THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF MATHEMATICS

Unveiling Gender Disparities in STEM Success: A Logistic Regression Analysis of Penn State Students

KATHERINE KELLY
SPRING 2024

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Mathematics
with honors in Mathematics (BS)

Reviewed and approved* by the following:

Laura Cruz
Research Professor at the Schreyer Institute for Teaching Excellence
Thesis Supervisor

Dr. Aissa Wade
Professor of Mathematics
Honors Advisor

* Electronic approvals are on file.

ABSTRACT

The topic of women in STEM-related fields has dominated many conversations about representation in the workplace. Some researchers have turned to investigating the gender breakdown in undergraduate STEM majors to see if the gender differential starts in college, where many students select their career path.

I wanted to investigate whether there were significant differences between those who succeed in STEM fields at the university level, where students typically decide on their future career paths. I was curious to know whether gender alone was a valid predictor of success.

In this paper, I conducted a few logistic regression models based on gender and ethnicity in order to predict the success of women in entrance to STEM major classes at Penn State University Park. Real-life data from Penn State Undergraduate Education was used to make these models. I completed the variable selection process, compared the models' performances and validity, and demonstrated if and how the models could be used to predict the success of an undergraduate STEM major based on these demographic factors.

For each entrance to major class, as well as the data set overall, four logistic regression models were created: Gender predicting Success, Race/Ethnicity predicting Success, Gender and Race/Ethnicity predicting Success, and the interaction between Gender and Race/Ethnicity predicting Success. Each model was compared by their McFadden R2 and AIC values. The best model for each data set was selected, and their predictive performances were evaluated using ROC curves and corresponding AUC values. Finally, I used the model to try to predict success on the test data and calculated each model's accuracy.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank Laura Cruz, Research Professor at the Schreyer Institute for Teaching Excellence for the inspiration for this project and unwavering encouragement. I have greatly enjoyed our time together. I am grateful for all the opportunities that working for the Schreyer Institute for Teaching Excellence has entailed.

Thank you to my advisor, Aissa Wade, for your expertise and guidance. Your leadership in developing this project was very valuable.

Thank you to my sister and roommate, Grace, for your technological expertise and steadiness when I needed it. I know I can always trust you.

Finally, I would like to thank my parents. Thank you for letting me work through ideas with you and for your constant love and support. Thank you for being there when I needed you. I could not have completed this project without you.

Chapter 1

Background

Gender & STEM

The topic of women in STEM has become a much debated and researched topic in recent history. The fields of Science, Technology, Engineering, and Math have been historically dominated by men. This has been attributed to underrepresentation, unequal pay, and implicit biases. There is a wealth of information available on these topics, like studies on the gender breakdown of workers in STEM, the wage gap between men and women, and the unconscious perceptions people have about women's abilities to succeed in STEM fields. Scholars still search for notable differences between genders in terms of performance, workplace demographics, and motivations in STEM disciplines.

According to an article published in 2023, "only 17% of the total population were women choosing a STEM career" (Ortiz-Martínez et al). The authors further investigated the reasons behind this statistic and found that women are turned away from STEM majors because of "students' interests and self-perceptions" (Ortiz-Martínez et al). A cultural belief persists that women are not capable or meant to go into STEM careers, and this perception influences students' decisions in college.

Recently, women have accounted for more of the STEM labor force. This has been attributed to many factors, including increased representation and targeted marketing strategies. (González-Pérez, Susana, et al). These approaches have even proven to decrease gender biases while inspiring more women to enter into STEM careers. However, must studies still conclude that women are less likely to choose STEM majors than men, and some even find that women are less likely to graduate on time (Vooren et al.).

STEM at Penn State University Park

In order to officially declare a major at Penn State, a student must take some required foundational classes. BIOL 110, CHEM 110, CHEM 111, CHEM 112, MATH 140, and MATH 141 were identified as the classes that most students in the Eberly College of Science at University Park must take in order to declare their chosen major. These classes are typically some of the first courses that students who intend to declare STEM majors take at University Park.

A grade of C or better in BIOL 110 is required for admittance to the Biology, Premedicine, and Science (BS) majors. A grade of C or better in BIOL 110 is required for admittance to the Biology, Astronomy and Astrophysics, Biochemistry and Molecular Biology, Biotechnology, Chemistry, Forensic Science, Microbiology, Physics, Premedicine, and Science (BS) majors. Grades of C's or better in CHEM 111 and CHEM 112 are required for admittance to the Biochemistry and Molecular Biology, Biotechnology, Chemistry, Forensic Science, Microbiology, and Premedicine majors. A grade of C or better in MATH 140 is required for admittance to the Biology, Astronomy and Astrophysics, Biochemistry and Molecular Biology, Biotechnology, Chemistry, Data Sciences, Forensic Science, Mathematics (BA and BS), Microbiology, Physics, Planetary Science and Astronomy, Premedicine, Science BS, and Statistics majors. A grade of C or better in MATH 140 is required for admittance to the Astronomy and Astrophysics, Chemistry, Data Sciences, Mathematics (BA and BS), Physics, Premedicine, and Statistics majors. Thirteen specialized Engineering majors require a C or better in CHEM 110, MATH 140, and MATH 141. ("Eberly College of Science")

In Fall 2023, 57.8% of Undergraduate enrollments in Eberly were women ("Undergraduate Enrollment"). This follows a generally increasing trend since Fall 2017. These statistics only include students who enter Penn State in a Science major. They do not include students who are Undecided that later declare in a science major. According to the Penn State Planning, Assessment, and Institutional Research Database, the 4-year graduation rate in 2019 for women was 76.5%, while the rate for men was 65.4%. Since 2013, the rate for women has been higher ("Graduation and Retention").

Logistic Regression

Regression models are a tool commonly used by mathematicians and statisticians to model relationships between a dependent (outcome) variable and one or more independent (predictor) variables. Linear regression tests whether a linear relationship exists between a dependent and one or more independent variables. Logistic regression is considered a “generalized regression model”. It quantifies the probability of a categorical outcome variable based on one or more predictor variables. In binary logistic regression, this outcome variable has two possible states, traditionally coded as “0” or “1”. Examples of dichotomous dependent variables include the presence or absence of a disease, passing or failing an exam, or a simple “yes” or “no” response. Today, logistic regression is a commonly used tool to “estimate the probability that a particular subject will develop the outcome” (Hosmer et. al 1) since the development of statistical software (like R, for example) has made interpreting the results much more accessible.

While the coefficients provided by a linear regression model are relatively simple to interpret, correctly understanding the coefficients of a logistic regression model takes an additional step. For predictor variables that are continuous, the coefficient provided by the model represents a change in the log odds of the outcome variable happening for every unit increase in the predictor. The sign of the coefficient implies whether the outcome variable will be more or less likely as the predictor variable increases by one unit (Menard).

For categorical (qualitative) predictor variables, each value that the variable can take is coded based around a reference level. For example, if the predictor variable is gender, the researcher could set “male” as the reference gender, so the coefficient returned by the model would signify a change in the log odds of the response variable if the subject were a woman. Here “male” would be coded as 0 while “female” would be coded as 1. A coefficient’s sign indicates whether that category is likely to happen (positive implying more likely, negative less). The intercept of the model, commonly denoted as β_0, describes the log odds of the outcome variable being true (equaling 1) before adding any predictors. (Ranganathan et. al).

Model Development

Data was provided by Penn State Undergraduate Education Research. The data set consisted of the grade code earned by Undergraduate students at University Park who took the most popular classes required to enter a Science major in the Eberly College of Science at Penn State between the years 2014 and 2022, as well as their self-reported gender and ethnicity.

For the purposes of this study, the responses for the variable *ScholarGender* were filtered to Man, Woman, Transgender Man, and Transgender Woman. The grade codes were limited to the traditional grading scale (A, B, B+, B-, C, C+, D, and F) as well as the alternative grading system offered during the years 2020-2021 due to the COVID-19 pandemic (SAT representing a C or better and V and Z representing D and F, respectively). Possible responses for the variable *ScholarIPEDSRaceEthnicity* given dummy codes based on their frequency in the dataset. Responses included White (coded as A), International (coded as B), Hispanic or Latino (C), Asian (D), Black or African American (E), Two or more races (F), Race/Ethnicity unknown (G), Native Hawaiian/Pac Islander (H), American Ind or Alaska Native (I).

