Executive Summary

Objective: This report aims to evaluate student engagement in the Mr. Kato Ophthalmology Module by modelling metrics such as module completion, usage and attempts. To further analyze the difference in engagement over time, variables of ‘time spent per page’ and ‘text scores’ will be examined through a Latent Profile Analysis (LPA). By doing so, we seek to uncover insights into how students interact with learning material.

Data Source:

The primary data source used in this analysis is xAPI data retrieved from the Learning Record Store (LRS) on Entrada, the Learning Management System (LMS) for the University of British Columbia’s Faculty of Medicine. Each event is stored as a JSON object containing key information such as:

- **Timestamp**: The precise time when the event occurred.
- **Person Name**: The person interacting with the module
- **Page Name**: The name or identifier of the page being interacted with.
- **Descriptors**: Verbs such as ‘Experienced’ or ‘Answered’
- **Responses**: Any relevant user input or actions, if applicable.

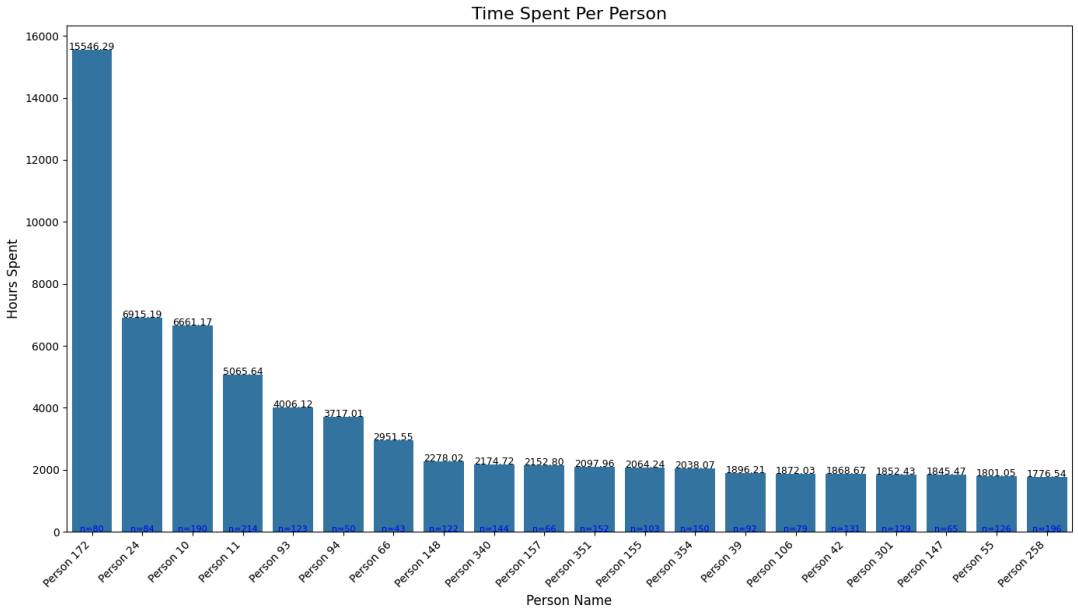
We collected the xAPI data from November 1, 2022, to November 1, 2024, with 404 participants.

Preliminary Metrics to Pursue:

