Using Ninth Grade and Sixth Grade Indicators within an Early Warning System to Predict Students at Risk for Graduating Late or Dropping Out

A Dissertation submitted
to the Graduate School
Valdosta State University

in partial fulfillment of requirements
for the degree of

DOCTOR OF EDUCATION

in Curriculum and Instruction

in the Department of Curriculum, Leadership, and Technology
of the Dewar College of Education and Human Services

September 2020

Dana Halley Morris

Ed.S., Georgia College & State University, 2005
M.A.T., Georgia College & State University, 2002
B.S., Mercer University, 1991

This dissertation, "Using Ninth Grade and Sixth Grade Indicators within an Early Warning System to Predict Students at Risk for Graduating Late or Dropping Out," by Dana Halley Morris, is approved by:

Dissertation Committee Chair	Lantry L. Brockmeier Lantry L. Brockmeier, Ph.D. Professor of Curriculum, Leadership, and Technology
Committee Members	J. L. Pate James L. Pate, Ph.D. Professor of Curriculum, Leadership, and Technology
	Gerald R. Siegrist, Ed.D. Professor of Curriculum, Leadership, and Technology
Associate Provost for Graduate Studies and Research	Becky K. da Cruz Becky K. da Cruz, Ph.D., J.D. Professor of Criminal Justice
Defense Date	9.23.2020

FAIR USE

This dissertation is protected by the Copyright Laws of the United States (Public Law 94-553, revised in 1976). Consistent with fair use as defined in the Copyright Laws, brief quotations from this material are allowed with proper acknowledgement. Use of the material for financial gain without the author's expressed written permission is not allowed.

DUPLICATION

I authorize the Head of Interlibrary Loan or the Head of Archives at the Odum Library at Valdosta State University to arrange for duplication of this dissertation for educational or scholarly purposes when so requested by a library user. The duplication shall be at the user's expense.

Signature Dara H. Morris

I refuse permission for this dissertation to be duplicated in whole or in part.

Signature ______________________________

Abstract

The problem of high school dropouts has been studied for decades, but utilizing readily obtainable student data and data mining can aid school leaders to more accurately detect which students will likely drop out. This early warning information can be used by educators as early as sixth grade to help identify potential high school dropouts and students who will not graduate on time and intervene more efficiently and effectively with those students. The purpose of this nonexperimental correlational study was to use longitudinal data from a mid-sized school district from two cohorts to support the creation of a dropout early warning system to predict both sixth-grade and ninth-grade students who are at risk for not graduating on time. The statistical models utilized to identify the most accurate indicators were logistics regression, linear discriminate analysis, and quadratic discriminate analysis. The variables identified in the ninth-grade models as able to predict students who would not complete high school within four years were: if a student did not receive enough credits to advance to the tenth grade, did not attend school at least 90% of the time, was suspended from school, had multiple school moves in the ninth grade, and gender. The sixth-grade variables identified as able to predict students who would not complete high school within four years were: if the student was suspended from school, had multiple school moves, did not pass English, and gender. The model identified as having the lowest false-positive rate and a relatively high accuracy for ninth grade was the downsampled QDA model with the lowest false-positive rate of 0.43 and an accuracy of 75%. For sixth grade, the downsampled LDA model had the lowest false-positive rate at 0.34 and an accuracy of 67%.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I wish to thank my committee members, Dr. Pate and Dr. Siegrist, for taking the time to provide me valuable feedback. A special thanks to Dr. Brockmeier, my committee chairman, for his countless hours of insight, reading, patience, and encouragement during this process. Often when I thought I simply could not move forward, he would give me a few words of encouragement that would help me to keep pushing ahead.

I would like to acknowledge and thank Mr. Tony Jones for approving my research within the school district used in this study. Also, thank you to Dr. Anthony Blasingame for pulling the school district's data and coding it to my specifications. The data he provided was the foundation of my research.

I am thankful to God for giving me the strength and wisdom to complete this process. I am blessed beyond measure. Finally, I would like to thank my supervisor, Dr. Eric Payne, and colleague, Mrs. Gretna Soltis, for their interest in my study and encouragement all along the way. They were great cheerleaders and accountability partners and for that, I am very appreciative.

DEDICATION

This work is dedicated to my family. To my husband, Clell, without your support and belief in me, this dream of mine would not have been possible. Thank you for your unconditional love and encouragement all along the way. To my daughters, Kalie, Kara, and Kenley, thank you for your understanding while I tried to balance life, work, and school. You all will always be my greatest accomplishment. To my mother, Argen, thank you for believing in me, instilling the love of learning, and always being my biggest cheerleader. To my sisters, Shannon and Michelle, thank you for being the best sisters and for always being there when I needed you. Finally, to my grandson, Duke, you have brought me immeasurable joy. What a gift you are. I love you all.

Chapter I

INTRODUCTION

With improvements in student information systems over the past decade, there has been a growth of early warning systems that address the dropout crisis more effectively (Jobs for the Future, 2014). An early warning system involves analysis of school data to monitor students at risk for falling off the path to graduation and implementing interventions to help them graduate (Davis, Herzog, & Legters, 2013). Statistical evidence is available to show that dropouts can be identified long before they fail to graduate (Allensworth & Easton, 2007; Balfanz, Herzog, & Mac Iver, 2007; Neild, 2009). A few key data points allow schools and districts to identify those students who are most likely to drop out of high school. As of 2014, the National Governors Association identified 16 states and a growing number of school districts that use some form of an early warning system to predict students, as early as elementary school, who are likely to struggle to graduate from high school within four years (Jobs for the Future, 2014). Most early warning systems include measurements of attendance, behavior, and course performance, also known as the "ABCs," because together they are strong predictors of high school graduation at all levels (U.S. Department of Education, 2016b). Identifying the best combination of early warning system indicators for each user's state, district, school, or grade is key to accurately detecting potential dropouts.

Some benefits of an early warning system include its use of readily available student data, its ability to cull through large amounts of data and focus only on the most important indicators, its use in real-time, and its ability to monitor throughout the school year (Davis et al., 2013). Another advantage of an early warning system is it accurately predicts a high percentage of students who will not graduate from high school within four years based only on academic data instead of background characteristics (Brundage, 2014). School personnel cannot impact all of students' background characteristics, but they can intervene with students' school academic performance. Interventions focused simply on attendance and behavior have been shown to be effective (Balfanz et al., 2007).

When developing and using an early warning system, stakeholder input is essential to improving and refining it. The system should be regularly revisited and risk models revaluated. Before the development of an early warning system, quality research should be performed to identify the best indicators or combinations of indicators. Proper training of the users and clear visuals help ensure effective use of the system. Although there is valuable information about risk factors for dropping out obtained from studies of large districts, such as Chicago and Philadelphia, school systems should look at their longitudinal data to identify factors most strongly associated with their student dropouts (Heppen & Therriault, 2008). To identify who is at risk of dropping out, districts can save time and money by investigating longitudinal data of past cohorts to predict what will happen to students in future cohorts (Jerald, 2006). This personal data can help a system more accurately predict students who are most at risk for dropping out.

The problem of high school dropouts has been studied for decades but, now more than ever, better research and data are available for school leaders to learn both who will likely drop out and what interventions to implement. There are several factors for why a student chooses to drop out, and there are variables along the way that can identify who is likely to drop out (Rumberger & Lim, 2008). Early warning systems could aid dropout prevention by testing local indicators to identify accurately students who are likely to drop out and to aid schools in identifying those who need interventions (Pinkus, 2008). In the past, dropout prevention efforts generated poor results (Jerald, 2006). Accurate and early identification of students at risk for not graduating can lead to appropriate and timely interventions that increase the likelihood of students completing high school (McKee & Caldarella, 2016). An early warning and multi-tiered response system is essential to ending the dropout crisis in schools, districts, and the nation.

Impact of Dropping Out

High school dropouts not only negatively impact themselves, they negatively impact their families and society. Society needs an educated and trained workforce to compete in the world marketplace (Neild, 2009). Members of society without even a high school diploma can become a burden because higher rates of unemployment and higher crime rates are associated with high school dropouts (Alliance for Excellent Education, 2011). During the third quarter of 2019, all full-time workers aged 25 and older had a median weekly income of $975 while full-time workers without a high school diploma earned 62% of that amount. A high school graduate earned 77% of that amount and employees with a bachelor's degree earned 131% of the $975 median weekly income (Bureau of Labor Statistics, 2019). Dropouts also experienced more unemployment, utilized more government aid, or spent more time incarcerated than their peers who graduated high school (Zvoch, 2006). Dropouts also report more health problems, and on average, die at a younger age than those who graduate (Laird, Kienzl, DeBell, &

Chapman, 2007). Students who drop out of high school simply limit their life opportunities and personal wellbeing (McKee & Caldarella, 2016).

The vast evidence of how dropping out of high school significantly impedes a person's quality of life cannot be ignored. Because of the detrimental consequences caused by dropping out of high school, it is imperative that schools take steps to ensure all students graduate. Not only do students' futures depend on it, but society's future as well. The dramatic economic benefit of improving the outcomes of academically at-risk students should be a wake-up call to the nation. With global competition and the grim outlook for dropouts, high schools must keep students in school and prepare them for life after high school, including college, the workforce, or possible military service (Amos, 2008). The social and economic contributions of these young people cannot be underestimated.

Early Warning Signs

Dropping out of high school is more of a process than an event. A student's decision to drop out of high school may be just one more event in a chain that may have begun years before (Finn, 1989). According to Allensworth (2013), before academic data was so prevalent, being able to prevent student dropouts seemed like an insurmountable problem caused by factors out of a school's control. Now, teachers and administrators can control dropout prevention by monitoring students' class performance and social behavior. Allensworth (2013) also stated that by reaching out to struggling students and keeping track of students who could fail to graduate, teachers and administrators can strategically target those who need help. This makes dropout prevention a school problem where the staff takes ownership, and it is not a problem that the staff feels it cannot control (Allensworth, 2013). Evidence lends weight to the argument that readily available data could give schools some control over their dropout problem. Statistical evidence is available to show that students who dropped out began their disengagement from school for reasons rooted in either academic struggle, including failure, or reactions to the school environment, such as high absenteeism (Balfanz et al., 2007). Jerald (2006) stated that the first step to reducing dropout rates is identifying the students most at risk for dropping out.

Much of the research for dropout identification has focused on the ninth-grade year because of the difficulty adolescents have transitioning to high school. Ninth grade is a critical juncture in a child's education because it is the time when the student's lack of knowledge and skills catches up with him or her (Neild, 2009). Students who struggled with reading and math in middle school become overwhelmed with the demands of high school, become discouraged and then truant, and may eventually drop out (Neild, 2009; Sparks, Johnson, & Akos, 2010). Also, part of what makes the transition so difficult is not just the adolescent age of the students, but the substantial academic and social differences between middle school and high school (McCallumore & Sparapani, 2010). The transition to a new school with a new environment, new classes, and new teachers, all while transitioning in their own emotional development, makes for a tough first year of high school (McIntosh, Flannery, Sugai, Braun, & Cochrane, 2008). When it comes to predicting the future graduation of ninth-grade students, academic performance has proven critical. Statistical evidence is available to show that when ninth graders fail to obtain enough credits to be promoted to the tenth grade, they have very low odds of graduating high school (Allensworth & Easton, 2005). Editorial Projects in

Education (EPE) Research Center analyzed student enrollment and diploma patterns, and in 2006 they reported that approximately 35% of the nation's high school students who dropped out were never promoted out of the ninth grade ("The High School Pipeline," 2006). Even more staggering, based on data reported in the household survey of education and economic indicators, the Current Population Survey (CPS), from 1996 to 2003, Black and Hispanic ninth graders were more than twice as likely as their White peers to repeat ninth grade (Neild, 2009). With performance during the first year of high school often being a deciding factor in whether or not a student drops out, reform initiatives should include the middle to high school transition and moving through ninth grade (Cohen & Smerdon, 2009).

Bowers, Sprott, and Taff (2013) examined dropout prediction literature from the last 30 years to identify which predictors of high school dropouts are most accurate and usable by schools and systems. They found that often the indicators used in early warning systems do a poor job of balancing the tradeoff between accurate identification of potential dropouts and false alarms. A possible reason why dropout indicators do a less than ideal job of accurately identifying potential dropouts is poor identification methods. This trade-off has rarely been addressed in dropout indicator literature in the past (Bowers et al., 2013). It is important for educators to know the difference between risk factors that are generally correlated with dropping out and those that predict dropouts well. Exhibiting a dropout risk factor means a student is in a group that is more likely to drop out but does not mean that a particular student will drop out (Jerald, 2007). An evaluation of federally funded dropout interventions found that programs often enroll the wrong students and serve those who would not have dropped out and do not serve those who would have dropped out (Dynarski & Gleason, 2002). Improper identification could lead to wasted time, money, and lives (Jerald, 2007). In this era of technology, schools do not need more data; they simply need the actionable information to help with applying effective interventions to drive real change. Here is where a good early warning system can be a game-changer for schools, districts, and states.

Statement of the Problem

The consequences of dropping out of high school can last a lifetime and create problems for the student, the student's family, and society (Lemon & Watson, 2011). High rates of dropouts in a community lead to social and economic woes because dropouts are more likely to be unemployed, use public assistance, commit crimes, and be incarcerated. They are also less likely to have health insurance, stay healthy, perform civic duties, and contribute to the tax base (Jerald, 2007). The United States had approximately 650,000 high school dropouts in 2014, and 33,000 of those dropouts were from Georgia (Alliance for Excellent Education, 2015). A problem with ending the dropout crisis is identifying those students who are likely to drop out. Identifying students at risk for dropping out requires finding the best dropout indicators at the appropriate time in a student's academic career.

Districts can take the first step in solving the problem by identifying likely high school dropouts using an early warning system. An early warning and multi-tiered response system is essential to ending the dropout crisis in schools, districts, and the nation. Creating an accurate dropout early warning system involves finding the best combination of indicators considering both true-positive and false-positive results and identifying when it is best to begin identifying the students. Nationwide, about half of public high schools implemented an early warning system in the 2014-2015 school year (U.S. Department of Education, 2016b). Presently, the state of Georgia does not offer a statewide dropout early warning system, but having one could help solve the problem of identifying students most at risk for dropping out of high school. Having a dropout early warning system with high out-of-sample predictive power could ease the burden of identifying at-risk students for schools and districts and allow them to focus on implementing effective interventions to put students back on the path to graduation. The dropout early warning system could play a significant role in helping school systems in Georgia raise their graduation rates and improve the lives of thousands of individuals for years to come.

Purpose of the Study

The purpose of this study is multifaceted. The foremost purpose of this nonexperimental, correlational study was to use data from a mid-sized school district from two cohorts to support the creation of a dropout early warning system to predict both middle school and high school students who are at risk for not graduating on time. Both sixth-grade and ninth-grade longitudinal data were used in this study.

This study identifies the most accurate indicators at each respective grade level that result in high levels of true classification and low levels of false identification at all middle schools and high schools in a mid-sized school district. Also, this study will identify the most accurate statistical models necessary to detect the most accurate indicators.

Research Questions

The research questions framing this study are:

1. Does one or more of the ninth-grade variables consisting of attending less than 90% of the time, earning sufficient credits to move to the tenth grade, receiving out-of-school suspension, number of school moves, standardized reading and math scores, failing no more than one semester of a core content course, school minority percentages, school poverty percentages, English Learner Status (ELL) status, Students with Disabilities (SWD) status, free/reduced meal status, race, and gender accurately predict students who will not complete high school within four years?
2. Which statistical model is the most accurate at predicting future dropouts or late graduates utilizing ninth-grade variables with high levels of true classification and low levels of false identification?
3. Does one or more of the sixth-grade variables consisting of failing English, failing math, attending less than 80% of the time, receiving out-of-school suspension, number of school moves, standardized reading and math scores, school minority and poverty percentages, ELL status, SWD status, free/reduced meal status, race, and gender accurately predict students who will not complete high school within four years?
4. Which statistical model is the most accurate at predicting future dropouts or late graduates utilizing sixth-grade variables with high levels of true classification and low levels of false identification?

Research Methodology

The statistical procedures utilized in the study included logistics regression, linear discriminate analysis, and quadratic discriminate analysis to identify the most accurate indicators at both the ninth-grade and sixth-grade levels. The goal was to use indicators and models that resulted in high levels of true classification and low levels of false identification outside of the training data used. Quantitatively examining whether a dropout early warning system can accurately predict which students are likely to drop out or not graduate on time will provide opportunities for interventions and potentially have meaningful life-changing results for some students.

Independent variables for this study were chosen based on a review of dropout indicators at both the ninth-grade and sixth-grade levels and, thus, were used as predictor variables in this study. Two different types of variables were analyzed during this study: first, those which schools have a direct relationship with and can impact, and second, those outside the influence of the local school. The most accurate ninth-grade dropout indicators that are both influenced and not influenced by schools were: attending school less than 90% of the time, earning sufficient credits to move to the tenth grade, number of days suspended out of school, number of school moves, standardized reading and math scores, failing no more than one semester of a core course, school minority percentages, school poverty percentages, and student characteristics, including ELL status, SWD status, free/reduced meal status, race, and gender (Allensworth & Easton, 2005; DePaoli et al., 2015; Kemple, Segeritz, & Stephenson, 2013; Lee, Cornell, Gregory, & Fan, 2011; Mac Iver & Mac Iver, 2010; Mac Iver & Messel, 2012, 2013; Neild, 2009; Zvoch, 2006).

The most accurate sixth-grade dropout indicators in these two categories were: failing English, failing math, attending less than 80% of the time, receiving out-of-school suspension, number of school moves, standardized reading and math scores, school minority percentages, school poverty percentages, and student characteristics, including ELL status, SWD status, free/reduced meal status, race, and gender (Balfanz, 2009; Balfanz et al., 2007; Jerald, 2006; Mac Iver, 2010; Rumberger, 2004; Silver, Saunders, & Zarate, 2008).

The dependent variable is on-time graduation. A relationship among the set of variables forecasted students at risk of not graduating within four years of entering high school or dropping out. If the model predicted accurately, future interventions could be developed, used, and tested by the school system.

The source of information for this study was longitudinal data gathered from two cohorts of students from all six high schools in a mid-sized school system in Georgia. These two cohorts had on-time graduation dates of 2017 and 2018, respectively. Data was pulled from both cohorts' sixth-grade and ninth-grade academic years. The total number of students in the study was approximately 1,800, while the entire district had over 20,000 students in each of the cohort years. All required sixth-grade and ninth-grade student data, except school minority data, school poverty data, and free/reduced meal status, were maintained in the Infinite Campus (IC) student data system. The Georgia Department of Education College and Career Performance Index (CCRPI) website contained information pertaining to school minority and poverty data (Georgia Department of Education, n.d.). Finally, student free/reduced meal data were housed in the Georgia School Nutrition database, Websmart, which could only be accessed by the district's school nutrition staff.

The models used for this study predicted students who were likely to drop out of high school or not graduate within four years by identifying the variables that yield the best performance. This research analysis used correlational methods to indicate the relationships between the predictor data and the criterion variable of students being either an on-time graduate or not.

Checks for outliers and missing data are essential for accurate data analysis. Outliers can cause severe problems that even the robustness of discriminant analysis cannot overcome. Therefore, data in this study were screened for outliers using Cook's distance. Cook's distance is used for regression analysis to find influential outliers in the predictor variables. Additionally, missing data could undermine the ability to make valid inferences for this study. Therefore, the missing data were imputed.

One of the statistical procedures that was used in this correlational study was logistic regression. Logistic regression was used to examine the relationship among the variables to yield a maximum correlation to predict which students would drop out or not finish within four years (Ary, Jacobs, Sorensen, & Walker, 2014). Assumptions for logistic regression include observations must be independent, and independent variables must be linearly related to the logit of the dependent variable (Leech, Barrett, & Morgan, 2008). The dependent variable must be dichotomous, and the independent variables must be continuous or categorical.

Two other alternative and widely utilized statistical procedures employed in this study were linear discriminate analysis (LDA) and quadratic discriminate analysis

(QDA). Linear discriminant analysis and quadratic discriminant analysis help to find the boundaries around the classification choices (James, Witen, Hastie, & Tibshirani, 2013). For the multiple variables, the models estimate the mean and variance from the data for each class. Both the LDA and QDA algorithms make predictions by estimating the probability that a new set of inputs belongs to a particular class or group. The class with the highest probability is the output class, and therefore, the prediction (James et al., 2013). LDA and QDA have assumptions that are often more restrictive than logistic regression. Both LDA and QDA assume that the predictor variables are drawn from a normal distribution. LDA assumes equality of covariances among the predictor variables X across all levels of Y, but QDA does not. Both LDA and QDA require the number of predictor variables to be less than the sample size ("Linear & quadratic discriminant analysis," n.d.).

Significance of the Study

Creating an accurate dropout early warning system is a critical component to combatting the dropout epidemic. This study identifies both middle school and high school students likely to drop out or not graduate within four years of entering high school. Identification of those students could allow schools to provide interventions to get them on track for on-time graduation. By accurately identifying the students most at risk for not graduating on time, schools could use their limited resources effectively. Dropout prevention initiatives are expensive and should not be wasted on students who do not require an intervention. In addition, by not providing an intervention to future dropouts who need it is a disservice because they could have been saved.

A secondary significance is the use of a dropout early warning system throughout a district, and potentially the state could help increase the graduation rate and, thus, improve the lives of thousands of people. This improvement could come in the form of higher earned income, less crime, higher employment rates, and even better quality of life for future generations. The most compelling reason to focus on early identification is that the benefits can be far-reaching for society and make a better quality of life in general. After all, education is central to the well-being of society.

Conceptual Framework of the Study

The conceptual framework for this study includes aspects of a student's academic and social engagement and background characteristics that influence academic achievement. Students with a strong educational foundation do not misbehave at school, they have high attendance, and they pass all their classes (Mac Iver & Mac Iver, 2009; Rumberger & Larson, 2007). The theory that students' educational performance is influenced by their background characteristics and engagement in school, as well as their families, their schools, and their communities is the foundation of the framework shown in Figure 1 (Rumberger, 2004).

Although each variable's influence on identifying potential dropouts varies, each plays a role in this important task. Sometimes, indicator systems will use a single variable. Still, often, it is a combination of indicators that are best associated with the likelihood of dropping out and increase the predictive accuracy. Successful early warning systems track multiple variables that are shown to be related to students' likelihood of dropping out (Heppen & Therriault, 2008). Some of these variables include poor grades in core subjects, high absenteeism, and failure to be promoted (Kennelly & Monrad, 2007).

The most commonly used student-level variables to make up an indicator are attendance, failing grades in core courses, misbehavior, credits earned, and promotion (Allensworth & Easton, 2005; Balfanz, 2009; Balfanz et al., 2007; Mac Iver & Mac Iver, 2010; Mac Iver & Messel, 2012). Less common, but noted in the literature are: socioeconomic status, special education status, English language learner status, and standardized test scores (Allensworth & Easton, 2005; DePaoli et al., 2015; Gwynne, Lesnick, Hart, & Allensworth, 2009; Rumberger, 2004; Zvoch, 2006).

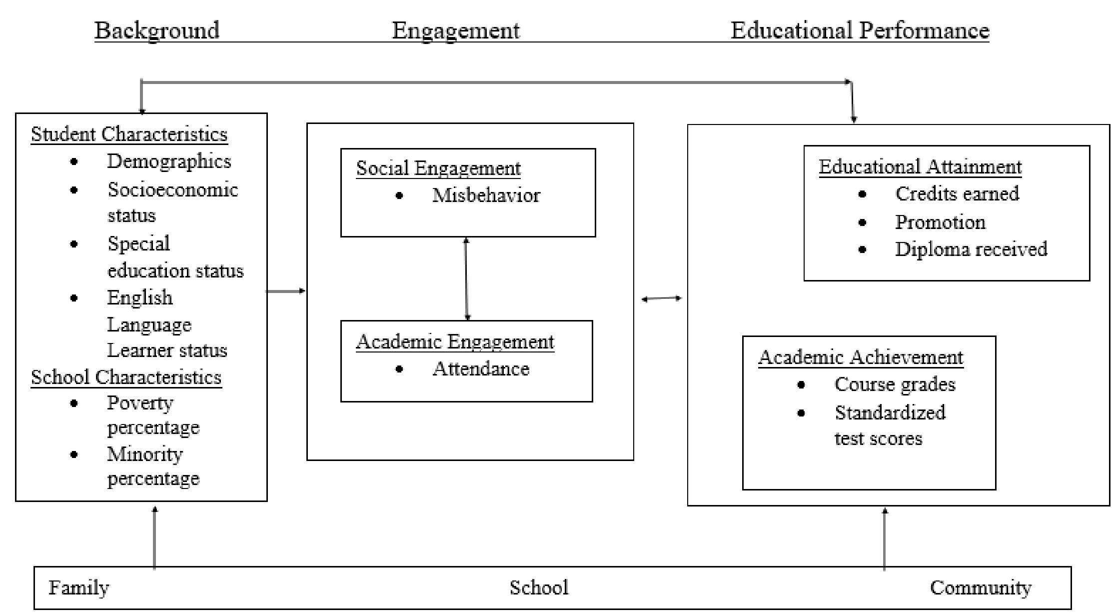

Figure 1. Students' educational performance is influenced by their background characteristics and engagement in school, as well as their families, their schools, and their communities.

Limitations of the Study

There were several limitations to this study, and they began with record keeping. Identification of actual dropouts likely had mistakes due to misinformation from families when they withdrew their student from school. Also, a simple human error could misrepresent the data. Data entry is often dependent upon human entry, resulting in likely errors. For example, a student could have been identified with the incorrect race, incorrect socioeconomic status, inaccurate attendance, or incorrect behavior infractions.

New access restrictions enacted by the federal government in 2016, put major limitations on who has access to free and reduced meal data, so it is not as readily accessible as it once was. Another limitation was that the analyses of this study may not apply to other districts. The overall demographics of the school district used in this study impacted which characteristics were more prevalent for graduating late or dropping out of school and not have as high of a predictive yield in other districts. Another limitation of this study was the analysis of all the student indicators connected with dropping out in the literature is beyond the scope of this study. Only those indicators most easily collected and those that capture the most dropouts were used. Lastly, some assumptions were not met for the statistical analyses used in this study.

Definition of Terms

The following terms or phrases are used in this research topic and are provided for consistency.

Adjusted Cohort Graduation Rate (ACGR). The *ACGR* is required by the U.S. Department of Education and is based on the year a student enters high school as a freshman. This rate is calculated by dividing the number of students in the cohort who graduate with a regular high school diploma within four years by the number of students who form the adjusted cohort for the graduating class. The cohort is adjusted by adding any students who transfer in and subtracting any students who transfer out (DePaoli et al., 2015).

Cohort. Students who begin school together for the first time form a *cohort*, and it can be adjusted by adding students who transfer into the cohort later.

Early warning system. An *early warning system* is based on indicators such as attendance, behavior, and course performance and helps schools identify students at risk for not graduating on time.

False-negative rate. The *false-negative rate* is the percentage of students not identified at risk of dropping out by the early warning system, but actually do drop out.

False-positive rate. The *false-positive rate* is the percentage of the graduates identified at risk of dropping out by the early warning system, but actually do not drop out.

Machine learning. *Machine learning* is a method used to create complex models and algorithms that are suitable for making predictions. Machine learning explores the construction of algorithms that utilize data and make predictions. The algorithms make data-driven predictions by building a model from sample inputs.

Out-of-sample predictive power. The ability of a dropout early warning system to predict accurately potential dropouts or students who will not complete high school in four years outside of the sample used to create the system is the *out-of-sample predictive power*.

Precision. The percentage of students identified at risk of dropping out who dropped out is *precision*.

Receiver operation characteristic (ROC) plot. A plot of true-positive rate and true-negative rate and the closest point to the left-upper corner of the plot is considered the most accurate.

Test set. The *test set* is used when evaluating the performance of the estimator of an algorithm.

Training set. The *training set* is data used to train or create the estimator of an algorithm.

True-negative rate/specificity/false alarm. The percentage of all the dropouts not identified as at risk of dropping out is known as *true-negative rate, specificity, and false alarm*.

True-positive rate/sensitivity. *True-positive rate or sensitivity* is the percentage of all the dropouts identified as at risk of dropping out.

Organization of the Study

This study of predicting likely dropouts and students who will not complete high school within four years using sixth-grade and ninth-grade indicators is organized into five chapters and a reference list. Chapter 1 states the problem to be analyzed and the purpose of the study. The research questions and methodology are listed, followed by the significance of the study, and conceptual framework, along with limitations and definitions of the study. A review of the related literature pertaining to: (a) the dropping out of high school, (b) dropout early warning systems, (c) indicators used to identify students at risk for dropping out or not graduating within four years, and (d) the grade levels best able to predict dropouts and late graduates are described in detail in Chapter 2. Chapter 2 also describes major studies related to dropout early warning indicators and their findings. A detailed description of the importance of looking at all aspects of an early warning system's accuracy is provided. Chapter 3 describes the research design and methodology for the study. The indicators used to predict likely dropouts and late graduates, the grade levels selected to pull the indicator data, and data collection procedures for the study are described in detail. Chapter 3 also describes the various statistical analysis options available based on the major studies in the literature. Specific statistical tests available for use in the analysis and their assumptions are described. The district characteristics, student characteristics, and data analyses are also described in detail. A description of the results of the statistical analyses for each grade level and the results of each statistical model used to answer the four research questions are provided in Chapter 4. The significance of the study and the implications of the results based on the research questions are described in Chapter 5.

Chapter II

LITERATURE REVIEW

The review of literature provides an examination of individual and school-level factors that are used by early warning systems at middle and high school levels to prevent dropouts. This chapter continues with describing studies that identify the benefits of utilizing an early warning system in the ninth grade and additional studies that describe the benefits of utilizing early warning systems in the sixth grade. Finally, the review of literature will conclude with descriptions of early warning systems currently in use in a variety of states and the indicators utilized within them.

National Graduation Rate

Since 2011, the nation and states have calculated their graduation rates using the adjusted cohort graduation rate (ACGR). This method is required by the U.S. Department of Education, and the ACGR is based on the year a student enters high school as a ninth grader. This rate is calculated by dividing the number of students in the cohort who graduate with a regular high school diploma within four years by the number of students who form the adjusted cohort for the graduating class. The cohort is adjusted by adding any students who transfer in and subtracting any students who transfer out (DePaoli et al., 2015).

With a renewed focus on academic achievement and the graduation rate in recent years, the national ACGR has increased each year since its inception in 2011. The national ACGR had increased from 79% in 2011 to 84% in 2016. In 2016, the state level

ACGR ranged from a low of 69% in the District of Columbia to a high of 91% in Iowa. In 2016, the national ACGR for White public school students was 12% higher than their Black peers at 88% compared to 76% ("Public High School Graduation Rates," 2018).

