

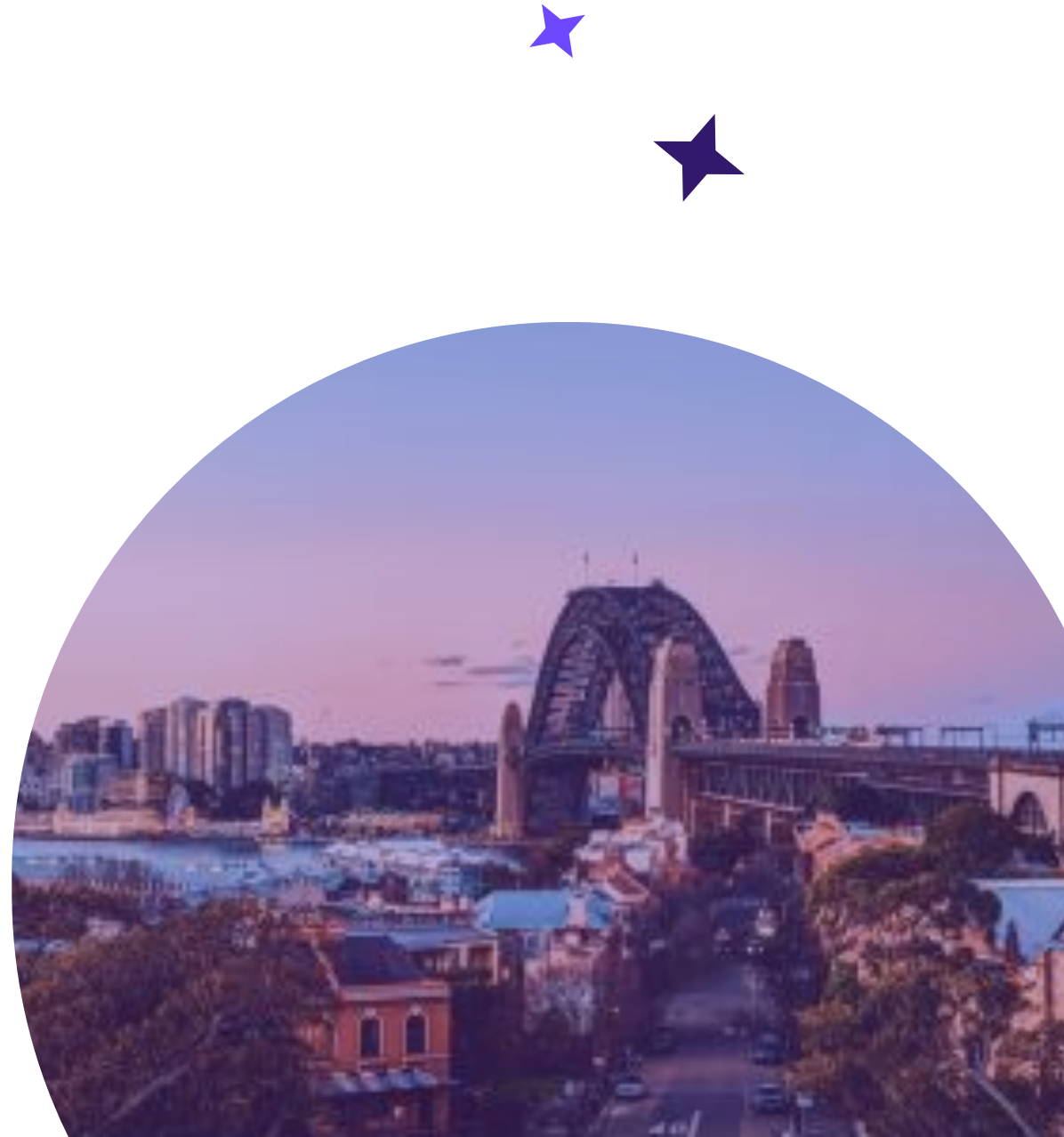


Transforming Education with AI-Driven Insights and Student Feedback

 University of New South Wales | Sydney, Australia

 May 13-14, 2024

Co-hosted



Data analytics from thin AAIR?