To build my logistic regression models, the outcome or dependent variable was determined to be success in one of these courses, meaning receiving a grade of a C or better (including SAT). Failure was defined by receiving a D, F, V, or Z, since a student who receives one of these grades would not get credit for the course. Failure was coded as 0, and Success was coded as 1. Late drops and withdrawals were not included in the study. The data was divided into train and test data in an 70/30 split. I developed the models with the training data and used the test data to check the accuracy and predictive power of the chosen models. R Studio was used for model creation and evaluation.

Chapter 2

LR Model with Gender as a Predictor

First, I created a logistic regression model using the *glm()* function in R to see if gender alone was statistically significant in predicting success across all classes in the dataset. The coefficient table provided by R is pictured below:

```
Coefficients:
                   Estimate Std. Error z value Pr(>|z|)
(Intercept)         1.67885    0.01314  127.77   <2e-16 ***
ScholarGenderWoman  0.23903    0.02042   11.71   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 66246  on 80382  degrees of freedom
Residual deviance: 66108  on 80381  degrees of freedom
AIC: 66112

Number of Fisher Scoring iterations: 4
```

Figure 1: Coefficient Table for Gender Predicting Success for Overall Dataset

This table shows that the log odds of a woman succeeding in an entrance to STEM major class is 0.23903. When plugged into the exponential function (e^x), the odds become 1.27. So, this implies that women are 1.27 times more likely to succeed in an entrance to STEM major class as an Undergraduate at University Park than men. They have an 87.19% chance of succeeding. The p-value is less than 0.05, which indicates that the variable *ScholarGender* is significant in modeling *Success*.

The intercept in this model represents the log odds of a man succeeding in an entrance to STEM major class. When inserted in the exponential function, this coefficient becomes 5.359, which implies that the chances of a man earning the credit for one of these classes is 5.359 times more likely than not earning the credit. When plugged into the probability formula, we conclude that men have an 84.28% chance of succeeding.

This process was repeated for each class in the overall dataset. Their corresponding coefficients and p-values for women's success are shown in the table below. The coefficient represents the log odds (the coefficient provided by R), the odds column represents the odds of success (the coefficient plugged into the exponential function), and the probability column represents the probability of women succeeding in each class.

Table 1: Coefficient Table for Gender as the Sole Predictor of Success

Course	Coefficient	Odds	Probability (W)	Probability (M)	p-value
BIOL 110	0.3967	1.4869	94.09%	91.46%	1.16e-6
CHEM 110	-0.1414	0.8681	81.84%	83.85%	0.0001
CHEM 111	0.6931	1.9999	96.22%	92.72%	<2e-16
CHEM 112	0.1685	1.1835	85.95%	83.79%	0.001
MATH 140	-0.0504	0.9508	76.05%	76.96%	0.266
MATH 141	0.0172	1.0173	82.31%	82.05%	0.756

From these results, we can see that *ScholarGender* is a significant predictor of modeling success for BIOL 110, CHEM 110, CHEM 111, and CHEM 112 at the 5% significance level. However, since the p-value is greater than 0.05 for MATH 140 and MATH 141, we can conclude that gender alone might not be a significant predictor of modeling success for those classes. It is also interesting to note that the negative coefficients for CHEM 110 and MATH 140 are negative, which implies that women are less likely to succeed in those classes than men. This is so evidenced in the probability columns, where the probability of success is lower for women than men.

Chapter 3

Logistic Regression Model with Ethnicity as a Predictor

The second model I developed evaluated the if ethnicity, or *ScholarIPEDSRaceEthnicity* as it is named in the data, was a significant predictor of success in Undergraduate students at University Park. The category White was selected as the reference category because it was (by far) the largest subset in the data. The coefficient table provided by R is pictured below:

```
Coefficients:
                           Estimate Std. Error z value Pr(>|z|)    
(Intercept)                 1.98981    0.01386 143.562  < 2e-16 ***
ScholarIPEDSRaceEthnicityB -0.27596    0.03128  -8.823  < 2e-16 ***
ScholarIPEDSRaceEthnicityC -0.72191    0.03380 -21.356  < 2e-16 ***
ScholarIPEDSRaceEthnicityD -0.24952    0.03858  -6.468 9.91e-11 ***
ScholarIPEDSRaceEthnicityE -1.22435    0.03935 -31.112  < 2e-16 ***
ScholarIPEDSRaceEthnicityF -0.49825    0.04831 -10.314  < 2e-16 ***
ScholarIPEDSRaceEthnicityG  0.35480    0.07206   4.923 8.51e-07 ***
ScholarIPEDSRaceEthnicityH -0.98651    0.24963  -3.952 7.75e-05 ***
ScholarIPEDSRaceEthnicityI -0.01573    0.47746  -0.033    0.974    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 66246  on 80382  degrees of freedom
Residual deviance: 64976  on 80374  degrees of freedom
AIC: 64994

Number of Fisher Scoring iterations: 4
```

Figure 2: Coefficient Table for Ethnicity Predicting Success for Overall Dataset

From this table, we can see that every category except G (Race/Ethnicity Unknown) returned a negative coefficient, implying odds less than 1 of succeeding compared to White students. Also, every category but I (American Ind or Alaska Native) was significant at the 5% confidence level. Specifically, International students have 24.12% lower odds of succeeding than White students, while students who identified as Race/Ethnicity Unknown have 1.426 times the odds of succeeding than White students. When plugged into the exponential function, we find that White students are 7.314 times more likely to succeed in an entrance to STEM major class than they are to fail. This equates to an 87.97% chance of success.

Once again, I ran the same model with Ethnicity as the sole Predictor for each class. For BIOL 110, every category was significant (p-value < 0.05) except for Asian, Race/Ethnicity Unknown, Native Hawaiian/Pac Islander, and American Ind or Alaska Native. Every category except Race/Ethnicity Unknown and American Ind or Alaska Native had a negative coefficient, which signifies that every other level of the categorical variable has lower odds of succeeding in BIOL 110 than White students. The coefficient of the intercept was 2.936, which means that the odds of a White student succeeding in BIOL 110 are 18.84 times more likely than failing.

For CHEM 110, every category in the model was significant at the 5% level except Race/Ethnicity Unknown. The only positive coefficient was again Race/Ethnicity Unknown, which means that every other level in the category is less likely to succeed than White students. The coefficient of the intercept was 1.803, which translates into White students having an 85.85% chance of succeeding in CHEM 110.

For CHEM 111, every level was significant at the 5% level except for Race/Ethnicity Unknown and American Ind or Alaska Native. Every level of *ScholarIPEDSRaceEthnicty* except Race/Ethnicity Unknown and American Ind or Alaska Native have lower odds of succeeding than White students. The coefficient was 3.1423, which becomes 23.157 when plugged into the exponential function. This means that White students are 23.157 times more likely to succeed than fail in CHEM 111.

For CHEM 112, the International, Asian, Native Hawaiian/Pac Islander, and American Ind or Alaska Native levels were not statistically significant. Race/Ethnicity Unknown was still the only level with higher odds of success than White students. The intercept was 1.855, which translates into White students being 6.393 times more likely to earn credit for taking CHEM 112 than not.

For MATH 140, the International, Asian, Race/Ethnicity Unknown, and American Ind or Alaska Native levels were not statistically significant. The coefficients for the International and American Ind or

Alaska Native levels were positive, which means that the odds of students in these groups have higher odds of succeeding in MATH 140 than White students. From the intercept, we know that White students are 1.314 times more likely to succeed in MATH 140 than fail.

For MATH 141, every level except International, Race/Ethnicity Unknown, Native Hawaiian/Pac Islander, and American Ind or Alaska Native were significant at 5%. The coefficients for these four levels alone were positive. The intercept was 1.625, which means that White students' odds of succeeding in MATH 141 are 5.078 times the odds of failing.

