UNITED STATES MILITARY ACADEMY

WEST POINT, NEW YORK

HONORS THESIS

Predicting Academic Success at West Point using the SAT and High School Class Rank Scores

by

Hailey Conger

May 2019

Thesis Advisors:	COL Lindquist LTC Arney MAJ Corson
Second Reader:	LTC Florkowski

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE		Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	25MAY19	Senior Thesis (Honors)

4. TITLE AND SUBTITLE Predicting Academic Success at West Point using the SAT and High School Class Rank Scores	5. FUNDING NUMBERS
6. AUTHOR(S) Hailey Conger	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) United States Military Academy West Point, NY 10996-1786	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
Approved for public release; distribution is unlimited.	

13. ABSTRACT (maximum 200 words)

Every year many high school students prepare essays, take standardized tests, interview for nominations, and complete physical assessments to improve their chances of being accepted to the United States Military Academy at West Point. It is the Admissions team's responsibility to determine the best assessment of academic potential. The Admissions team assesses academic potential through a College Entrance Examination Rank (CEER) score. The Admissions team currently utilizes a linear equation to predict academic success developed in 1995. They input standardized test scores from the SAT or ACT and class rank percentile score into the equation to determine a score out of 800. However, the CEER equation has not been updated since 1995. There have been major changes to standardized test scores and high school rank reporting in recent years. The CEER equation has not been updated to account for these changes. This thesis focuses specifically on the 2016 updated SAT's impact on the CEER equation formulation and it will explore possible solutions that offer better predictions of academic potential. West Point does not only use standardized tests for Admissions. The USMA English Department assesses writing ability directly from the SAT essay scores. This thesis also analyzed the SAT Essay scores' predictive power for success in EN101, a freshmen level English class.

KEY WORDS: College Admission, SAT, High School Rank, Linear Regression Modeling

14. SUBJECT TERMS	15. NUMBER OF PAGES
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

PREDICTING ACADEMIC SUCCESS AT WEST POINT USING THE SAT AND HIGH SCHOOL CLASS RANK SCORES

Hailey E. Conger
CDT, Aviation
B.S., United States Military Academy, 2019

Submitted in partial fulfillment of the
requirements for the degree of
BACHELOR OF SCIENCE
In Operations Research
with Honors
from the
UNITED STATES MILITARY ACADEMY
May 2019

Author: Hailey Conger

Approved by: Major William Corson
Thesis Advisor

Colonel Joseph Lindquist
Second Reader/Co-Advisor

Lieutenant Colonel Kristin Arney
Second Reader/Co-Advisor

Colonel Tina Hartley
Head, Department of Mathematical Sciences

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Every year many high school students prepare essays, take standardized tests, interview for nominations, and complete physical assessments to improve their chances of being accepted to the United States Military Academy at West Point. It is the Admissions team's responsibility to determine the best assessment of academic potential. The Admissions team assesses academic potential through a College Entrance Examination Rank (CEER) score. The Admissions team currently utilizes a linear equation to predict academic success developed in 1995. They input standardized test scores from the SAT or ACT and class rank percentile score into the equation to determine a score out of 800. However, the CEER equation has not been updated since 1995. There have been major changes to standardized test scores and high school rank reporting in recent years. The CEER equation has not been updated to account for these changes. This thesis focuses specifically on the 2016 updated SAT's impact on the CEER equation formulation and it will explore possible solutions that offer better predictions of academic potential. West Point does not only use standardized tests for Admissions. The USMA English Department assesses writing ability directly from the SAT essay scores. This thesis also analyzed the SAT Essay scores' predictive power for success in EN101, a freshmen level English class.

KEY WORDS: College Admission, SAT, High School Rank, Linear Regression Modeling

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I. INTRODUCTION....1
A. BACKGROUND....1
B. PROBLEM STATEMENT....2
C. CONSIDERATIONS....2
1. West Point Admissions Process....2
2. West Point Admissions Requests....3
D. LITERATURE REVIEW....4
1. SAT Background....4
2. Predictive Power of the SAT....5
3. Other Factors that Predict Academic Success....5

II. DATA....8
A. OVERVIEW....8
B. KEY VARIABLES....9
1. CQPA....9
2. SAT Math Score....10
3. SAT Reading Score....10
4. SAT Essay Scores....11
5. High School Class Rank Score....11
6. Faculty Appraisal Score....12
7. High School GPA....12
C. ASSUMPTIONS....12
D. THRESHOLDS....13

III. METHODOLOGY....15
A. MULTIPLE LINEAR REGRESSION....15
B. MODEL PROCESS STEPS....15
1. Identify Variables....16
2. Predict CQPA....16
3. Transform Predicted CQPA into a CEER Score....16
4. Create CEER Formula....17
5. Assess the Model....17
C. CONSIDERATIONS....18

IV. MODELS AND RESULTS....21
A. 1995 MODEL....21
B. BASIC MODEL....24
C. HS GPA MODEL....26
D. BEST PREDICTOR....30

V. CONCLUSION....35
A. RECOMMENDATION....35
B. FUTURE WORK....35

VI. PREDICTING EN101 SUCCESS WITH SAT ESSAY SCORES....37

A. INTRODUCTION....37
B. SAT ESSAY SCORING UPDATES....37
C. METHODOLOGY....37
1. Data....37
2. Procedure....38
D. MODELS AND RESULTS....39
1. Logistic Regression....39
a. Pre-2016 SAT Logistic Model....39
b. Post 2016 SAT Logistic Model....40
2. Multiple Linear Regression....42
a. Pre-2016 SAT Multiple Linear Regression Model....42
b. Post 2016 SAT Multiple Linear Regression Model....43
E. CONCLUSION....44
LIST OF REFERENCES....45

LIST OF FIGURES

FIGURE 2. Type of CEER score used in the WCS by year [9] 3
FIGURE 3. Pre-2016 SAT [15] 8
FIGURE 4. Post-2016 SAT Scoring [16] 9
FIGURE 5. West Point CQPA scale 10
FIGURE 6. SAT Reading Concordance Table [1] 11
FIGURE 7. Faculty Appraisal Survey [9] 12
FIGURE 8. Model Process 15
FIGURE 9. CQPA to CEER Transformation 17
FIGURE 10. Accuracy Example 18
FIGURE 11. At Risk Thresholds as of 2019 19
FIGURE 12. 1995 Model's Prediction Accuracy for the Class of 2021 21
FIGURE 13. QQ Plot of 1995 Model's Prediction of the Class of 2021 22
FIGURE 14. Residual Plot of 1995 Model's Prediction of the Class of 2021 23
FIGURE 15. 1995 Model Variable Contributions 23
FIGURE 16. The Basic Model's Prediction Accuracy for the Class of 2021 24
FIGURE 17. QQ Plot of the Basic Model's Prediction of the Class of 2021 25
FIGURE 18. Residual Plot of the Basic Model's Prediction of the Class of 2021 25
FIGURE 19. The Basic Model Variable Contributions 26
FIGURE 20. HS GPA Model's Prediction Accuracy for the Class of 2021 27
FIGURE 21. QQ Plot of HS GPA Model's Prediction of the Class of 2021 28
FIGURE 22. Residual Plot of HS GPA Model's Prediction of the Class of 2021 28
FIGURE 23. The HS GPA Model Variable Contributions 29
FIGURE 24. Comparison of ACEER and SAT CEER with HS GPA in the model 30
FIGURE 25. Best Predictor Model's Prediction Accuracy for the Class of 2021 31
FIGURE 26. QQ Plot of Best Predictor Model's Prediction of the Class of 2021 32
FIGURE 27. Residual Plot of Best Predictor Model's Prediction of the Class of 2021 .32
FIGURE 28. Best Predictor Model Variable Contributions 33
FIGURE 29. Model Summary Table 35
FIGURE 31. SAT Essay Scores Predicted Probability of Passing EN101 39
FIGURE 32. SAT Writing Score Predicted Probability of Passing EN101 40
FIGURE 33. SAT Essay Analysis 41
FIGURE 34. SAT Essay Reading 41
FIGURE 35. SAT Essay Writing 42
FIGURE 36. SAT Writing and SAT Essay Scores 43
FIGURE 37. SAT Essay Sub-scores 44

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to thank MAJ Bill Corson for his advice and assistance through the past two semesters. Thank you to COL Joseph Lindquist and LTC Kristin Arney for their enthusiasm and time. COL Lindquist, LTC Arney, and MAJ Corson made up a fantastic advisor team. Thank you to MAJ Dusty Turner for helping with the statistics and coding. Thank you to the Admissions team, COL Deborah McDonald and LTC Timothy Viles for their support and insight and Mr. John Wolek for the data. Thank you to LTC Stanley Florkowksi for being a second reader and offering excellent suggestions. A big thanks to cadets, Caroline McCusker and Jimmy Palmer for sharing their findings and R code. Finally, thank you to the Department of Mathematical Science for a challenging but rewarding four years.

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

The CEER score is a metric for assessing academic potential using long-term and short-term indicators. The Admissions team uses the SAT or ACT as a snap-shot of short-term indicators and a high school class rank score as the long-term gauge. The Admissions team uses a linear equation to compute a score out of 800 with a mean of 600. The purpose of this thesis was to address two challenges with the current CEER score. The first challenge is the recent update of the SAT. In 2016, the SAT underwent major changes in format and scoring. The CEER score equation has not been updated to reflect the changes to the SAT. The second challenge is the decrease in high school class rank reporting. The Admissions team must create an artificial class rank based on the candidate's standardized test scores. This method limits the ability to assess long-term academic potential.

This thesis built and analyzed three separate models that addresses the two challenges. The first step is to use multiple linear regression to predict plebe year GPA. The next step is to scale the predicted plebe year GPA to a CEER score. The first model updated the coefficients for the SAT scores but left the high school class rank score in the equation. The second model substituted the high school class rank score for a high school GPA ratio. The candidates' high school GPAs were divided by their respective GPA scales, capped at one. The final model attempted to find the best predictor of academic potential, regardless of usability. This thesis assessed the models based on fit and accuracy. The results suggest that the first model, updating the SAT coefficients is the best solution. However, future work would be to improve a metric to replace high school class rank.

Finally, this thesis explored the uses for the SAT Essay scores for the Admissions team as well as the English Department. This research found that the SAT Essay scores do not improve the accuracy of the CEER score equation and do not help predict performance in EN101.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

Every year many high school students prepare essays, take standardized tests, interview for nominations, and complete physical assessments to improve their chances of being accepted to the United States Military Academy at West Point. It is the Admissions team's responsibility "to recruit and enroll men and women from across the nation and admit diverse, high quality candidates who meet USMA's entry qualifications and are inspired to serve as Army officers" [3]. To assess every candidate's potential, the application file is broken down into three factors: academic, physical, and leadership. Each factor is scored independently. The sum of each factor's sub-score results in a Whole Candidate Score (WCS), shown in Figure 1.

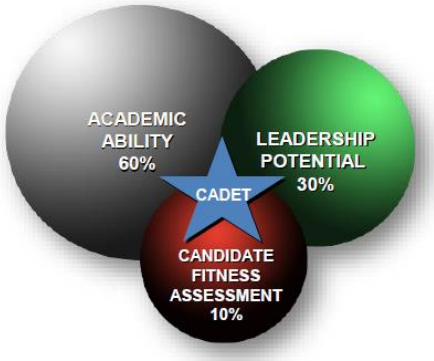

FIGURE 1. Whole Candidate Score Figure [3]

The purpose of this thesis focuses specifically on the academic ability section of the WCS, otherwise known as the College Entrance Examination Rank (CEER) score. This score is the most heavily weighted, accounting for 60% of the WCS. Therefore, it is critical that the CEER offers the best assessment of academic potential. The Admissions team currently uses a linear equation to predict academic success based on plebe year GPA. They input standardized test scores from the SAT or ACT and class rank percentile score into the equation to determine a score out of 800. This thesis focuses specifically on the SAT. The issue is that the current CEER score equation was written in 1995 [13]. Since 1995, the SAT changed numerous times, but a major change occurred in 2016. Additionally, the head of West Point Admissions reported only 60% of high school report class rank [10].

However, the CEER equation has not been updated to account for these changes. The SAT's scoring scale change and lack of high school class rank scores force the Admission's team to estimate numbers. The Admissions team converts the new scores into the old scores and they create an artificial class rank based on SAT scores and the respective national percentile ranking. The estimated numbers decrease the CEER score's accuracy.