Practical uses of Artificial Intelligence for institutional reporting



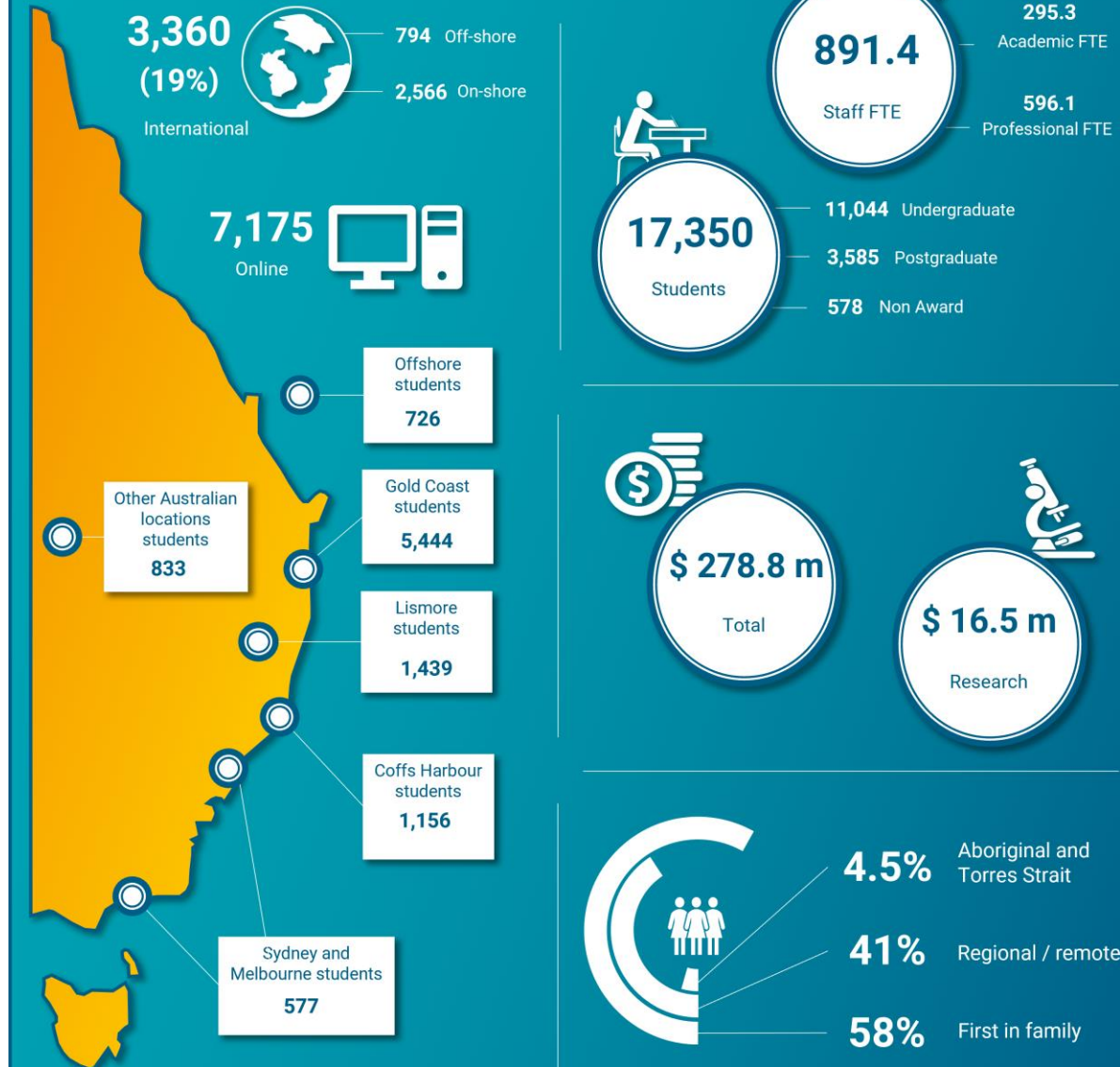
Southern Cross University

Our university

- Northern Rivers NSW-based
- 17,000 students / 11,000 EFTSL
- Largest numbers of students studying Online, Gold Coast campus
- Recently replaced traditional delivery with the Southern Cross Model: 6*6 week Terms
- BIQ responsible for data analytics, student surveys, quality management and government reporting