Georgia's Graduation Rate

Georgia's graduation rate has continued to climb since the inception of the adjusted cohort rate. Although all states must calculate the adjusted cohort rate using the same formula, each state sets its own requirements for students to earn a regular high school diploma. Georgia has some of the toughest requirements in the nation for a student to earn a regular high school diploma (Frick, 2018). Georgia's adjusted cohort graduation rate has steadily increased from 72.5% in 2014 to 78.8% in 2015, 79.4% in 2016, 80.6% in 2017, and 81.6% in 2018 (Frick, 2018). Even with Georgia's seemingly high graduation rate, in 2014, Georgia had the fifth-lowest graduation rate in the nation (U.S. Department of Education, 2016a). The graduation rate for the mid-sized school systems in this study has continuously climbed as well but has remained below the state average each year. Its adjusted cohort graduation rate had steadily increased from 58.9% in 2014 to 71.2% in 2015, 71.6% in 2016, 77.1% in 2017, and 78.5% in 2018 (Frick, 2018; Jones, 2016; "The Governor's Office of Student Achievement," 2018).

Consequences of Dropping Out

Dropping out of high school is a serious problem, not only for the individual but the school system, community, and society, as well. A key strategy for economic growth is addressing the high school dropout crisis (Tucci, 2011). Improving education not only benefits the individual, but the individual gains compound to benefit the economy at the local, state, and national levels. Economic benefits include increased tax revenues, increased individual earnings, increased sales of homes and autos, more jobs, and economic growth, along with more spending and investment (Tucci, 2011). Tucci calculated that if only half of the dropouts in 2010 graduated, the economic benefits would have included more than $7.6 billion in increased annual earnings and the creation of 54,000 new jobs.

Additionally, the net public benefit of raising the level of high school graduates is substantial. A 2007 study found that each high school graduate provides a public benefit of $209,000 in higher government tax revenue and lower government spending on health, crime, and welfare (Levin, Belfield, & Rouse, 2007). If the dropout rate for one cohort were cut in half, the government would gain $45 billion in increased tax revenue and lower costs of public health, crime, and welfare payments (Levin et al., 2007). It is easy to conclude that it is an economic necessity to help more students to graduate from high school.

Individuals who drop out of high school have fewer options for employment and often work at low-paying jobs with few opportunities for advancement (Christle, Jolivette, & Nelson, 2007). A report on high school graduates in the year 2000 stated that 56% of high school dropouts were unemployed, while only 16% of high school graduates were unemployed (Stanard, 2003). Employment rates vary across race-ethnic groups, as well. In 2008, Black dropouts were the least likely to be employed with a high 69% unemployment rate compared to a 46% unemployment rate for White dropouts (Sum, Khatiwada, & McLaughlin, 2009). Students who failed to graduate were more likely than their peers who graduated to be unemployed, live in poverty, receive public assistance, spend time in prison, be on death row, have poor health, and be a single parent with children who also drop out (Bridgeland, Dilulio, & Morison, 2006).

Statistical evidence by Sum et al. (2009) showed that specific groups of dropouts had bleak data, including young female dropouts who were six times as likely to have given birth and nine times as likely to be a single mother as their counterparts who were college-educated. These same females were likely to be poor and dependent upon government assistance to support themselves and their children. Sum et al. (2009) also found that high school dropouts have more than 63 times higher rates of incarceration than those with a bachelor's degree. Also, in 2006-2007, nearly one out of every 10 young male high school dropouts was incarcerated compared to one out of 33 high school graduates (Sum et al., 2009).

Individual and School Level Factors

The nation has focused attention on improving the graduation rates for all students, including those who have been underserved in the past or who present specific learning challenges (DePaoli et al., 2015). The increased attention from social, political, and governmental agencies has created pressure for educators to identify and intervene with students who seem likely not to graduate from high school (Bruce, Bridgeland, Fox, & Balfanz, 2011). Part of this pressure is the creation of the GradNation campaign that has set a common goal of a 90% national high school graduation rate by the year 2020. With the GradNation campaign, states must show real progress each year in meeting graduation goals (DePaoli et al., 2015). By identifying potential dropouts earlier, research-based interventions can be put in place for those identified students to prevent them from dropping out and to work to get them on track to graduate.

Key indicators associated with higher levels of dropouts must be analyzed when creating an early warning system. Key researchers in the field of dropout indicators from the Consortium on Chicago School Research at the University of Chicago, the University of Pennsylvania, and the Center for Social Organization of Schools at Johns Hopkins University have discovered that to identify who is most likely to drop out, schools need to identify who: (a) receives poor grades in core subjects, (b) has low attendance rates, (c) fails to be promoted to the next grade level, and (d) has behavior problems (Allensworth & Easton, 2005; Balfanz, 2016; Kennelly & Monrad, 2007; Neild, 2009). For students in high-poverty environments, having even one of these indicators from the sixth to ninth grades typically means having a 25% chance at best of graduating from high school (Balfanz, 2011). These indicators are considered to be better predictors of dropouts than fixed indicators, such as gender, race, and low socioeconomic status, although the fixed indicators are often associated with dropping out (Jerald, 2006; Rumberger, 2004). English language learner status, special education status, and standardized test scores are additional indicators to analyze in an early warning system because of their association with higher levels of dropouts (Jobs for the Future, 2014).

Academic Success

A Chicago study of data from both the 2001 and 2004 freshman cohorts, which included 23,564 and 26,562 students, respectively, revealed that ninth-grade-year course performance is strongly linked to high school graduation. To be considered on track, a student must have accumulated five full course credits, the number required to be promoted to the tenth grade, and receive no more than one-semester grade of F at the end of their freshman year (Allensworth & Easton, 2005). This on-track indicator is highly predictive of whether freshmen will eventually graduate. Students who were on track by the end of their ninth-grade year were more than three and one-half times more likely to graduate high school in four years than off-track students. Also, 81% of the students on track at the end of their freshman year graduated from high school in four years, but only 22% of the off-track students graduated in four years. On-track students were three times as likely to graduate within five years as off-track students (Allensworth & Easton, 2005). There is a need for interventions to ensure that more ninth graders pass and earn the credits needed to stay on track to graduate because waiting until the end of ninth grade to intervene is often too late (Mac Iver, 2013).

The Chicago study by Allensworth and Easton (2005) utilized a two-level hierarchical generalized linear model to predict graduation using student-level data including: on track at the end of freshman year, scores on the Iowa Tests of Basic Skills, race or ethnicity, gender, age of entering high school, level of neighborhood poverty from Census, and dummy variables. The relationship between being on track and graduating remained the same, even after accounting for the other differences between students (Allensworth & Easton, 2005).

In the Kemple et al. (2013) study, longitudinal data from 10 cohorts of first-time ninth-grade students in New York City between 2001 and 2010 were compiled to conduct a systematic analysis of several on-track indicators that predict the likelihood of students graduating from high school. A sample of more than 576,000 first-time ninth graders from more than 350 New York City high schools was used. This study considered both the overall correct prediction rate and the correct prediction rates for both on-track and off-track statuses. The New York City Department of Education (NYCDOE) considers a key ninth-grade on-track indicator as earning 10 or more credits in the ninth grade (44 credits are required for graduation). Earning 10 or more credits in the ninth grade was a reasonably accurate overall predictor that a student would graduate with a diploma within four years. When also considering if the student passed at least one end-of-course exam in the ninth grade along with earning 10 or more credits, together those two indicators showed an even higher degree of stability in the relationship with students' actual graduation status. Using both indicators together, the overall correct prediction rate increased to upwards of 82% (Kemple et al., 2013). A majority of this study's analysis involved a percentage of accuracy for the correct prediction of graduation or not. In addition, the significance was high, with $p = 0.05$ at 0.03 using logistic regression. Especially for the later cohorts, having earned 10 or more credits and passing at least one end-of-course exam in ninth grade provided a compelling prediction of the future graduation rate for successive cohorts of first-time ninth graders (Kemple et al., 2013).

A study of Baltimore City Public Schools in both the 2005 (n = 6,812) and 2006 (n = 7,729) school years set out to discover the extent to which ninth-grade, school-level factors are associated with non-graduation outcomes, including student attendance, behavior problems, and course failure (Mac Iver & Messel, 2012, 2013). Researchers focused on factors that schools can easily monitor and address through interventions aimed at changing students' behaviors and actions before they drop out. This study followed the two freshmen cohorts until their graduation year and one-year past. A dichotomous variable of displaying any early warning indicator in the ninth grade, including chronic absenteeism, being suspended for at least three days, and failing a core course was used to code students. Students with any of the three indicators were coded as a 1, and students with no indicator were coded as a 0. Graduation was also coded as a dichotomous variable with students being coded as either a graduate or a nongraduate (Mac Iver & Messel, 2012, 2013).

In the Mac Iver and Messel (2012, 2013) study, multivariate analysis of longitudinal student cohort data was used to analyze the impact of early warning indicators of dropping out identified in earlier research. For this study, logistic regression hierarchical linear models indicated that course failure and attendance in the ninth grade were the two most significant predictors of graduation. At level 1, the student level, student outcomes were modeled as a function of demographic characteristics and early warning indicators (attendance, suspensions, and course failure). At level 2, the school level, the impact of school characteristics were estimated on student outcomes. Analyses of graduation or nongraduation outcomes followed a pattern of analyses first with only demographic variables, then sequentially adding behavioral variables, then school-level variables (Mac Iver & Messel, 2012, 2013).

This study found that in the 2005 school year, 86% of those who passed all their ninth-grade core courses were on-time graduates. Of the students in the same cohort who failed two or more ninth-grade core courses, only 30% graduated on time (Mac Iver & Messel, 2012, 2013). Identifying and providing interventions for students who are failing core subjects is critical when trying to raise the national graduation rate. Identifying students failing courses is available in real-time by utilizing a school's student information system.

An additional study by Mac Iver (2010) of student-level data files from Baltimore City Public Schools using 2008-2009 graduate and dropout data took a different approach. Mac Iver's retrospective approach differed from more traditional cohort analyses in that it focused on all students with a particular outcome (dropout vs. graduate) in a particular year, and then followed them backward in time through district records. There were 1,646 students with dropout codes. On average, dropouts had accumulated only 5.2 credits over their entire high school career, compared to an average of 24.6 credits for graduates. A third of the dropouts did not earn any credits, and another quarter had earned between only one and five credits. Dropouts also differed dramatically from graduates in their first-time ninth-grade course performance. A majority (85.4%) of the dropouts failed at least one core course during their ninth-grade year, compared to roughly a third (36.8%) of the 2008-2009 graduates. Almost three in four dropouts (72.1%) failed at least two core courses, compared to only 18.1% of graduates with the same ninth-grade failure rate. Nearly half (43.6%) of dropouts failed four or more core courses in ninth grade compared to only 3.6% of graduates. Additionally, dropouts earned an average of 2.3 credits during their first ninth-grade year, compared to 5.6 for graduates (Mac Iver, 2010).

Balfanz, et al. (2007) conducted a study of a Philadelphia cohort of approximately 13,000 students who entered the sixth grade in September 1996. Their study focused their initial efforts on identification, prevention, and intervention on sixth graders because it was an official start of the middle grades. In addition, sixth graders must adapt to a variety of changes such as more departmentalized staffing, larger class sizes, different assessments, grading, testing, and more challenging instructional programs (Epstein & Mac Iver, 1990). In the Balfanz et al. (2007) study, the students were followed until one year after their expected graduation date. Longitudinal student data collected included attendance, demographic data, test data, and course credits. This data helped to identify flags that had both high predictive power and high yield. A flag had high predictive power if 75% or more of the sixth graders who were flagged did not graduate within one year of their expected graduation date. A flag would have a high yield if it identified 10% or more of future nongraduates. One goal of the study was to discover how early in middle grades a significant number of students in a high poverty school district—absent interventions—will fall off the path to graduation. Another goal of the study was to identify reliable indicators that are readily available to school personnel to identify students at risk of not graduating.

In the Balfanz et al.'s (2007) study, multivariate logistic regression techniques were used to identify which variables had both significant and independent predictive power after controlling for other flags and demographic variables. Five flags identified as having both high predictive power and high yield were: (a) attending school 80% or fewer days in the sixth grade; (b) failing math in the sixth grade; (c) failing English in the sixth grade; (d) being suspended while in sixth grade; and (e) receiving an unsatisfactory behavior grade in any course in the sixth grade. Each flag was statistically significant as a predictor ($p < .0001$) even after controlling for the other flags and race. The study found that students with chronic absenteeism are 68% less likely to graduate than students without chronic absenteeism. Students with unsatisfactory behavior are 56% less likely to graduate while students who fail math or English are 54% and 42%, respectively, less likely to graduate when compared to students who did not. Sixth-grade students with one or more of the indicators listed above may have only a 15% to 25% chance of graduating from high school on time or within one year of their expected graduation date (Balfanz & Fox, 2011). Using the five sixth-grade warning flags, 60% of nongraduates could be identified (Balfanz et al., 2007).

Schools staff members should pay special attention to students who send distress signals, such as high absenteeism and course failures in sixth grade because the earlier a student first sends a signal, the greater the risk that the student will drop out of school (Balfanz et al., 2007; Neild, Balfanz, & Herzog, 2007). Once a sixth grader demonstrates that they are unable to pass tests in math or English or lack the ability to complete assignments, the situation is unlikely to change on its own without an intervention. As a result, the student would continue to fail classes and might not be promoted to the next grade level. The student would then enter high school overage with a history of failure. Because the student would lack the skills, knowledge, and self-confidence to succeed in high school, the student would continue to fail and likely not earn a promotion to the tenth grade. By this time, the student would have reached the legal age to drop out of school (Balfanz, 2009).

The Baltimore Education Research Consortium (2011) studied 7,887 sixth-grade students in the 2000-2001 cohort and 5,816 sixth-grade students in the 2008-2009 cohort from Baltimore City Schools. Baltimore's decision to focus on dropout prevention was essential, and an early warning system helped inform prevention efforts. This study looked for indicators that could predict eventual dropouts with a reasonable level of certainty so that interventions could be put into place. Logistic regression analysis was used to determine which indicators were statistically significant, highly predictive, and practically meaningful. In addition, only variables with the highest yield were used. The variable criteria included projects with at least 70% probability that a student with the indicator will not graduate, and more than 20% of the students who eventually dropped out possessed the indicator. Four predictive sixth-grade early warning indicators for not graduating are: (a) missing 20 or more days of school; (b) failing English or math or both and/or failing all four core subjects; (c) being at least one-year overage; and (d) being suspended for three or more days. Failing a core course in sixth grade is strongly associated with a lower likelihood of graduating from high school. Less than a third of sixth graders who failed a core course in the study eventually graduated. Those who failed both English and math were particularly unlikely to graduate, as only 18.9% eventually graduated (Baltimore Education Research Consortium, 2011).

One purpose of the Los Angeles Unified School District's (LAUSD) seven-year longitudinal data set for the class of 2005 by Silver et al. (2008) was to find the individual student's and school's characteristics at both the middle and high school levels were associated with dropping out. The trajectories leading students to either high school graduation or dropping out begin years before those events happen. The causes of dropping out are many and complex. Identifying relevant school-related causes requires a comprehensive analysis of longitudinal district, school, and student data. This study's cohort consisted of 48,561 students who attended 163 LAUSD middle and high schools in the second-largest school district in the country. A series of multilevel logistic regression models were used to determine the independent contribution of several student-level and school-level factors. This study found that 69% of the students who graduated on time had never failed a class in middle school. These data were compared to less than a 50% graduation rate for students who failed at least one course in middle school. In fact, students who did not graduate on time failed four times as many courses in middle school as those who did graduate on time. Student characteristics such as gender, race or ethnicity, language, and socioeconomic status explained only 4% of the student-level variation in graduation rates. The chance of graduating dropped to less than half for students who were absent more than 10 days per year in seventh or eighth grade or while in high school. Finally, 42% of the between-school variation is explained by the school characteristics of the percentage of qualified teachers, the percentage of English language learners, and magnet school status (Silver et al., 2008).

Balfanz (2009) wrote a data brief summarizing more than a decade of research and development work at the Center for the Social Organization of Schools at Johns Hopkins University. This research was based on direct field experience in more than 30 middle schools in collaboration with the Philadelphia Education Fund. The purpose of the research of Philadelphia sixth graders through one year past their on-time graduation date was to find out how early in the middle grades there were clear indicators that students had fallen off the path to high school graduation and to identify those high-yield indicators. The high-yield indicators identified those students with 25% or lower graduation rates and identified 25% or more of future dropouts. This research found four indicators that identified sixth graders who met these criteria. Sixth graders who failed math, failed English, attended school less than 80% of the time, or received a poor final behavior grade in a course had only a 10% to 20% chance of graduating on time. Also, less than one out of every four sixth graders who had at least one indicator graduated within one extra year of their on-time graduation date. Together, the four indicators identified 59% of nongraduates in the cohort (Balfanz, 2009).

Low Attendance Rates

Varying critical attendance thresholds during transition years in the sixth and ninth grades help signal that a student is at risk of falling off the graduation path. The previously mentioned study of Baltimore City Public Schools in both the 2005 (n = 6,812) and 2006 (n = 7,729) school years involved research to determine if ninth-grade school-level factors were associated with non-graduation outcomes, including student attendance. A student was considered chronically absent if the student missed more than 20 days. Data from two freshmen cohorts revealed that attendance played a significant role in who graduated on time and who did not. While 82% of first-time ninth graders with attendance of 95% or higher graduated on time, only 26% of those who missed more than 20 days graduated (Mac Iver & Messel, 2012, 2013).

An analysis by Neild, Stoner-Eby, and Furstenberg (2008) took a sample of 1,457 ninth graders in the Philadelphia public school system during the 1997 school year and examined whether or not ninth grade represented a particularly vulnerable time for students on the path to graduation while controlling for a wide range of demographic factors. Specifically, did ninth-grade course failure and attendance add substantially to the ability to predict dropouts? Student data, including the students' ages, race, gender, percentage of courses a student received a grade of a D or an F in ninth grade, and percentage of days attended in ninth grade, were pulled from school records. A series of logistic regression models estimated the effects of various characteristics on dropping out, including the increased or decreased odds of dropping out based on each variable. Weighted data accounted for oversampling of students from smaller schools. This study found that each additional week of school attended decreased the dropout odds by 7%.

The negative experiences of ninth grade contributed considerably to the probability that a student will drop out. Particularly, ninth-grade attendance has a substantial impact on the probability of dropping out within six years of starting high school (Neild et al., 2008).

A study of the 2010 cohort of 19,963 first time ninth graders in Chicago public schools looked to identify which middle grade metrics predicted ninth-grade failure. Utilizing student transcripts and test scores, the study found that information about students only marginally improved the prediction of later outcomes after considering just two or three key middle school indicators (Allensworth, Gwynne, Moore, & de la Torre, 2014). A series of analyses examined combinations of potential indicators for each high school outcome. With each outcome, the indicator with the strongest bivariate relationship was used first, and then additional predictors were added one at a time to determine whether each added new information to improve the prediction. Regression models (or logistic regression models for dichotomous outcomes) were used, comparing the R-squared statistic, percent correct prediction, sensitivity, and specificity derived from each model. There was a focus on model statistics, rather than coefficients associated with individual variables, to decide whether including each additional potential indicator in the model improved the prediction of the high school outcome. Grades and attendance were the factors in middle grades most strongly connected with high school success. For later academic success, grades and attendance mattered more than test scores, race, poverty, or other background characteristics (Allensworth, Gwynne, Moore, & de la Torre, 2014).

The R-square statistic is a measure of how well a predictor, or a set of predictors, explains variation in an outcome. An R-squared statistic can range from 0 to 1, with higher values signifying a better prediction. This Chicago study considered R-squares under 0.10 a poor prediction of the outcome, between 0.11 and 0.30 a moderately good prediction for the outcome, and R-squares above 0.30 a very good prediction for the outcome. This study provided the R-squared value for eighth-grade core GPA and attendance at 0.21 (moderate) to predict a student being on track academically in the ninth grade (Allensworth et al., 2014). A very low GPA prior to eighth grade was the strongest signal that a student was at high risk of being off track in high school, and attendance in grades five through seven was the next best indicator. Students who have an attendance rate of less than 85% in any of the middle grades were very likely to fail classes when they arrive in high school and are likely to fall off track for graduation. Many of the students who were at high risk of ninth-grade failure could be identified as early as sixth grade, and without a dramatic change in their educational experience, those students had very little chance of graduating from high school (Allensworth et al., 2014).

In the Allensworth et al. (2014) study, students with 80% or lower attendance in middle school had a risk greater than 75% to drop out of school. Students who attended less than 90% of the time in middle school had less than a 50% chance of staying on track to graduate once they entered high school. In addition, in the Neild et al. (2007) September 1996 study of an entire cohort of approximately 14,000 sixth-grade students in Philadelphia previously described, a sixth grader with attendance less than 80% for the year had a three in four chances of dropping out of high school.

The purpose of a study of student data for ninth-grade students in the 2010 and 2011 graduating cohorts from three Ohio school districts that varied in size, urbanicity, and characteristics of their student populations was to identify locally tailored indicators that predict students who fail to graduate on time. One of the districts had more than 40,000 students, while the other two districts each had 5,000 to 10,000 students. In all three districts, the end-of-year attendance rate was the only consistent predictor of failure to graduate on time (Stuit et al., 2016).

The study team used stepwise logistic regression to identify the subset of indicators for each district that were statistically significant predictors of student failure to graduate on time. A validation test was used to verify that the identified indicators were consistent predictors of failure to graduate on time when applied to 100 subsamples of students included in the logistic regression analysis (Stuit et al., 2016). Additionally, a receiver operating characteristic (ROC) curve analysis was used to determine the cut point that maximized the correct off-track prediction rate and to minimize the false alarm rate. The ROC curve is a plot of true-positive rate and true-negative rate, and the closest point to the left-upper corner of the plot is considered the most accurate. The larger the area under the curve statistic, which ranges from 0.50 to 1.00, the better the indicator classified nongraduates as nongraduates and did not classify graduates as nongraduates. The overall accuracy of the area-under-the-curve analysis for attendance below 90% was 0.69. The study found that an attendance rate below 90% in one district and below 95% in the other two districts was a statistically significant predictor of failure to graduate on time in at least 50 of 100 randomly simulated cohorts (Stuit et al., 2016).

An analysis of 30,000 students' longitudinal data in a large, urban school district in a southeastern school district by Schoeneberger (2012) looked at attendance patterns and their respective dropout risks. This data analysis spanned grades one through 12 and pulled data between the school years of 1998 and 2009. The study used group-based trajectory modeling to categorize students into longitudinal groups based on students' attendance patterns (Schoeneberger, 2012). Dropout risk factors were found to be associated with certain patterns of attendance. As students progress through schooling, they could have either a positive outcome or become disengaged. Becoming disengaged from school often leads to low attendance patterns and a higher risk of dropping out of school.

In the Schoeneberger (2012) study, students were divided into four categories of attendance: (a) constant attendees, (b) developing truants, (c) early truants, and (d) chronic truants. Constant attendees rarely missed more than 10% of their school days. Developing truants displayed patterns of increased prevalence of absenteeism from late elementary throughout middle school. Early truants had a greater prevalence of missing more than 10% of days in early elementary school. Chronic truants had the highest prevalence of missing school across all grade levels. The developing truants had an increased likelihood of missing more than 10% of their school days in middle school. This absenteeism pattern aligns with literature where disinterest in school occurs over time and manifests itself as poor attendance and eventually dropping out of school. In this study, the developing truants had the highest dropout rate at close to 25%. The chronic truants had the next highest dropout rate at nearly 21%. Early truants and constant attendees had dropout rates of 11% and 4%, respectively. Tracking student-level attendance patterns longitudinally could help identify students on the path to disengagement early in their academic career and provide assistance to correct the problem (Schoeneberger, 2012).

While many of the factors leading to student disengagement are not school-related, the behavioral indicators of student disengagement leading to a dropout outcome, such as attendance, manifest themselves directly at school (Mac Iver & Mac Iver, 2010).

As seen in a variety of studies at both the sixth-grade and ninth-grade levels, getting kids to come to school every day is vital because they cannot learn if they are not at school. Students with very strong attendance in middle school got much better grades in high school than those with moderate attendance (Schoeneberger, 2012). A plus for educators is that attendance is a variable that is monitorable from the first day of school, and there are ways to influence it, which could change the potential outcome for students.

School Failure

The likelihood of a student dropping out was considerably higher for students who have been retained. A comprehensive literature review of dropout research that examined grade retention as a predictor variable by Jimerson, Anderson, and Whipple (2002) reported that students who are retained at least once were 40% to 50% more likely to drop out when compared to students who were never retained. Additionally, students retained more than once in their academic careers were 90% more likely to drop out (Jimerson, Anderson, & Whipple, 2002).

Using an event history analysis and controlling for differences in grades, background, and attendance, Fall River, Massachusetts' public school's 1980-1981 seventh-grade cohort of middle school students (n = 1,052) showed that students retained in grades kindergarten through eighth grade are more than three times as likely to drop out when compared to their non-retained peers (Roderick, 1993). By evaluating student transcripts from elementary school until high school graduation, dropping out, or transferring, this study looked at how the experience of repeating grades influenced students' chances of dropping out. Students in the Fall River cohort who repeated at least one grade had a dropout rate of 77% compared to a 25% dropout rate among students who had never been retained. Even after controlling for grades and attendance through the ninth grade, students who repeated grades were significantly more likely to drop out than those who had never repeated a grade. This study also found that students who experience grade retention were more likely to drop out regardless of whether they were retained early or late in elementary and middle school. However, taking the analysis further, Roderick (1993) looked at how the association between grade retention and school dropout could be explained by grades and attendance both before and after the student was retained. It was discovered that approximately one-third of the association between grade retention and school dropout behavior could be explained by academic performance and attendance. Yet, even after controlling for a student's background and grades, students who repeated grades were substantially more likely to drop out than students who never experienced retention (Roderick, 1993).

In the Fall River study, hard copies of transcripts were used to collect student data and included such data as attendance, courses, course grades, and grade retention (Roderick, 1993). A logit analysis used the calculation of the derivative of the logistic probability function when all variables included in the equation were set at their mean. When all values are set at their mean, the values reported for each equation can be interpreted as the change in the probability of dropping out for the change in the independent variable. One incidence of grade retention was associated with a 0.20

increase in the probability of a student dropping out, and repeating more than one grade was associated with a 0.35 increase in the likelihood of dropping out (Roderick, 1993).

In another study, a sampling of 1,202 senior transcripts at a southeast Texas high school was analyzed to measure ninth-grade retention and dropout rates among students (Bornsheuer, Polonyi, Andrews, Fore, & Onwuegbuzie, 2010). The purpose of this study was to examine the relationship between ninth-grade retention and on-time graduation. On-time graduation was defined as completion of high school within four years after entering as a freshman. Archived student transcript and graduation data from the 2006-2007, 2007-2008, and 2008-2009 academic years focused on the population of students who did not obtain a minimum of 5.5 credits to move forward with their class from ninth grade to tenth grade (Bornsheuer et al., 2010).

In Bornsheuer et al.'s (2010) study, a quantitative analysis was used to determine if there was a relationship between ninth-grade retention and on-time graduation. The independent variable was ninth-grade retention status, and the dependent variable was on-time graduation. The relationship was assessed using a chi-square analysis. Statistical significance was obtained at the $\alpha = 0.05$ level. Effect sizes also were obtained via a chi-square analysis. The 2 x 2 chi-square analysis revealed a statistically significant relationship between ninth-grade retention and on-time graduation. The analysis found that students retained in the ninth grade were likely to not graduate on time at 85.8%, and students who were not retained in the ninth grade were more likely to graduate on time at 85.5%. The effect size, as measured by Cramer's V, was 0.61. Using Cramer's V criteria, the magnitude of the effect size was very large based on the scale in which 0.1 is small, 0.3 is medium, and 0.5 is large (Cohen, 1988). The odds ratio revealed that students who are retained in the ninth grade were 6.01 times less likely to graduate on time (Bornsheuer et al., 2010). An analysis of the nation's recent high school dropouts suggests that approximately 30% of dropouts are never promoted past the ninth grade (Neild, 2009).

In the previously mentioned 1996-1997 Philadelphia longitudinal study that used a sampling of 1,457 first-time freshmen, the majority of the Philadelphia Education Longitudinal Study (PELS) dropouts were listed as ninth or tenth graders despite having been enrolled in high school for several years (Neild et al., 2008). Almost half of the dropouts were listed as ninth graders when they dropped out, and another 31% were never promoted beyond the tenth grade. A majority of the PELS dropouts were seriously behind in their course credits at the time of dropping out. Despite being in high school for a number of years, 85% of the dropouts had earned no more than three credits, and 75% had earned no more than two credits (Neild et al., 2008).

Behavior Problems

Being suspended from school can be a trigger that puts a student on the path to ultimately dropping out. A suspension can lead to more suspensions, higher absenteeism, and course failure in the future. This progression can eventually lead to the path of dropping out (Balfanz, Byrnes, & Fox, 2014). Balfanz et al. (2014) analyzed suspensions, absenteeism, and course failures of 181,897 Florida ninth graders from the 2000-2001 school year. The purpose of their study was to learn to what extent suspensions were connected to lower academic outcomes such as course failure and later high school dropout. A longitudinal cohort study followed students from ninth grade until two years past their expected graduation date. Logistic regression modeling was used to examine the impact of suspensions in conjunction with demographic and other off-track indicators. The analysis resulted in a strong correlation between student suspension, chronic absenteeism (defined as attendance under 90%), and course failure. Almost half of the students who were suspended in the ninth grade were chronically absent, and nearly three-quarters failed a course (Balfanz et al., 2014). The rates for chronic absenteeism and course failures were much higher for students suspended in the ninth grade compared to freshmen who were not suspended. Black students had a significantly higher suspension rate than White students. For students whose only off-track indicator in ninth grade was being suspended, they exhibited other academic or behavioral issues later in high school. Throughout the tenth through twelfth grade, 42% of those students became chronically absent, and 59% failed courses. It is possible that for about 20% of the students suspended in ninth grade, alternatives to suspensions could reduce dropout rates (Balfanz et al., 2014).

In 2013, the Los Angeles Unified School District (LAUSD) banned the use of suspensions for "willful defiance," which can include anything from dress code violations, cell phone violations, or class disruptions. Instead, LAUSD began implementing alternative discipline options to still hold the student accountable for their actions but keeping them in school. In the two years after the change, graduation rates rose by 12%, and suspensions dropped by 53% (DePaoli et al., 2015).

A study of dropout rates in Kentucky high schools (n = 196) found a positive relationship between suspension rate and dropout (Christle et al., 2007). Annual reports submitted by the Kentucky Department of Education and the Kentucky Center for School Safety provided quantitative data for the two consecutive academic years 2000-2001 and

2001-2002. A sample of 40 schools that had both the highest and lowest dropout rates in the study was compared using a multivariate analysis of variance. The Pearson product-moment correlation coefficients were computed for each pair of variables to identify school-level variables that had a strong relationship to dropouts. The Bonferroni approach was used to control for Type I error, and a *p*-value of less than 0.004 was required for significance. A significant positive correlation was found between the dropout rate and suspension with a value of 0.362 for the Pearson product-moment correlation. The coefficient of determination between the dropout rate and suspension was 13.1%. Schools that relied on discipline that was exclusionary and absolute, such as suspension, may be perpetuating the failure cycle by impeding the educational progress of students (Christle et al., 2007). Balfanz et al. (2014) discovered in the Florida study that students who are expelled have fewer opportunities to gain academic skills, and even one suspension in the ninth grade was associated with a double increase in the likelihood of dropping out. The probability of dropping out rose from 16% for those not suspended to 32% for those who were (Balfanz et al., 2014).