Chapter 4

Logistic Regression Model with Gender and Ethnicity as Predictors

This model was an additive model containing Gender and Ethnicity, created to test whether the combination of the two variables would be better indicators of success. The reference categories were White male students. First, I ran the model for all classes. The coefficient table is shown below:

```
Coefficients:
                             Estimate Std. Error z value Pr(>|z|)    
(Intercept)                  1.864911   0.016345 114.095  < 2e-16 ***
ScholarGenderWoman           0.283141   0.020840  13.586  < 2e-16 ***
ScholarIPEDSRaceEthnicityB -0.232346   0.031462  -7.385 1.52e-13 ***
ScholarIPEDSRaceEthnicityC -0.737388   0.033882 -21.763  < 2e-16 ***
ScholarIPEDSRaceEthnicityD -0.240769   0.038626  -6.233 4.57e-10 ***
ScholarIPEDSRaceEthnicityE -1.262820   0.039556 -31.925  < 2e-16 ***
ScholarIPEDSRaceEthnicityF -0.510122   0.048389 -10.542  < 2e-16 ***
ScholarIPEDSRaceEthnicityG  0.351950   0.072120   4.880 1.06e-06 ***
ScholarIPEDSRaceEthnicityH -0.967648   0.250081  -3.869 0.000109 ***
ScholarIPEDSRaceEthnicityI -0.007722   0.477945  -0.016 0.987110    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 66246  on 80382  degrees of freedom
Residual deviance: 64789  on 80373  degrees of freedom
AIC: 64809

Number of Fisher Scoring iterations: 4
```

Figure 3: Coefficient Table for Gender and Ethnicity Predicting Success for Overall Dataset

From this table, we can see that every level is significant except for American Ind or Alaska Native. Also, women and Race/Ethnicity Unknown were the only categories in which the likelihood of success was higher than failure given their positive coefficients. A female student whose Race/Ethnicity is unknown would 1.887 times more likely to succeed in an entrance to STEM major class than a White male. They will succeed in one of these classes with probability 92.41%. A White female (just the coefficient *ScholarGenderWoman*) is 1.327 times more likely to succeed than a White male. Their probability of success is 89.55%. The intercept implies that White men are 6.455 times more likely to succeed than fail, and they have an 86.59% chance of success.

I continued to run this model for each of the other classes. For BIOL 110, every level except Asian, Race/Ethnicity Unknown, Native Hawaiian/Pac Islander, and American Ind or Alaska Native were significant. Coefficients for women, Race/Ethnicity Unknown, and American Ind or Alaska Native were positive. A White female is 1.58 times more likely to succeed than a White male, with a 95.74% chance of success. A White male student is 14.222 times more likely to succeed in BIOL 110 than they are to fail, with a 93.43% chance of success.

For CHEM 110, every level of *ScholarIPEDSRaceEthnicity* was significant at 5% except American Ind or Alaska Native. Every category except Race/Ethnicity Unknown had a negative coefficient. White male students are 6.34 times more likely to succeed in CHEM 110 than they are to fail. In this model, White women are less likely to succeed than White males, like the first model with Gender as the sole predictor. White women have an 85.25% chance of success, while White men have an 86.38% chance of success.

For CHEM 111, every level was significant at the 5% level except for Race/Ethnicity Unknown and American Ind or Alaska Native. Women, students who reported their Race/Ethnicity as Unknown, and American Ind or Alaska Native students had positive coefficients, meaning that their odds of success were greater than the odds of success for White males. The odds of a White man succeeding in CHEM 111 is 16.82 times their odds of failure, resulting in a success probability of 94.39%.

For CHEM 112, every level was significant with the exception of International, Asian, Native Hawaiian/Pac Islander, and American Ind or Alaska Native at 5% significance. The coefficients for women, Asian students, and students who identified as Race/Ethnicity Unknown had positive coefficients. White women are 1.256 times more likely to succeed than White men, with an 87.74% chance of success. The intercept implied that White men are 5.7 times more likely to succeed in CHEM 112 than they are to fail, with a success probability of 85.07%.

For MATH 140, the only significant levels at 5% were Hispanic or Latino, Black or African American, Two or more races, and Native Hawaiian/Pac Islander. Coefficients for Asian students and American Ind or Alaska Natives had positive coefficients. White women have a 78.59% chance of success, and White men have a 78.93% chance of success.

For MATH 141, the only significant levels of the predictor were Hispanic or Latino, Asian, Black or African American, and Two or more races. Women, International students, students who reported their Race/Ethnicity as Unknown, Native Hawaiian/Pac Islanders, and American Ind or Alaska Natives had positive coefficients. White women are 1.05 times more likely to succeed in MATH 141 than White men, who are 5.01 times more likely to succeed than fail. White men also have an 83.37% chance of success in MATH 141, while White women have an 84.03% chance of success.

Chapter 5

Logistic Regression Model with the Interaction Between Gender and Ethnicity as a Predictor

To limit specification and non-additivity error in the models, I made a model that included the interaction between Gender and Ethnicity as a term. This decision meant that I could test whether the combination of gender and ethnicity would have a significant effect on success. This model determined whether the independent variables were interactive as opposed to additive (which was shown in the previous chapter). The reference categories were again White male students. First, I ran the model for all classes. The coefficient table is shown on the next page.

All levels except for and American Ind or Alaska Native, International women, Asian women, Black or African American women, female Native Hawaiian/Pac Islanders, and female American Ind or Alaska Natives were significant. The only positive coefficients were for women, students who reported their Race/Ethnicity as Unknown, American Ind or Alaska Natives, International women, Black or African American women, and women whose Race/Ethnicity is unknown. Again, I repeated this process for each class.

```
Coefficients:
                                             Estimate Std. Error z value
(Intercept)                                   1.85144    0.01805 102.598
ScholarGenderWoman                            0.31587    0.02823  11.188
ScholarIPEDSRaceEthnicityB                   -0.24482    0.03717  -6.587
ScholarIPEDSRaceEthnicityC                   -0.61675    0.04746 -12.995
ScholarIPEDSRaceEthnicityD                   -0.18069    0.05010  -3.607
ScholarIPEDSRaceEthnicityE                   -1.30726    0.05836 -22.399
ScholarIPEDSRaceEthnicityF                   -0.41571    0.06714  -6.192
ScholarIPEDSRaceEthnicityG                    0.23233    0.09019   2.576
ScholarIPEDSRaceEthnicityH                   -0.80547    0.32292  -2.494
ScholarIPEDSRaceEthnicityI                    0.04568    0.61940   0.074
ScholarGenderWoman:ScholarIPEDSRaceEthnicityB  0.06719    0.07054   0.952
ScholarGenderWoman:ScholarIPEDSRaceEthnicityC -0.25088    0.06783  -3.699
ScholarGenderWoman:ScholarIPEDSRaceEthnicityD -0.15041    0.07858  -1.914
ScholarGenderWoman:ScholarIPEDSRaceEthnicityE  0.07060    0.07960   0.887
ScholarGenderWoman:ScholarIPEDSRaceEthnicityF -0.20342    0.09681  -2.101
ScholarGenderWoman:ScholarIPEDSRaceEthnicityG  0.31018    0.15108   2.053
ScholarGenderWoman:ScholarIPEDSRaceEthnicityH -0.42357    0.50925  -0.832
ScholarGenderWoman:ScholarIPEDSRaceEthnicityI -0.13355    0.97295  -0.137
                                             Pr(>|z|)
(Intercept)                                   < 2e-16 ***
ScholarGenderWoman                            < 2e-16 ***
ScholarIPEDSRaceEthnicityB                   4.50e-11 ***
ScholarIPEDSRaceEthnicityC                    < 2e-16 ***
ScholarIPEDSRaceEthnicityD                   0.000310 ***
ScholarIPEDSRaceEthnicityE                    < 2e-16 ***
ScholarIPEDSRaceEthnicityF                   5.95e-10 ***
ScholarIPEDSRaceEthnicityG                   0.009995 **
ScholarIPEDSRaceEthnicityH                   0.012619 *
ScholarIPEDSRaceEthnicityI                   0.941210
ScholarGenderWoman:ScholarIPEDSRaceEthnicityB 0.340855
ScholarGenderWoman:ScholarIPEDSRaceEthnicityC 0.000217 ***
ScholarGenderWoman:ScholarIPEDSRaceEthnicityD 0.055605 .
ScholarGenderWoman:ScholarIPEDSRaceEthnicityE 0.375160
ScholarGenderWoman:ScholarIPEDSRaceEthnicityF 0.035628 *
ScholarGenderWoman:ScholarIPEDSRaceEthnicityG 0.040068 *
ScholarGenderWoman:ScholarIPEDSRaceEthnicityH 0.405549
ScholarGenderWoman:ScholarIPEDSRaceEthnicityI 0.890823
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 66246  on 80382  degrees of freedom
Residual deviance: 64759  on 80365  degrees of freedom
AIC: 64795

Number of Fisher Scoring iterations: 5
```