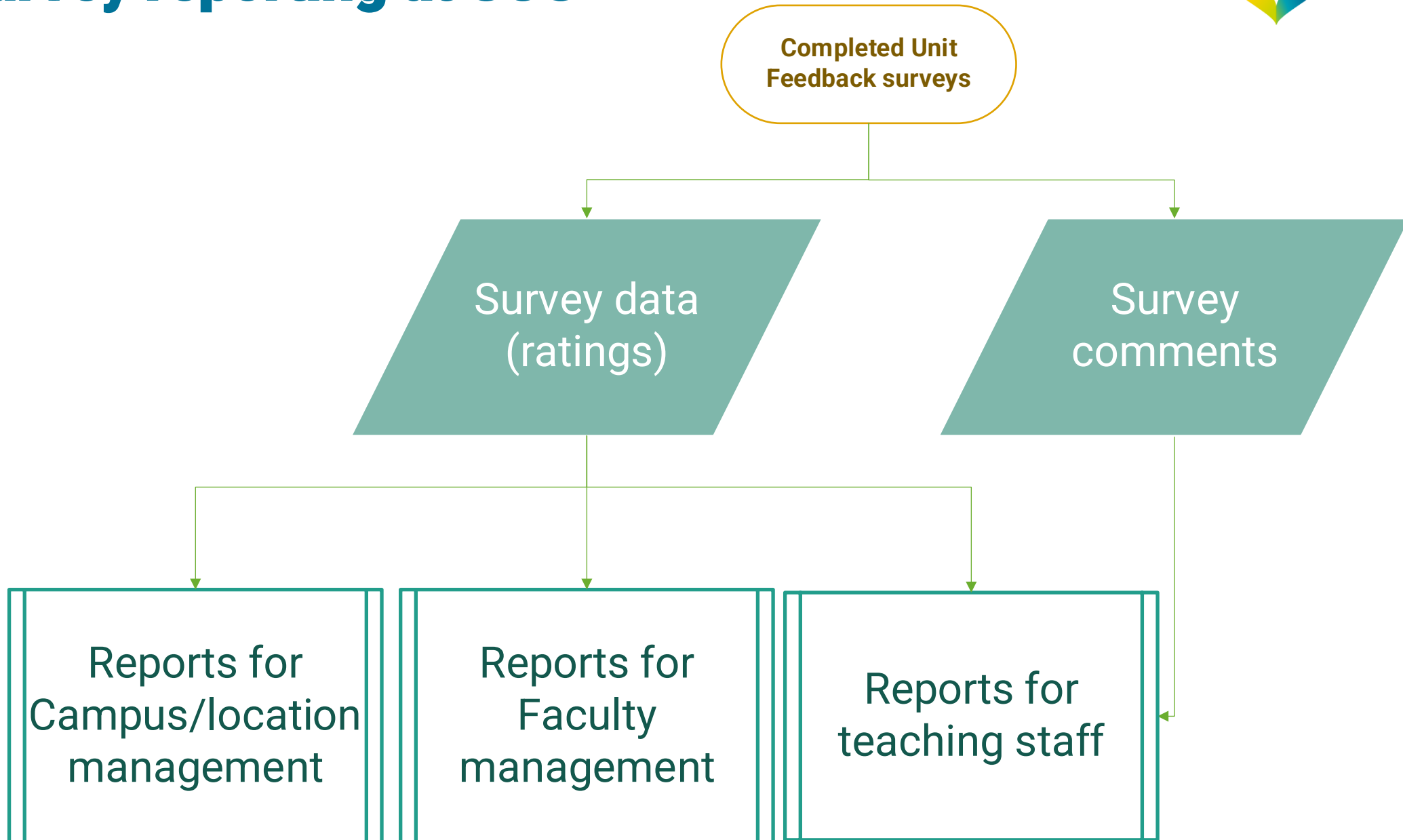
Southern Cross University at a Glance



These figures are subject to minor revisions throughout the year

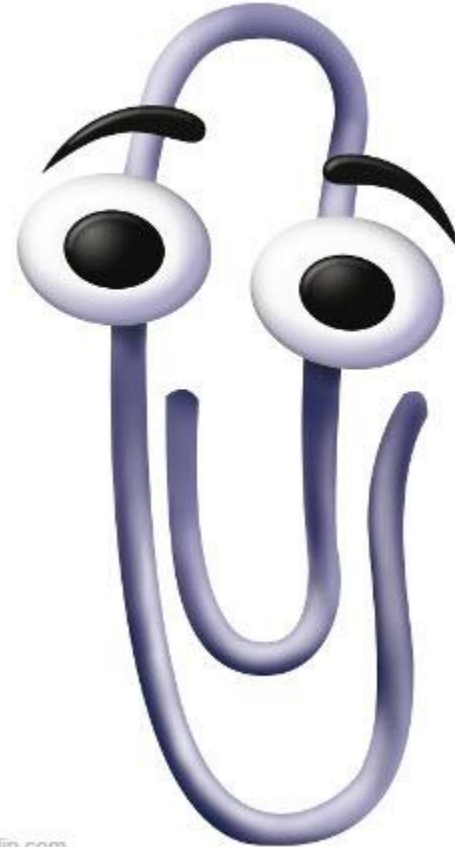
Source: Southern Cross University Management Information System

