CALBRIGHT

Data Analysis Tutoring Effect Evaluation

Term: 2025Q1

Masae Wen
Mentor: Bryan Queme
8/14/2025

Thank you all for being here today. I'm here to present the findings from my Data Analysis Peer Tutoring Effect Evaluation.

Agenda

❖ Executive Overview

    ➢ The Question

    ➢ The Answer

    ➢ The Evidences

    ➢ The Insights

❖ Comprehensive Statistical Overview (Appendix)

CALBRIGHT
COLLEGE

I'll be walking you through the executive overview of the analysis, covering the key questions, findings, and strategic insights. For those interested in the underlying statistical details, a comprehensive overview is available in the appendix.

Executive Overview

Agenda

- ❖ Executive Overview
  - ➢ The Question
  - ➢ The Answer
  - ➢ The Evidences
  - ➢ The Insights

CALBRIGHT
COLLEGE

The Big Question: Does Tutoring Work?

- **Our Question:** Are students who use our tutoring sessions making more progress than those who don't?

- **Our Goal:** To understand the effect of tutoring on student progress
  - Key metric of learner success: competency completion (0 - 12)

- **Data Used:** Cleaned Gradebook and Cleaned Post Tutor Form
  - Data Range: 2025 Q1
    - Overall: 2558 observations
    - Tutored Group: 137 observations
    - Untutored Group: 2421 observations

CALBR★GHT

We wanted to answer a simple but important question:
Are students who use our tutoring sessions making more progress than those who don't?
We decided to focus on competency completion as a key metric of learner success, since we have total of 12 competency to complete, the number can span from 0-12.

Using cleaned gradebook and post-tutor forms (also knowns as Post Session Support Form), we analyzed data for over 2,500 students in the Data Analysis Program. The analysis covers the first quarter of 2025, and as you can see on the slide, this includes 137 students in the tutored group and 2,421 in the untutored group. > Questions?

Agenda

- ❖ Executive Overview
  - ➢ The Question
  - ➢ **The Answer**
  - ➢ The Evidences
  - ➢ The Insights

CALBRIGHT
COLLEGE

The Answer is a Resounding YES!

Finding 1: Tutored students complete more than DOUBLE the coursework.

- Students who receive tutoring complete on average **2.16 times*** as many competencies as students who don't.
    - *95% Confidence Interval (CI): [1.56, 3.01]
    - p < 0.00001 (highly statistically significant)

Finding 2: Tutoring Appears Critical for Student Engagement

- **77%** of untutored students completed zero competencies (untutored: 2421)
- **33%** of tutored student completed zero competencies (tutored: 137)

The Answer is a Resounding YES!

Our analysis uncovered two key findings that show tutoring has a powerful, positive effect.

Finding 1: Tutored students complete more than DOUBLE the coursework.

- WIth appropriate statistical testing and model fitting, we found out that a student who receives tutoring completes **2.16 times** as many competencies as a student who doesn't. This shows tutoring significantly boosts overall progress.

Finding 2: Tutoring seems crucial for getting students engaged.

- A major challenge is students who don't complete anything. We found that **77%** of untutored students completed zero competencies.
- With tutoring, that number plummets to just **33%**. Tutoring is extremely effective at preventing total disengagement.
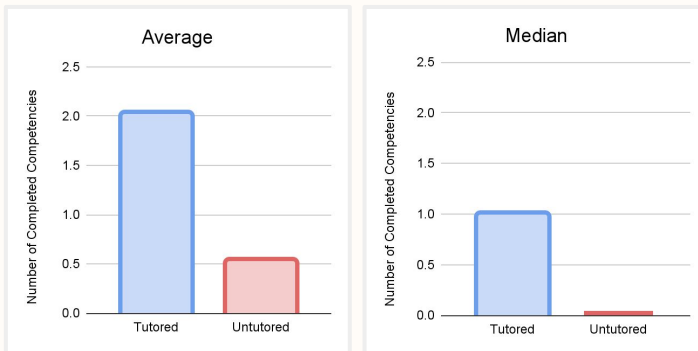
Agenda

- ❖ Executive Overview
  - ➢ The Question
  - ➢ The Answer
  - ➢ The Evidence
  - ➢ The Insights

CALBR★GHT
COLLEGE

A First Look: A Powerful Initial Effect

Competency Completion

Average Completions:
2.0 (Tutored) vs. 0.5 (Untutored)

Median Completions:
1.0 (Tutored) vs. 0.0 (Untutored)

Initial Conclusion:
The raw data shows a strong positive effect, but its complexity requires a deeper analysis for an accurate measure.