A 2004 study by Suh and Suh used data from the National Longitudinal Survey of Youth (NLSY97) database from the U.S. Department of Labor. Participants were a nationally representative sample of approximately 6,200 students who either completed high school or dropped out without receiving a diploma or a GED by December 31, 2000. The purpose of the study was to identify the most significant factors contributing to high school dropout and the extent of their impact on the likelihood of dropping out of school (Suh & Suh, 2004). Multiple logistic regression using the forward selection procedure was used to methodically screen all variables and arrive at a simple model with great explanatory power. The screening process yielded 16 statistically significant predictors of high school dropout, including being suspended from school at least once. For the suspension predictor, the statistical significance value of $p < .001$ points to a strong association between being suspended and the likelihood of dropping out. The probability values represented the expected change in the probability of dropping out of school for every one standard deviation increase in the predictor variable. The change in the probability is found by subtracting 1 from Exp(B) (1.775), where a positive value represented an increase in the likelihood of dropping out, and a negative value indicated a decrease. Thus, the behavioral risk of being suspended increases the likelihood of dropping out by 77.5% (Suh & Suh, 2004). Therefore, the risk factor of being suspended has a major impact on students' decision to drop out of school.

Lee et al. (2011) examined the association between school suspension rates and dropout rates using a statewide sample of 289 Virginia public high schools with almost 350,000 enrolled students. Multiple logistic regression analysis was used to identify statistically significant predictors of dropout. The study found that if a student had a history of suspension, it increased the likelihood of dropping out by 78%. Additionally, after controlling for school demographics, the contribution of school suspension rates on dropout rates was examined. An independent-samples t-test was used to compare the mean dropout rates between schools that have low and high suspension rates. Schools with low suspension rates ($M = 2.26$) had significantly lower whole-school dropout rates than those schools with high rates ($M = 3.52$) $t(122) = -2.79$, $p < .01$, and $d = 0.40$, indicating a moderate effect size (Cohen, 1988; Lee et al., 2011). These findings contribute to the growing evidence that school suspension may have an adverse effect on students completing high school.

Gender

Robst and de Vries (2010) used the National Longitudinal Survey of Youth (NLSY) data, which contained information about children born in 1979 to women in the 1979 NLSY cohort, including 281 boys and 289 girls. Information included childhood behavioral problems, academic achievement in adolescence, and demographic information. The purpose of the study by Robst and de Vries (2010) was to examine the relationship between childhood emotional and behavioral problems and the probability of graduating from high school while controlling for various sociodemographic characteristics. In 2004, graduation data was examined from a few years when the participants were between the ages of 24 and 27 years old. Logistic regressions were estimated to determine whether childhood behavior problems were associated with academic achievement when controlling for various sociodemographic and maternal characteristics. There were a few gender differences in educational outcomes between boys and girls. Seventy-two percent of the boys received their high school diploma compared to 77% of girls. In the study, boys were more likely than girls to drop out of school with boys having a 14.6% dropout rate compared to an 8.6% dropout rate for girls ($X^2 = 4.7$, $p = 0.029$) (Robst & de Vries, 2010).

In a study of Baltimore City Public Schools in both the 2004-2005 (n = 6,812) and 2005-2006 (n = 7,729) school years, McIver and Messel (2013) set out to discover the extent that ninth-grade demographic characteristics are associated with non-graduation outcomes, including gender, race, socioeconomic status, English language learner status, and special education status. The researchers followed the two ninth-grade cohorts until their graduation year and one year past. Although the strongest predictors of graduation were ninth-grade attendance and course failure, gender was still significant. In Baltimore's multivariate analysis of longitudinal student cohort data for first-time freshmen in the 2004-2005 and 2005-2006 school years, male gender was found to be a significant predictor of non-graduation when controlling for ninth-grade behavioral variables. In both cohorts, the likelihood of males not graduating was much higher than females by a factor of roughly two (the odds ratio varied between 1.59 and 1.82 for the various models) (Mac Iver & Messel, 2013).

For the Los Angeles Unified School District (LAUSD), the educational progress of all first-time ninth graders in the 2001-2002 school year was analyzed through their expected graduation year of 2005 (Silver et al., 2008). One purpose of the LAUSD's seven-year longitudinal data set for the class of 2005 by Silver et al. (2008) was to find the student and school characteristics at both the middle and high school levels associated with dropping out. This cohort consisted of 48,561 students who attended 163 LAUSD middle and high schools in the second-largest school district in the country. A series of multilevel logistic regression models were used to determine the independent contribution of several student-level and school-level factors. For the students attending the LAUSD schools during the study, a higher percentage of girls than boys in the cohort graduated high school on time. Fifty-four percent of the girls graduated within four years compared to 42% of the boys (female odds of on-time completion ratio varied between 1.18 and 1.52 for the various models and $p < .001$) (Silver et al., 2008).

The purpose of the Neild and Balfanz (2006) study of 130,000 students enrolled in Philadelphia schools in grades six through twelve during the 2003-2004 school year was to find the key characteristics of Philadelphia's 13,000 dropouts and those considered "near-dropouts." Near-dropouts are students who were absent from school for more than half the school year. This study pulled data from the Kids Integrated Data System (KIDS) housed at the University of Pennsylvania. This system merged student data from the School District of Philadelphia with the city's social services agencies. A hazard analysis was performed to determine what is the probability or "hazard" of not graduating (the number of dropouts divided by the total number of students). During the 2003-2004 school year, males were considerably more likely than females to drop out but only somewhat more likely to be near-dropouts. Additional analysis in Philadelphia found that first-time freshman cohorts with expected graduation dates of 2000, 2001, 2002, 2003, 2004, and 2005 consistently graduated females at a rate of at least 10% higher than males with an almost 15% advantage from 2000 through 2003. Key characteristics of dropouts in Philadelphia were being a Latino student, being a male student, a student aged 15 or older at the beginning of ninth grade, having high absenteeism, and earning few credits (Neild & Balfanz, 2006).

The United States Census Bureau administers an annual October Current Population Survey (CPS) nationally that involves interviewing approximately 60,000 households (Bartishevich et al., 2003). Some of the questions asked of each household include the school enrollment status of all household members, the high school graduation status of household members aged 16 and older, the year in which they graduated, and if a dropout, the last date of their attendance in a regular school. This survey is one of the data sources that the U. S. Department of Education uses to estimate the number of new 16 to 24-year-old high school dropouts and graduates. For the years 1988-1989 through 2000-2001, the estimated dropout rate for men was higher than for women, with the gap widening by the end of the decade. An average of 510,000 students in grades nine through 12 dropped out with a cumulative dropout rate of 13.3% during that time period. For the years 1997 through 2001, there was an average of 120 male dropouts for every 100 female dropouts. Dropout rates for men were 1.4 to 3.0 percentage points higher than for women from 1997 to 1999. Dropout rates were available from 35 states during the 1997-1998 school year, and males had a higher dropout rate in all 35 states (Bartishevich et al., 2003).

Low Socioeconomic Status

In 2010, America's Promise Alliance launched the GradNation campaign with a goal of graduating 90% of high school students in the United States by the year 2020 (Alliance for Excellent Education, Everyone Graduates Center, Civic Enterprises, 2015). The purpose of this effort was to put more young people on the path to success in school, work, and life by raising high school graduation rates. An annual update of the GradNation campaign goal came in the form of the state-level adjusted cohort graduation rate (ACGR), which indicates the proportion of public high school freshmen who graduate with a regular diploma four years after they start ninth grade (Stark, Noel, & McFarland, 2015). A 2016 update of the GradNation campaign is found in the *Building a Grad Nation* report by DePaoli, Balfanz, and Bridgeland (2016). This report provides an analysis of each state's progress towards meeting the 90% on-time graduation rate by 2020, along with subgroup data. The 2016 update noted that even with increasing graduation rates, persistent graduation rate gaps hold back large numbers of minority, low-socioeconomic, homeless, limited English proficient (LEP), and students with disabilities across the United States (DePaoli et al., 2016).

Per the DePaoli et al. (2016) report, for the national and state-level 2013-2014 adjusted cohort graduation rate (ACGR) data, nearly half of all 2014 graduates came from low-income families. Only 74.6% of all low-income students graduated compared to 89% of non-low-income students. Since 2011, graduation rates for low-income students have increased, but they are still significantly behind their peers who are not low income (DePaoli et al., 2016).

In a Baltimore study of 790 first graders in the fall of 1982 until 1996, after their expected graduation date, Alexander, Entwisle, and Horsey (1997) looked at early-schooling predictors of eventual dropout. The four kinds of explanatory variables evaluated were: (a) sociodemographic characteristics, (b) family-context factors, (c) children's personal resources, and (d) children's school experiences. The Beginning School Study (BSS) within Baltimore City Public Schools involved the academic progress and personal development of a random sample of students enrolled in the schools since the first grade in 1982. The sample of students reflected the makeup of Baltimore's public schools in 1982, including racial composition and socioeconomic status. Logistic regression analyses were used to identify predictors of dropout, and the analyses evaluated the predictors one at a time (zero-order associations) by subclusters of predictors for each of the four kinds of explanatory variables, and the four full groups separately. The study found that being from a lower socioeconomic family increased the risk of dropping out. The socioeconomic index follows a normal distribution closely.

Therefore, its odds-ratio metric of 0.19 indicates a fivefold difference in the likelihood of dropping out for students who were from families that were one standard deviation below the mean as compared to students whose families had average socioeconomic scores (Alexander et al., 1997).

A thorough review of 25 years of ERIC literature from 1980 until 2005 focused on risk factors that increase the likelihood of students dropping out of school. This review resulted in an analysis of 21 longitudinal studies in a technical report by Hammond, Linton, Smink, and Drew (2007). In the studies, a variety of types of predictors were examined, and multivariate statistical techniques or models were used to control independent relationships between student demographics, individual factors, and the dependent variable of dropout and/or high school graduation. Risk factors for dropping out of school stem from a variety of factors in four areas: (a) individual, (b) family, (c) school, and (d) community. In 10 of the 12 studies, a family's socioeconomic status was a major risk factor for dropping out. A family's socioeconomic status was a more powerful influence than other factors that may well prevent dropping out, such as good school performance. A student's family socioeconomic status is one of the factors that most consistently impacts a variety of student educational outcomes, including being a major risk factor for dropping out of school when it is low (Hammond et al., 2007).

While high school dropout rates have steadily declined nationally, dropout rates for students from families with a low socioeconomic status have steadily increased. Low-income students drop out at a rate four and a half times higher than the rate of higher-income youth (Stark et al., 2015).

English Language Learner Status

A 2015 update of the GradNation campaign is found in the *Building a Grad Nation* report by DePaoli et al. (2015). This report provides an analysis of each state's progress towards meeting the 90% on-time graduation rate by 2020, along with subgroup data such as English language learners (ELL). English language learners are students whose native languages are languages other than English or whose difficulties in speaking, reading, writing, or understanding the English language will impede their ability to meet the state's proficiency level of achievement on assessments and their ability to perform successfully in classrooms where the language of instruction is English (DePaoli et al., 2015).

In a National Education Association (NEA) policy brief, ELL students' academic performance was found to be below that of their non-ELL peers, and ELLs have very high dropout rates (National Education Association, 2008). A 2005 nationally representative sample of 162,000 eighth grade students took both the reading and math assessments given by the National Assessment of Educational Progress (NAEP). For over three decades, the NAEP has assessed what students know and can do in various subjects to help evaluate the condition and progress of U.S. education. For the reading assessment, only 29% of eighth grade ELL students performed at or above the basic level compared to 75% of non-ELL students (Ford et al., 2005b). For the math assessment, only 29% of eighth grade ELL students performed at or above the basic level compared to 71% of non-ELL students (Ford et al., 2005a). When a student scores at the basic achievement level, it means that the student has shown only partial mastery of the knowledge and skills, which are fundamental for proficient work in the eighth grade.

An analysis of 48,561 first-time ninth graders in the Los Angeles Unified School District (LAUSD) from 2001-2002 until their expected graduation date in 2005 showed a much lower graduation rate for students designated as limited English proficient (LEP). Only 33% of students designated at LEP graduated within four years compared to 54% of students not designated as LEP (Silver et al., 2008). Of all student subgroups measured by the National Center for Education Statistics, students with limited English proficiency graduate at the lowest rate (DePaoli et al., 2015). Working to close the identified math and reading achievement gaps should be a big step in helping to increase the graduation rate of ELLs.

Special Needs Status

In the United States, students identified as having a disability receive individualized services based on their needs and educational goals. Even with individualized support, many students with disabilities perform below their non-disabled peers (Gwynne et al., 2009). Closing the large graduation gap for the special education population is a necessary task. The state-level 2013 adjusted cohort national graduation rate data show that the graduation rate for students with disabilities was 61.9%, nearly 20 points behind the national average (DePaoli et al., 2015). States that consistently graduate 85% or more of their general population students struggle to graduate even 70% of their students with disabilities. Of all student subgroups measured by the National Center for Education Statistics, students with disabilities graduate second lowest, only behind English language learners. By rising only 2.9 percentage points since 2011, the adjusted cohort graduation rate for students with disabilities has been one of the student subgroups with the least amount of growth (DePaoli et al., 2015).

In a report by Stark et al. (2015) based on the 2012 national Current Population Survey (CPS), students with disabilities had a dropout rate three times the dropout rate of students without disabilities (10.0% vs. 3.2%). Reschly and Christenson (2006) analyzed data from the National Educational Longitudinal Study (NELS), which collected data from almost 25,000 eighth graders beginning in 1988 until the study was completed in 2000. This longitudinal study was the third conducted by the National Center for Education Statistics (NCES) in the U.S. Department of Education and used a clustered, stratified national probability sample. Approximately 26 students were randomly selected from each of the 1,052 public and private schools in the study. The information came from students, teachers, parents, and school administrators. A series of stepwise logistic regressions were used to determine how well certain variables measured in the eighth grade predicted dropout among students with learning disabilities (LD) or emotional or behavioral disorders (EBD). For students with EBD who had not been retained, they had a 73% higher chance of graduating than those who were retained. Students with LD who had not been retained had a 33% decrease in the odds of dropping out of school than retained students. The researchers found that students with disabilities have much higher dropout rates than students without disabilities. With dropout rates for students with learning disabilities and emotional or behavioral disorders at 26% and 50%, respectively, the dropout rate for students without disabilities was much lower at 15% (Reschly & Christenson, 2006). These eye-opening statistics reveal the importance of educating the whole child and providing multiple tiers of support.

Using the Allensworth and Easton's (2005) Chicago study of data from both the 2000-2001 and 2003-2004 freshman cohorts, Gwynne et al. (2009) focused on the students in the study who received special education services. This additional Chicago study focused on describing the early warning indicators, including absences, course failures, GPA, and on-track status, of the students who received special education services during their freshman year of high school. Additionally, the relationship between those early-warning indicators and graduation rates was explored to determine if the early-warning indicators were useful in identifying students at risk of dropping out. Non-nested and nested models were used to estimate the gaps in absences, study habits, course failures, and GPA between students with disabilities and students without disabilities. Two-level hierarchical linear models were used to take into account school effects and prior achievement. There was a strong relationship between freshman year indicators and graduation rates for all students, but students who received special education services were much less likely to graduate high school than non-disabled students (Gwynne et al., 2009).

Graduation rates vary for students based on their identified disabilities. Students with physical/sensory and speech/language disabilities graduated at higher rates than students with learning disabilities and mild cognitive disabilities. Students with emotional disabilities graduated at the lowest rates for students with disabilities. In the Chicago study, 45% of students with disabilities graduated within four years, and 50% graduated within five years as compared to the four-year and five-year graduation rates of non-disabled students at 67% and 70%, respectively. The lower four-year and five-year graduation rates resulted from students with disabilities having lower grade point averages, more course failures, and higher absenteeism (Gwynne et al., 2009).

Standardized Test Scores

Many studies have found that standardized tests do not predict dropout status after controlling for other variables (Allensworth & Easton, 2005; 2007; Baltimore Education Research Consortium, 2011; Burke, 2015). A 2007-2008 study of 6,118 freshmen at four Oregon school districts looked at the students' graduation outcomes. The researchers analyzed high school graduation outcomes associated with gender, race or ethnicity, special education status, English learner student status, attendance, and achievement on standardized tests in grades eight and nine, and behavior in grade nine after controlling for other variables and for characteristics of the schools that students attended. The relationship of the predictor variables with students' high school outcomes was analyzed using logistic regression while controlling for other variables and the schools that students attended. The variables of passing the eighth-grade reading and math standardized tests had high p values of 0.142 and 0.663, respectively. Thus, the variables were not statistically significant when testing the association between students' high school graduation outcomes and their eighth-grade standardized reading and math scores. However, valuable indicators of graduation outcomes were attendance rates below 80% in grades eight and nine and a GPA below 2.0 in grades eight and nine (Burke, 2015).

In the previously mentioned Chicago study, ninth-grade course achievement explained greater variance in students' predicted graduation status than did their eighth-grade standardized test scores (Allensworth & Easton, 2007). The 2007 Chicago study built upon the 2005 Chicago study, which identified the freshman on-track indicator of being on track at the end of the freshman year as a predictor of graduating four years later. The 2007 study looked to identify equally predictive freshman-year indicators that are available during the school year and not at the end. Logistic regression models utilized in the 2007 Chicago study used a variety of indicators, including students' failures, absences, and overall grades, to identify what matters for a successful freshman year.

For first-time ninth graders in the 2004-2005 school year (n = 24,894), on-track status, GPA, the number of semester course failures, and semester absences all correctly identified graduates and nongraduates over 77% of the time. In addition, course attendance was eight times more predictive of course failure in the freshman year than eighth-grade test scores, with 63% compared to 8% (Allensworth & Easton, 2007). The relative inadequacy of standardized test scores to predict graduation implies that the transition to high school places other demands on students besides acquiring academic skills (Allensworth & Easton, 2005).

Additionally, the Baltimore Education Research Consortium study (2011) of the 2000-2001 cohort of 7,887 sixth-grade students from the Baltimore City Schools sought to determine if there were indicators that predicted eventual dropouts with a reasonable level of certainty. Utilizing logistic regression analysis, this study revealed that sixth-grade math and reading standardized test scores did not meet the criteria as highly predictive and high-yield indicators of dropping out in the future. For indicators to be considered highly predictive and high-yield, the following principles had to be met: (a) at least 70% of students with the indicator did not graduate; and (b) more than 20% of students who eventually dropped out displayed the indicator. In this way, only indicators that were highly predictive and practically meaningful were identified. Even scoring the very lowest scale scores or national percentile ranks for math and reading did not predict future non-graduation (Baltimore Education Research Consortium, 2011).

School-level Variables

Schools themselves can play a part in the difficulty that some students experience in school. Once students reach high school, they often rush from one 50-minute class to another without establishing relationships with teachers. Students can feel anonymous and alienated (Zvoch, 2006). There is no one person to keep a watch on how individual students are doing in their classes, and teachers often do not have the expertise or time to work with students who enter high school with weak academic skills. In addition, ninth graders who experience scheduling issues and unorganized classrooms at the beginning of the school year have lower grade point averages (Neild, 2009).

School characteristics such as the organization, school climate, and overall socioeconomic status of the student body can impact student dropout (Neild, 2009). For example, a substantial minority of ninth-grade students who get off track test at or above grade level, have no prior course failures, and have good attendance. The cause could be the negative impact of high school's organization and climate. This argument is supported by analyses that show how ninth-grade failure rates vary considerably within a single school district, even after adjusting for characteristics of the schools' student populations (Sebring et al., 1996).

A study of a large southwestern school district's 2001-2002 ninth-grade cohort (n = 6,330) used multilevel logistic regression models to assess relations between school characteristics and dropping out of high school. This study found that the percentage of low socioeconomic students was a significant predictor of the school dropout outcomes.

As school poverty increased, the relation between individual economic disadvantage and dropping out increased. The data showed that the odds that an economically disadvantaged student would drop out of an affluent school were no greater than their economically advantaged peers. However, as school poverty increased, economically disadvantaged students were much more at risk of dropping out compared to their economically advantaged peers (odds ratio = 1.90) (Zvoch, 2006).

Major Studies

The identification of students at risk of dropping out of high school is based on solid research. Long-term studies in large urban school districts, including Chicago, Texas, Baltimore, and Philadelphia, provide information about powerful indicators that can predict during the first year of high school and even during the sixth grade, whether students will drop out before graduation (Allensworth & Easton, 2005; 2007; Balfanz et al., 2007; Baltimore Education Research Consortium, 2011; Bornsheuer et al., 2010; Neild et al., 2007). Research by the Consortium on Chicago School Research (CCSR) demonstrates that the most powerful indicators in high school were related to attendance, course performance, and earned credits (U.S. Department of Education, 2016b). For middle school, the most powerful indicators are associated with poor attendance, poor behavior, and failing math or English (Balfanz et al., 2007).

Performance in the ninth grade can be an important indicator of whether a student was likely to graduate from high school (Allensworth & Easton, 2007). One district at the forefront of developing ninth-grade on-track indicators of dropout is Chicago's public schools. Their school system, along with external partners, developed early warning indicators to improve student achievement (Allensworth, 2013). A 2005 study of over

26,000 high school freshmen in 2003-2004 by the University of Chicago Consortium on Chicago School Research (CCSR) found a simple indicator to determine if a student is on track to graduate within four years. Whether students earned enough credits by the end ninth grade to be promoted to the tenth grade and not failing more than one semester of a core subject could predict who would graduate with 80% accuracy (Allensworth & Easton, 2005).

Allensworth and Easton (2005) looked at the graduation rates of students in Chicago who failed no more than one core course and accumulated at least five full course credits during their freshman year, enough to be promoted to the tenth grade. These students were considered on track and were found to be nearly four times more likely to graduate in four years than freshmen who failed two or more courses and accumulated fewer than five credits. These freshman year on-track indicators could correctly predict, with very high accuracy, whether a student would graduate. In the Allensworth and Easton (2005) Chicago study, students who transferred out of Chicago public schools during their freshman year were not included in any of the statistics. Those who transferred after their freshman year were included in the freshman year on-track rates, but not in the graduation rate statistics. The report only included analysis of first-time high school freshmen. Students who repeated ninth grade were not counted as first-year students in any of the calculations.

The on-track indicator was timely and was a valuable tool for parents, schools, and the school system (Allensworth, 2013). The Chicago school district realized that students on track with credits in the ninth grade predicts eventual graduation better than standardized test scores. Freshmen with high test scores and off track were half as likely to graduate as those with low test scores and on track at the end of ninth grade. Also, students in the second quartile of achievement, with below-average skills, and who ended their ninth-grade year on track had a 76% graduation rate. Thus, it is not academic skills that are the best instrument to predict future graduation; it is freshman year failure (Allensworth, 2013).

It is important for school and district personnel to watch the ninth-grade pass rate because it can predict eventual graduation much better than standardized test scores. Educators often think that academic skills are a good predictor of graduation and are frequently used to identify students for intervention programs. Academic skills, like standardized test scores, and even background characteristics, are not as strong of a predictor of eventual graduation as freshmen year failure (Allensworth, 2013). In the Allensworth and Easton (2005) Chicago study, 22% of the freshmen with the top incoming standardized test scores finished ninth grade off track, and only 37% of those off-track students graduated four years later.

In the Chicago study, using ninth-grade students' background data of eighth-grade reading and math test information, gender, race, socio-economic status, mobility, and age in the ninth grade only predicted 65% of the graduates, while the on-track indicator alone predicted 80% of graduates. When combining the background data with the on-track indicator into one model, the prediction went up only one point to 81% (Allensworth, 2013). Once the on-track indicator is known at the end of ninth grade, the background information was essentially inconsequential. Because the on-track indicator was a categorical variable, either on track or off track, it was easy to report data over time. But, it is not timely or detailed. School personnel must wait until the end of the ninth-grade year to determine who is off track and at risk for not graduating. It also does not provide detailed information to target specific interventions for students (Allensworth, 2013).

A similar 2007 Chicago study of almost 25,000 freshmen in 2004-2005 looked at why ninth-grade students in Chicago failed a class and went off track. It was discovered that students' behaviors, especially not attending class, helped explain student failure. This study discovered that together, a student's grade point average (GPA), failure rate, and absences in the ninth grade are just as predictive of their graduation four years later as the on-track indicator but is available earlier in the school year (Allensworth & Easton, 2007). Educators do not have to wait until the end of the freshman year to determine who is most at risk for not graduating. Attendance is known immediately, and the GPA and failures are known after the first quarter or semester. For consistency, absences were counted on a course-by-course basis and then combined to total the number of days absent. For example, if a student missed one out of seven courses in a day, it was counted as one-seventh of a day of absence for that student. Data on grades were also provided separately for each course taken by each student each semester (Allensworth & Easton, 2007).

In the 2007 Chicago study, of the freshmen who entered in the 2000-2001 school year, almost all students with at least an earned B GPA at the end of their freshman year graduated within four years (Allensworth, 2013). The advantage of the additional indicators was information was available much earlier in the school year, and specific information was available to target programs for students struggling to pass or come to school. With such strong data that showed the importance of the freshmen year of high school, serious consideration should be given to providing more money, personnel, and interventions during ninth grade to ensure their success and eventual graduation. The interventions should focus on redirecting the indicators that identify freshmen most at risk for not graduating within four years. For the 2004-2005 freshman class in Chicago, 41% of students were off track (Allensworth, 2013). To improve graduation rates, students' ninth-grade course performance must improve. By using the on-track indicator, schools could focus their resources on individual students who are identified as having the highest risk for failure, rather than simply focusing on their background characteristics (Allensworth & Easton, 2005). It seems more productive to think about individual students who are at high risk of failure, rather than assuming that certain types of students will fail in high school.

In another study, Bornsheuer et al. (2010) looked at archived graduation data from 2007, 2008, and 2009 at a southeast Texas high school to determine if there was a relationship between ninth-grade retention and on-time graduation. The focus was on the students who did not obtain a minimum of 5.5 out of 7.0 credits required to move from the ninth grade to the tenth grade. The researchers wanted to see if the students who did not obtain the necessary 5.5 credits to move to the tenth grade were able to overcome the setback, catch up on credits, and graduate with their classmates within four years of entering high school. This study revealed a statistically significant relationship between ninth-grade retention and on-time graduation. In this study, students retained in ninth grade were six times less likely to graduate on time. Students retained in the ninth grade tended not to graduate on time (85.8%), whereas students not retained in the ninth grade were more likely to graduate on time (85.5%) (Bornsheuer et al., 2010). Considering both the Chicago studies and the southeast Texas study had similar results, the freshman year of high school is shown to be one of the most critical years when determining if a student will graduate on time from high school.

A Baltimore study of 6,812 2004-2005 first-time ninth graders and another 7,729 2005-2006 first-time ninth graders sought to determine if factors identified in previous research as dropout early warning indicators were also significant predictors of non-graduation in Baltimore (Mac Iver & Messel, 2012; 2013). For both cohorts, the probability of on-time graduation decreased with each ninth-grade course failure and with increased absences. In this study, students missing course grade data were excluded from analyses. While 82% of first-time ninth graders with attendance of at least 95% graduated on time in Baltimore, only 26% of students who missed more than 20 days in ninth grade graduated on time. Also, while 86% of the 2004-2005 ninth graders who passed all their core courses graduated on time, just 30% of those in the same cohort who failed two or more ninth-grade core courses graduated on time (Mac Iver & Messel, 2012; 2013).

Baltimore's graduation outcomes were also linked to student behavior. Students suspended for at least three days in the ninth grade were much less likely to graduate on time than those without the indicator (28% graduation rate compared to 63% graduation rate) (Mac Iver & Messel, 2012).

Ninety-two percent of nongraduates in the 2004-2005 Baltimore cohort demonstrated an early warning signal in ninth grade. Surprisingly, almost half the graduates also displayed at least one early warning signal. Additional research discovered those students had fewer course failures and/or higher rates of attendance than did the dropouts (Mac Iver & Messel, 2012).

Just as ninth grade has been identified as a tough transition year for students, so has the sixth grade. The first year of middle school, much like the first year of high school, looks to be a make-or-break year (Balfanz, 2009). Like results in Chicago, Texas, and Baltimore, a study of Philadelphia students yielded graduation indicators, but for sixth-grade students. In a Philadelphia study of over 12,000 sixth graders, four sixth-grade indicators were identified as having both high predictive power for those flagged and high ability to predict the future nongraduates. Those indicators were: (a) attend school less than 80% of the time; (b) fail math in the sixth grade; (c) fail English in the sixth grade; and (d) receive an out-of-school suspension in the sixth grade. Together, those four sixth-grade warning flags identified 60% of the students who would not graduate on time and had a false-positive rate of less than 40% (Balfanz et al., 2007). It should be noted that districts can have different thresholds for signaling danger of dropping out (Balfanz, 2009). For example, in some districts, a 90% absenteeism rate may signal that a student is in danger of not graduating, but in another district, the absenteeism rate is at 80% to signal danger.

To identify the most accurate sixth-grade indicators, the Philadelphia study tested a variety of variables related to attendance, demographics, behavior, course credits, and test data to determine their predictive power and high yield. If 75% or more of the sixth graders flagged did not graduate from high school either on time or one year late, and if the flag had a high yield by identifying 10% or more of the district’s future nongraduates then the variable experienced additional testing (Balfanz et al., 2007). Those flags with both high predictive power and high yield underwent logistic regression techniques to establish that each variable had significant and independent predictive power, even after controlling for other variables.

Attending school 80% or less of the time identified 50% of the cohort's future nongraduates, and students who had this flag along with a course failure flag were especially unlikely to graduate. Course failure predicted nongraduates better than low standardized test scores. Sixth graders who failed either a math or English course rarely graduated. Fourteen percent of the sixth graders failed math, and only 19% of those students eventually graduated within one year of their on-time graduation date. Eleven percent of students in the sixth grade failed an English course, and only 18% graduated within one year of their on-time graduation date. Of the students who received even one out of school suspensions in the sixth grade, only 20% of those students graduated within one year of their on-time graduation date (Balfanz et al., 2007).

Balfanz et al. (2007) discovered with all else being equal, students with chronic absenteeism were 68% less likely than other students to graduate, those with poor behavior grades were 56% less likely to graduate than others, those who failed math were 54% less likely to graduate than others, and those who failed English were 42% less likely to graduate than others. Students who had one or more of the warning flags had only a 29% graduation rate. A significant finding of the Philadelphia study is that the appearances of academic and behavioral problems that many students display at the start of middle school do not self-correct (Balfanz et al., 2007).

In another study of sixth graders, the Baltimore Education Research Consortium (2011) studied the Baltimore City School's 2000-2001 cohort of sixth-grade students who were scheduled to graduate in 2007. The purpose of the study was to determine if there were indicators that predicted eventual dropouts with enough reasonable certainty to justify intervention efforts. This study used logistic regression to identify the following highly predictive and practically meaningful indicators: (a) absent 20 or more days; (b) failing English, or math, or both and/or failing course average across all four content courses; (c) being at least one-year overage; and (d) being suspended three or more days (Baltimore Education Research Consortium, 2011). The study found that the more indicators a student possessed, the less likely he or she was to graduate. Also, less useful predictors found were: (a) gender; (b) English for speakers of other languages (ESOL) services; and (c) special education services (Baltimore Education Research Consortium, 2011). While only 30% of nongraduates had no indicators in the sixth grade, most had developed them by ninth grade. Because a majority of nongraduates showed signs of disengagement in sixth grade (70.8%), this study emphasized the need to monitor early warning indicators in both sixth and ninth grades (Baltimore Education Research Consortium, 2011).