Survey reporting at SCU



Artificial Intelligence and survey reporting

- Breakout year for public access to AI
- Fast and efficient writing
- Natural language generation, excellent spelling & grammar, and accessible writing
- But can we trust AI conclusions?
How can we verify AI report content?



imgflip.com

It looks like you're writing a report. Would you like some help with that?

Our test: Thematic analysis of survey comments



Aim: Replicate a manual qualitative analysis of survey comments using an AI tool. Cross-validate clerical results against the AI output.

Data: Anonymised test data simulating unit feedback comments

Clerical method: Latent Dirichlet Allocation (LDA) model conducted via R, followed by clerical identification and labelling of themes.

VS

AI method: Used a generative AI tool with least possible prompting. No specific methodology referenced in prompts

SCU Unit Feedback



The unit feedback survey is an internal SCU survey conducted across all unit enrolments across each Term. The survey consists of 13 rating questions and two qualitative comment questions.

Test data replicated feedback from across the university for a single Term.

Helpful aspects: What aspects of this unit were the most helpful for your learning?

Improvements: What would have improved your learning experience in this unit?



Unit feedback instrument

Southern Cross
University

The following statements relate to your unit. How strongly do you agree or disagree that:

ID	Summary	Question as shown in questionnaire
QU01	Clarity	The unit's expectations and learning outcomes were clear
QU02	Site content	The unit's learning site content helped me learn
QU03	Classes	The classes and/or their recording helped me learn
QU04	Workload	The unit's workload was reasonable
QU05	Unit engagement	The unit's learning experience was engaging
QU06	Assessment	The assessment tasks were relevant to the unit
QU07	Rubrics	The marking rubrics helped me to understand the assessment tasks
QU08	Overall unit	Overall, I am satisfied with the unit
CU01	Helpful aspects	What aspects of this unit were the most helpful for your learning?
CU02	Improvements	What would have improved your learning experience in this unit?

The following statements relate to the teaching in your unit. How strongly do you agree or disagree that:

ID	Summary	Question as shown in questionnaire
QT01	Learning	The teaching in the unit helped me learn
QT02	Feedback	The teacher feedback provided helped me learn
QT03	Teaching engagement	The teaching engaged me in the unit learning
QT04	Support	I was able to access support from my teacher when needed
QT05	Inclusiveness	The teaching was respectful and inclusive of diversity
QT06	Overall teaching	Overall, I am satisfied with the teaching in this unit

Select at least one of this unit's teaching staff to provide individual teacher feedback.

Please select only those staff who have taught you in this unit.

ID	Summary	Question as shown in questionnaire
QI01	Overall individual teacher	Overall, I am satisfied with the teaching of [TEACHER NAME] in this unit
CI01	Individual teacher comment	{Blank space underneath rating question}

Latent Dialect Allocation Using R



- Load test data
- Create a corpus
 - Remove empty comments
 - Remove punctuation
 - Remove symbols
- Create a Document Term Matrix
 - Remove general stop words
 - Remove survey-specific stop words
- Fitting a LDA Model
 - Selected five themes

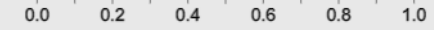
```
41 library(tidyverse)
42 library(tidytext)
43 library(quanteda)
44 library(topicmodels)
45 # Loading our data-----
46 feedback <- readRDS("test unit feedback.rds")
47
48 # Data Cleaning-----
49 # Creating a tokenised list of our feedback comments for analysis
50 feedback_corpus <- tokens(feedback %>%
51   filter(!is.na(CU01),
52         teaching_period == "Term 4") %>%
53   pull(CU01),
54   what = "word4",
55   remove_punct = TRUE,
56   remove_symbols = TRUE)
57
58 # Creating a document term matrix to feed to the LDA model
59 dtm <- dfm(feedback_corpus) %>%
60   dfm_remove(stopwords("english")) %>%
61   dfm_remove(c("the tutorials", "tutorials", "the workshops", "workshops",
62              "na", "n/a", "N/A", "nothing", "null", "Null", "nil",
63              "Nil", "none", "None")) %>%
64   convert(to = "topicmodels")
65
66 # Model Fitting-----
67 # Creating our LDA model
68 lda_positive_aspects <- LDA(dtm, k = 5)
69 lda_positive_aspects
```

