2018

A Statistical Analysis of Students' Attitudes and Achievement in Introductory Statistics Courses

Kaitlin M. Dio
University of Rhode Island, katiebell869@gmail.com

A STATISTICAL ANALYSIS OF STUDENTS' ATTITUDES AND ACHIEVEMENT IN INTRODUCTORY STATISTICS COURSES

BY

KAITLIN M. DIO

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

STATISTICS

UNIVERSITY OF RHODE ISLAND

2018

MASTER OF SCIENCE THESIS

OF

KAITLIN M. DIO

APPROVED:

Thesis Committee:

Major Professor Natallia Katenka

Steffen Ventz

Ashley Buchanan

Nasser H. Zawia

DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND

2018

ABSTRACT

The Millennial Generation has been phasing out of undergraduate classrooms since 2013 and is being replaced by the technologically savvy and visual learners of Generation Z. To help to increase our understanding of the learning needs and attitudes of this new population of students, a two-fold data collection design has been implemented in undergraduate statistics classes at the University of Rhode Island. In the first round of data collection during the spring 2016 semester, survey and grade data was collected from an introductory biostatistics class pertaining to 146 students. Results from the analysis including the use of longitudinal generalized linear mixed models, hierarchical linear models and regression trees indicated a relationship between time and student performance throughout the semester, as well as a relationship between students starting attitudes and their performance and a potential group structure in the class based on their attitudes.

This first round of data collection and analysis lead to interesting results about students starting attitudes and the effect on their performance. To further explore these results and extend them to more than one course, a second round of data collection was completed during the spring 2017 semester. Principal component analysis in connection with regression analysis indicated a relationship between students starting attitudes and their course performance. Cluster analysis indicated a two group structure in starting attitudes of the students in each course, with each cluster showing different achievement and learning preferences.

ACKNOWLEDGMENTS

I would first and foremost like to thank Dr. Natallia Katenka for her unending patience and guidance throughout this whole process. Without her advice and encouragement, this thesis would not be possible. I would also like to thank my committee members Dr. Ashley Buchanan and Dr. Steffen Ventz for their advice and involvement in helping me to complete this body of work. Along with my committee, a special thank you to Dr. Cornelis de Groot for agreeing to chair my defense.

This thesis would not be possible without the participation and support of several educators within the statistics department: Dr. Natallia Katenka, Dr. Steffen Ventz, Dr. Liliana Gonzalez, Catherine Robinson, and Prince Allotey. Thank you all for allowing me access to your classrooms and students.

A special thank you to my closest friends and family for all of their support throughout the past two years. I would especially like to thank my parents for their support and worry keeping me on track.

Lastly, to everyone within the computer science and statistics department who has helped make my time as a masters student so memorable- thank you!

PREFACE

This thesis was formatted with accordance to the manuscript format guidelines established by the graduate school at the University of Rhode Island. Manuscript one has been submitted and is under review at the Journal of Statistics Education. The second manuscript has been prepared with the intention to submit to the Journal of Statistics Education as well. The co-author on both papers is Dr. Natallia Katenka.

The work within this thesis is motivated by the changing landscape of learners within undergraduate classes, specifically those within statistics courses. Currently, there is a shift in generations and with it comes a shift in the learning styles, focus and interests of the students. The work began during the Spring 2016 semester when data was collected in an introductory biostatistics course at the University of Rhode Island. Survey and grade data were collected from 146 undergraduate students about their attitudes towards statistics, learning preferences and mathematical background. The description and results from this round of data collection are detailed in manuscript one.

Following the first round of data collection, a second, more broad round of data collection was completed during the Spring of 2017. During this round, data were collected from all introductory statistics courses at URI with surveys pertaining to the students' starting and ending attitudes towards statistics and introductory and exit surveys about their learning and study preferences. The results from this second phase are detailed in manuscript two.

Together, the body of research within this thesis aims to answer the following research questions: (1.) What factors affect student performance throughout the semester? (2.) Do students' starting attitudes towards statistics affect their course performance? (3.) Are there groups of students with similar attitudes towards statistics and if so, how do these groups differ in learning styles, study habits and course performance?

TABLE OF CONTENTS

Page

LIST OF FIGURES

Figure		Page

LIST OF TABLES

Table		**Page**

MANUSCRIPT 1

Increasing Feedback from Generation Z

This manuscript has been submitted to the Journal of Statistics Education and is currently under review. The paper was co-authored by Dr. Natallia Katenka.

Abstract

The Millennial Generation is phasing out of undergraduate classes and being replaced by the technologically savvy and visual learners of Generation Z. To help to increase our understanding of the learning needs and attitudes of this new population of students, we collected survey and grade data in an introductory biostatistics course pertaining to 146 students at the University of Rhode Island. Our purpose was three-fold. First, to increase the amount of immediate feedback collected from students by implementing weekly quizzes. These quizzes were analyzed using longitudinal mean response profiles and generalized linear mixed models to discover a significant effect of time on the student performance, but not of grade incentives. Next, students attitudes towards statistics were analyzed to determine how the starting attitudes effected performance using hierarchical linear models to find a significant effect of starting affect and cognitive competence on students final grades. Finally, regression trees were utilized to identify groups of learners who increased their attitude throughout the semester dependent on their starting attitude and final grade. These results lead to practical implications for instructors as they plan the timing of their instruction within a course, as well as the importance of identifying students' confidence and feelings towards the subject at the start of the course and to hopefully minimize the impact of these negative attitudes on their students' performance.

1.1 INTRODUCTION

1.1.1 Motivation

The Millennial Generation (all persons born from the early 1980s to the mid-1990s) is phasing out of undergraduate courses and the next generation is replacing them. Penned Generation Z these students were born into a world with technology, where their phones contained the answers to nearly any question they could ask. As described by Shatto and Erwin (2016), there are several consequences to this upbringing with instant gratification from technology. One result is the lowered attention span of eight seconds, which shows a sharp decrease from the Millennial Generation's time span of 12 seconds. Another result is an increased ability to understand visual imagery (Shatto and Erwin, 2016).

This shift in generations calls for an update to teaching methodology and techniques. This generation of students did not adapt to the introduction of technology like the Millennial Generation, but rather they grew up with applications like Snapchat, YouTube, Facebook and more available at a moments notice. When any question can be answered with a quick Google search or trip to Wikipedia, the idea of listening to lectures and reading from textbooks is not only unappealing, but very dissimilar to the normal way of learning for these students. This generation can get any answer they want in seconds, but their ability to validate and further interpret these answers may be absent (Shatto and Erwin, 2016). No longer do students look to books for answers, now the knowledge is held in their mobile devices. But how do we adapt teaching to this new generation? The first step to any unknown is to gather more information about these students and their needs.

Getting feedback from students about their understanding of ongoing topics and learning preferences can be incredibly difficult when many students are afraid to ask or answer questions in front of their peers and do not participate in office hours. This is especially true in large (>100 students) lecture sections, such as introductory statistics courses, which are filled with a diverse set of students from various mathematical backgrounds, many of whom have a low or neutral evaluation of the subject and little inclination to participate. In response to these challenges, an ongoing study to incorporate additional feedback strategies and garner more information about students attitudes and achievement has been implemented in undergraduate statistics courses here at the University of Rhode Island. In this study, an interactive feedback framework was implemented in the Spring 2016 Introduction to Biostatistics courses which included the use of weekly quizzes, an introductory survey and two attitude surveys.

1.1.2 Related Work

Many studies have evaluated the use of immediate feedback in large courses through the use of clickers. Typically, clickers (small electronic devices with educational software to collect student responses) are implemented in lectures or labs to allow for student to answer a small number of questions, either before, during or after instruction. According to Dunham (2009), clickers have traditionally been used in many large lecture courses with the potential benefits of improved attendance, immediate feedback for students, the ability to revisit challenging topics and continuous assessment throughout the lecture and semester (Dunham, 2009). In a study on student perception of clickers, Vaterlaus et al. (2012) found a positive perception overall and a significant effect of clicker usage on student recall on exams (Vaterlaus et al., 2012).

In a randomized experiment in an introductory statistics course, McGowan and Gunderson (2010) studied the effect of clicker use on engagement and learning during lab sections. In this study, there was little evidence that clicker use increased students engagement; however there was an effect on student learning if the number of questions were low and well assimilated with the material. The researchers also studied the effect of external incentives on student clicker participation and discovered that students were much more likely to participate in clicker questions when given external motivation (McGowan and Gunderson, 2010). This study did not however look at the actual responses or grades on the questions, but rather if students answered at least 1 or at least 50% of the questions.

Many of the studies used clickers during lecture periods, rather than in recitations after the weeks lectures. Also, they required the additional cost and hardware of using clicker software, whereas the use of smart phone technology using either the smart phone application or on-line website for Socrative in the spring 2016 course utilized a familiar device to students and did not add financial burden to the students (Soc, 2017).

In addition to the use of weekly quizzes, this study implemented surveys to measure students attitudes at the beginning and end of the course. Many instruments have been published to measure students attitudes towards statistics. In this study, the SATS-36 (Survey of Attitudes Toward Statistics) was chosen. The SATS-36 was created by Candace Schau, a professor who taught statistics courses for over 25 years, to help better understand the attitudes of students and the effect on teaching and learning. The first version of the SATS, called SATS-28, contained 28 Likert-type scale questions to assess four components of students attitudes: Cognitive Competence, Value, Affect and Difficulty. The newer version, SATS-36, contains 36 items that assess six components: the original four plus Effort and Interest (SAT, 2017).

There are many studies with reported results from the SATS-36 from various populations of students. In a study of approximately 2200 undergraduate students from many institutions across the United States, Schau and Emmioglu (2012) utilized the SATS-36 instrument to measure the students attitudes towards statistics. They found that, on average, students entered the courses with neutral Affect and Difficulty scores, positive Cognitive Competence, Value and Interest and very positive Effort attitudes. By the end of the semester they found that the attitudes stayed about the same in most categories, but decreased in Value, Interest and Effort (Schau and Emmioğlu, 2012). In a study of 47 students from a small liberal arts college, Bond et al. (2012) had students complete the SATS-36 along side a short perception of statistics survey at the beginning and end of the semester. They also observed a decrease in students attitudes over the course of the semester (Bond et al., 2012).

In a comparative review of these surveys, Nolan et al. (2012) explored the validity and reliability of the tools with published evidence of these measures (Nolan et al., 2012). From their summary, the SATS-36 scores appeared to have the strongest construct validity based on unparceled CFA and internal consistency ratings based on Cronbach's Alpha, assuming the construct validity evidence for the SATS-28 can be applied (Nolan et al., 2012). Several studies have documented the solid psychometric properties, including confirming the four factor structure of the SATS-28, including Dauphinee et al. (1997) and Hilton et al. (2004), however only two could be found for the SATS-36 (Dauphinee et al., 1997) (Hilton et al., 2004). The six factor structure was confirmed in studies by Vanhoof (2011) and Coetzee and Merwe (2010) (Vanhoof, 2011) (Coetzee and Merwe, 2010). Several authors have explored the relationship between students attitudes and course performance including Sorge and Schau (2002), Miller and Schau (2010) and Emmioglu (2011) (Sorge and Schau, 2002) (Millar and Schau, 2010) (Emmioğlu, 2011). Other researchers have used the SATS instruments to explore the differences in teaching environments and methods including Gundlach et al. (2015), DeVaney (2010), Carnell (2008) and Carlson and Winquist (2011) (Gundlach et al., 2015) (DeVaney, 2010) (Carnell, 2008) (Carlson and Winquist, 2011). Several studies also looked to explore the attitudes of students from different fields of study, such as Hannigan et al. (2013, 2014) and Mathew and Aktan (2014) (Hannigan et al., 2013) (Hannigan et al., 2014) (Mathew and Aktan, 2014).

This study utilized the SATS-36 to measure students' attitudes at the beginning and end of the course. The relationship between the attitude components and course performance is evaluated, as well as the use of regression trees to identify groups of students with a similar change in attitudes. The rest of the paper continues as follows: Section 1.2 describes the methods for data collection, design of experiment and data analysis. Next, the results are presented in Section 1.3. Finally, the main findings, limitations and practical recommendations are discussed in Section 1.4.

1.2 METHODS

1.2.1 Data Description

The data were collected for this work during the spring 2016 semester at the University of Rhode Island in an undergraduate introductory biostatistics course. This course had a total enrollment of 171 students and two professors. There were six recitation sections and three teaching assistants for the students included in the analysis, the distribution of students between sections is in Table 1.1.

Table 1.1. Distribution of Students by Professor, Teaching Assistant and Recitation Day

	Professor		**Teaching Assistant**			**Recitation Day**		
	1	**2**	**1**	**2**	**3**	**Monday**	**Tuesday**	**Wednesday**
n	71	66	45	43	49	54	38	45
%	51.8	48.2	32.8	31.4	35.8	39.4	27.7	32.8

This course covers many topics typical to most introductory statistics courses while using medical or health related examples, as seen in Table 1.2. The course also utilizes the statistical computing software SAS Studio during several recitations and homework assignments.

Table 1.2. Weekly Schedule of Topics in Introductory Biostatistics

Week	**Topic**
Week 1-2	Definitions, Population vs. Sample, Types of Variables.
Week 3	Descriptive Statistics and Graphical Data Summaries. Basic Probability.
Week 4	Combinations and Permutations. Random Variable. Binomial Distribution.
Week 5	Normal Distribution. Empirical Rule. Normal Approximation to Binomial.
Week 6	Sampling Distribution. Central Limit Theorem.
Week 7	Statistical Inference. Estimating Population Mean. Confidence Intervals. **Midterm 1.**
Week 8	One-sample Hypothesis Test for Population Mean. Sample size calculation.
Week 9	Two Independent Sample Inferences for Difference in Population Means. Paired Test.
Week 10	One Sample Tests for Population Proportion. **Midterm 2.**
Week 11	Difference in Population Proportion. Chi-Square Tests.
Week 12	Introduction to ANOVA.
Week 13	Introduction to Correlation and Regression.
Week 14	Final Review.

Of the 171 students enrolled in the course, 146 students signed the Institutional Review Board (IRB) consent form to allow their data included in the survey. An additional nine students' data were removed due to too small of a sample size for one teaching assistant. Of these 137 students eligible to be included in the analysis, only 114 students completed the attitude surveys at both the beginning and end of the semester. All available data for the 137 students were included in the analysis of course performance throughout the semester, however only 114 students' data were included in the analysis of the attitude data. These students included 29 male and 85 female students, the majority of students were 19 years old and most students are from the College of Pharmacy.

1.2.2 Design of Experiment

This experiment implemented several components throughout the semester, including two attitude surveys, one introductory survey and several graded assignments. The attitude surveys, the SATS-36, were implemented during the first and last homework assignment of the semester. This survey consists of a pretest and a posttest survey designed for students to answer 36 likert-like questions based on their attitudes at the beginning and end of the semester, with several additional questions to determine other characteristics of the students, such as age, mathematical background and other demographics. The survey is constructed to measure six attitude components- Affect, Cognitive Competence, Value, Difficulty, Interest and Effort as described in Table 1.3 (SAT, 2017).

Table 1.3. Description of SATS-36 Attitude Components

Attitude Component	**#**	**Description**	**Example Question**
Affect	6	Students feelings concerning statistics	"I will like statistics"
Cognitive Competence	6	Students attitudes about their intellectual knowledge and skills when applied to statistics	"I will understand statistics equations"
Value	9	Students attitudes about the usefulness, relevance, and worth of statistics in personal and professional life	"I use statistics in my everyday life"
Difficulty	7	Students attitudes about the difficulty of statistics as a subject	"Statistics formulas are easy to understand"
Interest	4	Students level of individual interest in statistics	"I am interested in learning statistics"
Effort	4	Amount of work the student expends to learn statistics	"I plan to work hard in my statistics course"

Along with the pretest SATS-36 survey, the students were asked to complete an introductory survey created for this study. This survey included questions about students' study habits, learning preferences and extracurricular activities.

The study habits inquired about students' use of office hours, the library and preferences for working in groups. The learning preferences inquired about students' feelings towards different types of assignments, such as SAS coding, homework, presentations and exams. The extracurricular activities inquired about students' physical activity levels, stress reduction techniques and hobbies.