The additional part of this research is to analyze the SAT Essay scores. The SAT Essay scores are not used in the CEER score but they are used an indicator for risk. If a cadet candidate's essay score falls under a given threshold, then the West Point English Department reads his or her SAT essays and then offers an admission recommendation. The Admission team is interesting in if the SAT Essay scores should be included in the CEER equation and the English Department is interested in the if the thresholds are appropriate.

B. PROBLEM STATEMENT

The purpose of this thesis is to update the current CEER equation so that it offers a better prediction of academic success. The updated CEER equation will account for the recent SAT score change and for the lack of high school class rank. The process is first to assess the 1995 equation, second to update the coefficients of the CEER equation, third to create the best CEER model to predict success, and finally recommend possible updated CEER equations.

The secondary purpose of this thesis is to assess the predictive power of SAT Essay scores for EN101, an introductory level English class. The Admissions team and the English department work together to assess the admissions and placement of candidates based on their writing abilities and potential.

C. CONSIDERATIONS

1. West Point Admissions Process

As mentioned before, West Point uses the CEER score to measure academic aptitude. These numbers come from a linear equation from a 1995 study using high school rank and SAT scores. The CEER equation was later updated to include the ACT. For students who take both exams, the Admissions team compares the ACT and SAT CEER scores and takes the higher score of the two. There has been a steady decrease in SAT

CEER scores used in Admissions. Notice the clear decrease in SAT CEER scores used in Figure 2.

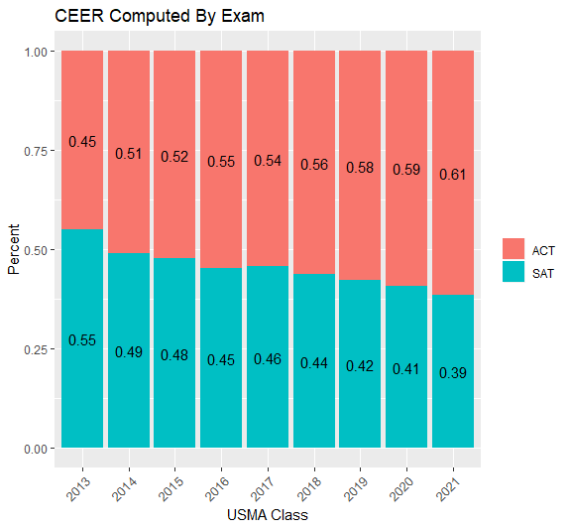

FIGURE 2. Type of CEER score used in the WCS by year [9]

This may be due to a shift in preferred standardized tests or perhaps the CEER scores calculated from the ACT and SAT may differ. There may be a difference in the distributions, so a 520 SAT based CEER score may not mean the same as 520 ACT based CEER score.

The SAT CEER equation uses the SAT Critical Reading, SAT Math, and high school class rank score. The Admissions team wants to account for long and short-term academic success [10]. The standardized test scores represent short-term success and the high school class rank represents long-term success. In addition, if a high school does not report class rank then the Admissions team predicts a rank based on a candidate's standardized test scores. This method hinders the ability to assess long term academic success and double counts standardized test scores.

2. West Point Admissions Requests

The Admissions team asked for an easily implementable equation update for the CEER. This will increase its use and understanding. For example, in 2012, USMA Mathematics Department built a new CEER equation. This equation was better at determining success in at-risk cadets and accounts for the quality of the school in measuring class rank [6]. Although CEER12 was not implemented, this thesis will consider

continuity and ease of implementation [10]. The new equation should be easily understood and easy to explain so that it helps the Admissions team and the cadet candidates.

D. LITERATURE REVIEW

College admission teams are tasked with a daunting challenge of predicting success. Success can be defined by many different measures, such as graduation, college GPA, post-graduation success, and so on. The challenge is how to predict academic achievement. Common measurements are high school rank, high school GPA, and standardized test scores, such the ACT and SAT. It is important to consider what each number informs the Admissions team. The head of Admissions stated in a personal interview that the Admissions team is concerned about long term and short-term academic achievements [10]. She explained that a standardized test score is snap shot view of a candidate. There are many confounding factors that could have influenced the scores, such as fatigue and personal distractions. Class rank and GPA provide indicators of long-term success because the influence of outside factors diminishes over time. Class rank and GPA are accumulations of successes and failures, averaging out to capture a candidate's academic capabilities. Admissions team defines academic success as a passing plebe year GPA, over a 1.67 GPA [10]. The goal of the CEER is to consider both short-term and long-term success in candidates to accurately predict if they will be successful at West Point. The CEER score is also used as a flag for at risk candidates. Because the CEER score is standardized across all years, the Admissions team can have an idea of where the candidate may fall academically in his or her class.

1. SAT Background

According to a PBS article, the SAT was based on an Army intelligence test from WWI, called the Army Alpha [4]. The Army Alpha was revised to be used for college admissions. It was first administered in 1926 but grew into a Harvard scholarship exam in 1933, and a couple years later was adopted by the College Board in 1944. The Harvard president liked the test for its focus on intelligence, rather than quality of high school education. Since then, the SAT has become a well-known test and a given requirement for most college applications. It has been updated, revised, and remodeled over the last 20 years. The most prevalent was in 2016. This revision altered the point system, eliminated the penalty for guessing, and made the essay optional [5]. College Board published

Concordance tables to convert new and old scores. However, there may be a discrepancy between the new scores and their predictive power. Therefore, it is important to avoid converting scores, hence the need for an updated CEER formula.

2. Predictive Power of the SAT

In a 2014 study of prediction variables and college success, Zahner, Ramsaran, and Steedle analyzed high school GPA, college entrance exams (SAT/ACT), and an open-ended, performance-based assessment of critical thinking and writing skills [19]. They compared the predictive power of these factors to each college year's GPA. This study suggests that standardized tests might not be the best predictor of success. They determined that high school GPA was the best predictor of overall college GPA. However, the critical writing assessment, when coupled with the high school GPA, is better at predicting senior year college GPA than standardized tests. They concluded that a critical writing assessment may offer different information than standardized tests. Yet, when considering change, Admission teams are concerned about ease of implementation. The predictive power of other variables, such as a critical writing assessment, should be significantly more powerful than standardized tests. This study supports the significance of the essay scores and suggests their place in the CEER formula. It also supports high school GPA's potential as a replacement for the high school class rank score.

3. Other Factors that Predict Academic Success

This section explains a series of studies that either relate to updating the CEER formula or offer insight to other considerations about predicting academic success that may be useful for future research.

Geiser conducted a study that analyzed the relationship between race and SAT scores [7]. He looked at socioeconomic background factors, family income, parental education, and race/ethnicity to predict SAT score. He studied data from 1994-2011 of California residents who applied to the University of California. He found that a third of the SAT score variation can be attributed to characteristics known at birth, with race being the most significant. He acknowledged that this study pertains solely to California and more research must be done to determine if the trends expand to the rest of the country. This study is significant because it supports the importance of updating the CEER formula.

If the SAT has some correlation with non-academic factors, such as race, then the SAT's use in the CEER equation may unknowingly influence the diversity of the incoming class.

While the SAT scores may be skewed information, the high school rank percentile may not offer enough about academic potential. It may be necessary to consider more than just high school rank percentile. Young conducted a study that looked at the quality of high schools and the trend of college admission into the various campuses of the University of California [18]. His goal was to challenge colleges to consider the quality of the high school. There are two points of view. One, if a student goes to a higher achieving school, he or she will be more likely to apply to and attend a high achieving college. Two, if a student goes to an under achieving high school, they would be more competitive for a selective college because of their inflated class rank. However, there are not very many studies that address the actual impact of using high school quality in admissions. The University of California Admissions team uses 14 criteria, which includes quality of senior year program, recent improvement in academic performance within the last two years of high school, academic accomplishments considering life experiences and special circumstances, such as disabilities, low family income, or first-generation college attendant, and the location of high school and residence. He suggested that parents and policymakers consider how school quality and size affect admissions. This study provides insight that the Admissions team may need to consider the quality of high school or the size of the class regarding high school class rank. A student who is in the top third of a high school with 1,000 students per grade may have different academic abilities than a student who is in the top third of a high school with 12 students.

The significance of these studies demonstrate that SAT scores and high school rank are not the sole predictors of academic success and it is important to understand the potential impact on admission, especially diversity.

High school GPAs may provide the same information of long-term academic success and could be used to predict college academic success. The issue is the numerous types of high school GPAs, ranging in scale and weight versus unweighted. This makes it difficult to input GPAs into a standard linear equation. It also does not account for grade inflation. However, class rank solves the problem of grade inflation because it compares

them to other people in their class. The issue with using class rank is that more high schools are not reporting it. The Admission team creates an artificial class rank, derived from the SAT or ACT scores. This technique drifts further and further from the actual rank and therefore does not appear to be an optimal technique.

In 2010, Walker conducted a study, researching the impact of not having class rank on admission to college, the number of schools who no longer rank their students, and if the lack of class rank reduces stress and allows better course selections [17]. National Association for College Admission Counseling reported that more than 50 percent of high school no longer use class rank and less than 20 percent of colleges say it is important in their decisions. The movement began with small private schools because they wanted their students to try difficult courses and has since expanded to large public schools [17].

Therefore, the introduction of high school GPA into the equation would challenge the Admission team's current formula structure but may ultimately be a better predictor of success.

II. DATA

A. OVERVIEW

This thesis had access to Admissions data from the Classes of 2013-2021.This consisted of 21,330 cadets and over 50 variables. While it was de-identified, it contained standardized test scores and high school class ranks, as well as information about nominations, high school sport and extracurricular participation, leadership positions, community service, and prior service. However, the data used in this project is from the Class of 2021 because they are the first class to have taken the updated SAT. 1,254 candidates were admitted into the class. Of the Class of 2021, this research only used the data from candidates whose CEER score came from the SAT.

This project considered much of the provided Admissions information. The most important variables are described in this section. The SAT updated scores are outlined below. The pre-2016 SAT originally had three sections, Math, Writing, and Critical Reading, as shown in Figure 3. They were all scored out of 800, with a total score of 2400.

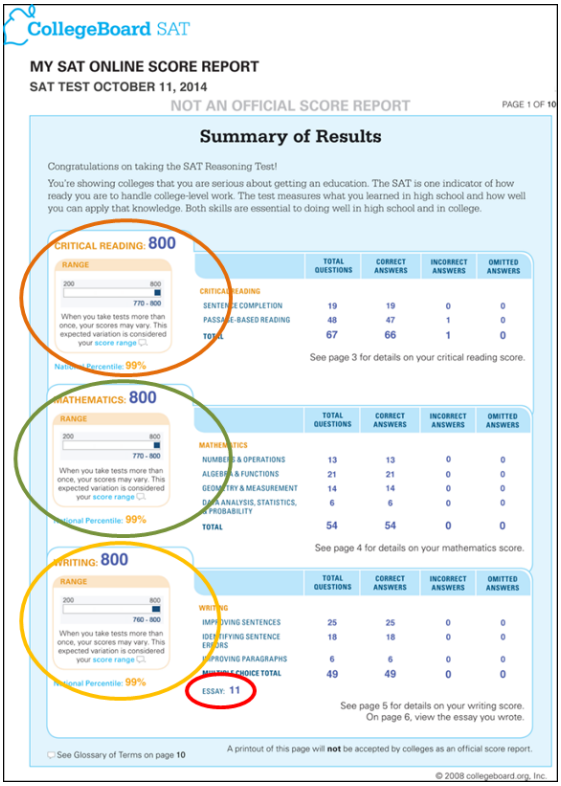

FIGURE 3. Pre-2016 SAT [15]

The post 2016 SAT now has two sections scored out of 800, Math and Evidence-Based Reading and Writing, with a total score of 1600. Each of the two sections has sub scores, as shown in Figure 4. The Admissions team would like to use the Reading score instead of Evidence-Based Reading because previous analysis showed it is a better option [2].