Helpful Aspects LDA results

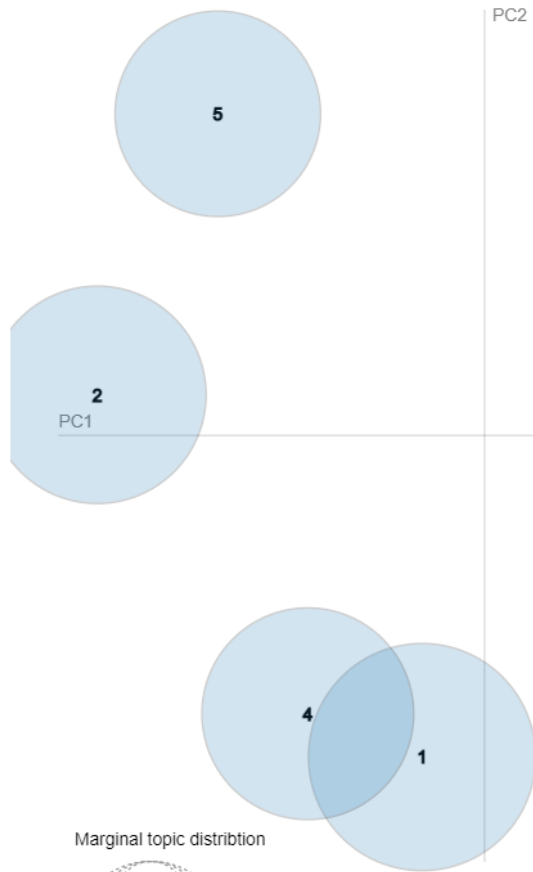
Selected Topic:

Slide to adjust relevance metric:⁽²⁾

$\lambda = 1$



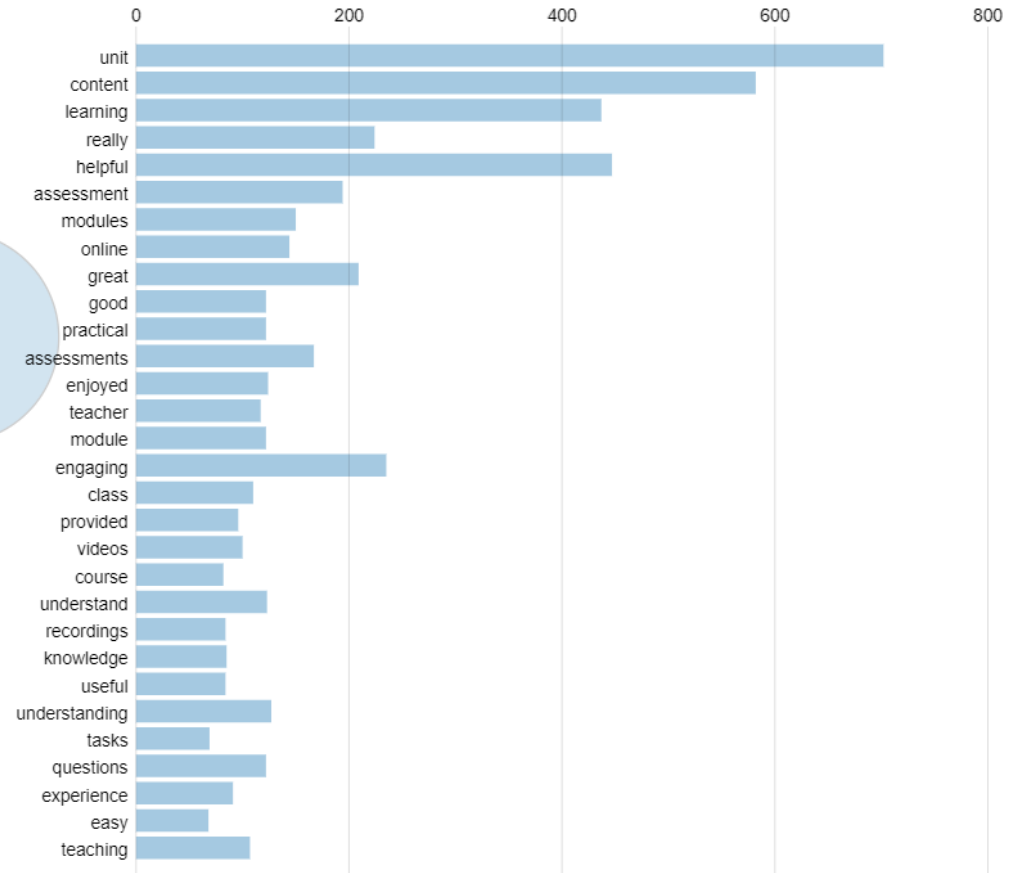
Intertopic Distance Map (via multidimensional scaling)