This course had several assessment methods, including 11 weekly homework assignments, nine weekly quizzes, three exams and recitation attendance. The homework assignments were collected weekly, after the second week which suffered from several snow days. The first two exams were during the regular semester 50 minute classes and the third exam was completed during the three hour final exam period. Attendance was recorded during all recitation sessions throughout the semester.

Table 1.4. Summary Statistics for Each Grade Component

Grade Item	**Mean**	**Median**	**Standard Deviation**
Homework	95.94	98.88	15.29
Quiz	67.27	66.67	18.57
Exam 1	90.54	96.11	15.56
Exam 2	88.85	89.00	10.76
Exam 3	86.43	89.25	11.20

The quizzes were implemented in recitation sections using the Socrative online quiz environment, starting the third week of the semester (Soc, 2017). Each quiz consisted of approximately six multiple choice questions relating to the prior week's lecture material. There were three different grading schemes possible for the quizzes: graded personal (GP), graded competition (GC) and non-graded (NG). The graded personal quizzes were individually graded based on the students' performance. The non-graded quizzes were used strictly for student feedback and attendance and the graded competition quizzes were graded with a bonus for the team who completed all questions correctly first. Each recitation section had each quiz scheme for three consecutive weeks with the order based on the recitation day. The Monday recitations had a NG-GC-GP rotation, whereas the Tuesday recitations had a GC-GP-NG rotation and the Wednesday recitations had GP-NG-GC. This rotation allowed each professors' section to have one of each rotation order and every TA to have two different rotations.

Table 1.5. Summary Statistics for Each Grading Scheme of the Weekly Quizzes, out of 100%

Grading Scheme	Mean	Median	Standard Deviation
Graded Competition	55.28	50.00	27.34
Graded Personal	56.70	66.60	26.79
Not Graded	53.94	50.00	26.96

1.2.3 Data Analysis Tools

Longitudinal Models

Mean response profiles were used to graphically and analytically display patterns of change in the mean quiz grades over time for each grading structure. This method is primarily used to address the null hypothesis of no group by time interaction effect, represented graphically by parallel response profiles between groups. The null hypotheses of no time effect and no group effect can also be graphically shown by flat or overlapping lines respectively. This method can be utilized due to the balanced design of the study, with the timing of the repeated measures common to all subjects.

To model the students' quiz performance over time, piecewise quadratic generalized linear mixed models were utilized. The quiz grades were recorded as a count of correct responses out of six questions. This count variable can be modeled with a mixed effects log-linear regression model with a random intercept for each student. To incorporate the hierarchical structure of the course where groups of students are in the same professors' section and in smaller classes with teaching assistants, random effects for these grouping structures were modeled. The full hierarchical model to represent the quiz grade in terms of the grading scheme and time is:

$$\begin{aligned} logE(Y_{ij}|b_i) &= \beta_0 + \beta_1 Time_{ij} + \beta_2 Time6_{ij} + \beta_3 Time9_{ij} + \beta_4 GP_i + \beta_5 GC_i \\ &\quad + \beta_6 Time^2_{ij} + \beta_7 Time6^2_{ij} + \beta_8 Time9^2_{ij} + \beta_9 GP_i * Time_{ij} \\ &\quad + \beta_{10} GC_i * Time_{ij} + \beta_{11} GP_i * Time6_{ij} + \beta_{12} GC_i * Time6_{ij} \\ &\quad + \beta_{13} GP_i * Time9_{ij} + \beta_{14} GC_i * Time9_{ij} \end{aligned} \tag{1.1}$$

where Y_{ij} is the number of quiz questions answered correctly for individual i at time j. The variables GP_i and GC_i refer to the quiz types graded personal and graded competition, with a reference of not graded. The variables $Time$, $Time6$ and $Time9$ refer to the piecewise time variables cut before each of the first two exams at weeks seven and ten. The quadratic terms allow for the quiz grades to change in a non-linear trend between exams. There is a random intercept for each student and a random effect for professor and teaching assistant. Given b_i, it is assumed that the Y_{ij} are independent and have a Poisson distribution, with $Var(Y_{ij}|b_i) = E(Y_{ij}|b_i)$. The random intercepts are assumed to have a bivariate normal distribution, with a mean of zero and a 2x2 covariance matrix G (Fitzmaurice et al., 2012).

Linear Models

Correlation analysis was utilized to explore the relationship between the attitude scores and each of the grade book items. Pearson correlations were calculated between each pretest and posttest component score and the quiz, exam, homework and final grades.

Next, hierarchical linear regression was utilized to model the relationship between the final course grade and starting attitude components. Once again, the hierarchical structure of the course needed to be modeled using random effects to account for the dependence between students in similar professors' and teaching assistants' sections. The full hierarchical linear model is:

$$Y_{ijk} = \beta_0 + \beta_1 Attitude_{ijk} + b_k + b_{jk} + \epsilon_{ijk} \tag{1.2}$$

where Y_ijk is the final grade for the i^{th} student from the j^{th} recitation nested in the k^{th} professor's section. The final grade is predicted by each of the pretest attitude components. The term b_k represents the random effect for professor and b_{jk} is the random effect for the recitation section, resulting in a three-level model. The error term is assumed to follow a normal distribution with a mean of zero and constant variance.

Regression Trees

Regression trees (also called decision trees) are a nonparametric method for segmenting the feature space based on a set of covariates. The algorithm to build the regression tree partitions the feature space to minimize least squares criterion and continues to create splits until the error can no longer be reduced. The resulting nodes are the means for each partition. The tree must then be pruned to avoid over fitting the data and reduce the variance of the final model.

Regression trees were used to model:

$$f(X) = \sum_{m=1}^{M} c_m * I(x \in R_m) \tag{1.3}$$

where $f(X)$ represents the change in each attitude component, c_m is the constant mean change in attitude for each partition of final grade and starting attitude and $I(x \in R_m)$ is an indicator which equals one if the student is in partition m and zero otherwise. This method does not require the assumption of a linear relationship and has easy to interpret results (Henderson and Parmeter, 2015).

1.3 RESULTS

1.3.1 Quiz Performance

The longitudinal analysis of the quiz grading schemes began with a visual representation of the mean response profiles over time using the software SAS Studio 3.6 Enterprise Edition (SAS Institute Inc.,). The mean quiz, homework and exam grades are plotted in Figure 1.1. The homework grades are plotted against recitation day and time and the quiz grades are plotted against grading scheme and time. The figure displays a clear effect of time for both the homework and quizzes, as evidenced by the slope in the lines. The quiz grading scheme effect has several lines overlapping at between certain intervals which indicates support for the null hypothesis of no group effect. The group by time effect also suggests to support the null hypothesis of no effect represented by the parallel slopes between many intervals.

The plot for the weekly homework suggests a significant effect for recitation day as the Monday recitation section had a higher mean for all weeks, except for a tie at week ten. The group by time effect for the quizzes is also not consistent throughout the weeks, as the trend over time appears similar between the groups. The homework plot also indicates a change in performance over time, indicated by the differing slopes.

Also of note is the effect of the exams, which occurred during weeks seven and ten of the semester. Leading up to the exams in weeks six and nine, the quiz grades appear to have local maximums. The week of each exam, student quiz grades drop noticeably, especially in week seven where the quiz related to the new topic of statistical inference and confidence intervals. The homework grades show a similar

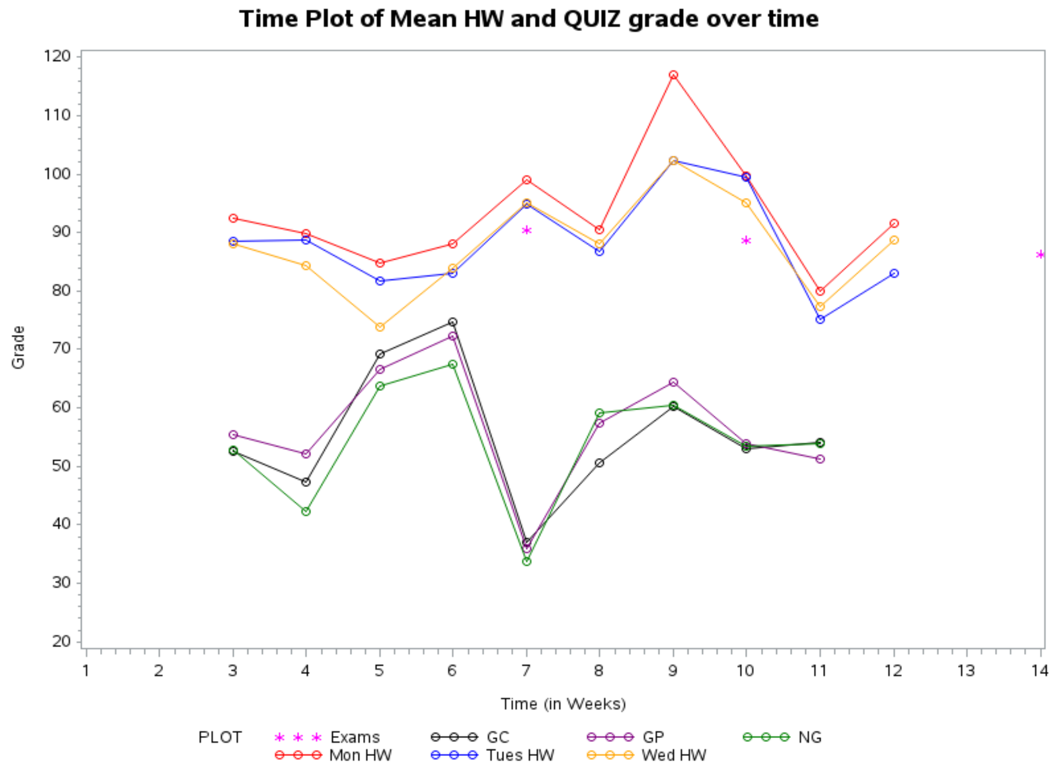

Figure 1.1. Time plot of mean homework, exam and quiz grades, plotted out of 100%. The homework grades are plotted by recitation day and quiz grades are plotted by grading scheme. An increase and subsequent decrease in quiz performance is observed leading up to each exam, while homework grades follow this trend, one week delayed.

drop following the exams with the new material, however the spike in homework grades appear during the week of the first exam and the two weeks leading up to the second. The ninth homework allowed for extra credit, which explains the higher peak for the Monday sections.

The hierarchical log-linear regression model was also run in SAS Studio 3.6 Enterprise Edition using the GLIMMIX procedure (SAS Institute Inc.,). The model was fit using maximum likelihood and approximated using adaptive Gaussian quadrature with ten quadrature points for each random effect during the evaluation of the integrals for the marginal likelihood (Fitzmaurice et al., 2012). The interaction terms were found to be non-significant and were removed from the model. The nested random effects for professor and teaching assistant were also found to be non-significant, based on covariance estimates not significantly different than zero, indicating no effect of clustering within the data. The random intercepts were non-zero and included in the model. The final model resulted in estimates for fixed effects in Table 1.6.

Table 1.6. Results of the Longitudinal Model for Weekly Quiz Grades

Fixed Effect	**Estimate**	**Std Err**	**95% C.I.**
Intercept	**7.14**	**0.17**	**(6.81, 7.46)***
Time	**-2.60**	**0.03**	**(-2.66, -2.54)***
Time6	**-0.99**	**0.09**	**(-1.16, -0.81)***
Time9	0.08	0.13	(-0.17, 0.33)
GP	0.05	0.04	(-0.03, 0.13)
GC	0.03	0.03	(-0.05, 0.11)
$Time^2$	**0.27**	**0**	**(0.27, 0.27)***
$Time6^2$	**-0.21**	**0.02**	**(-0.26, -0.17)***
$Time9^2$	**-0.12**	**0.03**	**(-0.17, -0.06)***

All effects for time and quadratic time were significant, except Time9. The fixed effects for the quiz types were not significant, nor were the interaction effects between quiz types and time. The over dispersion parameter indicated a good fit to the conditional distribution of the model, based on the Pearson Chi-Square value of 0.56 and the assumption of conditional normality of residuals was also met. These results support the mean response profiles graphical findings. There is a trend in quiz performance over time, influenced by the timing of the exams, however the quiz grading structure had no significant effect on student performance each week.

1.3.2 Survey of Attitudes

Before analyzing the results of the SATS-36 attitude survey, the internal consistency of the attitude components had to be investigated to explore the extent to which each parcel was actually measuring the same construct. Cronbach's coefficient alpha is one of the most frequently reported measures of internal consistency, however the assumptions can be difficult to meet and an alternative method, Omega, is available that assumes fewer assumptions and holds fewer restrictions on the dataset (Tavakol and Dennick, 2011). The pretest and posttest Omega values for all attitude components are within the acceptable range (above 0.70), except pretest difficulty measured at 0.64 and can be found in Table 1.7 showing both the point estimate and bootstrap confidence interval using the MBESS package in R Studio (R Core Team, 2017; Kelley, 2017).

Table 1.7. Omega Values for Each Attitude Component

Attitude Component	**Test**	**Omega**	**Bootstrap 95% C.I.**
Affect	Pretest	0.77	(0.67, 0.84)
Cognitive Competence	Pretest	0.82	(0.74, 0.87)
Value	Pretest	0.89	(0.85, 0.92)
Difficulty	Pretest	0.64	(0.51, 0.74)
Interest	Pretest	0.91	(0.87, 0.94)
Effort	Pretest	0.9	(0.81, 0.95)
Affect	Posttest	0.86	(0.80, 0.89)
Cognitive Competence	Posttest	0.87	(0.82, 0.91)
Value	Posttest	0.92	(0.9, 0.94)
Difficulty	Posttest	0.81	(0.75, 0.86)
Interest	Posttest	0.93	(0.9, 0.95)
Effort	Posttest	0.81	(0.72, 0.88)

1.3.3 Performance and Attitudes

The correlation analysis results between the attitude components and each gradebook item are presented in Table 1.8. Results indicate that the pretest attitudes are most highly correlated to the quiz, final exam and final grades, albeit none of the correlations are very strong. The posttest attitude scores show a stronger correlation to the gradebook items, especially in the Affect, Cognitive

Competence and Difficulty components. The Effort component shows the weakest correlation in both the pretest and postest values. The homework grades appear to have the lowest correlation to the attitude components, compared to the other gradebook items.

Table 1.8. Pearson Correlations for Each Attitude Component and the Gradebook Items

Attitude Component	**Test**	**Quiz**	**Homework**	**Exam 1**	**Exam 2**	**Exam 3**	**Final Grade**
Affect	Pretest	0.24	0.14	0.16	0.14	0.31	0.23
Cognitive Competence	Pretest	0.23	0.14	0.14	0.15	0.26	0.22
Value	Pretest	0.12	-0.08	-0.04	0.03	0.02	-0.01
Difficulty	Pretest	0.13	0.15	0.13	0.02	0.13	0.14
Interest	Pretest	0.13	-0.03	0.004	0.11	0.12	0.06
Effort	Pretest	0.04	-0.07	-0.12	0.07	0.03	-0.02
Affect	Posttest	0.51	0.37	0.51	0.57	0.59	0.62
Cognitive Competence	Posttest	0.5	0.34	0.50	0.53	0.56	0.59
Value	Posttest	0.39	0.12	0.25	0.29	0.29	0.30
Difficulty	Posttest	0.47	0.35	0.51	0.36	0.47	0.52
Interest	Posttest	0.34	0.17	0.21	0.31	0.29	0.30
Effort	Posttest	0.22	0.11	-0.04	0.25	0.15	0.14

The hierarchical linear models for modeling the final grade using the starting attitude component values were fit using maximum likelihood in the lme4 package in R Studio (R Core Team, 2017; Bates et al., 2015). Six models were fit, one for each starting attitude component. After fitting the model, the model assumptions were checked and the error terms were shown to be non-normal. This effects the validity of the inference on the fixed effects, so bootstrapped confidence intervals were produced instead by refitting the model to 2000 re-sampled datasets. The estimates and bootstrapped confidence intervals for the fixed effects are in Table 1.9 (Pek et al., 2017).