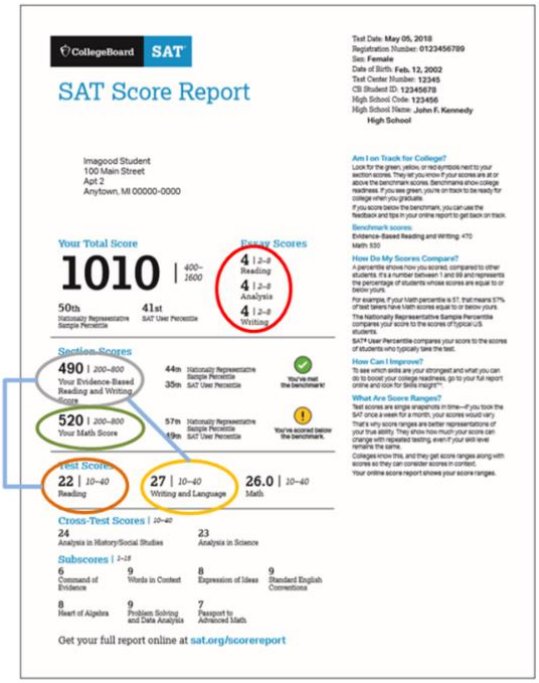

FIGURE 4. Post-2016 SAT Scoring [16]

The Admissions team converts new scores to old scores to fit the current CEER model. The conversion process loses valuable information, supporting the need for an updated CEER formula using raw SAT scores.

B. KEY VARIABLES

1. CQPA

CQPA is a cumulative academic GPA on 4.00 scale. The CEER score is based on a predicted plebe year CQPA. The CQPA measures a weighted grade point average based on credit hours. While the CQPA does not consider honor or accelerated classes, an A+ is a 4.33.

Letter Grade	GPA	Letter Grade	GPA
A+	4.33	**C+**	2.33
A	4.00	**C**	2.00
A-	3.67	**C-**	1.67
B+	3.33	**D**	1.00
B	3.00	**F**	0.00
B-	2.67		

FIGURE 5. West Point CQPA scale

2. SAT Math Score

The raw SAT Math score is on the same 800-point scale as the pre-2016 SAT. The content of the SAT Math test changed and the scaling of the scores shifted slightly. Therefore, the Admissions team must convert the new SAT Math scores to the old scores even though both scores are out of 800.

3. SAT Reading Score

The raw SAT reading score is a subsection of the Evidence-Based Reading and Writing section that replaced the SAT Critical Reading Score. The Evidence-Based Reading and Writing score is comprised of a reading and writing score. Both the reading and writing scores are out of 40. The Admission's team would like to use the SAT Reading score instead of the Evidence-Based Reading and Writing score because previous analysis showed that it was the strongest indicator of academic potential [2]. Currently, the Admissions teams converts the SAT Reading score to the SAT Critical Reading Score using Concordance tables as shown in Figure 6 [1].

Old SAT Critical Reading Section to New SAT Reading Test Concordance Table	
Old SAT Critical Reading Section (200-800)	New SAT Reading Test (10-40)
200	10
210	11
220	12
230	13
240	13
250	14
260	15
270	16
280	17
290	17
300	18
310	18
320	18
330	19
340	19
350	19
360	20

FIGURE 6. SAT Reading Concordance Table [1]

4. SAT Essay Scores

The new SAT Essay score are comprised of three individual scores that measure different writing skills. The three individual scores are reading, writing, and analysis. The scores are out of 8. Each essay has two graders that give a grade out of 4. The two scores are added to provide a combined score out of 8 for each section. Each individual score measures something different about the essay. The reading score assesses how well the student understood the passage. The writing score measures the student's writing style and handling of the English language. The analysis score determines how well the student created and supported his or her argument [14].

The CEER currently does not consider the essay scores. However, the Admissions team set a threshold to identify candidate at risk for failing EN101, the entry level English class. If a candidate scores under the essay score threshold, a member from the English Department will read the essay and offer a recommendation. The Admissions team also wondered if the SAT Essay should be included in the CEER calculation. The SAT Essay is now optional for the SAT but mandatory for Admissions.

5. High School Class Rank Score

The High School Class Rank Score (HSR) is a score out of 800, calculated from a candidate's rank within his or her high school class. The current method uses an extensive computer program to assign a score. If a candidate's high school does not rank, the Admissions team creates an artificial class rank based on their standardized test percentiles. As more and more high schools stop ranking their students, it is becoming more important to find an alternative method to measuring high school success.

6. Faculty Appraisal Score

The Faculty Appraisal Score (FAS), out of 800, is calculated from high school teacher's survey responses about the candidate's performance as a student and member of a classroom. The scoring may be subjective and are frequently very positive, comments offer helpful insight [2]. The score is not currently being used in the CEER. Rather, the Admissions team is using it to measure grit in the community leadership score. However, given the its relation to academic potential, it is worth considering. An example faculty appraisal survey is shown in Figure 7.

1= Agree Strongly; 2= Agree; 3= Neither agree nor disagree; 4= Disagree; 5= Disagree Strongly

This candidate has demonstrated an ability to:

answer is 1	1. Show interest and concern for the welfare of others.
answer is 1	2. Work effectively with others toward group goals.
answer is 1	3. Influence others in a positive manner.
answer is 1	4. Communicate effectively in face to face discussion.
answer is 1	5. Communicate effectively in written work.
answer is 2	6. Set an example of good conduct for others.
answer is 1	7. Set high standards for own performance in a number of activities.
answer is 1	8. Maintain composure and perform effectively under pressure.
answer is 2	9. Adjust to demanding schedule of activities without neglecting school work.
answer is 2	10. Seek academic challenge beyond that required by normal course work.
answer is 1	11. Reach sound logical conclusions based on analysis of facts.
answer is 1	12. Accept full responsibility for own actions.

FIGURE 7. Faculty Appraisal Survey [9]

7. High School GPA

High school GPA is not part of the CEER score other studies suggest this is an important measure of high school success. High school GPAs are complex because of numerous scales and different calculation methods. The weight and unweight GPAs make it difficult to measure success. For this project, the high school GPAs were standardized by taking a ratio of the candidate's high school GPA over his or her high school's GPA scale. This project does not consider weighted GPAs. The GPA ratios are capped at one.

C. ASSUMPTIONS

Out of 1,254 cadets admitted into the Class of 2021, the final dataset considered is 310 cadets. The significance decrease was due to missing information and filtering for SAT CEERs. Because of the decreased used SAT CEERs, this thesis only considers to cadets who's higher CEER score came from their SAT. This is to filter for discrepancies between standardized tests and for cadet candidates' half-hearted testing attempts due to having a

high enough ACT score. This project assumes the 310 cadets provide an accurate representation of the Class of 2021.

Another assumption is that a 1.67 cumulative GPA (C-) or greater is considered passing. This is an assumption because there is some instructor discretion to fail a cadet. That means that if a cadet has an average of a D or F GPA, then multiple instructors felt the cadet did not earn a passing grade. It is worth considering a passing grade as higher than a C- because to earn a C-, a cadet may still have failed a class. Still, for the purposes of this thesis, a passing GPA is a C- average.

Finally, an additional assumption is that the relationship between the variables and academic success is linear. A linear equation is easier to explain and justify. It is easier to understand that high test scores and high school class rank score mean higher CEER score.

D. THRESHOLDS

The Admissions team sets identifiers of academic risk based on standardized test scores and CEER scores. If a candidate falls under the threshold, his or her file is reviewed with greater scrutiny. He or she is then either recommended for prep school, direct admission, or rejection. According to Admissions, the CEER thresholds should roughly represent the bottom ten percent of the incoming class. The standardized test score thresholds are identifiers for risk. These numbers do not necessarily apply to the ten percent goal [2]. This research will examine these thresholds in conjunction with the new proposed models.

THIS PAGE INTENTIONALLY LEFT BLANK

III. METHODOLOGY

A. MULTIPLE LINEAR REGRESSION

The methodology used in this project follows the 1995 linear model creation. Every model is built using a standard linear model,

$$y = a_1 * x_1 + a_2 * x_2 + .. + a_n * x_n + b$$

where *y* is the output value, in this case the CQPA. The a_n is the coefficient of a variable x_n which could be a standardized test score, high school class rank score, or any combination of the previously explained key variables and *b* is some constant. The CEER equation is a multiple linear regression model. A multiple linear regression is a model that takes known inputs and forms a linear relationship to predicts an outcome. The goal of multiple linear regression is to account for as much of the output's variance as possible with the known inputs.

$$\hat{y} = \beta_1 * x_1 + \beta_2 * x_2 + .. + \beta_n * x_n + \alpha$$

Therefore, $\hat{y}$ is the predicted output value. Any deviation from the actual and predicted outputs stems from a lack of known information.

B. MODEL PROCESS STEPS

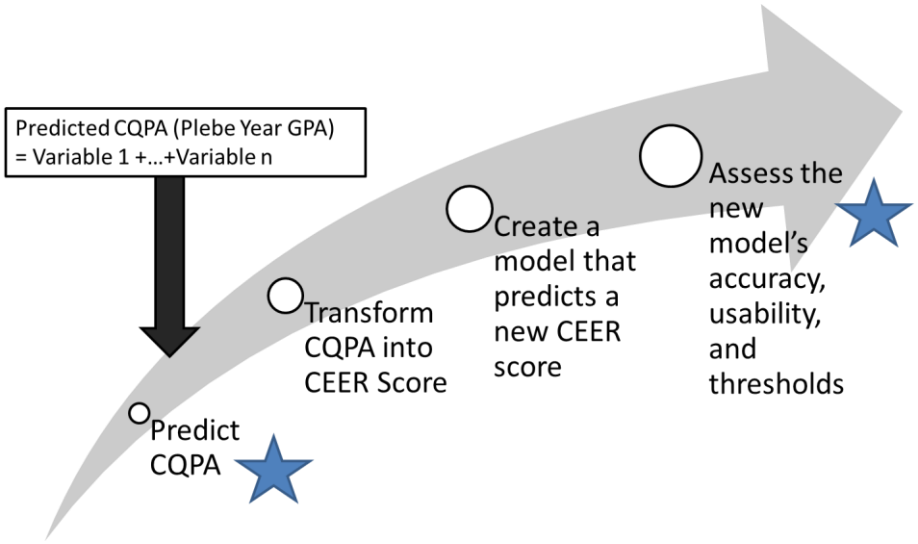

FIGURE 8. Model Process

There are five steps to creating a CEER formula. See Figure 8. This process was used for each of the models. It is the same steps that created the 1995 model. The stars indicated the key steps in the process.

1. Identify Variables

The first step is to identify the variables that will be part of predicting the candidate's CQPA. The current options include SAT scores and high school class rank. This research also considers faculty appraisal score and high school GPA. The Admissions team asked for updated coefficients to the current model. Therefore, the first model must include SAT Math, SAT Reading, and high school class rank score. The remaining models are open to exploration.

2. Predict CQPA

The second step is to predict CQPA using the previously identified variables. This is an important step because the CEER score's accuracy is currently measured by the model's ability to predict a candidate's academic performance. The predicted CQPA is a linear equation like the CEER score equation, based on the inputted variables. There will be some variation between the two different scores. Linear regression tries to account for the maximum variation of the outputs. The fit of this model is important because it is the basis for the CEER score.

3. Transform Predicted CQPA into a CEER Score

Next, the predicted CQPAs are transformed into a CEER score. The predicted CQPAs are linearly transformed into CEER scores and then substituted into a linear model to create the final equation. There are known CEER parameters, such as a maximum of 800, a mean of 600, and a minimum of 200. Given these boundaries, the CQPAs are assigned a CEER score. The highest CQPA will receive an 800 and average CQPA will receive a 600. The rest of the CQPAs fall in according to where they rank between the maximum and mean.

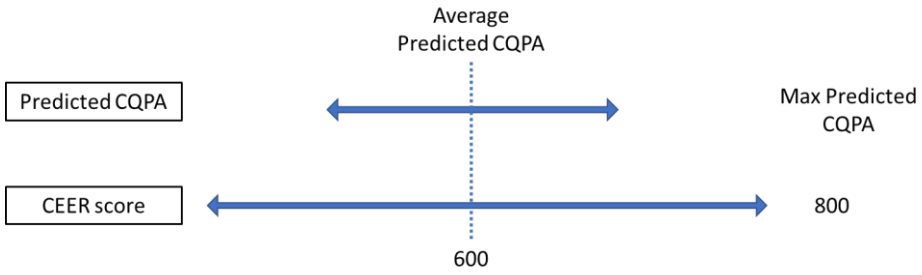

FIGURE 9. CQPA to CEER Transformation

While the CEER is bounded at 200, this project did not consider the lower boundary because all the candidates admitted scored well above 200. The predicted CQPA, after being transformed into a CEER sore, is now a predicted CEER score.

