UNITED STATES
MILITARY ACADEMY

WEST POINT, NEW YORK

HONORS THESIS

A COMPREHENSIVE STUDY OF MA104: INFORMING COURSE DESIGN DECISIONS WITH PERFORMANCE, ATTITUDE, AND SURVEY DATA

by

Cadet Stephanie McDermott

May 2018

Primary Thesis Advisor:	LTC Michael Scioletti
Advisory Team:	CPT Andrew Plucker
	CPT Dusty Turner

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
	May 2018	Senior Thesis (Honors)

4. TITLE AND SUBTITLE A Comprehensive Study of MA104: Informing Course Design Decisions with Performance, Attitude, and Survey Data

5. FUNDING NUMBERS

6. AUTHOR(S) Cadet Stephanie McDermott

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
Distribution Statement (mix case letters)

12b. DISTRIBUTION CODE

13. ABSTRACT (maximum 200 words)

Each year, over 900 students at the United States Military Academy enroll in MA104: Single Variable Calculus. Students complete pre- and post-course surveys designed to capture their feedback on quality of instruction, course design, learning, and overall satisfaction. Data from these surveys comes in varied forms: continuous, integer or Likert Scale, categorical, and free-text responses. Data is then paired with a student's performance data, which includes SAT/ACT scores, MA103 grades, and other similar pre-MA104 scores. In this study, using survey and performance data from the graduating class of 2020, we employ a variety of statistical techniques to identify significant relationships between cadet feedback and performance. Not only do we apply linear regression to identify the significant predictors for MA104 grades, we also employ logistic regression to predict cadet's attitudes towards mathematics. Results highlight the most significant indicator of success in MA104 is a cadet's MA103 grade and performance in MA104 is positively correlated with a favorable attitude towards mathematics. These findings are used to better understand the students in order to make well-informed course design decisions.

14. SUBJECT TERMS
Model Selection, Linear Regression, Logistic Regression, Predictors, Survey

15. NUMBER OF PAGES 39

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

THIS PAGE INTENTIONALLY LEFT BLANK

A COMPREHENSIVE STUDY OF MA104: INFORMING COURSE DESIGN DECISIONS WITH PERFORMANCE, ATTITUDE, AND SURVEY DATA

Stephanie R. McDermott
Cadet, Engineers
B.S., United States Military Academy, 2018

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 2018

Author: Cadet Stephanie McDermott

Approved by: LTC Michael Scioletti
Thesis Advisor

CPT Andrew Plucker
Co-Advisor

CPT Dusty Turner
Co-Advisor

Colonel Steven B Horton
Head, Department of Mathematical Sciences

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

A Comprehensive Study of MA104: Informing Course Design Decisions with Performance, Attitude, and Survey Data

Cadet Stephanie McDermott
Department of Mathematical Sciences
United States Military Academy
West Point, New York

Advisor: LTC Michael Scioletti
Professor
Department of Mathematical Sciences
United States Military Academy
West Point, New York

ABSTRACT:

Each year, over 900 students at the United States Military Academy enroll in MA104: Single Variable Calculus. Students complete pre- and post-course surveys designed to capture their feedback on quality of instruction, course design, learning, and overall satisfaction. Data from these surveys comes in varied forms: continuous, integer or Likert Scale, categorical, and free-text responses. Data is then paired with a student's performance data, which includes SAT/ACT scores, MA103 grades, and other similar pre-MA104 scores. In this study, using survey and performance data from the graduating class of 2020, we employ a variety of statistical techniques to identify significant relationships between cadet feedback and performance. Not only do we apply linear regression to identify the significant predictors for MA104 grades, we also employ logistic regression to predict cadet's attitudes towards mathematics. Results highlight the most significant indicator of success in MA104 is a cadet's MA103 grade and performance in MA104 is positively correlated with a favorable attitude towards mathematics. These findings are used to better understand the students in order to make well-informed course design decisions.

KEYWORDS: Model Selection, Linear Regression, Logistic Regression, Predictors, Survey

CONTACT: CDT Stephanie McDermott, Department of Mathematical Sciences, USMA, West Point, NY, 10997 Tel: 845.608.2139; E-mail: Stephanie.McDermott@usma.edu

LTC Michael Scioletti, Department of Mathematical Sciences, USMA, West Point, NY, 10997 Tel: 845.938.0061; E-mail: Michael.Scioletti@usma.edu

CPT Andrew Plucker, Department of Mathematical Sciences, USMA, West Point, NY, 10997 Tel: 845.938.2654; E-mail: Andrew.Plucker@usma.edu

CPT Dusty Turner, Department of Mathematical Sciences, USMA, West Point, NY, 10997 Tel: 845.938.6419; E-mail: Dusty.Turner@usma.edu

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES AND FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to thank my thesis advisor LTC Michael Scioletti for everything he has done for me these past two semesters. I was excited for the opportunity to work with an engaging, challenging, and positive advisor – LTC Scioletti did not disappoint. From the beginning stages of our research, to include the literature review and reviewing the fundamentals of coding, to the final paper, LTC Scioletti has guided me patiently through it all. He always believed in my ability to succeed, and provided instrumental mentorship during my final year at the Academy.

I would also like to thank my two secondary advisors – CPT Andrew Plucker and CPT Dusty Turner. Without CPT Plucker, I would not have had access to the data which made this research possible. He also provided insight and guidance during the initial understanding and analysis of the data. CPT Turner was the coding expert for this project. His passion and enthusiasm taught me not only about the depth and power of R, but also about expectations for a young Engineer officer.

THIS PAGE INTENTIONALLY LEFT BLANK

1. INTRODUCTION

Calculus is the primary gateway for most students heading into the technical and scientific fields that will drive the economy of the 21st century. It is celebrated as one of the greatest intellectual achievements of the western civilization – yet this feeling is often lost on its students. Why is this? Is it reversible?

The leading hypothesis infers Calculus has an accessibility issue. To study this, the Mathematical Association of America (MAA) conceived a national survey titled Characteristics of Successful Programs of College Calculus (CSPCC), which has five goals (Bressoud et al., 2013):

(1) To improve our understanding of the demographics of students who enroll in Calculus

(2) To measure the impact of the various characteristics of Calculus classes that are believed to influence student success

(3) To analyze course feedback in order to determine which programs are successful, and why

(4) To develop a model that identifies the most statistically significant predictor variables for student success in Calculus

(5) To use the results of these studies and the influence of the MAA to leverage improvements in Calculus instruction across the United States.

To meet these goals, the MAA administered a series of surveys to over 10,000 students to observe significant relationships between student performance and attitude/survey data. Surveys were conducted both prior to and following the Calculus course. The pre-course surveys gathered demographic information about the students, to include race and gender, as well as secondary school background, such as standardized testing math scores and mathematics courses taken prior to university. Information gathered prior to the course was selected to identify the factors influencing student persistence and achievement not only in college, but in science, technology, engineering, and mathematics (STEM). The post-course surveys collected information regarding student feedback on the course, along with their final grades and whether there is a desire to continue studying mathematics.

The pre-Calculus surveys found most students were motivated and confident in their ability to do well in Calculus, regardless of their background or previous mathematics experience. Students who reported an interest in pursuing a STEM discipline also reported they would pursue further study in mathematics. However, this changed on the post-Calculus surveys, which showed a dramatic decrease in student confidence, enjoyment, and the desire to continue in mathematics.

These results validated the initial hypotheses of the MAA, i.e. students are highly motivated and consider themselves well-prepared prior to Calculus, but lose momentum somewhere in the execution of the course. To address this issue, the MAA continues to explore survey data in partnership with other research teams across academia, while also inspiring related, smaller-scale projects like the one in this study. Before describing further the application to West Point, we must first explain why surveys are used and how they are analyzed.

2. SURVEY ADMINISTRATION

Surveys are a powerful feedback tool that academic stakeholders employ to measure satisfaction, quality of instruction, and general student attitudes. Traditionally, student surveys are composed of both open-ended questions and Likert-scale based questions, which are a common ratings format for surveys that was developed in 1932 to measure attitudes on a five to seven point ordinal scale (Allen & Seaman, 2007). An ordinal scale ranks observations in some measure of magnitude. For example, on a five point ordinal scale, a 1 means Strongly Disagree, 2-Disagree, 3-Neutral, 4-Agree, and 5-Strongly Agree. These values express a ‘greater than’ relationship, i.e. Strongly Agree is a stronger opinion than Agree; however, how much greater is not known (Boone & Boone, 2012). As a result, Likert scales are often truncated to an even number of categories to eliminate the neutral option in a forced choice survey scale (Allen & Seaman, 2007). For example, students participating in the MAA study responded to Likert-scale based questions and statements on a scale of 1 to 6, where a 1 meant Strongly Disagree, 2-Disagree, 3-Somewhat Disagree, 4-Somewhat Agree, 5-Agree, and 6-Strongly Agree. The MAA posed questions such as “Do your friends see you as a person who is good at mathematics?” and statements such as “This course has increased my interest in taking more math courses.”

Likert scale-based questions are often used in academic survey research to gather information relative to general attitudes, emotions, personalities, and descriptions of peoples’ environment (Gliem & Gliem, 2003). They eliminate the difficulties associated with measuring attitudes, character, and personality traits by transferring these qualities into a numerical measure (Boone & Boone, 2012). While the use of Likert response questions is common in surveys, the methods for analyzing and interpreting results vary. Research from 2003 states “many individuals invalidate research findings due to improper data analysis,” and over ten years later, the controversy regarding the *best* way to analyze Likert-type data still exists (Gliem & Gliem, 2003, and Sullivan & Artino, 2013). Likert scales and questions are practical and effective in many fields of survey research, but interpreting and analyzing the data involves a high degree of complexity that could ultimately result in costly conclusions.

An alternative to Likert-scale based questions is to ask open-ended questions and allow survey respondents to provide free-text responses. Open-ended questions from the MAA study include

“What did you like the most about this course?” and “For me, making unsuccessful attempts when trying to solve a problem is…” The students’ answers to these questions help academic stakeholders improve the course because they are honest and capture general student opinions which may not be otherwise voiced. As meaningful as these answers might be, the results are typically cumbersome, and difficult to objectively analyze (Hood & Kuiper). While these results are also prone to response biases, such as sampling bias (drawing conclusions from an inaccurate sample of the population under study) or acquiescence bias (tendency for survey respondents to agree positively to a question in lieu of a deliberate answer), they are a tool that measures student feedback. To capture this same feedback, minus the response biases, Hood and Kuiper created Student Directed Discussion Surveys, or SDDS.