In addition to major studies, a unique warning system created for statewide use in grades six through nine is the Wisconsin Dropout Early Warning System (DEWS) (Knowles, 2014). DEWS is a predictive model of student dropout risk. Using only the student's prior year's data, Wisconsin's DEWS creates a risk score from 0-100 that signifies the probability of that student graduating within four years. The risk score represents how often similar students across the state graduated high school on schedule in the past (Wisconsin Department of Public Instruction, 2015). For example, a student with a score of 70 has similar reported data as students in previous cohorts who graduated on time 70% of the time. A higher score means a student has a better chance of graduating on time. Along with a score, students are given a risk level of "low," "moderate," or "high" for each of the four domains, including attendance, discipline, mobility, and annual state assessment (Knowles & White, 2013). DEWS used the Receiver Operation Characteristic (ROC) plot to show the tradeoff between true-positive classifications and false alarms or true-negative rates. The ROC plot showed a comparison of an accurate identification of dropouts and false identification of eventual graduates as dropouts (Knowles, 2014).

According to the DEWS Data Brief, Wisconsin's DEWS used the R statistical software to run a state-of-the-art machine-learning algorithm, instead of a checklist indicator approach, to test and combine up to 50 statistical models per grade for grades six through nine. This model predicted between 60% and 65% of future dropouts and late graduates and had a low false-positive rate (Wisconsin Department of Public Instruction, 2015). Wisconsin's DEWS did a good job of balancing the trade-off between correctly identifying likely dropouts and false-alarms by using the ROC plot. A separate model of fit was used for each grade level and used data indicators related to attendance, suspensions, school moves, state assessment performance, school characteristics, and student characteristics (Wisconsin Department of Public Instruction, 2015). Some of the categorical variables included race, gender, free/reduced lunch status, discipline, English language learner status, special education status, and school moves. Some of the continuous variables used with DEWS included individual and cohort standardized math and reading test scores and individual and cohort attendance rate. The goal of DEWS was to maximize the accuracy of predictions for individual students using data available on a statewide basis. DEWS will ultimately be used to predict the likelihood of current middle grade students finishing high school many years later, so it must be designed to have predictive power outside of the cohort of students used to create it (Knowles, 2014). The advantage of the DEWS model is that it can help identify students before they begin to show signs of low performance or disengagement in school.

Summary

States, districts, and schools are increasingly interested in using early warning systems to identify students most at risk for dropping out of high school. Utilizing academic and behavioral predictors at the right time in a student's school career will help school systems accurately identify at-risk students and provide the best interventions to put the students back on the path to graduation. Students often signal that they are on or off track on the path to graduation through their attendance, behavior, and course performance. An early warning system can be used at the school level as well as the district level to identify the students who are in need of guidance and interventions in order to graduate within four years of starting high school. The impact of putting students back on track to graduation could be life changing for the students and result in a better quality of life for both the students, their families, and society.

Chapter III

METHODOLOGY

This chapter describes the research design, participants, and instrumentation of this nonexperimental study using longitudinal data from two cohorts of a mid-sized Georgia school district. The study's purpose was to create a dropout early warning system to predict both middle school and high school students who are at risk for dropping out or not graduating on time. This chapter also describes the methods for collecting the data. This study utilized the statistical procedures of logistics regression, linear discriminate analysis, and quadratic discriminate analysis to identify the most accurate indicators at each grade level that result in high levels of true classification and low levels of false alarms.

The research questions framing this study are:

1. Does one or more of the ninth-grade variables consisting of attending at least 90% of the time, earning sufficient credits to move to the tenth grade, receiving out-of-school suspension, number of school moves, standardized reading and math scores, failing no more than one semester of a core content course, school minority percentages, school poverty percentages, ELL status, SWD status, free/reduced meal status, race, and gender accurately predict students who will not complete high school within four years?

2. Which statistical model is the most accurate at predicting future dropouts or late graduates utilizing ninth-grade variables with high levels of true classification and low levels of false identification?
3. Does one or more of the sixth-grade variables consisting of failing English, failing math, attending at least 80% of the time, receiving out-of-school suspension, number of school moves, standardized reading and math scores, school minority and poverty percentages, ELL status, SWD status, free/reduced meal status, race, and gender accurately predict students who will not complete high school within four years?
4. Which statistical model is the most accurate at predicting future dropouts or late graduates utilizing sixth-grade variables with high levels of true classification and low levels of false identification?

Research Design

This study employed a nonexperimental, ex post facto, multivariate correlational research design. Logistic regression analysis, linear discriminate analysis, and quadratic discriminate analysis were used to predict individual students' risk of not graduating on time. This design was necessary since the data have been collected by the school district, and the variables cannot be manipulated.

The independent variables for this study were chosen based on a thorough literature review of dropout indicators at both the ninth-grade and sixth-grade levels and thus, used as predictor variables in this study. Two different types of variables were analyzed during this study, those which schools have a direct relationship with and can impact and those outside the influence of any school. The ninth-grade dropout indicators that are both influenced and not influenced by schools are: attending school less than 90% of the time, earning sufficient credits to move to the tenth grade (5 Carnegie units), number of days suspended out of school, number of school moves, End of Course Test (EOCT) standardized reading and math scores (ranges from 200 to 600), failing no more than one semester of a core course, school minority percentages (from 1 to 100), school poverty percentages (from 1 to 100), ELL status (yes or no), SWD status (yes or no), free/reduced meal status (yes or no), race, and gender (Allensworth & Easton, 2005; DePaoli et al., 2015; Kemple et al., 2013; Lee et al., 2011; Mac Iver & Mac Iver, 2010; Mac Iver & Messel, 2012; 2013; Neild, 2009; Zvoch, 2006).

The sixth-grade dropout indicators in these two categories are: failing English with an average below a 70, failing math with an average below a 70, attending school less than 80% of the time, receiving out-of-school suspension (yes or no), number of school moves, Criterion-Reference Competency Test (CRCT) standardized reading and math scores (ranges from 650 to 900), school minority percentages (from 1 and 100), school poverty percentages (from 1 and 100), ELL status (yes or no), SWD status (yes or no), free/reduced meal status (yes or no), race, and gender (Balfanz, 2009; Balfanz et al., 2007; Jerald, 2006; Mac Iver, 2010; Rumberger, 2004; Silver et al., 2008).

The dependent variable was either graduating high school within four years or not. A relationship among the set of variables will forecast students at risk of not graduating within four years of entering high school or dropping out. If the model predicts accurately, interventions could be developed, used, and tested by the school system.

Participants

This study was conducted using student data from a mid-sized school district in Georgia. According to the Governor's Office of Student Achievement website (https://gosa.georgia.gov/report-card-dashboards-data), this school district supports approximately 26,500 students consisting of 73% Black, 18% White, 5% Hispanic, 2% Asian, and 2% multiracial. There are six middle schools and six high schools, with each of the schools having a free and reduced lunch percentage close to 99%. Of the student population, 2% are English language learners (ELL), and 10% are students with disabilities (SWD).

This study included all students who entered sixth grade in the 2010 and 2011 school years and ninth grade in the 2013 and 2014 school years in a mid-sized school district in Georgia. Their on-time graduation years were 2017 and 2018. The analysis was conducted for each cohort's sixth-grade and ninth-grade school years. The data set was split so that the 2017 cohort data were used for training, and the 2018 cohort data were used to evaluate the models. Each year the total number of students in each cohort was approximately 1,000 students.

When looking at the respective cohort beginning grades (sixth and ninth), student data was excluded if the student transferred into the district after the start of that grade level. Data was also removed for students who transferred out of the district and enrolled in another school system during or after the grade level being studied. The data warehousing system, Infinite Campus, tracked student data and determined when either of the two previously mentioned scenarios occurred. A student's death was also a reason to remove the student from the sample. Based on research, the following criteria were used to identify dropouts (Balfanz, Bridgeland, Bruce, & Fox, 2012):

- Students who did not officially withdraw but were removed from the school rosters for lack of attendance;
- Students who were removed from the school rosters but did not enroll in another school;
- Students who were incarcerated in a facility not associated with a public school system; and
- Students who were expelled from school and failed to return after the expulsion was completed.

Instrumentation

For both ninth graders and sixth graders in both cohorts of this study, data were related to the following predictor variables: (a) course failures, (b) credits earned, (c) out-of-school suspension, (d) number of school moves, (e) standardized reading and math scores, (f) school minority percentages and school poverty percentages, (g) attendance, (h) student ELL status, (i) SWD status, (j) free/reduced meal status, (k) race, and (l) gender.

The district uses a standard numeric grading scale for grades, and teachers record students' numerical grades in Infinite Campus. Infinite Campus is the district's online student information system created by the company Transforming K12 Education. Students' final numerical grades are posted to their transcripts in Infinite Campus at the end of both fall and spring semesters. Each class period's teacher records attendance in Infinite Campus daily. Student discipline referrals are documented in Infinite Campus after each occurrence by each school's respective assistant principals. Schools follow a progressive discipline plan to keep discipline consistent based on the discipline infraction. This discipline plan provides building-level administrators a guide to use when addressing discipline referrals. For this study, the district-level analyst created reports within Infinite Campus and exported data related to grades, attendance, and discipline into Excel. Additional reports from outside sources containing school minority data, school poverty data, and free/reduced meal data were created and exported to Excel.

School-level registrars enter data for student ELL status and SWD status. The coordinators of the ELL and SWD departments attend regular registrar meetings conducted by the Assessment and Accountability Department to provide any missing data to the respective registrars. Each school's registrar also enters race and gender information. To validate student data in Infinite Campus, registrars used the web-based application FTE Track to check for potential reporting errors quarterly. FTE Track was a browser-based, multi-user student information system (SIS) by USHA software. Building-level personnel could also use the FTE Track student information system to check data. Monthly training meetings were held by the Student Information Services Manager and Student Information Services Coordinators with the registrars pertaining to upcoming work and data (T. Jones, personal communication, June 20, 2018). Each school's registrar contacts parents, teachers, administrators, and district staff to gather the necessary information.

Additional tools available to help ensure continuity across the district are a common pacing guide in courses, frequent common assessments in courses, and a progressive discipline plan. These commonalities provide consistency among the schools in the district and are reflected in overall grades and similar punishments for similar discipline infractions. Also, school administrators from across the district attend yearly discipline training to be consistent with both the dispensing of punishment and entering discipline data in Infinite Campus.

Georgia's two state standardized tests given to students in the 2017 and 2018 graduation cohorts were Georgia's Criterion-Referenced Competency Test (CRCT) and the End-of-Course Test (EOCT) for sixth and ninth grades, respectively. For the CRCT, the cut scores are: (a) below 800, do not meet the requirements; (b) 800-849, meet requirements, and (c) 850 or above, exceed requirements. Performance levels were identified as Level 1, did not meet standards; Level 2, met standards; and Level 3, exceeded standards. Like the CRCT, the EOCT was a state-mandated, standardized test administered at the completion of certain core content courses. Students who scored below 400 on the EOCT were identified at Performance Level 1 and did not meet standards. Scores from 400 to 449 met standards and were identified at Performance Level 2. Finally, scores of 450 or higher identified a student at Performance Level 3, exceeded standards. For ease in computing students' grades, the scale EOCT scores were converted to a 100-point scale and entered as 20% of the final course grade. For this study, standardized reading and math scores were coded as scale scores. Both the CRCT and EOCT had validity and reliability data provided by the Testing and Assessment Division of the Georgia Department of Education (J.D. Rollins, personal communication, July 27, 2018).

Validity

The evidence for the validity of the CRCT and EOCT shows how well the assessments match the intended curriculum and how the score reports inform stakeholders about the students' performance. By paying careful attention to each phase of the test development process, the GaDOE ensures that both the CRCT and EOCT are valid instruments. Validity is associated with a varied process and a collection of evidence over time. Questions of validity cannot be summed up in a single statistic because the answer to a valid question for a test lies in the documentation of the test development process. The test development process includes: (a) clearly identifying the purpose of the test; (b) committees of educators reviewing the curriculum and deciding which concepts, knowledge, and skills were assessed and how they were assessed; (c) aligning the test with curriculum; (d) creating content domain and test item specifications; (e) using professional assessment specialist to write test items; (f) teams of Georgia educators review items; (g) field test items; (h) Georgia educators again examine test items along with data from field tests for final acceptance, revision, or rejection; (i) develop the actual test and equate it to ensure equal difficulty for multiple test forms and additional administrations; (j) determine the number of items correct to decide if a score meets or exceeds expectations; and (k) score tests and distribute results. In addition to the content validity, the GaDOE collects construct validity evidence for the EOCT. Construct validity and criterion-related validity address the psychological characteristic of interest and how accurately criterion performance can be predicted from test scores, respectively (J.D. Rollins, personal communication, July 27, 2018).

.

The GaDOE carefully assessing each phase of the test development process provides evidence that the CRCT and EOCT are valid instruments for the purpose for which they were intended—to measure student mastery of the state's curriculum. Documentation was produced at every phase of the test development, and the state collects evidence through separate independent alignment studies. Additionally, external validity was shown by comparing how the CRCT and EOCT measure compared to other well-established assessments including the ITBS (J.D. Rollins, personal communication, July 27, 2018).

Reliability

The reliability of the CRCT and EOCT means they produce stable scores if the same group of students took the same test repeatedly. For the CRCT and EOCT, Cronbach's alpha reliability coefficient, which expresses the consistency of test scores, and the standard error of measurement (SEM), which quantifies the precision of a test, are reported. Additionally, the Rasch-based conditional standard errors of measurement (CSEMs), which express the degree of measurement error in scale score units, was reported for the CRCT (J.D. Rollins, personal communication, July 27, 2018).

The reliability coefficient is a unitless index and can be compared from test to test and ranges from 0 to 1. For the CRCT in the years 2010 and 2011, the reliability coefficients ranged from .85 to .94. The values were consistent with previous administrations and suggested that the CRCT assessments were sufficiently reliable for their intended purpose. For the EOCT in years 2013 and 2014, the reliability coefficients ranged from .58 to .94. The values are in the range of industry standards for a criterion-referenced test like the EOCT, signifying that the EOCT assessments are sufficiently reliable for their intended purpose. The standard error of measurement value ranged from 3.10 and 3.82 for the EOCT in years 2013 and 2014, respectively, and 2.10 and 3.38 for the CRCT in the years 2010 and 2011, respectively. These values resulted in reasonably small error bands, indicating high reliability across various EOCT and CRCT. Additionally, for the CRCT in the years 2010 and 2011, CSEMs associated with the cut scale scores that define the performance levels ranged from 6 to 11 for the meets cut scores and from 9 to 14 for exceeds cut score. The CSEMs are consistent with prior administrations, meaning that the scores reported to students were well estimated and provided an accurate picture of their performance (J.D. Rollins, personal communication, July 27, 2018).

Data Collection

Once the Institutional Review Board (IRB) provided an exemption from IRB oversight as shown in Appendix A and the researched school district granted permission to collect research data, data were obtained from the district-level analyst employed by the school district. All required sixth and ninth-grade student data, with the exception of school minority data, school poverty data, and free/reduced meal status, were maintained in the Infinite Campus (IC) student data system. The studied school district had used Infinite Campus for over 10 years to house its student data and thus had sufficient longitudinal student data for this study. The Georgia Department of Education College and Career Performance Index (CCRPI) contains information pertaining to school minority and poverty data (Georgia Department of Education, 2019). Finally, student free/reduced meal data were housed in the Georgia School Nutrition database, Websmart. Students qualify for free or reduced-priced school meals based on family income or if any family member qualifies for public assistance, such as Supplemental Nutrition Assistance Program (SNAP) or Temporary Assistance for Needy Families (TANF). To maintain student confidentiality, students were identified using random numbers without the researcher ever knowing students' identities.

For this study, coding for both sixth and ninth-grade, non-school predictors included English language learner status, students with disability status, free or reduced meal price status, race, and gender were all coded dichotomously as either a 1 or 0. Students who qualified as an English language learner were coded as a 1 while those who were not were coded as a 0. Students who have a disability have an Individualized Education Plan (IEP). If a student was identified as having a disability, that student was coded with a 1 while a student without an IEP was coded as a 0. Students who qualify for free or reduced-priced meals were coded as a 1 while students who do not qualify were coded as a 0. For race, White students were coded as a 0 and non-White students were coded as a 1. For gender, males were coded as a 0 and females were coded as a 1.

Coding for both sixth and ninth-grade school predictors including attendance, course failure, number of school moves, standardized reading and math scores, school minority percentages, school poverty percentages, credits earned, and discipline infractions had a variety of coding processes. The detailed descriptions of their respective codes or values are listed in Tables 1 and 2 below.

Within the attendance policy of the studied school district were the following attendance guidelines that applied to unexcused absences:

- Students in kindergarten through eighth grades with 10 or more unexcused absences in one school year will not be promoted to the next grade level.

- Students in grades nine through 12 with five or more unexcused absences in any semester-long class will not receive credit for the course. Students in grades nine through 12 who take year-long classes with 10 or more unexcused absences will not receive credit for the course.
- Students who miss more than 10 consecutive days due to unexcused absences and who are not subject to compulsory attendance laws, who have not responded to efforts to get them to return to school, and who are not receiving instructional services from the school system through homebound instruction or instructional services required by the federal Individuals with Disabilities Act, withdrawn.

According to the school calendar of the studied school district, students attend school 180 days per year. Ninth graders who had an annual attendance rate of less than 90% per year were coded as a 0. Ninth graders who had an annual attendance rate of 90% and higher were coded as a 1. Sixth graders who had an annual attendance rate of less than 80% per year were coded as a 0. Sixth graders who had an annual attendance rate of 80% and higher were coded as a 1.

Variables

For the statistical analysis of the categorical dependent variable in this study, binary variables were utilized. The binary variables were coded as one of two values, a zero or a one. Table 1 has a list of the categorical variables and their respective binary codes.

Table 1

Binary Codes for Categorical Variables

Independent Variable	Binary Code - 0	Binary Code - 1
Graduate	Dropout or not graduate within four years	Graduate
6^{th} yearly English average	Fail	Pass
6^{th} yearly math average	Fail	Pass
Percent attendance in 6^{th}	< 80% attendance	≥ 80% attendance
Received OSS	No	Yes
Received free/reduced meal	No	Yes
Percent attendance in 9^{th}	< 90%	≥ 90% attendance
Earn 5 or more credits in 9^{th}	No	Yes
Failed < 2 core courses in 9^{th}	No	Yes
Received OSS	No	Yes
An English language learner	No	Yes
Student with disability	No	Yes
Race of student	White	Non-White
Gender of student	Male	Female

All categorical variables in both cohorts were coded as a 0 or 1 before the statistical analyses were performed. In addition, actual numeric values were used for continuous variables, and they are listed in Table 2.

Table 2

Continuous Variables That Utilized Actual Numeric Values

Independent Variables
<u>Ninth Grade</u>
Number of school moves in 9th grade
Scaled score on 9th grade Lit EOC test in 9th grade
Scaled score on Algebra EOC test in 9th grade
Percentage of minority students at the school attended in 9th grade
Percentage of economically disadvantaged students at the school attended in 9th grade
<u>Sixth Grade</u>
Number of school moves in 6th grade
Scaled score on reading CRCT in 6th grade
Scaled score on math CRCT in 6th grade
Percentage of minority students at the school attended in 6th grade
Percentage of economically disadvantaged students at the school attended in 6th grade

For ninth-grade course performance, the Chicago study of data from both the 2001 and 2004 freshman cohorts revealed that freshman-year course performance is strongly linked to high school graduation. An on-track freshman must have accumulated five full course credits, the number required to be promoted to the tenth grade, and receive no more than one semester F at the end of their freshman year (Allensworth & Easton, 2005). This on-track indicator is highly predictive of whether freshmen will eventually graduate. Students who were on track by the end of their freshman year were more than three and one-half times more likely to graduate high school in four years than off-track students (Allensworth & Easton, 2005). In the district participating in this study, five Carnegie units are also required for promotion to the tenth grade. If a student has accumulated five course credits and received no more than one semester F at the end of their freshman year, the student was coded as a 1. If this did not occur, the student was coded as a 0.

For sixth-grade course performance, the Baltimore Education Research Consortium (2011) studied sixth-grade students in the 2001 cohort and the 2009 cohort and found failing a core course in sixth grade is strongly associated with a lower likelihood of graduating high school. Less than a third of sixth graders who failed a core course in the study eventually graduated. Those who failed both English and math were particularly unlikely to graduate, as only 18.9% eventually graduated. For this study, if a student failed math during any grading period in the sixth grade, the student was coded as a 0, and students who did not fail math in the sixth grade were coded as a 1. Additionally, if a student failed English during any grading period in the sixth grade, the student was coded as a 0, and students who do not fail English in the sixth grade were coded as a 1. In the district of this study, a student fails a course if the grade is below 70.

Being suspended from school can be a trigger that puts a student on the path to ultimately dropping out. A suspension can lead to more suspensions, higher absenteeism, and course failure in the future. This progression can eventually lead to the path of dropping out (Balfanz et al., 2014). In the district of this study, a level III, IV, or V behavior violation could result in an out-of-school suspension from school. Some examples of these violations include fighting, drug or alcohol possession, and bullying. For both ninth graders and sixth graders, if the student was suspended out of school, the student was coded as a 1. If this particular punishment did not occur, the student would be coded as a 0.

Coding for both ninth and sixth-grade school predictors, including a student's number of school moves, along with their school minority and poverty percentages, were listed as a continuous variable because the actual number of occurrences or values were entered. Standardized reading and math scores were coded as scale scores. Finally, because a student's end status was either an on-time graduate or not, the criterion variable was a dichotomous variable and coded as a 1 for on-time graduates and 0 for those who were not on-time graduates. Each of the predictor variables with their respective code or value were entered in separate columns in Excel to prepare for their statistical analysis.

Data Analysis

The models used for this study predicted students who were likely to drop out of high school or not graduate within four years by identifying the predictor data that yielded the best performance. While utilizing training data, different models were compared by looking at their prediction when knowing the outcome. Descriptive statistics reported for the independent variables in each cohort include mean, standard deviation, skewness, and kurtosis. Student demographic characteristics were also reported.

This research analysis used correlation to indicate the relationships between the predictor data and the criterion variable of students being either a graduate or not. Pearson's correlation coefficient was used because the variables were measured on a combination of both interval and ratio scale. Pearson's correlation coefficient measures the strength of association between two variables. The measure of the strength ranges from -1 to +1, where -1 point to a perfect negative association, +1 indicates a perfect positive association, and 0 shows no association at all ("Conduct and Interpret a Point-Biserial Correlation," n.d.).

Developing the most accurate dropout early warning system involved quantifying the strength of the relationship between the predictors and the outcome. With a large number of attributes, an exploratory analysis of all the predictors may be less than feasible. An effective triaging strategy was to concentrate on predictors with strong relationships with the outcome and rank them based on their strength. Measuring the strength or relevance of the predictors helped to decide which ones to use as inputs in the model. The results of the filtering process helped create an accurate predictive model. Many of the ready-made predictive models available in R statistical software and available for use to answer the research questions in this study have intrinsic measurements of predictor importance (Kuhn & Johnson, 2013).

When assessing the quality of a model, accurately measuring its prediction error is important. One method used to determine the accuracy of the model involved splitting the data. Data splitting partitions the data into an explicit training data set used to prepare the model and another test data set used to evaluate the model's performance. Typically, the data set is split so that 60% is used for training, and 40% is used to evaluate the model's performance. This study used 2017’s data for training and 2018’s data to evaluate each model. Another method used to determine accuracy is plotting the receiver operating characteristic curve or ROC curve. The ROC curve is a graphical representation that plots the true-positive rate against the false-positive rate at various threshold settings (James et al., 2013).

Additionally, a confusion matrix is a table that allows visualization of the performance of an algorithm. A confusion matrix is used to describe the performance of a classification model on a set of test data for which the true values are known. The measure of model error that is used should be one that helps make a prediction model that most accurately predicts the desired target value for new data (Fortmann-Roe, 2012).

Logistic regression predicts the probability that the dependent variable belongs to one group or another with the probability falling somewhere between 0 and 1. The following equation is used in logistic regression to calculate probability:

$$P = \frac{e^{a + bX}}{1 + e^{a + bX}}$$

where:

p = the probability that a case is in a particular category,

e = the base of the natural logarithms (2.72),

a = the constant of the equation, and

b = the coefficient of the predictor variables.

By using logistic regression, schools can predict the probability of a student not graduating within four years of starting high school given the model of predictor variables, and have an early warning tool to help intervene appropriately.

Two other alternative and widely used statistical procedures utilized in this study were linear discriminate analysis and quadratic discriminate analysis. Linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) help to find the boundaries around the classification choices. Both models consist of statistical properties of the data, calculated for each class. For the multiple variables, the models estimate the mean and variance from the data for each class. Both the LDA and QDA algorithms make predictions by estimating the probability that a new set of inputs belongs to a particular class. The class with the highest probability is the output class and therefore, the prediction. Quadratic discriminate analysis allows for each class in the dependent variable to have its own covariance rather than a shared covariance as linear discriminate analysis does. This allows for quadratic terms in the development of the model (James et al., 2013).

Both LDA and QDA use Bayes' theorem to estimate the probability of the output class (k) given the input (x) using the probability of each class and the probability of the data belonging to each class (James et al., 2013). The formula for Bayes' theorem is:

$$\Pr(Y = k|X = x) = \frac{\pi_k f_k(x)}{\sum_{l=1}^{k} \pi_l f_l\,(x)}$$

Statistical learning allows for the rapid testing of a variety of statistical models while identifying their accuracy. This testing helps identify the best possible models and combinations of models given available data. The purpose of the selected models is to generate the most accurate prediction of the output on a variety of data sets and not just the one used to fit the model (Knowles, 2014). Overfitting models to sample data causes inaccurate out-of-sample predictions. Applied predictive modeling is well suited for creating a dropout early warning system because of its purpose to accurately classify new students each year based on ever-changing data (Knowles, 2014).

When using statistical modeling, data must first be trained to develop accurate predictive models and then score those models using additional data (Knowles, 2014). This was done at both the sixth and ninth-grade levels to ensure better accuracy. Finding the best predictive model to classify students as dropouts or not graduating on time focuses on identifying both the variables X and identifying the best *f (X)* such that:

$$Y = f\,(X\,) + \varepsilon$$

where:

f = is a function of X where $X = (X_1, X_2, \ldots, X_p)$ or p different predictors and

ε = is a random error term independent of X and has a mean of zero.

This study identified the subset of X that best predicts y given the function $f(X)$ (James et al., 2013).

Class Imbalance

Accuracy is not always a reliable metric for the real performance of a classifier because it will yield misleading results if the data set is unbalanced. Imbalanced classes are a common problem in machine learning classification. Algorithms in machine learning work best when the number of samples in each class are about equal. This balance is necessary because algorithms are designed to maximize accuracy and reduce errors (Boyle, 2019). To help with the imbalance of graduates and nongraduates, upsampling of the minority class and downsampling of the majority class of the data were performed before training the model. In upsampling, for every observation in the majority class, a randomly selected observation from the minority class with replacement was used. In downsampling, the majority class was randomly downsampled to be of the same size as the smaller class (Prabhakaran, 2017). The end result for both upsampling and downsampling is the same number of observations from the minority and majority classes.

In the 2017 cohort, which was used as training data, there were 973 students, and of those, 835 graduated within four years, and 138 either dropped out or did not graduate within four years. Thus, there were 835 observations of four-year graduates and 835 non-four-year graduates when the data set was upsampled and 138 observations of each when the data set was downsampled.

Assumptions

Assumptions for logistic regression include that the independent variables are linearly related to the log odds, and that the model should have little or no multicollinearity (Kassambara, 2018b). Linear discriminant analysis assumptions require multivariate normality for each level of the grouping variable, homogeneity of the variance and covariance (homoscedasticity), and little or no multicollinearity (Brownlee, 2016). Quadratic discriminant analysis has the same assumptions as linear discriminant analysis except that each class is required to have its own covariance matrix (Kassambara, 2018a). Additionally, both linear discriminant analysis and quadratic discriminant analysis are sensitive to outliers.

Logistic Regression Assumptions

Continuous Predictors Linearly Correlation to Logit of Outcome

The sixth-grade continuous variables consisting of mathcrct6 (actual scaled score on math CRCT), minority6 (actual percent of minority students in the school attended), readcrct6 (actual scaled score on reading CRCT), and trans6 (actual number of school moves) were all checked to determine if they were linearly related to the log odds. Both mathcrct6 (actual scaled score on math CRCT) and readcrct6 (actual scaled score on reading CRCT) were linearly correlated with the logit of the outcome. The trans6 variable (actual number of school moves) seemed to be nonlinearly correlated with the logit of the outcome, while minority6 (actual percent of minority students in the school attended) was not correlated in any way. The ninth-grade continuous variables consisting of algss9 (actual scaled score on Algebra EOC test), minority9 (actual percentage of minority students in the school attended), litss9 (actual scaled score on ninth Lit EOC test), and trans9 (actual number of school moves) were all checked to determine if they were linearly related to the log odds. Algss9 (actual scaled score on Algebra EOC test) and litss9 (actual scaled score on ninth Lit EOC test) showed weak linear correlation with the logit of the outcome, but minority9 (actual percent of minority students in the school attended) showed little to no relation with the logit of the outcome. The variable trans9 (actual number of school moves) showed a non-linear correlation with the logit of the outcome.

Pearson Correlation Coefficients

To determine if there was a correlation between quantitative variables, the Pearson correlation coefficient between all variables was calculated for both the ninth-grade data and sixth-grade data. The value of the Pearson correlation coefficient is a number between -1 and 1, and it indicates the extent to which two variables are related linearly. Each data set had one very high intercorrelation (i.e., multicollinearity) among the predictors, meaning one of the predictor variables was nearly perfectly predicted by one of the other predictor variables. Because of the problematic amounts of collinearity between the percentage of economically disadvantaged students and the percentage of minority students in the school along with the 100% value of economically disadvantaged students in the 2018 ninth-grade cohort, it was decided to eliminate the percentage of economically disadvantaged students variable from all calculations in the study.

Assumptions for Pearson correlation were met between the percentage of economically disadvantaged students and the percentage of minority students in the school, and those assumptions include that the variables were continuous, each observation had a pair of values, each variable did not have outliers, and the relationship between the two variables formed a straight line and was not curved.

Variance Inflation Factor (VIF)

Another method used to determine multicollinearity is the variance inflation factor (or VIF), which measures how much the variance of a regression coefficient is inflated due to multicollinearity in the model. The absence of multicollinearity has a VIF value of one. Typically, a VIF value that exceeds five to 10 indicates a problematic amount of collinearity, and the troubled variable should be removed. The troubled variable should be removed because the information the variable provides about the response is redundant while in the presence of the other variables.