Number of Events: 32584

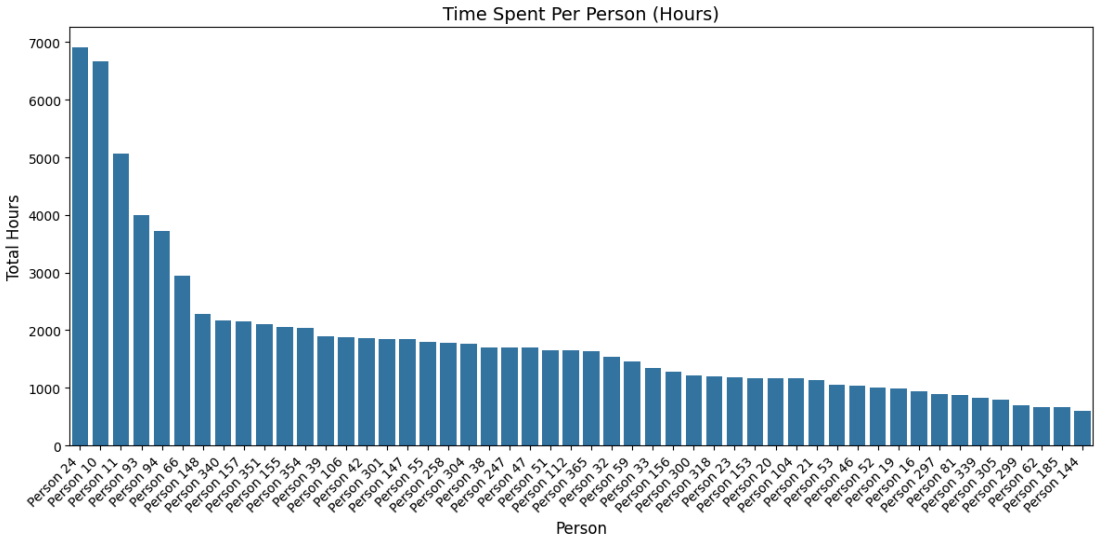
Unique users: 404

Time spent per user (first timestamp to last timestamp):



- This bar graph was meant to show the top 20 people who spent the most time on the module by finding the difference between each person's first and last time stamp. However, this is much more time than expected for an approximately 30-minute module.
- This chart seems to indicate more people who have revisited the module for review, with one person looking back at the module over a year after first viewing it.

Time spent per user (aggregate of sessions)



This bar graph displays the total time the top 50 users spent in the module, aggregated by sum across all their sessions. While the module is designed to take approximately 30 minutes to complete, the recorded totals are significantly higher, likely due to repeated access, extended idle time, or multitasking during sessions.

- **Extreme outlier removed:** One user with 15,546.3 hours (~1.8 years) of recorded time was excluded, as this was anomalous—possibly due to a system error or sessions left running indefinitely.
- **Two-tiered usage pattern:** A small group of heavy users logged between 3,000 and 7,000 hours, while the majority of users recorded under 2,500 hours.
- **Top users' substantial engagement:** Person_24 recorded nearly 7,000 hours (~292 days), with Person_10 and Person_11 surpassing 5,000 hours. These numbers suggest unusually prolonged or frequent access.
- **Plateau effect:** After the top 6 users, a middle tier emerges with more stable totals between 1,500 and 2,200 hours, showing relatively consistent patterns among mid-level users.

Interpretation: These values represent accumulated session durations, not continuous interaction. While more representative than simple timestamps, they likely still include passive time and should not be interpreted as fully active learning. Nevertheless, the consistently high totals suggest this module is part of an extended educational workflow.

Sessions (attempts) per user count the ‘Mr. Kato’s Vision Loss’

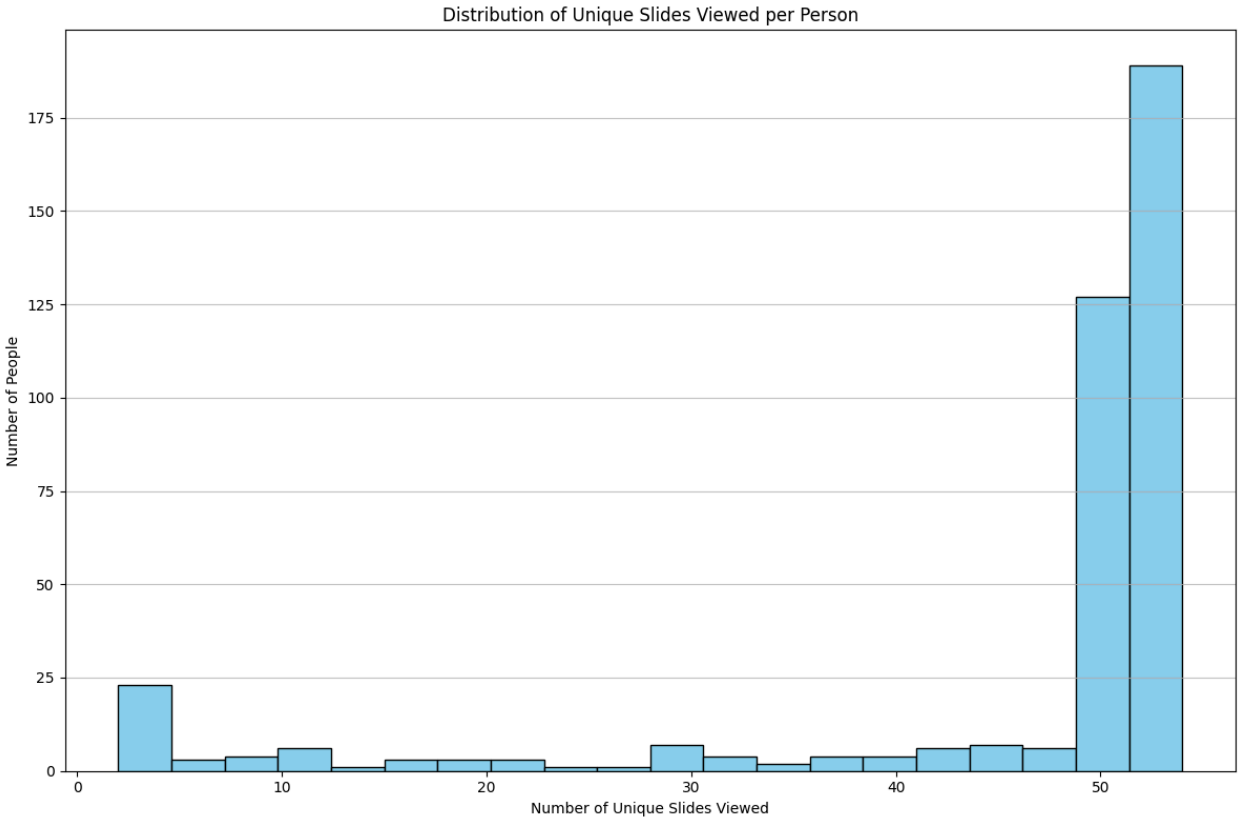
This chart shows how many times each person in the system has registered. Most users (259) registered only once, making single registrations the most common.

