# An analysis of undergraduate major attrition rates within the department of computer science at CSUCI

**MATH 398 (Fall 2025)**

Monique Calatayud, Robert McEachron, and Morgan McMurray

# Table of contents



## Introduction

Despite ongoing efforts to increase diversity in the technology sector, computer science programs continue to be disproportionately represented by white males. Limited minority representation within the profession contributes to the exclusion of diverse perspectives in product design, which can result in algorithmic harm—technology-driven threats to civil rights and democratic values. Understanding the factors contributing to underrepresentation in computer science is therefore critical to fostering equitable and inclusive technological development. To the best of our knowledge, no such analysis has been conducted on students currently attending California State University Channel Islands (CSUCI). This is especially important considering CSUCI is a a Hispanic-Serving Institution.

Thus, the current study will examine demographics among student's enrolled in the Computer Science Program at CSUCI, while painting a holistic picture of the department. Additionally, factors predicting graduation rates among students in the program will expose problem areas necessary for intervention. The findings aim to help CSUCI and similar institutions design targeted interventions to support at-risk computer science students, ultimately contributing to a more diverse and inclusive technology workforce that better serves society as a whole.

# Methods

## Dataset

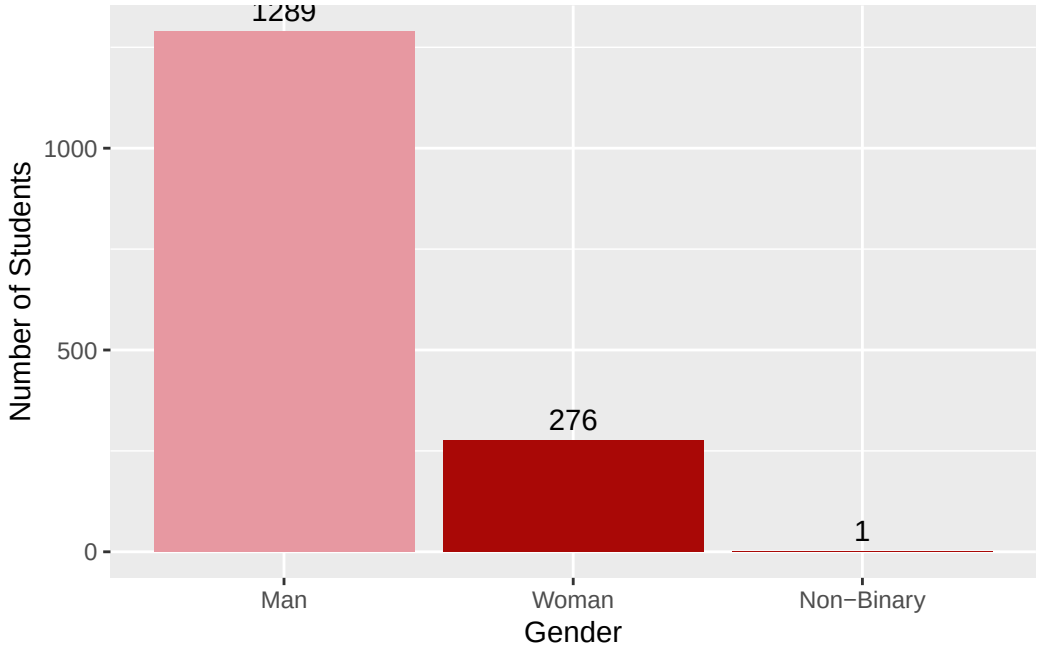
The data set given by the Office of Institutional Research included students enrolled in the Department of Computer Science at California State University Channel Islands (CSUCI). Data included 1,691 students enrolled between Fall 2015 and Fall 2025.

# Results

## Descriptive Statistics

### Gender

Computer science programs have historically under-represented women and non-binary populations. This same relationship can be seen from the data from Channel Islands. Due to the small sample size of non-binary ($n = 1$), their value was re-coded as being female for future analyses.




```
mdata <- ddata[ddata$Gender_Binary == 0, ]
fdata <- ddata[ddata$Gender_Binary == 1, ]
```

It was found that both males and females stay at Channel Islands roughly the same amount of terms with males staying 5.11 terms and females staying 5.59 terms on average. A t-test determined a non-significant difference between the two groups ($p > .05$).