Figure 4: Coefficient Table for Overall Data with the Interaction Between Gender and Ethnicity as a Predictor

Chapter 6

Model Comparisons

Once I had created all of these models, I wanted to see which one best fit the data. To do this, I utilized the minimum AIC procedure. The AIC, or Akaike information criterion, of a model is “an estimate of minus twice the expected log likelihood of the model” (Akaike 1). Common practice dictates that the model with the lowest AIC is the best fitting. Using the *tab_model()* function in R, I was able to compare my models side-by-side with their AIC values and levels of significance. The table for the overall dataset is shown on the next page. Based on the minimum AIC procedure, the model that fits the data best is the fourth model, which includes the interaction between Gender and Ethnicity. The AIC value was 65,279.608.

I repeated this process for every class. These tables can be found in Appendix A. For BIOL 110, CHEM 112, MATH 140, and MATH 141, the fourth model was also the best fitting. For CHEM 110 and CHEM 111, the third model was the best fitting, which was the additive model with Gender and Ethnicity as predictors.

To confirm these results, I ran likelihood ratio tests to test for a significant difference between models three and four. The null hypothesis in each test was that the two models were equivalent. Using the *anova()* function in R, the tests returned significant results at the 5% level for the overall data, BIOL 110, MATH 140, and MATH 141. So, we can reject the hypothesis that there is no interaction between age and sex for those datasets. The tests for CHEM 110, CHEM 111, and CHEM 112 returned insignificant values, which means that we cannot reject the null hypothesis.

	Success		Success		Success		Success	
Predictors	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*
(Intercept)	5.35 ***	<0.001	7.23 ***	<0.001	6.46 ***	<0.001	6.35 ***	<0.001
ScholarGender [Woman]	1.24 ***	<0.001			1.29 ***	<0.001	1.35 ***	<0.001
ScholarIPEDSRaceEthnicity [B]			0.77 ***	<0.001	0.80 ***	<0.001	0.80 ***	<0.001
ScholarIPEDSRaceEthnicity [C]			0.48 ***	<0.001	0.47 ***	<0.001	0.54 ***	<0.001
ScholarIPEDSRaceEthnicity [D]			0.75 ***	<0.001	0.75 ***	<0.001	0.80 ***	<0.001
ScholarIPEDSRaceEthnicity [E]			0.29 ***	<0.001	0.28 ***	<0.001	0.28 ***	<0.001
ScholarIPEDSRaceEthnicity [F]			0.63 ***	<0.001	0.62 ***	<0.001	0.67 ***	<0.001
ScholarIPEDSRaceEthnicity [G]			1.39 ***	<0.001	1.38 ***	<0.001	1.29 **	0.006
ScholarIPEDSRaceEthnicity [H]			0.36 ***	<0.001	0.36 ***	<0.001	0.41 **	0.004
ScholarIPEDSRaceEthnicity [I]			0.62	0.175	0.61	0.157	0.63	0.356
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [B]							1.05	0.480
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [C]							0.75 ***	<0.001
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [D]							0.87	0.085
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [E]							1.02	0.823
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [F]							0.84	0.087
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [G]							1.18	0.246
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [H]							0.75	0.558
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [I]							0.93	0.916
Observations	80383		80383		80383		80383	
R^2 Tjur	0.001		0.018		0.020		0.021	
AIC	66599.750		65441.646		65290.458		65279.608	

• *p<0.05 ** p<0.01 *** p<0.001*

Figure 5: Model Comparison for Overall Dataset

Chapter 7

Model Evaluations

McFadden's R^2

Now that we have the best model for each dataset, we can determine exactly how well the model fits and test its accuracy in predicting success. For this paper's purposes, I used McFadden's R^2 to assess the goodness of fit, and the receiver operating characteristic (ROC) along with its quantitative area measure AUC.

In linear regression, the R^2 value that the model summary returns measures how well the model estimates the data. Since the traditional R^2 is not recommended for use in logistic regression, several other "pseudo"-R^2 measures have been developed. McFadden's R^2 compares the log-likelihoods of the null model (null meaning without any predictors) with the selected model (Hemmert et al.). The closer a linear regression model's R^2 is to 1, the better the model approximates the data. For logistic regression, "values from 0.2 to 0.4 are tolerable and values higher than 0.4 are a good fit" (Hemmert et al.). The McFadden's R^2 values for each optimal model (as determined in the previous section) can be found in the table below:

Table 2: McFadden R2 for Every Optimum Model

Dataset	McFadden R^2
Overall	0.0223
BIOL 110	0.0497
CHEM 110	0.0273
CHEM 111	0.0457
CHEM 112	0.0220
MATH 140	0.0202
MATH 141	0.0193

ROC Curves and AUC

One method of evaluating a model's predictive performance is the ROC curve, which is a graph that allows researchers to plot a model's false-positive rate versus its true-positive rate (Huang et al.). We use the "testing" data, coming from the initial 70/30 splitting of the dataset, to test how well these optimal models can accurately predict success. The ROC provides a visual of the accuracy, while the AUC (Area Under the Curve) indicates how well the model performs numerically. A perfect ROC curve, where a model is totally accurate, would look like a 90° angle, with a straight vertical line from 0 to 1 from the origin. This would signify a perfect true positive rate while the false positive rate remains 0 until the true positive rate reaches 100%. Accordingly, the AUC for a perfect predictor model would be 1. If the curve is a 45° angle, or the shape of the line y = x, that would signify a model whose true positive rate consistently equaled the false positive. This would mean that the model is no better than random chance. The AUC for this model would be 0.5. (Huang et al.)

The ROC curve for the overall model is pictured below:

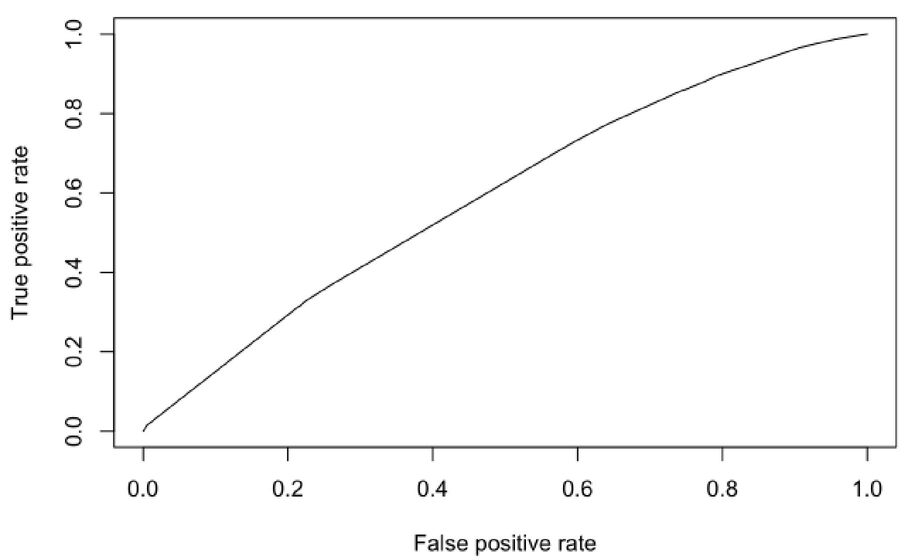

Figure 6: ROC Curve for Overall Model with Interaction Between Gender and Ethnicity

This model is relatively close to the like y = x, which indicates that this model is not a very good predictor. To confirm, I calculated the Area Under the Curve, or AUC, which was 0.593. Based on the flatness of the curve and the AUC, I conclude that this model is likely not a very good predictor of success. I plotted the ROCs for the rest of the models (see Appendix B). Their AUC values are in the table below:

Table 3: AUC Scores for Class Datasets

Data set	AUC
BIOL 110	0.669
CHEM 110	0.594
CHEM 111	0.614
CHEM 112	0.590
MATH 140	0.564
MATH 141	0.565

Since they are all close to 0.5, none of these models are particularly good predictors of success. This implies that Gender and Ethnicity (and the interaction between them) might not be the best predictors of success for Undergraduate students at University Park.