The bootstrapped confidence intervals for each pretest attitude components' model show whether this term is significant in predicting the final grade in the course. Only pretest Affect and Cognitive competence did not include zero in the confidence interval, and were both positive. This is consistent with the parametric confidence interval, as well as the Pearson correlation estimates. These findings

Table 1.9. Results of the Hierarchical Linear Models for Final Grade by Pretest Attitude

Fixed Effect	Estimate	Bootstrap 95% C.I.
Intercept	81.76	(73.38,88.92)
Pretest Affect	**2.10**	**(0.45,3.91)***
Intercept	81.12	(71.5,89.28)
Pretest Cog. Comp.	**1.93**	**(0.32,3.78)***
Intercept	91.30	(83.25,98.58)
Pretest Value	-0.11	(-1.82,1.61)
Intercept	83.99	(74.19,92.26)
Pretest Difficulty	1.78	(-0.32,4.11)
Intercept	88.95	(83.90,93.34)
Pretest Interest	0.44	(-0.73,1.58)
Intercept	91.84	(83.23,99.09)
Pretest Effort	-0.17	(-1.33,1.16)

indicate that students' with more positive feelings towards statistics and confidence in their own computational abilities at the start of the semester performed better in the course overall.

1.3.4 Change in Attitudes

After analyzing the quiz performance throughout the semester and the relationship between the course performance and students' attitudes, the next analysis of interest in this study is the change in attitudes throughout the semester. To begin, the summary statistics for each attitude during the pretest and posttest are in Table 1.10. The decrease in the attitudes throughout the semester is similar to other studies using the SATS-36 survey and has been hypothesized to be caused by an increase in students' understanding of what statistics is and the details involved in the subject throughout the semester (Schau and Emmioğlu, 2012) (Bond et al., 2012) .

Table 1.10. Summary Statistics for Pretest and Posttest Attitude Components

Attitude Component	**Pretest Mean**	**Pretest S.D.**	**Posttest Mean**	**Posttest S.D.**
Affect	4.30	1.01	4.09	1.36
Cognitive Competence	5.00	1.02	4.88	1.36
Value	4.77	1.06	4.49	1.29
Difficulty	3.82	0.70	3.53	1.01
Interest	4.21	1.28	3.66	1.52
Effort	6.29	1.14	6.09	1.08

To begin to the explore the students' change in attitude throughout the semester, regression trees were the chosen technique due to the resulting partitions and ease of interpretation of the model. This model was used in more of an exploratory nature to identify groups of students with similar changes in attitude within the course, in relation to their final grade and starting attitude. The regression trees were built in R Studio using the packages rpart and partykit (R Core Team, 2017; Therneau et al., 2015; Hothorn and Zeileis, 2015). The full trees were first built, then pruned to reduce the variance of the model.

The resulting trees are shown in Figure 1.2. The final course grade was the most impactful first split for all trees, except effort. The effort tree shows only one split at a prior attitude score of 4.375. Students with a low effort showed an increase, on average, by the end of the course and students with a high effort score showed little change, on average. The other trees are a bit more complicated to interpret. For all other trees, the left most node displays the change in attitude for the students with lower final grades, and shows that students that did poorly in the course had a lower affect, cognitive competence, difficulty, interest and value than at the start of the course.

The partitions that showed the greatest average increase in each attitude component contained students that performed well in the course (at least a B average) and were in the lower partition for starting attitude. The other partitions contain groupings of students with higher starting attitudes that lowered or students who did okay in the course and showed very little change in their attitude. Each tree also shows that the largest partitions result in a group with little change in attitude throughout the semester. Many students whose attitude were higher than their peers at the start of the course showed a decrease throughout the semester. There are many factors that could influence students' attitudes towards statistics during the semester, including the chosen covariates of starting attitude and course performance.

1.4 MAIN FINDINGS

1.4.1 Limitations

As with any observational or survey experimental design, there are several limitations to consider. First, due to the nature of the data and the use of human subjects, Institutional Review Board (IRB) consent was needed from the students. This process reduced the potential sample size from 171 to 146 students. The students who did not consent to be in the study were not eligible to be included in any analysis and may have differing characteristics than those who did consent. Another reduction to the sample size was the removal of nine students form a small Thursday recitation that had both a different day of the week and teaching assistant than the other students. This sample size was too small to be modeled for the day of week or TA effects.

Beyond these sample restrictions, another limitation was missing data in the form of non-response from students on one or both of the attitude surveys. There were 23 students that did not complete both surveys and were removed from the analysis of attitudes. These 23 students' information was included, whenever available, in the course performance analysis.

These reductions to the sample used for the analysis limits the amount of data available and leads to potential biases. The non-response bias, from both non-consent and missing surveys, leads to a potential missing subset of the class. The students who were willing to not complete a survey required for their course homework could potentially share similar attitudes about the course that are now not available in this analysis. The students who did not want to consent could potentially share feelings of discontent with the course or their course performance. Students who did consent and complete the surveys could also potentially share more positive feelings towards the subject and have better course outcomes.

Another potential bias with any survey is response bias. There is no way of knowing if students were completely truthful in their responses to the attitude surveys. The survey was designed to have positively and negatively worded questions to help reduce the tendency to answer the same way to every question which helps to check if students were paying attention to the survey. However, the possibility that students were answering how they felt the professor or their peers would want cannot be measured, but must be considered. Students may feel they should answer more positively if they wanted to align with their professor's wants or views. Conversely, they may have answered more negatively than their true feelings to conform with other students. Similarly, there's always the chance that students were not taking the survey seriously and did not answer truthfully due to their desire to quickly complete the survey. There are many possible factors that could lead to different results in the survey responses and the nature of survey data needs to be taken into consideration when drawing conclusions from the results.

1.4.2 Practical Recommendations

From the analysis of student performance throughout the semester, the main conclusion was that time effects students performance, specifically the timing of exams and new concepts effects students' homework and quiz performance significantly. The lowest performance was when the new topic of inference was presented after the first exam for both homework and quizzes. The best quiz performance was the week before the exam and the best homework performances were the week of the exams. These conclusions are important for practical implementation and planning of course materials. Taking into consideration when students are devoting the most time to the course and conversely the least amount of time is important for students' understanding of topics. If a certain topic is only presented immediately after the exam, there is the possibility that students will not fully understand or study this topic at all before the next exam because they have less time and focus devoted to this course immediately after an exam. If possible, taking at least one class to review the exam results and problem areas and then switching to new material could help students with this transition. Similarly, including the newest material as a review on the second homework after the exam could help students recall and focus on the new topics, rather than having only new material on each assignment.

The next main conclusion is the relationship between course performance and starting attitudes. Students who started the course with more positive feelings towards the subject of statistics, as well as higher confidence in their statistical abilities performed better in the course. This shows that students' confidence levels early in the semester have a strong impact on their performance throughout the semester. The converse is also true; students who had more negative feelings towards the subject, as well as less confidence in their computational abilities performed worse in the course. These students are an important subset of the class and identifying them early in the semester could be an integral step towards increasing the success of the course, both through the performance and attitude metrics.

Finally, the change in attitudes throughout the semester showed an interesting trend depending on the student's final course grade. In general, students who performed poorly in the course showed a decrease in each attitude component. Students who performed decently in the course, but started with lower attitudes showed an increase in the attitude component. Students who started with high values for the attitude component also showed a decrease. This shows that the change in attitudes is related to student performance and the starting attitude. This relationship is especially important for students who did poorly in the course. On average, these students showed a decrease in every attitude component and left the course with low interest and value in statistics. This could prevent students from enrolling in consequent statistics courses and if they have to retake the course, having much less interest and confidence the second time around. Beyond statistics courses, these feelings could emanate into their professional experience with statistics as well. The majority of students in this course were Pharmacy majors, a field that will undoubtedly have to work with statistical measures in their profession. Having a positive outlook and trust in the subject is important, even beyond the class.

Overall, this study shows that students' attitudes are an important measure in relation to student outcomes and motivation in a course. Not only do students who perform poorly in the course show low starting attitudes in certain components, they also leave with a negative change throughout the semester. The question now is, how can we increase students attitudes towards statistics? How can we help students who are struggling still see value in learning the subject?

To attempt to get closer to these answers, a second round of data collection has been collected in the Spring 2017 semester in all undergraduate courses at the University of Rhode Island. This new dataset includes a broader set of students, a more detailed collection of exit learning preferences and a more diverse set of professors. This new data will hopefully help us get a closer look into the learning preferences and teaching styles that effect the change in attitudes throughout the semester, as well as how different subsets of students view statistics.

List of References

(2017). Cs consultants, llc. http://www.evaluationandstatistics.com/. Accessed: 2017-09-10.

(2017). Socrative by mastery connect. https://www.socrative.com/index.html. Accessed: 2017-09-10.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Bond, M. E., Perkins, S. N., and Ramirez, C. (2012). Students'perceptions of statistics: An exploration of attitudes, conceptualizations, and content knowledge of statistics. *Statistics Education Research Journal*, 11(2).

Carlson, K. A. and Winquist, J. R. (2011). Evaluating an active learning approach to teaching introductory statistics: A classroom workbook approach. *Journal of Statistics Education*, 19(1):1–23.

Carnell, L. J. (2008). The effect of a student-designed data collection project on attitudes toward statistics. *Journal of Statistics Education*, 16(1):1–15.

Coetzee, S. and Merwe, P. v. d. (2010). Industrial psychology students' attitudes towards statistics. *SA Journal of Industrial Psychology*, 36(1):1–8.

Dauphinee, T. L., Schau, C., and Stevens, J. J. (1997). Survey of attitudes toward statistics: Factor structure and factorial invariance for women and men. *Structural Equation Modeling: a multidisciplinary journal*, 4(2):129–141.

DeVaney, T. A. (2010). Anxiety and attitude of graduate students in on-campus vs. online statistics courses. *Journal of Statistics Education*, 18(1):1–15.

Dunham, B. (2009). Statistics clicks: Using clickers in introductory statistics courses. *University of British Columbia.*

Emmioğlu, E. (2011). A structural equation model examining the relationships among mathematics achievement, attitudes toward statistics, and statistics outcomes. *Unpublished doctoral dissertation). Middle East Technical University, Ankara, Turkey.*

Fitzmaurice, G. M., Laird, N. M., and Ware, J. H. (2012). *Applied longitudinal analysis*, volume 998. John Wiley & Sons.

Gundlach, E., Richards, K. A. R., Nelson, D., and Levesque-Bristol, C. (2015). A comparison of student attitudes, statistical reasoning, performance, and perceptions for web-augmented traditional, fully online, and flipped sections of a statistical literacy class. *Journal of Statistics Education*, 23(1):1–33.

Hannigan, A., Gill, O., and Leavy, A. M. (2013). An investigation of prospective secondary mathematics teachers conceptual knowledge of and attitudes towards statistics. *Journal of Mathematics Teacher Education*, 16(6):427–449.

Hannigan, A., Hegarty, A. C., and McGrath, D. (2014). Attitudes towards statistics of graduate entry medical students: the role of prior learning experiences. *BMC medical education*, 14(1):70.

Henderson, D. J. and Parmeter, C. F. (2015). *Applied nonparametric econometrics*. Cambridge University Press.

Hilton, S. C., Schau, C., and Olsen, J. A. (2004). Survey of attitudes toward statistics: Factor structure invariance by gender and by administration time. *Structural Equation Modeling*, 11(1):92–109.

Hothorn, T. and Zeileis, A. (2015). *partykit: A Modular Toolkit for Recursive Partytioning in R*. R package version 1.1-1.

Kelley, K. (2017). *MBESS: The MBESS R Package*. R package version 4.4.1.

Mathew, L. and Aktan, N. M. (2014). Nursing student attitudes toward statistics. *Journal of Nursing Education*, 53(4):233–237.

McGowan, H. M. and Gunderson, B. K. (2010). A randomized experiment exploring how certain features of clicker use effect undergraduate students' engagement and learning in statistics. *Technology Innovations in Statistics Education*, 4(1).

Millar, A. M. and Schau, C. (2010). Assessing students attitudes: the good, the bad, and the ugly.

Nolan, M. M., Beran, T., and Hecker, K. G. (2012). Surveys assessing students'attitudes toward statistics: A systematic review of validity and reliability. *Statistics Education Research Journal*, 11(2).

Pek, J., Wong, O., and Wong, C. (2017). Data transformations for inference with linear regression: Clarifications and recommendations. *Practical Assessment, Research & Evaluation*, 22.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

SAS Institute Inc. Sas studio.

Schau, C. and Emmioğlu, E. (2012). Do introductory statistics courses in the united states improve students'attitudes? *Statistics Education Research Journal*, 11(2).

Shatto, B. and Erwin, K. (2016). Moving on from millennials: Preparing for generation z. *The Journal of Continuing Education in Nursing*, 47(6):253–254.

Sorge, C. and Schau, C. (2002). Impact of engineering students attitudes on achievement in statistics: A structural model.

Tavakol, M. and Dennick, R. (2011). Making sense of cronbach's alpha. *International journal of medical education*, 2:53.

Therneau, T., Atkinson, B., and Ripley, B. (2015). *rpart: Recursive Partitioning and Regression Trees*. R package version 4.1-10.

Vanhoof, S. (2011). Statistics attitudes in university students: Structure, stability and relationship with achievement.

Vaterlaus, J. M., Beckert, T. E., Fauth, E. B., and Teemant, B. (2012). An examination of the influence of clicker technology on college student involvement and recall. *International Journal of Teaching and Learning in Higher Education*, 24(3):293–300.

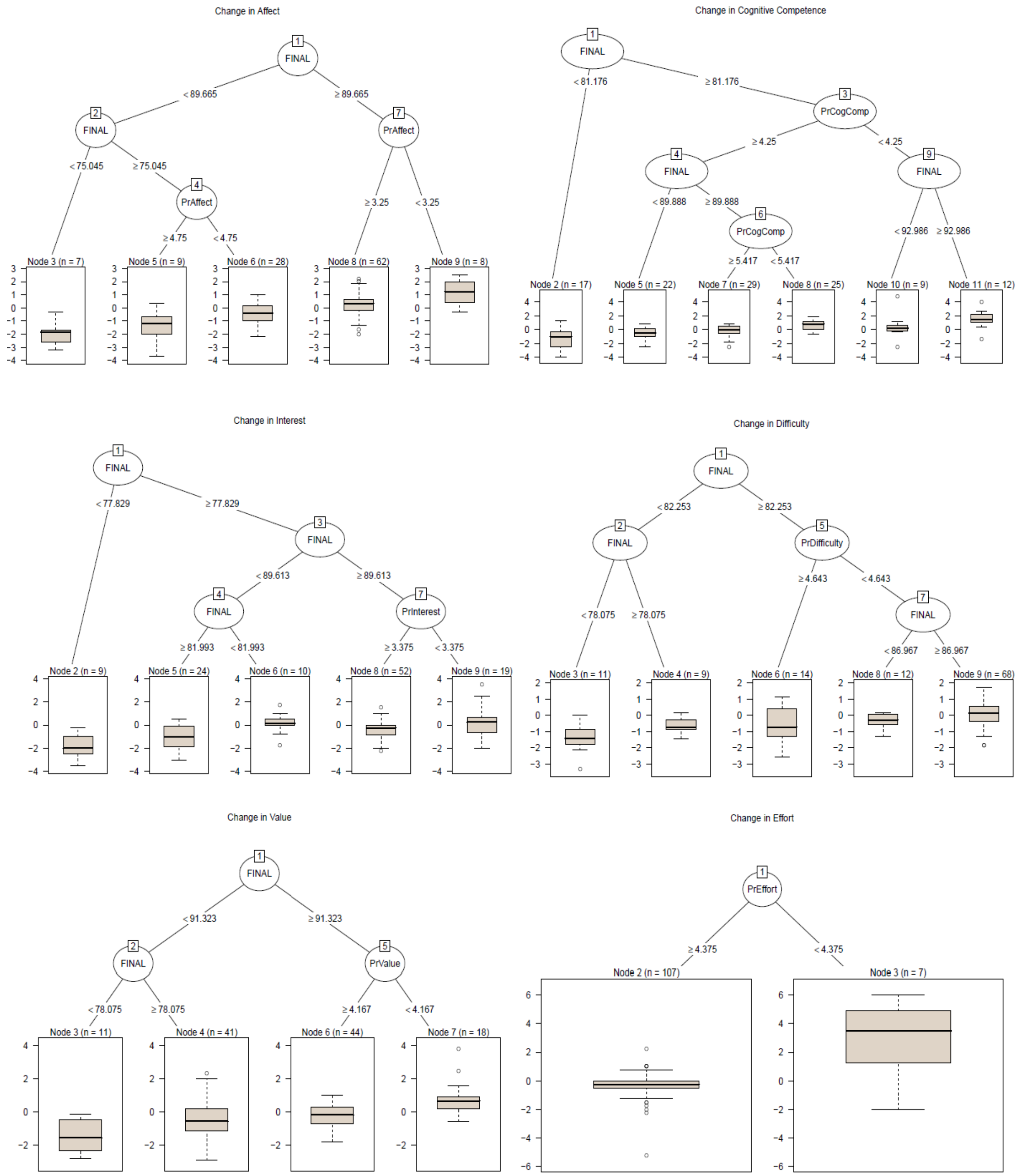

Figure 1.2. Regression trees for the change in each attitude component dependent on the pretest attitude and final course grade. Partitions show groups of students ranging from negative to neutral to positive changes in attitudes. Students who performed poorly in the course tended to leave with lower attitudes, while students who did well and started with lower attitudes than their peers left with a more positive outlook on statistics. Students who did well and started with higher attitudes, typically left the course with the same attitude towards statistics.