- **Decreasing pattern:** As the number of registrations increases, the number of users sharply declines:
    - 259 users registered once
      95 users registered twice
    - 35 users registered three times
    - 7 users registered four times
    - 6 users registered five times
    - Only two users registered six times
- **Total user count:** These values represent around 404 individuals in the dataset.
- **Engagement breakdown:**
    - Most users (~64%) registered only once
    - About 36% returned for multiple registrations
    - Fewer than 4% registered four or more times

Interpretation:
This pattern suggests that while most users engage just once, a meaningful portion returns

multiple times—potentially for review. The steep drop-off beyond three registrations implies that frequent re-registration is rare and that most returning users do so only once or twice. These trends could help identify engagement opportunities and understand how users cycle through the system.

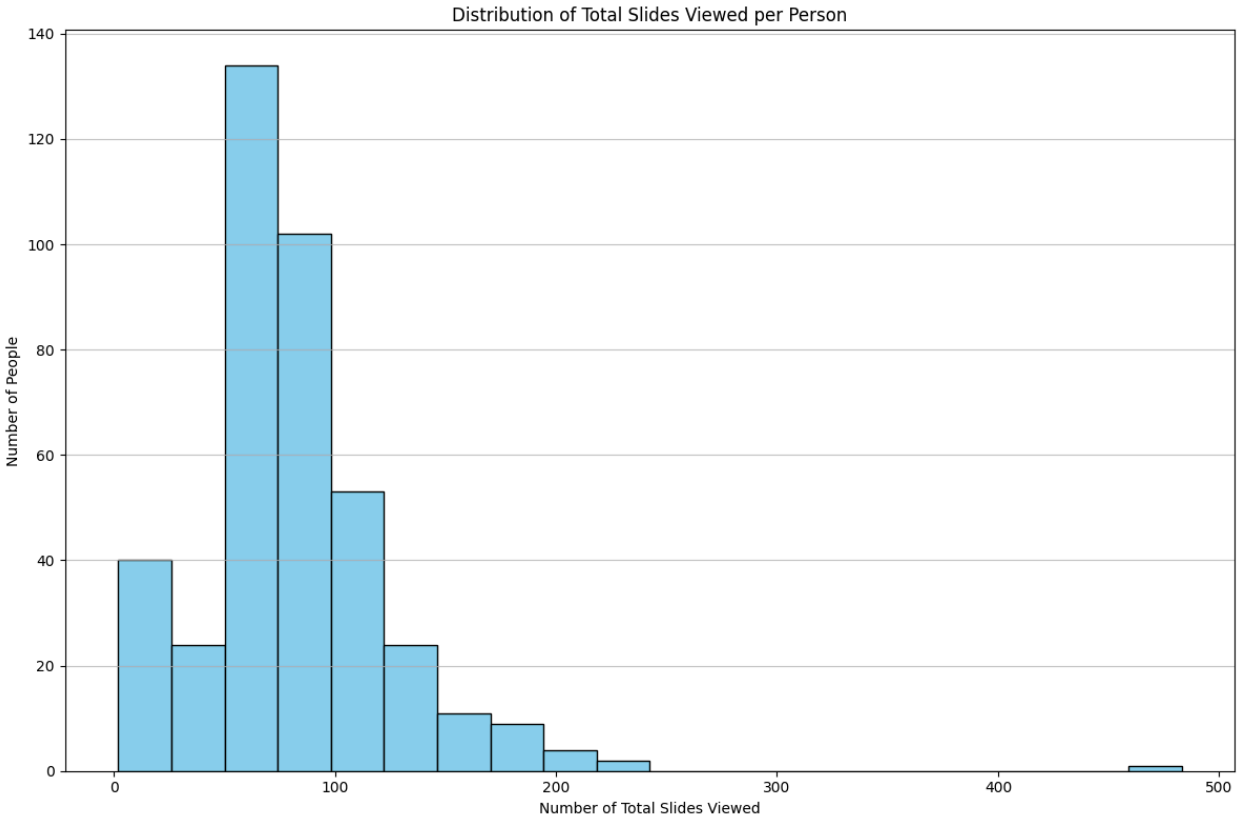
Depth of module usage (number of slides they went through)