The SDDS provides one to three broad and undirected questions, allowing the student to lead the discussion. For example, instead of “What did you like most (or least) about this course?” the SDDS asks the student, “Please discuss your thoughts on this calculus course.” The resultant data is analyzed using the sentiment analysis methodology known as Natural Language Processing (NLP) – implementable in two Python libraries: Natural Language Toolkit (NLTK) and VADER sentiment analysis (VADER). Analyzing the results of three separate surveys (taken at three periods throughout the semester) indicate SDDS produces comparable results to similar traditional student survey questions, but without the response biases and the artificial variance of negative or polar responses. Hood and Kuiper contend SDDS could be a useful supplement to traditional surveys because the results produce a more information-rich data set, providing stakeholders with more useful information.

3. SURVEY ANALYTICS

Surveys can generate vast amounts of data from which researchers are expected to extract important patterns and trends. A popular technique for mining survey data is with supervised learning – where the goal is to predict the value of an outcome measure based on a number of input measures (Hastie, 2009, p. xi). The most common of these learning techniques is linear regression, which results in a function of the form$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$. These functions are simple, yet they provide an adequate and interpretable description of how the inputs affect the output (Hastie, 2009, p. 43).

Linear regression is well suited for analyzing survey data because the independent variables can be both categorical and numeric. For prediction purposes this technique can sometimes outperform more complex nonlinear models, especially in situations with small testing samples or sparse data (Hastie, 2009, p.43).

Since its inception in 2009, the open-source integrated development environment software R has become a common tool for statisticians. Linear regression models in R apply the least squares method, which assigns coefficients to the independent variables in order to predict the value of the dependent variable, while minimizing the residual error (Hastie, 2009, p. 44). R finds the most statistically significant predictors as a function of the model.

Similar to linear regression, logistic regression models use the least squares method. However, the output of logistic regression is binary. Also, logistic regression models require a manual split of the data into a training and testing set. The training set, normally set as 85% of the data, is used to fit the model in R. The remaining 15% of the data, the testing set, is then run through the model to test its predictive ability. By using the predictive model (built with the training set) on other similar data (the testing set), statisticians can effectively test the accuracy of their model, and refine their model further, if necessary.

A logistic regression model is best evaluated with a receiver operating characteristic (ROC) curve that plots the true positive rate against the false positive rate at various decision thresholds. Thus, the predictive score of a logistic regression model is the area under the ROC curve (AUC). The AUC can be interpreted in the same manner as adjusted R-squared in linear regression, i.e. an AUC of 1.0 is a perfect fit because it means there is a 100% discrimination rate, the model always predicts correctly the binary outcome.

Linear and logistic regression are useful tools to handle quantitative data collected by surveys. Qualitative data, on the other hand, is difficult to analyze mathematically. When considering open-ended textual responses, less common statistical tools like sentiment analysis must be used. In addition to the Natural Language Processing mentioned earlier, one way to measure the opinion or emotion in a student answer is to analyze the sum of the sentiment content of the individual words. The first step in this procedure is to format the data so that each response is separated into several rows so that each word is its own row. Second, stop words are extracted, or words with presumably zero sentiment or emotion, such as "is," "the," "are," etc. The remaining words are scored according to a lexicon. In R, these lexicons are; "AFINN," which scores words on a scale of -5 to 5, where negative scores mean negative sentiment, and vice versa; "bing," which scores words on a binary scale, positive or negative; and "nrc," which places words into categories, to include positive, negative, anger, anticipation, disgust, fear, joy, sadness, surprise, and trust.

Responses are then scored by summing each individual word sentiment. A less mathematically rigorous, but insightful, technique to exploring large amounts of textual data is with a Word Cloud

– visual representations of selected pieces of text, where the most frequently occurring words are the largest in the graphic.

4. APPLICATION TO USMA

Like most liberal arts colleges, the United States Military Academy (USMA) at West Point requires students to complete a core curriculum, comprised of mathematics, basic sciences, engineering sciences, information technology, humanities, behavioral sciences and social sciences. Three of these courses satisfy the mathematical sciences requirement. The first mathematics course in this sequence is MA103: Mathematical Modeling and Introduction to Calculus, which is followed by MA104: Single Variable Calculus (termed Calculus I at most colleges). The third course in the sequence is MA206: Probability and Statistics. The focus of our efforts are on MA104.

Upon completion of MA104, as they would for any course at the Academy, cadets take an end-of-course survey. These surveys are comprised of several Likert-scaled questions and open ended free-text questions, designed to capture feedback on the course, instructor, and overall experience. For example, cadets respond Strongly Disagree, Disagree, Neutral, Agree, or Strongly Agree to statements such as “This instructor used effective techniques for learning, both in class and for out-of-class assignments.” and “The homework assignments, papers, and projects in this course could be completed within the USMA time guideline of two hours preparation for each class attendance.” They are also asked open-ended questions such as “What suggestions would you like to provide to the Department of Mathematical Sciences that would help enhance the learning experience in this course?”

The course end feedback captured in these surveys is then used to better shape the course for the next academic term. However, despite adjustments made each year, results seem to mimic the national trend identified by the MAA. That is, cadets are highly motivated and consider themselves well-prepared prior to MA104, but the momentum fades somewhere in the execution of the course. In an effort to address this stigma, surveys similar to those administered in the CSPCC (see Appendix A) were given to cadets in the Class of 2020 to solicit feedback on:

(1) Information which can be used to predict a cadet's success in MA104
(2) Relationships between a cadet's performance and their sentiment towards mathematics
(3) Sentiment towards electronic classroom resources in comparison to the traditional textbook

with the goal to use student feedback to adjust how course concepts are defined, implemented, and assessed in order to foster a favorable commitment to learning that extends beyond the classroom.

There is limited research on assessment of West Point cadets; however, two Senior Theses by cadets Betzel and Lindsay involve similar data analysis as ours. Specifically, Betzel analyzed ten years of USMA's admission data, applying linear regression to determine which criteria from a cadet's initial entry data correlated to their final academic GPA. His original model included over forty predictor variables, but using step-wise linear regression to remove insignificant variables one at a time, he found the following three to be the most statistically significant: SAT Math, ACT Math, and the Faculty Appraisal Score (a cumulative score given to a cadet candidate based on English, math, and lab science instructor evaluations).

Similarly, Lindsay used linear regression to predict cadet performance, which she defined as Cumulative Grade Point Average (CQPA), given cadet candidate entry scores. Furthermore, she observed the relationship between cadet candidate, cadet, and officer performance. Lindsay found the most statistically significant predictor of CQPA to be the Faculty Appraisal Score, but that cadet performance records (those measured while they are enrolled in the Academy) prove far better predictors of CQPA than do cadet candidate records. That is, the linear regression models built to predict CQPA as a function of other cadet scores had a much higher adjusted R-squared (0.910) than did the models which were a function of cadet candidate scores (.336). Overall, Lindsay's results highlight one key trend: the further removed the predictor variables are from the target parameter, the lower the predictive power of that model, e.g. cadet candidate records cannot accurately predict performance as an officer. The more time between datasets, the less of a correlation will exist between them. We apply this research to help in our first research objective, predicting MA104 grades. However, before applying any analysis to our data, we must first prepare and explore the data to focus our efforts.

5. METHODOLOGY

5.1. DATA PREPARATION

In 2017, various types of data were collected on cadets in MA104 through the surveys mentioned above, found in Appendix A. Then, additional data was paired with each cadet that included their MA103 grade and cadet candidate records, e.g. SAT Math score, the USMA College Entrance Examination and Class Rank (CEER) score, sex, and race.

The total dataset is 233 columns (hereafter referred to as fields) by 897 rows (hereafter referred to as cadets.) Any field where data was not captured for more than twenty percent of the cadets was removed from the dataset. A total of 132 fields were thus eliminated from use – nearly splitting the number of fields in half. Remaining data was treated with the following imputation techniques.

SAT Math and ACT Math scores were imputed using concordance tables (found in Appendix B), which use historical performance data from both exams to estimate a score on one exam, given the other. For example, if a cadet candidate scored a 750 on the SAT Math section, but did not take the ACT, their ACT math score is then estimated based on the concordance table, which correlates a 750 on the SAT Math section to a 33 on the ACT Math section. Other fields, such as a cadet's score on the gateway Calculus exam or on the first Fundamental Concepts Exam (FCE), were imputed using the K-Nearest Neighbor Algorithm (KNN). The purpose of KNN is to assign a best estimate value for a given field, based on the values of that field for the K number of most similar cadets. In this study, we set K equal to five. For example, let the FCE score of Cadet X be missing. Then, this approach considers the five cadets whose data is most similar with that of Cadet X, averages their FCE Score, and assigns that value to Cadet X.

After data imputation and scrubbing, we performed initial data exploration which began with statistical graphics created in R. First, a pairs graph (see Figure 1) was created to compare the relationships between data fields. Note, that in Figure 1 below, it appears the strongest, most positive linear relationship exists between MA103 Grade and MA104 Grade, as depicted in the top right and bottom left plots.

