# Understanding Attrition Rates: Why the Drop-Out Rate Within Computer Science Degree Is so High

## Project Objective

This project aims to study the quality of courses within the undergraduate Computer Science degree and how it could affect the high attrition rates found in higher education. Although abundant resources exist for Computer Science students, the material can be difficult to learn due to its conceptuality and the student's learning environment. This study will examine the Computer Science courses and labs by surveying students to see if there are any underlying differences between Computer Science and other degree fields that might be holding students back.

## Project Background and Significance

In recent years, the drop-out rate of students within the Computer Science degree has been relatively high. Most of these drop-outs occur during the first two years of the degree program averaging a 40% drop-out rate and ranging from 30% to 60% depending on the university. This is significant because jobs requiring skills taught within the Computer Science curriculum are becoming more prevalent and in high demand, yet students are hesitant to pursue the field (Giannakoset 2). This may be due partially to the little exposure students from K-12 have to computing skills. Although initiatives like "CS for All" exist to help students in primary and secondary education systems acquire computing skills, the efforts are underdeveloped and not widely used. This dramatically affects the ability of incoming college students within Computer Science as they have little knowledge or background of computer technologies, limiting their growth (Jun 2). Due to this lack of exposure, incoming college students are unfamiliar with the concept of computational thinking, which employs the application of other concepts like "abstraction, problem decomposition, algorithms design, data collection and analysis, data representation, and simulation" (Jun 2). By familiarizing oneself with computational thinking, many challenges faced by incoming undergraduate students are alleviated, which could result in lower drop-out rates (Takács 12). Many approaches have been made to help improve Computer Science attrition rates in higher education by enacting curriculum adjustments or by adding supplementary lab courses. One method investigated was the implementation of group project presentations, where students collaboratively worked together to learn about conceptual Computer Science problems and presented their findings. This allowed the students to get active feedback from their instructors and peers, bolstering their confidence and their knowledge (Bakhry 4). Additional methods include using adaptive-learning programs online that aid students in how to code. The online programs allow the students access to instant feedback fostering quicker growth in their development of programming languages and being familiarized with the syntax (Anindyaputri, 1). Another study shows that there is also an impactful difference depending on the teachers' teaching styles and the students' learning styles. It shows that students would perform much better when the teacher provides multiple presentation styles of the same information (Chetty 5). However, many of these methods are not widely utilized and are primarily used in study cases. It is clear that there are variables that affect the student's ability to perform, which affects the attrition rates of Computer Science in higher education. That is why an analysis of students is necessary to help determine any avoidable determinants in the Computer Science field.

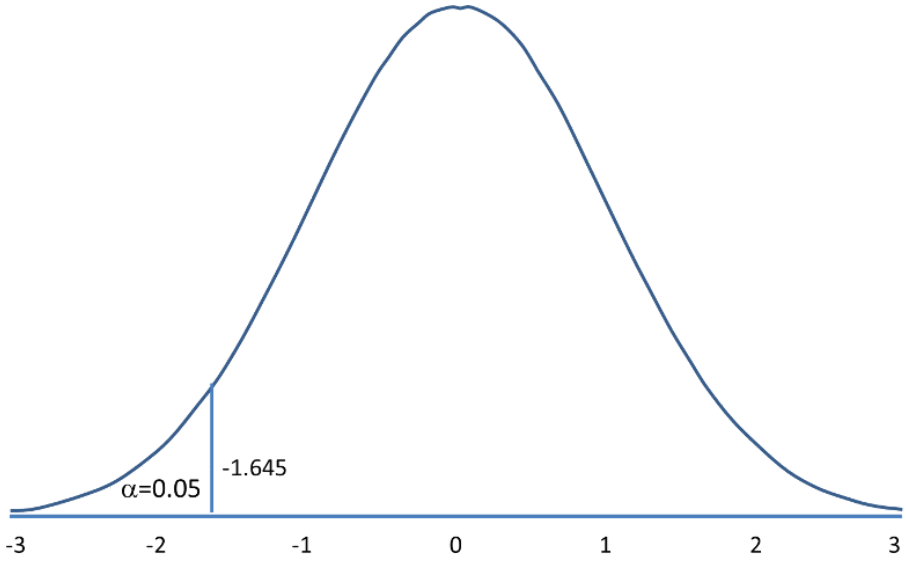
## Research Methods

This proposal will be conducted with a survey and a statistical analysis, examining random samples of students in Computer Science and observing what obstacles may be affecting them. One of the factors that students often cite is poor lecture quality, which can be measured by conducting a campus-wide survey. The survey will be conducted by utilizing a Google Form where students will rate their course experiences. From those that responded, one hundred students will be randomly selected to have the chance to win a $10 Amazon gift card. The advantage of sampling this data, as opposed to using a third-party source such as ratemyprofessor.com, would be that it avoids some voluntary response bias due to the incentivization of an award. In a conventional survey, a surveyor could only obtain data from people who responded willingly; however, they miss out on all the data from those who decided to abstain. Fortunately, with an award, students are more likely to fill out professor reviews, which means that the population sample that is gathered would be a more accurate representation of all students without being subject to any extreme voluntary response bias.