The VIF value for the percentage of economically disadvantaged students and the percentage of minority students in the school exceeded five and thus had a problematic amount of collinearity for three of the cohorts and had a calculation error in the fourth cohort. There was a VIF value of 5.93 for the percentage of economically disadvantaged students for ninth graders in the 2017 cohort, and there was an error with the ninth-grade 2018 cohort calculation because of "aliased coefficients" in the model indicating at least one of the columns of the model matrix for the model was perfectly collinear with another. Additionally, there was a VIF value of 6.37 for the percentage of minority for ninth graders in the 2017 cohort, and there was also an error with the ninth-grade 2018 cohort calculation because of "aliased coefficients" in the model. No other variables had problematic amounts of collinearity. There was a VIF value of 7.52 for the percentage of economically disadvantaged students for sixth graders in the 2017 cohort, 5.71 for sixth graders in the 2018 cohort. Additionally, there was a VIF value of 8.14 for the percentage of minority students for sixth graders in the 2017 cohort and 6.11 for sixth graders in the 2018 cohort.

Because of the problematic amounts of collinearity with the percentage of economically disadvantaged students in the school along with the 100% value of that same variable in the 2018 ninth-grade cohort, it was decided to eliminate the percentage of economically disadvantaged students variable from all calculations in the study. To substantiate the decision to remove the economically disadvantaged percentage variable, the VIF calculations were rerun without the variable, and the results were good with no problematic amounts of collinearity among the remaining variables. The new VIF value for the percentage of minority students for ninth graders in the 2017 cohort was 1.40, for ninth graders in the 2018 cohort the value was 1.44, for sixth graders in the 2017 cohort the value was 1.42, and for sixth graders in the 2018 cohort the value was 1.43. With the economically disadvantaged percentage variable removed, the new VIF values fell well below the problematic threshold of collinearity for the percentage of minority students' variable.

Linear Discriminant Analysis and Quadratic Discriminant Analysis Assumptions

Multivariate Normality

The ninth-grade data and sixth-grade data were checked for multivariate normality utilizing a multivariate chi-square q-q plot. The plots showed that the data were not distributed in a multivariate normal fashion because for both grade levels there was a heavy concentration of data in the mean regions, but there were also heavy tails and skewness. Histograms confirmed the nonnormality of the data, as the vast majority of indicator variables were heavily skewed either left or right for both grade levels of data. Of the continuous variables, the histograms for algss9 (actual scaled score on Algebra EOC test), litss9 (actual scaled score on ninth Lit EOC test), readcrct6 (actual scaled score on reading CRCT), and mathcrct6 (actual scaled score on math CRCT) all showed some normality, but minority6 (actual percent of minority students in the school attended) and minority9 (actual percent of minority students in the school attended) had heavy tails on both ends, and the trans6 (actual number of school moves), and trans9 (actual number of school moves) variables were heavily skewed left.

Homoscedasticity

To check for homoscedasticity, the spread plot was used for both ninth-grade data and sixth-grade data. The data for ninth-grade variables and sixth-grade variables displayed heteroscedasticity, as the residuals nonlinearly trended downwards for both true-positive and true-negative observations. Therefore, the assumption of homoscedasticity was not met for linear discriminant analysis, but the assumption of each class having its own covariance matrix was met for quadratic discriminant analysis.

Outliers

Cook's distance was used to identify outliers for both sets of data. There was one observation with standardized residuals greater than three, and it was evident that the nongraduate observations yielded much larger residual terms than the graduate ones did. The model tended to see nongraduate observations as outliers because of their relative infrequency.

Multicollinearity

The Pearson correlation coefficient between all variables was calculated for both the ninth-grade data and sixth-grade data. Each data set had one very high intercorrelation—i.e., multicollinearity, among the predictors, meaning one of the predictor variables was nearly perfectly predicted by one of the other predictor variables. In addition, the VIF calculations done for the logistic regression assumptions apply to the linear discriminant analysis and quadratic discriminant analysis assumptions. Because of the problematic amounts of collinearity with the percentage of economically disadvantaged students in the school, along with the 100% value of that same variable in the 2018 ninth-grade cohort, it was decided to eliminate the percentage of the economically disadvantaged students variable from all calculations in the study. With the economically disadvantaged percentage variable removed, the new VIF values fell well below the problematic threshold of collinearity for the percentage of minority students' variable.

Summary

This chapter presented the details of the method and procedures for this nonexperimental, correlational study with the purpose of supporting the creation of a dropout early warning system to predict both middle school and high school students who are at risk for not graduating on time. Additionally, the most accurate indicators and statistical models will be identified to help pinpoint those students most at risk for not graduating high school within four years. The independent variables for this study were chosen based on a thorough literature review of dropout indicators at both the ninth-grade and sixth-grade levels and, thus, used as predictor variables in this study.

This study included all students who entered sixth grade in the 2010 and 2011 school years and ninth grade in the 2013 and 2014 school years in a mid-sized school district in Georgia and had on-time graduation years of 2017 and 2018. The data set was split so that the 2017 cohort data were used for training, and the 2018 cohort data were used to evaluate the models. The binary variables were coded as one of two values, a zero or a one, and actual numeric values were used for the continuous variables.

When utilizing the training data, different models were compared by looking at their prediction when knowing the outcome. To determine the accuracy of a model, two methods were used and were described in this chapter. One was plotting the receiver operating characteristic curve or ROC curve, which is a graphical representation that plots the true-positive rate against the false-positive rate at various threshold settings (James et al., 2013). Another method used to determine the accuracy of a model was a confusion matrix, which describes the performance of a classification model on a set of test data for which the true values are known.

The statistical procedures utilized in this study were logistic regression, linear discriminate analysis, and quadratic discriminate analysis. It should be noted that before those statistical models were used and to help with the imbalance of graduates and nongraduates, upsampling of the minority class and downsampling of the majority class of the data were performed. This chapter presented the assumptions for each of those statistical models and their respective results for this study.

Chapter IV

RESULTS

The purpose of this nonexperimental, correlational study was to use longitudinal data from a mid-sized school district from two cohorts to create a dropout early warning system to predict both middle school and high school students who are at risk for not graduating on time. Both sixth-grade and ninth-grade longitudinal data were used to create this dropout early warning system. A dropout early warning system can help school personnel plan for future dropout interventions. Analyzing longitudinal data allows educators to build an effective early warning system to identify potential high school dropouts and intervene more effectively and efficiently.

The following research questions were addressed in this study:

1. Does one or more of the ninth-grade variables consisting of attending at least 90% of the time, earning sufficient credits to move to the tenth grade, receiving out-of-school suspension, number of school moves, standardized reading and math scores, failing no more than one semester of a core content course, school minority percentages, school poverty percentages, ELL status, SWD status, free/reduced meal status, race, and gender accurately predict students who will not complete high school within four years?
2. Which statistical model is the most accurate at predicting future dropouts or late graduates utilizing ninth-grade variables with high levels of true classification and low levels of false identification?

3. Does one or more of the sixth-grade variables consisting of failing English, failing math, attending at least 80% of the time, receiving out-of-school suspension, number of school moves, standardized reading and math scores, school minority and poverty percentages, ELL status, SWD status, free/reduced meal status, race, and gender accurately predict students who will not complete high school within four years?
4. Which statistical model is the most accurate at predicting future dropouts or late graduates utilizing sixth-grade variables with high levels of true classification and low levels of false identification?

In general, this chapter includes descriptive statistics of student data in two cohorts, examinations of statistical results for each research question, as well as auxiliary findings. Statistical results presented include assumptions, logistic regression results, linear discriminant analysis results, quadratic discriminant analysis results, and accuracy results using ROC curve and the confusion matrix.

The results of this study helped determine if particular school-related variables and nonschool-related variables identified with high accuracy the students who would not graduate within four years, and if so, identify those variables. Accurately identifying potential dropouts is an important step for increasing the high school graduation rate. This study could add to the current body of knowledge about the variables that help identify students at risk for dropping out of high school and, thus, could have a positive impact on students' educational outcomes for many years to come.

Missing Data

A few variables in both cohorts had missing data. Because this missing data could undermine the ability to make valid inferences for this study, the missing data were imputed. The missing data were imputed utilizing the bagged tree model within the Caret package in R, a programming language for statistical computing and data visualization. For each predictor in the data, a bagged tree was created using all the other predictors in the data set. When a new sample had a missing value, the bagged tree model was used to predict it. Bagged trees helped to reduce overfitting by fitting several models on different samples with replacement. The 25 bagged tree models were then aggregated by using their average to derive the proper missing value.

To improve the statistical inferences, it was decided to impute all of the missing data. Per the district's Director of Assessment and Accountability who gathered the data for this study, standardized test scores were missing most likely because the students did not take the tests. The school minority and economically disadvantaged data were missing because there were a few programs in the district that were not considered a school. When the data were pulled for the school of enrollment, students enrolled in programs such as the alternative school were not defaulted to their home school enrollments, resulting in the missing values. Table 3 shows the percentages of missing data by cohort for the respective independent variables in this study. In both the 2017 and 2018 cohorts, for independent variables with missing data, reading and math CRCT scores had the highest percentages of missing data at around 10%. For both cohorts, school minority percentages and school economically disadvantaged percentages had the lowest amount of missing data at around 1%.

Table 3

Ninth-grade and Sixth-grade Percentage of Missingness by Cohort and Variable

	2017		2018	
Independent Variable	N	% of Students	N	% of Students
Ninth Grade				
Literature score	27	2.8	20	2.6
Algebra score	39	4.0	57	7.3
School minority percentage	23	2.4	24	3.1
School eco. disadv. percentage	23	2.4	0	0.0
Sixth Grade				
Reading CRCT score	91	9.4	72	9.3
Math CRCT score	98	10.1	78	10.1
School minority percentage	11	1.1	4	0.5
School eco. disadv. percentage	11	1.1	4	0.5

Descriptive Statistics

Up until the end of their fourth year of high school, students in both cohorts who left school and did not enroll in another school, did not graduate and were still enrolled, or whose status was unknown were coded as nongraduates. The reasons for not graduating within four years could include leaving to complete a GED, multiple course failures, illness, pregnancy, suspension, overage, or incarceration. Students who received their high school diploma within four years of starting ninth grade were coded as graduates. General student demographic data were also collected on the students in the two cohorts, including gender, race or ethnicity, ELL status, special education status, and free/reduced lunch status.

This study utilized data from two cohorts, one consisting of 2017 graduates and the other consisting of 2018 graduates. Each of the two cohorts contains both ninth-grade and sixth-grade data for the students in those cohorts. In the 2017 graduate cohort, consisting of 973 students, there were 835 (85.8%) graduates and 138 (14.2%) nongraduates, and in the 2018 graduate cohort, consisting of 776 students, there were 677 (87.2%) graduates and 99 (12.8%) nongraduates. The demographic characteristics for the 2017 and 2018 cohorts are presented in Table 4, and female-to-male ratios were relatively equal for both groups. Females outnumbered males in the 2017 cohort (51.9% vs. 48.1%) while males outnumbered females (50.8% vs. 49.2%) in the 2018 cohort. For both cohorts, the race or ethnicity majority in each group was Black, with 83.6% and 79.8% in the 2017 and 2018 cohorts, respectively. In the 2017 and 2018 cohorts, race of White made up 11.8% and 13.9%, respectively. There was less than 1% English language learners (ELL) in each cohort. There was also less than 10% of special education students in both cohorts, with 8.6% in the 2017 cohort and 7.7% in the 2018 cohort. Finally, a large majority of students in the 2017 and 2018 cohorts qualified for free or reduced meals at 85.7% and 84.3%, respectively.

Table 4

Demographic Characteristics for Students in the 2017 and 2018 Cohorts Containing Both Ninth-grade and Sixth-grade Data

	2017		2018	
Demographic Characteristics	N	% of Total Graduates and Nongraduates	N	% of Total Graduates and Nongraduates
Gender				
Female	505	51.9	382	49.2
Male	468	48.1	394	50.8
Race or ethnicity				
Black	813	83.6	619	79.8
White	115	11.8	108	13.9
Asian	5	0.5	13	1.7
Hispanic	27	2.8	13	1.7
Multiracial	13	1.3	22	2.8
Native American	0	0.0	1	0.1
ELL Status				
No	968	99.5	769	99.1
Yes	5	0.5	7	0.9
Special Education Status				
No	889	91.4	716	92.3
Yes	84	8.6	60	7.7
Free/Reduced Lunch Status				
No	139	14.3	122	15.7
Yes	834	85.7	654	84.3

The descriptive statistics for each ninth-grade and sixth-grade cohort are listed in Tables 5 through 8. Table 5 shows that in the 2017 ninth-grade cohort consisting of 973 students, students were a majority non-White and living in poverty, which is displayed by the district's data that had a large school minority percentage and a large school poverty percentage. The data also indicated that a large majority of students attended school over 90% of the time, earned sufficient credits to advance to the tenth grade, and received free or reduced lunch because those variables had mean values of 0.83, 0.87 and 0.86, respectively. Those three variables were all coded as a one for the affirmative and zero if not. There was also a low percentage of students who qualified for either ELL services or special education services because those variables had mean values of 0.01 and 0.09, respectively. The ELL and special education variables were also coded as a one for the affirmative and zero if not. The cohort had almost an even split of males and females, with females slightly outnumbering males. Gender had a mean value of 0.52, with females coded as a one and males codes as a zero.

Table 5

Ninth-grade 2017 Cohort Descriptive Statistics by Independent Variable

Independent Variable	*M*	*SD*	Skewness	Kurtosis
Graduate in four years (grad)	0.86	0.35	-2.05	2.21
Receive free or reduced meals (frl)	0.86	0.35	-2.04	2.16
Qualify as English learner (ell)	0.01	0.07	13.82	189.21
Qualify for special education services (swd)	0.09	0.28	2.94	6.66
Race (race)	0.88	0.32	-2.36	3.58
Gender (gend)	0.52	0.50	-0.08	-2
Attend at least 90% of days (abs9)	0.83	0.38	-1.74	1.03
Earn sufficient credits to advance to 10th grade (cred9)	0.87	0.33	-2.22	2.92
Failed less than two core courses (fail9)	0.71	0.45	-0.94	-1.11
If suspended out of school (oss9)	0.22	0.41	1.37	-0.12
Number or school moves in 9th grade (trans9)	0.39	0.93	3.81	21.75
Scaled score of ninth literature standardized test (litss9)	421.27	31.36	-0.13	2.16
Scaled score of algebra standardized test (algss9)	374.70	23.46	0.68	0.67
School minority percentage (minority9)	84.17	15.89	-0.85	-0.97
School poverty percentage (ed9)	80.08	10.93	-0.45	-1.14

Note. n = 973.

Table 6 shows that the 2018 ninth-grade cohort consisting of 776 students had almost 200 fewer students than the 2017 cohort. The descriptive statistics were very similar to the 2017 ninth-grade cohort except for the ed9 variable. The ed9 variable of 100.0 does not provide an accurate account of the proportion of economically disadvantaged students. In this ninth-grade cohort, the reported percentage of economically disadvantaged students in the respective schools was 100% because of the Community Eligibility Percentage (CEP) grant awarded to the district in 2018. The CEP grant resulted in all students in the schools receiving free meals.

For this ninth-grade cohort, like the 2017 cohort, the data indicated that a large majority of students attended school over 90% of the time, earned sufficient credits to advance to the tenth grade, and received free or reduced lunch because those variables had mean values of 0.87, 0.87, and 0.84 respectively. There was also a low proportion of students who qualified for either ELL services or special education services because those variables had mean values of 0.01 and 0.08, respectively. This cohort also had almost an even split of males and females, but males slightly outnumbered females. Gender had a mean value of 0.49.

Table 6

Ninth Grade 2018 Cohort Descriptive Statistics by Independent Variable

Independent Variable	*M*	*SD*	Skewness	Kurtosis
Graduate in four years (grad)	0.87	0.33	-2.23	2.97
Receive free or reduced meals (frl)	0.84	0.36	-1.88	1.54
Qualify as English learner (ell)	0.01	0.09	10.37	105.59
Qualify for special education services (swd)	0.08	0.27	3.16	7.99
Race (race)	0.86	0.35	-2.08	2.33
Gender (gend)	0.49	0.50	0.03	-2.00
Attend at least 90% of days (abs9)	0.87	0.33	-2.23	2.97
Earn sufficient credits to advance to 10th grade (cred9)	0.87	0.33	-2.25	3.05
Failed less than two core courses (fail9)	0.73	0.45	-1.02	-0.95
If suspended out of school (oss9)	0.25	0.43	1.14	-0.71
Number or school moves in 9th grade (trans9)	0.46	1.02	3.32	14.82
Scaled score of ninth literature standardized test (litss9)	485.11	46.60	-0.46	1.76
Scaled score of algebra standardized test (algss9)	475.56	34.05	0.60	1.62
School minority percentage (minority9)	83.27	15.36	-0.62	-1.33
School poverty percentage (ed9)	100.00	0.00	NA	NA

Note. n = 776.

Table 7 shows that for the 2017 sixth-grade cohort consisting of 973 students, a majority of the students were non-White and living in poverty, which is displayed by the district's data that showed a large school minority percentage and a large school poverty percentage. The data also indicated that a large majority of students attended school over 80% of the time, passed their English and math courses, and received free or reduced lunch because those variables had mean values of 0.99, 0.98, 0.97, and 0.86, respectively. Those four variables were all coded as a one for the affirmative and zero if not. There was also a low proportion of students who qualified for either ELL services or special education services because those variables had mean values of 0.01 and 0.09, respectively. In addition, a mean value of 0.17 indicated that only a few students were suspended out of school. The suspension variable was also coded as a one for the affirmative and zero if not. The cohort had almost an even split of males and females, with females slightly outnumbering males. Gender had a mean value of 0.52.

Table 7

Sixth Grade 2017 Cohort Descriptive Statistics by Independent Variable

Independent Variable	*M*	*SD*	Skewness	Kurtosis
Graduate in four years (grad)	0.86	0.35	-2.05	2.21
Receive free or reduced meals (frl)	0.86	0.35	-2.04	2.16
Qualify as English learner (ell)	0.01	0.07	13.82	189.21
Qualify for special education services (swd)	0.09	0.28	2.94	6.66
Race (race)	0.88	0.32	-2.36	3.58
Gender (gend)	0.52	0.50	-0.08	-2.00
Passed English in 6th grade (ela6)	0.98	0.15	-6.26	37.25
Passed math in 6th grade (math6)	0.97	0.18	-5.06	23.6
Attend at least 80% of days (abs6)	0.99	0.11	-8.82	75.93
If suspended out of school (oss6)	0.17	0.37	1.78	1.16
Number or school moves in 6th grade (trans6)	0.28	0.73	4.06	24.53
Scaled score of 6th reading standardized test (readcrct6)	823.02	22.32	0.49	0.74
Scaled score of 6th math standardized test (mathcrct6)	807.62	24.24	0.81	1.15
School minority percentage (minority6)	84.39	17.17	-0.97	-0.81
School poverty percentage (ed6)	82.64	13.49	-0.6	-1.15

Note. n = 973.

Table 8 shows that the 2018 sixth-grade cohort with 776 students had statistics that looked very similar to the 2017 sixth-grade cohort. The majority of the students were non-White and living in poverty, which is displayed by the district's data that showed a large school minority percentage and a large school poverty percentage. The data also indicated that a large majority of students attended school over 80% of the time, passed their English and math courses, and received free or reduced lunch because those variables had mean values of 0.99, 0.97, 0.98, and 0.84, respectively. There was also a low proportion of students who qualified for either ELL services or special education services because those variables had mean values of 0.01 and 0.08, respectively. In addition, a mean value of 0.18 indicated that only a few students were suspended out of school. The cohort had an almost even split of males and females, with males slightly outnumbering females. Gender had a mean value of 0.49.

Table 8

Sixth Grade 2018 Cohort Descriptive Statistics by Independent Variable

Independent Variable	*M*	*SD*	Skewness	Kurtosis
Graduate in four years (grad)	0.87	0.33	-2.23	2.97
Receive free or reduced meals (frl)	0.84	0.36	-1.88	1.54
Qualify as English learner (ell)	0.01	0.09	10.37	105.59
Qualify for special education services (swd)	0.08	0.27	3.16	7.99
Race (race)	0.86	0.35	-2.08	2.33
Gender (gend)	0.49	0.50	0.03	-2
Passed English in 6th grade (ela6)	0.97	0.18	-5.07	23.71
Passed math in 6th grade (math6)	0.98	0.15	-6.14	35.77
Attend at least 80% of days (abs6)	0.99	0.09	-11.22	124.01
If suspended out of school (oss6)	0.18	0.38	1.66	0.75
Number or school moves in 6th grade (trans6)	0.31	0.76	3.17	12.13
Scaled score of 6th reading standardized test (readcrct6)	826.09	21.43	0.52	1.09
Scaled score of 6th math standardized test (mathcrct6)	806.84	20.87	0.6	0.82
School minority percentage (minority6)	82.75	16.96	-0.75	-1.12
School poverty percentage (ed6)	82.56	13.89	-0.3	-1.59

Note. n = 776.

Pearson Correlation Coefficients

To determine if there was a correlation between quantitative variables, the Pearson correlation coefficient between all variables was calculated. The value of the Pearson correlation coefficient is a number between -1 and 1, and it indicates the extent to which two variables are related linearly. For both the sixth-grade data and ninth-grade data, each had one very high intercorrelation (i.e., multicollinearity) among the predictors, meaning one of the predictor variables was nearly perfectly predicted by one of the other predictor variables. A visual representation of the correlation matrix known as a correlogram is helpful and was created using the corrgram package in R and will be shown with each respective correlation matrix created in this study.

For the 2017 cohort of ninth-grade variables, there was a strong correlation between the percentage of minority students and the percentage of economically disadvantaged students in the school the students attended in ninth grade. The Pearson correlation coefficient indicated that the percentage of minority students and the percentage of economically disadvantaged students in the school were highly positively correlated, $r(971) = .91$, $p < .001$, which is indicative of multicollinearity. There was also a moderately positive correlation between literature and algebra standardized test scores, $r(971) = .65$, $p < .001$, and the correlation between race and percentage of school minority, $r(971) = .47$, $p < .001$. The Pearson correlation coefficient values and correlogram for the 2017 cohort of ninth-grade variables are shown in Table 9 and Figure 2.

Table 9

Pearson Correlation Coefficients Among Variables for the Ninth Grade 2017 Cohort

Variables	Grad	Frl	Ell	Swd	Race	Gend	Abs9
Graduate in four years (grad)	1.00						
Receive free or reduced meals (frl)	-0.10**	1.00					
Qualify as English learner (ell)	-0.01	0.03	1.00				
Qualify for special education services (swd)	-0.10**	0.05	-0.02	1.00			
Race (race)	-0.02	0.42***	0.03	0.02	1.00		
Gender (gend)	0.14***	0.03	-0.02	-0.13***	0.04	1.00	
Attend at least 90% of days (abs9)	0.35***	-0.15***	-0.01	-0.05	-0.07*	0.00	1.00
Earn sufficient credits to advance to 10th grade (cred9)	0.50***	-0.11***	-0.06	-0.16***	-0.04	0.13***	0.40***
Failed less than two core courses (fail9)	0.32***	-0.15***	-0.02	-0.11***	-0.06	0.16***	0.28***
If suspended out of school (oss9)	-0.33***	0.16***	0.00	0.06	0.11**	-0.10**	-0.36***
Number or school moves in 9th grade (trans9)	-0.28***	0.11***	-0.01	0.08*	0.10**	0.01	-0.33***
Scaled score of ninth literature standardized test (litss9)	0.27***	-0.33***	-0.05	-0.33***	-0.22***	0.09**	0.25***
Scaled score of algebra standardized test (algss9)	0.24***	-0.30***	-0.04	-0.23***	-0.23***	-0.05	0.21***
School minority percentage (minority9)	-0.04	0.31***	-0.08	0.03	0.47***	0.06	-0.12***
School poverty percentage (ed9)	-0.05	0.28***	-0.06	0.04	0.39***	0.00	-0.10**

Note. p < 0.001 '***', p < 0.01 '**', p < 0.05 '*'.

Table 9 (continued)

Pearson Correlation Coefficients Among Variables for the Ninth Grade 2017 Cohort

Variables	Cred9	Fail9	Oss9	Trans9	Litss9	Algss9	Minority9	Ed9
Earn sufficient credits to advance to 10th grade (cred9)	1.00							
Failed less than two core courses (fail9)	0.48***	1.00						
If suspended out of school (oss9)	-0.34***	-0.26***	1.00					
Number or school moves in 9th grade (trans9)	-0.28***	-0.10**	0.27***	1.00				
Scaled score of ninth literature standardized test (litss9)	0.34***	0.32***	-0.25***	-0.17***	1.00			
Scaled score of algebra standardized test (algss9)	0.30***	0.34***	-0.21***	-0.16***	0.65***	1.00		
School minority percentage (minority9)	0.03	0.04	0.03	0.10**	-0.21***	-0.21***	1.00	
School poverty percentage (ed9)	0.07*	0.10	0.04	0.08*	-0.20***	-0.16***	0.91***	1.00

Note. $p < 0.001$ '***', $p < 0.01$ '**', $p < 0.05$ '*'.

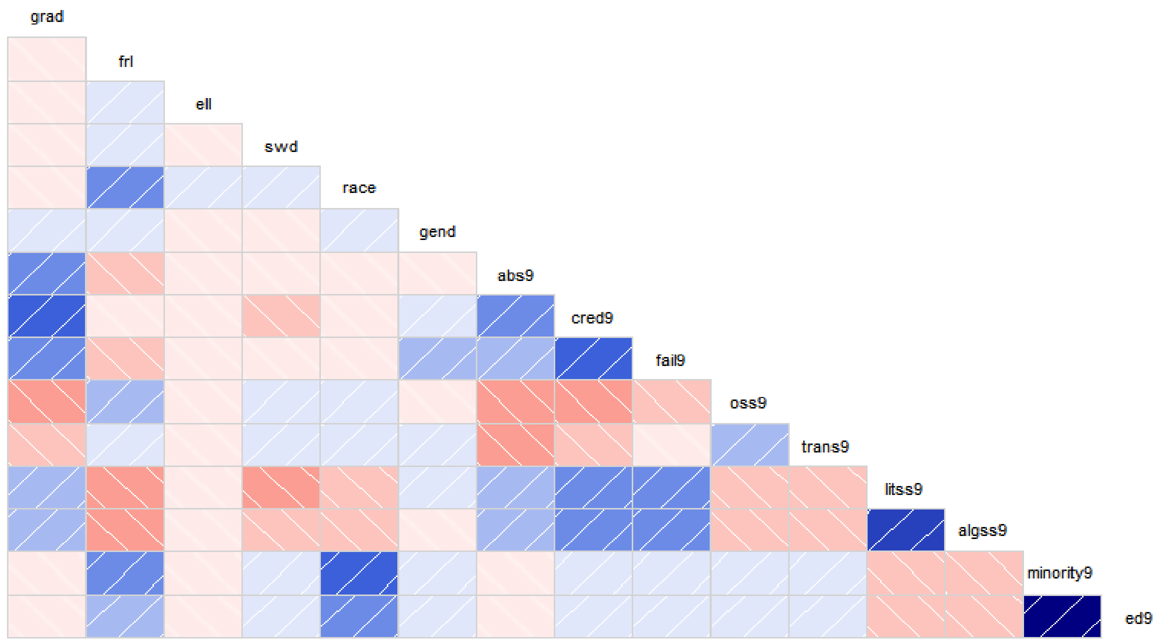

Figure 2. 2017 ninth-grade correlogram. The correlogram provides a visual display that represents the pattern of relations among the set of variables in terms of their correlations. Blue is for positive, and red is for negative. Each cell in the table shows the correlation between the two variables. The darker the hue, the greater the magnitude of the correlation.

For the 2018 cohort of ninth-grade variables, the percentage of economically disadvantaged students in the school was 100% for all schools because of the Community Eligibility Provision (CEP) grant. This grant gave all students free meals regardless of their household income, resulting in a value of 100% economically disadvantaged students in the school. This common value resulted in an error in the Pearson correlation coefficient calculation between the percentage of economically disadvantaged students in the school and all other variables. Additionally, the 2018 ninth-grade cohort showed a moderate correlation between literature and algebra standardized test scores with r(774) = .63, $p < .001$, and the correlation between race and percentage of school minority was weak to moderate at r(774) = .49, $p < .001$. The Pearson correlation coefficient values and correlogram for the 2018 cohort of ninth-grade variables are shown in Table 10 and Figure 3.

Table 10

Pearson Correlation Coefficients Among Variables for the Ninth Grade 2018 Cohort

Variables	Grad	Frl	Ell	Swd	Race	Gend	Abs9
Graduate in four years (grad)	1.00						
Receive free or reduced meals (frl)	-0.10*	1.00					
Qualify as English learner (ell)	0.00	0.04	1.00				
Qualify for special education services (swd)	-0.12**	0.09*	-0.03	1.00			
Race (race)	-0.01	0.49***	0.04	0.02	1.00		
Gender (gend)	0.13***	0.02	-0.01	-0.15***	0.04	1.00	
Attend at least 90% of days (abs9)	0.35***	-0.12***	0.04	-0.09*	-0.08*	0.04	1.00
Earn sufficient credits to advance to 10th grade (cred9)	0.40***	-0.09*	0.04	-0.09**	-0.05	0.13***	0.44***
Failed less than two core courses (fail9)	0.30***	-0.12***	0.03	-0.03	-0.08*	0.10**	0.23***
If suspended out of school (oss9)	-0.27***	0.19***	-0.06	0.06	0.17***	-0.11**	-0.31***
Number or school moves in 9th grade (trans9)	-0.24***	0.10**	-0.02	0.09**	0.13***	-0.04	-0.34***
Scaled score of ninth literature standardized test (litss9)	0.35***	-0.29***	0.02	-0.36***	-0.27***	0.18***	0.23***
Scaled score of algebra standardized test (algss9)	0.22***	-0.30***	0.06	-0.27***	-0.29***	0.03	0.16***
School minority percentage (minority9)	-0.08*	0.36***	-0.06	0.02	0.49***	-0.02	-0.13***
School poverty percentage (ed9)	NA	NA	NA	NA	NA	NA	NA

Note. p < 0.001 '***', p < 0.01 '**', p < 0.05 '*'.

Table 10 (continued)

Pearson Correlation Coefficients Among Variables for the Ninth Grade 2018 Cohort

Note. $p < 0.001$ '***', $p < 0.01$ '**', $p < 0.05$ '*'.