This histogram shows how many unique slides each user viewed in the module (out of a maximum of 54). The distribution is distinctly bimodal, with two clear peaks—one near the beginning and one at full completion.

- **Complete engagement:** The largest group of users (~190) viewed all or nearly all 54 slides. This suggests a high completion rate and strong sustained engagement.
- **Minimal engagement:** A smaller cluster (~23 users) viewed only a few slides (~5 or fewer), indicating early drop-off or brief content exploration.
- **Limited middle ground:** Very few users viewed between 10 and 45 slides. This gap reinforces the idea of a binary engagement pattern—most users either commit fully or disengage early.

Interpretation:
This kind of bimodal distribution is often seen in online learning, where users either briefly test the waters or fully engage. Understanding where and why users drop off could help optimize early slide content to improve retention further.



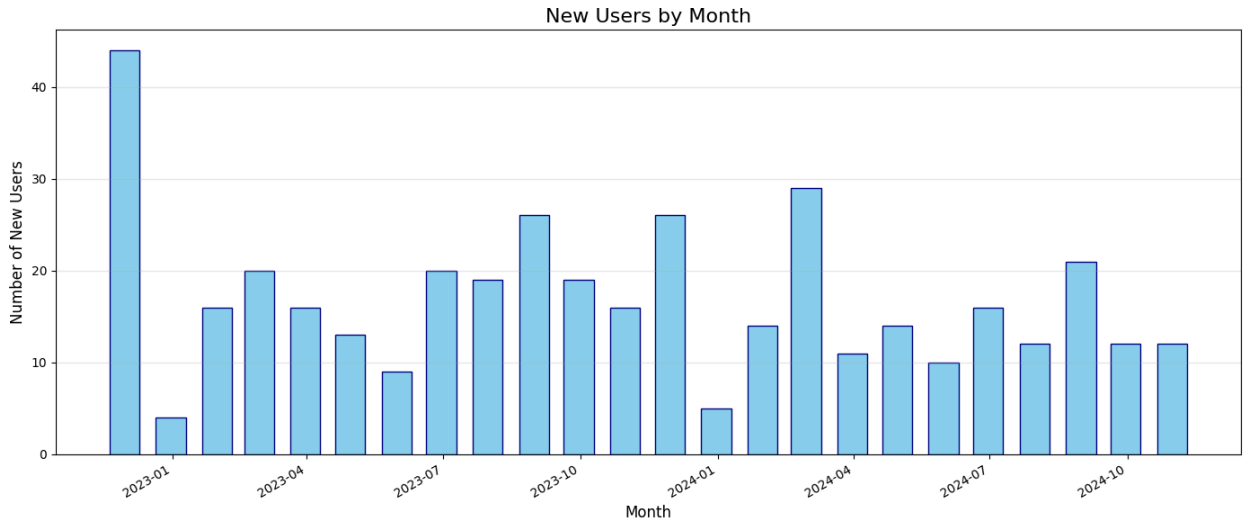
This histogram shows the total number of slides viewed per user, counting repeated views of the same slide. It contrasts with the previous chart, which showed the number of unique slides viewed.