4. Create CEER Formula

Finally, the newly predicted CEER score is inputted back into the linear model so that the coefficients on the variables are updated. Consider the current CQPA equation and the resulting CEER score equation.

$$Pred\ CQPA = 0.002283 * SAT\ Math + 0.001421 * SAT\ Critical\ Reading + 0.001926 * HSR - 0.6865$$

$$CEER = 0.432 * SAT\ Math + 0.269 * SAT\ Critical\ Reading + 0.364 * HSR - 48$$

The CQPA equation is multiplied by 189.303 to create the CEER equation.

5. Assess the Model

There are a couple measures to assess the model's fit. A model often does not account for all possible factors. There will always be variation. The goal is to find where the model falters and determine whether that is acceptable for the model. R^2 is commonly used to assess a model's fit and it used in this research.

Another measure of fit is accuracy. For the project, accuracy is designated as the model's ability to predict a candidate's CQPA within 0.33 of his or her actual CQPA. This means that if a cadet has a 3.33 CQPA, an accurate model will predict the cadet having a CQPA between 3.0 and 3.67.

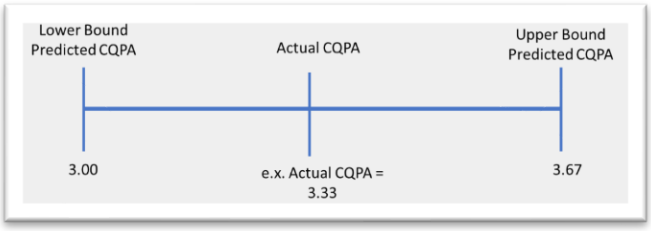

FIGURE 10. Accuracy Example

Therefore, if the difference between predicted CQPA, $\hat{y}$ and the real CQPA, y is Δy, then Δy must be less than the tolerance, ε. For this project, ε is 0.33. If Δy is less than 0.33, the model is considered to accurately predict the CQPA.

$$\Delta y = \hat{y} - y$$

$$\Delta y < \varepsilon$$

This measure of accuracy accounts for expected variation. The goal of the CEER score is to estimate general academic performance. The nuances of the CQPA's second decimal place are not important to the model's purpose. However, mathematically the nuances do affect other measures, like r^2. That is why it is important to have multiple measures of fit.

Finally, residuals are another factor to consider when assessing a model's fit. Residuals are the difference between the predicted and actual outputs. Ideally, the model is inaccurate across the spread of data. That means a model's accuracy does not favor certain combinations of data. It is important to know if the model favors certain candidates. Common ways to assess residuals are QQ plots and a scatter plot to see if the data points are scattered evenly above and below the predicted linear line. QQ Plots describe the distribution of Δy and they will reveal if the model skews the outputs.

C. CONSIDERATIONS

There are a couple considerations that fall outside mathematical implications. The first is the ease of implementation or usability of the model. The Admissions team must consider the available information and how the updated model affects their process. The model's ease of implementation affects access of required variables, understanding of the score and coding. This project does not consider information that is currently not being

collected. However, the outputted new CEER score should ideally follow the Admission's team familiar distribution of CEER scores.

Second, the Admissions team must be aware of any change in the new CEER score distribution. A lower score from the old model may not mean the same in the new model. Therefore, it is important to know if the Admissions team's bottom 10% thresholds changed with the new models.

For reference, the current thresholds are given below.

At Risk Thresholds (< 10% of Admitted Class)	
SAT Reading	31
SAT Math	580
CEER	520

FIGURE 11. At Risk Thresholds as of 2019 [2]

Third, the SAT CEER must align with the ACEER. A SAT produced CEER of 500 must mean the same as an ACT produced CEER of 500. This is not an issue if the candidate took both tests because the Admissions team takes the higher CEER. However, if the candidate only took one of the standardized tests, he or she would be at a disadvantage. The main comparison of the SAT and ACT CEERs is their score distribution. There should be similar proportion of candidates for each score. For example, if the Admissions team would like 10% of the incoming class under 520, then the new SAT CEER and ACT CEER equations should predict about 10% of the class scoring under a 520.

Finally, another consideration is the contribution of the variables to the CEER score. Some variables may contribution more than others. The variables that contribute more to the CEER score help identify what the CEER score is measuring. If high school class rank contributes 60% of the CEER score than it is important to make sure the high school class rank is an accurate representation of a candidate's academic potential.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. MODELS AND RESULTS

A. 1995 MODEL

The first model is the current model. It was created in 1995, following the same steps described in the Methodology section. The predictors are SAT Math, SAT Critical Reading and high school class rank score. The following equations are listed below.

$$Pred\ CQPA = 0.002283 * SAT\ Math + 0.001421 * SAT\ Critical\ Reading + 0.001926 * HSR - 0.6865$$

$$CEER = 0.432 * SAT\ Math + 0.269 * SAT\ Critical\ Reading + 0.364 * HSR - 48$$

Given an equation already exists, the next step is to assess the fit of the model. First is the r^2. The r^2 from 1995 is 0.62. That means this model can predict 62% of the CQPA's variation. However, using the current data from 2021, the r^2 is 0.47. The drop in the model's fit of the data further supports the need for an updated model.

The current model predicts a candidate's score within 0.33 of their actual GPA 46% of the time. This means that the model does offer some information, but it will have a GPA prediction error greater than 0.33 more time than not. Note Figure 12. The x-axis is the Class of 2021's actual CQPA and the y-axis is the Class of 2021's predicted CQPA using the 1995 model. The red dots reflect the cadets who's predicted CQPA was outside the 0.33 threshold.

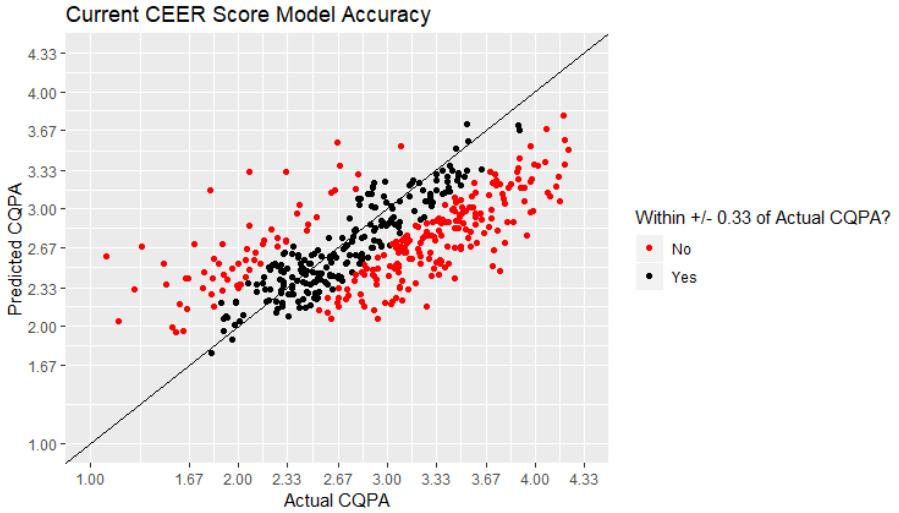

FIGURE 12. 1995 Model's Prediction Accuracy for the Class of 2021

Notice the model tends to under predicted cadets who achieved higher CQPAs and overpredict cadets who achieved lower CQPAs

This observation is validated through the QQ plot as shown in Figure 13. Ideally, all the points would be within the dashed lines which represent what is acceptable variation. Note the collection of points at the plot's corners.

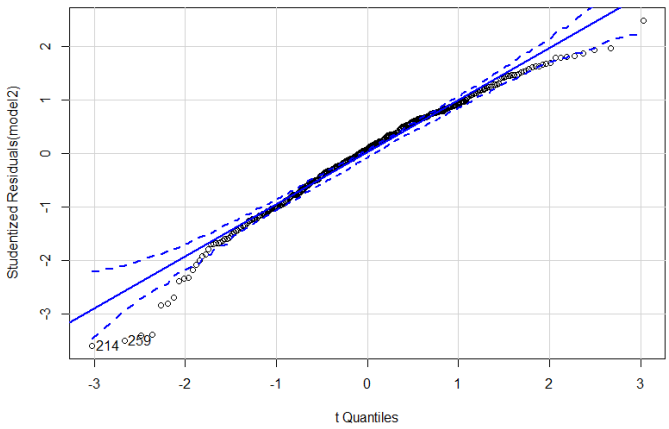

FIGURE 13. QQ Plot of 1995 Model's Prediction of the Class of 2021

Another way to assess the residuals distribution is using a residual plot. The plot below is a residual plot of the 1995 model. The x-axis is actual CQPA and the y-axis is the standardized differences between predicted and actual. The blue line shows how the model tends to underpredict because it is curved downward. Ideally, the blue line would be straight along the 0 line and the point evenly distributed above and below the line and across the fitted values.

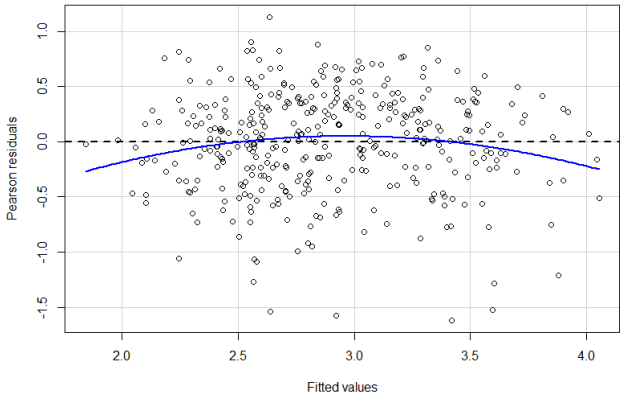

FIGURE 14. Residual Plot of 1995 Model's Prediction of the Class of 2021

The next consideration is the SAT Math, SAT Critical Reading, and high school class rank score contributions to the CEER Score.

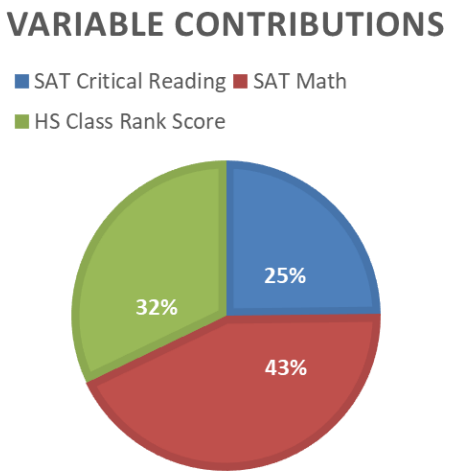

FIGURE 15. 1995 Model Variable Contributions

As shown in Figure 15, the SAT Math offers the most weight with high school class rank score being the second most weighted factor. These contribution percentages provide good information about what is most important under the old model to predict academic success and offer a good way to compare the other models we are analyzing.

The thresholds of 1995 SAT CEER also shifted. The bottom ten percent of the class has CEER scores under 500. The overall class CEER score distribution is still at the

Admission's goal of 520. This means that the SAT CEER score distribution may be different than the ACT CEER score distribution. This discovery supports the need for a new CEER equation.

B. BASIC MODEL

The purpose of the Basic Model is to update the coefficients of the current equation. The variables are the SAT Math, SAT Reading, and the high school class rank score.

$$Pred\ CQPA = 0.0040 * SAT\ Math + 0.\ 0340 * SAT\ Reading + 0.00174 * HSR - 1.8383$$

$$CEER = 0.6315 * SAT\ Math + 5.33 * SAT\ Reading + 0.2724 * HSR - 136.3$$

Overall, the model has a better accuracy and fit than the current model. The r^2 is 0.49. The accuracy improved significantly from the 1995 model, jumping from 46% to 58%.