Variables	Cred9	Fail9	Oss9	Trans9	Litss9	Algss9	Minority9	Ed9
Earn sufficient credits to advance to 10th grade (cred9)	1.00							
Failed less than two core courses (fail9)	0.40***	1.00						
If suspended out of school (oss9)	-0.31***	-0.24***	1.00					
Number or school moves in 9th grade (trans9)	-0.38***	-0.11**	0.22***	1.00				
Scaled score of ninth literature standardized test (litss9)	0.34***	0.30***	-0.33***	-0.20***	1.00			
Scaled score of algebra standardized test (algss9)	0.24***	0.24***	-0.26***	-0.18***	0.63***	1.00		
School minority percentage (minority9)	-0.01	-0.13***	0.21***	0.06	-0.27***	-0.23***	1.00	
School poverty percentage (ed9)	NA	NA	NA	NA	NA	NA	NA	NA

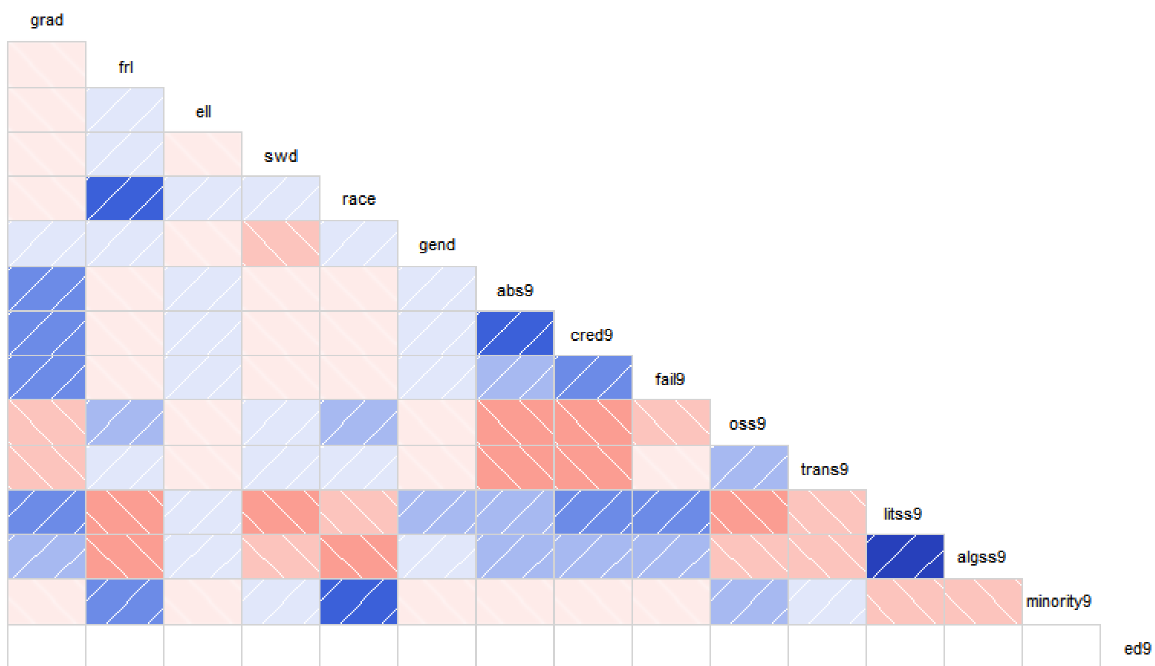

Figure 3. 2018 ninth-grade correlogram. The correlogram provides a visual display that represents the pattern of relations among the set of variables in terms of their correlations. Blue is for positive, and red is for negative. Each cell in the table shows the correlation between the two variables. The darker the hue, the greater the magnitude of the correlation.

For the 2017 cohort of sixth-grade variables, there was a strong correlation between the percentage of minority students and the percentage of economically disadvantaged students in the school the students attended in sixth grade. Using the Pearson correlation coefficient for the 2017 cohort coefficient, the percentage of minority students and the percentage of economically disadvantaged students in the school were found to be highly positively correlated at $r(971) = .93$, $p < .001$, which is indicative of multicollinearity. There was a moderate correlation between reading CRCT scores and math CRCT scores at $r(971) = .69$, $p < .001$. The correlation between race and school minority percentage was weak to moderate at $r(971) = .50$, $p < .001$. The Pearson correlation coefficient values and correlogram for the 2017 cohort of sixth-grade variables are shown in Table 11 and Figure 4.

Table 11

Pearson Correlation Coefficients Among Variables for the Sixth Grade 2017 Cohort

Variables	Grad	Frl	Ell	Swd	Race	Gend	Ela6
Graduate in four years (grad)	1.00						
Receive free or reduced meals (frl)	-0.10**	1.00					
Qualify as English learner (ell)	-0.01	0.03	1.00				
Qualify for special education services (swd)	-0.10**	0.05	-0.02	1.00			
Race (race)	-0.02	0.42***	0.03	0.02	1.00		
Gender (gend)	0.14***	0.03	-0.02	-0.13***	0.04	1.00	
Passed English in 6th grade (ela6)	0.21***	-0.04	0.01	-0.17***	-0.04	0.04	1.00
Passed math in 6th grade (math6)	0.12**	-0.05	-0.06*	-0.02	-0.03	0.07*	0.23***
Attend at least 80% of days (abs6)	0.19***	-0.05	0.01	-0.07*	-0.04	0.02	0.17***
If suspended out of school (oss6)	-0.20***	0.12***	-0.03	0.12***	0.10**	-0.10**	-0.15***
Number or school moves in 6th grade (trans6)	-0.26***	0.10**	-0.01	0.05	0.10**	0.01	-0.21***
Scaled score of 6th reading standardized test (readcrct6)	0.17***	-0.30***	-0.05	-0.26***	-0.23***	0.07*	0.11***
Scaled score of 6th math standardized test (mathcrct6)	0.18***	-0.30***	-0.01	-0.24***	-0.25***	0.00	0.10**
School minority percentage (minority6)	-0.07*	0.33***	-0.06	0.01	0.50***	0.08*	-0.02
School poverty percentage (ed6)	-0.07*	0.33***	-0.04	0.00	0.44***	0.05	-0.03

Note. p < 0.001 '***', p < 0.01 '**', p < 0.05 '*'.

Table 11 (continued)

Pearson Correlation Coefficients Among Variables for the Sixth Grade 2017 Cohort

Variables	Math6	Abs6	Oss6	Trans6	Read crct6	Math crct6	Minority6	Ed6
Passed math in 6th grade (math6)	1.00							
Attend at least 80% of days (abs6)	0.18***	1.00						
If suspended out of school (oss6)	-0.12***	-0.20***	1.00					
Number or school moves in 6th grade (trans6)	-0.15***	-0.38***	0.26***	1 .00				
Scaled score of 6th reading standardized test (readcrct6)	0.09**	0.04	-0.15***	-0.12***	1.00			
Scaled score of 6th math standardized test (mathcrct6)	0.17***	0.06	-0.21***	-0.12***	0.69***	1.00		
School minority percentage (minority6)	-0.02	-0.05	0.12***	0.12***	-0.22***	-0.23**	1.00	
School poverty percentage (ed6)	-0.05	-0.07*	0.10**	0.12***	-0.20***	-0.18**	0.93*	1.00

Note. $p < 0.001$ '***', $p < 0.01$ '**', $p < 0.05$ '*'.

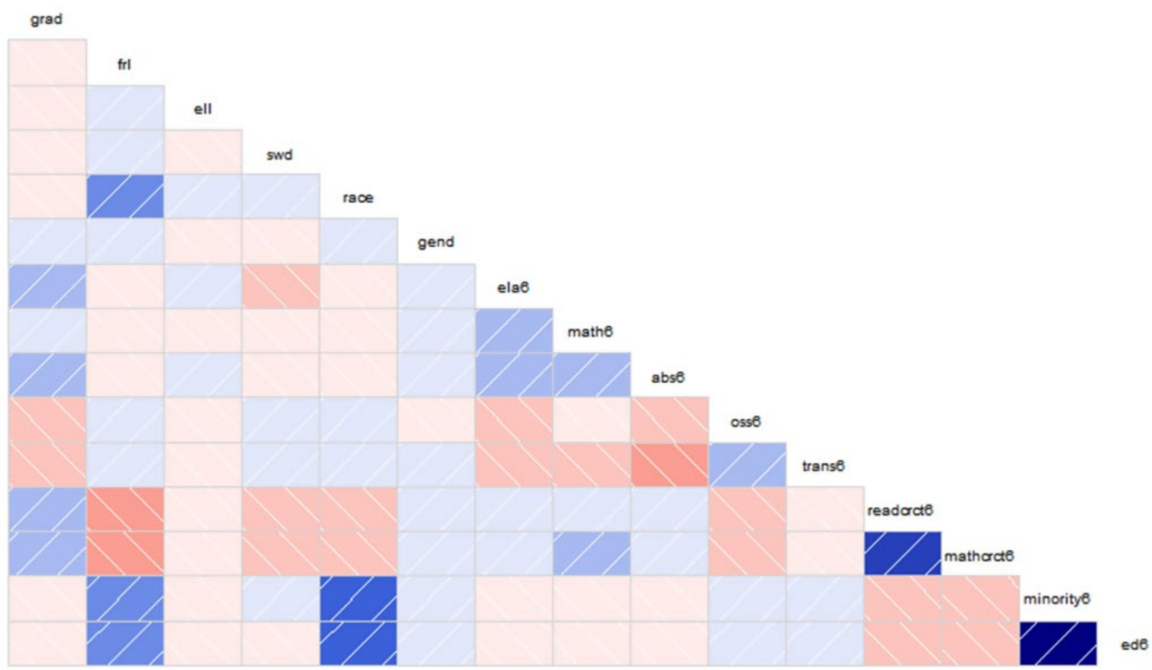

Figure 4. 2017 sixth-grade correlogram. The correlogram provides a visual display that represents the pattern of relations among the set of variables in terms of their correlations. Blue is for positive, and red is for negative. Each cell in the table shows the correlation between the two variables. The darker the hue, the greater the magnitude of the correlation.

For the 2018 cohort of sixth-grade variables, there was again a strong correlation between the percentage of minority students and the percentage of economically disadvantaged students in the school the students attended in sixth grade. Again, using the Pearson correlation coefficient for the 2018 cohort coefficient, the percentage of minority students and the percentage of economically disadvantaged students in the school were found to be highly positively correlated at r(774) = .91, $p < .001$, which is indicative of multicollinearity. Just like in the 2017 cohort, there was a moderate correlation between the reading CRCT and math CRCT in the 2018 sixth-grade cohort at r(774) = .68, $p < .001$. There was also a weak to moderate correlation between race and school minority percentage at r(774) = .49, $p < .001$. The Pearson correlation coefficient values and correlogram for the 2018 cohort of sixth-grade variables are shown in Table 12 and Figure 5.

Table 12

Pearson Correlation Coefficients Among Variables for the Sixth Grade 2018 Cohort

Variables	Grad	Frl	Ell	Swd	Race	Gend	Ela6
Graduate in four years (grad)	1.00						
Receive free or reduced meals (frl)	-0.10*	1.00					
Qualify as English learner (ell)	0.00	0.04	1.00				
Qualify for special education services (swd)	-0.12***	0.09*	-0.03	1.00			
Race (race)	-0.01	0.49***	0.04	0.02	1.00		
Gender (gend)	0.13***	0.02	-0.01	-0.15***	0.04	1.00	
Passed English in 6th grade (ela6)	0.14***	-0.04	0.02	-0.10**	-0.02	0.12**	1.00
Passed math in 6th grade (math6)	0.14***	-0.07	0.02	-0.08*	0.01	0.07*	0.38***
Attend at least 80% of days (abs6)	0.05	-0.04	0.01	0.03	0.01	0.00	0.14***
If suspended out of school (oss6)	-0.25***	0.17***	-0.04	0.05	0.14***	-0.14***	-0.15***
Number or school moves in 6th grade (trans6)	-0.18***	0.08*	-0.02	0.07*	0.12**	-0.06	-0.13***
Scaled score of 6th reading standardized test (readcrct6)	0.19***	-0.34***	-0.03	-0.22***	-0.38***	0.15***	0.13***
Scaled score of 6th math standardized test (mathcrct6)	0.22***	-0.30***	0.01	-0.14***	-0.32***	0.05	0.20***
School minority percentage (minority6)	-0.07	0.39***	-0.04	0.06	0.49***	0.01	-0.02
School poverty percentage (ed6)	-0.07	0.37***	-0.05	0.06	0.43***	0.01	-0.06

Note. $p < 0.001$ ‘***’, $p < 0.01$ ‘**’, $p < 0.05$ ‘*’.

Table 12 (continued)

Pearson Correlation Coefficients Among Variables for the Sixth Grade 2018 Cohort

Variables	Math6	Abs6	Oss6	Trans6	Read crct6	Math crct6	Minority6	Ed6
Passed math in 6th grade (math6)	1.00							
Attend at least 80% of days (abs6)	0.18***	1.00						
If suspended out of school (oss6)	-0.16***	-0.07*	1.00					
Number or school moves in 6th grade (trans6)	-0.08*	-0.10**	0.22***	1.00				
Scaled score of 6th reading standardized test (readcrct6)	0.13***	-0.01	-0.15***	-0.14***	1.00			
Scaled score of 6th math standardized test (mathcrct6)	0.18***	0.07	-0.22***	-0.15***	0.68***	1.00		
School minority percentage (minority6)	0.03	-0.05	0.18***	0.07	-0.30***	-0.30***	1.00	
School poverty percentage (ed6)	0.02	-0.05	0.19***	0.09*	-0.27***	-0.27***	0.91***	1.00

Note. $p < 0.001$ '***', $p < 0.01$ '**', $p < 0.05$ '*'.

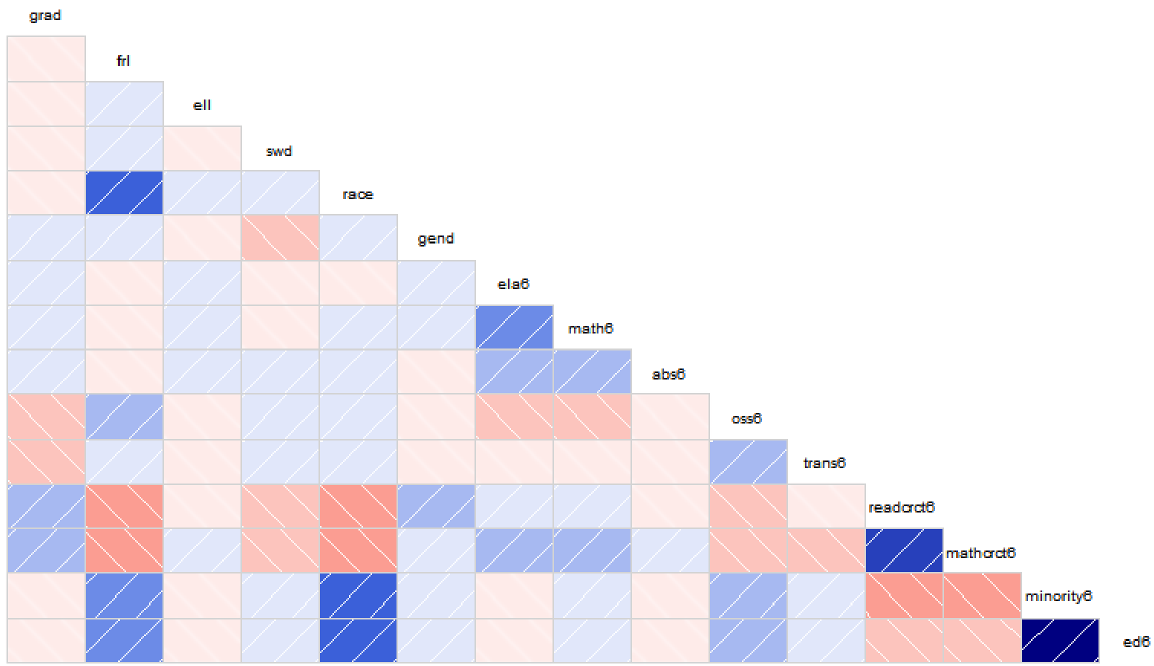

Figure 5. 2018 sixth-grade correlogram. The correlogram provides a visual display that represents the pattern of relations among the set of variables in terms of their correlations. Blue is for positive, and red is for negative. Each cell in the table shows the correlation between the two variables. The darker the hue, the greater the magnitude of the correlation.

The Pearson correlation coefficient between all variables was calculated for both the ninth-grade data and sixth-grade data. Each data set had one high intercorrelation (i.e., multicollinearity) among the predictors, meaning one of the predictor variables was nearly perfectly predicted by one of the other predictor variables. Because of the high collinearity between the school minority percentage and school economically disadvantaged percentage, the school economically disadvantaged percentage variable was eliminated from all calculations in the study. To further support this decision, there was a 100% value of economically disadvantaged students in the 2018 ninth-grade cohort. This high percentage was a result of all high schools in the district qualifying for the CEP grant, which provided free meals to all students in those schools.

RQ1**:** Does one or more of the ninth-grade variables consisting of attending at least 90% of the time, earning sufficient credits to move to the tenth grade, receiving out-of-school suspension, number of school moves, standardized reading and math scores, failing no more than one semester of a core content course, school minority percentages, school poverty percentages, and student characteristics (ELL status, SWD status, free/reduced meal status, race, and gender) accurately predict students who will not complete high school within four years?

Three statistical analyses were utilized to answer this question. The statistical analyses used to provide confirmation that the correct variables were identified were logistic regression, linear discriminant analysis, and quadratic discriminant analysis. Even though the three analyses often reveal the same patterns in a set of data, they do so in different ways and require different assumptions. Logistic regression calculates the odds of the outcome as the ratio of the probability of having the outcome divided by the probability of not having it. Linear discriminant analysis and quadratic discriminant analysis are similar classification methods that are used to determine which set of variables discriminate between two groups and to classify an observation into one of the groups. Quadratic discriminant analysis is more flexible than linear discriminant analysis. Each analysis method helped identify which independent variables had the most impact on the dependent variable and understand how each variable contributed towards the classification. The results from all three statistical analyses provided a more complete picture of which variables are able to predict students who will not complete high school within four years.

To help with the imbalance of graduates and nongraduates, upsampling of the minority class and downsampling of the majority class of the data were performed before training the model. The end result for both upsampling and downsampling is the same number of observations from the minority and majority classes. When running each analysis, the two cohorts were used as training and testing sets. Each model could train using roughly half the available data and then test on the other half. For the ninth-grade cohort, the 2017 variable data was used as the training data, and the 2018 variable data was used to test the accuracy of the model.

Logistic Regression

The logistic regression analysis finds the best fitting model to describe the relationship between a student's graduation status and a variety of student and school variables identified through the review of literature. This statistical model can estimate the probability of certain events occurring based on some previous data. In this study, logistic regression was used to predict whether or not a student would graduate within four years based on known student ninth-grade data. The results of the upsampled logistic regression prediction for the ninth-grade cohort are listed in Table 13, and the downsampled logistic regression prediction results are listed in Table 14.

Using the upsampled logistic regression model to increase the size of the training set resulted in an accuracy level of 89.0%, and its statistical output can be seen in Table 13. Another measure of the quality of the model is the Akaike Information Criterion (AIC) value. If two similar models are being compared, then the model with the lower

AIC is superior. The upsampled logistic regression model had an AIC level of 1545.20 and can be compared to the downsampled logistic regression model's AIC level to identify which model is preferred.

The Hosmer and Lemeshow goodness of fit test results (chi-square 52.52, df = 11, $p < 0.001$) indicate that the observed nongraduates differ significantly from the expected value of nongraduates using the predictor variables based on their proportion of the population. The chi-square value is above the critical value with a degrees of freedom value of 11 at the $p < 0.05$ level.

The Nagelkerke pseudo R-squared value is helpful when it is compared to another pseudo R-squared of the same type and predicting the same outcome. This value can be helpful when deciding which logistic regression model is best. A higher pseudo R-squared indicates which model does a better job of predicting the outcome (UCLA: Statistical Consulting Group., n.d.). The upsampled logistic regression model had a Nagelkerke pseudo R-squared value of 0.492 and can be compared to the downsampled logistic regression model's pseudo R-squared level to identify which model is preferred.

Of the 13 predictor variables, eight were statistically significant: race, gender, attend at least 90% of days, earn sufficient credits to advance to tenth grade, failed less than two core courses, suspended out of school, number of school moves in ninth grade, and scaled score of algebra standardized test. By earning sufficient credits to advance to the tenth grade, a student increases their odds of graduating by a factor of 5.2, given all other variables are unchanged. If a student received an out-of-school suspension, the odds them graduating decreased by 57.9% (0.421 – 1 = -0.579), keeping other variables constant.

Interestingly, the estimate for algebra standardized test score variable (z = 5.905, $p < .001$, odds ratio = 1.022, 95% CI = 1.014 to 1.030) was positive (with low magnitude), while the ninth-grade literature standardized test score variable (z = 1.190, $p = 0.234$, odds ratio = 1.003, 95% CI = 0.997 to 1.009) was not significant. This difference suggests that algebra scores were much more indicative than literature scores of a student's likelihood to graduate high school.

To identify the variables that are stronger in their role as a predictor, z-statistic values that are farther away from zero identify the strongest predictors. The variables of race (z = 3.103, $p = 0.002$, odds ratio = 2.472, 95% CI = 1.511 to 4.084) and failing less than two core courses (z = 3.018, $p = 0.003$, odds ratio = 1.608, 95% CI = 1.195 to 2.162) were associated with a slight increase in graduating high school.

Even higher graduation percentages were associated with the variables of gender (z = 6.656, $p < .001$, odds ratio = 1.927, 95% CI = 1.491 to 2.495), attending school at least 90% of days (z = 4.333, $p < .001$, odds ratio = 1.960, 95% CI = 1.415 to 2.719), earning sufficient credits to advance to the tenth grade (z = 7.939, $p < .001$, odds ratio = 5.234, 95% CI = 3.527 to 7.876), and scaled score of algebra standardized test (z = 5.905, $p < .001$, odds ratio = 1.022, 95% CI = 1.014 to 1.030). In addition, lower chances of graduating high school were associated with the fact that a student was suspended out of school (z = -5.993, $p < .001$, odds ratio = 0.421, 95% CI = 0.314 to 0.562) and the number of school moves a student made in ninth grade (z = -4.994, $p < .001$, odds ratio = 0.656, 95% CI = 0.572 to 0.749).

Table 13

Upsampled Ninth-Grade Variables Used to Predict Graduation Utilizing Logistic Regression (Testing Data)

Independent Variable	Regression Coefficient	*SE*	*z*	Pr(>|z|)		*OR*
(Intercept)	-12.235	1.588	-7.704	*p* < .001	***	-
Receive free or reduced meals (frl)	-0.220	0.242	-0.913	0.361		0.650
Qualify as English learner (ell)	0.345	0.835	0.413	0.680		1.143
Qualify for special education services (swd)	0.328	0.222	1.480	0.139		1.468
Race (race)	0.790	0.255	3.103	0.002	**	2.472
Gender (gend)	0.881	0.132	6.656	*p* < .001	***	1.927
Attend at least 90% of days (abs9)	0.720	0.166	4.333	*p* < .001	***	1.960
Earn sufficient credits to advance to 10th grade (cred9)	1.619	0.204	7.939	*p* < .001	***	5.234
Failed less than two core courses (fail9)	0.449	0.149	3.018	0.003	**	1.608
If suspended out of school (oss9)	-0.888	0.148	-5.993	*p* < .001	***	0.421
Number of school moves in 9th grade (trans9)	-0.344	0.069	-4.994	*p* < .001	***	0.656
Scaled score of ninth literature standardized test (litss9)	0.004	0.003	1.190	0.234		1.003
Scaled score of algebra standardized test (algss9)	0.024	0.004	5.905	*p* < .001	***	1.022
School minority percentage (minority9)	-0.008	0.005	-1.635	0.102		0.991

Note. AIC: 1545.20. p < 0.001 '***', p < 0.01 '**', p < 0.05 '*'.

Using the downsampled logistic regression model to decrease the size of the training set decreased model accuracy to 88.1%, and, as a result, the model detected only a few significant coefficients. The downsampled logistic regression model had an AIC level of 272.36, and when compared to the upsampled logistic regression model's AIC level of 1545.2, the downsampled model is identified as the preferred model. The downsampled logistic regression model's statistical output can be seen in Table 14. This declaration seems to contradict the accuracy percentage and number of significant variables.

The Hosmer and Lemeshow goodness of fit test results (chi-square 28.18, df = 11, $p = 0.003$) indicate that the observed nongraduates differ significantly from the expected value of nongraduates using the predictor variables based on their proportion of the population. The chi-square value is above the critical value with a degrees of freedom value of 11 at the $p < 0.05$ level.

Remember, a higher Nagelkerke pseudo R-squared indicates which model does a better job of predicting the outcome (UCLA: Statistical Consulting Group., n.d.). The downsampled logistic regression model had a pseudo R-squared value of 0.518. When compared to the upsampled logistic regression model's pseudo R-squared level of 0.492, the downsampled model is again identified as the preferred model.

Of the 13 predictor variables, three variables were statistically significant: gender, earning sufficient credits to advance to tenth grade, and the number of school moves a student makes in the ninth grade. By earning sufficient credits to advance to the tenth grade, a student increases their odds of graduating by a factor of 4.3, given that all other variables are unchanged. If a student had multiple moves in the ninth grade, the odds of them graduating decreased by 28.9% (0.711 – 1 = -.289), keeping all other variables constant.

The z-statistic was again used to identify the variables that were stronger in their role as a predictor. Z-statistic values that are farther away from zero identify the strongest predictors. The variables identified as gender (z = 2.700, $p = 0.007$, odds ratio = 2.060, 95% CI = 1.097 to 3.928) and earning sufficient credits to advance to the tenth grade (z = 4.298, $p < .001$, odds ratio = 4.321, 95% CI = 1.690 to 1.184) were associated with an increased chance of graduating high school. Having a lower chance of graduating high school was associated with the number of school moves a student made in ninth grade (z = -1.960, $p = .050$, odds ratio = 0.711, 95% CI = 0.510 to 0.961).

Table 14

Downsampled Ninth-Grade Variables Used to Predict Graduation Utilizing Logistic Regression (Testing Data)

Independent Variable	Regression Coefficient	*SE*	*z*	Pr(>|z|)		*OR*
(Intercept)	-12.996	3.989	-3.258	0.001	**	-
Receive free or reduced meals (frl)	-0.706	0.665	-1.062	0.288		0.429
Qualify as English learner (ell)	2.982	1.522	1.959	0.050		0.000
Qualify for special education services (swd)	0.622	0.590	1.055	0.291		1.765
Race (race)	1.353	0.754	1.794	0.073		4.238
Gender (gend)	0.910	0.337	2.700	0.007	**	2.060
Attend at least 90% of days (abs9)	0.666	0.413	1.613	0.107		1.774
Earn sufficient credits to advance to 10th grade (cred9)	2.686	0.625	4.298	*p* < .001	***	4.321
Failed less than two core courses (fail9)	0.297	0.363	0.818	0.414		1.580
If suspended out of school (oss9)	-0.513	0.366	-1.402	0.161		0.498
Number of school moves in 9th grade (trans9)	-0.320	0.163	-1.960	0.050	*	0.711
Scaled score of ninth literature standardized test (litss9)	0.011	0.008	1.460	0.144		1.007
Scaled score of algebra standardized test (algss9)	0.014	0.010	1.364	0.173		1.027
School minority percentage (minority9)	-0.002	0.015	-0.140	0.889		0.994

Note. AIC: 272.36. p < 0.001 ‘***’, p < 0.01 ‘**’, p < 0.05

Linear Discriminant Analysis

Like logistic regression, linear discriminant analysis (LDA) predicts a dependent variable based on the values of independent variables. Linear discriminant analysis makes predictions by estimating the probability that new inputs belong to a particular class. In this study, graduating or not graduating within four years was the prediction being made.

The 2017 cohort was used as training data, and the 2018 cohort was used as testing data. The results of the upsampled linear discriminant analysis prediction are listed in Table 15, and the results for the downsampled linear discriminant analysis prediction are listed in Table 16.

The coefficients of linear discriminants were the values used to predict whether or not each student would graduate within four years. The coefficients are similar to regression coefficients. The category with the highest probability was, therefore, the prediction. The higher the coefficient, the more weight it had.

As shown in Table 15, the upsampled LDA model with an accuracy of 89.0% had some results that closely aligned with those of the upsampled logistic regression models. The coefficients of linear discriminants for race (0.466), gender (0.568), attend at least 90% of days (0.519), earning sufficient credits to advance to tenth grade (0.985), failing less than two core courses (0.499), and suspended out of school (-0.550) show that those predictors were the most influential in determining if a student graduated high school, as they were the ones with the highest magnitude.

Earning sufficient credits to advance to the tenth grade was still the most influential predictor with a coefficient of 0.985. The group means confirmed this, as only 47.1% of nongraduates earned sufficient credits to advance to the tenth grade, while 93.9% of graduates did. Gender had the second-highest weighted coefficient of 0.568. The difference in gender suggests that resampling the training set's nongraduates consisted predominantly of males and that the model utilized that difference heavily in determining a decision boundary in this model. Another strong predictor in the upsampled LDA model was fail9 (coefficient of 0.499), suggesting that the students who fail less than two classes are much more likely to graduate. The group means confirmed this, as 77.1% of graduates failed less than two classes, and only 35.6% of nongraduates failed less than two classes in the ninth grade. Race also had a high weighted coefficient of 0.466, but the group means did not confirm it because the nongraduate and graduate means were very close at 0.920 and 0.879, respectively.

There was a drastic difference in group means for earning sufficient credits to advance to tenth grade, as the nongraduates had a group mean of 0.471, and the graduates had a group mean of 0.939. The group means showed that 93.9% of the graduates earned sufficient credits to advance to the 10th grade, while only 47.1% of the nongraduates did. Similarly, only 16.2% of graduates received out-of-school suspension while 53.4% of nongraduates received out-of-school suspension, and 88.1% of graduates attended at least 90% of school days while only 48.9% nongraduates attended at least 90% of school days. These results mean that students who earned sufficient credits to advance to the 10th grade had fewer suspensions and attended more school were all more likely to be on track to graduate.

What is surprising is that the coefficients of linear discriminants for both test score variables were near-zero, suggesting that testing performance had little effect on graduation likelihood in this model. The effort-based attributes, such as attendance and passing class, were much more influential.

Table 15

Upsampled Ninth-Grade Variables Used to Predict Graduation Utilizing LDA (Testing Data)

Independent Variable	Nongraduate Means	Graduate Means	Coefficients of Linear Discriminants: LD1
Receive free or reduced meals (frl)	0.951	0.843	-0.151
Qualify as English learner (ell)	0.012	0.005	-0.356
Qualify for special education services (swd)	0.149	0.075	0.197
Race (race)	0.920	0.879	0.466
Gender (gend)	0.337	0.547	0.568
Attend at least 90% of days (abs9)	0.489	0.881	0.519
Earn sufficient credits to advance to 10th grade (cred9)	0.471	0.939	0.985
Failed less than two core courses (fail9)	0.356	0.771	0.499
If suspended out of school (oss9)	0.534	0.162	-0.550
Number of school moves in 9th grade (trans9)	1.059	0.289	-0.147
Scaled score of ninth literature standardized test (litss9)	399.969	424.662	0.004
Scaled score of algebra standardized test (algss9)	360.599	376.995	0.013
School minority percentage (minority9)	86.465	83.913	-0.008

Note. Prior probabilities of groups: nongraduate: 0.5, graduate: 0.5.

As shown in Table 16, the downsampled LDA model had an accuracy score of 88.7%, which was slightly less than the upsampled LDA model. This model's coefficients generally confirm the results of the upsampled model, with several notable exceptions. The most interesting observation was if students qualified as an English learner. The English learner predictor had the largest coefficient in this model (-1.892), while it was far from being the most influential in the other model. This change was a result of the downsampling technique, as 0.7% of nongraduates were English learners, and 0% of graduates were English learners. From this data, we can see that, because downsampling reduces the size of the graduate class significantly, it weeded out all of the observations of English learner graduates, leading the model to lean on this predictor heavily in drawing its decision boundary. In fact, with the given sample, if the model predicted non-graduation for every English learner, it would have relatively high accuracy.