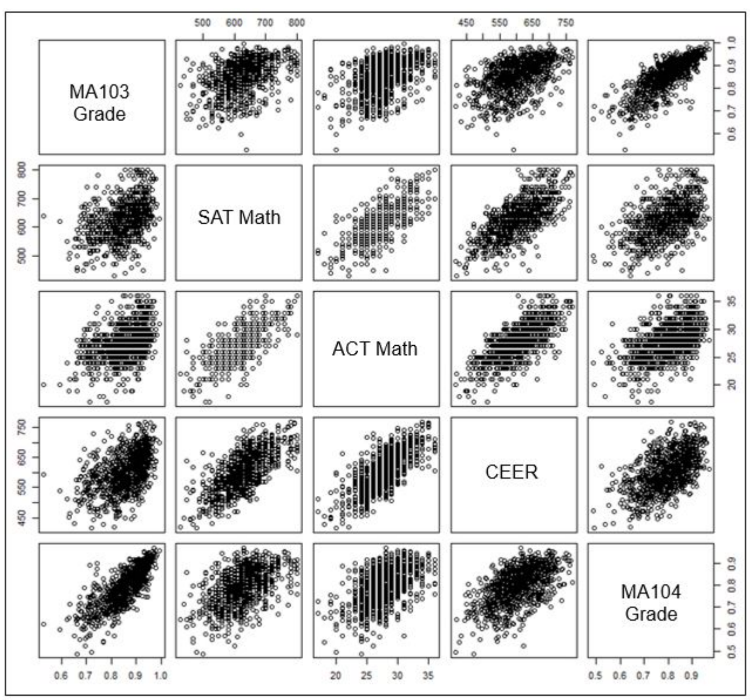


Figure 1: *A matrix of scatterplots which depict the relationships between data collected on the MA104 cadets from the Class of 2020.*

Second, we examined perception versus reality in order to determine if cadets had realistic expectations upon beginning the course. On the pre-course survey, students were asked what grade they expected to receive in MA104, which we compared to their actual grade. The highest numbers in this matrix do not lie all on the diagonal, indicating a disconnect between cadet expectation and reality. To mitigate this disconnect, instructors can more clearly outline their standards and expectations so MA104 cadets can meet their expected grades. Only seven cadets who did not expect an A actually received an A, and the largest number in the matrix corresponds to cadets who expected an A, but got a B. This indicates MA104 may be harder than the average cadet assumes, cadets are not merely handed an A, or cadets who expect a high grade may not be as good at math as they think or maybe they do not work hard enough to deserve an A.

		Actual Grade				
		A	B	C	D	F
Expected Grade	A	114	212	93	7	6
	B	5	94	208	26	19
	C	2	1	24	6	16
	D	0	0	0	0	1

Figure 2: *Expected Versus Actual Grade – On the pre-course survey, cadets were asked what grade they expected to receive in MA104, which was then compared to their actual grade.*

We also examined a cadet's enjoyment of mathematics, as they indicated on the pre-course survey, in comparison to their final grade in MA104. Cadets responded to the statement, "I enjoy doing mathematics," on a scale of 1 to 6, where a 1 is Strongly Disagree, 2-Disagree, 3-Somewhat Disagree, 4-Somewhat Agree, 5-Agree, and 6-Strongly Agree. As shown below, the cadets who reported enjoying math the most, responding with a 5 or 6, also had the highest averages in MA104. This is useful because instructors can predict which cadets will do well in MA104 by their positive sentiment towards math.

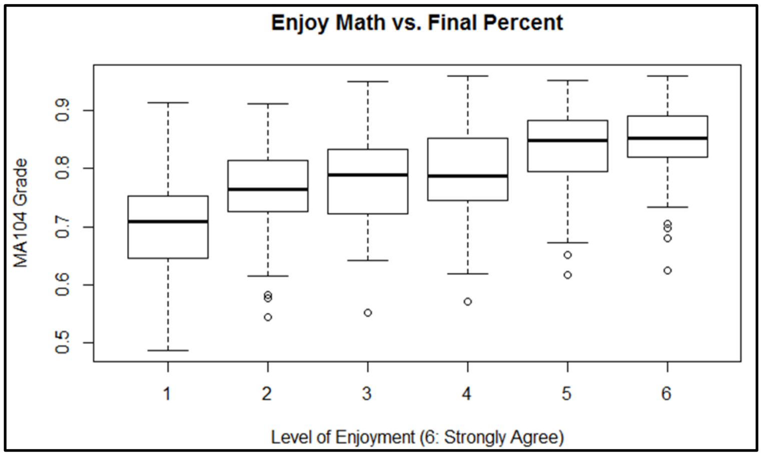


Figure 3: *Cadets' response to "I enjoy doing mathematics," where 1 is Strongly Disagree and 6 is Strongly Agree, plotted against their final MA104 Grade.*

Last, we examined the demographics of cadets, in particular the math courses they took prior to attending the Academy. Over ninety percent of cadets reported taking Pre-Calculus, while less than half took Advanced Placement (AP) Calculus AB, and not even ten percent took AP Calculus BC. AP Calculus BC is an extension of AP Calculus AB: the difference between them is scope, not level of difficulty. AP Calculus AB is equivalent to a semester of Calculus at most colleges and universities, while BC is equivalent to one year (AP Central). From Figure 4, we conclude the majority of MA104 cadets have limited, to no, experience with Calculus instruction.

Math Courses Cadets Took Prior to Attending West Point		
Course	*Cadets*	*Percent*
Pre-Calculus	770	93.7%
Calculus (Non AP)	189	23.0%
AP Calculus AB	344	41.8%
AP Calculus BC	68	8.3%
International Baccalaureate (IB) Calculus	20	2.4%
Online calculus course	8	1.0%
Calculus course taught at my high school for which I also received college credit	36	4.4%
Calculus course taught at a two-year college for which I also received high school credit	13	1.6%
Calculus course while I was enrolled in college	65	7.9%
Calculus course taught at the USMA Prep School	46	5.6%

Figure 4: *Math Courses cadets from the Class of 2020 took prior to attending West Point, as reported on the pre-course survey.*

5.2. LINEAR REGRESSION

Once the data was well prepared and explored, it is ready for analysis. First, we used linear regression to predict cadet success in MA104. When deciding which predictor variables to use, we considered many factors; scrubbing the data eliminated over half the fields from use, which meant we were working with a smaller data set; conclusions drawn from the above figures; and predictor variables used by Betzel and Lindsay. We thus chose 16 fields to include as predictor variables, outlined in Table 1 below. Then, we considered three separate linear models, with (1) Numeric predictor variables, (2) Categorical (to include binary) predictor variables, and (3) A combination of both. The response variable was a numerical value between 0 and 100 representing a cadet's final grade in MA104. In model 1, there were six predictor variables, X_1 throughX_6. Model 2 included the next ten predictor variables, X_7 throughX_{16}. Note we treated the Calculus Validation, AP Calculus AB, and AP Calculus BC exam scores as binary variables, yes or no, because a majority of the cadets in the dataset had not taken any of these three exams. Instead of using their scores, we used a 1 if the cadet took the exam, and 0 if not. Model 3 had all sixteen predictor variables, X_1 throughX_{16}.

Models 1 – 3 were compared against one another by their adjusted R-squared values, and statistically significant predictor variables were identified by low P-values. The closer the P-value is to 0, the more meaningful the addition of that predictor variable is to the model. (To emphasize variables of significance, any P-value in table 1 with eight or more leading zeros after the decimal was displayed as 0.) Model 3 has the highest adjusted R-squared value at 0.744, indicating nearly three-quarters of the variance in the data is accounted for with Model 3. (With MA103 grade being the most statistically significant predictor variable, removing it did not raise the significance of other variables in the model.)

Field Descriptions				Results		
Variable	**Class**	**Field Name**	**Description**	**Model 1**	**Model 2**	**Model 3**
Response Variable				*Adjusted R-Squared*		
Y	Numeric	MA104	Final grade in MA104	0.713	0.398	0.744
Predictor Variables				*P-Value*		
x_1	Numeric	MA103	Final grade in MA103	***0**		***0**
x_2	Numeric	SAT-M	SAT Math score	0.0718		0.761
x_3	Numeric	ACT-M	ACT Math score	0.991		0.411
x_4	Numeric	CEER	USMA Admissions CEER score	***0**		***0.00173**
x_5	Numeric	Gateway	Gateway Exam score	0.000282		***0.00116**
x_6	Numeric	FCE	Fundamental Concepts Exam 1 score	***0**		***0**
x_7	Categorical	Instructor	MA104 Instructor		0.455	0.899
x_8	Binary	Sex	Sex		0.349	0.811
x_9	Categorical	Race	Race		0.245	0.533
x_{10}	Binary	StandM	Scored in 90th percentile or above on either the SAT or ACT Math Section		***0**	0.132
x_{11}	Binary	CVE	Took the Calculus Validation Exam		***0**	0.0131
x_{12}	Binary	APAB	Took the AP Calculus AB Exam		***0.000247**	0.0299
x_{13}	Binary	APBC	Took the AP Calculus BC Exam		0.545	0.274
x_{14}	Binary	USMAPS	Attended the United States Military Academy Preparatory School		***0**	***0.00241**
x_{15}	Binary	STEM	Intent to major in a Science, Technology, Engineering, or Mathematics (STEM) Discipline		***0.0000121**	0.0264
x_{16}	Binary	RECATH	Recruited Athlete		***.000102**	0.209

Table 1: *Linear Regression Field Descriptions and Results**.** The most statistically significant variables (those with the lowest P-Values) per model are indicated by an asterisks.*

5.3. LOGISTIC REGRESSION

Next, we applied logistic regression to predict whether a cadet agrees with a pre-course survey question, "Do you see yourself as a person who is good at mathematics?" The predictor variables are the same numeric predictor variables as in the first linear model, X_1 throughX_6, and the MA104 grade. As previously discussed, the data was split into a training and testing set. There were 480 cadets who responded to this question, so an approximate 85%/15% split of the data meant the model was built using data from 400 cadets, and tested on data from 80 cadets.

The receiver operating characteristic (ROC) curve was created to evaluate the predictive ability of this logistic regression model. We determined our model has about an eighty percent discrimination rate because the AUC is 0.821. The most statistically significant predictor variables found were MA103 Grade, MA104 Grade, and the USMA Admissions CEER Score. Naturally, the cadets who did well in both MA103 and MA104 likely consider themselves good at mathematics.