- **Repeat viewing pattern:** While most users viewed around 50 unique slides, the peak here is in the 75–100 range, suggesting that users typically revisit some slides. On average, each slide is viewed 1.5 to 2 times per user.
- **Right-skewed distribution:** The distribution has a long right tail. Most users viewed between 50–150 slides.
- **Low-engagement group:** Roughly 40 users viewed very few slides (0–25), aligning with the early dropoff pattern in the unique slides chart.
- **Review users:** A small number of users (in the 475–500 range) repeatedly engaged with the content—potentially treating it as a reference resource or engaging in intensive review.

- **Content consumption pattern:** The peak at 75–100 suggests that most users go through the module once, with some light review, rather than doing full repeat passes through all slides.

Interpretation:
This view reveals how users interact beyond a single pass compared to the unique slides chart. Most complete the module and revisit selected slides, while a few engage at a much deeper level. This pattern highlights a flexible use case—some users consume the content linearly, others revisit specific sections, and a few use it heavily as a study or reference tool.

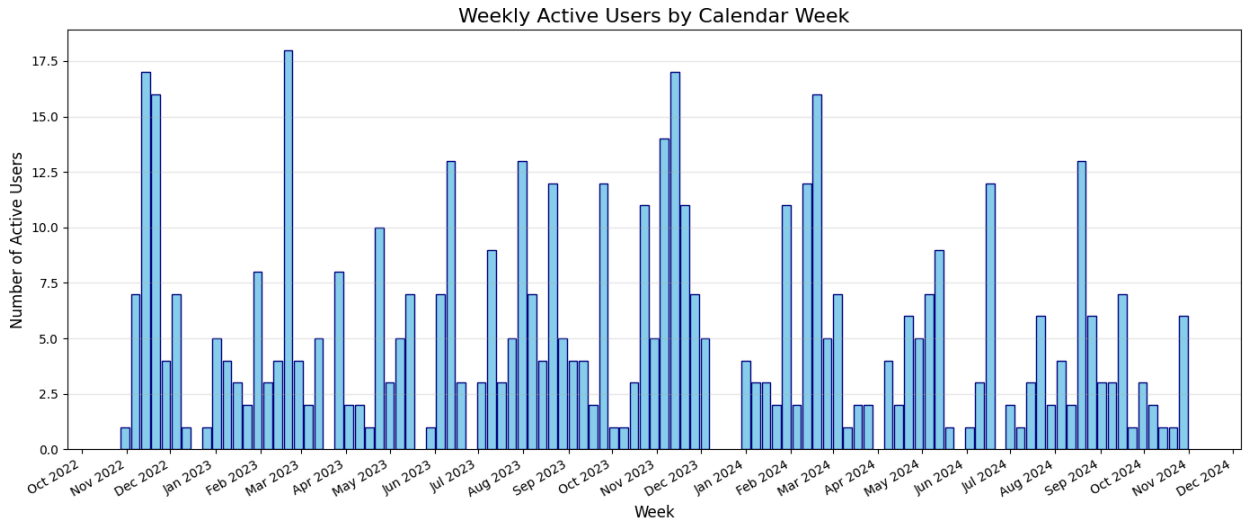
Plot the number of users over time.



This bar chart shows how many new users monthly over two years. The trend reflects an initial launch surge followed by a stable new user acquisition rate, with seasonal fluctuations likely tied to academic or institutional cycles.