Another interesting observation was that within all three models, the coefficient for gender had its largest value in this downsampled LDA model (0.724), suggesting that the sampling technique shrank the number of either female nongraduate or male graduate observations or both. Thus, this model could more reliably use gender as a predictor of a student's likelihood to graduate.

A strong predictor in the downsampled LDA model was earning sufficient credits to advance to tenth grade, which had a coefficients of linear discriminants value of 1.064. The nongraduates had a group mean of 0.464, and the graduates had a group mean of 0.906. The group means showed that 90.6% of the graduates earned sufficient credits to advance to the 10th grade, while only 46.4% of the nongraduates did. Other predictors with high coefficients of linear discriminants are qualifying for special education (0.588), race (0.636), gender (0.724), and suspended out of school (-0.762). Those predictors were the most influential in determining a student's likelihood of graduation, as they were the variables with the highest values. Again, the group means of race did not support its high predictability because the nongraduate and graduate means were very close at 0.899 and 0.870, respectively.

Table 16

Downsampled Ninth-Grade Variables Used to Predict Graduation Utilizing LDA (Testing Data)

Independent Variable	Nongraduate Means	Graduate Means	Coefficients of Linear Discriminants: LD1
Receive free or reduced meals (frl)	0.942	0.826	-0.328
Qualify as English learner (ell)	0.007	0.000	-1.892
Qualify for special education services (swd)	0.152	0.116	0.588
Race (race)	0.899	0.870	0.636
Gender (gend)	0.348	0.565	0.724
Attend at least 90% of days (abs9)	0.507	0.804	0.059
Earn sufficient credits to advance to 10th grade (cred9)	0.464	0.906	1.064
Failed less than two core courses (fail9)	0.362	0.710	0.255
If suspended out of school (oss9)	0.551	0.159	-0.762
Number of school moves in 9th grade (trans9)	1.036	0.297	-0.146
Scaled score of ninth literature standardized test (litss9)	400.759	424.510	0.003
Scaled score of algebra standardized test (algss9)	360.845	377.177	0.016
School minority percentage (minority9)	85.722	81.801	-0.013

Note. Prior probabilities of groups: nongraduate: 0.5, graduate: 0.5.

Quadratic Discriminant Analysis

Using quadratic discriminant analysis (QDA), each observation was classified in the group that had the least squared distance. In this study, quadratic discriminant analysis was also used to predict whether or not a student would graduate within four years based on known ninth-grade student data. Quadratic discriminant analysis was used to analyze the research data to understand the relationship between the dependent variable and different independent variables. The output contains the group means, but because the analysis involved a quadratic, rather than a linear function of the predictors, QDA does not provide the coefficient of the linear discriminant for each variable like the LDA did (Crouser, n.d.). The coefficients of linear discriminants are used to determine the likelihood that each student will graduate within four years. For the QDA analysis, the prior probability provided was the chance of being in a particular group before the analysis began, and the posterior probability for each group was provided after all background information was taken into account.

Again, the 2017 cohort was used as training data, and the 2018 cohort was used as testing data. The group means and the accuracy of the upsampled quadratic discriminant analysis prediction are listed in Table 17, and the group means and the accuracy of the downsampled quadratic discriminant analysis prediction are listed in Table 18. The accuracy was only slightly higher for upsampled ninth-grade data compared to the downsampled data, but that is not surprising because of the imbalance of graduates and nongraduates. There were many more graduates than nongraduates. The class imbalance was managed by both downsampling the majority class and upsampling the minority class of the data set before training the model.

In the upsampled QDA model, the accuracy score was 87.4%. The statistical output for the upsampled QDA model can be seen in Table 17. There was still a drastic difference in group means for students earning sufficient credits to advance to the tenth grade, as the nongraduates had a group mean of 0.483, and the graduates had a group mean of 0.939. The group means show that 93.9% of the graduates earned sufficient credits to advance to the 10th grade, while only 48.3% of the nongraduates did. Similarly, only 16.2% of graduates received out of school suspension while 53.4% of nongraduates received out of school suspension, and 88.1% of graduates attended at least 90% of school days while only 49.7% nongraduates attended at least 90% of school days. These results confirm that students who earn sufficient credits to advance to the 10th grade, have fewer suspensions, and attend more school are all more likely to be on track to graduate.

Table 17

Upsampled Ninth-Grade Variables Used to Predict Graduation Utilizing QDA (Testing Data)

Independent Variable	Nongraduate Means	Graduate Means
Receive free or reduced meals (frl)	0.947	0.843
Qualify as English learner (ell)	0.005	0.005
Qualify for special education services (swd)	0.150	0.075
Race (race)	0.886	0.879
Gender (gend)	0.329	0.547
Attend at least 90% of days (abs9)	0.497	0.881
Earn sufficient credits to advance to 10th grade (cred9)	0.483	0.939
Failed less than two core courses (fail9)	0.368	0.771
If suspended out of school (oss9)	0.534	0.162
Number of school moves in 9th grade (trans9)	1.051	0.289
Scaled score of ninth literature standardized test (litss9)	401.553	424.662
Scaled score of algebra standardized test (algss9)	360.921	376.995
School minority percentage (minority9)	85.854	83.913

Note. Prior probabilities of groups: nongraduate: 0.5, graduate: 0.5.

In the downsampled QDA model, the accuracy score was 85.4%, which performed slightly worse than the upsampled models. The statistical output for the downsampled QDA model can be seen in Table 18. The drastic difference between the means of graduates and nongraduates for the variables of earning sufficient credits to advance to tenth grade, attending at least 90% of school days, and if suspended from school continued for this downsampled QDA model just as it did for the upsampled QDA model.

Table 18

Downsampled Ninth-Grade Variables Used to Predict Graduation Utilizing QDA (Testing Data)

Independent Variable	Nongraduate Means	Graduate Means
Receive free or reduced meals (frl)	0.942	0.804
Qualify as English learner (ell)	0.007	0.007
Qualify for special education services (swd)	0.152	0.058
Race (race)	0.899	0.891
Gender (gend)	0.348	0.565
Attend at least 90% of days (abs9)	0.507	0.877
Earn sufficient credits to advance to 10th grade (cred9)	0.464	0.957
Failed less than two core courses (fail9)	0.362	0.790
If suspended out of school (oss9)	0.551	0.138
Number of school moves in 9th grade (trans9)	1.036	0.225
Scaled score of ninth literature standardized test (litss9)	400.759	427.609
Scaled score of algebra standardized test (algss9)	360.845	379.949
School minority percentage (minority9)	85.722	82.174

Note. Prior probabilities of groups: nongraduate: 0.5, graduate: 0.5.

Unlike the LDA model, the QDA model allows for the classes to have different covariance matrices, and thus the discriminant function is a quadratic. In order for the QDA models to outperform the LDA models, it must be the case that the decision boundary between graduates and nongraduates is moderately nonlinear, allowing a quadratic decision boundary to more accurately separate the classes without the addition of bias to the model. Here, we observed that the accuracies of the QDA models with upsampled data and downsampled data were 87.4% and 85.4%, respectively. These were both lower than the accuracies of the best-performing logistic regression and LDA models, suggesting that the decision boundary is better estimated with a linear discriminant function. This difference was also a result of the fact that there were relatively few observations in the training set for the model to train on, so reducing variance was a large concern for the models in question, and model flexibility was less of a concern.

It should be noted that even though race was identified as a significant predictor in the upsampled logistic regression analysis, it does not look to be a strong predictor considering the conflicting data among all of the models. Even though the upsampled logistic regression identified race as a significant predictor of dropouts, the downsampled logistic regression model was identified as the preferred model, and it did not identify race as a significant predictor. Race does not look to be a significant variable as it was first identified.

In addition, for both the upsampled and downsampled LDA and QDA analyses, the difference between the nongraduate means and graduate means is minimal, signifying that race has little effect on the models. Overall, the variable of race did not make a significant difference for ninth-grade data when predicting graduation.

Model Comparisons and Variable Evaluations

Of the three statistical models, the upsampled logistic regression and upsampled QDA models tied for the highest accuracy at 89.0%. The downsampled QDA model had the lowest accuracy at 85.4%. That low performance aligns with the prior perceptions of the QDA model, namely that being a more flexible model requires a larger number of training observations in order to estimate the decision boundary better. The downsampled QDA model had the smallest training set and thus performed the worst. In addition, because the QDA model had lower accuracies than those of the logistic regression and LDA models, it could be that the decision boundary is better estimated with a linear discriminant function.

Of all of the variables included in the ninth-grade data, there were several that were consistently shown to have a strong relationship with predicting high school graduation—specifically, if a student received enough credits to advance to the tenth grade. This variable was a significant predictor for all three models, including both upsampled and downsampled data. The analysis showed that if a student did not earn sufficient credits to advance to 10th grade, then the student was very unlikely to graduate from high school. Other variables identified in most every model as having a strong relationship with predicting high school graduation were: attending school at least 90% of the days, if suspended from school, and gender.

Lower chances of graduating high school were associated with being suspended out of school and making a higher number of school moves while in ninth grade. Both variables were shown to be statistically significant predictors for graduation in the logistic regression model using both upsampled and downsampled data.

Overall, variables consistently identified in a majority of the ninth-grade models as able to predict students who would not complete high school within four years were if a student: (a) did not receive enough credits to advance to the tenth grade, (b) did not attend school at least 90% of the time, (c) was suspended from school, (d) had multiple school moves in the ninth grade, and (f) male gender.

RQ2**:** Which statistical model is most accurate at predicting future dropouts or late graduates utilizing ninth-grade variables?

Three statistical analyses were utilized to predict future dropouts or late graduates utilizing ninth-grade variables. The statistical analyses used were logistic regression, linear discriminant analysis, and quadratic discriminant analysis. To help with the imbalance of graduates and nongraduates, upsampling of the minority class and downsampling of the majority class of the data were performed before training the model. The end result for both upsampling and downsampling is the same number of observations from the minority and majority classes. When each analysis was run, the data set was split into training and testing sets, so each model could train using roughly half the data set and then test on the other half of the data set. For the ninth-grade cohort, the 2017 variable data was used as the training data, and the 2018 variable data was used to test the accuracy of the model.

Receiver Operating Characteristic (ROC) Curve and Confusion Matrix

To help identify the model that is most accurate at predicting future dropouts or late graduates, both the receiver operating characteristic curve, or ROC curve, and confusion matrix were utilized. The ROC curve is a good way to see how a predictive model can distinguish between the true-positives (sensitivity) and true-negative (specificity). The ROC curve does this by plotting sensitivity, the probability of predicting a true-positive will be a positive, against one minus specificity, the probability of predicting a true-negative will be positive. The closer the curve follows the top left-hand border of the ROC space, the more accurate the test. The best performance is one with high amounts of true-positives (sensitivity) and few true-negatives predicted to be positive (one minus specificity). The area under the ROC curve is a measure of the usefulness of a test in general, where a greater area means a more useful test.

The confusion matrix is a layout that aids with the visualization of the performance of an algorithm. It is another performance measurement for machine learning classification. It is a table with two rows and two columns that report the number of false-positives, false-negatives, true-positives, and true-negatives. The confusion matrix shows how the statistical model was confused when it made the predictions. It not only gives insight into the errors that were made by each statistical model but, more importantly, the types of errors that were made.

The areas under the curve for the various statistical models for ninth grade are listed in Table 19. In addition, the statistical analysis results of the confusion matrix, including the accuracy, kappa, sensitivity, specificity, F1, and balanced accuracy, for each statistical model are provided. All six models had relatively high accuracy levels, and kappa values from fair to moderate told how much better each classifier performed over a classifier that only guessed at random. Downsampled logistic regression had the highest kappa value of 0.43. Sensitivity is the true-positive rate, while specificity is the true-negative rate, meaning those predicted to graduate actually did graduate, and those predicted to not graduate actually did not graduate. The true-positive rates clearly outperformed the true-negative rates for all of the models. The F-score or F1 is another measure of a model's accuracy. Each of the models had a relatively high F-score, with 0.91 being the lowest accuracy where one is considered a perfect F-score. The balanced accuracy's best possible value is one, and the worst possible value is zero. Balanced accuracy measures the average of the proportion correct using an equal number of trials in each class. The balanced accuracy performed decently with values ranging from 0.62 to 0.73. When considering all of the statistical analysis results, the models all performed relatively close, but two models equally outperformed the others. Those models were downsampled logistic regression and downsampled quadratic discriminant analysis performances.

Table 19

Analysis Results Based on Ninth Grade Variables for all Data Types and Statistical Models Using Test Data

Statistical Model Used for Ninth Grade Data	AUROC	Accuracy	Kappa	Sensi-tivity	Speci-ficity	F1	Balanced Accuracy
Upsampled LR	0.842	.89	.39	.97	.34	.94	.66
Downsampled LR	0.842	.88	.43	.94	.46	.93	.70
Upsampled LDA	0.841	.89	.39	.97	.34	.94	.66
Downsampled LDA	0.841	.89	.31	.98	.25	.94	.62
Upsampled QDA	0.812	.87	.34	.94	.37	.92	.66
Downsampled QDA	0.812	.85	.41	.90	.57	.91	.73

Note. AUROC is Area Under the ROC Curve.

Figures 6 and 7 show the ROC curves for ninth-grade logistic regression. The two logistic regression models were the models with the highest area under the ROC curve, with both scoring 0.842. The two ROC curves for the logistic regression models with upsampled and downsampled data were of the same shape, showing the false-positive rate hovering around zero until the sensitivity exceeds 0.8. That is, with the logistic regression models, if one wishes to obtain a true-positive rate greater than 0.8, then one must accept a significant increase in the percentage of false-positive classifications. The false-positive rate rises to 0.5 for sensitivity of around 0.9.

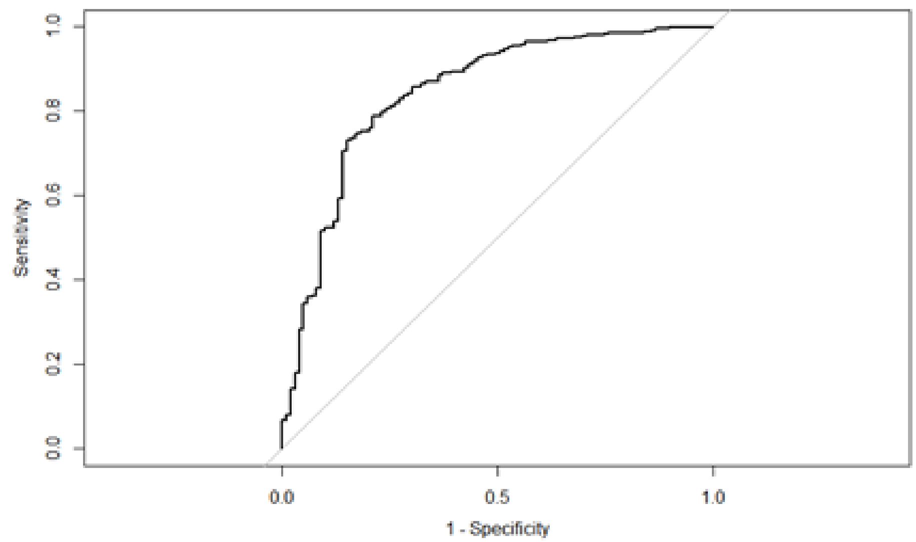

Figure 6. ROC curve results based on ninth-grade variables used to predict graduation utilizing upsampled logistic regression. Area under the curve: 0.842.

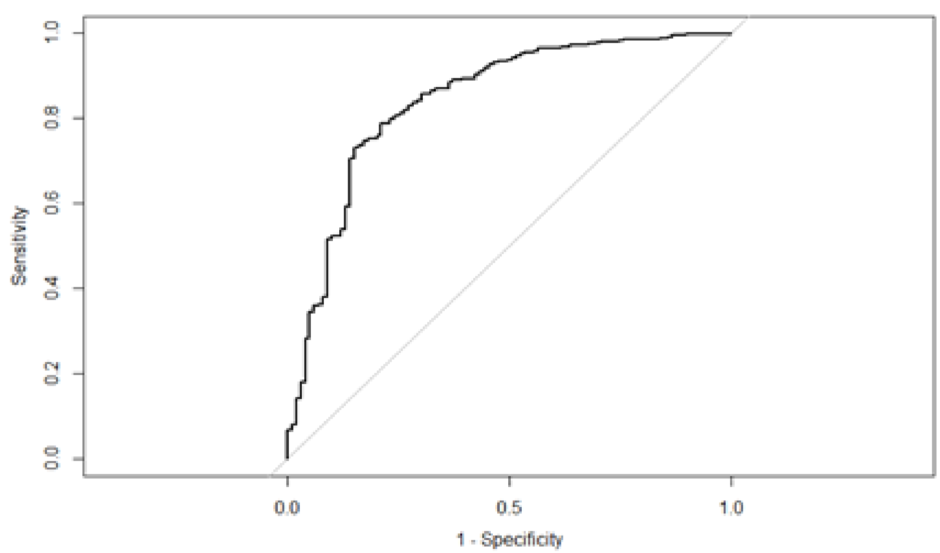

Figure 7. ROC curve results based on ninth-grade variables used to predict graduation utilizing downsampled logistic regression. Area under the curve: 0.842.

Examining the confusion matrices in Tables 20 and 21 gives more insight into the nuances of the tradeoff of correctly predicting a higher number of graduates and incorrectly identifying a higher number of nongraduates. As seen with the logistic regression models' ROC curves, if one wishes to obtain a true-positive rate (sensitivity) greater than 0.8, then one must accept a significant increase in the percentage of false-positive classifications (specificity subtracted from one).

Table 20

Confusion Matrix Results Based on Ninth-grade Variables Used to Predict Graduation Utilizing Upsampled Logistic Regression

	Actual no	Actual yes
Predicted no	34	20
Predicted yes	65	657

The confusion matrix results using the upsampled logistic regression model (see Table 20) had a true-positive rate (sensitivity) of 0.97 and a false-positive rate (specificity subtracted from one) of 0.66. That is, the model was able to identify almost all true graduates correctly. Still, in doing so, it made the mistake of incorrectly predicting that 66% of nongraduates would end up graduating. But, by sampling more nongraduate observations with replacement, the model is able to more accurately predict nongraduates compared to a non-upsampled data set, as it is trained with an equal number of observations of each class. Thus, although the true-positive rate slightly decreased from that of non-upsampled data set, the false-positive rate decreased drastically.

Along with the upsampled logistic regression accuracy value, the kappa identifies how well the model is predicting. The higher the kappa value, the better the model. An accuracy value of 89% and kappa of 0.39 are moderately good results. The F1 value of 0.94 also reflects a good performance because for the F1 score, which is the weighted average score of recall and precision, a value at one is the best performance, and at zero is the worst. The balanced accuracy measures the accuracy using an equal number of trials in each class, and the upsampled logistic regression model had a 66% balanced accuracy, brought down by the low true-negative performance (0.34).

Table 21

Confusion Matrix Results Based on Ninth-grade Variables Used to Predict Graduation Utilizing Downsampled Logistic Regression

	Actual no	Actual yes
Predicted no	46	39
Predicted yes	53	638

The Confusion Matrix results using the downsampled logistic regression model (see Table 21) had a true-positive rate of 0.94 and a false-positive rate of 0.54. This method had the better false-positive rate of the two logistic regression models. This better performance was due to the fact that downsampling reduced the training set size significantly by removing positive observations, preventing the model from learning more about graduates, but allowing it to see a proportionally higher number of nongraduates in its training.

The downsampled logistic regression model had an accuracy value of 88% and kappa of 0.43, which are moderately good results and a bit better than the upsampled performance. Remember, the higher the kappa value, the better the model. The F1 value of 0.93 also reflects a good performance because it is close to 1, which is the best performance. The balanced accuracy measures the accuracy using an equal number of trials in each class, and the downsampled logistic regression model had a 70% balanced accuracy. This balanced accuracy value was also brought down by its low true-negative performance of 0.46, but higher than the upsampled logistic regression's true-negative performance. The higher true-negative performance resulted in a higher balanced accuracy when compared to the upsampled logistic regression's value.

Logistic regression and LDA upsampled and downsampled data had very similar conclusions and high performance, with only a 0.001 difference for each of their areas under the ROC curve. Their model accuracies were virtually identical, and they identified most all the same statistically significant variables when predicting future dropouts or late graduates.

As seen in Figures 8 and 9, the ROC curves for LDA are nearly identical to those of the logistic regression models. That is, with the LDA models, if one wishes to obtain a true-positive rate (where graduation is a positive classification) greater than 0.8, then one must accept a significant increase in the percentage of false-positive classifications (false-positive rates rise to 0.5 for sensitivity of around 0.9). The accuracy of the LDA is just slightly below the accuracy of the logistic regression model.

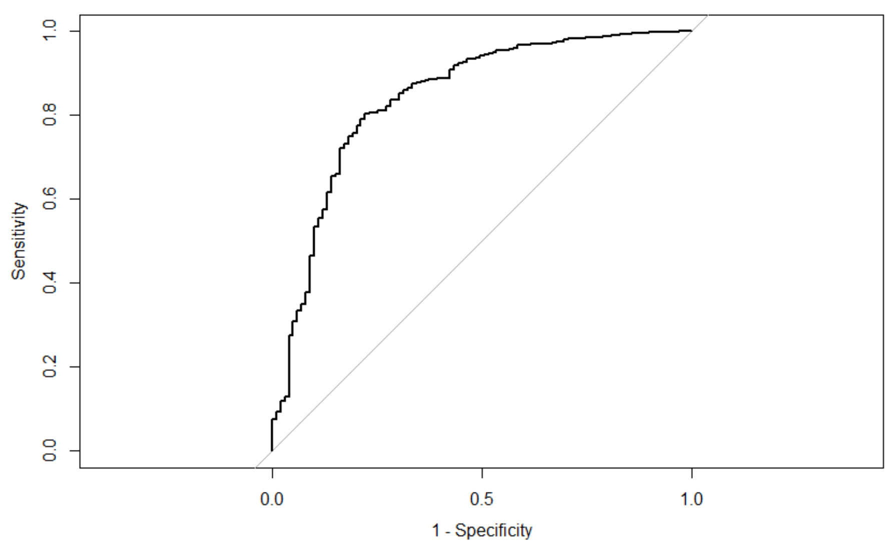

Figure 8. ROC curve results based on ninth-grade variables used to predict graduation utilizing upsampled linear discriminant analysis. Area under the curve: 0.841.

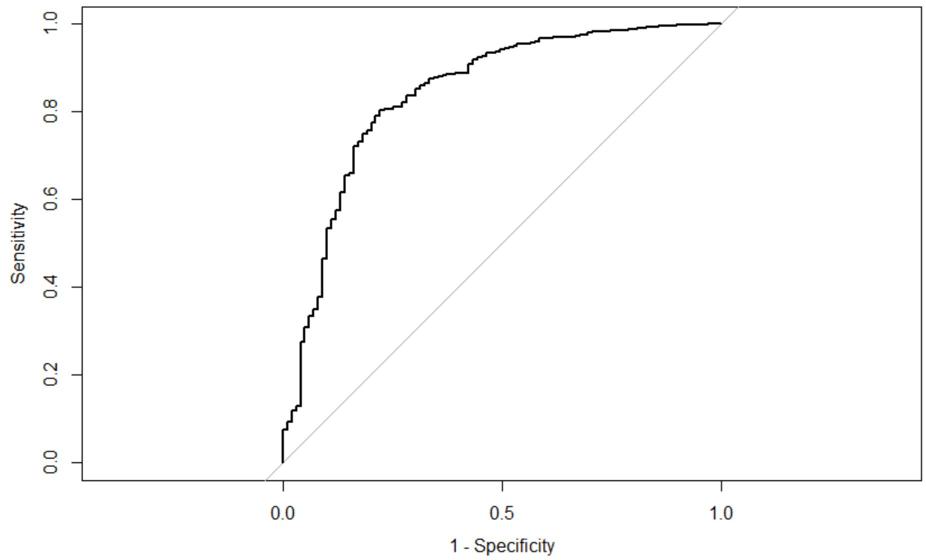

Figure 9. ROC curve results based on ninth-grade variables used to predict graduation utilizing downsampled linear discriminanat analysis. Area under the curve: 0.841.

The confusion matrix results using the upsampled LDA model (see Table 22) had an accuracy of 0.89, a true-positive rate of 0.97, and a false-positive rate of 0.66. This performance was identical to the upsampled logistic regression model.

Table 22

Confusion Matrix Results Based on Ninth-grade Variables Used to Predict Graduation Utilizing Upsampled Linear Discriminant Analysis

	Actual no	Actual yes
Predicted no	34	20
Predicted yes	65	657

The upsampled LDA model had a kappa level of 0.39, which is considered good and was identical to the upsampled logistic regression kappa value. The F1 value of 0.94 also reflects a good performance and the same as the upsampled logistic regression performance. The balanced accuracy measures the accuracy using an equal number of trials in each class, and the upsampled LDA model had a 66% balanced accuracy. This balanced accuracy value was identical to the upsampled logistic regression's value and was also brought down by its low true-negative performance of 0.34.

The confusion matrix results using the downsampled LDA model (see Table 23) had the highest true-positive rate at 0.98, but also the highest false-positive rate at 0.75. This change showed that the model classified almost all observations as graduates and that it performed just marginally better than a model that classified all observations as graduates. The downsampling technique shrunk the training set to a size that was evidently too small for the LDA to fit an adequate decision boundary.

Table 23

Confusion Matrix Results Based on Ninth-grade Variables Used to Predict Graduation Utilizing Downsampled Linear Discriminant Analysis

	Actual no	Actual yes
Predicted no	25	14
Predicted yes	74	663

The downsampled LDA model had an accuracy value of 89% and kappa of 0.31, which are good results but just not quite as good as the upsampled LDA's performance. The F1 value of 0.94 also reflects a good performance because it is close to one, which is the best performance. The balanced accuracy measures the accuracy using an equal number of trials in each class, and the downsampled LDA model had a 62% balanced accuracy. This balanced accuracy value was lower than the upsampled LDA's value and also brought down by its low true-negative performance of 0.25, which was lower than the upsampled LDA's true-negative performance.

The ROC curves for the QDA models shown in Figures 10 and 11 are similar in shape to those of the logistic regression and LDA models, but they are further from the upper left corner. They thus have less area under the ROC curve, signifying less accuracy. The QDA model is the least accurate of the three models.

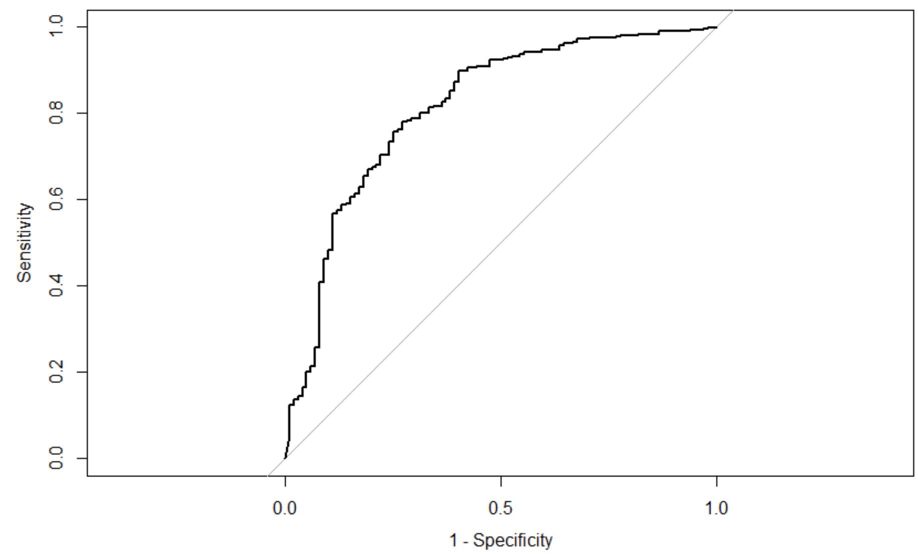

Figure 10. ROC curve results based on ninth-grade variables used to predict graduation utilizing upsampled quadratic discriminant analysis. Area under the curve: 0.812.

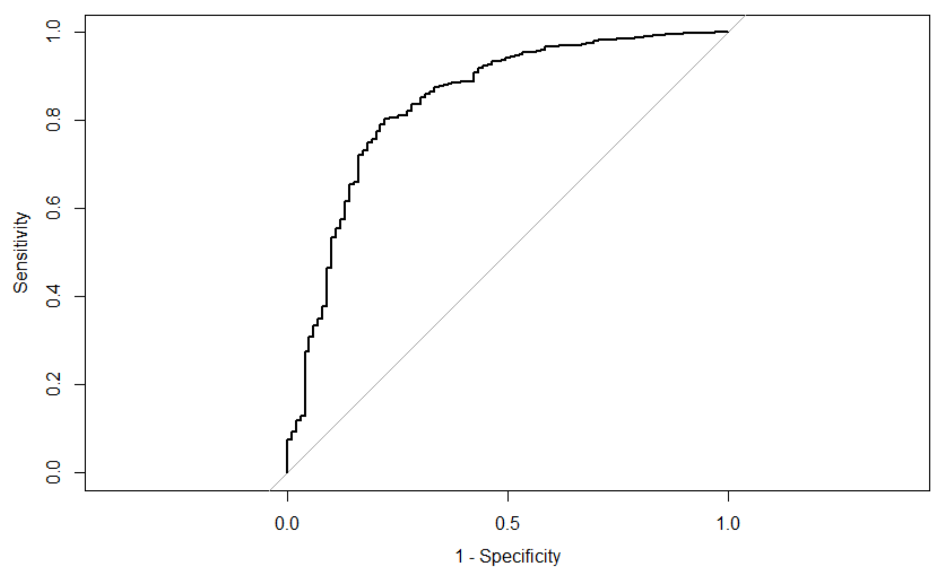

Figure 11. ROC curve results based on ninth-grade variables used to predict graduation utilizing downsampled quadratic discriminanat analysis. Area under the curve: 0.812.

The confusion matrix results using the upsampled QDA model (see Table 24) had a true-positive rate of 0.94 and a false-positive rate of 0.63, which was quite similar to the regular QDA model but with a slight decrease in both the true-positive rate and the false-positive rate. This performance suggests that upsampling caused the QDA model to more accurately predict nongraduates and less accurately predict graduates, suggesting that the quadratic nature of the decision boundary that it determined allowed for more students to be categorized as nongraduates.

Table 24

Confusion Matrix Results Based on Ninth-grade Variables Used to Predict Graduation Utilizing Upsampled Quadratic Discriminant Analysis

	Actual no	Actual yes
Predicted no	37	42
Predicted yes	62	635

The upsampled QDA model had an accuracy value of 87% and kappa of 0.34, which are also good results but not quite as good as either the upsampled logistic

Table 25

Confusion Matrix Results Based on Ninth Grade Variables Used to Predict Graduation Utilizing Downsampled Quadratic Discriminant Analysis

	Actual no	Actual yes
Predicted no	56	70
Predicted yes	43	607

regression or upsampled LDA values. The F1 value of 0.92 also reflects a good performance but lower than the upsampled logistic regression or upsampled LDA F1 values. The balanced accuracy measures the accuracy using an equal number of trials in each class, and the upsampled LDA model had a 66% balanced accuracy. This balanced accuracy value was identical to the upsampled logistic regression and upsampled LDA's balanced accuracy values.