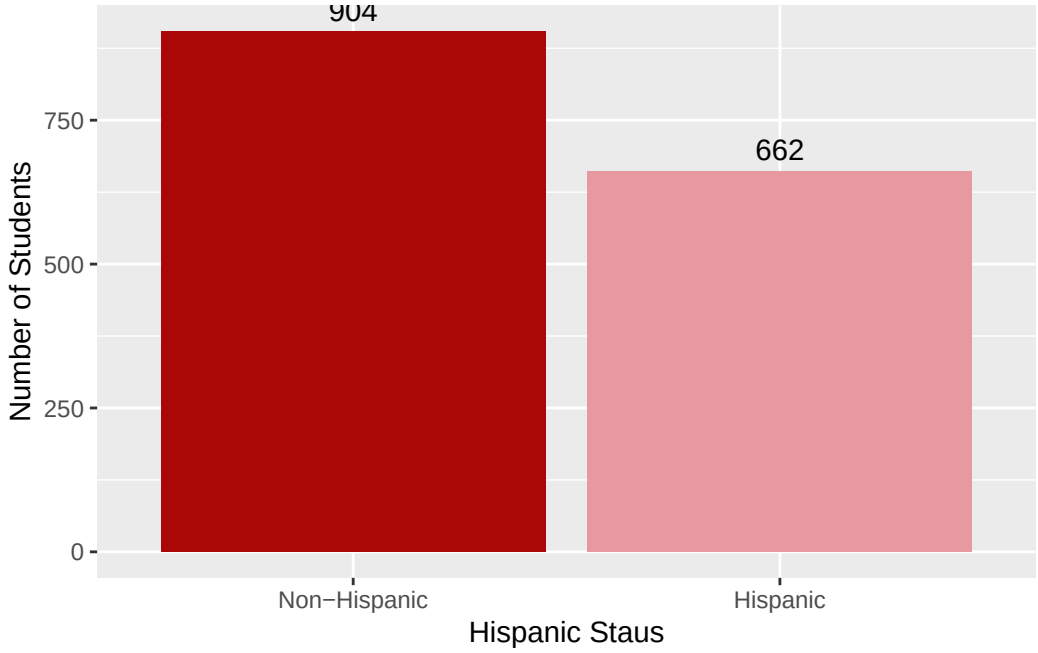
```r
mean(mdata$Terms_Attended) # Male Terms Mean
mean(fdata$Terms_Attended) # Female Terms Mean
t.test(mdata$Terms_Attended, fdata$Terms_Attended, alternative = "two.sided", var.equal = FA
```

Graduation rates were found to be non-significantly different ($p > .05$), with males graduating 44.76% of the time and females 42.24% of the time across all terms.

```r
mean(mdata$Did_Graduate_All) # Male Graduation Mean
mean(fdata$Did_Graduate_All) # Female Graduation Mean
t.test(mdata$Did_Graduate_All, fdata$Did_Graduate_All, alternative = "two.sided", var.equal =
```

### Hispanic Status

Due to Channel Islands being a Hispanic Serving Institution (HSI), exploration of student's Hispanic Status was important to us. The following chart showcases the makeup of Hispanic and Non-Hispanic student's in the Computer Science Department.




```r
nhisdata <- ddata[ddata$Is_Hispanic == 0, ]
hisdata <- ddata[ddata$Is_Hispanic == 1, ]
```

Student's from both identity statuses attended same amount of terms (*Winter, Fall, Spring, Summer*) with non-Hispanic students attending a mean of 5.638 terms and Hispanic students 5.370. A two sided t-test determined no significant difference between the two groups ($p > .05$).

```
mean(nhisdata$Terms_Attended) # Non-Hispanic Terms Mean
mean(hisdata$Terms_Attended) # Hispanic Terms Mean
t.test(hisdata$Terms_Attended, nhisdata$Terms_Attended, alternative = "two.sided", var.equal
```