Table 2: Logistic Regression

Variable	Name	P-Value
x_1	MA103	***0.0222**
x_2	SAT-M	0.892
x_3	ACT-M	0.535
x_4	CEER	***0.0147**
x_5	Gateway	0.695
x_6	FCE	0.129
x_7	MA104	***0.0282**

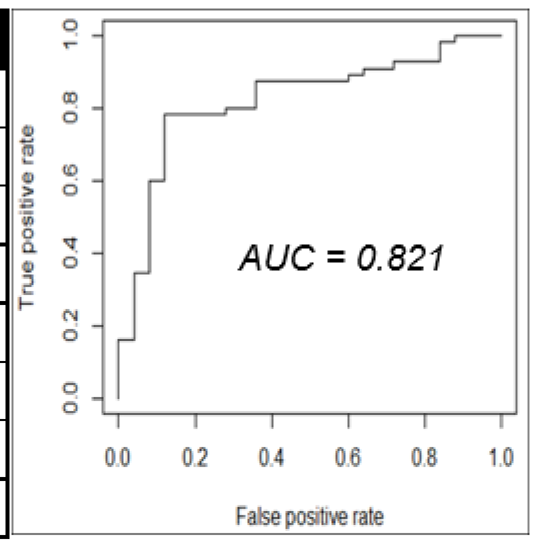


Table 2: *Linear Regression Results. The most statistically significant variables (those with the lowest P-Values) per model are indicated by an asterisks.*

5.4. SENTIMENT ANALYSIS

After performing linear and logistic regression on the numeric data, we then considered the fifteen free text survey responses. These textual entries were converted to numerical values in accordance with sentiment analysis as explained previously. For example, consider one cadet's response to a pre-course survey question,

> "For me, making unsuccessful attempts when attempting to solve a problem is…"

The cadet replied,

> "…tedious. I tend to learn better when I start with simple problems and gradually deal with more difficult problems. I get discouraged when facing a difficult problem when I'm not good at it yet."

After removing the stop words from this response, the words remaining to be analyzed for sentiment were*: tedious, tend, learn, start, simple, gradually, deal, difficult* (twice), *discouraged*, and *facing*. AFNIN scored the response -2 for *discouraged*, -1 for *difficult* (twice), and the overall score is -4. Bing scored the response Negative for *difficult* (twice) and Negative for *tedious*, the overall score is -3. Nrc scored, surprisingly, *deal* as Positive, *learn* as Positive, and *tedious* as Negative, and so the overall score is a +1.

Similarly, we assigned numeric scores to cadet replies to two questions, selected among the fifteen for two purposes: first to verify this sentiment analysis procedure and second to gauge the general attitude cadets have towards problem solving. Thusly, we first considered the scores assigned to cadets' answers to the question, "What did you enjoy the most about this class?" Because this question requires a cadet to think of what they enjoyed most about MA104, their overall response is likely to be more positive. Accordingly, as seen on the left in Figure 5 below, the majority of response scores for this question were positive. Second, we analyzed the response scores to the statement, "For me, making unsuccessful attempts when attempting to solve a problem is…" As illustrated on the right in Figure 5 below, the clear majority of cadets appear to have a negative sentiment towards problem solving.

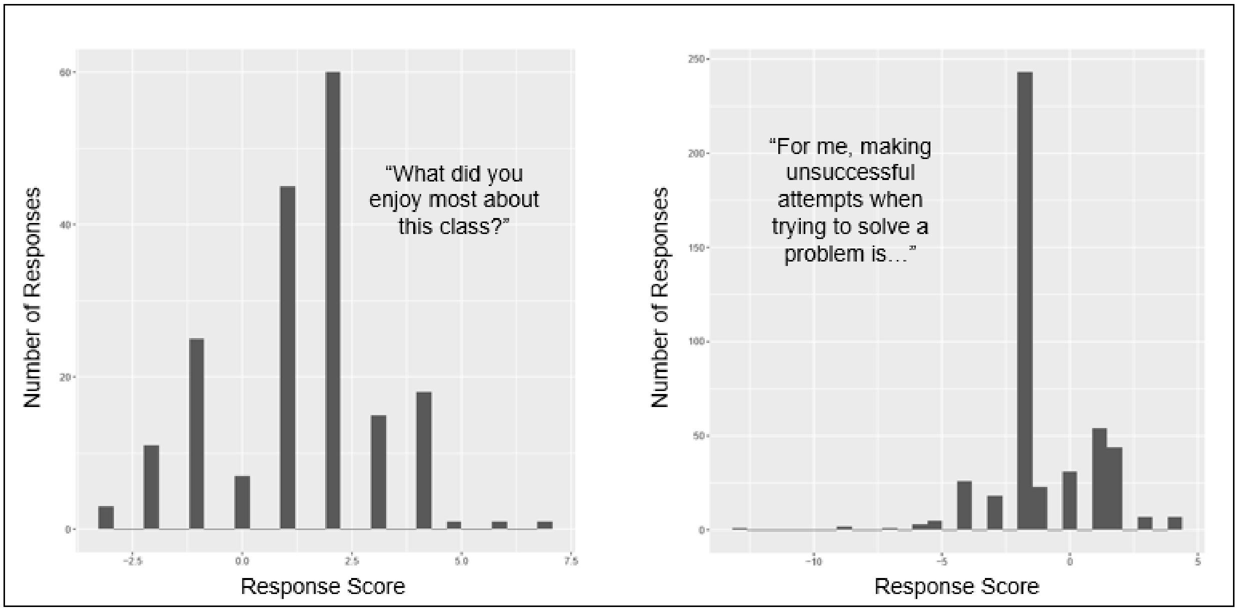


Figure 5: *Histograms depicting the numeric sentiment score assigned to cadet survey responses and the number of responses per score.*

In addition to assigning numerical scores to cadet free-text survey responses, Word Clouds were built to further explore cadet responses to open-ended questions. Two word clouds were built: one for the question used above,

“For me, making unsuccessful attempts when attempting to solve a problem is…,”

the other for the question,

“Do you prefer online videos or textbooks? Why?”

In Figure 6 below, the first is on the left, the second on the right. Not surprisingly, the most dominant word in the first word cloud is frustrating. However, also popular were the words “learning” and “process.” This suggests that while cadets may be irritated encountering difficulties, there is a general appreciation that it is part of the problem solving process. From the second word cloud, we conclude the majority of cadets prefer online videos – which helps inform future course design decisions regarding technological aids and use of the textbook.

Figure 6: *Word Clouds for Cadet Replies to "For me, "For me, making unsuccessful attempts when attempting to solve a problem is…" (left) and "Do you prefer online videos or textbooks? Why?"(right).*

6. CONCLUSIONS

From our results we conclude that the dominant predictor for a cadet receiving a high final grade in MA104 is the cadet's MA103 grade. While the Fundamental Concepts Exam and the USMA Admissions CEER score were also found to be significant predictor variables, MA103 accounted for the majority of the variance in each model built. This relationship is likely not due to similarities in course material. Rather, consider the progression of these two courses – most cadets take MA103 first semester their freshman year, followed by MA104 the second semester. Cadets who have higher MA103 grades have likely adapted well to the academy – they have learned effective study habits and time management skills as they work to overcome the academic rigor of USMA.

We also use our results to suggest course design decisions. Future MA104 sections could be built around MA103 grades, CEER scores, and the Fundamental Concepts Exam scores, due to their prominence as statistically significant predictor variables. At the course director discretion, sectioning options could include assigning the stronger teachers to the weaker sections or carefully balancing the stronger and weaker cadets across sections to encourage collaboration and teamwork. The sentiment analysis performed leads us to believe the accessibility issue in Calculus exists, and certainly at West Point. While some cadets accept difficulties as part of the problem solving process, the overwhelming majority of cadets find it frustrating. Also, we conclude cadets have a large appreciation for online videos as compared to the textbook. Course leadership should keep this in mind as they consider different technological aids to supplement their lectures.

Regarding survey administration, we conclude shorter surveys should be designed to allow the cadets to lead the discussion. As verified in the histograms above, the current survey questions are very pointed and almost lead the cadets to the answers course leadership expects, or wants. Instead, surveys should ask general questions, which prevent biases and polar responses. The use of Likert scale-based questions should also be reduced, because of the inherent difficulties in analyzing their results. That is, the difference between an agree and a strongly agree is difficult to quantify, and can even vary from person to person.

Our goal was to solicit cadet feedback for use to adjust how course concepts are defined, implemented, and assessed in order to foster a favorable commitment to learning that extends beyond the classroom. Instructors foster learning with an exciting learning environment as they develop cadets towards a growth mindset. The desired end state is for each cadet to view MA104 as both a rewarding and positive experience, in which they are developed as problem solvers and thinkers.

APPENDIX A: MA104 INITIAL & FINAL COURSE SURVEYS AY17-02

MA104 ▸
MA104 Initial Course Survey AY17-02: Respond to this Survey

Finish Cancel

* indicates a required field

In high school I was allowed to use graphing calculators on exams. *

- Always
- Sometimes
- never

Which of the following calculus courses did you take prior to West Point? (Mark all that apply) *

- ☐ Pre-Calculus
- ☐ Calculus (Non AP)
- ☐ AP Calculus AB
- ☐ AP Calculus BC
- ☐ International Baccalaureate (IB) Calculus
- ☐ Online calculus course
- ☐ Calculus course taught at my high school for which I also received college credit
- ☐ Calculus course taught at a two-year college for which I also received high school credit
- ☐ Calculus course while I was enrolled in college
- ☐ Calculus course taught at the USMA Prep School

My mathematics courses in high school have prepared me to: *

	Strongly Disagree 1	2	Neutral 3	4	5	Strongly Agree 6
a. complete calculations without a calculator						
b. solve word problems						
c. factor expressions						
d. solve equations						
e. solve inequalities						

The teacher of my last mathematics course in high school: *

	Strongly Disagree 1	2	Neutral 3	4	5	Strongly Agree 6
a. lectured most of the time						
b. primarily showed us how to get answers to specific questions						
c. frequently had us work in groups						
d. frequently had us solve challenging problems						
e. cared that I was successful in the course						

Please select the most appropriate answer *

	Strongly Disagree 1	2	Neutral 3	4	5	Stronly Agree 6
a. I am comfortable using a graphing calculator						
b. I am comfortable in using a computer algebra system (Mathematica, etc.)						

In high school I was allowed to use the following on exams *

	Never 1	Sometimes 2	Always 3
a. graphing calculator (TI83, TI89, etc.)			
b. scientific calculator (TI32X, etc.)			
c. computer algebra system (Mathematica, Matlab, Maple, etc.)			
d. web-based solver (Wolfram-Alpha, etc.)			