MANUSCRIPT 2

A Multivariate Analysis of Generation Z Students

This manuscript has been prepared with the intent to submit to the Journal of Statistics Education. The paper was co-authored by Dr. Natallia Katenka.

Abstract

Undergraduate classrooms have quietly been filled with a new generation of students. No longer are current undergraduate students from the Millenial Generation, but rather they are from Generation Z. So, what do we know about this generation of students? In response to the changing landscape of learners and to gather more information about the learning styles and attitudes of Generation Z students, we implemented a survey design within all undergraduate statistics courses at the University of Rhode Island. Students were asked to complete introductory and pretest attitudes towards statistics surveys at the beginning of the course and exit and posttest attitude surveys at the end of the semester. Principal component analysis in connection with regression analysis indicate a relationship between students starting attitudes and their course performance. Cluster analysis indicates a two group structure in starting attitudes of the students in each course, with each cluster showing different achievement and learning preferences. These results lead to interesting practical considerations for instructors to consider how their students' views on the subject can impact their performance, as well as how to implement students' learning preferences into their lesson plans.

2.1 INTRODUCTION

2.1.1 Motivation

In 2013, the first incoming class of Generation Z students walked onto college campuses around the world. This generation, filled with all persons born since approximately 1995, has quietly entered the undergraduate educational landscape, with much less fanfare than the Millennial Generation before them. Now, all undergraduate courses are filled with predominantly Generation Z students. These students were born into a world where the Internet was a reality and growing up any question could be answered with a simple Google search. Most Generation Z students also grew up in a post-9/11 world hearing about various mass shootings all over the country. Information about these events was immediately available through their social media accounts and news websites adding to a sense of global connectivity and spread of information unfamiliar to previous generations (Seemiller and Grace, 2017). While generational characteristics do not perfectly describe all people within the generation, they can help to understand their background, upbringing and how their worldview was shaped.

Understanding the background of a group of students can also help to understand their learning characteristics and educational needs. This shift in generations leads to an interesting dilemma for educators: How can we alter current teaching methods to better suit this new generation of undergraduate students? In response to this shift in generations in undergraduate statistics courses, we implemented a survey study within all introductory statistics courses at the University of Rhode Island to gather more information about the learning preferences, study habits, attitudes towards statistics and interpersonal study habits of this new population of students. The research presented in this paper aim to answer the following questions: (1.) What are Generation Z students' learning preferences and attitudes towards statistics? (2.) Are students' starting attitudes towards statistics related to their course outcomes? and (3.) How do students' starting attitudes affect their study habits within a statistics course?

2.1.2 Background

Several studies have begun to characterize Generation Z students' learning preferences in higher education settings. These students tend to learn better from observation, such as through watching a video or demonstration of how to perform a certain action before attempting it themselves. Many Generation Z students will prefer to watch a YouTube video rather than reference a textbook or written media while learning something new (Shatto and Erwin, 2016). These students also value the applicability and practice of new skills very highly. They desire to apply what they are learning in a variety of settings, with internships being a very important learning opportunity for them (Seemiller and Grace, 2017).

An interesting development from their technology driven upbringing is a desire to work independently to find answers to their questions and work through their assignments. The individual nature of the Internet allows Generation Z students to take entire classes without interacting with peers or instructors, find resources for research papers without traveling to the library and complete many instructional activities without the aid of others (Seemiller and Grace, 2017). This leads to an interesting preference to work independently and utilize those around them as a resource, rather than a requirement.

Generation Z students are typically accustomed to instantaneous answers to their questions and have almost too many sources available to them with a simple Google search. It has been observed that students from this generation may lack the ability to parse through these results and critique their validity. This instant gratification is also important to Generation Z students as they have been found to have a decreased attention span from previous generations and are accustomed to being surrounded by visual imagery, multiple sources and almost too much information at any time (Shatto and Erwin, 2016).

Specifically relating to STEM education, Hora et al. (2017) completed a descriptive study of students study habits in real-world situations. Students in the sample were from biology, physics, earth science and mechanical engineering courses. Their study found that students' studying habits had several stages: cues, timing, resources, setting and method of study. They found that students' most common cue to study was the instructor mentioning an upcoming exam and that the timing of study was split between the sample, with several students studying for days leading to the exam, while others crammed the night before and some studied throughout the semester. The most common resources for studying were found to be the course website, google, the textbook and lecture notes while the least used were the human resources and cue cards. The setting of study seemed to vary depending on the assignment, as many students reported studying in both groups and alone. Finally, the method of study was most commonly a review of the notes and textbook, while the least used was reviewing homework and weekly quizzes (Hora and Oleson, 2017).

This research also involves looking at students attitudes towards statistics as a subject and tool to use in their future fields. In order to measure students attitudes, a latent construct, an appropriate instrument needed to be chosen. There have been several instruments developed and published to measure students attitudes towards statistics. In this study, the SATS-36 (Survey of Attitudes Towards Statistics) was chosen. The SATS-36 was developed by Candace Schau and contains 36 Likert-type scale questions to assess six components (as opposed to the previous version which had four components) of students attitudes: Cognitive Competence, Value, Affect, Difficulty, Effort and Interest (SAT, 2017).

To choose an appropriate attitude survey, several different instruments were considered. In a comparative review of these surveys, Nolan et al. (2012) compared the validity and reliability measures of several tools by collecting published evidence from various studies. The SATS-36 appeared to have the strongest construct validity and internal consistency, as long as the validity of the SATS-28 can be applied as there were no measures available for the SATS-36 (Nolan et al., 2012). The SATS-36 has been used in several studies from various populations of students. In a study of 47 students from a small liberal arts college, Bond et al had students complete the SATS-36 alongside a short perception of statistics survey at the beginning and end of the semester. They also observed a decrease in students attitudes over the course of the semester (Bond et al., 2012). Another study of approximately 2200 undergraduate students from several institutions across the United States, Schau and Emmioglu used the SATS-36 instrument to measure the students attitudes towards statistics. Their study found that, on average, students entered the courses with neutral Affect and Difficulty scores, positive Cognitive Competence, Value and Interest and very positive Effort attitude scores. By the end of the semester, found no change in Affect, Cognitive Competence and Difficulty, but a decrease in Value, Interest and Effort (Schau and Emmioğlu, 2012).

This study utilized the SATS-36 to measure students' attitudes at the beginning and end of the course, along with original introductory and exit surveys about students' learning, teaching and collaboration preferences. Extending on the qualitative analysis of the study habits, this paper also explores the relationship between the attitude components and course performance, as well as between the attitude components and learning preferences. The rest of the paper continues as follows: first, Section 2.2 describes the methods for data collection, design of experiment and data analysis. Next, the results are presented in Section 2.3. Finally, the main findings, limitations and practical recommendations are discussed in Section 2.4.

2.2 METHODS

2.2.1 Design of Experiment

This study includes data that were collected from all four introductory statistics courses at the University of Rhode Island: STA220 Statistics in Modern Society, STA307: Introductory Biostatistics, STA308: Introductory Statistics and STA409: Statistical Methods in Research I. Each course is designed to be a first course in statistics, with students from various levels of mathematical backgrounds and majors. STA220 is a general education course which focuses on descriptive statistics, probability and does not cover inference. STA307 is an introductory biostatistics course covering the typical introductory statistics topics with a focus on health and biological applications. STA308 is a typical introductory statistics course and STA409 is an introductory statistics course designed for students with a stronger mathematics background. Only STA307, STA308 and STA409 are included in this paper, as they are considered prerequisites for further statistical study at the university.

This experiment consisted of two rounds of survey collection within each of the three courses. Students were asked to complete two surveys at both the beginning and the end of the semester. One of the surveys each time was the SATS-36 survey and the other surveys were the introductory and exit surveys. Each course included the surveys for this research as a part of a homework or extra credit assignment, however the choice to have their data included in the study was voluntary and indicated by signing the Institutional Review Board (IRB) consent form.

The pretest and posttest attitude surveys were developed by Candace Schau and chosen based on the review in Section 2.1.2. These surveys have 36 Likert-like scale items designed to measure six attitude components: Affect, Cognitive Competence, Value, Difficulty, Interest and Effort which are described in Table 2.1. There are additional questions on each survey about students' demographics, mathematical background and academic performance. The pretest and posttest surveys differ in their wording tense and in some of the additional questions at the end of the surveys (SAT, 2017).

Table 2.1. Description of SATS-36 Attitude Components

Attitude Component	**#**	**Description**	**Example Question**
Affect	6	Students feelings about taking a statistics course	"I will enjoy taking statistics courses"
Cognitive Competence	6	Students attitudes about their ability to learn statistics	"I will find it difficult to understand statistical concepts"
Value	9	Students attitudes about the usefulness and worth of statistics	"Statistics should be a required part of my professional training"
Difficulty	7	Students attitudes about the difficulty of statistics	"Statistics is a subject quickly learned by most people"
Interest	4	Students level of interest in statistics	"I am interested in understanding statistical information"
Effort	4	Amount of work a student plans to expend on statistics	"I plan to complete all of my statistics assignments"

Along with the pretest attitude survey, students were asked to complete an introductory survey. This survey included questions about students' outside habits such as hobbies, stress and physical exercise. A section of questions about study habits asked students' where they prefer to study, if they prefer to work alone or in groups, if they attend office hours and if they complete all of their homework and practice exams. Each question has an interval scale of choices of Never, Rarely, Sometimes, Often and Always. Another section of questions asks students to rank various learning activities from least to most beneficial (1 to 8) to their learning. The activities include homework, group projects, computer analysis, weekly quizzes, recitation practice problems, note taking/lectures, textbook and exams. The last section of the introductory survey asks students to rank teaching techniques in the same way. The methods here include timely feedback, timely response to email, office hours, in-depth knowledge, good pace, clear explanations and use of real life examples.

When the posttest SATS-36 survey was given to students, they were also asked to complete an exit survey. The exit survey had the same set of questions about study habits, learning activities and teaching techniques as the introductory survey, with the questions worded to pertain to study habits in this statistics course. An additional set of questions about using various resources was included in the exit survey. The resources included the textbook, printable notes, email, practice exams, TA office hours and professor office hours. Each resource had an option of 0-1 times, 2-3 times, Monthly, Weekly and Daily. A final question was added to the exit survey asking students about their collaborators throughout the semester. Each student was asked to report the names of each student they worked with throughout the semester in this class, as well as what they worked on and how they met.

The surveys were the only additions made to each course. At the end of the semester, final course grades were requested from each professor for all of the consenting students in the study.

2.2.2 Data Description

Each course had two different professors and at least one teaching assistant. Students in STA307 and STA308 had lecture three times a week and have one weekly recitation class led by a teaching assistant where practice problems were solved. STA409 students did not have recitation classes, however their lecture classes were smaller than those in STA307 and STA308. Excel was used on several homework assignments for STA308 and SAS was the technology chosen for the STA307 students to practice.

In order to be included in any analysis students needed to sign the IRB consent form and to be included in any analysis beyond the descriptive plots and tables, students' needed to have completed both the pretest and posttest SATS-36 surveys. The total enrollment, consent totals and study participation totals can be found in Table 2.2.

Table 2.2. Enrollment and Participation Totals for Each Course

	STA307	**STA308**	**STA409**
Total Enrollment	153	248	113
Consent Total	128 (83.7%)	170 (68.5%)	64 (56.6%)
Survey Completion Total	110 (71.9%)	76 (30.65%)	59 (52.2%)

Each course had it's own population of students with differing mathematical preparation and majors. STA307 can be characterized by predominantly sophomore students from the College of Pharmacy. STA308 also had predominantly sophomore students and the most common major from students in the sample was Biology, followed by Pre-Med. STA409 had mostly junior and senior students and most students were engineering majors. Most students in STA307 had only taken one previous college mathematics or statistics course, while most STA308 students had taken between one and three classes and STA409 students had taken between 4 and 5 prior mathematics or statistics courses. The survey sample for STA307 and STA308 both show about 70% female students, while STA409 sample has 57% male students.

The summary statistics for each course's pretest and posttest SATS-36 surveys are in Table 2.3. Consistent with other studies, each class showed a decrease in each average attitude component by the end of the course. STA409 had the highest pretest attitudes for each component, except difficulty. STA307 had the lowest posttest averages for all components, except effort. STA308 showed the smallest change in average attitude throughout the semester for all components.

Table 2.3. Summary Statistics for Pretest and Posttest Attitude Components for Each Course

Attitude Component	**Course**	**Pretest Mean (S.D)**	**Posttest Mean (S.D)**
Affect	STA307	4.93 (1.03)	4.04 (1.36)
Cognitive Competence	STA307	5.49 (1.36)	4.93 (1.23)
Value	STA307	5.25 (1.04)	4.58 (1.18)
Difficulty	STA307	3.98 (0.6)	3.70 (0.77)
Interest	STA307	4.77 (0.63)	3.67 (1.57)
Effort	STA307	6.58 (0.63)	6.32 (0.74)
Affect	STA308	4.80 (1.14)	4.77 (1.22)
Cognitive Competence	STA308	5.45 (1.03)	5.30 (1.10)
Value	STA308	4.95 (0.99)	4.85 (1.09)
Difficulty	STA308	4.04 (0.62)	3.96 (0.68)
Interest	STA308	4.26 (1.22)	4.20 (1.31)
Effort	STA308	6.48 (0.65)	6.36 (0.76)
Affect	STA409	5.01 (0.88)	4.71 (1.14)
Cognitive Competence	STA409	5.61 (0.87)	5.29 (0.95)
Value	STA409	5.61 (0.82)	5.47 (0.98)
Difficulty	STA409	4.00 (0.63)	3.71 (0.72)
Interest	STA409	5.02 (1.02)	4.76 (1.31)
Effort	STA409	6.60 (0.53)	6.31 (0.68)

To begin to explore the study habits of students in each class, visualizations and summary statistics were generated. In Figure 2.1, bar plots for each resource surveyed in the exit survey for STA307 are displayed. The most used resources were the online notes and practice exams. The professor and teaching assistant office hours were the least used resources, followed by the textbook. It appears that more students use email to contact their instructors than in person meetings. The plots for STA308 show a similar trend, however more students report using the textbook. STA409 also has an overall similar trend, however more students report using the textbook weekly, more students report emailing their instructor and less use the teaching assistant's office hours.