Marginal topic distribution



Top-30 Most Salient Terms¹



Overall term frequency

Estimated term frequency within the selected topic

1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)

2. relevance(term w | topic t) = $\lambda * p(w | t) + (1 - \lambda) * p(w | t)/p(w)$; see Sievert & Shirley (2014)

Clerically identified feedback themes



Helpful Aspects

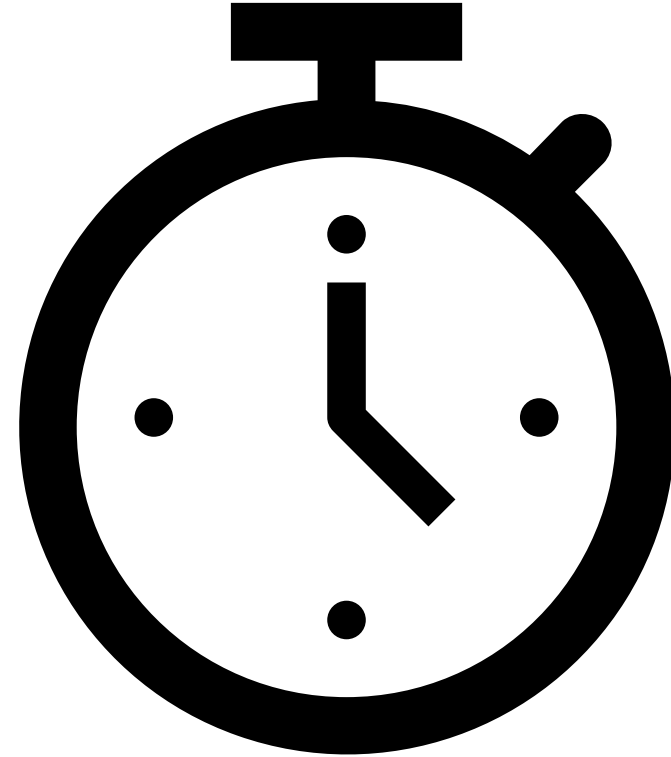
1. Teachers and teaching
2. Practical, information, activities, resources and experience
3. Assignments feedback understanding/learning
4. Helpful and engaging, online modules, classes and lectures delivery
5. Content delivery, Delivery resources - Videos, recordings, tutorials, discussions

Improvements

1. Workload and time management, assessments
2. Assessments and assignments feedback and content
3. Timing of delivery and length of study period
4. Assessments and materials, developing understanding
5. Online classes and study

How long did that take?

- File Preparation – 30 minutes
- Data Cleaning – 30 minutes
- LDA Creation – 20 minutes
- LDA Validation – 20 minutes
- Theme Labelling – 1 hour
- Report Preparation – 1 hour



Now it's the machine's turn...

Transferred as much processing as possible to AI responsibility.

We made three attempts based on variations on our data preparation.

We used a single prompt:

"Extract the top five themes from the Helpful Aspects column"

Attempt 1: AI themes from raw qualitative data



Helpful Aspects

1. Everything - mentioned 10 times.
2. Workshops - mentioned 9 times.
3. The workshops - mentioned 8 times.
4. Tutorials - mentioned 8 times.
5. Unit content - mentioned 7 times.

Improvements

1. "nothing" - mentioned 22 times.
2. "Nothing" (with a capital 'N') - also mentioned 22 times.
3. "Nil" - mentioned 10 times.
4. "More time" - mentioned 7 times.
5. "." (a period) - mentioned 6 times.

