Closing the Gap in Higher Education: Factors in Student Success on CLEP Exams

Isabelle Petersen

Fordham College at Lincoln Center Honors Program

Advisor: Dr. Kei Kobayashi

May 1, 2025

I.   Introduction

When thinking about what barriers exist that prevent people from higher education, cost is one that seems to trump the rest. In 2024, about 59% of adults not enrolled in a postsecondary program say that they have considered pursuing some form of degree or credential. Of these unenrolled adults, 56% cited cost as a "very important" reason for not being enrolled, and a further 29% cited cost as "moderately important" (Gallup & Lumina Foundation, 2024, p. 13). Over the years, many programs have popped up—grants, scholarships, online universities—to help those who have all the time and effort needed to get a degree, but are missing the funds to do so. One of these is a nonprofit which, for the sake of anonymity, I'll be calling Online College Access Network (OCAN). OCAN is an organization founded by a businessman looking to have an impact in the field of education. At its core, the organization believes that education is a fundamental human right, and that a college education should be affordable and accessible to anyone who seeks one. To further this goal, Online College Access Network (OCAN) offers a variety of completely free online courses, and allocates vouchers for students to take College Board CLEP (College Level Examination Program) exams that can be transferred as college credit at accredited postsecondary institutions. Later, I will discuss what CLEP is and why it's important, but as for OCAN, the goal is to give every student the opportunity to obtain this college credit for free and to offer courses that prepare them for success on the exam. It is for this reason that I have chosen some aspects of a learner profile and aim to use them to answer the following research questions:

Question 1: To what extent does each of the chosen components relate to student success on the test?

Question 2: Given a particular student and their demographic information, what is the probability that they will pass? (or what is their projected score) How reliable is this prediction?

In this research paper, I will use a statistical analysis to attempt to find answers to these questions, and discuss my findings in relation to the impact they could have on improving CLEP scores for learners who use Online College Access Network.

II.   Background

Before talking about the significance of CLEP Exams, it makes sense to explain their place in the current landscape of post-secondary education. CLEP offers exams in 34 different subjects, including typical general education college classes such as College Composition and College Algebra. CLEP exam credits are transferable at 2,900 accredited institutions across the U.S., and there are over 2,000 testing centers throughout the country to offer access to anyone who can use it ("Benefits for Everyone", n.d.). But why is it significant that these courses are available online? As technology develops and we find new ways to communicate across physical divides, we are also finding new ways to teach across boundaries. Traditional 4-year university education is rapidly increasing in price, and more jobs are requiring some form of post-secondary education. A report from Georgetown University's Center on Education and the Workforce predicts that by 2031, 72% of jobs will require more than a high school diploma, up from 28% in the 1970s (Carnevale et al, 2023, p. 10). Federal financial programs such as the GI Bill and Pell Grant help students with lower-income or non-traditional backgrounds (e.g. students who have served in the military) to access education without the barrier of cost. However, federal funding is not able to reach every single person that could benefit from it, so in recent decades, programs have popped up that reshape the traditional post-secondary learning

experience. For example, public online institutions such as Southern New Hampshire University and Arizona State University Online give learners from all over the country access to a post-secondary degree at a reduced cost, including learners that fall outside of the expected demographic of college students (i.e. learners older than the age of 22, learners with full-time jobs, and learners in low-income households). Since early 2020, the ongoing COVID-19 pandemic and its resulting lockdowns have shown students, teachers, and parents alike that educational processes can adapt to the digital world, and specifically for universities, have exposed a wider student body to learning methods outside of traditional brick-and-mortar institutions. However, 2020 was certainly not the beginning of this shift: even in 2013, studies noted that "students today are more likely to be older, working, attending part time, and learning outside of traditional credit-bearing classrooms than students in the past" (Laitinen, 2013, p. 63). For the modern-day post-secondary learner, it's not always possible to follow the traditional pathway where it costs hundreds of thousands of dollars to get a degree and students must spend a minimum of 12 hours a week (and up to 20) physically in the classroom. OCAN seeks to address these challenges while still working within this existing system and playing off of its strengths.

This is where the CollegeBoard College Level Examination Program (CLEP) comes in. Learners can choose to take the CLEP exam either remotely through a CollegeBoard proctoring service, or in person at a CLEP testing center. Test takers receive their score immediately after taking the exam: out of 80 possible points, any score 50 or higher generally is considered a passing score ("Benefits for Everyone", n.d.), though individual institutions may set different acceptance criteria. In theory, CLEP exams are designed to be more accessible than other dual enrollment courses, as they are open to everyone regardless of background, and the fee to take

the test is approximately $95, a significantly lower cost than the charge per credit even at a community college or online institution. However, there are some learners for whom this test fee is still a barrier to their education, and this is the gap that Online College Access aims to bridge. When a learner registers and completes a course with OCAN, they can request a voucher to waive the fee for the CLEP exam. Every one of OCAN's online courses is completely free of charge and open to anyone, even learners outside of the U.S.

Learners have lots to gain from using a program like OCAN to take CLEP exams. Traditionally, dual enrollment programs like AP, IB, or other college credit programs have existed for the purpose of further accelerating high-achieving students by giving them the opportunity to get a head start on their postsecondary education, where they were already believed to be headed. However, as more schools attempt to include their low and middle performing students in the rigorous push, some educators find that "exposing lower achieving students to college early can improve their access to college and their success once they are there" (Bailey & Karp, 2003, p. 16). As college gets increasingly expensive and necessary, it's important for lower-income students to find ways to decrease the cost of a bachelor's or associate degree. CLEP exams offer just that: in a study that compared previous CLEP exam-takers and students that did not receive credit from CLEP, "CLEP students graduated in less time, enrolled in fewer semesters, maintained a higher GPA, and graduated with fewer credits when compared to a matched sample of students who did not pass the CLEP exam" (Barry, 2013, p. 14). Online College Access breaks down the final barrier towards higher education, as there are some learners for whom even the cost of a CLEP exam is not feasible and means that they are barred access to the benefits of transferable college credits.