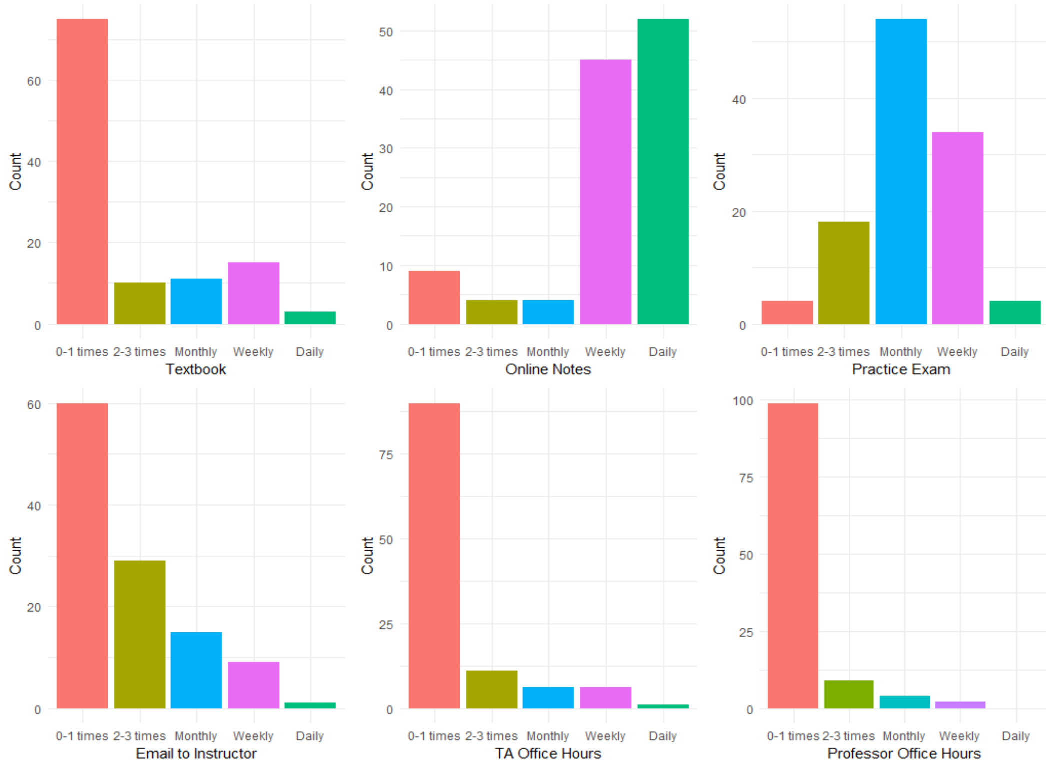

Figure 2.1. Bar plots of each of the resources surveyed on the exit survey in STA307. The most used resources were the online notes and practice exams.

The summary statistics for the rankings (1-8) of various learning activities from the exit survey are in Table 2.4. The survey questions asked students to rank the activities from least (1) to most (8) beneficial to their learning this semester. From the table, STA307 students report recitation problems, exams and lecture notes as the most beneficial activities towards furthering their learning, while reading the textbook and using SAS were the least beneficial. For STA308, lecture notes and recitation problems were the most beneficial, while the exams and Excel were somewhere in the middle and reading the textbook was the least beneficial. For STA409, the course had no recitation sections, nor use of statistical software, so

Table 2.4. Mean and standard deviation for the exit survey rankings of various learning activities. STA409 did not use statistical software, nor did the course have recitation sections.

	STA307	**STA308**	**STA409**
Stat Software	1.76 (1.46)	4.15 (2.22)	-
Recitation Problems	5.67 (2.08)	6.15 (1.86)	-
Lecture Notes	5.47 (1.92)	6.16 (1.82)	4.53 (1.70)
Textbook	3.83 (2.94)	3.42 (2.21)	4.77 (2.76)
Exams	5.57 (1.89)	4.55 (1.73)	6.20 (1.69)

their most beneficial activity was taking and studying for the exams, followed by reading the textbook and finally the lecture notes.

2.2.3 Data Analysis Tools

Principal Component Analysis

To begin to analyze the relationship between the attitude components and course outcomes, principal component analysis (PCA) was used. Due to the highly correlated nature of the attitude components, as seen in Figure **??** typical multiple regression analysis cannot be used on the raw attitude components simultaneously. Past studies have dealt with this issue by performing separate regression equations for each attitude component as a predictor of course performance, however multivariate techniques such as principal component analysis exist to combat this issue (Millar and Schau, 2010). Principal component analysis is a multivariate technique that aims to reduce the dimensionality of a dataset, while also retaining as much of the original variation as possible. PCA builds new variables, the principal components, that are linear combinations of the original variables which are uncorrelated and ordered to account for decreasing amounts of the variation. PCA is typically used when there are too many explanatory variables in relation to the number of observations or when the explanatory variables are highly correlated, the latter of which is the issue with this dataset (Everitt and Hothorn, 2011).

Principal component analysis aims to build a set of uncorrelated principal components. Each principal component, y_i is the linear combination: $y_i = a_{11}x_1 + a_{12}x_2 + ...a_{16}x_6$ where each vector $\mathbf{x}$ is an attitude component. The coefficients, a_i are chosen to maximize the variance of y_i subject to a constraint that the sum of squares of the coefficients should equal one and each principal component is uncorrelated with each other component before it. The maximization is performed using Lagrange multipliers and each a_i is the eigenvector of the sample covariance matrix corresponding to the matrix's largest eigenvalue under the constraints. Another more numerically sound method to calculate the principal components is by using singular value decomposition of the data matrix. Once the principal components are determined, the number of components necessary is chosen by analyzing a scree plot of the variance explained. This is necessary, as one of the motivations for using PCA is to reduce the number of variables (Everitt and Hothorn, 2011).

Once the principal components are determined, they can also be used to plot the pretest attitudes in lower dimension using a biplot. This plot allows the 6 attitude component vectors and individual student scores to be plotted in the dimensions of two of the principal components. These plots are helpful for viewing potential groups within the data that are not visible in the original multivariate dimensions of the data. They are also able to show the correlation between individual attitude components within each principal component.

Hierarchical Linear Model

Once the principal components were determined for the pretest attitude scores for each class, they were used as explanatory variables in a hierarchical multiple regression model to explain course performance. The aim of this model was to determine if the pretest attitude survey could be used to identify students at risk of performing poorly in the course at the beginning of the year. A hierarchical model was necessary due to the multiple sections within each course, where the grades of students in similar sections were not independent of the section. A separate model was built for each course, dependent on the course's principal components and section. The full model is:

$$Y_{ij} = \beta_0 + \beta_1 PC_{1ij} + \beta_2 PC_{2ij} ... + \beta_6 PC_{6ij} + b_j + \epsilon_{ij} \tag{2.1}$$

where Y_{ij} is the i^{th} student's final numerical course grade from professor j's course as predicted by the principal components and b_j is the random effect for professor. The error term is assumed to follow a normal distribution with a mean of zero and constant variance. A reduced model with the number of sufficient principal components, as determined by the scree plot, was used.

Cluster Analysis

After performing principal component analysis and plotting the biplots for each class, a group structure within each class was explored. Groups of students with similar starting attitudes are of interest to look for differences in performance and learning strategies between the groups. The groups of students were compared for their course performance, leaving attitude and study habits throughout the semester to see if there is any relationship between the starting attitudes and their activities throughout the semester. The method utilized to uncover these groups was cluster analysis.

Cluster analysis attempts to uncover groups or clusters that are homogeneous within a dataset. There are several methods for performing cluster analysis. The method determined to be most suitable (based on the cohesion within the clusters) for this dataset was k-means clustering. K-means clustering attempts to partition the classes of students into k groups $(G_1, G_2, ...G_k)$ where G_i represents the group of n_i students in group i. There are several ways to determine the clustering criterion, with the most common method involving choosing the partitions that minimizes the within-group sum of squares (WGSS) over all variables:

$$WGSS = \Sigma_{j=1}^{q} \Sigma_{l=1}^{k} \Sigma_{i \in G_l} (x_{ij} - \bar{x}_j^{(l)})^2 \tag{2.2}$$

where $\bar{x}_j^{(l)}$ is the mean of the students in group G_l on variable j (Everitt and Hothorn, 2011).

While this method sounds fairly straightforward, in practice it is impractical to search every possible partition of the individuals into k clusters. Instead, algorithms exist to search for improvements in a clustering criterion after some starting partitions are made. With k-means clustering, k has to be determined before running the algorithm. Choosing k can be done several ways, including by running k-means for several values of k and analyzing a scree plot of the WGSS. The choice of k is where the "elbow" in the plot is (Everitt and Hothorn, 2011).

Canonical Correlation Analysis

Once cluster analysis was used to find groups of students with similar starting attitudes in each class, statistical tests were performed to compare learning characteristics between the groups. To further characterize the relationship between learning preferences and starting attitudes, canonical correlation analysis was used. Canonical correlation analysis looks for relationships between two sets of variables, like multiple regression, but with multiple response variables. CCA attempts to quantify the association between two sets of variables, $x^T = (x_1, x_2, ..., x_{q1})$ and $y^T = (y_1, y_2, ..., y_{q2})$ as the largest correlation between two single variables u_1 and v_1 where u_1 is a linear combination of $x_1, x_2, ..., x_{q1}$ and v_1 is a linear combination of $y_1, y_2, ..., y_{q2}$. Often, one pair of variables (u_1, v_1) cannot sufficiently quantify the relationship and several pairs are necessary. The pairs (u_i, v_i) are chosen such that the u_1 are mutually uncorrelated, as are the v_i, the correlation between u_i and v_i

is R_i where the correlation decreases as i increases and the u_i is uncorrelated with all v_j except v_i (Everitt and Hothorn, 2011). Here, the set of variables will be the change in attitudes related to the rankings of learning activities to see if there is a relationship between the change in attitudes and the way students preferred to learn in the class.

2.3 RESULTS

2.3.1 Omega Internal Consistency

The internal consistency of the attitude component structure of the SATS-36 instrument must first be checked to ensure that each group of questions is measuring the intended construct. Typically, Cronbach's coefficient alpha is reported for the internal consistency, however the alternative method, Omega, is reported as this measure holds fewer restrictions on the data (Tavakol and Dennick, 2011). The pretest and posttest Omega values for all attitude components for each course can be found in Table .9 showing both the point estimate and bootstrap confidence interval using the MBESS package in R Studio (R Core Team, 2017; Kelley, 2017). The Omega values for all attitudes are within the acceptable range (above 0.70), except the difficulty values.

2.3.2 Principal Component Analysis

The pretest attitude scores are of the most interest to this research due to their timing within the course. Learning more about students at the beginning of the course, especially something with potential to predict student success, is very important. In order to use the pretest attitude scores in a regression model, their correlation amongst themselves needs to be taken into consideration. To remedy this problem, principal component analysis was applied to the pretest attitudes using the prcomp function in the stats package in R Studio (R Core Team, 2013). This function takes the scaled pretest attitudes and uses singular value decomposition to determine the value of the principal components.

PCA was applied to all three courses separately. The results from STA307 are included here, with significant deviations from the other courses included for comparison. From the screeplot in Figure 2.2 the number of principal components necessary to explain a sufficient amount of the original variation in the STA307 pretest attitude scores appears to be 3 components. These first 3 principal components explain 86.7% of the original variance. The coefficient values for these principal components are in Table 2.5. The first PC appears to be a weighted average of Affect, Cognitive Competence, Value and Interest, with smaller contributions from Difficulty and Effort. This component seems to consist of students starting perception of the subject, rather than their attitude towards the amount of work it will be to complete the material. The second PC is most heavily dominated by Difficulty. The third PC is dominated by Effort. The results for STA308 and STA409 show similar results, with 3 PCs being sufficient to explain the majority of the variation in the pretest attitudes. The interpretation of the PCs are also similar, except for PC2 for STA308 where Difficulty is still the largest contributor, however Value and Effort contribute more than in the other classes.

The biplots in Figure 2.3 and Figure 2.4 display a projection of the students and attitude vectors onto the first and second or first and third principal components respectively. The students are represented by the symbols and colored by their final grade as either an A (93.5+) or not. Figure 2.3 shows the separation of Difficulty from the other attitude factors in the space of PC2. The ellipses also indicate the potential difference in performance between students who have more positive attitudes (in the direction of the vectors) and those with lower starting attitudes. Figure 2.4 shows the separation of the Effort vector from the other attitude components in PC3. The separation between the students who performed well

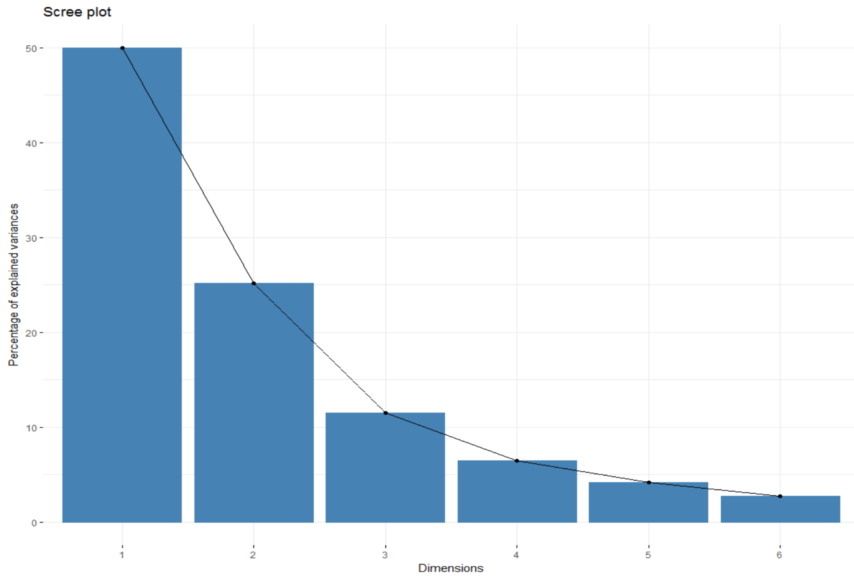

Figure 2.2. Scree plot of variance explained by each principal component for STA307 pretest attitudes. The "elbow" of the plot appears to be at 3 principal components to explain a sufficient amount of the original variation.

being closer to the positive attitude vector direction is also clear in these dimensions. The biplots for STA308 and STA409 showed similar results with the attitude vectors, however the separation between the grade ellipses is not as apparent in those classes.

Table 2.5. Coefficients for each pretest attitude component in each principal component.

Attitude Component	PC1	PC2	PC3
Affect	0.50	-0.23	0.11
Cognitive Competence	0.47	-0.35	-0.14
Value	0.48	0.31	0.17
Difficulty	0.17	-0.71	-0.03
Interest	0.41	0.35	0.51
Effort	0.33	0.33	-0.83

2.3.3 Hierarchical Linear Model

Once the principal components were calculated and the number of components sufficient to explain the pretest attitude variation were chosen, they were

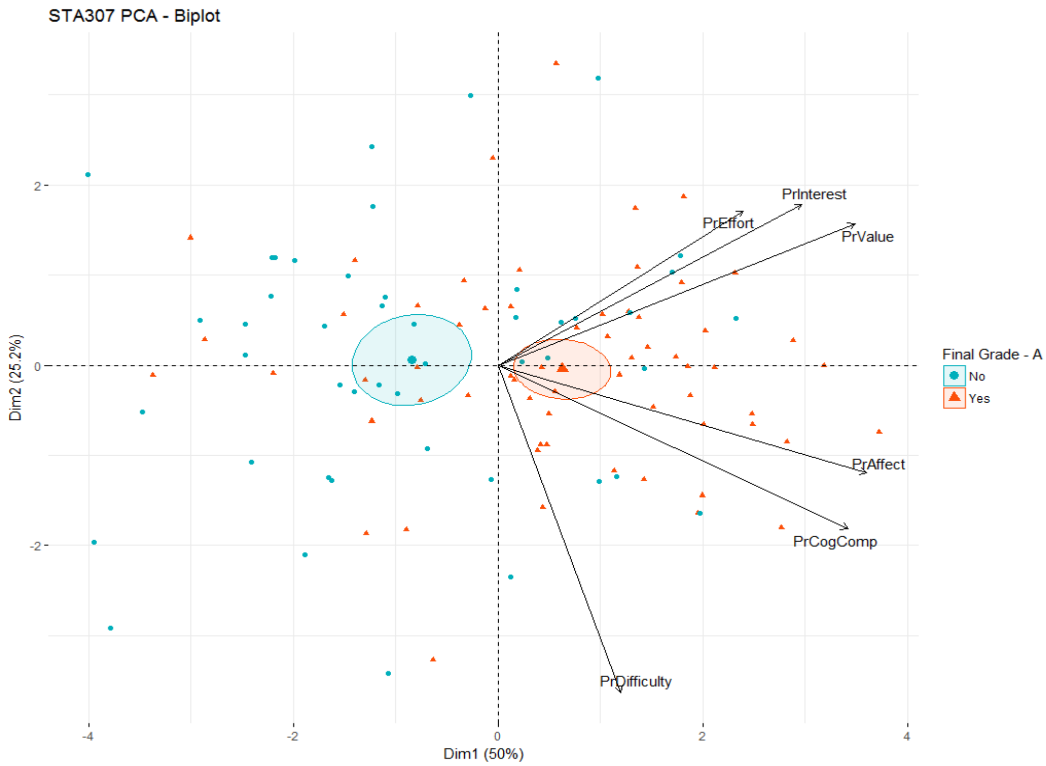

Figure 2.3. Biplot of the first two principal components for the pretest attitudes for STA307 with students plotted as points colored by their final grade as an A (93.5 and up) or not.

implemented in the hierarchical linear regression model to predict students' final numerical grades. Results from the regression run in R Studio using the lmer function in the lme4 package (Bates et al., 2015). The resulting model did not meet the assumption of normality of the residuals, so the mean and confidence intervals for the coefficients were bootstrapped by re-sampling the data. The regression models for STA308 and STA409 met the assumption of normality. The results for STA307 can be found in Table 2.6 and show that only PC1 was significant in predicting students' final grades.