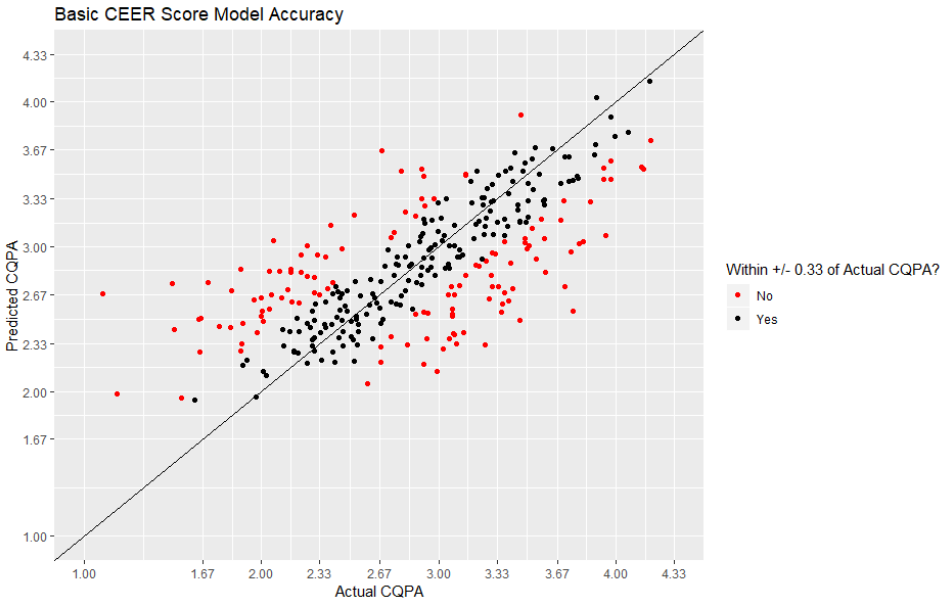

FIGURE 16. The Basic Model's Prediction Accuracy for the Class of 2021

The spread of the points shown in Figure 16 demonstrates a more even distribution. There is still a tendency to under predict CQPA. However, it is less pronounced. This also could be due to decrease in data points.

The residuals are better than the 1995 model. A few points fall outside the acceptable boundaries. However, this may not be a problem when the model is tested with new data.

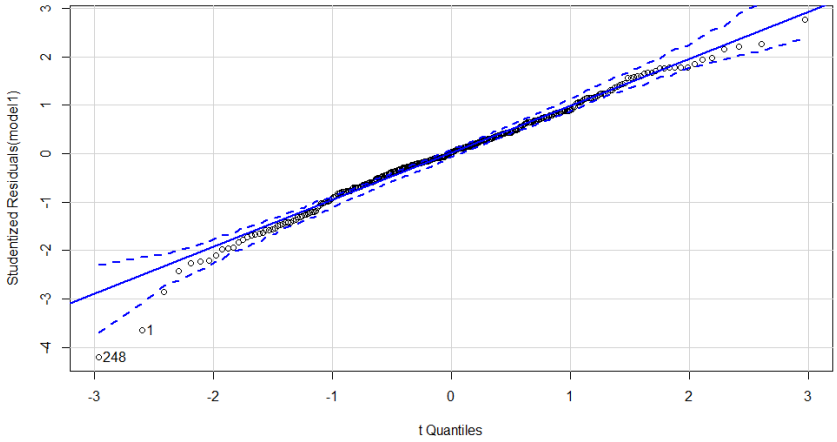

FIGURE 17. QQ Plot of the Basic Model's Prediction of the Class of 2021

Regarding each of the variable' residuals, ideally they would be evenly distributed. Because the blue lines are straight, there is not any issue with the variables.

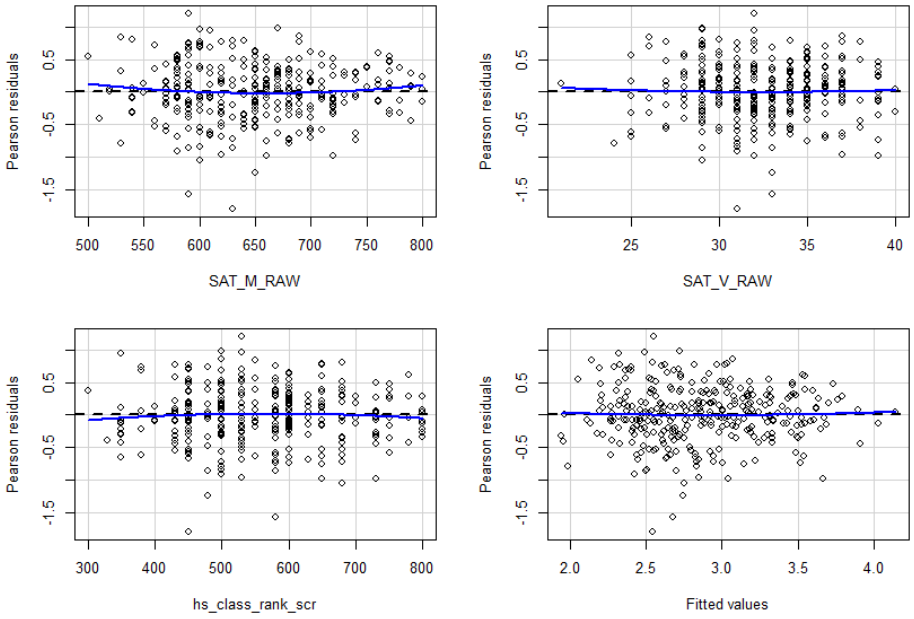

FIGURE 18. Residual Plot of the Basic Model's Prediction of the Class of 2021

This means that this model is an improvement from the 1995 model and is mathematically sound. The main issue with this model is its sustainability. This model solves the first goal of the updated SAT scores, but it does not address the high school class rank score. It would be easy to implement because it is a simple matter of updating the coefficient in the code.

Regarding the variables' contributions, this model is like the 1995 model. The SAT Math is the greatest contributor, followed by the SAT Reading and high school class rank score. The high school class rank score is still contributing about a quarter of the CEER score.

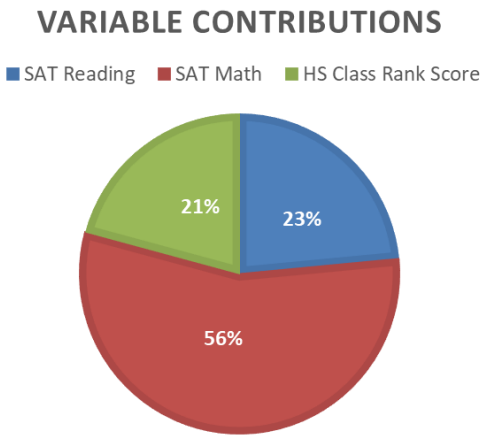

FIGURE 19. The Basic Model Variable Contributions

Regarding the new thresholds, the bottom 10% have a 520 or under. This raises the threshold from 500 back to the expected 520. Using this equation may even out the use of ACT and SAT CEERs because the scores are similar again.

C. HS GPA MODEL

This next model attempts to address the high school class rank issue by replacing the high school class rank score with high school GPA. The challenge with using high school GPA is the diversity of GPA. Some high schools use a 4.0, 5.0, 6.0, or 100. Some weight honors and AP classes and other do not. To standardize high school, we took the ratio of each cadet's high school GPA over his or her school's GPA scale. We capped the ratios at one, and so we did not consider weighted GPAs. The resulting equations are shown below.

$$Pred\ CQPA = 0.0043 * \ SAT\ Math + 0.\ 0362 * SAT\ Reading + 1.275 * HS\ GPA - 2.2996$$

$$CEER = 0.861 * \ SAT\ Math + 7.235 * SAT\ Reading + \ 255.0 * HS\ GPA - 432.8$$

The model's fit decreased with a r^2 of 0.43, which is lower than the current model at 0.47. However, the accuracy did not decrease as much. This model can predict a candidate's CQPA within 0.33 56% of the time.

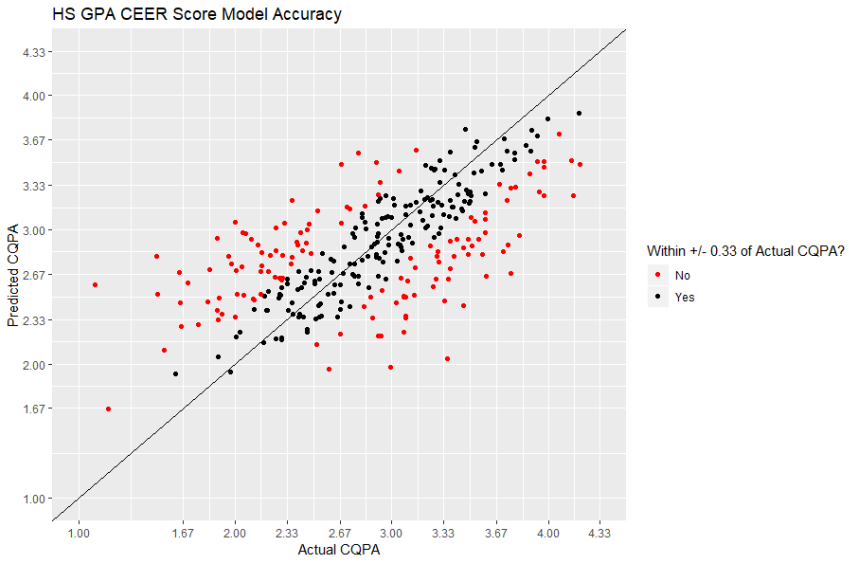

FIGURE 20. HS GPA Model's Prediction Accuracy for the Class of 2021

The spread of points is like the other two models. However, the residuals are not as optimal. The QQ plot does not appear different from the other models. The variables' residuals are more concerning.

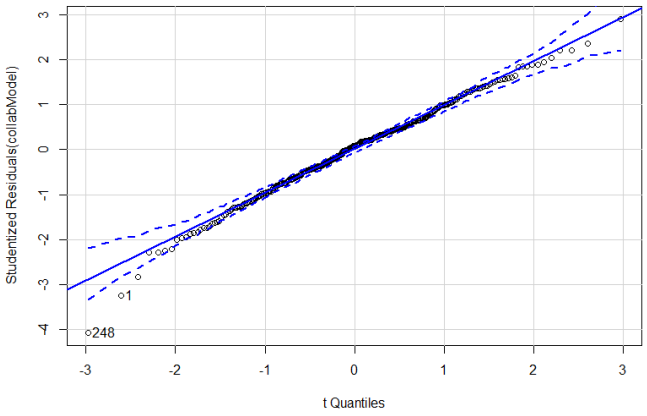

FIGURE 21. QQ Plot of HS GPA Model's Prediction of the Class of 2021

The SAT scores are curved more than the other models. This means that the model does not predict with the same accuracy for each combination of variables. This accounts for the lower r^2 value.

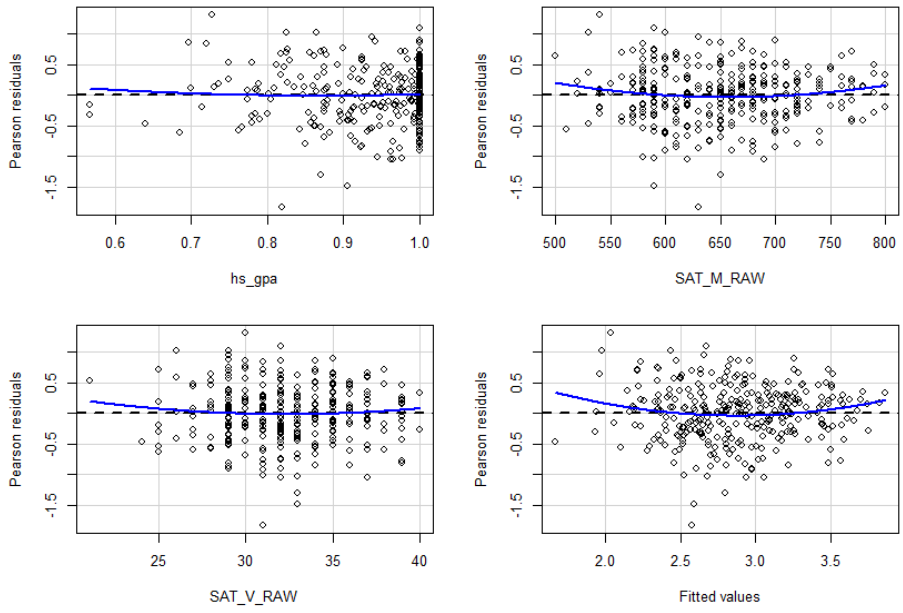

FIGURE 22. Residual Plot of HS GPA Model's Prediction of the Class of 2021

However, the relatively high accuracy justifies the potential of high school GPA as a replacement for the high school class rank score. The variable contributions remain

similar to the other models. The SAT Math offers the most information, followed by SAT Reading and high school GPA.