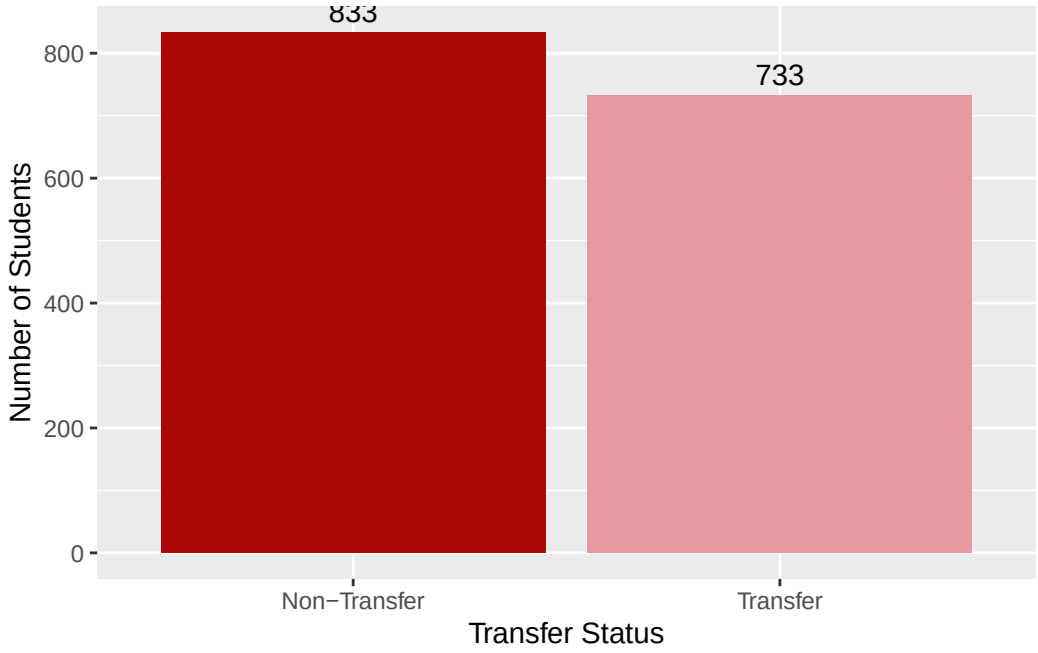
Hispanic student's were found to graduate 37.5% of the time, whereas their Non-Hispanic counterparts graduated 49.3% of the time. Further examination revealed a significant difference in graduation rate among the two groups, $t(1449) = 4.73$, $p < .001$.

```
summary(nhisdata$Did_Graduate_All) # Non-Hispanic Graduation Mean
summary(hisdata$Did_Graduate_All) # Hispanic Graduation Mean
t.test(nhisdata$Did_Graduate_All, hisdata$Did_Graduate_All, alternative = "two.sided", var.e
```

The intersection of the aforementioned findings indicate that Hispanic and Non-Hispanic students attend the same amount of terms, but Hispanic students graduate at a significantly lower rate.

### Transfer Status

Channel Islands receives lots of transfer student's due to programs in place at the local community colleges and the institutions relatively low cost. This is supported by the data we received below.




```
hdata <- ddata[ddata$Is_Transfer == 0, ]
tdata <- ddata[ddata$Is_Transfer == 1, ]
```

It was found that student's who started their academic journey at Channel Islands attend more terms ($M = 5.96$) as those who transfer ($M = 5.15$) with a significant difference, $t(1282) = 4.957$, $p < .001$. Interestingly, this difference was not as large as expected as those who transfer should theoretically need far less classes then those who start at the institution.

```
summary(hdata$Terms_Attended) # Home Terms Mean
summary(tdata$Terms_Attended) # Transfer Terms Mean
t.test(tdata$Terms_Attended, hdata$Terms_Attended, alternative = "two.sided", var.equal = FA
```

When taking into account graduation rates, the lack of disparity between terms attended for transfer and non-transfer student becomes apparent. Student's who started at Channel Islands graduate 35.61% of the time whereas those who transfer graduate 51.98% of the time. A t-test determined that this was a significant difference, $t(1552) = 6.61$, $p > .001$.

```
summary(hdata$Did_Graduate_All) # Home Graduation Mean
summary(tdata$Did_Graduate_All) # Transfer Graduation Mean
t.test(tdata$Did_Graduate_All, hdata$Did_Graduate_All, alternative = "two.sided", var.equal =
```