CALBRIGHT
COLLEGE

When we take a first look at the raw data, the effect of tutoring appears incredibly strong. On average, tutored students complete about **2 competencies**, while untutored students complete only **0.5**. This simple comparison suggests a nearly four-fold increase.

However, we know that raw averages don't always tell the full story, especially with complex student data. To ensure our conclusions are accurate, we needed to dig deeper.

How Tutoring Changes the Path to Completion
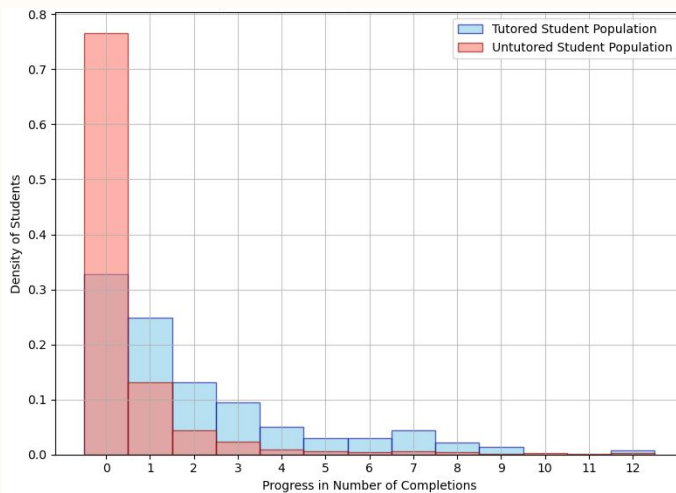
Untutored group (in red)

- Majority at **Zero Completions**.

Tutored group (in blue)

- Many students showed **1 or more Completions**.

So, the initial averages look promising. Next, we looked at the overall distribution of the data to understand the student journey. This graph shows a substantial difference. The untutored group, in red, is heavily concentrated at zero. The tutored group, in blue, shows a much healthier path toward completion. This visual confirms a strong positive effect, but it also reveals the data's complexity, it's not a simple bell curve. To prove this effect is statistically significant and to measure it accurately, we needed to use more advanced methods. > Questions?

Measuring the Impact: A Rigorous Approach

Method 1: Confirming the Difference

- A **Wilcoxon Rank-Sum test** confirms the group difference is **highly statistically significant** (p < 0.00001).

Because of the data's complexity, our analysis involved two key steps. First, to prove the difference we saw wasn't just random chance, we used a robust statistical test called a Wilcoxon test, which confirmed the difference is highly significant. > Questions?

Note: confirmed both mean and median differences are highly significant. The median is particularly reliable here due to data's excess zeros and skewed distribution.

Rationale: Non-parametric test chosen due to highly skewed and non-normal count data, robustness to outliers, and ability to handle unequal sample sizes.

A Rigorous Approach to Measuring the Impact

**Method 2: Quantifying the Effect Size**

- A **Zero-Inflated Negative Binomial (ZINB) model** was used to accurately measure the impact, accounting for the data's excess zeros and skewed distribution.

- ✅**Finding 1 Confirmation:** Tutored students complete **2.16 times*** as many competencies as untutored students.
    - *95% Confidence Interval (CI): [1.56, 3.01]
    - $p < 0.00001$

CALBRIGHT

Second, to calculate size of the effect, we used a specialized model called a Zero-Inflated Negative Binomial (ZINB) model. This is the proper-practice approach for this type of data with excess zeros and skewed distribution. This rigorous model, which gives us our more accurate and defensible outcome, led to finding one: tutoring leads to students completing **2.16 times** more competencies. > Questions?