What grade do you expect to get in this course? *

- A
- B

- C
- D
- F

Do you intend to take Calculus II (MA205)? *

- Yes
- No
- I don't know yet

How important is a good grade in this course in influencing your decision whether or not to take Calculus II (MA205)? *

- Very Important
- Important
- Slightly Important
- Slightly unimportant
- Unimportant
- Not Important al all

Is Calculus II (MA205) required for your intended major? *

- yes
- No
- I don't know

Please select the most appropriate answer. *

	Strongly Disagree 1	2	3	Neutral 4	5	Strongly Agree 6
a. I believe I have the knowledge and abilities to succeed in this course	O	O	O	O	O	O
b. I am confident in my mathematical abilities	O	O	O	O	O	O
c. I understand the mathematics that I have studied	O	O	O	O	O	O
d. I enjoy doing mathematics	O	O	O	O	O	O

Select the best answer: *

	Not at all 1	Maybe 2	3	Yes 4
Do you think taking Calculus I (MA104) will influence you to become a STEM major?	O	O	O	O

When experiencing a difficulty in math class: *

- I try hard to figure it out on my own
- I quickly seek help
- I give up trying

A. Please select the most appropriate answer below: *

	A natural part of solving the problem 1	2	3	An indication of my weakness in mathematics 4
For me, making unsuccessful attempts when solving a mathematics problem is:	O	O	O	O

B. Please select the most appropriate answer below: *

	Solve specific kinds of problems 1	2	3	Make connections and form logical arguments 4
My success in mathematics PRIMARILY relies on my ability to:	O	O	O	O

C. Please select the most appropriate answer below: *

	I understnad the covered material 1	2	3	I can do things the way the teacher wants 4
My score on my mathematics exam is a measure of how well	O	O	O	O

D. Please select the most appropriate answer below: *

	I would never take another mathematics course 1	2	3	I would continue to take mathematics 4
If I had a choice,	O	O	O	O

E. Please select the most appropriate answer below: *

	Memorize it the way it is presented 1	2	3	Make sense of the material so that I understand it 4
When studying mathematics using a textbook or course materials, I tend to:	O	O	O	O

F. Please select the most appropriate answer below: *

	understand underlying mathematical ideas 1	2	3	find answers to problems 4
When solving mathematics problems, graphing calculators or computers help me to:	O	O	O	O

G. Please select the most appropriate answer below: *

	work problems so students know how to do them 1	2	3	help students learn to reason through problems on their own 4
The primary role of a mathematics instructor is to:	O	O	O	O

H. Please select the most appropriate response below: *

	Strongly disagree 1	2	3	Neutral 4	5	Strongly Agree 6
Mathematics instructors should show students how mathematics is relevant	O	O	O	O	O	O

I. Please select the most appropriate response below: *

	Personal Interest 1	2	3	Points associated with the assignment 4
My motivation to complete assignments is mostly influenced by	O	O	O	O

J. Please select the most appropriate response below: *

	Strongly Disagree 1	2	3	Neutral 4	5	Strongly Agree 6
a. In order to succeed in college calculus, I must have taken calculus before	O	O	O	O	O	O
b. I feel prepared to take Calculus I (MA104)	O	O	O	O	O	O
c. I would take this course if it were not a requirement	O	O	O	O	O	O

K. Please select the most appropriate response below: *

	I really don't enjoy it 1	2	3	4	5	I really enjoy it 6
The extent at which I enjoy mathematics is	O	O	O	O	O	O

Please select the best answer for each of the following.

	Strongly Disagree 1	2	3	Neutral 4	5	Strongly Agree 6
a. If I am unable to solve a problem within a few minutes, it is an indication of my weakness in mathematics.	O	O	O	O	O	O
b. Mathematics is about getting exact answers to specific problems.	O	O	O	O	O	O
c. The process of solving a problem that involves mathematical reasoning is a satisfying experience.	O	O	O	O	O	O

I learn mathematics best by *

	Strongly Disagree 1	2	3	Neutral 4	5	Strongly Agree 6
a. group work on problems	O	O	O	O	O	O
b. individual work on problems	O	O	O	O	O	O
c. lecture	O	O	O	O	O	O

What was your home zip code at the time you graduates from high school? If you graduated from a high school outside of the U.S., enter 00000 *

Do the following people see you as a person who is good at mathematics? *

	Not at all 1	2	3	4	5	Very Much 6
a. Yourself	○	○	○	○	○	○
b. Parents/Relatives	○	○	○	○	○	○
c. High school mathematics teacher	○	○	○	○	○	○
d. Friends	○	○	○	○	○	○

In your own words (and without a reference) define a limit. Use one or two sentences. *

In your own words (and without a reference) define an integral. Use one or two sentences. *

In your own words (and without a reference) define a tangent line. Use one or two sentences. *

In your own words (and without a reference) define a function. Use one or two sentences. *

In your own words (and without a reference) describe the difference between continuous and discrete data. Use one or two sentences. *

If you have taken calculus before, how do you think MA104 will be different from other calculus courses you have taken? If you have not taken calculus before, write "N/A". *

Finish   Cancel

MA104 ▸
MA104 Cadet Final Survey AY 17-2:
Respond to this Survey

Finish Cancel

* indicates a required field

Please select the most appropriate response to the following: *

	Strongly disagree 1	2	3	Neutral 4	5	Strongly Agree 6
1) I am a strategic learner (i.e. I learn to perform well on tests, but do not retain the material afterward)	○	○	○	○	○	○
2) I enjoy math	○	○	○	○	○	○
3) I am a math person	○	○	○	○	○	○
4) I primarily learn math from algebraic examples	○	○	○	○	○	○
5) I primarily learn math from graphic examples	○	○	○	○	○	○
6) I primarily learn math from numeric examples	○	○	○	○	○	○
7) I am a visual learner	○	○	○	○	○	○
8) I am a deep thinker	○	○	○	○	○	○
9) I am confident in my algebra skills	○	○	○	○	○	○

I took calculus A/B in High School. *

○ Yes
○ No

I took calculus B/C in High School. *

○ Yes
○ No

I first learned about derivatives and how to apply them from a course I took prior to attending West Point. *

○ Yes
○ No

I first learned about integrals and how to apply them from a course I took prior to attending West Point. *

○ Yes
○ No

Please select the best response. *

	Memorize how to solve specific kinds of problems 1	2	3	Make connections and form logical arguments 4
My success in math primarily relies on my ability to	○	○	○	○

Please respond to the following. *

	Strongly Disagree 1	2	3	Strongly Agree 4
I enjoy solving difficult problems on my own.	○	○	○	○

For me, making unsuccessful attempts when attempting to solve a problem is: *

I prepare for class by: *

	Never 1	2	Sometimes 3	4	Always 5
1) Reading the textbook	○	○	○	○	○
2) Completing WebAssign	○	○	○	○	○
3) Working out problems by hand on paper	○	○	○	○	○
4) Asking a tutor questions	○	○	○	○	○
5) Using mathematica	○	○	○	○	○

6) Collaborating with a classmate	O	O	O	O	O
7) Viewing online resources (such as Khan Academy)	O	O	O	O	O

Please provide a response. *

	Not at all effective 1	2	3	Very effective 4
How effective is your class preparation in facilitating your understanding of the material?	O	O	O	O

During class I: *

	Never 1	2	Sometimes 3	4	Always 5
1) ask questions.	O	O	O	O	O
2) contribute to class discussions.	O	O	O	O	O
3) attempt to work problems on my own.	O	O	O	O	O
4) take notes.	O	O	O	O	O
5) follow the lecture using my textbook.	O	O	O	O	O
6) collaborate with classmates.	O	O	O	O	O

Please provide a response below. *

	Not effective 1	2	3	Very effective 4
How effective is your use of class time in facilitating your understanding of the material?	O	O	O	O

After class, I solidify my grasp of the material by: *

	Never 1	2	Sometimes 3	4	Always 5
1) reading the textbook.	O	O	O	O	O
2) completing WebAssign.	O	O	O	O	O
3) working problems out by hand on paper.	O	O	O	O	O
4) watching online tutorials.	O	O	O	O	O
5) asking a tutor questions.	O	O	O	O	O
6) collaborating with a classmate.	O	O	O	O	O
7) seekeing AI from an instructor.	O	O	O	O	O

Please respond below. *

	Not effective 1	2	3	Very effective 4
How effective after class are you in cementing your understanding of the material?	O	O	O	O

When I use WebAssign, I: *

	Never 1	2	Sometimes 3	4	Always 5
1) work on scratch paper and then discard it.	O	O	O	O	O
2) work on scratch paper and then retain it.	O	O	O	O	O
3) rely on collaboration with classmates.	O	O	O	O	O
4) rely on the textbook.	O	O	O	O	O
5) rely on watch-it videos.	O	O	O	O	O
6) rely on mathematica.	O	O	O	O	O
7) rely on instructor AI.	O	O	O	O	O
8) rely on example problem solutions.	O	O	O	O	O

I like WebAssign, because... *

I dislike WebAssign, because... *

If WebAssign were voluntary, I would: *

Strongly disagree	Undecided	Strongly Agree

	1	2	3	4	5	6
1) Try to complete it prior to the lesson.	O	O	O	O	O	O
2) Try to complete it after the lesson.	O	O	O	O	O	O
3) Never look at it or use it.	O	O	O	O	O	O

Please provide a response to the following: *

	Strongly Disagree 1	2	Neutral 3	4	5	Strongly Agree 6
1) My High School math courses prepared me for Blocks I and II (differentiation and applications of differentiation) of MA104.	O	O	O	O	O	O
2) My High School math courses prepared me for Blocks III and IV (integration and applications of integration) of MA104.	O	O	O	O	O	O
3) MA103 prepared me for derivatives in MA104.	O	O	O	O	O	O
4) MA103 prepared me for integrals in MA104.	O	O	O	O	O	O
5) I had the knowledge and ability to succeed in this course.	O	O	O	O	O	O

The material in MA104 was better understood through: *

	Strongly Disagree 1	2	Neutral 3	4	5	Strongly Agree 6
1) Suggested board problems.	O	O	O	O	O	O
2) Instructor AI.	O	O	O	O	O	O
3) Classroom instruction.	O	O	O	O	O	O
4) Classmate collaboration.	O	O	O	O	O	O
5) Use of the textbook.	O	O	O	O	O	O
6) Use of Mathematica.	O	O	O	O	O	O
7) Previous experience (Algebra).	O	O	O	O	O	O
8) Previous experience (Calculus).	O	O	O	O	O	O