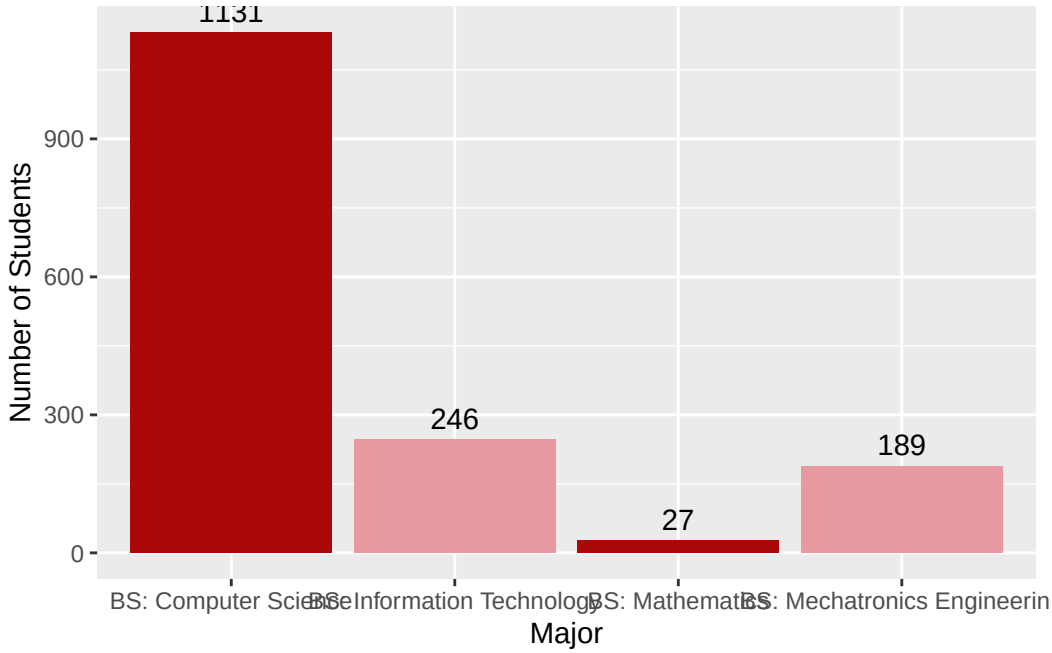
## Majors

The Computer Science Department at Channel Islands includes the majors of Computer Science, Mechatronics and Information Technology. Due to some student's double majoring or changing programs the data includes majors that are outside this category.

Our major count was:

| | Major | n |
|---|---|---|
| 1 | BA: Art | 1 |
| 2 | BA: Communication | 1 |
| 3 | BA: Economics | 2 |
| 4 | BA: Performing Arts | 1 |
| 5 | BA: Political Science | 1 |
| 6 | BA: Psychology | 4 |
| 7 | BS: Applied Physics | 1 |
| 8 | BS: Applied Physics, Tech | 2 |
| 9 | BS: Biology | 1 |
| 10 | BS: Business | 3 |
| 11 | BS: Chemistry | 2 |
| 12 | BS: Computer Science | 1131 |
| 13 | BS: Information Technology | 246 |
| 14 | BS: Mathematics | 27 |
| 15 | BS: Mechatronics | 1 |
| 16 | BS: Mechatronics Engineering | 189 |

Due to the small sample size, majors with under 20 student's were excluded for the following chart:




### Gender and Hispanic Status

It was found that Hispanic males graduate **39.44%** of the time, whereas their Non-Hispanic male counterparts graduate **48.54%** of the time. Similarly, Hispanic females graduate **29.13%** of the time and Non-Hispanic females graduate **53.33%** of the time. This showcases a large disparity between Hispanic and Non-Hispanic students. In addition, there is a ~**10%** difference in graduation rates among Hispanic males and females, whereas this disparity is not seen in their Non-Hispanic counterparts.

```
mhisdata <- mdata[mdata$Is_Hispanic == 1, ] # Male Hispanic
fhisdata <- fdata[fdata$Is_Hispanic == 1, ] # Female Hispanic

mnhisdata <- mdata[mdata$Is_Hispanic == 0, ] # Male Not Hispanic
fnhisdata <- fdata[fdata$Is_Hispanic == 0, ] # Female Not Hispanic

# Gender by Hispanic Status
  summary(mhisdata$Did_Graduate_All) # Male Hispanic 0.3944
  summary(fhisdata$Did_Graduate_All) # Female Hispanic 0.2913

  summary(mnhisdata$Did_Graduate_All) # Male Not Hispanic: 0.4854
  summary(fnhisdata$Did_Graduate_All) # Female Not Hispanic: 0.5333
```

## T-Tests

## Logistic Regression

### Predicting Graduation Rates

### Background of Logistic Regression

# Discussion