is when a group of things have similar properties or act in a predictable way.

Thematic analysis is a method used to identify patterns or themes that emerge from qualitative data. The process contains six steps:

- Familiarisation
- Generating initial codes
- Identifying themes
- Reviewing themes
- Naming themes
- Writing about themes

Excel is a useful tool for analysing qualitative data.

Identifying themes can be done using excel with column headers:

- Source #
- Note #
- Author
- Note
- Theme
- Initial Code



Unit 4 Topic 2 completed!

You should now:

- Have demonstrated knowledge and understanding of how to organise qualitative data ready for analysis.
- Have figured out some techniques to identify patterns and themes in your data.
- Have improved the understanding of the use of Excel as a qualitative analysis software.

[Back](#)

[Exit](#) ✕

If you have taken notes using the in-built notebook, you can now export them.

[Export my notes](#) ↑

Here are listed the references that are used in this topic.

[References](#) ↗