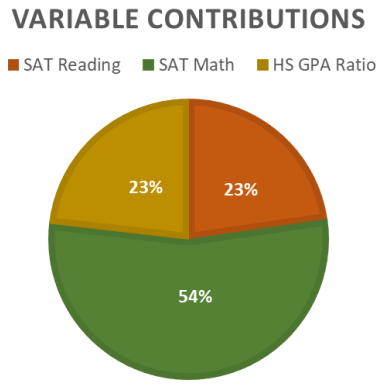

FIGURE 23. The HS GPA Model Variable Contributions

This model's thresholds are different than the current thresholds. The bottom 10% of the class have under a 500. However, this model is comparable to an updated ACEER that includes high school GPA [12]. The resulting CEER score distributions are shown below.

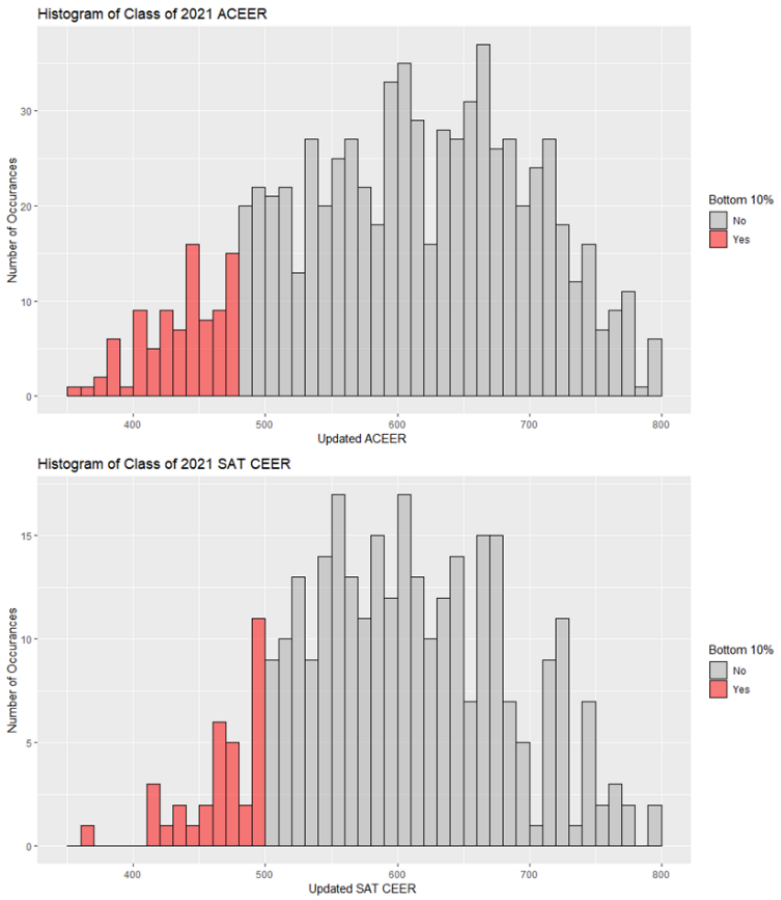

FIGURE 24. Comparison of ACEER and SAT CEER with HS GPA in the model

The new ACEER distribution 10% threshold is 480. There is slight difference between the distributions of the ACEER and SAT CEER. The ACEER has a greater variation and the SAT CEER is narrower. More research should be done to compare the two distributions.

D. BEST PREDICTOR

The final model is the best predictor of academic success. The model is built with the sole purpose of creating a model with most accurate predicted CQPA, regardless of usability and interpretability. The variables are SAT Math, SAT Reading, high school class rank score, and faculty appraisal score. The faculty appraisal score comes from a survey high school teacher fill out, measuring classroom participation and desire to learn. It is currently used in the community leadership score. It would require the Admissions team to alter the WCS scoring and so it is not preferable. The following equations are listed below.

$$pred\ CQPA^{1.667} = 0.0000001 * (SAT\ Math^2) + 0.00002\ * (SAT\ Reading^2) + 0.00004 * HSR + 0.00021 * FAS - 1.763$$

$$CEER = 0.0000044 * (SAT\ Math^2) + 0.00073 * (SAT\ Reading^2) + 0.191 * HSR + 0.918 * FAS - 472.9$$

This model looks different from the other models. Starting with predicted CQPA raised to 1.667. This is to help with residual skewing. The 1.667 came from a R program that optimized residuals when the output raised to a λ. It returned the optimal λ. The SAT Math and SAT Reading are squared to also help with the residuals. As expected, the model has the best r^2 at 0.54. The accuracy is 62% which is also higher than the other models. The spread is like the others as well.

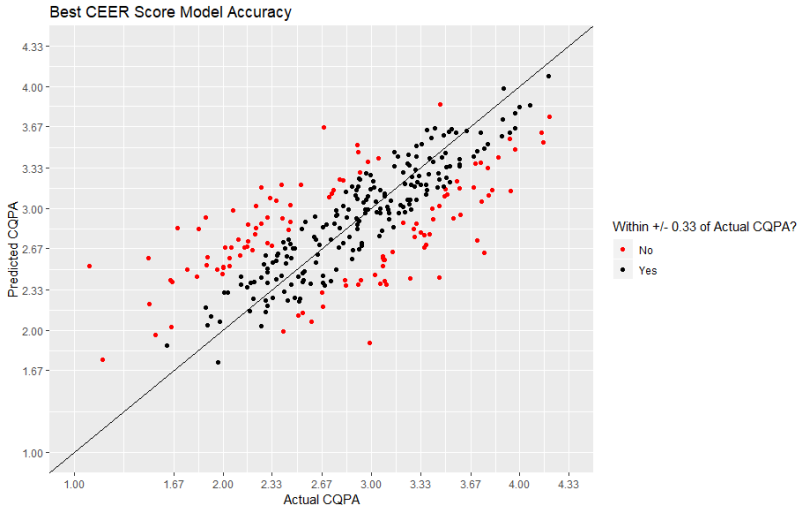

FIGURE 25. Best Predictor Model's Prediction Accuracy for the Class of 2021

This model has the best QQ Plot because the points lie along the line and do not deviate at the edges. See Figure 26.

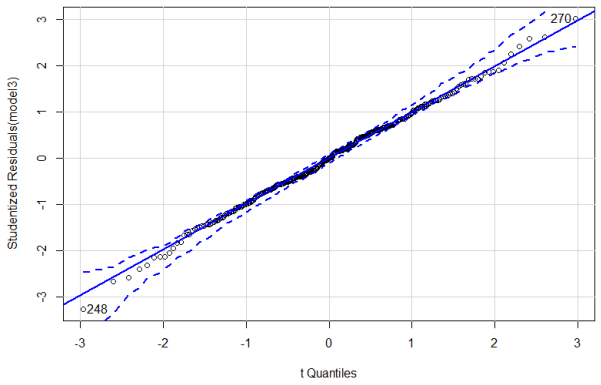

FIGURE 26. QQ Plot of Best Predictor Model's Prediction of the Class of 2021

The variables also have acceptable residuals. There is very little curving as shown in Figure 27.

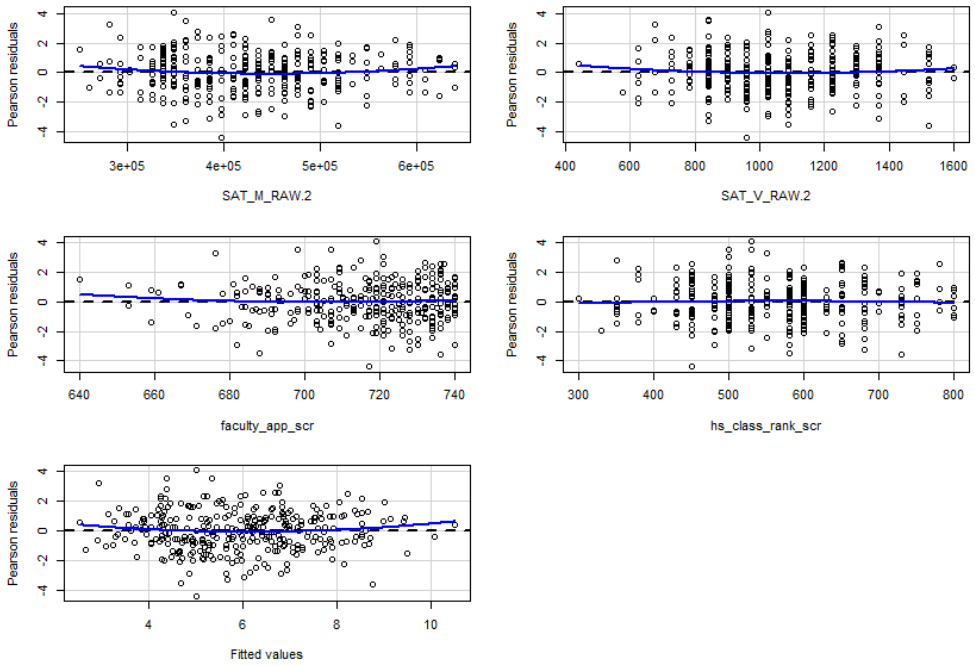

FIGURE 27. Residual Plot of Best Predictor Model's Prediction of the Class of 2021

The variable contributions are interesting because the faculty appraisal score is by far the highest contributor. There is something about the score that offers a significant amount of predictive power.

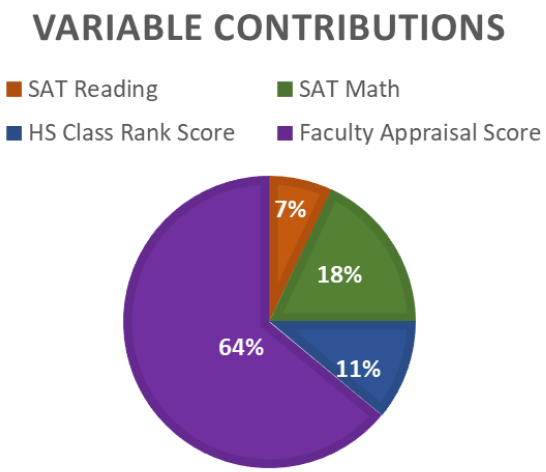

FIGURE 28. Best Predictor Model Variable Contributions

Finally, the new thresholds follow the same thresholds as the current goals. The bottom 10% of the class have under a 520 CEER score. The main challenge with this model is the interpretability and usability. It is difficult to consider a predicted CQPA raised to 1.667 and even more difficult to update the coding. It is no longer a simple update of the coefficients. This coding would require more steps to achieve a slightly better CQPA prediction and CEER score.

THIS PAGE INTENTIONALLY LEFT BLANK

V. CONCLUSION

A. RECOMMENDATION

Based on the results, the Best Predictor model has the highest accuracy and fit. However, it has a very low ease of implementation. The Basic Model has slightly lower accuracy and fit but it is highly implementable. The HS GPA has a poor fit, a slight decrease inaccuracy but it has the highest sustainability because of the HS GPA.

Model	Accuracy (within 0.33 of GPA)	Fit of Model (r^2)	Ease of Implementation	Sustainability
Current '95	46%	0.47	------	Low
Basic (revised)	58%	0.49	High	Medium
HS GPA	56%	0.43	Medium	High
Best Predictor	62%	0.54	Low	Medium

FIGURE 29. Model Summary Table

Therefore, we recommend the Admissions team adopts the Basic model because it has the highest ease of implementation without sacrificing too much accuracy or fit.