     Rationale: A Zero-Inflated Negative Binomial (ZINB) model is a statistical model used for analyzing count data that exhibits both excess zeros and overdispersion (variance greater than the mean).

Tutored Students Show Drastically Less Disengagement

Tutored

- No Progress
- Progress

33.3%

66.7%

Untutored

- No Progress
- Progress

23.5%

76.5%
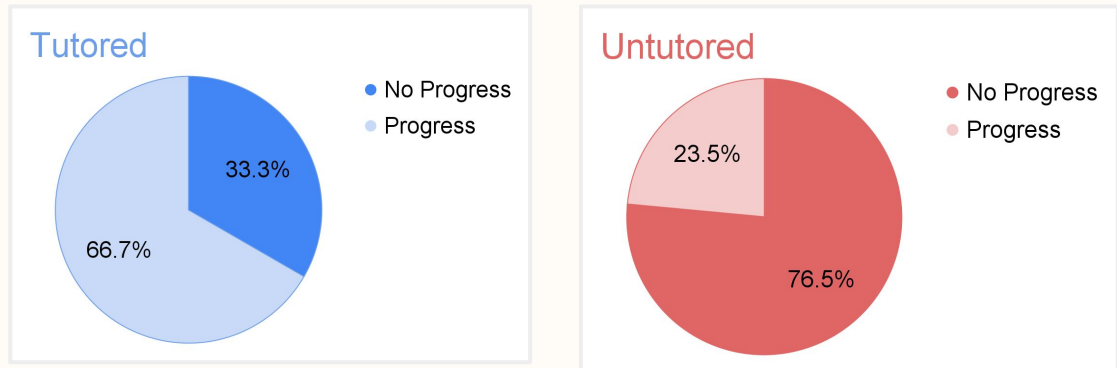
No Progress: Number of Completed Competencies = 0

Progress: Number of Completed Competencies > 0 (1 or more)

✅**Finding 2 Confirmation: 33%** No Progress (Tutored) vs. **77%** No Progress (Untutored)

CALBRIGHT

13

This visual highlights tutoring's impact on getting students started and completing at least one competency. The direct visual evidence of Finding 2.

- A major program challenge is initial disengagement. **77%** of students who don't get tutoring end up completing zero competencies.
- Tutoring is a powerful solution. For students who receive tutoring, that number is cut by more than half, with only **33%** completing no competencies.

The takeaway: Tutoring is a key strategy to ensure students engage with the material and make critical first steps toward completion.

Agenda

- ❖ Executive Overview
  - ➢ The Question
  - ➢ The Answer
  - ➢ The Evidences
  - ➢ The Insights

CALBRIGHT
COLLEGE

Our Findings Are Robust and Reliable

Robust Analysis: Advanced statistical methods ensure accurate and reliable conclusions.

Strong Evidence: Differences are not due to chance and are statistically significant.

Clear Impact: Tutoring is a major factor in observed improvements.

Bottom Line: Tutoring Works!

The data confirms tutoring more than doubles competency completions

CALBRIGHT
COLLEGE

Our finding are robust and reliable because...

- We used advanced statistical methods to ensure our conclusions were accurate and reliable.

- The differences we found are not due to chance; there is **strong statistical evidence** that tutoring is driving this positive change.

- While we know motivated students might be more likely to seek help, the improvement we see is so large that the tutoring itself is clearly a major factor.
- **Bottom Line: The Data Confirms Tutoring Works.**
    - **tutored students complete more than double the competencies of untutored students!**

Identifying Key Student Groups

Tutoring Impact on Student Progress	Tutored	Untutored
Made Progress (≥1)	67% Successful Engagement	23% Self-Starters
Made No Progress (0)	33% Opportunity Group "Quick Win"	77% Challenge

Tutoring Success:
67% of tutored students achieved progress.

Challenge:
77% of untutored students made no progress.

The "Opportunity Group":
33% of tutored students made no progress, showing a quick win chance.