- **Initial launch spike:**
    - The highest number of new users (~44) was recorded in January 2023.
    - However, since the data begins in November 2022, it's possible that this spike does not reflect the true number of new users, but rather includes a wave of users returning to review the module.
- **Steady acquisition:** After the initial surge, monthly new registrations generally settled into 10–20 users per month, indicating consistent but moderate organic growth.
- **Seasonal peaks:** Noticeable increases in registrations occur around:
    - October 2023 (~26 users)

    - January 2024 (~26 users)
    - February 2024 (~29 users)
    - September 2024 (~21 users)
- These spikes may coincide with the start of the academic term, supporting the idea that the platform serves an educational purpose.
- **Low points:**
    - February 2023 (~4 users)
    - January 2024 (~5 users)
    - April 2024 (~11 users)
- These dips could reflect holiday periods or breaks in institutional activity.
- **Year-over-year comparison:** No significant upward or downward trend is evident between 2023 and 2024, suggesting the module is stable

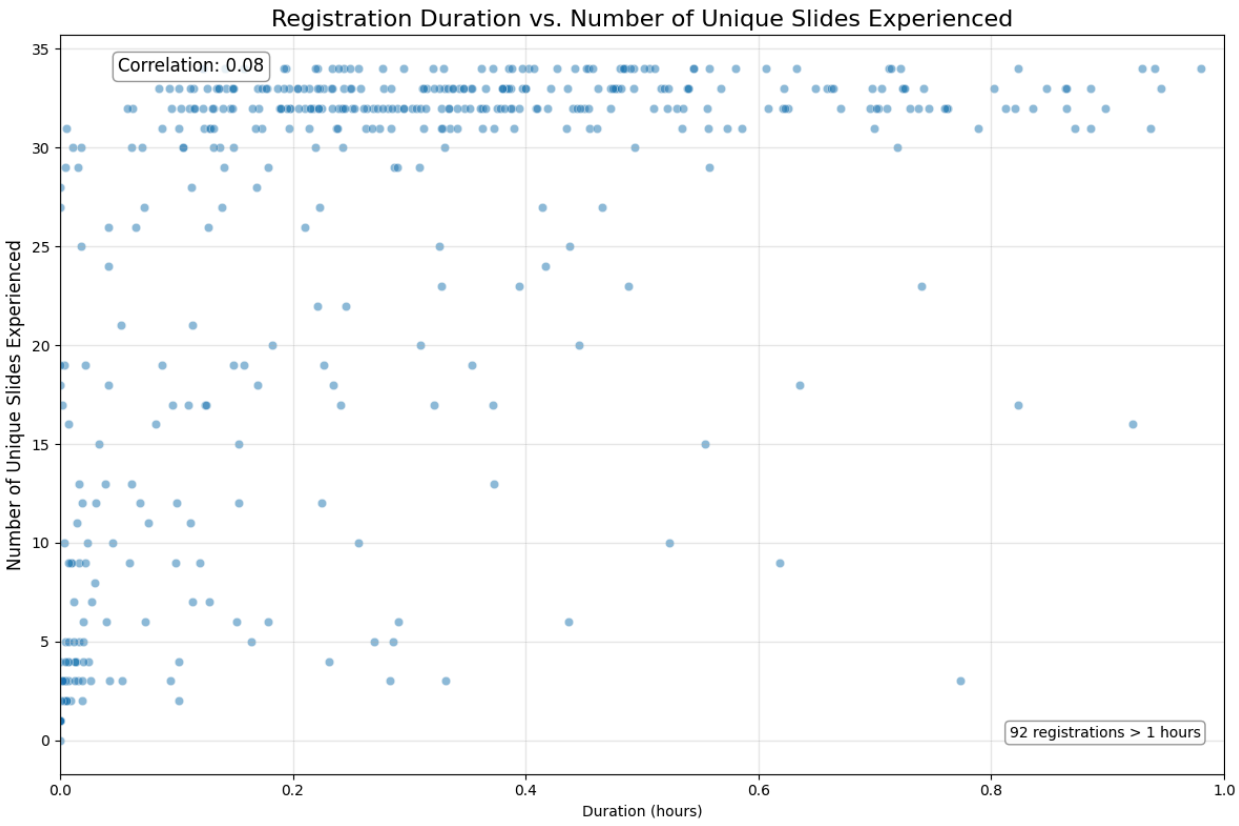


This chart shows how many users were active each calendar week over two years. The pattern reveals cyclical engagement tied to possible academic or operational cycles, with a steady usage baseline throughout.