B. FUTURE WORK

Because of the high school GPA's potential, we recommend working on other high school GPA standardizations. The Air Force Academy accounts for the differences in GPA scales by recalculating each candidate's GPA. This method requires significantly more work than a simple ratio. However, it does solve the challenge of different scales, weighted versus unweighted, and class difficulty. There may also be a way to account for the high school's location by using zip code and the percentage of students who attend college. We recommend looking for similarities between students with similar scales. There may be some relationship between the type of high school, high school GPA scale, and location. Finally, the Admission team could consider asking for an unweighted STEM GPA or other variations of GPAs. We also recommend analyzing and adjusting for the differences between the ACEER and SAT CEER scoring distributions so that candidates who do not

have access to both standardized tests are not at a disadvantage. Another area for more research is identifying the at-risk candidates. Because the models tend to underpredict CQPA, is there something that connects “at-risk” candidates?

VI. PREDICTING EN101 SUCCESS WITH SAT ESSAY SCORES

A. INTRODUCTION

The West Point English Department uses the SAT Essay scores as an identifier of candidates at risk for failing EN101, the first English class plebes take at West Point. Currently SAT Essay scores under a certain threshold are read by the English Department faculty to ensure writing ability. The English Department is searching for a better understanding of the predictive power of SAT scores towards EN101 success. They are especially interested in the correlation between the new SAT Essay sub-scores and success in EN101.

B. SAT ESSAY SCORING UPDATES

Among other key changes of the 2016 updated SAT was the adoption of an optional essay and new scoring for all sections. Instead of a mandatory essay, the SAT offers an optional essay at the end of the test. The essay is given three sub-scores. The SAT Essay sub-scores are reading, writing, and analysis. The purpose of each score is to capture the nuanced skills in the essay. The SAT Essay Reading score measures the student's ability to understand the passage accurately. The SAT Essay Writing score measures the student's ability to write correctly, using proper syntax and word choice. The SAT Essay Analysis score measures the student's ability to make a logical, supported argument based on the evidence in the passage [14].

C. METHODOLOGY

1. Data

The data used in this analysis is from the Class of 2021. The EN101 grades are from their plebe year and include both semesters. The SAT scores from the Classes of 2012-2018 are also used to compare predictive powers to see if converting the scores is effective. It is important to note that the passing rate for both sets of data is very high. 96% of the cadet who took the pre-2016 SAT passed EN101 and 97% of the cadets who took the post 2016 SAT passed EN101. This is significant because it makes it more challenging to find the factors that account for the variance.

Figure 30 provides a clear representation of the relationship between SAT scores and EN101 grades. The data is from the Classes of 2012-2018 who took the pre-2016 SAT.

The data points illustrate the SAT Essay Score (out of 12) compared whether they passed EN101. The color corresponds to the grades achieved in EN101. This pattern is like the post 2016 SAT and EN101 relationships. This relationship is consistent with the other SAT Essay scores.

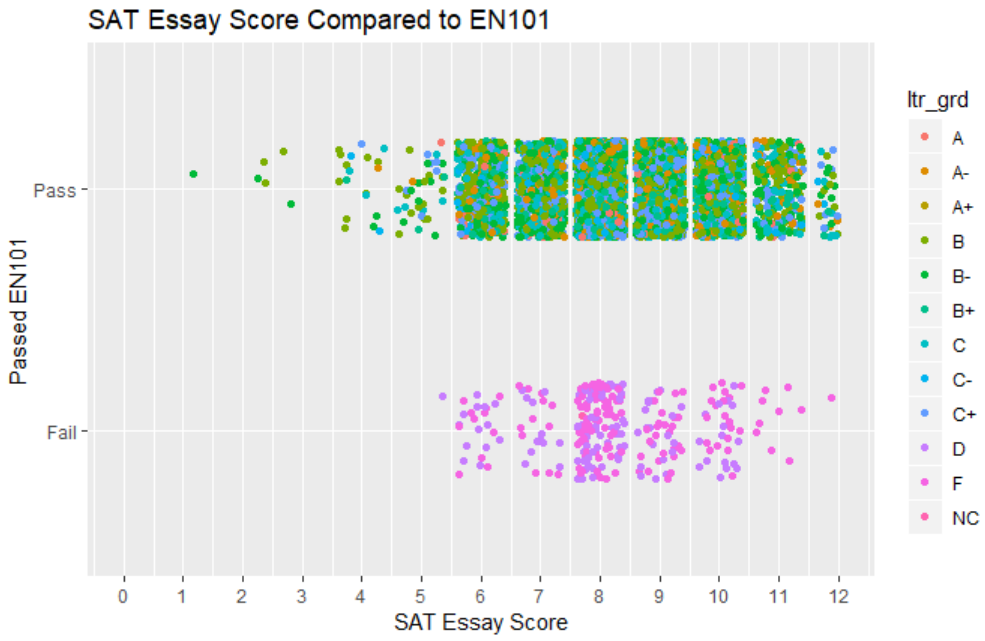

FIGURE 30. Pre-2016 SAT Essay Score Compared to EN101 Grade

It would be expected that the cadets who failed would have lower SAT Essay scores and we would also expect the colors to appear somewhat as a defined rainbow in blocks of colors and minimal overlap. However, that is not the case. Figure 30 shows that it is difficult to determine a relationship between SAT scores and EN101 grades. This challenge is evident in the following more in depth analysis.

2. Procedure

The analysis consists of two parts. The first part is a logistic regression and the second is a multiple linear regression. The logistic regression model predicts a candidate's probability of passing EN101 based on his or her SAT scores. The multiple linear regression model predicts a candidate's grade in EN101 based on his or her SAT scores. Both models consider pre and post 2016 SAT scores.

A logistic regression model gives the probability of a binary event. In this case the binary event was passing or failing EN101. The purpose of this logistic regression model

is to help predict a new "at-risk" score based on the probability of passing. To test the accuracy of the model, we set a 95% threshold. This 95% threshold means that if the probability of passing was greater than 95%, the candidate passed, else the candidate would be predicted to fail.

The multiple linear regression models predict a candidate's EN101 grade, using the SAT Writing and SAT Essay scores. The purpose of the multiple linear regression model is to determine if there is relationship between the EN101 grades and SAT scores.

D. MODELS AND RESULTS

1. Logistic Regression

a. Pre-2016 SAT Logistic Model

We used the SAT Writing and SAT Essay scores to build two models, predicting EN101 success. The SAT Writing score was significant, with a 96% accuracy which is the same as the EN101 passing rate. While the accuracy is high, the SAT Writing logistic model does not offer a better prediction of passing EN101 than a naïve guess.

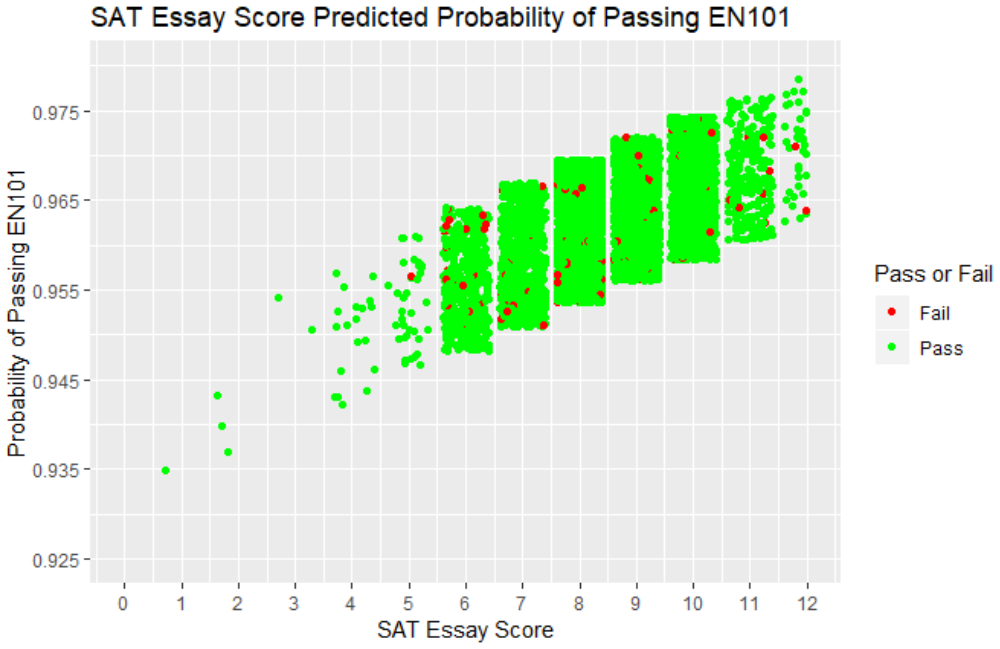

FIGURE 31. SAT Essay Scores Predicted Probability of Passing EN101

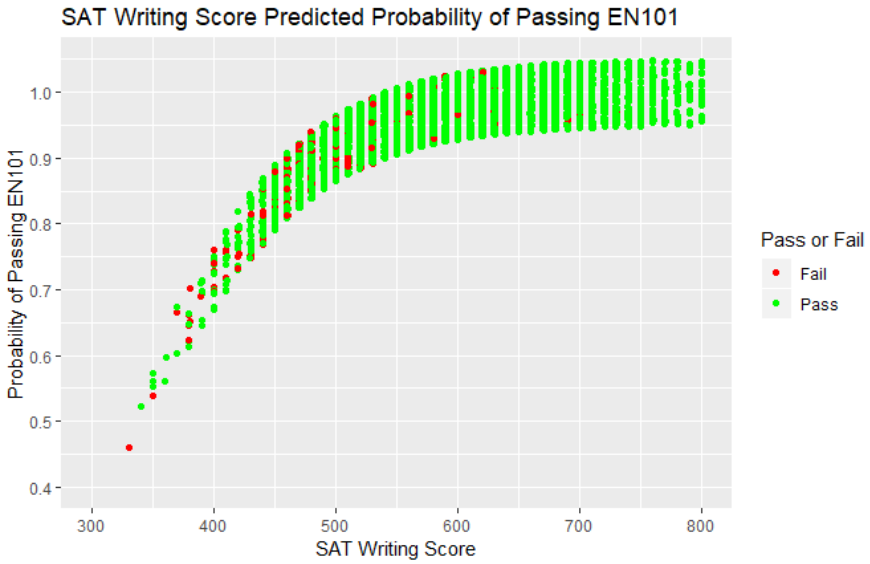

FIGURE 32. SAT Writing Score Predicted Probability of Passing EN101

The SAT Essay data was not significant, with 80% accuracy. This accuracy is significantly lower than the naïve model. This means that the English Department would be more accurate using historical data, than the SAT Essay logistic model. The SAT Essay does not offer new information and the English Department should be skeptical of the candidate's essay scores. It is not surprising that the SAT Writing scores were better at predicting success because the scoring is objective. The SAT Writing section is multiple choice while the SAT Essay score is more subjective due to human graders.

b. Post 2016 SAT Logistic Model

We used each SAT Essay sub-scores (Reading, Writing, and Analysis) to predict EN101 success. We used the same 95% threshold to determine accuracy. The SAT Essay sub-scores are not significant, but the models are between 97.5- 98% accurate. This greater than the naïve model which states that 97% of the candidates will pass. However, these models predict that everyone will pass, and the range of probabilities is very minuscule. Note Figures 33 thru 35. Therefore, the English Department should not put much weight behind the SAT Essay sub-scores.