CALBR★GHT
COLLEGE

Now that we've confirmed the impact of tutoring, this table helps us move from analysis to strategy by breaking our student population into four groups. We have our successes in green, and our long-term 'Challenge' in red.

But our most actionable insight is here in yellow: the **'Opportunity Group.'** This is the **33% of students who have already taken the crucial step of seeking help but are still stuck at zero**. Because we're already in contact with these motivated students, they represent our best chance for a quick and targeted win.

Based on this insight, we have two key recommendations. > Questions?

What This Means for Us: Actionable Insights

Tutoring is a Proven Success: Effective tool for student progress.

Recommendation 1: Use Tutoring as a **Targeted** Intervention.

- Proactively support the **"Opportunity Group"** with personalized outreach.
- Implement an onboarding survey to identify **"at-risk" students** even earlier.

Recommendation 2: Promote and Expand Tutoring.

- Increase awareness and accessibility for all students via Canvas and Email.

.

CALBR★GHT
COLLEGE

As the data has clearly shown, **tutoring is a proven success** and one of our most effective tools for student progress. To build on that success, we propose two key recommendations.

First, and most importantly, we need to **Use Tutoring as a Targeted Intervention** for that 'Opportunity Group.' We can empower our tutors to perform personalized outreach during their downtime, checking in on students who attended once or twice but have since disappeared. From my own experience as both a tutor and a student, a personal encouragement message can make a huge difference in preventing disengagement, and this is a task tutors can perform without needing support from other departments.

To get ahead of the problem, we can also identify at-risk students *before* they fall behind. By implementing a **brief survey after the mandatory onboarding session** about their background and comfort with data analysis and AI tools, we can strategically invite them to PASS sessions and other support services right from the start. This aligns with Greg's interest in creating pilot groups for PASS recruitment and would be a highly effective way to do so.

Secondly, we must continue to **Promote and Expand Tutoring** for all students, leveraging the proven effectiveness of Canvas announcements and email notifications to reach the entire student body, which are currently led by the instructors and TLC staff. > Questions?

Additional Tutoring Impact: Supports Our Instructors

- Effective tutoring reduces repetitive questions for instructors.
- Frees up instructor time for faster grading and feedback.
- Leads to higher overall student satisfaction.

.

CALBRIGHT
COLLEGE

Beyond the direct impact on students, tutoring service have an important indirect benefit. We know a key student complaint can be slow grading times. Effective tutoring reduces the number of common questions instructors have to answer, freeing up their valuable time to focus on grading and providing quality feedback. This can lead to faster turnaround times and higher overall student satisfaction with the Calbright program

Looking Ahead: Next Steps

- Continue this analysis for future terms to track ongoing impact.

  - Identify "Opportunity Group" students sooner than Quarterly analysis if possible.

- Explore applying this analytical framework to other Calbright programs.

.

SMALL STEPS ARE STILL PROGRESS

Finally, to ensure we continue building on this success, we recommend conducting this analysis for future terms, ideally perhaps in shorter intervals than quarterly if possible, to identify opportunity group students sooner. We should also explore applying this powerful analytical framework to other Calbright programs to measure and improve their effectiveness across the college.

Acknowledgement

- **Byan Queme:** Data Analysis Program Mentor

- **Richard Garcia:** Exll Intern Supervisor

First, I'd like to thank my mentor, Bryan. His expertise and guidance were invaluable in navigating the statistical analysis and its interpretation. I especially appreciate the time he dedicated to helping me select and optimize the appropriate statistical model, which ensures the integrity of the findings I am presenting today.

I would also like to thank my supervisor, Dr. Garcia, for the opportunity to perform this analysis as part of my internship. I'm grateful that he shared with me what kind of analysis is currently performed for tutoring and TLC services in general. This allowed me to pivot our project's focus from a redundant attendance measurement to an analysis of tutoring's effectiveness, which I believe adds additional value to our understanding of the tutoring's benefits.

Thank You!

Questions?