III.   Descriptive Statistics

Being a program that seeks to close the gap that the barrier of cost creates, OCAN aims to reach learners from backgrounds that are traditionally underrepresented at the higher education level. I claim that Online College Access Network disproportionately serves learners from low-income families and people of color. Here, I provide a breakdown of student demographics in order to support this claim.

The data that I have consists of demographic data points and login data from six months of learner registrations. Upon registration, each user self-reports information about their age, gender, ethnicity, and income status. Measured against US Census Data, OCAN learners are made up of approximately 15% more non-white people (U.S. Census Bureau, 2023) and at least 13% more individuals below the federal poverty line (Guzman & Kollar, 2024, p. 16). 46% of learners are age 18 or younger, however as shown in Figure 1, learners as old as 86 have taken a course with Online College Access. Figures 1 and 2 offer a visual representation of the distributions of age, gender, ethnicity, and income. This displays that while not being exclusive to those who wouldn't typically be able to access higher education, as this program is open for use by anyone from any background, it is indeed reaching its intended demographic of those who need it most.

It is important to acknowledge that since this data is self-reported, many students were not able to answer the question about income. Since there is an option for "I don't know," I suspect that this is because a majority of learners are under the age of 18 and therefore did not have information about their family income status. Even ignoring these instances, it is evident that the largest proportion of OCAN learners come from low income families, with nearly 34%

of learners that self-reported disclosing income levels below the federal poverty line for a family of four.

Figure 3 displays a map of where students are located throughout the United States. The top three U.S. states where vouchers were requested were New York, Louisiana, and Tennessee. The reason I bring this up is because while a large portion of the vouchers and exams come from Louisiana, this state largely underperforms when compared to other states. There are a number of potential explanations for this discrepancy which do not necessarily fall into the scope of this paper, but for mathematical soundness, Louisiana is something that I will consider going forward as something to account for.

Another important thing to note is that outside of the US, there are a significant number of learners coming from places where higher education may be harder to access, such as Gaza and the West Bank. Figure 4 displays the top other locations, after the 50 U.S. States and District of Columbia, along with the portion of learners that reported these as their locations.

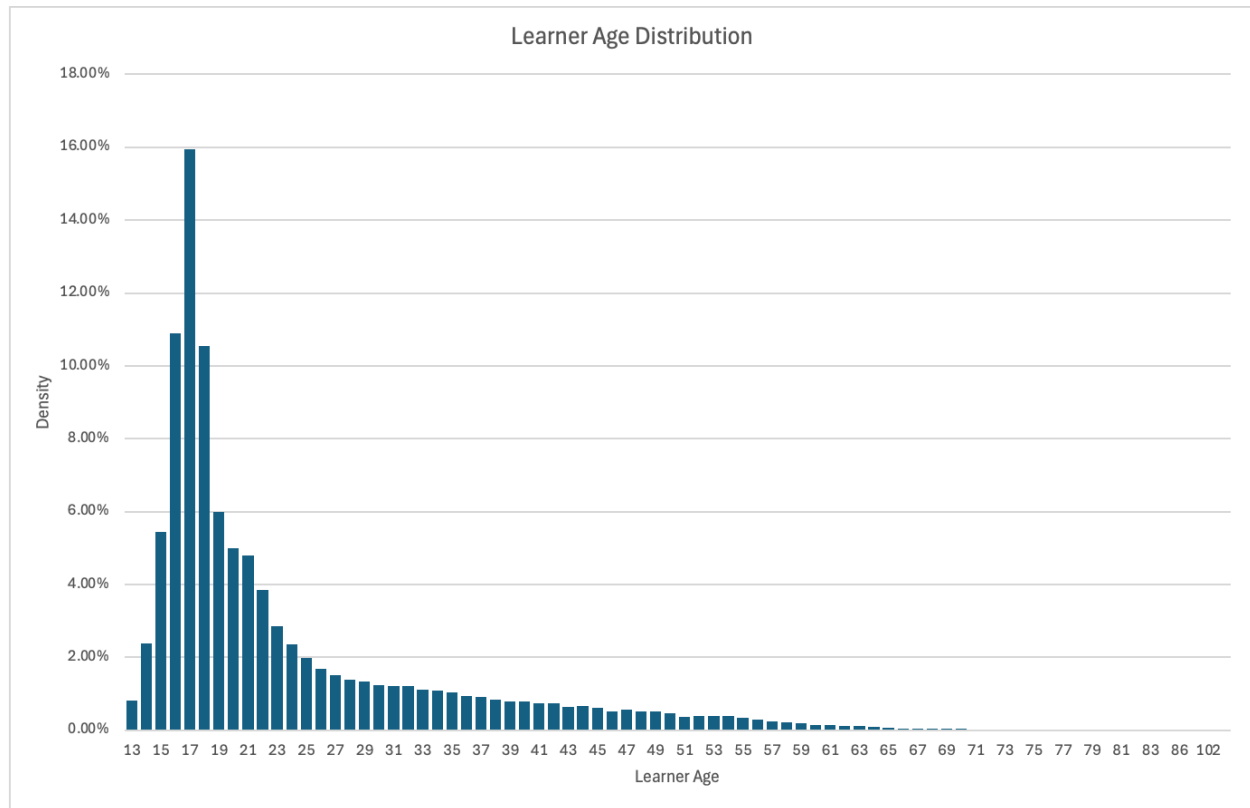*Figure 1*: *Learner Ages range from 13 to over 80, with the majority of learners between the ages of 14-18.*