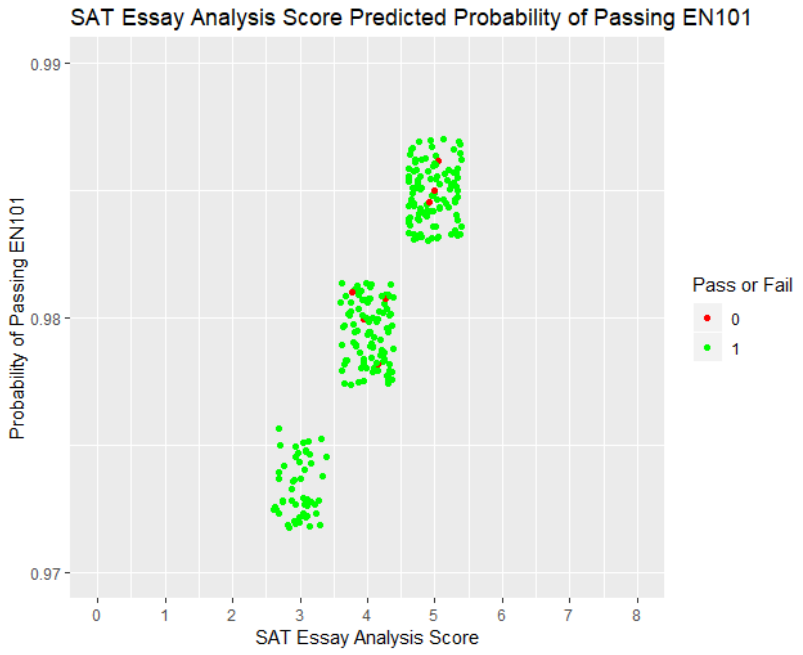

FIGURE 33. SAT Essay Analysis

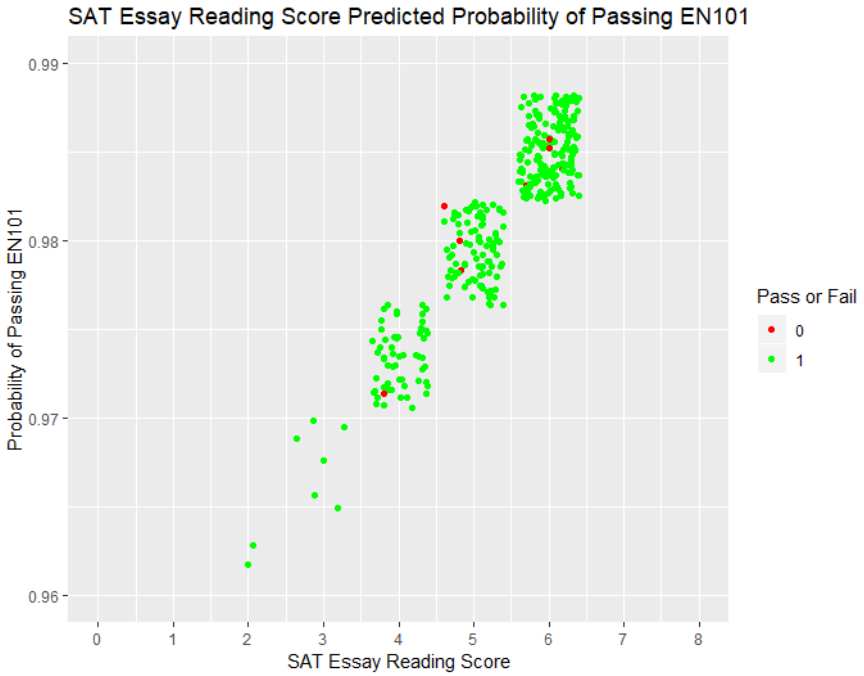

FIGURE 34. SAT Essay Reading

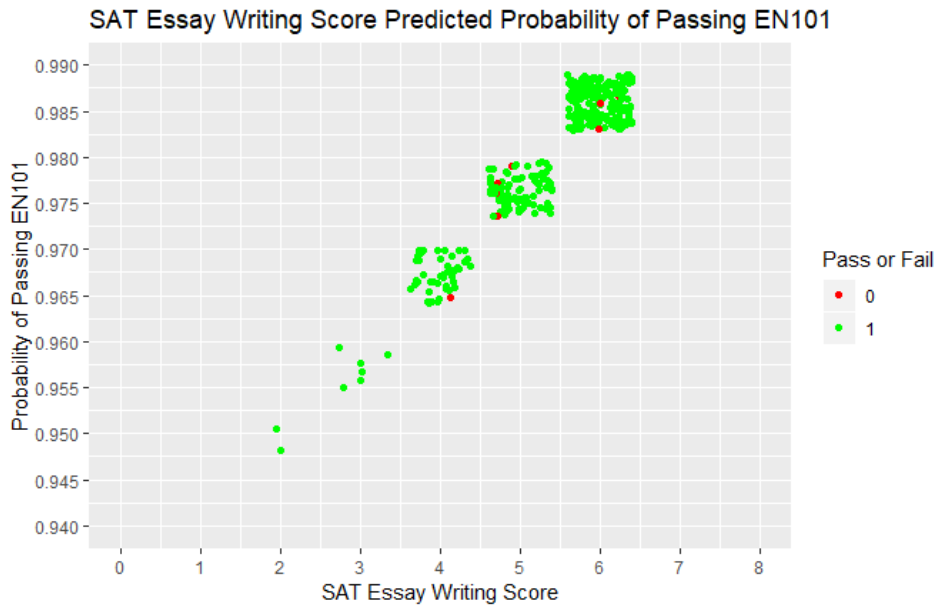

FIGURE 35. SAT Essay Writing

2. Multiple Linear Regression

a. Pre-2016 SAT Multiple Linear Regression Model

This model uses the SAT Writing, SAT Essay scores, and their interaction to create the model.

$$Grade\ (4.0\ scale) = 0.00474 * SAT\ Writing^{*} + 0.00413 * SAT\ Essay - 0.0000475 * SAT\ Writing * SAT\ Essay - 0.254$$

Asterisks denote significant components. The SAT Writing score was significant which is expected from the logistic regression model. The adjusted r^2 is 0.19, meaning the model accounts for 19% of the variance of EN101 grades. This is not a strong model for predicting EN101 grades, as shown in Figure 36 of actual compared to predicted grades.

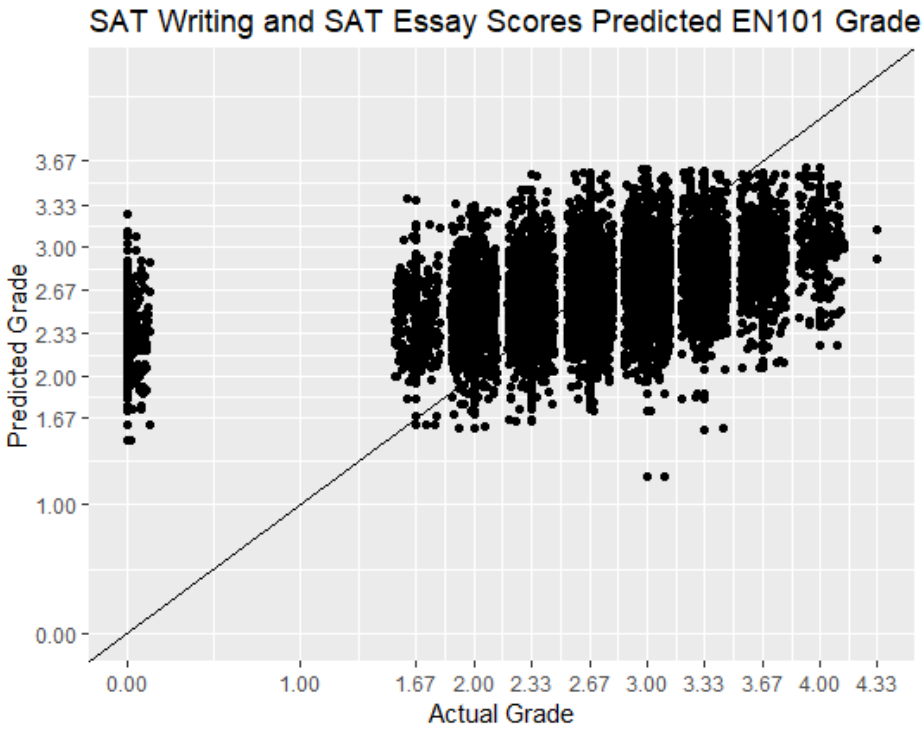

FIGURE 36. SAT Writing and SAT Essay Scores

b. Post 2016 SAT Multiple Linear Regression Model

This model uses the SAT Essay sub-scores and their interaction.

$$Grade\ (4.0\ scale) = -0.356 * SAT\ Essay\ Analysis\ - 0.352 * SAT\ Essay\ Reading - 0.462 * SATEssay\ Writing + \beta_4 * SAT\ Essay(s) + 4.84^{*}$$

Like the logistic regression models, none of the SAT Essay sub-scores were significant and the adjusted r^2 was negative. Note that the coefficients of SAT Essay sub-score are negative which means that as the essay sub-score increases, the predicted grade decreases. This is not logical. Observe Figure 37. The predicted grades range between a B- and B. There is little variation which does not reflect reality. Therefore, there is not a significant relationship between the SAT Essay sub-scores and the EN101 grades.

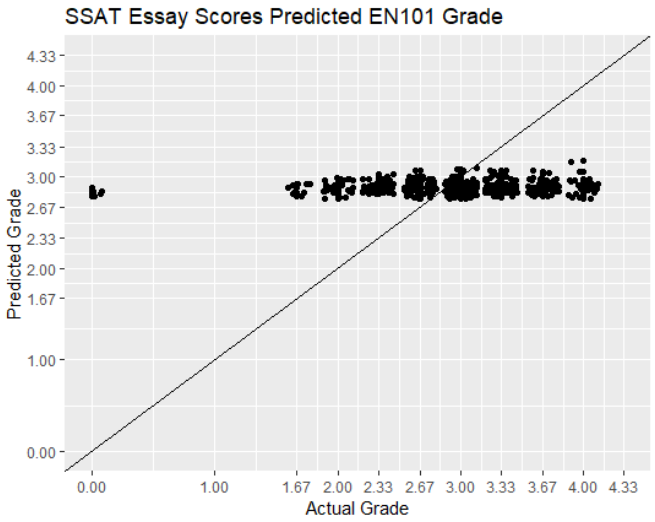

FIGURE 37. SAT Essay Sub-scores

E. CONCLUSION

Based on the analysis, the SAT scores do not provide a better predictor of EN101 success. The English Department should continue to read "at risk" essays because there is currently a high pass rate. That means that the candidates admitted tend to do well. This conclusion is supported by the Class of 2021 failures. According to the EN101 Course Director, 11/18 of the cadets who failed EN101, failed for a reason other than writing skill. Most of the failures were honor related [20].

LIST OF REFERENCES

[1] SAT Concordance Tables. (2016, May 9). PDF/EXCEL. April 2019.

[2] Arney, K. LTC. Personal conversation and email. West Point, NY. August 2018- May 2019.

[3] Whole Candidate Score Picture. New Admissions Committee Training 16-17. Power Point. January 2019.

[4] A Brief History of the SAT. (n.d.). Retrieved from https://www.pbs.org/wgbh/pages/frontline/shows/sats/where/history.html

[5] Changes to the SAT in 2016. (2018, August 29). Retrieved from https://www.manhattanreview.com/new-sat-2016/

[6] (CEER12) Dean's Time Study AY13-1 Report, Department of Mathematical Sciences. (2013, January 16). [Memorandum]. West Point, NY: United States Military Academy

[7] Geiser, S., (2015). The Growing Correlation between Race and SAT Scores: New Findings from California. *University of California, Berkley Research & Occasional Paper Series*: Center for Studies in Higher Education (CSHE). 10.15.

[8] Gipson, J. (2016). Predicting academic success for students of color within STEM majors. *Journal for Multicultural Education, 10*(2), 124-134. doi:10.1108/jme-12-2015-0044

[9] Lindquist, J. COL. Personal conversation and email. West Point, NY. August 2018- May 2019.

[10] McDonald, Deborah COL. Personal Interview. West Point, NY. 13 September 2018.

[11] McCusker, C CDT. Collaboration. West Point, NY. August 2018- May 2019.

[12] Palmer, J CDT. Collaboration. West Point, NY. August 2018- May 2019.

[13] Recalculations of CEER/HPA/MSE Predictors Using Recentered SAT I Scores. (1995, March 22). [Memorandum]. West Point, NY: United States Military Academy

[14] SAT Essay Scoring. (2018, March 12). Retrieved from https://collegereadiness.collegeboard.org/sat/scores/understanding-scores/essay

[15] SAT Score Report. (n.d.) Retrieved from https://www.kwellerprep.com/sat-test-prep/. December 2018.

[16] Understanding SAT Scores 2018. (2018). PDF.

[17] Walker, K., (2010). Rank in Class and College Admission. Education Partnerships, Inc. and Lebanon Valley College. [Research Paper].

[18] Young, S., (2016). Do Colleges Favor Students from Lower-Achieving High Schools? Evidence from the University of California. San Diego, CA: University of California Department of Economics. Research Paper.

[19] Zahner, D., Ramsaran, L., & Steedle, J. (2014) Comparing Alternatives in the Prediction of College Success. Lexington, NY: Council for Aid to Education.

[20] Zupancic, Anthony LTC. Personal Interview. West Point, NY. 27 November 2018.