CALBRIGHT
COLLEGE

Appendix

Comprehensive Statistical Overview

Evaluating the Impact of Peer Tutoring on Competency Completion in the Calbright Data Analysis Program

A Statistical Analysis Using Zero-Inflated Negative Binomial Regression

CALBRIGHT
COLLEGE

Agenda

❖ Comprehensive Statistical Overview

➢ Introduction

➢ Exploratory Data Analysis

➢ Modeling

➢ Conclusion

➢ Future Direction

CALBRIGHT
COLLEGE

Introduction & Research Question

- **Objective:** To evaluate the effect of a peer tutoring program on student competency completion within the Calbright Data Analysis Program.

- **Dataset:** The analysis utilizes cleaned gradebook data, separating learners into two groups: those who received tutoring (n=137) and those who did not (n=2,421).
  - Data Range: 2025 Q1
    - Overall: 2558 observations
      - Key Identifier: Calbright Student Email
  - Tools: Google Sheets, Python (Jupyter Notebook)

- **Response Variable:** The primary outcome is Completion Count Progress (Competency Completion), an integer count variable ranging from 0 to 12.

CALBRIGHT
COLLEGE

Exploratory Data Analysis & Modeling Rationale

Data Characteristics: The response variable exhibits properties that violate assumptions of standard linear models:

- **Count Data:** The outcome is a non-negative integer.

- **Excess Zeros:** A large proportion of students have zero completions, particularly in the untutored group where the median is 0 and the 75th percentile is also 0. For tutored students, the median is 1.

- **Overdispersion:** The variance in completion counts is substantially larger than the mean for both groups, as seen in the descriptive statistics.

CALBR★GHT
COLLEGE

Exploratory Data Analysis & Modeling Rationale

ECDF with Annotated Max Gap Between Groups
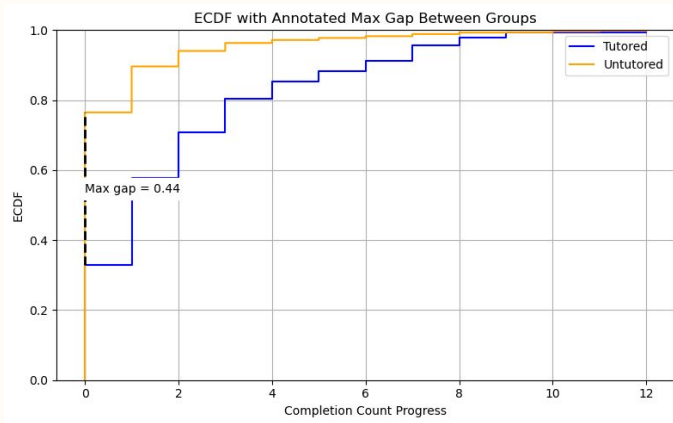
Visual Justification 1:

- The **ECDF plot** shows a maximum gap of 0.44 between the two groups at a completion count of 0, where 76.5% of untutored students have zero completions compared to only 32.9% of tutored students.

CALBRIGHT

Exploratory Data Analysis & Modeling Rationale
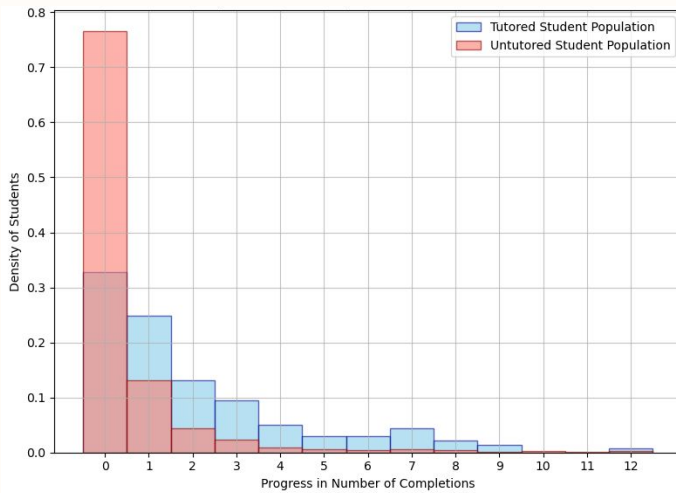
Visual Justification 2:

- A **density histogram** visually confirms the high concentration of zeros in the untutored population (~77%) versus the tutored population (~33%).