For a final evaluation, I tested the model's accuracy in predicting success with the testing datasets in a contingency table. The table for the overall dataset is shown below:

Table 4: Predicting Success Using Overall Model 4

	Predicted	
Observed	Failure	Success
Failure	0	5,105
Success	0	29,346

The model correctly predicted the failure of 0 students and the success of 29,346 students. The model incorrectly predicted the failure of 0 students and the success of 5,105 students. The accuracy of the model is 85.18%.

For the BIOL 110 test dataset, the fourth model correctly predicted the failure of 0 students and the success of 3,918 students. The model incorrectly predicted the failure of 269 students and the success of 0 students. Its accuracy was 93.58%.

For the CHEM 110 test dataset, the third model correctly predicted the failure of 0 students and the success of 7,454 students. The model incorrectly predicted the failure of 0 students and the success of 1,506 students. Its accuracy was 83.19%.

For the CHEM 111 dataset, the third model correctly predicted the failure of 0 students and the success of 5,730 students. The model incorrectly predicted the failure of 0 students and the success of 340 students. Its accuracy was 94.40%.

For the CHEM 112 dataset, the fourth model correctly predicted the failure of 1 student and the success of 4,370 students. The model incorrectly predicted the failure of 33 students and the success of 746 students. Its accuracy was 85.37%.

For the MATH 140 dataset, the fourth model correctly predicted the failure of 34 students and the success of 4,058 students. The model incorrectly predicted the failure of 46 students and the success of 1,2237 students. Its accuracy was 76.13%.

For the MATH 141 dataset, the fourth model correctly predicted the failure of 0 students and the success of 3,905 students. The model incorrectly predicted the failure of 0 students and the success of 836 students. Its accuracy was 82.37%.

Overall, the models were moderately successful at predicting success. However, the AUC values are still low, and the ROC curves are relatively flat. Therefore, we cannot conclude that Gender and Ethnicity are good predictors of success for Undergraduate students at University Park.

Chapter 8

Discussion

For a model with a very low R2, my models were able to predict many successes correctly. This could be attributed to other correlated variables that were missing from the model. There could be other variables that can predict success in STEM majors at University Park better. My predictors could also just be well fit for this data but cannot be used for larger generalizations. This conclusion is supported by the relatively high prediction accuracy when I tested the trained data.

While the models themselves were not the best predictors of success, I believe that there are still valuable insights from my research. I did not know that women were actually more likely to succeed than men overall, as well as in a few of the classes. I have taken a lot of Math classes in my time at Penn State, and I had felt that there was a gender imbalance. This could be attributed to a more pronounced gender difference in higher-level classes, or in the Math major specifically.

One concern I had regarding the diagnostics of these models was the possibility I was missing relevant variables, which could lead to specification error. I worked with the data I had to test my hypothesis, which was that Gender could predict success in Undergraduate students at University Park. In the future, I would like to add other demographic variables to see if those would make better predictors. For example, some other variables in LionPath include whether the student is an Honors student, whether the student is a student athlete, or whether the student is a first-generation college student.

Another way I would like to expand the model would be to include a wider respondent pool. These models were very specific to Undergraduates at Penn State University Park, where the success and graduation rates are usually higher than other campuses or universities. I wonder if there would be sharper differences between men and women at other institutions. I would be interested to see if there were similarities among other like schools, for example, other schools in the Big Ten Conference.

Appendix A

Model Comparison Tables

	Success		Success		Success		Success	
Predictors	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*
(Intercept)	9.43 ***	**<0.001**	17.55 ***	**<0.001**	12.07 ***	**<0.001**	10.07 ***	**<0.001**
ScholarGender [Woman]	1.74 ***	**<0.001**			1.86 ***	**<0.001**	2.68 ***	**<0.001**
ScholarIPEDSRaceEthnicity [B]			0.43 ***	**<0.001**	0.46 ***	**<0.001**	0.56 *	**0.014**
ScholarIPEDSRaceEthnicity [C]			0.40 ***	**<0.001**	0.39 ***	**<0.001**	0.69	0.067
ScholarIPEDSRaceEthnicity [D]			0.89	0.511	0.93	0.695	1.24	0.384
ScholarIPEDSRaceEthnicity [E]			0.26 ***	**<0.001**	0.24 ***	**<0.001**	0.44 ***	**<0.001**
ScholarIPEDSRaceEthnicity [F]			0.69	0.052	0.69	0.057	1.01	0.977
ScholarIPEDSRaceEthnicity [G]			2.39 *	**0.016**	2.45 *	**0.013**	4.07 *	**0.017**
ScholarIPEDSRaceEthnicity [H]			0.40	0.391	0.47	0.488	210422.34	0.975
ScholarIPEDSRaceEthnicity [I]			120695.00	0.965	114953.02	0.965	210422.34	0.981
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [B]							0.66	0.231
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [C]							0.37 ***	**<0.001**
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [D]							0.57	0.099
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [E]							0.38 ***	**0.001**
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [F]							0.49	0.071
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [G]							0.39	0.202
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [H]							0.00	0.970
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [I]							0.37	0.999
Observations	9767		9767		9767		9767	
R^2 Tjur	0.005		0.020		0.025		0.028	
AIC	4923.025		4818.173		4763.210		4749.379	

• p<0.05 ** p<0.01 *** p<0.001

Figure 7: Model Comparison for BIOL 110

	Success		Success		Success		Success	
Predictors	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*
(Intercept)	5.17 ***	**<0.001**	6.11 ***	**<0.001**	6.30 ***	**<0.001**	6.34 ***	**<0.001**
ScholarGender [Woman]	0.89 **	**0.001**			0.94	0.084	0.92	0.118
ScholarIPEDSRaceEthnicity [B]			0.93	0.275	0.92	0.206	0.87	0.078
ScholarIPEDSRaceEthnicity [C]			0.38 ***	**<0.001**	0.38 ***	**<0.001**	0.40 ***	**<0.001**
ScholarIPEDSRaceEthnicity [D]			0.80 **	**0.002**	0.79 **	**0.002**	0.87	0.166
ScholarIPEDSRaceEthnicity [E]			0.27 ***	**<0.001**	0.27 ***	**<0.001**	0.23 ***	**<0.001**
ScholarIPEDSRaceEthnicity [F]			0.62 ***	**<0.001**	0.62 ***	**<0.001**	0.61 ***	**<0.001**
ScholarIPEDSRaceEthnicity [G]			1.51 **	**0.002**	1.51 **	**0.002**	1.59 *	**0.017**
ScholarIPEDSRaceEthnicity [H]			0.33 *	**0.016**	0.32 *	**0.015**	0.28 *	**0.024**
ScholarIPEDSRaceEthnicity [I]			0.29 *	**0.029**	0.30 *	**0.029**	0.21 *	**0.041**
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [B]							1.21	0.193
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [C]							0.89	0.328
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [D]							0.83	0.200
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [E]							1.37 *	**0.026**
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [F]							1.04	0.827
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [G]							0.91	0.729
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [H]							1.50	0.686
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [I]							2.03	0.533
Observations	20905		20905		20905		20905	
R^2 Tjur	0.001		0.031		0.031		0.031	
AIC	19028.924		18499.974		18498.992		18503.753	