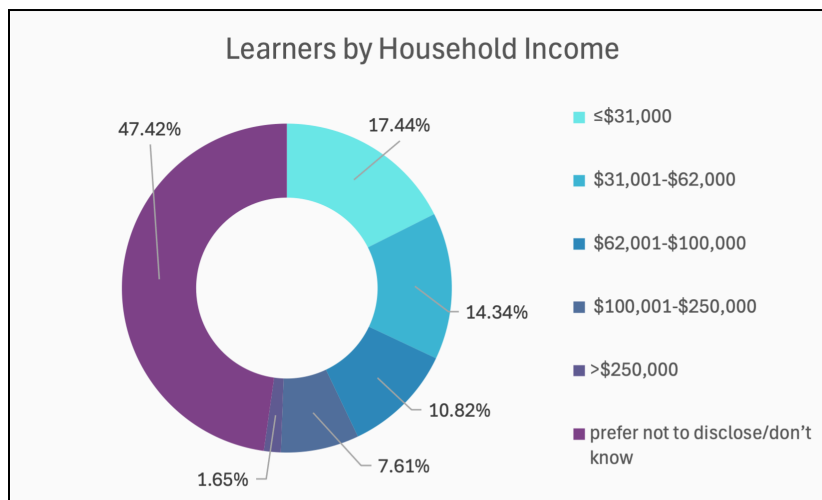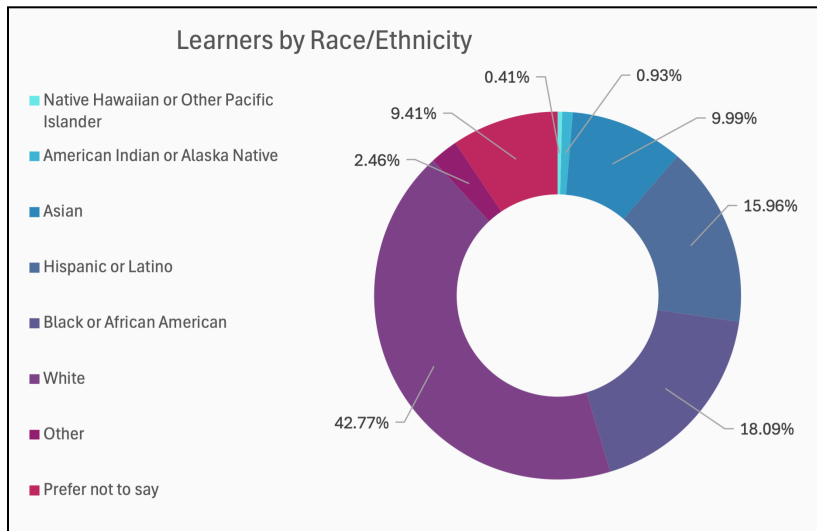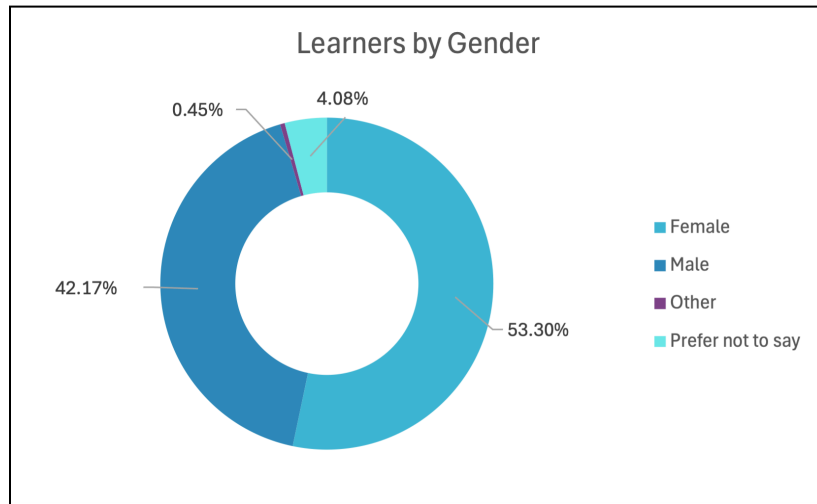*Figure 2*: *Learners by Gender, Race/Ethnicity, and Household Income*

Learners by Gender

4.08%
0.45%
53.30%
42.17%

- Female
- Male
- Other
- Prefer not to say

Learners by Race/Ethnicity

- Native Hawaiian or Other Pacific Islander
- American Indian or Alaska Native
- Asian
- Hispanic or Latino
- Black or African American
- White
- Other
- Prefer not to say

0.41%
0.93%
9.41%
9.99%
2.46%
15.96%
42.77%
18.09%

Learners by Household Income

47.42%
17.44%
14.34%
10.82%
1.65%
7.61%

- ≤$31,000
- $31,001-$62,000
- $62,001-$100,000
- $100,001-$250,000
- >$250,000
- prefer not to disclose/don't know

Figure 3: *Learners are present in every state in the US, but heavily concentrated in Louisiana, Florida, New York, Tennessee, and Texas*