Attempt 2: AI themes after minimal data cleaning



Helpful Aspects

1. Tutor Worked
2. Date Despite Complex
3. Learn Felt Confident
4. Thoughts Ideas
5. Responsibilities

Improvements

1. Needed QGIS
2. Guide Learning Expectation
3. LI Honest Think
4. Unit Required
5. OA Discussed Evidence

Preparation for attempt 3: Intensive data cleaning



- Removed special characters
- Removed punctuation
- Removed stop words (General and survey-specific)
- Removed whitespaces

```
comments <- read_csv("test feedback comments.csv")

SCU_lexicon <- tibble(word = c("units", "felt"),
                     lexicon = "SCU Lexicon")

sw_list <- stop_words %>%
  bind_rows(SCU_lexicon) %>%
  pull(word)

comments_clean_stop <- comments %>%
  mutate(across(everything(), ~iconv(., "UTF-8", "UTF-8", sub='')),
         orig_helpful = `Helpful Aspects`,
         orig_improve = Improvements,
         across(`Helpful Aspects`:`Improvements`, ~str_remove_all(tolower(.), paste0(sw_list, collapse = "\\b|\\b"))),
         across(`Helpful Aspects`:`Improvements`, ~str_remove_all(., "[[:punct:]]")),
         across(`Helpful Aspects`:`Improvements`, ~str_squish()))
```


Attempt 3 results: Helpful Aspects



AI Helpful Aspects themes

1. Interactive Teaching and Workshops
2. Resourceful and Comprehensive Learning
3. Teacher Engagement and Learning Support
4. Interactive Online Learning
5. Practical Application and Task Management

Clerical Helpful Aspects themes

1. Teachers and teaching
2. Practical, information, activities, resources and experience
3. Assignments feedback understanding/learning
4. Helpful and engaging, online modules, classes and lectures delivery
5. Content delivery, Delivery resources - Videos, recordings, tutorials, discussions

Attempt 3 results: Improvements



AI Improvements themes

1. Assessment Clarity and Feedback
2. Enhanced Learning Experience
3. Engaging and Timely Content Delivery
4. Content Management and Understanding
5. Online Learning Effectiveness

Clerical Improvements themes

1. Workload and time management, assessments
2. Assessments and assignments feedback and content
3. Timing of delivery and length of study period
4. Assessments and materials, developing understanding
5. Online classes and study

The results: Behind the Scenes



"Explain the process you went through to identify those themes."

To identify the themes from the 'Helpful Aspects' column of your dataset, I employed a method known as Latent Dirichlet Allocation (LDA), a type of unsupervised machine learning technique used for topic modelling. Here's an overview of the process:

1. Data Preparation

- Extraction
- Vectorization

2. Model Training with LDA

- Latent Dirichlet Allocation
- Random State

3. Theme Extraction

- Topic Identification
- Word Lists for Each Topic

4. Labelling Themes

The results: Behind the Scenes Part 2



"Can you please explain step 4. Labelling Themes in more detail?"

Absolutely! Step 4, "Labelling Themes," is a crucial part of the topic modelling process where we interpret the statistical results of the LDA model and translate them into meaningful, understandable labels. This step requires a blend of analytical skills and contextual understanding. Here's a detailed breakdown:

1. Interpreting Topic Keywords
2. Contextual Analysis
3. Synthesizing a Theme Label
4. Ensuring Clarity and Relevance
5. Examples from Your Data

This step is more art than science, requiring a mix of analytical insight and interpretive skill to translate statistical data into meaningful, contextually relevant themes.

Project conclusion: It's ok.



Strengths of AI analysis

- Successfully replicated the clerical analysis with minimal prompting
- Excellent theme labelling and text summaries
- Accurate selection of example comments
- Saved some time particularly in reporting findings

Weaknesses of AI analysis

- Required intensive data cleaning before AI analysis produced workable results
- AI analysis didn't always choose an LDA method when given the choice
- Small changes in prompts created vastly different outcomes

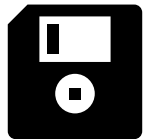
Boundaries to implementing general AI



Integration with existing
report infrastructure



Clerical validation still
required



Software upgrades and
unexpected changes



Legal and privacy
implications

Thanks for listening – Questions?