Continuing with this methodology, a unit of measurement is required. In this case, the reviews require each student to numerically rank, on a 1-10 scale, the quality of different aspects of each course. Some of these aspects would be lecture quality, professor explanations, material familiarity, the usefulness of labs, or homework fairness. These metrics would be tested to see if there is a significant drop in the average quality of C courses in comparison to other majors. In order to do this, the experiment would compile data and focus on two primary groups: Computer Science Courses and non-Computer Science courses. The Computer Science group would be the experimental group, while the non-Computer Science group would be the control group. Then, random samples of each group would be taken, and their means and adjusted standard deviations would be recorded. The optimal hypothesis test for this experiment would be a 1-tail, 2-sample null hypothesis test, with an alpha value of 0.05, to see if there is a statistically significant drop in ratings when comparing the C review samples with the non-Computer Science review samples.

## Expected Outcome

After performing the null hypothesis test, the findings would generate a p-value. Because the null hypothesis states that there is no difference between the sample groups, the findings would have to reject this notion to conclude that there is a difference between the groups. In order for this to happen, the p-value would have to exceed a z-score of -1.64 in order to hold any statistical significance. The p-value represents the probability, thus if the probability is lower than the z-score threshold, then that means that the difference observed between the random samples for the control and experimental group was so improbable that it could not have occurred by chance. The z-score threshold, in this case, would be -1.64 because the hypothesis test was a 1-tail test with an alpha value of 0.05. A z-score of -1.64 represents the 5th percentile in the normal model (see diagram below). This means that the p-value would have to fall in the lower 5% of all possible samples measured. The characteristic of a 1-tail test ensures that the upper and lower 5% are not measured, only the lower 5%. The alpha value of 0.05 represents the lower 5%, because 5% can be written as a probability of 0.05.




Once the testing is complete, the resulting p-value would determine whether rejecting the null hypothesis is appropriate. In order to get accurate results, the experiment would have to be repeated multiple times with different random samples. During this process, type I and type II errors would also be considered to correct any false assessment of the p-value. It is expected that the findings reject the null hypothesis in favor of the alternative hypothesis, which states that there exists a significant drop in quality in Computer Science courses. This would mean that there are obstacles that stand in the way of many Computer Science students, and that there would be a further need to counteract these obstacles and better the quality of the Computer Science curriculum. The findings would be published in a research article and shared with the University of Central Florida and other institutions.

## Literature Review

1. Jun, Andrew, et al. "CS for All: Introducing Computational Thinking with Hands-on Experience in College." *2017 International Conference on Computational Science and Computational Intelligence (CSCI)*, 2017, https://doi.org/10.1109/csci.2017.187.

2. Giannakoset, Michail N., et al. "Understanding Student Retention in Computer Science Education: The Role of Environment, Gains, Barriers and Usefulness." *Education and Information Technologies*, vol. 22, no. 5, 19 Oct. 2016, pp. 2365–2382., https://doi.org/10.1007/s10639-016-9538-1.

3. Anindyaputri, Natasha A., et al. "Enhancing Students' Ability in Learning Process of Programming Language Using Adaptive Learning Systems: A Literature Review." *Open Engineering*, vol. 10, no. 1, 2020, pp. 820–829., https://doi.org/10.1515/eng-2020-0092.

4. Chetty, Nithya D., et al. "Learning Styles and Teaching Styles Determine Students' Academic Performances." *International Journal of Evaluation and Research in Education (IJERE)*, vol. 8, no. 4, Dec. 2019, pp. 610–615., https://doi.org/10.11591/ijere.v8i4.20345.

5. Takács Rita, et al. "Successful Steps in Higher Education to Stop Computer Science Students from Attrition." *Interchange*, 18 Oct. 2022, https://doi.org/10.1007/s10780-022-09476-2.

6. Bakhru, Sunil A., and Rashesh P. Mehta. “Assignment and Project Activity Based Learning Systems as an Alternative to Continuous Internal Assessment.” *Procedia Computer Science*, vol. 172, 2020, pp. 397–405., https://doi.org/10.1016/j.procs.2020.05.073.

## Preliminary Work and Experience

As Computer Science majors, we have adequate experience regarding the quality of the Computer Science degree curriculum. Many incidents, such as professors giving inadequate lectures that are difficult to understand, make it very difficult for aspiring Computer Science majors. Additionally, some professors also give unforgiving assignments with strict grading policies. This creates a challenging learning environment for students as they feel abused. Quite often, these difficult classes provide little supplemental support enhancing the difficulty of the course and resulting in many students using third-party sources to get solutions to homework and assessments. The assignments are also unstimulating and monotonous. Many students are uninspired by the assigned problems and find difficulty being motivated as they do not see much real-world application for the problems. The laboratories associated with the Computer Science classes also provide a negative experience because many of the TAs do not appear appropriately trained and cannot provide adequate help. Although it is necessary to research and be disciplined independently, the Computer Science curriculum could give a better learning experience.

## IRB/IACUC statement

An IRB statement will be required as individuals will be surveyed

## Budget

$1000 - 100 $10 Amazon gift cards

Total Funding Amount: $1000