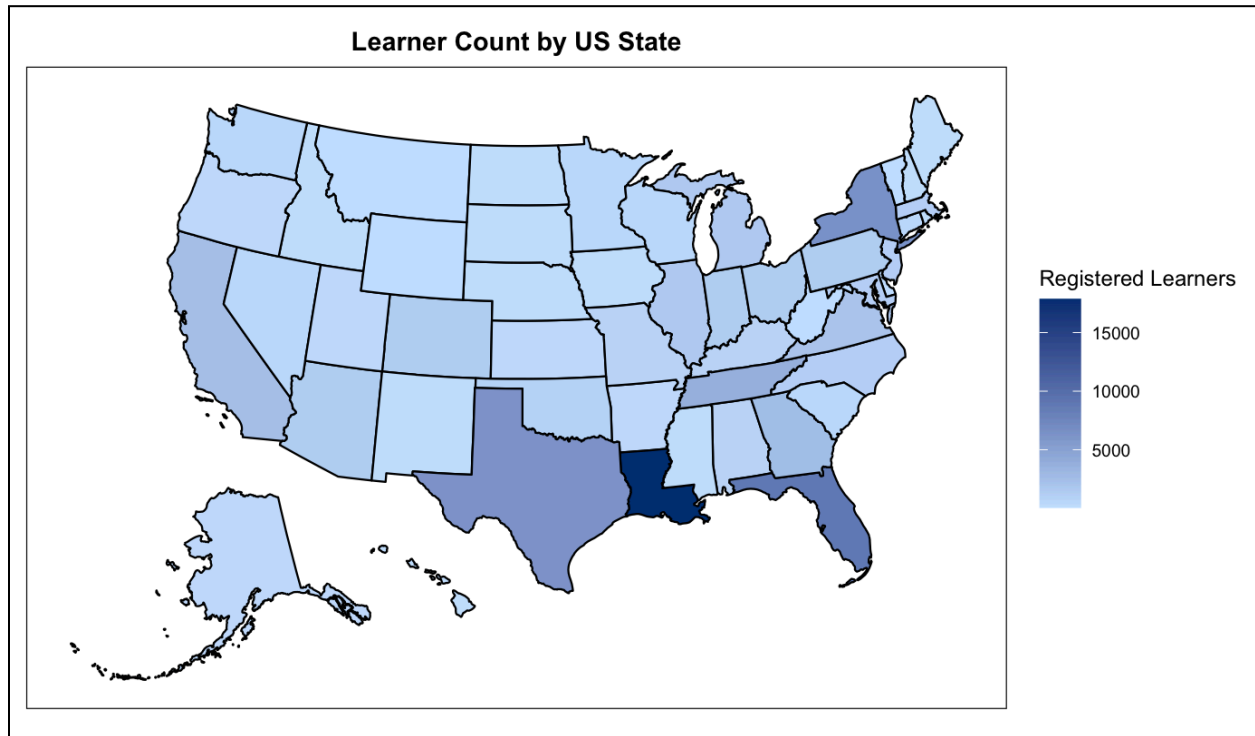**Learner Count by US State**

Figure 4: *Top 5 Non-U.S. Learner Locations*

State/Region	Number of Learners	% of Learners outside US
Palestine (including Gaza, West Bank, and other Palestinian territories)	200	9.17%
Ontario, Canada	115	5.27%
British Columbia, Canada	42	1.93%
Puerto Rico	34	1.56%
Alberta, Canada	31	0.40%

IV.   Methods

It is safe to assume that the majority of students who pass the course exam will also pass the respective CLEP exam, as the course exam has been designed as such. In order for OCAN to be effective in their work and not waste resources, it matters that students who take the CLEP exam succeed in passing. Therefore, Online College Access would like to be able to predict CLEP exam outcomes sooner than that—if a student is unlikely to succeed on the exam, OCAN can intervene and/or provide additional support resources that push the student towards a more successful exam outcome. In practice, this is something that OCAN already has begun doing: students who aren't performing as well in the online course cannot progress to the course exam until they revisit the content and retake the quiz. However, a challenge historically faced by analysts is how to determine whether a student is actually learning or just memorizing quiz content.

When a learner registers for an Online College Access course, they first sign up for an account and give OCAN an overview of their demographic information. Only once they have completed all of the unit quizzes and obtained an average score of at least 75% can a student take the course exam written by OCAN (note: this is not yet the CLEP exam—from here on out when I refer to "course exam" this is the test I am referring to). Once the student receives a score of at least 75% on the course exam, they may request a voucher to take the CLEP exam for free. Once the learner has taken the CLEP exam, CollegeBoard reports their score back to OCAN for record keeping.

As students progress through the course, OCAN receives information about their status at several crucial points. Specifically, there are three that I intuited might be important in this analysis: whether or not a student has taken a CLEP exam before, the student's quiz scores, and

time spent in content, that is, the time spent engaging with course materials. For each course, the content modules and curriculum look slightly different, so in order for consistency in this model, I set up variables to describe quiz scores and engagement as a z-score compared to all of the other students who took the course. This way, these variables describe a student's performance and engagement as either higher or lower than the rest of the learners on average.

To come to a concrete mathematical conclusion, I set up the following null and alternative hypotheses for the analysis:

H_0 (null hypothesis): None of these variables are related to testing outcomes for Online College Access learners who took CLEP exams.

H_1 (alternative hypothesis): At least one of these variables is related to testing outcomes.

My first step in setting up the analysis was to plot each independent variable on its own individual scatter plot against the dependent variable (CLEP exam outcome). Figure 5 displays the individual scatter plots for each of these variables. While there doesn't appear to be a clear relationship between each predictor variable and the outcome, I nonetheless decided to try a simple linear regression to test for correlation.

There was one predictor variable among the three that stuck out as being categorically different: the variable that describes whether or not a student has taken a CLEP exam before (I call it "Returning Learner") sorts learners into two groups: those who have seen a CLEP exam before and those who haven't. So, in addition to the regression, I decided to use a two-sample t-test to check for a considerable difference in scores between students who have taken an MS Course/CLEP exam before and those who haven't. Those who haven't are categorized with a 0, and those who have are categorized with a 1. Let *X* represent the scores of the population of first-time learners, and let *Y* represent the scores of the population of returning learners. Before

doing a t test, I must assume normality of these two populations. Figure 6 displays the distribution of scores for both groups—learners who have taken a CLEP exam before and learners who haven't. Seeing as both follow a rough bell curve, I assume the populations are approximately normal and proceed with the two-sample t-test.

If X describes the population of first-time learners and Y describes the population of returning learners, then μ_X is the expected score of a first time learner and μ_Y is the expected score of a returning learner. Then we have two sample means, one for each group: $\overline{X} = 53.558$ and $\overline{Y} = 58.118$. The t-test will determine whether or not it is reasonable to assume a significant difference in the averages of scores of the entire population of students who have/haven't taken a CLEP exam before.

The null and alternative hypotheses for this specific test are as follows:

H₀ (null hypothesis): There is no difference in average scores between the two groups of learners ($\mu_X = \mu_Y$).

H₁ (alternative hypothesis): Returning learners who have taken a CLEP exam before on average score higher than those who haven't ($\mu_X < \mu_Y$).