CALBRIGHT
COLLEGE

30

Exploratory Data Analysis & Modeling Rationale

Chosen Analytical Framework:

- A **Wilcoxon Rank-Sum test** confirmed a highly significant difference between the groups ($p < 0.00001$), supporting further investigation.

- Based on these characteristics, Negative Binomial (NB) and Zero-Inflated Negative Binomial (ZINB) regression models were selected as appropriate tools for the analysis.

CALBRIGHT
COLLEGE

Model Selection & Diagnostics

Systematic Comparison: A systematic model selection process was undertaken to identify the best-fitting and most stable model. Configurations tested included:

- **Model Type:** Standard Negative Binomial vs. ZINB.

- **ZINB Inflation Structure:** A full model (with 'Tutored' as a predictor in both count and zero-inflation components) vs. an intercept-only inflation model.

- **Optimizers:** lbfgs, bfgs, and newton were tested to ensure convergence.

CALBR★GHT
COLLEGE

Model Selection & Diagnostics

Selection Criteria: Models were evaluated based on Log-Likelihood, AIC, BIC, Pseudo R^2, and successful convergence without warnings (e.g., Hessian inversion failures).

Winning Model: The ZINB model with a full inflation structure, using the lbfgs optimizer and 100 iterations, was selected. It demonstrated superior fit over the standard NB model:

Model	Log-Likelihood	AIC	BIC
NB	-2470.26	4946.51	4964.06
ZINB	**-2456.79**	**4923.57**	**4952.81**

CALBR★GHT
COLLEGE

Final Model Interpretation: ZINB Results

The selected ZINB model (Full_Infl_LBFGS_Iter100) provides the following insights:

Count Model Component (Negative Binomial):

- The coefficient for **Tutored** is **0.7722 (p < 0.00001)**. The Incidence Rate Ratio (IRR) is e0.7722≈**2.16***.
    - *95% Confidence Interval (CI) of [1.56, 3.01]

- **Interpretation:** For learners not in the "always-zero" latent class, receiving tutoring is associated with a **116% increase** in the expected number of completed competencies, holding other factors constant.

CALBR★GHT
COLLEGE

Final Model Interpretation: ZINB Results

Zero-Inflation Component (Logit):

- The coefficient for **inflate_Tutored** is **-12.0146 (p = 0.879)**.

- **Interpretation:** The effect of tutoring on the log-odds of being in the "always-zero" group is not statistically significant. There is no evidence that tutoring affects whether a student belongs to a latent class of structural non-completers.

CALBR★GHT

Final Model Interpretation: ZINB Results

Dispersion Parameter:

- **alpha** is **1.9521 (p < 0.00001)**.

- **Interpretation:** The significant and positive alpha confirms the presence of overdispersion, validating the use of a Negative Binomial distribution over a Poisson.

CALBR★GHT
COLLEGE

Conclusion & Summary of Findings

- **Primary Finding:** Tutoring is strongly associated with higher competency completion. The model estimates that tutoring more than doubles the expected completion count for students who are actively participating in the course.

- **Secondary Finding:** The mechanism of this effect appears to be through increasing progress among active learners, not by changing the status of structurally disengaged learners.

- **Model Validation:** The data strongly supports the use of a ZINB model over simpler alternatives, as evidenced by the significant overdispersion parameter and the superior AIC/BIC fit statistics compared to a standard Negative Binomial model.

CALBR★GHT
COLLEGE

Limitations & Future Directions

- **Observational Data:** This analysis is observational, not experimental. Causality cannot be definitively established due to potential self-selection bias (i.e., more motivated students may be more likely to seek tutoring).

- **Future Work:** A more robust analysis could include controlling for covariates such as program start date, demographics, or prior academic history. A randomized controlled trial would be the ideal design to isolate the causal effect of tutoring.

CALBRIGHT
COLLEGE

Questions?

CALBR★GHT

COLLEGE

END