

# REIT6811 Tutorial 5

## Mixed-method data analysis

Please go through this handout before the tutorial session. **Some activities are meant to be done before the class, these are indicated in blue.** You will not be given time to do these activities during the session.

Tutorials are evaluated for each student; therefore, all students are required to participate. Every student should be actively involved in discussions among your groups and participate when your tutor discusses the answers. The marking rubric for tutorials is on Blackboard.

In this tutorial, you will analyse a hypothetical dataset from a study involving the language learning software LinguaLift. LinguaLift is designed to teach new languages to users through engaging features such as AI-driven conversation practice, gamification to boost learning motivation, and adaptive learning paths tailored to the individual and their progress.

Twenty-one users participated in a short study to test the effectiveness of LinguaLift. The participants chose a language to learn, and initially tested their proficiency using the software. They then spent 2 months using and testing the software. The total number of hours each user spent in the software was automatically logged. At the end of the 2 month period, participants re-took the proficiency test and completed a survey. The survey asked the participants to rate aspects of the software on a 0-5 scale, and provide qualitative feedback about their experiences.

The survey questions and data from the study are at the end of this tutorial in Appendices A and B. **Before class, download the data from <https://github.com/reit6811/data-analysis-tut>, and perform some exploratory data analysis in your software of choice.**

Work through the following exercises, discussing the answers in your groups. Where specified, record your answers in your journals.

### Activity 1 – Initial Impressions (15 mins)

1. What insights are the researchers hoping to gain by collecting this data? What research questions are they hoping to answer? *(In journal 1-2 sentences)*
2. Participant 13's data differs from the others in the data set.
  - a) What do you think could have happened and how might this affect the overall data analysis? *(In journal 1-2 sentences)*
  - b) What strategies could you use to handle outliers in this data set? *(In journal 1-2 sentences)*

### Activity 2 – Quantitative Data (30 mins)

3. What are some simple ways researchers might summarize the survey results to present to a team of software developers? *(In journal 1-2 sentences)*
4. Using Datawrapper (<https://www.datawrapper.de/>), Excel, Python (matplotlib.pyplot) or any other software you are familiar with, create a plot that shows
  - a) Participant satisfaction with the app
  - b) How participant satisfaction is related to learning outcomes
  - c) The relationship between actual learning outcome and perceived learning outcomes of the participants

*Paste an image of your plots in your journals. Include a one sentence justification for the visualisation/plot used and one sentence describing the trend in the data revealed by the plot.*

### Activity 3 – Qualitative Data (30 mins)

5. The research team has identified the following codes from the qualitative data:
  - 1) Usability Issues
  - 2) Technical Issues
  - 3) Engagement Features
  - 4) Learning Effectiveness
  - 5) AI Interaction
  - 6) Content Limitations and Diversity

In your groups, determine how you would code each of the participant comments (i.e., which code does each comment relate to?)

6. One theme that could be developed from the qualitative data is “User Experience”. The User Experience theme encompasses both the issues that hinder experience as well as features that enhance it, including codes such as Usability Issues, Technical Issues, AI Interaction and Engagement Features.

Discuss in your groups other themes that could be developed from the data. *(In journal describe the theme and codes it relates to in 2-3 sentences).*

### Activity 4 – Improving the Study Design (15 mins)

7. Your group is tasked with re-designing the data collection methodology for the next round of testing. Discuss any changes you would make to the data collected through the software or in the feedback survey to improve the comprehensiveness and accuracy of the study.

As a group, share the changes you would make with the class.

8. As a class, discuss the benefit of the mixed-method data set in this example. What information was gained from the mixed-method data that couldn't be obtained from quantitative or qualitative data alone?

### Activity 5 (optional) – do not do this exercise before completing the analyses above!

Repeat the exercises above using generative AI. What differences does the LLM find compared to your own and your table's findings?

Appendix A. Survey Questions:

**LinguaLift Participant Feedback**

Thank you for participating in our LinguaLift software testing phase. Your feedback is crucial in helping us improve and tailor our product to better meet user needs. Please answer the following questions based on your experience with the software.

1. **Ease of Use:** How easy is it to navigate the software’s interface?

☐ 0  
Strongly disagree

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5  
Strongly agree

2. **Effectiveness:** How effective were the lessons in improving your language skills?

☐ 0  
Strongly disagree

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5  
Strongly agree

3. **Engagement:** How engaging are the language games and AI conversations?

☐ 0  
Strongly disagree

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5  
Strongly agree

4. **Overall Satisfaction:** Overall, how satisfied are you with the software?

☐ 0  
Strongly disagree

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5  
Strongly agree

5. **Most Useful Feature:** What feature did you find most useful in the software?

6. **Biggest Issue:** What is the biggest issue you encountered while using the software?

## Appendix B. Study Data:

Participant	Hours Used	Test Score Before /100	Test Score After /100	Test Score Change /100	Q1	Q2	Q3	Q4	Q5	Q6
1	45	53	82	29	4	3	5	4	Interactive quiz game	Occasional crashes during lessons
2	18	57	65	8	2	4	3	3	AI conversations	Complicated to get around
3	38	56	83	27	3	5	4	4	Personalized learning paths	Needs more language options
4	15	48	56	8	1	3	2	2	Daily progress tracker	User interface is not intuitive
5	50	58	85	27	5	4	5	5	Gamification elements	Too few examples in foreign language
6	25	54	62	8	3	3	2	3	Quiz game and feedback	Slow to load new lessons
7	30	49	59	10	4	2	3	3	Multiple language support	Speech recognition needs improvement
8	35	60	70	10	4	3	4	4	AI-driven practice conversations	Lacks advanced content for proficient users
9	28	46	60	14	3	2	4	3	Variety of exercises	Repetitive tasks
10	10	70	68	-2	2	1	1	1	AI pronunciation feedback	Inaccurate translations sometimes
11	40	51	69	18	4	3	4	4	Progress tracking and learning path	Not enough cultural context
12	55	61	83	22	5	5	5	5	Custom learning pathway	Nothing, it's great
13	52	58	69	11	1	1	0	1	The lessons are really engaging and fun	Not much, a few games were a bit too easy
14	22	55	57	2	2	2	3	2	Interactive games and quizzes	Limited number of stories
15	26	51	54	3	3	3	2	3	Grammar tips from AI	Some lessons are too long
16	32	59	65	6	3	4	5	4	Real-time feedback	Occasional technical glitches
17	12	49	51	2	1	2	1	1	Cultural insights	Frequent crashes were annoying
18	29	65	68	3	4	5	3	4	Learning games	Some games were too easy
19	48	60	87	27	5	4	5	5	Extensive vocabulary lists	More examples needed in some sections
20	20	47	56	9	2	2	3	2	Flashcard game	Would have loved more flashcard variety
21	5	71	73	2	3	2	2	3	Real-time feedback on pronunciation	A lot of the time AI didn't understand me

## Frequency of Score Responses:

Score	Q1	Q2	Q3	Q4
0 Strongly disagree	0	0	1	0
1	3	2	2	3
2	4	6	4	3
3	6	6	5	6
4	5	4	4	6
5 Strongly agree	3	3	5	3