- *p<0.05 ** p<0.01 *** p<0.001*

Figure 8: Model Comparison for CHEM 110

	Success		Success		Success		Success	
Predictors	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*
(Intercept)	12.23 ***	**<0.001**	22.60 ***	**<0.001**	16.23 ***	**<0.001**	15.55 ***	**<0.001**
ScholarGender [Woman]	2.05 ***	**<0.001**			2.13 ***	**<0.001**	2.41 ***	**<0.001**
ScholarIPEDSRaceEthnicity [B]			0.45 ***	**<0.001**	0.50 ***	**<0.001**	0.52 ***	**<0.001**
ScholarIPEDSRaceEthnicity [C]			0.48 ***	**<0.001**	0.47 ***	**<0.001**	0.51 ***	**<0.001**
ScholarIPEDSRaceEthnicity [D]			0.64 **	**0.001**	0.63 **	**0.001**	0.76	0.137
ScholarIPEDSRaceEthnicity [E]			0.26 ***	**<0.001**	0.23 ***	**<0.001**	0.26 ***	**<0.001**
ScholarIPEDSRaceEthnicity [F]			0.61 **	**0.005**	0.57 **	**0.001**	0.68	0.121
ScholarIPEDSRaceEthnicity [G]			1.81	0.054	1.78	0.063	1.59	0.204
ScholarIPEDSRaceEthnicity [H]			0.18 *	**0.029**	0.16 *	**0.024**	0.06 **	**0.006**
ScholarIPEDSRaceEthnicity [I]			0.53	0.544	0.52	0.535	50109.69	0.961
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [B]							0.93	0.791
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [C]							0.80	0.346
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [D]							0.63	0.110
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [E]							0.73	0.225
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [F]							0.66	0.247
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [G]							1.41	0.619
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [H]							322792.40	0.954
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [I]							0.00	0.954
Observations	14162		14162		14162		14162	
R^2 Tjur	0.007		0.013		0.021		0.021	
AIC	6034.776		5985.887		5888.196		5894.187	

* $p<0.05$ ** $p<0.01$ *** $p<0.001$

Figure 9: Model Comparison for CHEM 111

	Success		Success		Success		Success	
Predictors	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*
(Intercept)	5.36 ***	**<0.001**	6.58 ***	**<0.001**	5.90 ***	**<0.001**	5.82 ***	**<0.001**
ScholarGender [Woman]	1.17 **	**0.002**			1.25 ***	**<0.001**	1.28 ***	**<0.001**
ScholarIPEDSRaceEthnicity [B]			1.00	0.996	1.05	0.620	0.98	0.868
ScholarIPEDSRaceEthnicity [C]			0.54 ***	**<0.001**	0.53 ***	**<0.001**	0.64 ***	**0.001**
ScholarIPEDSRaceEthnicity [D]			0.95	0.606	0.96	0.656	0.93	0.567
ScholarIPEDSRaceEthnicity [E]			0.31 ***	**<0.001**	0.30 ***	**<0.001**	0.29 ***	**<0.001**
ScholarIPEDSRaceEthnicity [F]			0.65 ***	**<0.001**	0.64 ***	**<0.001**	0.90	0.593
ScholarIPEDSRaceEthnicity [G]			1.89 ***	**<0.001**	1.88 ***	**0.001**	1.54	0.067
ScholarIPEDSRaceEthnicity [H]			0.42	0.136	0.42	0.141	1.20	0.863
ScholarIPEDSRaceEthnicity [I]			0.15	0.060	0.14	0.053	0.00	0.918
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [B]							1.32	0.194
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [C]							0.72	0.069
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [D]							1.08	0.712
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [E]							1.04	0.868
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [F]							0.55 *	**0.017**
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [G]							1.57	0.227
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [H]							0.15	0.147
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [I]							60554.97	0.927
Observations	11946		11946		11946		11946	
R^2 Tjur	0.001		0.018		0.019		0.021	
AIC	9972.033		9812.522		9797.107		9795.553	

- * $p<0.05$ ** $p<0.01$ *** $p<0.001$

Figure 10: Model Comparison for CHEM 112

	Success		Success		Success		Success	
Predictors	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*
(Intercept)	3.30 ***	**<0.001**	3.70 ***	**<0.001**	3.73 ***	**<0.001**	3.72 ***	**<0.001**
ScholarGender [Woman]	0.96	0.345			0.98	0.678	0.99	0.844
ScholarIPEDSRaceEthnicity [B]			1.02	0.765	1.02	0.773	0.93	0.288
ScholarIPEDSRaceEthnicity [C]			0.51 ***	**<0.001**	0.51 ***	**<0.001**	0.56 ***	**<0.001**
ScholarIPEDSRaceEthnicity [D]			0.96	0.632	0.96	0.629	0.99	0.896
ScholarIPEDSRaceEthnicity [E]			0.30 ***	**<0.001**	0.30 ***	**<0.001**	0.31 ***	**<0.001**
ScholarIPEDSRaceEthnicity [F]			0.61 ***	**<0.001**	0.61 ***	**<0.001**	0.81	0.129
ScholarIPEDSRaceEthnicity [G]			0.99	0.942	0.99	0.941	0.90	0.526
ScholarIPEDSRaceEthnicity [H]			0.39	0.054	0.39	0.053	0.38	0.096
ScholarIPEDSRaceEthnicity [I]			28469.97	0.917	28567.99	0.917	28357.33	0.941
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [B]							1.39 *	**0.012**
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [C]							0.77	0.115
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [D]							0.92	0.627
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [E]							0.90	0.599
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [F]							0.51 **	**0.002**
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [G]							1.38	0.306
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [H]							1.08	0.940
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [I]							1.01	1.000
Observations	12541		12541		12541		12541	
R^2 Tjur	0.000		0.020		0.020		0.022	
AIC	13678.694		13466.760		13468.588		13460.629	

• *p<0.05 ** p<0.01 *** p<0.001*

Figure 11: Model Comparison for MATH 140

	Success		Success		Success		Success	
Predictors	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*
(Intercept)	4.53 ***	**<0.001**	5.11 ***	**<0.001**	4.96 ***	**<0.001**	5.05 ***	**<0.001**
ScholarGender [Woman]	1.08	0.166			1.12 *	**0.046**	1.05	0.556
ScholarIPEDSRaceEthnicity [B]			1.15 *	**0.036**	1.15 *	**0.037**	1.03	0.714
ScholarIPEDSRaceEthnicity [C]			0.54 ***	**<0.001**	0.54 ***	**<0.001**	0.60 ***	**<0.001**
ScholarIPEDSRaceEthnicity [D]			0.79 **	**0.005**	0.79 **	**0.005**	0.77 **	**0.010**
ScholarIPEDSRaceEthnicity [E]			0.28 ***	**<0.001**	0.28 ***	**<0.001**	0.29 ***	**<0.001**
ScholarIPEDSRaceEthnicity [F]			0.67 **	**0.002**	0.67 **	**0.002**	0.70 *	**0.028**
ScholarIPEDSRaceEthnicity [G]			0.88	0.401	0.88	0.381	0.76	0.119
ScholarIPEDSRaceEthnicity [H]			0.46	0.256	0.45	0.248	0.40	0.286
ScholarIPEDSRaceEthnicity [I]			56083.77	0.929	55121.78	0.929	56794.74	0.946
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [B]							1.58 **	**0.004**
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [C]							0.71	0.121
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [D]							1.08	0.699
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [E]							0.88	0.605
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [F]							0.86	0.595
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [G]							1.66	0.148
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [H]							1.43	0.804
ScholarGender [Woman] × ScholarIPEDSRaceEthnicity [I]							0.95	1.000
Observations	11060		11060		11060		11060	
R^2 Tjur	0.000		0.017		0.017		0.018	
AIC	10350.639		10206.330		10204.314		10204.192	