With a significance level of 0.01, the constructed critical region for t-values is (-∞, -3.514). The calculated t-value for these two populations is -10.157. Therefore, at this significance level, we reject the null hypothesis and conclude that returning learners on average score higher than first time learners.

Lastly, over the course of my research, I decided that the discrepancy between learners from Louisiana and the rest of the learner population is too significant to ignore. Therefore, I decided to once again split the learner population into two categories: those not from Louisiana (X') and those from Louisiana (Y'). Figure 7 shows the distributions of these two populations:

since both are roughly normal, I can proceed with a two-sample t-test. The respective sample means are $\overline{X'} = 58.4168$ and $\overline{Y'} = 48.3313$. Then I performed this t-test using the following null and alternative hypotheses:
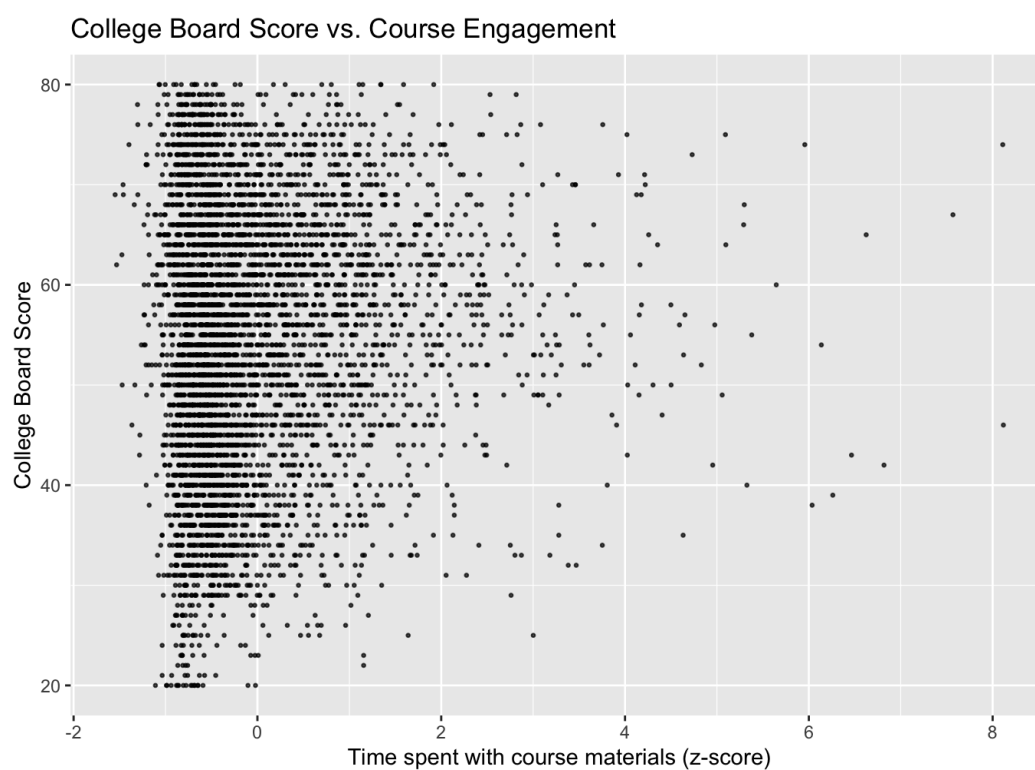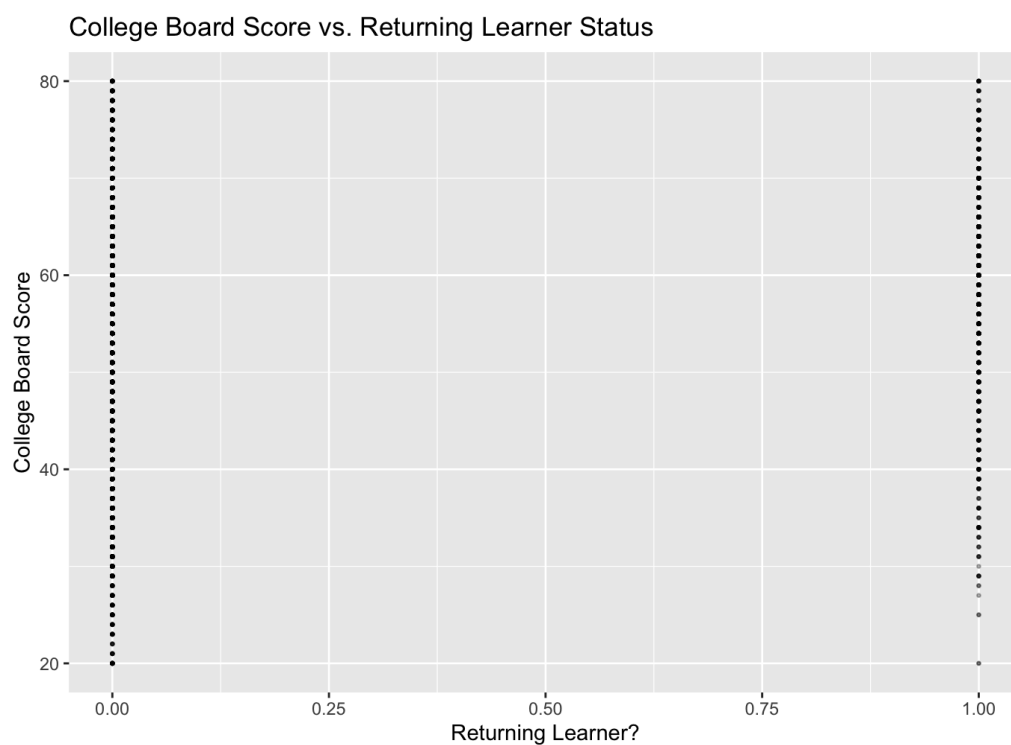
    H₀ (null hypothesis): There is no difference in average scores between the two groups of learners ($\mu_{X'} = \mu_{Y'}$)

    H₁ (alternative hypothesis): Learners from Louisiana on average score lower than those not from Louisiana. ($\mu_{X'} > \mu_{Y'}$)

    With a significance level of 0.01, the constructed critical region for t-values is (9.3028, inf). The calculated t value for these two populations is 29.984. Therefore, at this significance level, we reject the null hypothesis and conclude that learners from Louisiana score lower on average than learners outside of Louisiana.