Please provide a response to the statement below: *

	Strongly disagree 1	2	Neutral 3	4	5	Strongly agree 6
1) The primary role of a math instructor is to work problems so that students know how to do them.	O	O	O	O	O	O
2) The primary role of a math instructor is to help students reason through problems on their own.	O	O	O	O	O	O

My instructor's use of the following was effective: *

	Strongly Disagree 1	2	Neutral 3	4	5	Strongly Agree 6	N/A
1) Lectures.	O	O	O	O	O	O	O
2) Assignments (including quizzes).	O	O	O	O	O	O	O
3) Problem Solving Labs.	O	O	O	O	O	O	O
4) Board Problems.	O	O	O	O	O	O	O
5) Class discussions.	O	O	O	O	O	O	O
6) Instructor/cadet rapport.	O	O	O	O	O	O	O
7) Mentorship.	O	O	O	O	O	O	O
8) Online resources.	O	O	O	O	O	O	O
9) The course textbook.	O	O	O	O	O	O	O
10) Technology.	O	O	O	O	O	O	O
11) Games (i.e. Kahoot)	O	O	O	O	O	O	O

Please respond to the following with respect to the Problem Sets in MA104: *

	Strongly Disagree 1	2	Neutral 3	4	5	Strongly Agree 6
1) The problem sets were interesting.	O	O	O	O	O	O
2) I enjoyed working on the problem sets.	O	O	O	O	O	O
3) The problem sets were challenging, but beneficial to reinforce concepts.	O	O	O	O	O	O
4) I was able to do the problem sets with no assistance.	O	O	O	O	O	O
5) I was able to do the problem sets with some assistance.	O	O	O	O	O	O
6) I was only able to do the problem sets with a lot of assistance.	O	O	O	O	O	O

7) The technology required on the problem sets was appropriate to my skills.						

Please respond to the following with respect to technology used in MA104: *

	Strongly Disagree 1	2	3	Strongly Agree 4
1) The use of technology increased my motivation to use computers to solve more difficult problems.				
2) Technology facilitates my understanding of the material in this course.				

I used technology in this course to: *

	Never 1	2	Sometimes 3	4	Always 5
1) Complete my daily homework/WebAssign.					
2) Check my answers after doing them by hand.					
3) Complete the Problem Sets.					
4) Visually understand key concepts.					

My instructor used technology in class to: *

	Never 1	2	Sometimes 3	4	Always 5
1) Illustrate difficult concepts.					
2) Find answers to problems.					
3) Check answers after we worked them out by hand.					

Please respond to the following statements. *

	Strongly disagree 1	2	3	Neutral 4	5	Strongly agree 6
1) I am confident in my Mathematica abilities.						
2) I enjoy using Mathematica.						

Please give a response to the following questions and statements: *

	Not at all 1	Maybe 2	3	Absolutely 4
1) Do you think your current grade reflects your level of effort in this course?				
2) Do you think your grade reflects your understanding of the material?				
3) I understand what a derivative is and when to use it solving problems.				
4) I understand what an integral is and when to use it in solving problems.				

I could have done better in this course had we spent more time on: *

	Strongly Disagree 1	2	Neutral 3	4	Strongly Agree 5
1) Block 1: Differentiation					
2) Block 2: Applications of Differentiation					
3) Block 3: Integration					
4) Block 4: Applications of Integration					

Please respond to the following statements about your classroom experience in MA104: *

	Strongly disagree 1	2	3	Neutral 4	5	Strongly Agree 6
1) I enjoyed MA104.						
2) I enjoyed my instructor.						
3) MA104 was an exciting class.						
4) MA104 was a worthwhile class.						
5) If MA104 were not a core course, I would recommend to other cadets to take it.						

Please indicate the level at which you agree or disagree with the following: *

	Strongly Disagree 1	2	3	Neutral 4	5	Strongly Agree 6
1) I am interested in the material from MA104.						

2) This course has increased my interest in taking more math courses.	O	O	O	O	O	O
3) Calculus is an exciting field in math.	O	O	O	O	O	O
4) Calculus is about getting exact answers to specific problems.	O	O	O	O	O	O

What did you enjoy the most about this class? *

What did you enjoy the least about this class? *

Please respond to the following question. *

	No 1	2	Maybe 3	4	Yes 5
If we offered an algebra review course, would you take it?	O	O	O	O	O

Please respond to the following question regarding the course textbook: *

	Less likely to read it 1	2	I would probably read it as much as the current textbook 3	4	More likely to read it 5
If we changed the textbook to a book that is less technical (more conversational tone, but still rigorous), would you be more or less likely to read it?	O	O	O	O	O

Please respond to the following question about online videos versus textbooks. *

	No, I prefer textbooks 1	2	3	Yes, I prefer online videos over textbooks 4
Do you prefer online videos to a textbook?	O	O	O	O

Please explain your answer to the previous question. *

In your own words, what is a derivative? *

In your own words, what is an integral? *

In your own words, what is the Fundamental Theorem of Calculus? *

Finish Cancel

APPENDIX B: CONCORDANCE TABLES

Given ACT Math, Estimate SAT Math:	
ACT Math	SAT Math
36	800
35	790
34	780
33	740
32	720
31	700
30	680
29	650
28	640
27	620
26	600
25	580
24	560
23	540
22	520
21	500
20	480
19	460
18	440
17	420
16	390
15	360
14	330
13	290
12	250
11	220

Given SAT Math, Estimate ACT Math:			
SAT Math	ACT Math	SAT Math	ACT Math
800	36	500	21
790	35	490	20
780	34	480	20
770	34	470	19
760	33	460	19
750	33	450	18
740	33	440	18
730	32	430	18
720	32	420	17
710	31	410	17
700	31	400	17
690	31	390	16
680	30	380	16
670	30	370	15
660	29	360	15
650	29	350	15
640	28	340	14
630	28	330	14
620	27	320	14
610	27	310	14
600	26	300	13
590	26	290	13
580	25	280	13
570	25	270	13
560	24	260	12
550	23	250	12
540	23	240	12
530	22	230	11
520	22	220	11
510	21		

APPENDIX C: R CODE

```
##### Linear Regression #####

# Set Up workspace: Set Working Directory and Open Libraries
setwd("C:/Users/x86964/Desktop/Academics/18-2/Thesis")
library(tidyr)
library(janeaustenr)
library(tidyselect)
library(tidytext)
library(dplyr)
library(stringr)
library(ggplot2)

# Read the CSV file to work with
# Fixed Data1.csv = Fully scrubbed and imputed data set

info <- read.csv("FixedData1.csv",header=T)
attach(info)

# Next, we create 3 linear models: Numeric, Categorical, Both
# For each model, the response variable will be the cadet's final MA104
grade, or "ma104pct"

# Model 1: Numeric

# Use the 6 numeric predictors
# Confirm each variable is of the right class, numeric/integer
class(ma104pct)
class(ma103pct)
class(SATmath)
class(ACTmath)
class(ceer)
class(gateway)
class(fce1)

# Confirmed. Build the dataset.
table1 <- info[,c(35,8,11,13,16,17,19)]

# Create the model
model1 <- lm(ma104pct~.,data = table1)
# Analyze results of the model
summary(model1)

# Create Validation Plots
yhat1<-model1$fitted.values
standard.resid1<-rstandard(model1)
plot(yhat1,standard.resid1,main="Standard Residuals vs. Y",xlab="MA104
Percent",ylab="Standardized Residuals", ylim=c(-3,3))
abline(h=c(0,-2,2),lty=3)

qqnorm(standard.resid1)
qqline(standard.resid1)

par(mfrow=c(1,2))
```

```
hist(standard.resid1, breaks = 10, freq=T, xlim = c(-4,2),xlab =
"Standardized Residuals",main = "Histogram", col = "red")
resid.density1 <- density(standard.resid1)
plot(resid.density1, type = "l", xlab = "Standardized Residuals",main =
"Smoothed Histogram", col = "black", lwd = 2)
par(mfrow = c(1,1))


# Model 2: Categorical

# Use the 10 categorical predictors
# Confirm each variable is of the right class, factor
# If the variable is not, use 'as.factor' to correct
# Convert Calculus Validatiion (Gateway), AP Calc AB, and AP Calc BC to
binary values

class(MA104InstructorID)
MA104InstructorID<-as.factor(info$MA104InstructorID)
class(sex)
class(race)
class(standm)
info$standm<-as.factor(info$standm)
info$calcvalidation <- ifelse(is.na(info$svscore),0,1)
class(info$calcvalidation)
info$calcvalidation<-as.factor(info$calcvalidation)
info$tookcalcAB <- ifelse(is.na(info$svcalcAB),0,1)
info$tookcalcBC <- ifelse(is.na(info$svcalcBC),0,1)
class(info$tookcalcAB)
info$tookcalcAB<-as.factor(info$tookcalcAB)
class(info$tookcalcBC)
info$tookcalcBC<-as.factor(info$tookcalcBC)
class(info$usmaps)
info$usmaps<-as.factor(info$usmaps)
class(stem)
info$stem<-as.factor(info$stem)
class(recathlete)
info$recathlete<-as.factor(info$recathlete)

# Confirmed. Build the dataset.
table2 <- info[,c(35,3,4,5,10,236,237,238,26,30,31)]

# Create the model
model2 <- lm(ma104pct~.,data = table2)
# Analyze results of the model
summary(model2)

# Create Validation Plots
yhat2<-model2$fitted.values
standard.resid2<-rstandard(model2)
plot(yhat2,standard.resid2,main="Standard Residuals vs. Y",xlab="MA104
Percent",ylab="Standardized Residuals", ylim=c(-3,3))
abline(h=c(0,-2,2),lty=3)

qqnorm(standard.resid2)
qqline(standard.resid2)

par(mfrow=c(1,2))
```

```
hist(standard.resid2, breaks = 10, freq=T, xlim = c(-4,2),xlab =
"Standardized Residuals",main = "Histogram", col = "red")
resid.density2 <- density(standard.resid2)
plot(resid.density2, type = "l", xlab = "Standardized Residuals",main =
"Smoothed Histogram", col = "black", lwd = 2)
par(mfrow = c(1,2))


# Model 3: Combination of Numeric & Categorical

# Use all 16 predictor variables
# Variables are already confirmed to be of the right class

table3<-info[,c(35,3,4,5,10,236,237,238,26,30,31,8,11,13,16,17,19)]

# Create the model
model3 <- lm(ma104pct~.,data = table3)
# Analyze results of the model
summary(model3)

# Create Validation Plots
yhat3<-model3$fitted.values
standard.resid3<-rstandard(model3)
plot(yhat3,standard.resid3,main="Standard Residuals vs. Y",xlab="MA104
Percent",ylab="Standardized Residuals", ylim=c(-3,3))
abline(h=c(0,-2,2),lty=3)

qqnorm(standard.resid3)
qqline(standard.resid3)

par(mfrow=c(1,2))
hist(standard.resid3, breaks = 10, freq=T, xlim = c(-4,2),xlab =
"Standardized Residuals",main = "Histogram", col = "red")
resid.density3 <- density(standard.resid3)
plot(resid.density3, type = "l", xlab = "Standardized Residuals",main =
"Smoothed Histogram", col = "black", lwd = 2)
par(mfrow = c(1,2))
```