-

* $p<0.05$ ** $p<0.01$ *** $p<0.001$

Figure 12: Model Comparison for MATH 141

Appendix B

ROC Curves

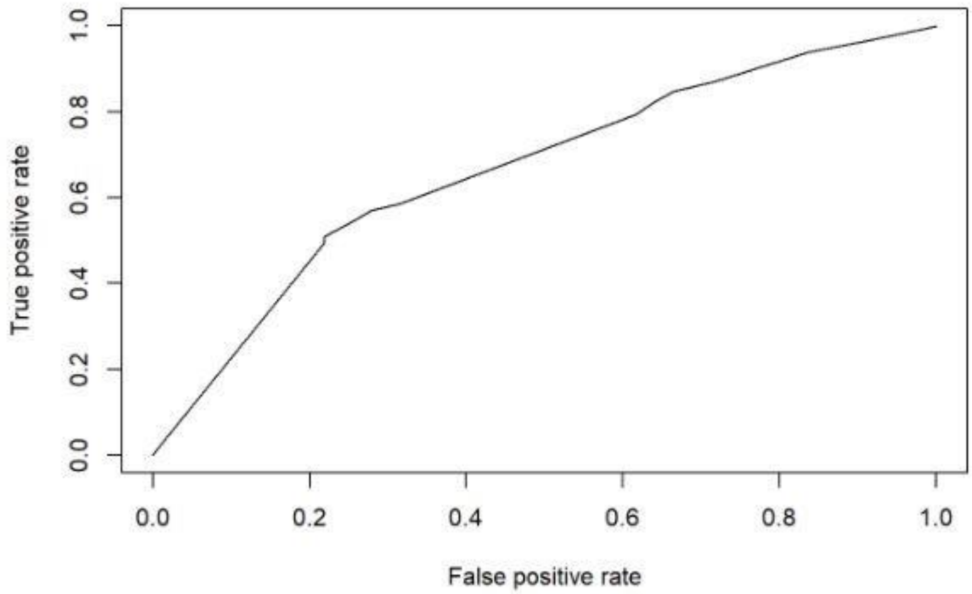

Figure 13: ROC Curve for BIOL 110 Model 4

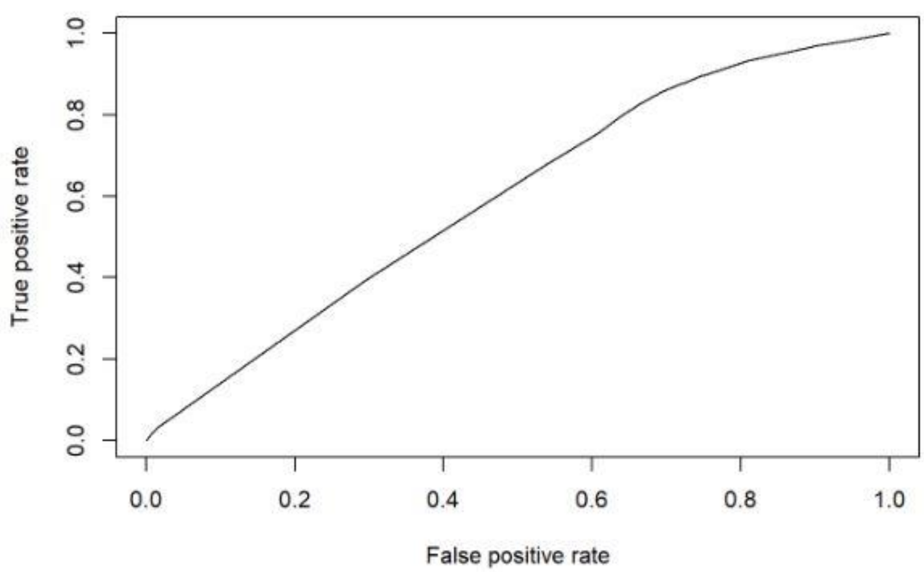

Figure 14: ROC Curve for CHEM 110 Model 3

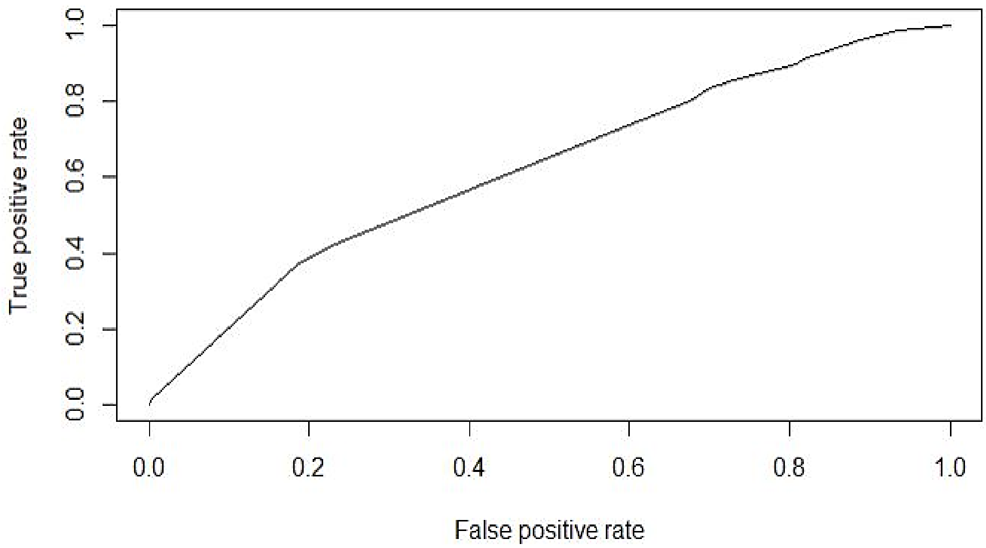

Figure 15: ROC Curve for CHEM 111 Model 3

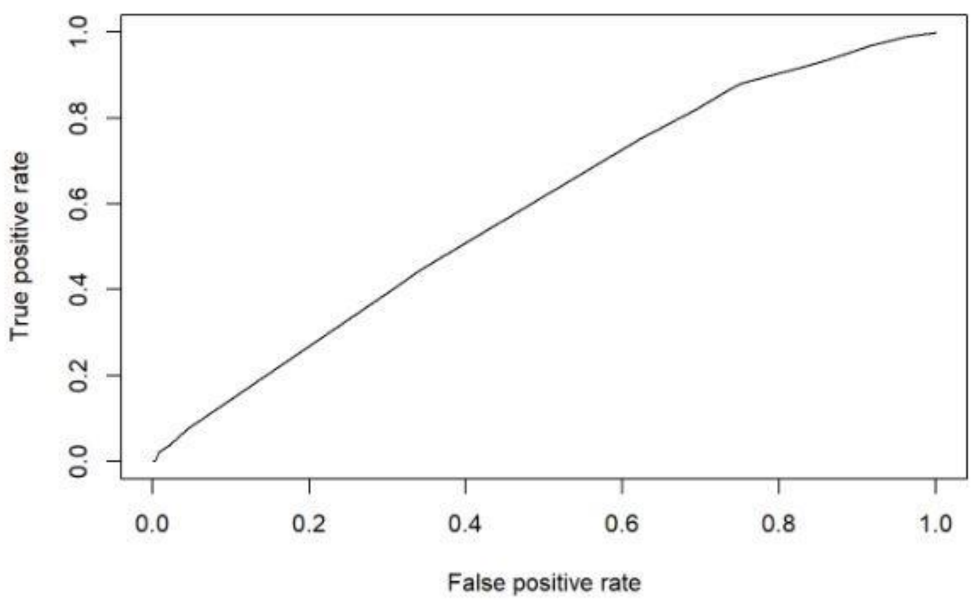

Figure 16: ROC Curve for CHEM 112 Model 4

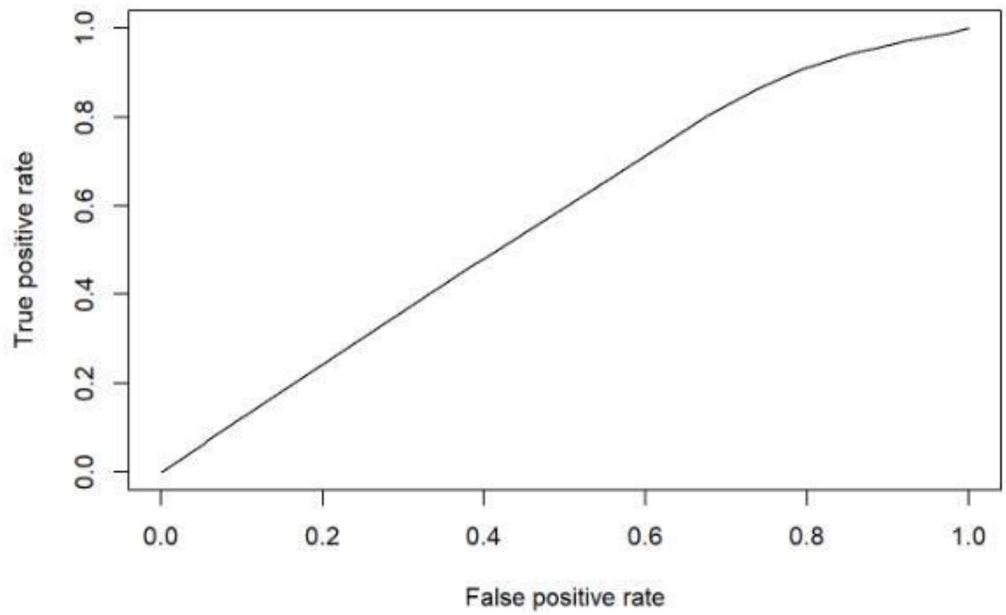

Figure 17: ROC Curve for MATH 140 Model 4

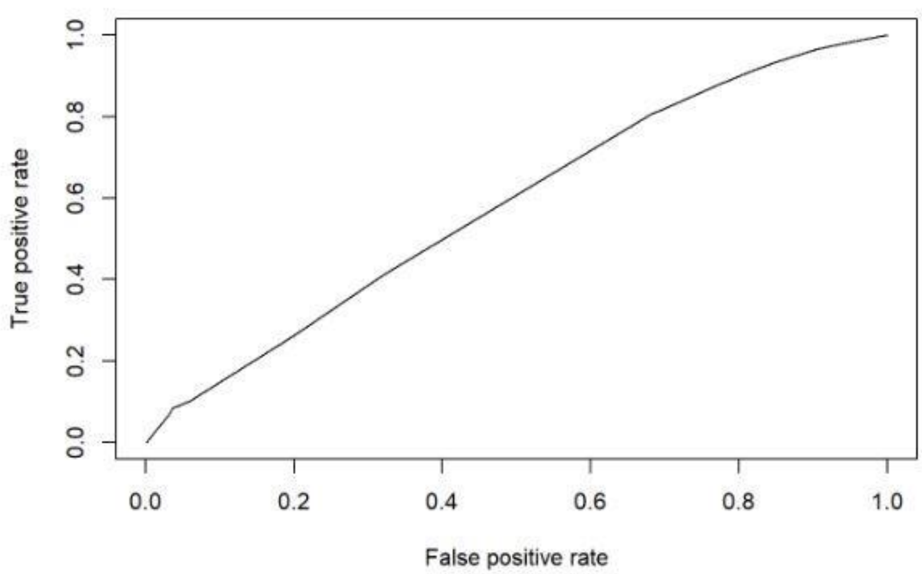

Figure 18: ROC Curve for MATH 141 Model 4

BIBLIOGRAPHY

Akaike, H. “A New Look at the Statistical Model Identification .” IEEE Xplore, Dec. 1974, ieeexplore.ieee.org/document/1100705/.

Bloodhart, Brittany et al. “Outperforming yet undervalued: Undergraduate women in STEM.” PloS one vol. 15,6 e0234685. 25 Jun. 2020, doi:10.1371/journal.pone.0234685

“Eberly College of Science.” Eberly College of Science at Penn State, bulletins.psu.edu/undergraduate/colleges/eberly-science/#majorsminorsandcertificatestext. Accessed 2 Apr. 2024.

“Graduation and Retention.” Data Digest, datadigest.psu.edu/graduation-and-retention/. Accessed 2 Apr. 2024.

González-Pérez, Susana, et al. “Girls in STEM: Is It a Female Role-Model Thing?” Frontiers, Frontiers, 5 Aug. 2020, www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2020.02204/full.

Hemmert, Giselmar A.J., et al. “Log-Likelihood-Based Pseudo-R2 in Logistic Regression: Deriving Sample-Sensitive Benchmarks .” Sage Journals, 18 Mar. 2016, journals.sagepub.com/doi/10.1177/0049124116638107.

Hosmer, David W., et al. Applied Logistic Regression . Wiley, 2013.

Huang, Ying et al. “LOGISTIC REGRESSION ANALYSIS WITH STANDARDIZED MARKERS.” The annals of applied statistics vol. 7,3 (2013): 10.1214/13-AOAS634SUPP. doi:10.1214/13-AOAS634SUPP

Menard, Scott W. Applied Logistic Regression Analysis. Sage Publications, 2002.

Ortiz-Martínez, Gabriela et al. “Analysis of the retention of women in higher education STEM programs.” Humanities & social sciences communications vol. 10,1 (2023): 101. doi:10.1057/s41599-023-01588-z

Peng, C. Y. J., Lee, K. L., & Ingersoll, G. M. An Introduction to Logistic Regression Analysis and Reporting. The Journal of Educational Research, 2002 https://doi.org/10.1080/00220670209598786