    Based on these analytical processes, I was able to proceed and look at the results in order to draw concrete mathematical conclusions.

Figure 5: *Scatter plots displaying the predictor variables against the response variables do not denote a clear relationship between each of the predictor variables and the response variable.*
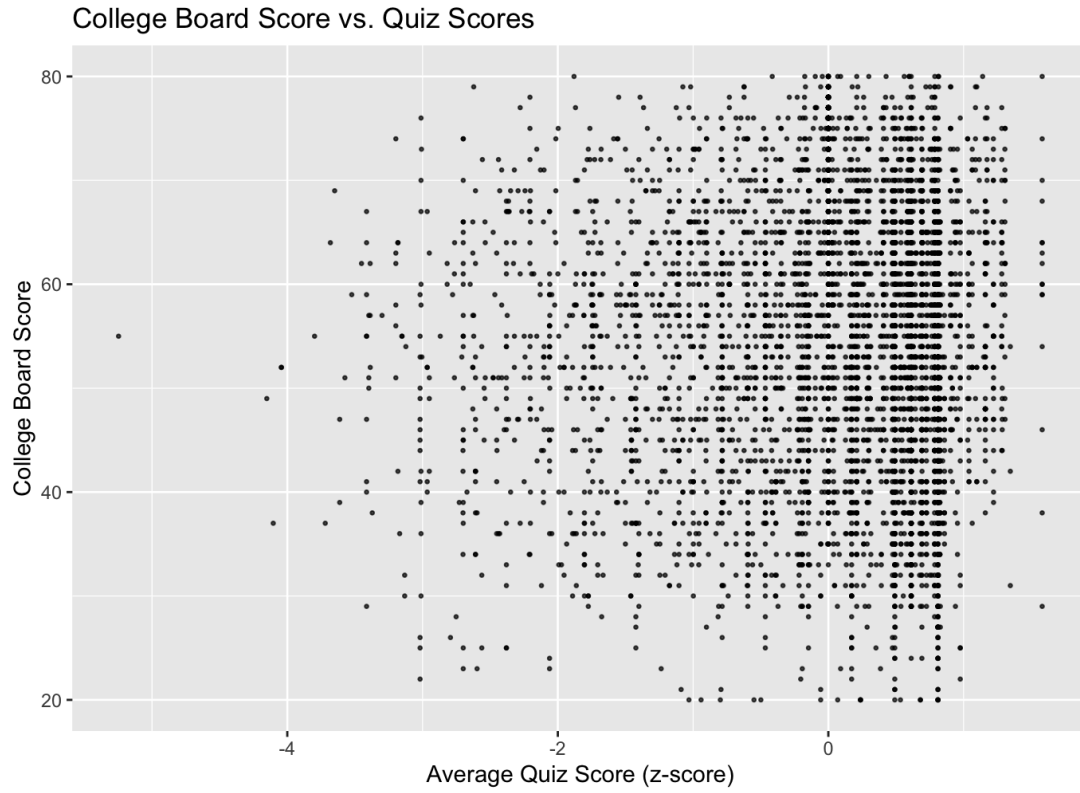
College Board Score vs. Returning Learner Status

College Board Score vs. Course Engagement

College Board Score vs. Quiz Scores

Figure 6: *Distributions of populations X and Y are roughly normal.*

Scores from first time learners (X)

Scores from returning learners (Y)
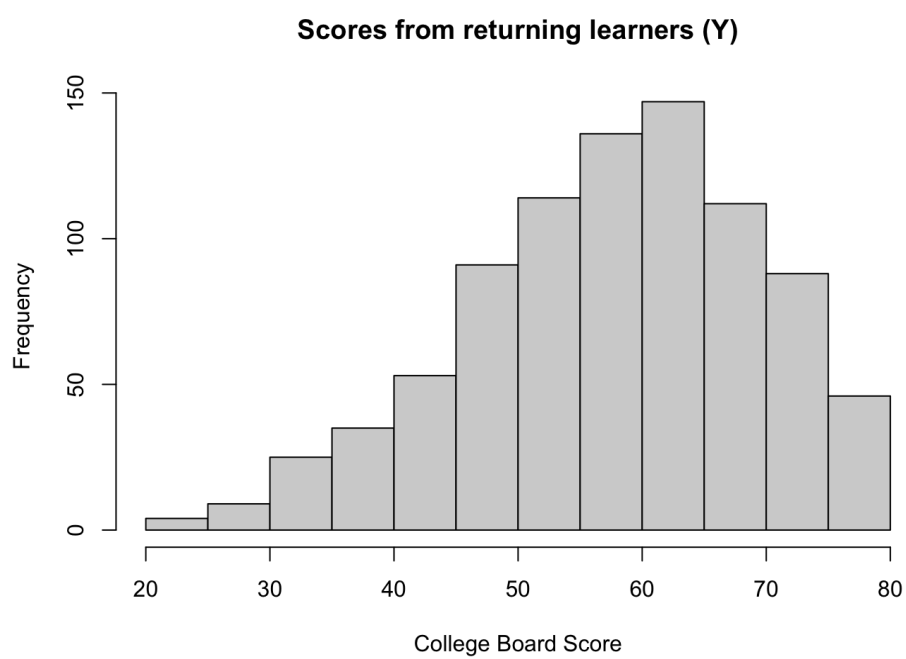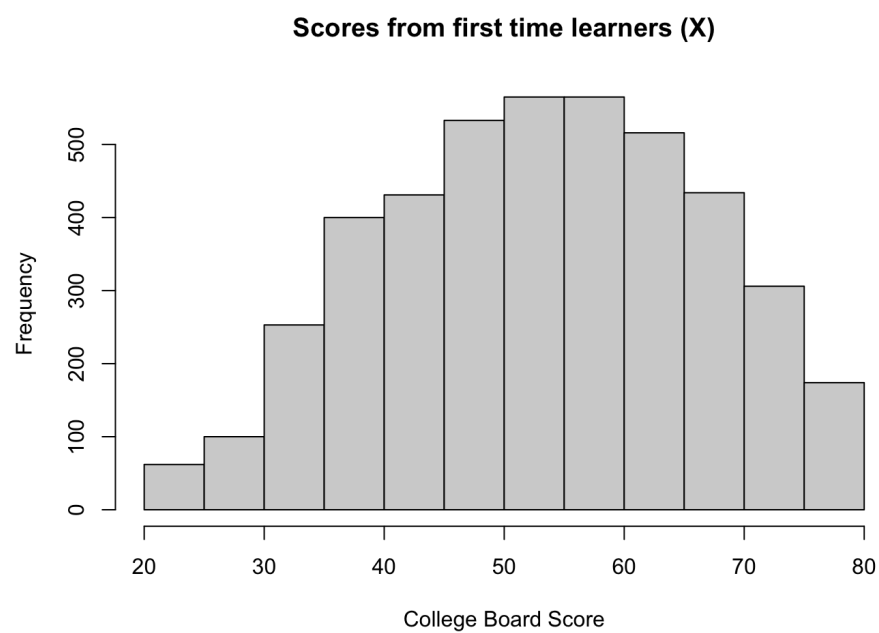
Figure 7: *Distributions of populations X' and Y' are roughly normal.*