The effect of PC1 is an increase in final grade of approximately 1.97 points. From the interpretation of PC1, this predicts that students' with higher pretest Affect, Cognitive Competence, Value and Interest scores have higher grades, on

Figure 2.4. Biplot of the first and third principal components for the pretest attitudes for STA307 with students plotted as points colored by their final grade as an A (93.5 and up) or not.

average. The results for STA308 showed similar results, with PC1 being the only significant predictor of final grade. The regression for STA409 did not show any significant predictors of final grade.

Table 2.6. Results from the regression of final grade on the first three PCs for STA307 pretest attitudes.

	Mean	**95% C.I.**
Intercept	90.82	(88.46, 92.96)
PC1	1.97	(0.88, 3.12)
PC2	-1.09	(-2.67, 0.29)
PC3	-0.15	(-2.30, 1.86)

2.3.4 Cluster Analysis

Principal component and regression analysis have found an indication that students' pretest attitude scores may have a significant relationship to their course performance. The next step in the analysis was to explore the grouping structure within the pretest attitudes to see if there are groups of students with similar attitudes. These groups were then compared for their average attitudes at the beginning and end of the semester, differences in final grades and in learning activities throughout the semester.

The cluster analysis was performed in R Studio using the eclust function in the factoextra package (Kassambara and Mundt, 2017). The chosen clustering method was kmeans, with k chosen through applying the cluster algorithm to $k = 1 : 10$ and plotting the Total Within Group Sum of Squares, as seen in Figure 2.5. The graph stops decreasing as sharply at $k = 2$. The cluster analysis was then applied to the pretest attitudes with 2 clusters specified. The resulting clustering result can be seen in Figure 2.6 plotted in the dimensions of the first two principal components. The clusters appear to have some overlap. Similar plots were made for STA308 and STA409 as both also showed $k = 2$ as the best solution based on the WGSS. The plot for STA409 shows greater separation between the clusters.

Once the students were clustered based on their pretest attitudes, differences between these groups were explored. The first measure looked at was the pretest attitudes themselves to see the composition of the clusters. Table 2.7 shows the mean and standard deviation for each attitude component for each cluster. For all three classes, cluster one has the higher averages with the exception of STA308's Effort mean and STA409's Difficulty average. This indicates that cluster one was built to consist of students with more positive starting attitudes towards statistics. The average change in attitude throughout the semester was also investigated

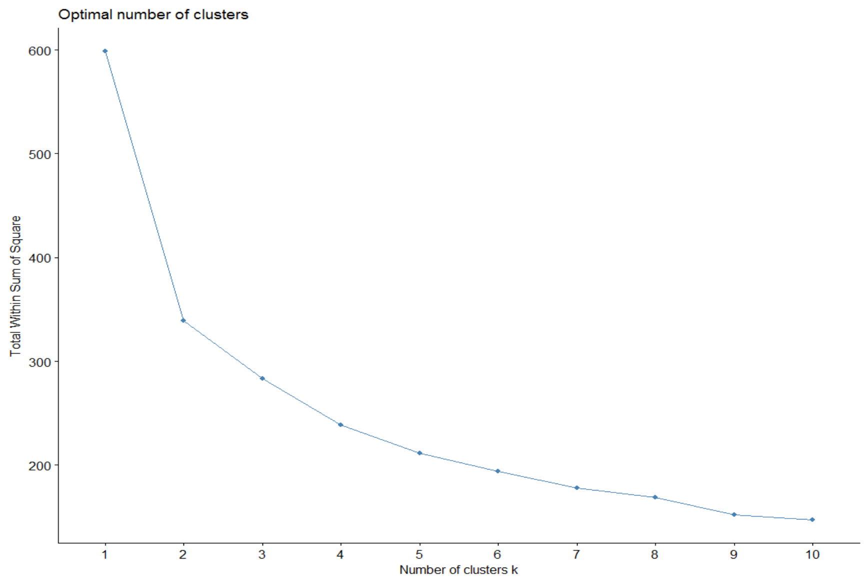

Figure 2.5. Cluster screeplot to determine choice of k for the k-means clustering of STA307 students based on their pretest attitudes. Based on the total within sum of squares, the best choice of k appears to be 2.

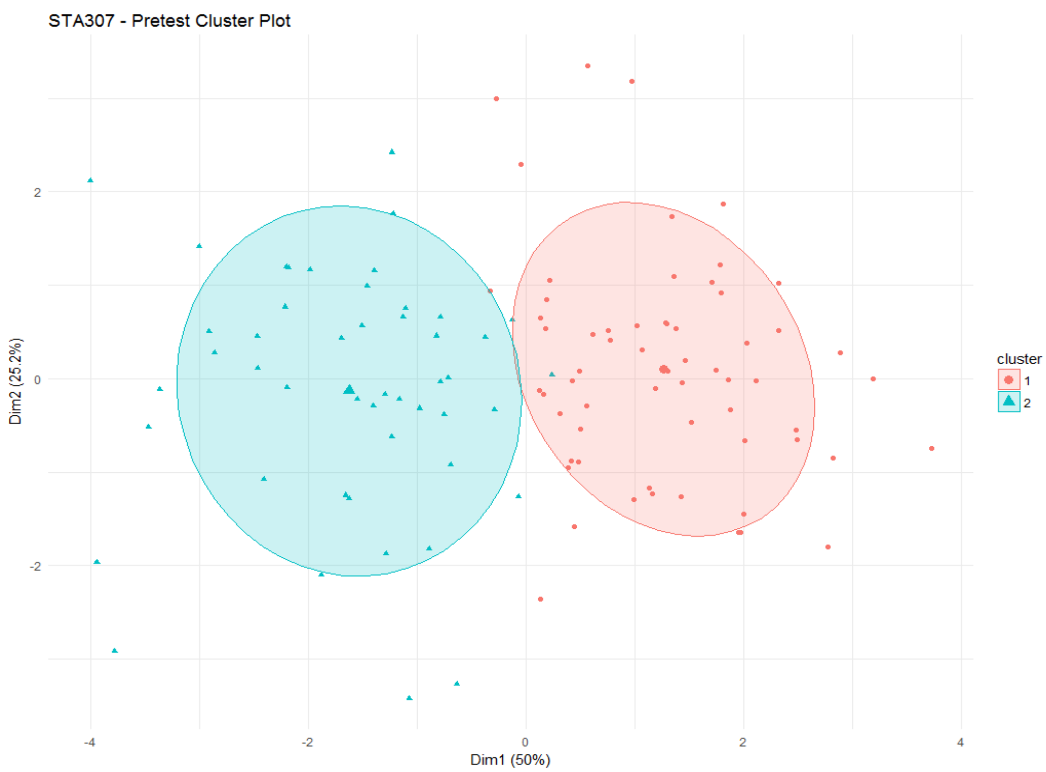

Figure 2.6. Cluster scatterplot for the k-means clustering of STA307 students' prestest attitudes plotted versus the first two principal components.

Table 2.7. Pretest attitude averages for each cluster of each class.

		STA307		**STA308**		**STA409**	
		C1	**C2**	**C1**	**C2**	**C1**	**C2**
Affect	Mean	5.54	4.14	5.02	4.51	5.34	4.51
	S.D.	0.76	0.75	1.20	1.01	0.72	0.87
Cognitive Competence	Mean	6.0	4.8	5.61	5.23	6.01	5.02
	S.D.	0.64	0.85	1.07	0.95	0.58	0.70
Value	Mean	5.96	4.32	5.05	4.86	6.05	4.97
	S.D.	0.57	0.71	1.11	0.84	0.58	0.70
Difficulty	Mean	4.07	3.87	4.09	3.96	3.95	4.07
	S.D.	0.64	0.54	0.63	0.61	0.55	0.74
Interest	Mean	5.57	3.75	4.41	4.09	5.64	4.12
	S.D.	0.88	1.06	1.31	1.12	0.67	0.74
Effort	Mean	6.79	6.32	6.43	6.58	6.79	6.33
	S.D.	0.37	0.77	0.73	0.48	0.28	0.68

for each set of clusters. Both STA307 and STA409 showed very little difference between the clusters, STA308 showed cluster 2 having an increase in attitude throughout the semester. The posttest averages follow the same pattern, with STA307 and STA409 scores higher for cluster 1 and STA308 having higher posttest scores for cluster 2.

Next, a qualitative analysis of the clusters based on the grades, demographics, learning preferences and study habits indicated in the exit survey was conducted. Selected results are in Table 2.8. The questions involving use of resources were grouped into two categories: Rare Use ($<$ 3 times) or Frequent Use ($\geq$ 3 times). The bold numbers represent a significant result on either a Wilcoxon Rank Sum Test or a Chi-Square Test of Independence between groups at the 5% level and italics represent a significant result at the 10% level. The distribution of gender between each cluster show a greater proportion of females in cluster one for STA307 and STA409, while STA308 has a significantly larger proportion of females in cluster 2. No classes show a significant difference between the clusters in terms of using the textbook as a resource and only STA307 shows a significant result

Table 2.8. Qualitative analysis of each cluster.

	STA307		**STA308**		**STA409**	
	C1	**C2**	**C1**	**C2**	**C1**	**C2**
n	62	48	41	35	35	24
Gender (% Female)	74%	66%	**56%**	**83%**	49%	29%
Use Textbook	24%	28%	39%	36%	65%	44%
Study in Groups	*55%*	*72%*	53%	40%	59%	70%
Mean Rank Recitation Problems	5.92	5.55	**5.94**	**6.91**	-	-
Mean Rank Lecture Notes	5.69	5.32	6.39	6.55	**4.97**	**4.04**
Mean Rank Exams	**6.07**	**5.19**	4.25	4.91	6.18	6.30
Mean Final Grade	**94.87**	**85.68**	91.34	90.95	87.34	84.15

for studying in groups. It is interesting to see how many more students used the textbook in STA409 than the other courses. For the average rank for the resources, STA308 showed a significantly higher rank for cluster 2 valuing the recitation problems, STA409 showed a significant difference for the mean rank of the value of the lecture notes as a benefit to their learning and STA307 showed a difference in the valuation of the exams. Finally, the final grade averages within each cluster were compared. The average final grade is higher in cluster one for all classes, which corresponds with the results from the principal component analysis, however only STA307 was significantly different.

2.3.5 Canonical Correlation Analysis

Canonical correlation analysis was used to explore the relationship between the change in attitudes and the learning activity ranks for each class. The process was implemented in R Studio. The first two canonical variates for STA307 are:

$$u_1 = -0.26Diff + 0.53CogComp + 0.30Effort - 0.13Value + 0.70Interest + 0.24Affect$$

$$v_1 = -0.54SAS + 0.34Recitations - 0.62Lecture + 0.45Textbook - 0.10Exams$$

$$u_1 = 0.54Diff - 0.66CogComp + 0.26Effort + 0.44Value + 0.09Interest - 0.00Affect$$

$$v_1 = 0.63SAS + 0.12Recitations - 0.17Lecture + 0.42Textbook - 0.62Exams$$

The correlation between the first variates is 0.27 and between the second variates is 0.10. An interpretation of this result is that a positive change in interest and cognitive competence is weakly correlated with low ranking of SAS and lecture notes, but a higher ranking of the textbook. The second variates show that a a positive change in difficulty and value and a negative change in cognitive competence is very weakly correlated with a higher ranking of SAS and the textbook and a lower ranking of the exams. This could show that students who left with higher cognitive competence and interest than they entered put a lower value on SAS and lecture notes, but learned independently from the textbook.

The results of CCA for the STA308 indicate that a positive change in affect is 0.16 correlated with a higher ranking of Excel and the exams. This indicates that students who left with a more positive feeling towards the subject placed a high value on the Excel assignments and exams. The results for STA409 indicate that an increase in cognitive competence and decrease in affect is 0.31 correlated with a higher ranking of the exams and lower ranking of the textbook.

2.4 MAIN FINDINGS

2.4.1 Limitations

As with any survey design with human subjects, there are several biases that present themselves in the data. First, there is the non-response bias from the students who did not consent or did not complete the surveys. There were 25, 78, and 49 students in STA307, STA308 and STA409 respectively that did not sign the IRB consent form to have their data included in the research. Some of these students may have been absent from the classroom of the day the consent forms were distributed, others may have refused to participate. The sample size decreased again when looking at the number of students that completed both of the SATS-36 surveys. There were 18, 94, and 5 additional students from each course that were not included in the analysis. An explanation for the low involvement in STA308 is a lack of incentive for students to complete the surveys. Only one section offered an incentive to participate, biasing the results towards that section. STA307 and one section of STA409 included the surveys on a homework assignment which helps to convince students to participate.

These reductions to the sample size may not have been random. The students who completed the surveys, either due to their own accord or due to the incentive very well may be from a different population of students than those that decided not to complete the surveys. The students who were absent from the class when the consent forms were distributed may have a different relationship between their attitudes and course performance than those present.

In addition to the nonresponse bias, every survey has the potential for response bias. Students are self reporting and answering a variety of questions. There is no way to know that the responses are entirely truthful, especially if the students were rushing to complete the survey just to get the task done. To limit this, there are negatively worded questions on the SATS-36 to identify students answering every question the same way. This will not totally protect against response bias due to the possibility that students were answering how they thought they should answer, instead of how they actually felt. It is important to keep these biases in mind when interpreting the results of this paper. The responses being analyzed can not be guaranteed to represent the entire course populations.

2.4.2 Practical Recommendations

Beyond the results discussed above, there are several practical implications to this work. The principal component analysis in conjunction with the regression analysis concluded that there is a potential relationship between pretest attitude scores and the course performance within two of the three classes. This is important because it shows that students' Value, Interest, Affect and Cognitive Competence at the start of the semester can affect their performance. Students with lower confidence in their technical skills and lower perception of the subject will not perform as well throughout the semester. This can be an important finding, if confirmed with further studies, to develop a way to identify students at risk of performing poorly at the beginning of the semester and possibly implement intervention methods.

The cluster analysis identified a potential two group structure to the class in terms of the pretest attitudes. Cluster one, containing students with higher average pretest attitudes, showed several interesting characteristics depending on the class. In STA307, cluster one had a higher average final grade and was more likely to find the exams beneficial to their learning. In STA308, cluster one was more likely to find the recitation problems helpful and contained a smaller proportion of females. In STA409, cluster one was more likely to find the lectures beneficial to their learning. These findings are important because they indicate that there are learning style differences between each group of students and one group is performing better than the other.