Ranganathan, Priya et al. “Common pitfalls in statistical analysis: Logistic regression.” Perspectives in clinical research vol. 8,3 (2017): 148-151. doi:10.4103/picr.PICR_87_17

Stoltzfus, Jill C. “Logistic Regression: A Brief Primer.” Wiley Online Library, 13 Oct. 2011, onlinelibrary.wiley.com/doi/pdf/10.1111/j.1553-2712.2011.01185.x.

“Undergraduate Enrollment.” Undergraduate Enrollment | Eberly College of Science, science.psu.edu/about/facts/undergraduate-enrollment. Accessed 2 Apr. 2024.

Vooren, Melvin, et al. “Comparing Success of Female Students to Their Male Counterparts in the STEM Fields: An Empirical Analysis from Enrollment until Graduation Using Longitudinal Register Data -

International Journal of STEM Education." SpringerOpen, Springer International Publishing, 3 Jan. 2022, stemeducationjournal.springeropen.com/articles/10.1186/s40594-021-00318-8.

Academic Vita

Katherine Kelly

katiiekelly@gmail.com

Education

The Pennsylvania State University University Park, PA

Schreyer Honors College *May 2024*

Bachelor of Science in Mathematics

Smeal Certificate in Business

Work Experience

Schreyer Institute for Teaching Excellence State College, PA

Research Assistant April 2022 – Present

- Worked on multiple research projects and collaborated with faculty from a wide range of disciplines to answer questions about teaching methods and student learning
- Conducted data wrangling, analysis, and visualization in Excel and R from surveys in order to find any statistically significant results
- Created reports on findings from data analysis and communicated the results to faculty in a way that they could understand

Leadership Experience

Eberly College of Science Student Council University Park, PA

Treasurer October 2021 - Present

- Managed all financial, fundraising, and budgeting records to keep track of the Council's finances and make sure it operated efficiently
- Prepared financial reports for Council meetings, helped plan fundraising events for the college, and made suggestions regarding Council budgeting and allocating resources

Processed all requests for funding for science clubs to serve as a filter for Eberly administration and streamline the funding request process

Skills and Certifications

- Proficiency in Microsoft Office, R, GitHub, Python, SQL, and Mathematica
- CITI Program Certification in Human Subjects Research and Social and Behavioral Human Subject Research Course