Scores from learners outside of Louisiana (X')

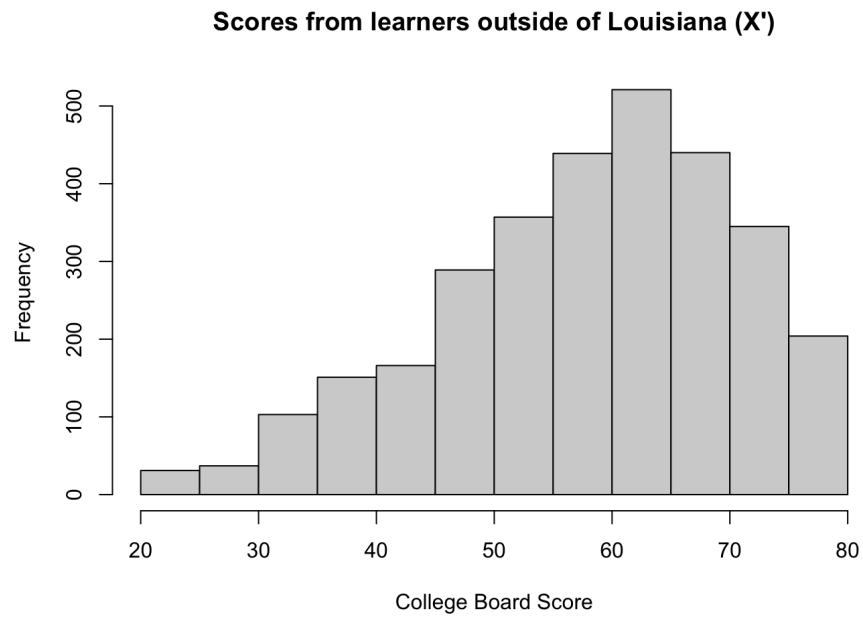**Scores from Louisiana learners (Y')**
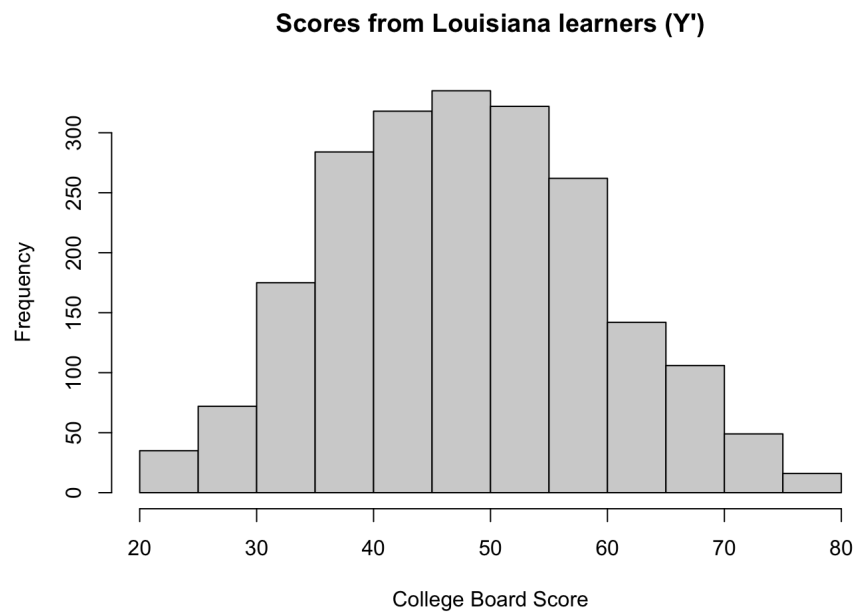
V.   Results/Discussion

For the individual regressions, the following linear equations were constructed, where $\widehat{y}$ is the predicted CLEP exam score, and the variables x, w, and z represent whether or not a student is a returning learner, their quiz scores as a z-score, and their engagement as a z-score, respectively:

Returning Learners: $\widehat{y} = 53.5577 + 4.5597x$

Quiz Scores: $\widehat{y} = 54.3120 + 0.4489w$

Course Engagement: $\widehat{y} = 54.3120 + 1.2946z$

Respectively, each had a correlation coefficient of 0.129, 0.033, and 0.099. Interpreted, I conclude that there does not appear to be a linear relationship between these three predictor variables and the CLEP exam score outcome.