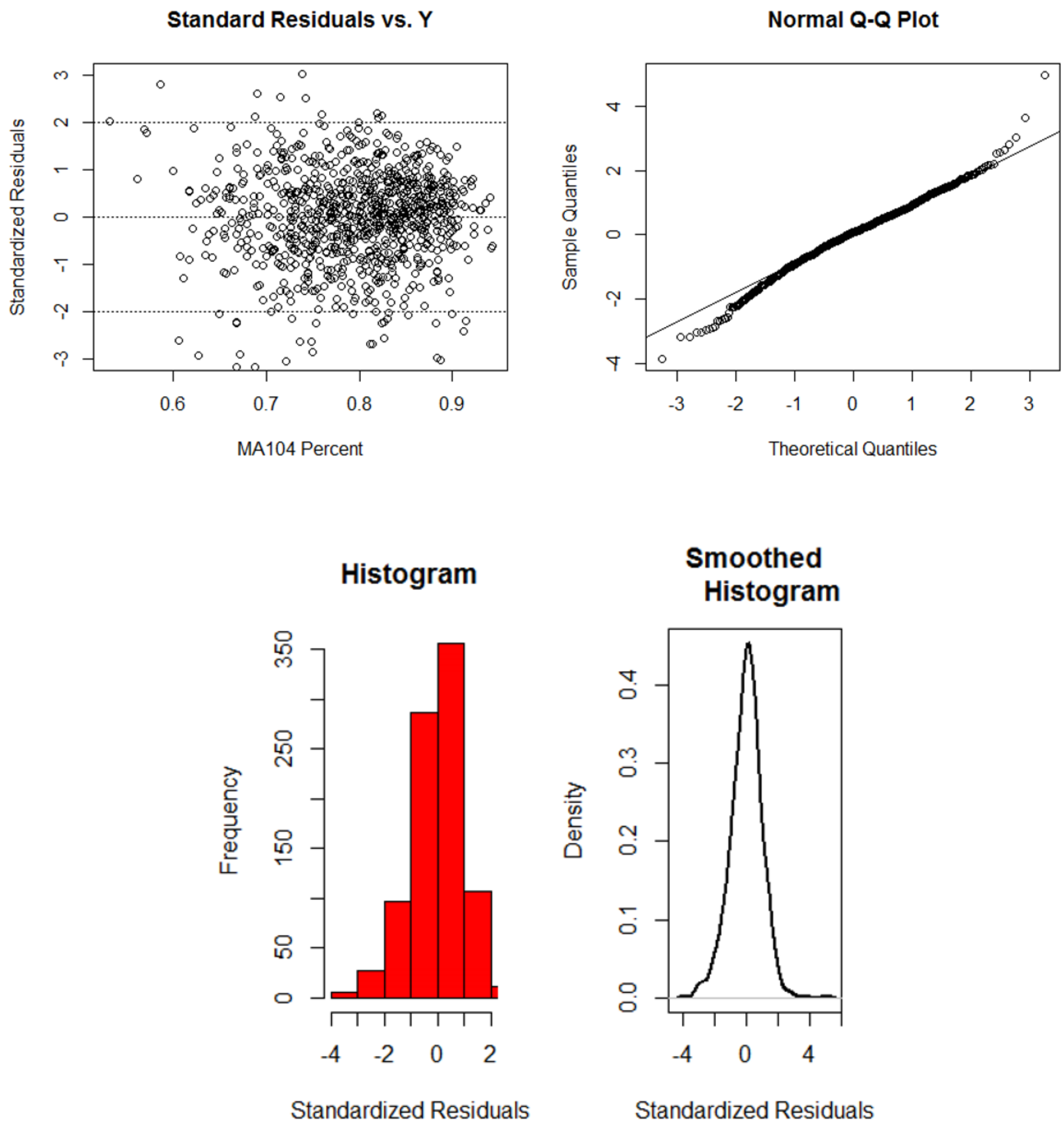
```
##### Logistic Regression #####

# Set Up workspace: Set Working Directory and Open Libraries
setwd("C:/Users/x86964/Desktop/Academics/18-2/Thesis")
library(pscl)
library(ROCR)

# Read the CSV file to work with
# LogitData.csv = Fully scrubbed and imputed data set

info <- read.csv("LogitData.csv",header=T)
attach(info)

# Next, create a logistic model where the predictor variables are the six
numeric predictor variables
# previously used, plus MA104 Grade
# The response variable is a 0 or 1 - the cadet disagrees or agrees

# Use the 6 numeric predictors
# Confirm each variable is of the right class, numeric/integer
class(ma104pct)
class(ma103pct)
class(SATmath)
class(ACTmath)
class(ceer)
class(gateway)
class(fce1)

# the response variable must be factor (2 levels, 0 or 1)
info$ask1 <- as.factor(info$ask1)

# Build the dataset.
table1 <- info[,c(95,35,8,11,13,16,17,19)]

# Split the data into a Training/Testing Set (~85%/15% split)
train <- table1[1:400,]
test <- table1[401:480,]

# Build the model.
model1 <- glm(ask3 ~., family = binomial(link = 'logit'),data = train)
# Analyze the results of the model.
summary(model1)
exp(coef(model1))
anova(model1,test="Chisq")

# Create the ROC curve and find AUC
p <- predict(model1, newdata = test, type = 'response')
pr <- prediction(p, test$ask1)
prf <- performance(pr, measure = "tpr",x.measure = "fpr")
plot(prf)

auc <- performance(pr, measure = "auc")
auc <- auc@y.values[[1]]
auc
```

```
##### Sentiment Analysis #####

# Set Up workspace: Set Working Directory and Open Libraries
setwd("C:/Users/x86964/Desktop/Academics/18-2/Thesis")
library(wordcloud)

# Assign numeric scores to cadet replies to 'What did you enjoy the most about
this class?'

# Step 1: call the data set
data<- read.csv("CalculusSurveyDataCleanwInstructor.csv", header=T)

# 2: Create the data set in tidytext format
text <- data[,118]
text <- as.character(text)
str(text)
text_df <- data_frame(line = 1:897, text = text)
text_df <- text_df %>% unnest_tokens(word(), text)

# 3: Get rid of the "stop" words
text_df_wostopwords <- text_df %>%
  anti_join(stop_words, by = c("word()" = "word"))

# 4: Use AFINN to get scores per line
afinn <- get_sentiments("afinn")
text_df_afinn <- text_df %>%
  inner_join(get_sentiments("afinn"), by = c("word()" = "word"))

# 5: Sum words per response to get one total response score
text_df_afinn2 <- text_df_afinn %>%
  group_by(line) %>%
  summarise(sum=sum(score)) %>%
  arrange(-sum)

# 6: Create visual - histogram
ggplot(data=text_df_afinn2, mapping=aes(x=sum)) +
  geom_histogram()


# Assign numeric scores to cadet replies to 'For me, making unsuccessful
attempts when attempting to solve a problem is…'
# Repeat steps 2-6

# 2: Create the data set in tidytext format
text <- data[,115]
text <- as.character(text)
str(text)
text_df <- data_frame(line = 1:897, text = text)
text_df <- text_df %>% unnest_tokens(word(), text)

# 3: Get rid of the "stop" words
text_df_wostopwords <- text_df %>%
  anti_join(stop_words, by = c("word()" = "word"))

# 4: Use AFINN to get scores per line
afinn <- get_sentiments("afinn")
text_df_afinn <- text_df %>%
```