- **Usage spikes:** There are several noticeable peaks in activity:
    - Nov–Dec 2022 (15–17 users)
    - March 2023 (highest peak, ~18 users)
    - November 2023 (~17 users)
    - Feb–Mar 2024 (~16 users)
- **Seasonal patterns:** Higher activity tends to cluster in:
    - Late fall / early winter (Nov–Dec)
    - Early spring (Feb–Mar)

- This likely reflects academic semester cycles.
- **Declining peaks:** Peak weekly activity in 2024 is slightly lower than in 2023, suggesting a possible slight tapering of engagement over time.
- **Baseline consistency:** Despite variability, most weeks maintain at least 2–5 active users.
- **Quiet periods:** Lower activity is seen in:
    - Early January (2023 & 2024) — likely post-holiday lulls
    - Summer months (July–August) — typical downtime in academic or work calendars
    - Late 2024 — possibly a result of waning platform use or external factors

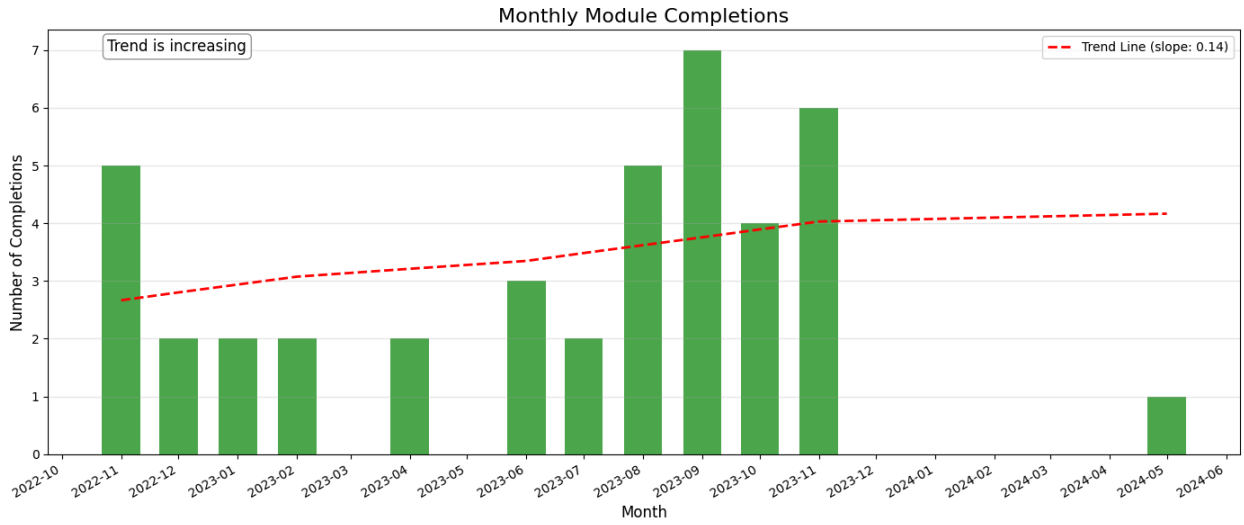
Other Visualizations:



This scatter plot explores how long users remain registered (in hours) versus how many unique slides they view.