For the multivariate regression using all three of these variables at the same time, I obtained the following equation:

Multiple regression equation: $\widehat{y} = 54.0837 + 4.3212x + 0.5274w + 1.4286z$

With a correlation coefficient of 0.163, I conclude that there does not appear to be a multilinear relationship between these three predictors and the CLEP exam score outcome.

With the inconclusive results of the preceding tests, I was not able to definitively reject or accept my null hypothesis that none of these variables are related to testing outcomes for learners. I was only able to conclude that there is not a linear relationship—there may very well be a different type of regression that better describes the relationship between these variables. My inconclusive results were part of what prompted me to continue with the t-tests, as being able

to definitively conclude a difference between these categories of students will mean deciding when to intervene for these learners.

For the t-test involving learners who have or have not taken a CLEP exam before, we rejected the null hypothesis and concluded that indeed first-time CLEP exam takers on average score lower than learners who have seen a CLEP exam before. For Online College Access Network, this means more intervention for first-time learners—practice questions that look more like the ones on the exam or simulated exam experiences may be one way to achieve this.

For the t-test involving learners from Louisiana, we rejected the null hypothesis and concluded that indeed students from Louisiana score lower on average than the rest of learners. This means that learners from Louisiana require further examination into their teaching practices and intervention techniques.

VI.   Limitations

While the data I have seems to be a fair representation of OCAN learners across the board, I ran into a few barriers over the course of my analysis that should be considered when drawing conclusions. First, this data came during a transition between learning management systems. While the learning management system afforded me the opportunity to have more pieces of information on each learner's demographic and activity during the course, there is missing information about learners' activity in the previous learning management system. Since the data covers about eight months in the new management system, there was enough information to consider students who took multiple courses and exams over that time, but there are likely a number of learners who were designated as first time learners that have in fact taken a course before in the previous learning management system.

Additionally, I was limited by the time and content constraints of this paper. In the time that I had, I was only able to perform and evaluate a linear regression analysis. With more time and a deeper understanding of advanced statistics, I would have tried to fit a different type of regression to the data in hopes of finding something that might better describe the relationship between the chosen predictor variables and response variable. It would be beneficial to consider other models that might better fit this data. With more time, it would also be beneficial to collect even more data, as trends in learner data may be skewed towards the beginning of the transition (such as the discrepancy in returning learners).

With the t-tests that I performed, I was able to reject the null hypothesis and conclude that Louisiana and being a returning learner are indeed related to exam outcomes. However, for OCAN, it would be more useful to know exactly how much (in terms of points) this variable affects the outcome. For example, we have concluded that learners in Louisiana score lower on average than learners outside of Louisiana, but as to how much lower, I have not been able to draw any conclusions. It would be useful to look further into the distribution of Louisiana learners' scores to determine how likely a learner from Louisiana is to pass (as the average score is below passing), and which other factors contribute to their likelihood of a passing score.

Finally, in general, I only chose four factors to look at as contributors to learner success. In reality, there are a myriad of other factors that could contribute to a learner's performance. Time spent engaging with the content outside of the online course with a teacher or tutor, time spent studying the course content before the exam, and even individual differences in learning styles and information retention are among the other factors that could influence learner performance. In further analysis, it would be interesting to consider these factors in conjunction

with the ones I've already considered in order to have a complete picture of the path to learner success.

VII.   Conclusion

Revisiting my two original research questions, I notice that the path of my research deviated slightly from my original plan. Recall the two research questions I began with:

Question 1: To what extent does each of the chosen components relate to student success on the test?

Question 2: Given a particular student and their demographic information, what is the probability that they will pass? (or what is their projected score) How reliable is this prediction?

Over the course of my research, I found that these questions were not so definitively answered given the limitations of my research. The four components I chose were not enough to determine a student's score—I suspect that there are other factors at play that should be taken into consideration. However, I was able to definitively say that being from Louisiana and never having seen test material before lead to lower scores on average, and this is a good starting place for OCAN to focus their intervention.

With my research as a starting point, scholars can continue to look into teaching and learning strategies that encourage learner success not just in the course and on the exam, but further along as learners pursue higher education at accredited institutions. What this may look like for now is extra tutoring resources for students who are underperforming. With the rise of AI and large language models, it's now possible for students to have a private AI tutor trained to explain content in a different way more suited to the learner, or provide the learner with additional study guide resources, flash cards, and practice quizzes. It may also be worth considering the effectiveness of course content for each individual course—if overall learners are

underperforming in a specific subject, perhaps content needs to be shifted to better reflect the CLEP exam content.

As time goes on, OCAN steadily gains more learners, and as a college education becomes more important and more people are accessing higher education via this route, it's important that OCAN can follow through on the claim that taking a course through them will reliably result in transferable college credit.

References

Bailey, Thomas, and Melinda Mechur Karp. Promoting College Access and Success: A Review of Credit-Based Transition Programs., 2003.

Barry, Carol L. A Comparison of CLEP® and Non-CLEP Students with Respect to Postsecondary Outcomes, 2013.

Carnevale, A. P., Smith, N., Van Der Werf, M., & Quinn, M. C. (2023). (rep.). *After Everything: Projections of Jobs, Education, and Training Requirements through 2031*. Washington, DC: Georgetown University Center on Education and the Workforce.

CLEP benefits for everyone. CLEP. (n.d.). https://clep.collegeboard.org/clep-benefits-for-everyone

Gallup, & Lumina Foundation. (2024). The State of Higher Education 2024 [Review of The State of Higher Education 2024]. https://www.luminafoundation.org/resource/the-state-of-higher-education-2024/

Guzman, G., & Kollar, M. (2024). Income in the United States: 2023. United States Census Bureau. https://www2.census.gov/library/publications/2024/demo/p60-282.pdf

Laitinen, Amy. "Changing the Way We Account for College Credit." Issues in Science and Technology 29, no. 2 (2013): 62–68. http://www.jstor.org/stable/43315720

U.S. Census Bureau. (2023). QuickFacts: United States. Census Bureau QuickFacts; United States Census Bureau. https://www.census.gov/quickfacts/