In practice, professors' should bear in mind that students rank the lectures and exams highly and teaching assistants should consider the importance of the recitations to students learning. It is also indicated that students in two of the classes do not view the textbook as a valuable resource. Other resources that students value in their learning process regardless of cluster include the online notes, practice exams and studying alone. The resources students don't utilize often include their TA and Professor's office hours and email to contact their instructors. Knowing the study habits of the students is important for educators, especially if students are not utilizing a valuable resource. Potential recommendations here include evaluating the choice of textbook for each course, advertising for office hours and promoting yourself as a resource for students, as well as making office hours at least once during the semester required for the course. However, based on the characteristics of Generation Z students, their independent learning styles and technological skills may be able to explain the choices in resources.

Future work includes looking into student collaboration networks within each course to see how students work together and choose which peers to work with. Other future work could consist of following up in future classes with similar pretest attitude surveys and exit surveys to see if the results are consistent, as well as to ask additional questions about other resources and learning preferences. Adding an option to have students self describe their own learning styles and give suggestions at the end of the semester could also be very beneficial to understanding the needs of the current generation of students.

List of References

(2017). Cs consultants, llc. http://www.evaluationandstatistics.com/. Accessed: 2017-09-10.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Bond, M. E., Perkins, S. N., and Ramirez, C. (2012). Students'perceptions of statistics: An exploration of attitudes, conceptualizations, and content knowledge of statistics. *Statistics Education Research Journal*, 11(2).

Everitt, B. and Hothorn, T. (2011). *An introduction to applied multivariate analysis with R.* Springer Science & Business Media.

Hora, M. T. and Oleson, A. K. (2017). Examining study habits in undergraduate stem courses from a situative perspective. *International Journal of STEM Education*, 4(1):1.

Kassambara, A. and Mundt, F. (2017). *factoextra: Extract and Visualize the Results of Multivariate Data Analyses.* R package version 1.0.5.

Kelley, K. (2017). *MBESS: The MBESS R Package.* R package version 4.4.1.

Millar, A. M. and Schau, C. (2010). Assessing students attitudes: the good, the bad, and the ugly.

Nolan, M. M., Beran, T., and Hecker, K. G. (2012). Surveys assessing students'attitudes toward statistics: A systematic review of validity and reliability. *Statistics Education Research Journal*, 11(2).

R Core Team (2013). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

R Core Team (2017). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Schau, C. and Emmioğlu, E. (2012). Do introductory statistics courses in the united states improve students'attitudes? *Statistics Education Research Journal*, 11(2).

Seemiller, C. and Grace, M. (2017). Generation z: Educating and engaging the next generation of students. *About Campus*, 22(3):21–26.

Shatto, B. and Erwin, K. (2016). Moving on from millennials: Preparing for generation z. *The Journal of Continuing Education in Nursing*, 47(6):253–254.

Tavakol, M. and Dennick, R. (2011). Making sense of cronbach's alpha. *International journal of medical education*, 2:53.

.1 Additional Tables and Figures

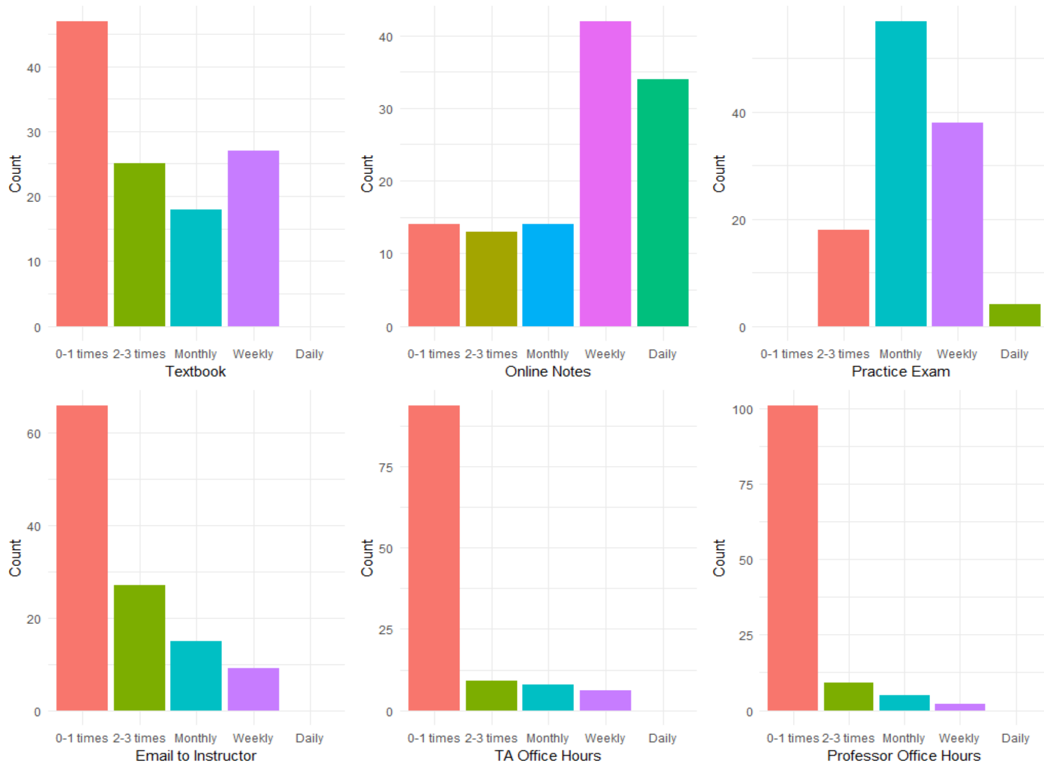

Figure .7. Bar plots of each of the resources surveyed on the exit survey in STA308. The most used resources were the online notes and practice exams.

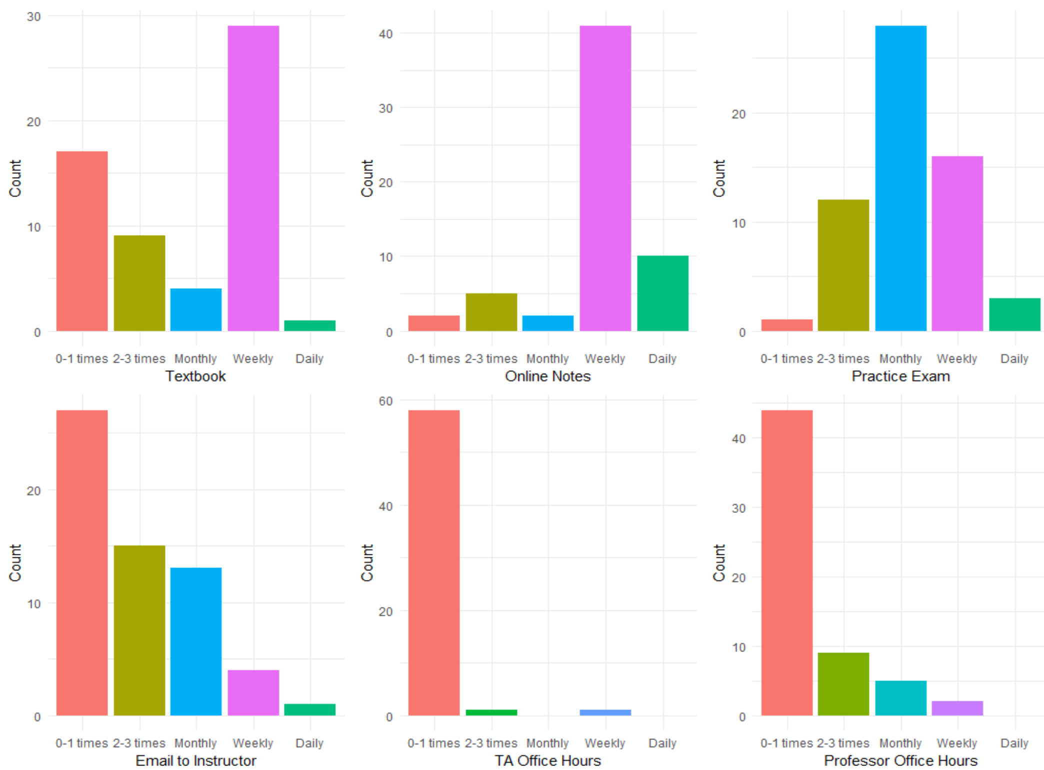

Figure .8. Bar plots of each of the resources surveyed on the exit survey in STA409. The most used resources were the online notes and practice exams.

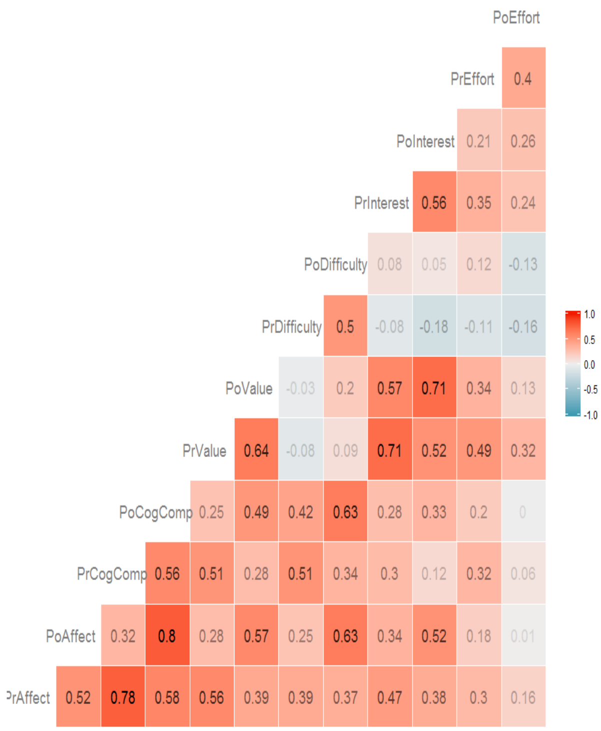

Figure .9. Correlation plot for pretest and posttest attitude components for STA307.

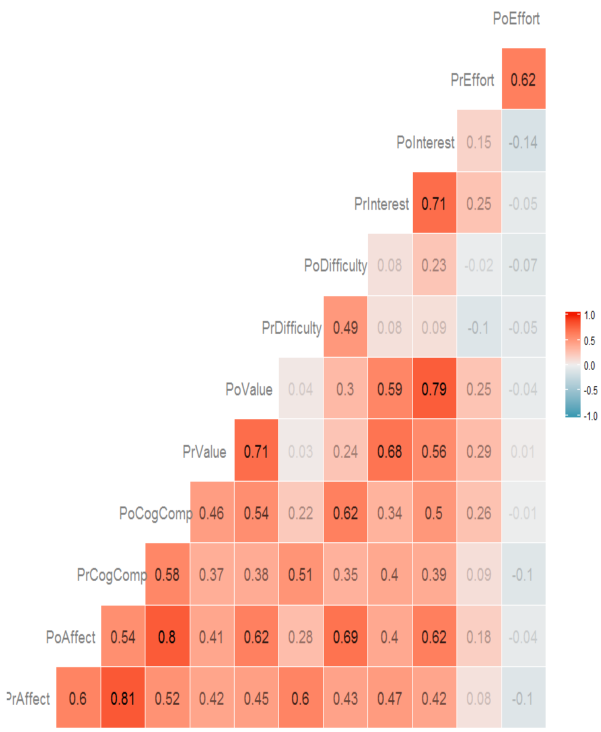

Figure .10. Correlation plot for pretest and posttest attitude components for STA308.

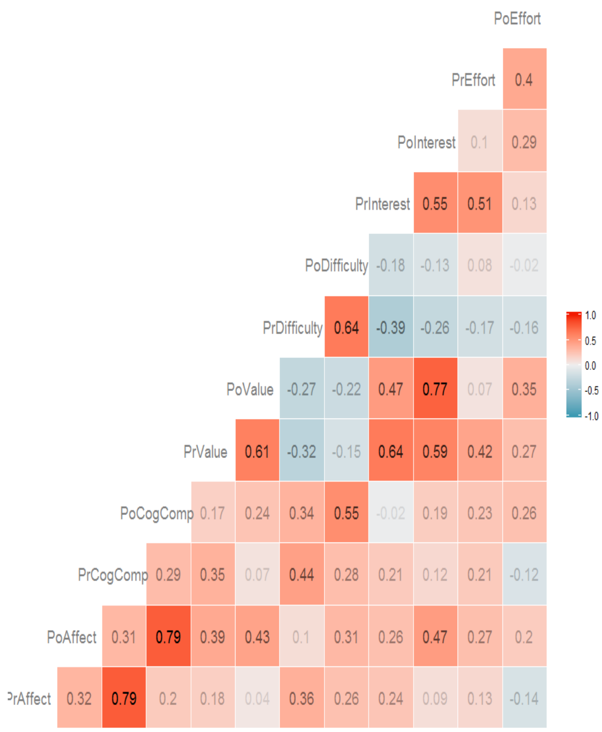

Figure .11. Correlation plot for pretest and posttest attitude components for STA409.

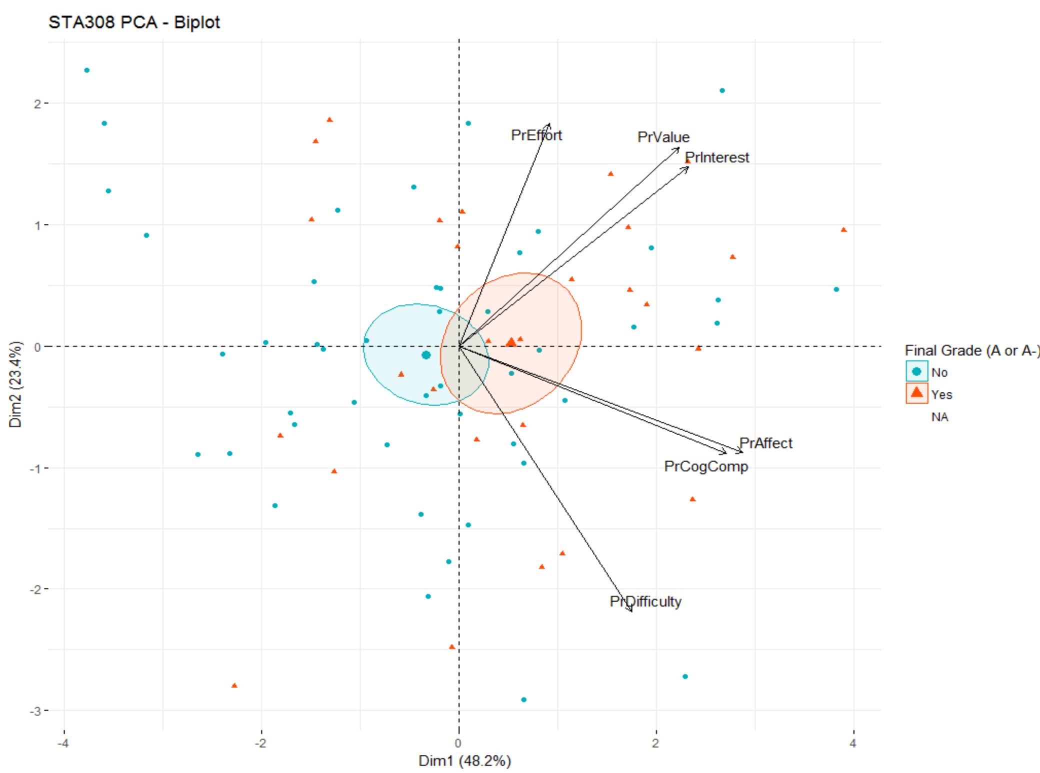

Figure .12. Biplot of the first two principal components for the pretest attitudes for STA308 with students plotted as points colored by their final grade as an A or A- (89.5 and up) or not.