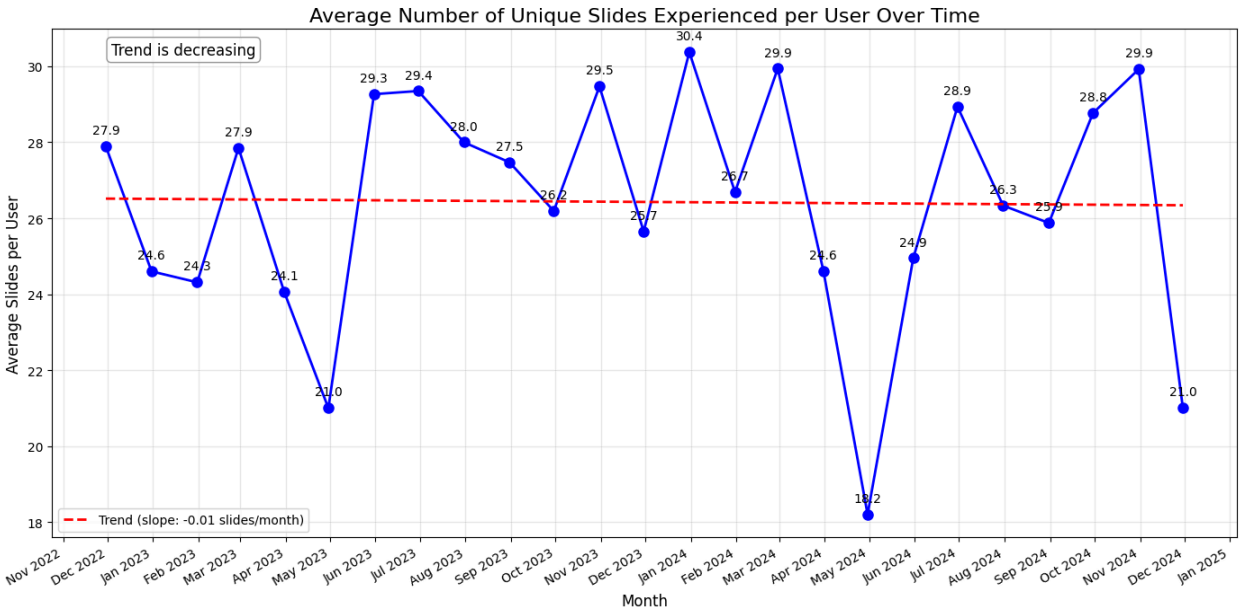
- **No linear correlation**

    - The correlation coefficient is 0.08, indicating no significant relationship between session duration and slide engagement.
- **Low duration - low slide clustering**
    - This may indicate that student may skim through the first few slides to determine the worth of completing this module to properly manage their time.
- **Registration duration is highly skewed**
    - 532 out of 624 registrations lasted less than 1 hour. Only 92 sessions exceeded that.
    - The median duration is just 0.26 hours (~15 minutes).
    - The mean duration (26.36 hours) is inflated by outliers, including one extreme case of 15,546 hours (~1.8 years)—likely due to system idling or a technical issue.
- **Slide engagement is consistently high.**
    - Most users viewed 31 of 34 unique slides (median = 31).
    - There's a tight cluster in the upper-left of the plot: short-duration sessions with high slide completion, suggesting users are efficiently consuming content, or skimming module pages.
- **Engagement efficiency**
    - Many users complete or nearly complete the module in 15–30 minutes.
- **System behavior considerations**
    - The lack of correlation and extreme outliers imply that "registration duration" likely includes idle time.



This bar chart tracks the number of module completions each month over 20 months. While the pattern is variable, a red dashed trend line indicates a slow upward trajectory (slope: 0.14), reflecting modest overall growth.

- **Completion peaks:** Several months stand out for elevated completion activity:
    - September 2023 (7 completions)
    - November 2023 (6 completions)
    - August 2023 and November 2022 (5 completions each)
- **Seasonal pattern:** Higher completion counts tend to cluster in late summer through fall (August–November), aligning with potential academic cycles or onboarding periods.
- **Recent decline:** Following a strong performance in late 2023, 2024 shows a notable drop
    - Only one completion in June 2024
    - Multiple months earlier in the year, with zero completions
- **Baseline consistency:** From December 2022 through mid-2023, completions typically hovered around 2 per month, indicating a steady engagement floor.
- **Low absolute numbers:** Monthly completions range from 1 to 7



This line chart tracks the average number of unique slides experienced per user each month over 26 months. A slight negative trend line (slope: –0.01 slides/month) is present, but it's minor.

- **High variability:** Month-to-month averages fluctuate considerably, with engagement levels ranging from:
    - Lows:
        - May 2024: 18.2 slides
        - May 2023 & January 2025: ~21 slides

    - Highs:
        - January 2024: 30.4 slides
    - December 2024: ~29.8 slides
- **Seasonal patterns:** Engagement appears to follow a cyclical pattern, with:
    - Peaks typically in mid-year (June–July) and winter (December–January)
    - Dips occurring in May (both 2023 and 2024)
- **Recovery after dips:** Following notable lows—especially May 2024's drop to 18.2 slides—average engagement bounces back strongly, reaching nearly 30 slides by December 2024 before a decline again in January 2025.
- **Overall engagement:** Despite the downward-sloping trend line:
    - Monthly averages generally hover between 25–30 slides
    - Given a maximum of 34 slides, this indicates strong content completion rates
    - The long-term average remains robust (~26–27 slides/user)

Interpretation:

Although labelled with a declining trend, the visualization reveals cyclical—not deteriorating—behaviour. User engagement remains resilient, with consistent recovery following predictable seasonal dips. Most users continue to consume a substantial portion of the content, suggesting effective design and sustained interest. This pattern may reflect academic rhythms or recurring professional development cycles.