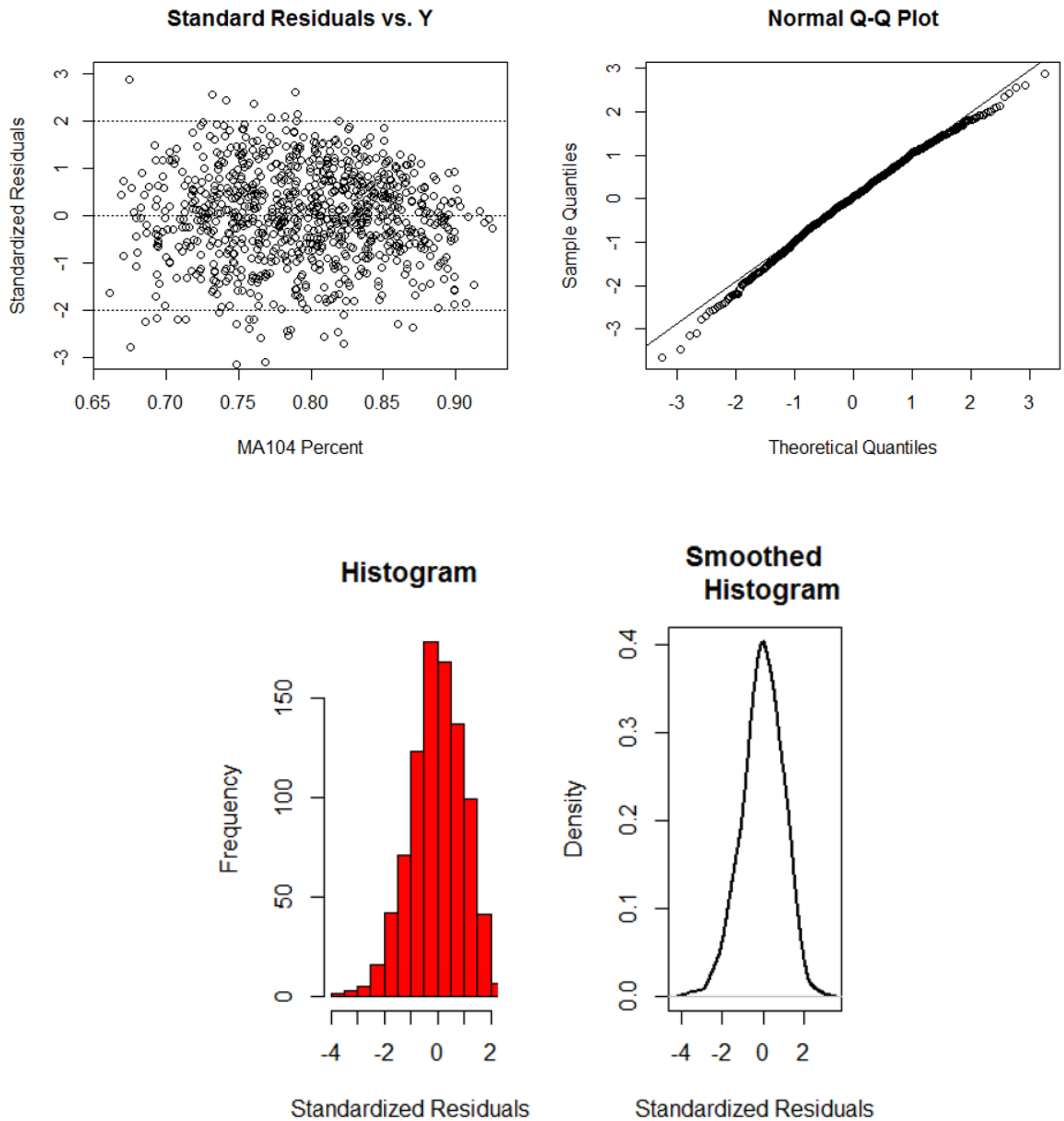
```
  inner_join(get_sentiments("afinn"), by = c("word()" = "word"))

# 5: Sum words per response to get one total response score
text_df_afinn2 <- text_df_afinn %>%
  group_by(line) %>%
  summarise(sum=sum(score)) %>%
  arrange(-sum)

# 6: Create visual - histogram
ggplot(data=text_df_afinn2, mapping=aes(x=sum)) +
  geom_histogram()


# Build Word Cloud for Cadet Replies to 'For me, making unsuccessful attempts
when attempting to solve a problem is…'

# Step 1: Create the data set in tidytext format
text115 <- data[,115]
text115 <- as.character(text115)
text115_df <- data_frame(line = 1:897, text = text115)
text115_df <- text115_df %>% unnest_tokens(word(), text)

# 2: Get rid of the "stop" words
text115_df <- text115_df %>%
  anti_join(stop_words, by = c("word()" = "word"))

# 3: Create word cloud
text115_df %>%
  count(`word()`) %>%
  with(wordcloud(`word()`, n, max.words = 100))


# Build Word Cloud for Cadet Replies to 'Do you prefer online videos or
textbooks? Why?'
# Repeat steps 1-3

# Step 1: Create the data set in tidytext format
text123 <- data[,123]
text123 <- as.character(text123)
text123_df <- data_frame(line = 1:897, text = text123)
text123_df <- text123_df %>% unnest_tokens(word(), text)

# 2: Get rid of the "stop" words
text123_df <- text123_df %>%
  anti_join(stop_words, by = c("word()" = "word"))

# 3: Create word cloud
text123_df %>%
  count(`word()`) %>%
  with(wordcloud(`word()`, n, max.words = 100))
```

APPENDIX D: LINEAR REGRESSION VALIDATION PLOTS

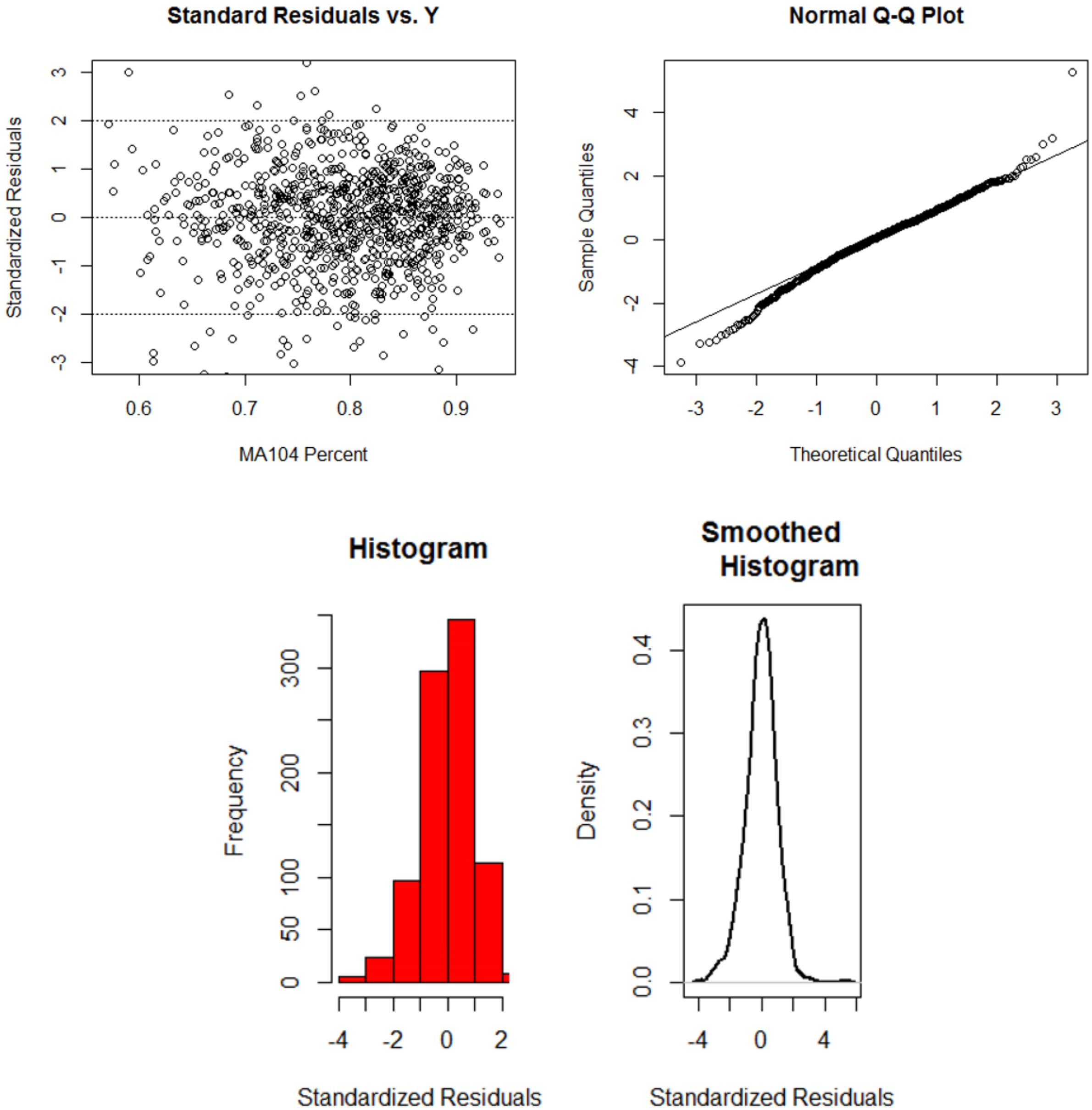
Model 1:



Model 2:


Standard Residuals vs. Y
Standardized Residuals
MA104 Percent
0.65 0.70 0.75 0.80 0.85 0.90
3 2 1 0 -1 -2 -3
Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles
-3 -2 -1 0 1 2 3
Histogram
Frequency
0 50 100 150
-4 -2 0 2
Standardized Residuals
Smoothed Histogram
Density
0.0 0.1 0.2 0.3 0.4
-4 -2 0 2
Standardized Residuals

Model 3:


Standard Residuals vs. Y
Standardized Residuals
MA104 Percent
Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles
Histogram
Frequency
Standardized Residuals
Smoothed Histogram
Density
Standardized Residuals

REFERENCES

Allen, I. E. & Seaman, C.A. (2007). Likert Scales and Data Analyses. *Quality Progress*, *40* (7), 64 – 65.

AP Central. AP Calculus AB Freqeuntly Asked Questions. Retrieved from: https://apcentral.collegeboard.org/courses/ap-calculus-ab/course/frequently-asked-questions?course=ap-calculus-ab

Betzel, I. M. (2017). *Scholastic Excellence Beginning to End: A Study of Academic Success Indicators for West Point Cadets.* United States Military Academy, West Point, NY.

Boone, H.N. & Boone, D. A. (2012). Analyzing Likert Data. *Journal of Extension, 50* (2), 1 – 5.

Bressoud, D.M., Carlson, M.P., Mesa, V., & Rasmussen, C. (2013). The calculus student: insights from the Mathematical Association of America national study. *International Journal of Mathematical Education in Science and Technology, 44* (5), 685 – 698.

Gliem, J.A. & Gliem, R. R. (2003). Calculating, Interpreting, and Reporting Cronbach's Alpha Reliability Coefficient for Likert-Type Scales. *Presented at the Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education,* October, Columbus, OH.

Hastie, T., Tibshirani, R., & Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction (*2nd ed.). Retrieved from: https://web.stanford.edu/~hastie/Papers/ESLII.pdf

Hood, K.M. & Kuiper, P.K. (2017). *Improving Student Surveys with NLP* (Unpublished). United States Military Academy, West Point, NY.

Lindsay, S. R. (2017). *An Exploration Into and Statistical Analysis of Cadet-Candidate Applications and Academy Performance*. United States Military Academy, West Point, NY.

Silge, J. & Robinson, D. (2018). *Text Mining with R: A Tidy Approach*. Retrieved from: https://www.tidytextmining.com/

Sullivan, G. M. & Artino, A.R. (2013). Analyzing and Interpreting Data From Likert-Type Scales. *Journal of Graduate Medical Education, 5,* 541 – 542.