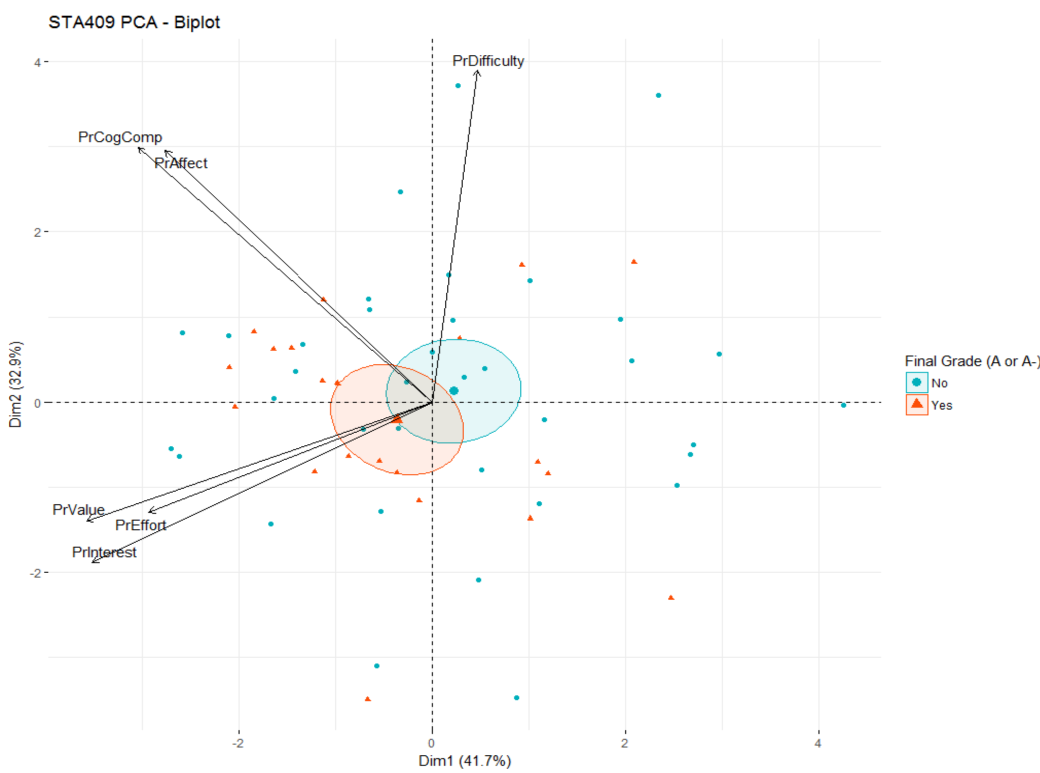

Figure .13. Biplot of the first two principal components for the pretest attitudes for STA409 with students plotted as points colored by their final grade as an A or A- (89.5 and up) or not.

Table .9. Omega Values for Each Attitude Component in Each Course

Attitude Component	**Course**	**Test**	**Omega**	**Bootstrap 95% C.I.**
Affect	STA307	Pretest	0.77	(0.69,0.83)
Cognitive Competence	STA307	Pretest	0.80	(0.68,0.87)
Value	STA307	Pretest	0.90	(0.87,0.92)
Difficulty	STA307	Pretest	0.21	(0,0.52)
Interest	STA307	Pretest	0.92	(0.90,0.95)
Effort	STA307	Pretest	0.85	(0.71,0.92)
Affect	STA307	Posttest	0.83	(0.76,0.88)
Cognitive Competence	STA307	Posttest	0.85	(0.77,0.89)
Value	STA307	Posttest	0.90	(0.87,0.93)
Difficulty	STA307	Posttest	0.67	(0.56,0.75)
Interest	STA307	Posttest	0.92	(0.89,0.94)
Effort	STA307	Posttest	0.70	(0.52,0.80)
Affect	STA308	Pretest	0.86	(0.79,0.91)
Cognitive Competence	STA307	Pretest	0.87	(0.80,0.91)
Value	STA308	Pretest	0.89	(0.84,0.92)
Difficulty	STA308	Pretest	0.53	(0.30,0.73)
Interest	STA308	Pretest	0.87	(0.78,0.91)
Effort	STA308	Pretest	0.79	(0.58,0.90)
Affect	STA308	Posttest	0.86	(0.79,0.91)
Cognitive Competence	STA308	Posttest	0.86	(0.77,0.92)
Value	STA308	Posttest	0.91	(0.87,0.93)
Difficulty	STA308	Posttest	0.57	(0.40,0.69)
Interest	STA308	Posttest	0.91	(0.86,0.94)
Effort	STA308	Posttest	0.80	(0.66,1)
Affect	STA409	Pretest	0.75	(0.51,0.84)
Cognitive Competence	STA409	Pretest	0.82	(0.72,0.88)
Value	STA409	Pretest	0.84	(0.73,0.89)
Difficulty	STA409	Pretest	0.63	(0.35,0.80)
Interest	STA409	Pretest	0.88	(0.82,0.92)
Effort	STA409	Pretest	0.78	(0.60,0.90)
Affect	STA409	Posttest	0.84	(0.75,0.89)
Cognitive Competence	STA409	Posttest	0.84	(0.73,0.89)
Value	STA409	Posttest	0.90	(0.84,0.94)
Difficulty	STA409	Posttest	0.70	(0.50,0.81)
Interest	STA409	Posttest	0.93	(0.88,0.96)
Effort	STA409	Posttest	0.65	(0.43,0.83)

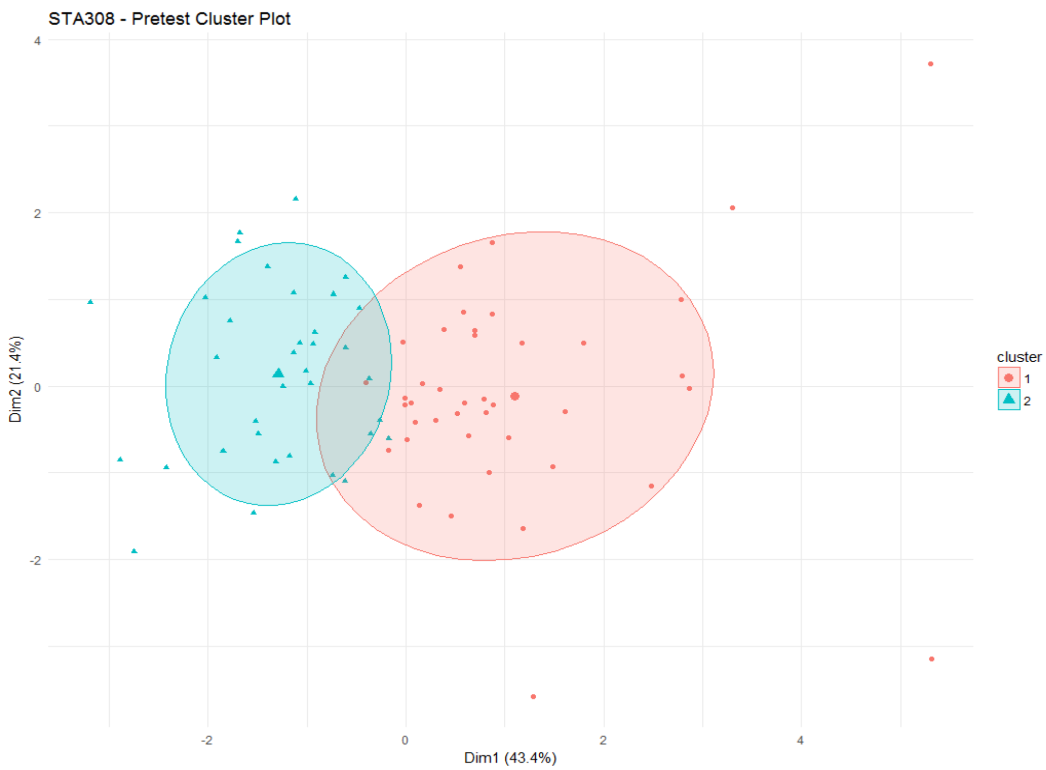

Figure .14. Cluster scatterplot for the k-means clustering of STA308 students' prestest attitudes plotted versus the first two principal components.

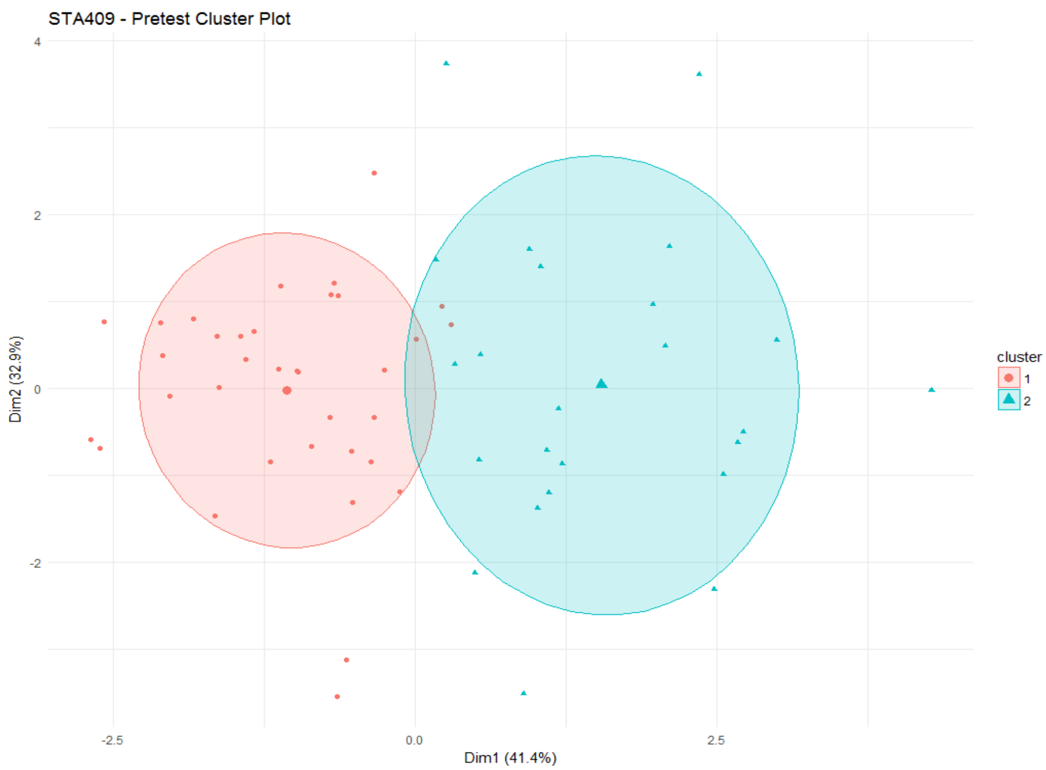

Figure .15. Cluster scatterplot for the k-means clustering of STA409 students' prestest attitudes plotted versus the first two principal components.

BIBLIOGRAPHY

"Cs consultants, llc," http://www.evaluationandstatistics.com/, accessed: 2017-09-10.

"Socrative by mastery connect," https://www.socrative.com/index.html, accessed: 2017-09-10.

Bates, D., Mächler, M., Bolker, B., and Walker, S., "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

Bond, M. E., Perkins, S. N., and Ramirez, C., "Students'perceptions of statistics: An exploration of attitudes, conceptualizations, and content knowledge of statistics." *Statistics Education Research Journal*, vol. 11, no. 2, 2012.

Carlson, K. A. and Winquist, J. R., "Evaluating an active learning approach to teaching introductory statistics: A classroom workbook approach," *Journal of Statistics Education*, vol. 19, no. 1, pp. 1–23, 2011.

Carnell, L. J., "The effect of a student-designed data collection project on attitudes toward statistics," *Journal of Statistics Education*, vol. 16, no. 1, pp. 1–15, 2008.

Coetzee, S. and Merwe, P. v. d., "Industrial psychology students' attitudes towards statistics," *SA Journal of Industrial Psychology*, vol. 36, no. 1, pp. 1–8, 2010.

Dauphinee, T. L., Schau, C., and Stevens, J. J., "Survey of attitudes toward statistics: Factor structure and factorial invariance for women and men," *Structural Equation Modeling: a multidisciplinary journal*, vol. 4, no. 2, pp. 129–141, 1997.

DeVaney, T. A., "Anxiety and attitude of graduate students in on-campus vs. online statistics courses," *Journal of Statistics Education*, vol. 18, no. 1, pp. 1–15, 2010.

Dunham, B., "Statistics clicks: Using clickers in introductory statistics courses," *University of British Columbia*, 2009.

Emmioğlu, E., "A structural equation model examining the relationships among mathematics achievement, attitudes toward statistics, and statistics outcomes," *Unpublished doctoral dissertation). Middle East Technical University, Ankara, Turkey*, 2011.

Fitzmaurice, G. M., Laird, N. M., and Ware, J. H., *Applied longitudinal analysis.* John Wiley & Sons, 2012, vol. 998.

Gundlach, E., Richards, K. A. R., Nelson, D., and Levesque-Bristol, C., "A comparison of student attitudes, statistical reasoning, performance, and perceptions for web-augmented traditional, fully online, and flipped sections of a statistical literacy class," *Journal of Statistics Education*, vol. 23, no. 1, pp. 1–33, 2015.

Hannigan, A., Gill, O., and Leavy, A. M., "An investigation of prospective secondary mathematics teachers conceptual knowledge of and attitudes towards statistics," *Journal of Mathematics Teacher Education*, vol. 16, no. 6, pp. 427–449, 2013.

Hannigan, A., Hegarty, A. C., and McGrath, D., "Attitudes towards statistics of graduate entry medical students: the role of prior learning experiences," *BMC medical education*, vol. 14, no. 1, p. 70, 2014.

Henderson, D. J. and Parmeter, C. F., *Applied nonparametric econometrics.* Cambridge University Press, 2015.

Hilton, S. C., Schau, C., and Olsen, J. A., "Survey of attitudes toward statistics: Factor structure invariance by gender and by administration time," *Structural Equation Modeling*, vol. 11, no. 1, pp. 92–109, 2004.

Hothorn, T. and Zeileis, A., *partykit: A Modular Toolkit for Recursive Partytioning in R*, 2015, r package version 1.1-1. [Online]. Available: http://jmlr.org/papers/v16/hothorn15a.html

Kelley, K., *MBESS: The MBESS R Package*, 2017, r package version 4.4.1. [Online]. Available: https://CRAN.R-project.org/package=MBESS

Mathew, L. and Aktan, N. M., "Nursing student attitudes toward statistics," *Journal of Nursing Education*, vol. 53, no. 4, pp. 233–237, 2014.

McGowan, H. M. and Gunderson, B. K., "A randomized experiment exploring how certain features of clicker use effect undergraduate students' engagement and learning in statistics," *Technology Innovations in Statistics Education*, vol. 4, no. 1, 2010.

Millar, A. M. and Schau, C., "Assessing students attitudes: the good, the bad, and the ugly," 2010.

Nolan, M. M., Beran, T., and Hecker, K. G., "Surveys assessing students'attitudes toward statistics: A systematic review of validity and reliability." *Statistics Education Research Journal*, vol. 11, no. 2, 2012.

Pek, J., Wong, O., and Wong, C., "Data transformations for inference with linear regression: Clarifications and recommendations." *Practical Assessment, Research & Evaluation*, vol. 22, 2017.

R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2017. [Online]. Available: https://www.R-project.org/

SAS Institute Inc., “Sas studio.” [Online]. Available: https://www.sas.com

Schau, C. and Emmioğlu, E., “Do introductory statistics courses in the united states improve students’attitudes?” *Statistics Education Research Journal*, vol. 11, no. 2, 2012.

Shatto, B. and Erwin, K., “Moving on from millennials: Preparing for generation z,” *The Journal of Continuing Education in Nursing*, vol. 47, no. 6, pp. 253–254, 2016.

Sorge, C. and Schau, C., “Impact of engineering students attitudes on achievement in statistics: A structural model,” 2002.

Tavakol, M. and Dennick, R., “Making sense of cronbach’s alpha,” *International journal of medical education*, vol. 2, p. 53, 2011.

Therneau, T., Atkinson, B., and Ripley, B., *rpart: Recursive Partitioning and Regression Trees*, 2015, r package version 4.1-10. [Online]. Available: https://CRAN.R-project.org/package=rpart

Vanhoof, S., “Statistics attitudes in university students: Structure, stability and relationship with achievement,” 2011.

Vaterlaus, J. M., Beckert, T. E., Fauth, E. B., and Teemant, B., “An examination of the influence of clicker technology on college student involvement and recall.” *International Journal of Teaching and Learning in Higher Education*, vol. 24, no. 3, pp. 293–300, 2012.