The confusion matrix results using the downsampled QDA model (see Table 25) had a true-positive rate of 0.90 and a false-positive rate value of 0.43, which was the lowest false-positive rate of all the models, but also the lowest true-positive rate value. Downsampling leads the model to have the best success in decreasing the number of nongraduates labeled as graduates. Still, it also suffered as it labeled more graduates as nongraduates. This change was likely due to the equality of class sizes with unique observations in each class.

The downsampled QDA model had an accuracy value of 85% and kappa of 0.41, which are fair results when compared to the performance of the upsampled logistic regression and LDA models. The F1 value of 0.91 reflects a decent performance because it is close to 1. The balanced accuracy measures the accuracy using an equal number of trials in each class, and the downsampled QDA model had a 73% balanced accuracy. This balanced accuracy value was higher than the balanced accuracy of all of the other upsampled and downsampled models. The boost in overall performance occurred because of its high true-negative performance of 0.57, which was also higher than any other model's true-negative performance.

The accuracies of the QDA models were lower than the accuracies of the logistic regression and LDA models. Unlike the LDA models, the QDA models' discriminant function is a quadratic. For the QDA models to outperform the LDA models, the decision boundary between graduates and nongraduates would need to be moderately nonlinear, allowing a quadratic decision boundary to more accurately separate the classes without the addition of bias to the model. The lower accuracies of the QDA compared to those of the logistic regression and LDA models indicated that the decision boundary is probably better estimated with a linear discriminant function.

Choosing the Best Model

All six models had moderate accuracy levels, and kappa values from poor to fair told how a classifier performed over a classifier that only guessed at random. Upsampled quadratic discriminant analysis had the highest kappa value of 0.26. Sensitivity is the true-positive rate, while specificity is the true-negative rate, meaning those predicted to graduate actually did graduate, and those predicted to not graduate actually did not graduate. The true-positive rates outperformed the true-negative rates for all of the models. The F-score or F1 is another measure of a model's accuracy. Each of the models had a relatively high F-score, with 0.78 being the lowest accuracy where one is considered a perfect F-score. The balanced accuracy's best possible value is one, and the worst possible value is zero. Balanced accuracy measures the average of the proportion correct using an equal number of trials in each class. The balanced accuracy performed decently with values ranging from 0.64 to 0.66. When considering all of the statistical analysis results, the models all performed relatively close, but two models did equally outperform the others. Those models were upsampled quadratic discriminant analysis and downsampled quadratic discriminant analysis performances.

Within all of these models, performance metrics is a tradeoff between the true-positive rate and the false-positive rate. This tradeoff exists in almost all classification problems. Some models are excellent at predicting positive observations (high true-positive rate), but that is almost always associated with a low detection rate (high false-positive rate). In this study, because there was a significant class imbalance in the training data, the models all had difficulty in detecting nongraduates, as there were relatively fewer observations of nongraduates. Even when using upsampling and downsampling, the false-positive may decrease, but it still remained quite high in most of the models. One must choose a model based on the tradeoffs presented and the problem at hand.

Because a main goal of the task is to identify nongraduates early in their schooling, it is wiser to choose a model with the lowest false-positive rate, even if the true-positive rate suffers slightly from that choice. It is a better option to offer help to more students than need it than to fail to identify students who need help. This judgment leads to the conclusion that the downsampled QDA model is the best model for minimizing false-positive observations, as its false-positive rate was the lowest at 0.43, and its true-negative was the highest at 0.56. The downsampled QDA model also had the second-highest kappa value and the highest balanced accuracy. However, in terms of accuracy and true-positive rate, the downsampled LDA model is best, with an accuracy level of 0.89 and a true-positive rate of 0.98. However, the accuracy level and true-positive rate come at a price because the downsampled LDA also had the lowest kappa value, lowest balanced accuracy, and highest false-positive rate.

RQ3: Does one or more of the sixth-grade variables consisting of failing English, failing math, attending less than 80% of the time, receiving out-of-school suspension, number of school moves, standardized reading and math scores, school minority and poverty percentages, ELL status, SWD status, free/reduced meal status, race, and gender accurately predict students who will not complete high school within four years?

Three statistical analyses were utilized to answer this question. The statistical analyses used to provide confirmation that the correct variables were identified were logistic regression, linear discriminant analysis, and quadratic discriminant analysis. To help with the imbalance of graduates and nongraduates, upsampling of the minority class and downsampling of the majority class of the data were performed before training the model. The end result for both upsampling and downsampling is the same number of observations from the minority and majority classes. Like with the ninth-grade data, the sixth-grade data set was split into training and testing sets. The data was split so each model could train using roughly half the data and then test on the other half of the data. For the sixth-grade cohort, the 2017 variable data was used as the training data, and the 2018 variable data was used to test the accuracy of the model.

Logistic Regression

The logistic regression analysis was used to explain the relationship between the students' graduation status and the various independent variables. This statistical model can estimate the probability of certain events occurring based on previous data. In this study, logistic regression was used to predict whether or not a student would graduate high school within four years based on sixth-grade student data. The results of the upsampled logistic regression prediction for the sixth-grade cohort are listed in Table 26, and the downsampled logistic regression prediction results are listed in Table 27.

As seen in Table 26, the upsampled logistic regression model had an accuracy level of 67.9%. Another measure of the quality of the model is the Akaike information criterion (AIC) value. If two similar models are being compared, then the model with the lower AIC is superior. The upsampled logistic regression model had an AIC level of

1944.0 and can be compared to the downsampled logistic regression model's AIC level to identify which model is preferred.

The Hosmer and Lemeshow goodness of fit test results (chi-square 48.34, df = 11, $p < 0.001$) indicate that the observed nongraduates differ significantly from the expected value of nongraduates using the predictor variables based on their proportion of the population. The chi-square value is above the critical value with a degrees of freedom value of 11 at the $p < 0.05$ level.

A higher Nagelkerke pseudo R-squared indicates which model does a better job of predicting the outcome (UCLA: Statistical Consulting Group., n.d.). The upsampled logistic regression model had an pseudo R-squared value of 0.312 and can be compared to the downsampled logistic regression model's pseudo R-squared level to identify which model is preferred.

Of the 13 predictor variables, eight were statistically significant: receiving free or reduced meals, race, gender, passing English in the sixth grade, being suspended from school, moving multiple times in the sixth grade, and scaled scores on the reading and math standardized tests. By passing English in the sixth grade, a student increases his or her odds of graduating by a factor of 4.7, given all other variables are unchanged. If a student received free or reduced lunch, the odds of a student graduating decreased by 70.2% (0.298 – 1 = -0.702), keeping other variables constant.

Interestingly, the estimate value for passing English was positive (z = 2.778, p = 0.005, odds ratio = 4.074, 95% CI = 0.678 to 2.242), while passing math (z = 0.103, p = 0.918, odds ratio = 1.220, 95% CI = 2.053 to 9.042) was not significant. This difference suggests that passing English in the sixth grade is much more indicative of a student graduating high school than passing math.

To identify the variables that are stronger in their role as a predictor, z-statistic values that are farther away from zero identify the strongest predictors. The variables identified as passing English ($z = 2.778$, $p = 0.005$, odds ratio = 4.074, 95% CI = 2.053 to 9.042) and scaled score on the reading standardized test ($z = 2.000$, $p = 0.046$, odds ratio = 1.010, 95% CI = 1.003 to 1.018) were associated with a slight increase in graduating high school.

Even higher graduation percentages were associated with the variables of race ($z = 5.205$, $p < .001$, odds ratio = 3.181, 95% CI = 2.042 to 5.004), gender ($z = 6.432$, $p < .001$, odds ratio = 1.886, 95% CI = 1.508 to 2.362), and higher scaled score on the math standardized test ($z = 5.448$, $p < .001$, odds ratio = 1.015, 95% CI = 1.008 to 1.023). In addition, lower chances of graduating high school were associated with receiving free or reduced lunch ($z = -4.493$, $p < .001$, odds ratio = 0.298, 95% CI = 0.186 to 0.468), suspended out of school ($z = -4.546$, $p < .001$, odds ratio = 0.504, 95% CI = 0.382 to 0.661) and a higher number of school moves made by the student ($z = -7.570$, $p < .001$, odds ratio = 0.558, 95% CI = 0.475 to 0.651).

Table 26

Upsampled Sixth-Grade Variables Used to Predict Graduation Utilizing Logistic Regression (Testing Data)

Independent Variable	Regression Coefficient	*SE*	*z*	Pr(>|z|)		*OR*
(Intercept)	-23.534	2.961	-7.947	p<.001	***	-
Receive free or reduced meals (frl)	-0.989	0.220	-4.493	p<.001	***	0.298
Qualify as English learner (ell)	-0.170	0.743	-0.229	0.819		0.660
Qualify for special education services (swd)	-0.082	0.195	-0.420	0.675		0.928
Race (race)	1.175	0.226	5.205	p<.001	***	3.181
Gender (gend)	0.732	0.114	6.432	p<.001	***	1.886
Passed English in 6th grade (ela6)	1.083	0.390	2.778	0.005	**	4.074
Passed math in 6th grade (math6)	0.032	0.308	0.103	0.918		1.220
Attend at least 80% of days (abs6)	0.884	0.651	1.357	0.175		2.556
If suspended out of school (oss6)	-0.635	0.140	-4.546	p<.001	***	0.504
Number of school moves in 6th grade (trans6)	-0.605	0.080	-7.570	p<.001	***	0.558
Scaled score of 6th reading standardized test (readcrct6)	0.007	0.004	2.000	0.046	*	1.010
Scaled score of 6th math standardized test (mathcrct6)	0.020	0.004	5.448	p<.001	***	1.015
School minority percentage (minority6)	-0.006	0.004	-1.486	0.137		0.998

Note. AIC: 1944. P < 0.001 '***', p < 0.01 '**', p < 0.05 '*'.

As shown in Table 27, using the downsampled logistic regression model to decrease the size of the training set increased the model accuracy to 72.8%, and the model detected only a couple of significant coefficients. The downsampled logistic regression model had an AIC level of 346.4, and when compared to the upsampled logistic regression model's AIC level of 1944.0, the downsampled model is identified as the preferred model.

The Hosmer and Lemeshow goodness of fit test results (chi-square 21.25, df = 11, $p = 0.031$) indicate that the observed nongraduates differ significantly from the expected value of nongraduates using the predictor variables based on their proportion of the population. The chi-square value is above the critical value with a degrees of freedom value of 11 at the $p < 0.05$ level.

Remember, a higher Nagelkerke pseudo R-squared indicates which model does a better job of predicting the outcome (UCLA: Statistical Consulting Group., n.d.). The downsampled logistic regression model had a pseudo R-squared value of 0.366. When compared to the upsampled logistic regression model's pseudo R-squared level of 0.312, the downsampled model is again identified as the preferred model.

Of the 13 predictor variables, only two were statistically significant: race and the number of school moves a student makes in the sixth grade. Race increases a student's odds of graduating by a factor of 3.0, given that all other variables are unchanged. If a student had multiple moves in the sixth grade, the odds of a student graduating decreased by 48.9% (0.511 – 1 = -0.489), keeping all other variables constant.

The z-statistic was again used to identify the variables that were stronger in their role as a predictor. Z-statistic values that are farther away from zero identify the strongest predictors. The variable of race (z = 2.117, $p = 0.034$, odds ratio = 3.012, 95% CI = 1.091 to 8.645) was associated with an increased chance of graduating high school. Having a lower chance of graduating high school was associated with multiple school moves a student made in sixth grade (z = -2.055, *p* = .040, odds ratio = 0.511, 95% CI = 0.324 to 0.764).

Table 27

Downsampled Sixth-Grade Variables Used to Predict Graduation Utilizing Logistic Regression (Testing Data)

Independent Variable	Regression Coefficient	*SE*	*z*	Pr(>|z|)	*OR*
(Intercept)	-57.510	1,477	-0.039	0.969	-
Receive free or reduced meals (frl)	-.048	0.574	-0.083	0.934	0.385
Qualify as English learner (ell)	-17.610	3,956	-0.004	0.996	0.000
Qualify for special education services (swd)	0.253	0.454	0.557	0.578	0.959
Race (race)	1.157	0.547	2.117	0.034 *	3.012
Gender (gend)	0.470	0.287	1.636	0.102	2.305
Passed English in 6th grade (ela6)	16.740	953.5	0.018	0.986	1.616
Passed math in 6th grade (math6)	-0.055	0.712	-0.077	0.938	0.503
Attend at least 80% of days (abs6)	15.690	1,127	0.014	0.989	0.981
If suspended out of school (oss6)	0.478	0.331	-1.445	0.148	0.521
Number of school moves in 6th grade (trans6)	-0.379	0.184	-2.055	0.040 *	0.511
Scaled score of 6th reading standardized test (readcrct6)	0.015	0.010	1.502	0.133	1.012
Scaled score of 6th math standardized test (mathcrct6)	0.016	0.010	1.658	0.097 .	1.022
School minority percentage (minority6)	-0.007	0.010	-0.637	0.524	0.995

Note. AIC: 346.4. p < 0.001 '***', p < 0.01 '**', p < 0.05 '*'.

Linear Discriminant Analysis

Like logistic regression, linear discriminant analysis (LDA) explains a categorical variable by the values of independent variables. Linear discriminant analysis makes predictions by estimating the probability that new inputs belong to a particular class. In this study, graduating or not graduating within four years was the prediction being made. It was the class with the highest probability that was the output class, and a prediction was made.

For the sixth-grade data, the 2017 cohort was again used as training data, and the 2018 cohort was used as testing data. The results of the upsampled linear discriminant analysis prediction for the sixth-grade cohort are listed in Table 28, and the results for the downsampled linear discriminant analysis prediction are listed in Table 29.

The coefficients of linear discriminants were the values used to predict whether or not each student would graduate within four years. The coefficients were similar to regression coefficients. Whichever category had the highest probability was selected. The higher the coefficient, the more weight it had.

The upsampled LDA model shown in Table 28 had an accuracy of 69.3% and had some results that closely aligned with those of the upsampled logistic regression models. The coefficients of linear discriminants for race (0.794), gender (0.768), and passing English in the sixth grade (0.733) show that those predictors were the most influential in determining a student's likelihood of graduation, as they were the ones with the highest magnitude. Having a lower chance of graduating from high school was associated with the student receiving free or reduced lunch (-0.772) and being suspended out of school (-0.847).

A student being suspended out of school was the most influential predictor with a coefficient of -0.847. The high suspension coefficient suggests that students who are suspended out of school are less likely to graduate. The group means confirmed this, as only 13.7% of graduates were suspended out of school in the sixth grade, while 40.1% of nongraduates were suspended. Race had the second-highest weighted coefficient at 0.794, but the group means did not confirm it because the nongraduate and graduate means were very close at 0.919 and 0.879, respectively.

What is surprising is that the coefficients for attending at least 80% of school days had low coefficient of 0.040, suggesting that sixth-grade attendance had little effect on graduation likelihood in this model, unlike when a student is in ninth grade.

Table 28

Upsampled Sixth-Grade Variables Used to Predict Graduation Utilizing LDA (Testing Data)

Independent Variable	Nongraduate Means	Graduate Means	Coefficients of Linear Discriminants: LD1
Receive free or reduced meals (frl)	0.949	0.843	-0.772
Qualify as English learner (ell)	0.005	0.005	-0.200
Qualify for special education services (swd)	0.154	0.075	-0.061
Race (race)	0.919	0.879	0.794
Gender (gend)	0.353	0.547	0.768
Passed English in 6^{th} grade (ela6)	0.908	0.989	0.733
Passed math in 6^{th} grade (math6)	0.923	0.974	0.120
Attend at least 80% of days (abs6)	0.931	0.996	-0.040
If suspended out of school (oss6)	0.401	0.137	-0.847
Number of school moves in 6^{th} grade (trans6)	0.798	0.205	-0.411
Scaled score of 6^{th} reading standardized test (readcrct6)	813.143	824.536	0.010
Scaled score of 6^{th} math standardized test (mathcrct6)	796.608	809.403	0.016
School minority percentage (minority6)	87.699	83.895	-0.004

Note. Prior probabilities of groups: nongraduate: 0.5, graduate: 0.5.

Table 29 shows that the downsampled LDA model had an accuracy level of 0.670, performing worse than the upsampled model. The poor performance was likely a result of the significant decrease in the number of graduate observations in the training set, thus preventing the model from learning more about the key characteristics of graduates. In this model, the most important features were passing English (1.416), race (1.22), and gender (1.02). It is important to note, though, that the downsampling technique used here eliminates all graduates who failed English, causing the passing of English coefficient to be the highest. Also interesting to note is that the group means for race are identical, and yet the coefficient for race is the second-largest in magnitude. This indicates that there might be a high amount of randomness contributing to the determination of the coefficients in this model, as there is no real reason for this coefficient to be high if it is not a determining factor in a student's graduation likelihood (which it clearly is not in this sample). The low accuracy and small sample size make the estimates presented in this table questionable.

Table 29

Downsampled Sixth-Grade Variables Used to Predict Graduation Utilizing LDA (Testing Data)

Independent Variable	Nongraduate Means	Graduate Means	Coefficients of Linear Discriminants: LD1
Receive free or reduced meals (frl)	0.942	0.906	-0.032
Qualify as English learner (ell)	0.007	0.014	0.314
Qualify for special education services (swd)	0.152	0.087	0.270
Race (race)	0.899	0.899	1.222
Gender (gend)	0.348	0.580	1.017
Passed English in 6th grade (ela6)	0.899	1.000	1.416
Passed math in 6th grade (math6)	0.913	0.978	0.074
Attend at least 80% of days (abs6)	0.935	1.000	0.422
If suspended out of school (oss6)	0.355	0.159	-0.403
Number of school moves in 6th grade (trans6)	0.754	0.261	-0.320
Scaled score of 6th reading standardized test (readcrct6)	813.855	824.982	0.011
Scaled score of 6th math standardized test (mathcrct6)	796.843	808.679	0.020
School minority percentage (minority6)	87.364	82.101	-0.017

Note. Prior probabilities of groups: nongraduate: 0.5, graduate: 0.5.

Quadratic Discriminant Analysis

Using quadratic discriminant analysis (QDA), each observation was classified in the group that had the least squared distance. In this study, quadratic discriminant analysis was also used to predict whether or not a student would graduate within four years based on known student data. Quadratic discriminant analysis was used to analyze the research data to understand the relationship between the dependent variable and different independent variables. QDA does not provide the coefficient of the linear discriminant for each variable like the LDA did (Crouser, n.d.). The coefficients of linear discriminants were used to determine the probability each student would graduate within four years. For the QDA analysis, the prior probability provided was the chance of being in a particular group before the analysis began, and the posterior probability for each group was provided after all background information was taken into account. Again, the 2017 cohort was used as training data, and the 2018 cohort was used as testing data. The group means of the upsampled quadratic discriminant analysis prediction for the sixth grade are listed in Table 30, and the group means of the downsampled quadratic discriminant analysis prediction are listed in Table 31. For the upsampled QDA prediction, the accuracy was 0.830, and the downsampled version had an accuracy of 0.746.

Table 30 shows that with an accuracy of 0.830, the upsampled model performed surprisingly well, particularly given that the QDA model fit so poorly to the ninth-grade data. There was a big difference in group means for being suspended out of school, as the nongraduates had a group mean of 0.357, and the graduates had a group mean of 0.137. The group means showed that 13.7% of the graduates had been suspended in the sixth grade, while 35.7% of the nongraduates did. These results mean that students who are suspended in the sixth grade are less likely to graduate or graduate on time than sixth-grade students who are not suspended in the sixth grade.

Table 30

Upsampled Sixth-Grade Variables Used to Predict Graduation Utilizing QDA (Testing Data)

Independent Variable	Nongraduate Means	Graduate Means
Receive free or reduced meals (frl)	0.945	0.843
Qualify as English learner (ell)	0.004	0.005
Qualify for special education services (swd)	0.132	0.075
Race (race)	0.901	0.879
Gender (gend)	0.328	0.547
Passed English in 6th grade (ela6)	0.913	0.989
Passed math in 6th grade (math6)	0.915	0.974
Attend at least 80% of days (abs6)	0.932	0.996
If suspended out of school (oss6)	0.357	0.137
Number of school moves in 6th grade (trans6)	0.716	0.205
Scaled score of 6th reading standardized test (readcrct6)	814.343	824.536
Scaled score of 6th math standardized test (mathcrct6)	797.965	809.403
School minority percentage (minority6)	88.028	83.895

Note. Prior probabilities of groups: nongraduate: 0.5, graduate: 0.5.

In the downsampled QDA model shown in Table 31, the accuracy score was 74.6%, which performed worse than the upsampled model. The drastic difference between graduates and nongraduates averages for the attendance and behavior variables indicates that sixth-grade students who miss more than 20% of school days and are suspended are less likely to graduate or graduate on time than students who miss 20% or fewer school days and are not suspended in the sixth grade.

Table 31

Downsampled Sixth-Grade Variables Used to Predict Graduation Utilizing QDA (Testing Data)

Independent Variable	Nongraduate Means	Graduate Means
Receive free or reduced meals (frl)	0.942	0.804
Qualify as English learner (ell)	0.007	0.007
Qualify for special education services (swd)	0.152	0.058
Race (race)	0.899	0.891
Gender (gend)	0.348	0.565
Passed English in 6th grade (ela6)	0.913	0.978
Passed math in 6th grade (math6)	0.935	0.993
Attend at least 80% of days (abs6)	0.355	0.152
If suspended out of school (oss6)	0.754	0.145
Number of school moves in 6th grade (trans6)	813.855	827.787
Scaled score of 6th reading standardized test (readcrct6)	796.843	814.287
Scaled score of 6th math standardized test (mathcrct6)	87.364	83.415
School minority percentage (minority6)	0.942	0.804

Note. Prior probabilities of groups: nongraduate: 0.5, graduate: 0.5.

The QDA models outperformed the upsampled and downsampled LDA models, suggesting that the decision boundary for that data might be better fit by a quadratic curve rather than a linear one. That performance was a surprising result, as the QDA model in the ninth-grade data was far inferior to the LDA model. This performance may suggest that the predictors observed in the sixth-grade students have a more nonlinear relationship with graduation likelihood than ninth-grade predictors do.

It should be noted that even though race was identified as a significant predictor in the upsampled and downsampled logistic regression analyses, it does not look to be a strong predictor considering the conflicting data among all of the models. However, even though the upsampled logistic regression identified race as a significant predictor of dropouts, the downsampled logistic regression model was identified as the preferred model, and race is a significant predictor at a much lower level. Race does not look to be as significant of a variable as it was first identified.

In addition, for both the upsampled and downsampled LDA and QDA analyses, the difference between the nongraduate means and graduate means is minimal, signifying that race has little effect on the models. Overall, the variable of race did not make a significant difference for sixth-grade data when predicting graduation.

Model Comparisons and Variable Evaluations

Of the three statistical models, the upsampled and downsampled QDA models had the highest accuracy at 83.0% and 74.6%, respectively. The downsampled LDA model had the lowest accuracy at 67.0%. That high performance contradicts prior perceptions of the QDA model—namely, that being a more flexible model, it requires a larger number of training observations in order to estimate the decision boundary better. The downsampled QDA model had the smallest training set and still performed well.

Unlike ninth grade, sixth grade did not have many variables that consistently had a strong relationship with predicting high school students who would not graduate on time. The variable of gender was a significant predictor in a few of the models, particularly logistic regression. Additionally, whether or not a student passed English was a strong predictor in the upsampled logistic regression model and upsampled and downsampled LDA models. The analysis showed that, if a student did not pass English, then the student was unlikely to graduate from high school. Two other sixth-grade variables identified in a few models as having a strong relationship with predicting high school graduation were: being suspended from school and making multiple school moves. Those two variables were identified in the upsampled data sets for both the logistic regression and QDA models. In addition, being suspended from school was identified as a strong predictor in the upsampled LDA and downsampled QDA models. Making multiple school moves in the sixth grade was also identified as a strong predictor in the downsampled logistic regression model.

Overall, variables consistently identified in most of the sixth-grade models as able to predict students who would not complete high school within four years were: (a) male gender, (b) if the student was suspended from school, (c) if the student had multiple school moves, and (d) if the student did not pass English.

RQ4 Which statistical model is most accurate at predicting future dropouts or late graduates utilizing sixth-grade variables?

Three statistical analyses were utilized to predict future dropouts or late graduates, using sixth-grade variables. The statistical analyses used were logistic regression, linear discriminant analysis, and quadratic discriminant analysis. To help with the imbalance of graduates and nongraduates, upsampling of the minority class and downsampling of the majority class of the data were performed before training the model. The end result for both upsampling and downsampling is the same number of observations from the minority and majority classes. When each analysis was run, the data set was split into training and testing sets, so each model could train using roughly half the data set and then test on the other half of the data set. For the sixth-grade cohort, the 2017 variable data were used as the training data, and the 2018 variable data were used to test the accuracy of the model.

Receiver Operating Characteristic (ROC) Curve and Confusion Matrix

To help identify the model that is most accurate at predicting future dropouts or late graduates, both the receiver operating characteristic curve, or ROC curve, and confusion matrix were utilized. The ROC curve is a good way to see how a predictive model can distinguish between the true-positive (sensitivity) and true-negative (specificity). The ROC curve does this by plotting sensitivity—the probability of predicting a true-positive will be a positive—against one minus specificity—the probability of predicting a true-negative will be positive. The closer the curve follows the top left border of the ROC space, the more accurate the test. The best performance is one with high amounts of true-positives (sensitivity) and few true-negatives predicted to be positive (one minus specificity). The area under the ROC curve is a measure of the usefulness of a test in general, where a greater area means a more useful test.

The confusion matrix is a layout that aids with the visualization of the performance of an algorithm. It is another performance measurement for machine learning classification. It is a table with two rows and two columns that report the number of false-positives, false-negatives, true-positives, and true-negatives. The confusion matrix shows how the statistical model was confused when it made the predictions. It not only gives insight into the errors that were made by each statistical model but, more importantly, the types of errors that were made.

The areas under the curve for the various statistical models for sixth grade are listed in Table 32. In addition, the statistical analysis results of the confusion matrix, including accuracy, kappa, sensitivity, specificity, F1, and balanced accuracy, for each statistical model, are included.

Table 32

Analysis Results Based on Sixth Grade Variables for all Data Types and Statistical Models Using Test Data

Statistical Model Used for Sixth Grade Data	AUROC	Accuracy	Kappa	Sensi-tivity	Speci-ficity	F1	Balanced Accuracy
Upsampled LR	0.752	.68	.17	.71	.60	.79	.65
Downsampled LR	0.736	.68	.17	.68	.64	.79	.65
Upsampled LDA	0.754	.69	.19	.69	.62	.80	.66
Downsampled LDA	0.736	.67	.18	.67	.66	.78	.66
Upsampled QDA	0.700	.83	.26	.88	.39	.90	.64
Downsampled QDA	0.703	.75	.20	.78	.51	.84	.64

Note. AUROC is Area Under the ROC Curve.

The upsampled and downsampled logistic regression models yield area under the ROC curve results of 0.752 and 0.736, respectively. The average of these values is the highest of the three modeling types, but it is considerably lower than the highest achieved by the logistic models of the ninth-grade students. Accordingly, the ROC curves for the logistic regression models were quite far from the upper left corner, as seen in Figures 12 and 13.

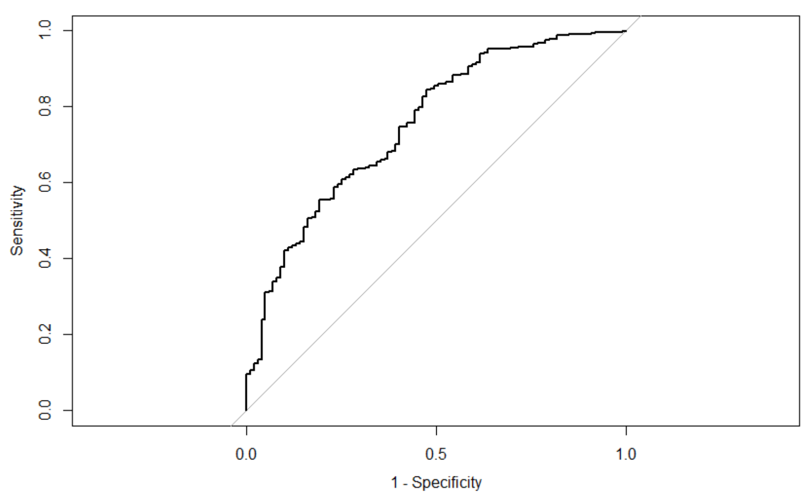

Figure 12. ROC curve results based on sixth-grade variables used to predict graduation utilizing upsampled logistic regression. Area under the curve: 0.752.

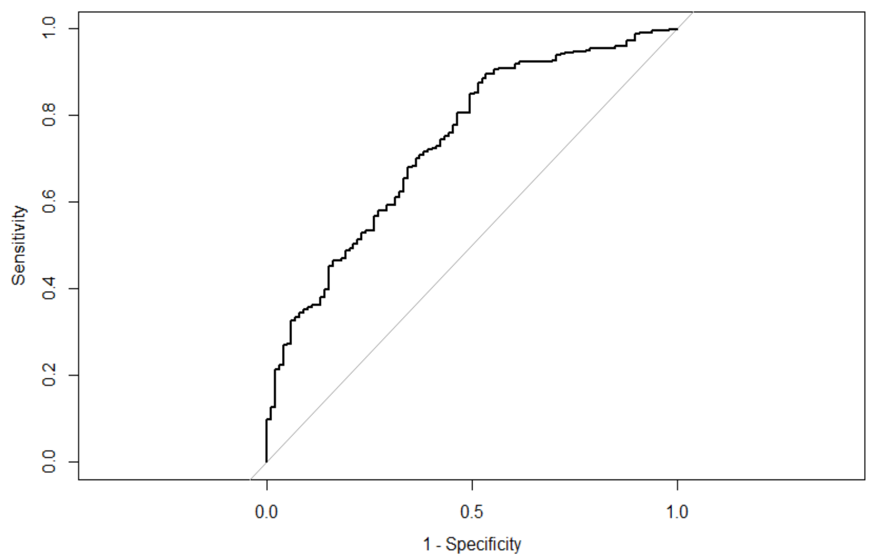

Figure 13. ROC curve results based on sixth-grade variables used to predict graduation utilizing downsampled logistic regression. Area under the curve: 0.736.

Examining the confusion matrices in Tables 33 and 34 gives more insight into the nuances of the tradeoff of correctly predicting a higher number of graduates and incorrectly identifying a higher number of nongraduates. As seen with the logistic regression models' ROC curves, if one wishes to obtain a higher true-positive rate (sensitivity), then one must accept a significant increase in the percentage of false-positive classifications (specificity subtracted from one).

Table 33

Confusion Matrix Results Based on Sixth-grade Variables Used to Predict Graduation Utilizing Upsampled Logistic Regression

	Actual no	Actual yes
Predicted no	59	192
Predicted yes	40	485

The confusion matrix results using the upsampled logistic regression model (see Table 33) had an accuracy of 0.68. When compared to the original data set, the upsampled model saw a large decrease in the true-positive rate to 0.71, accompanied by a large decrease in the false-positive rate to 0.40. Upsampling allowed the model to learn quite a bit more about nongraduates and, thus, it labeled far fewer nongraduates as graduates. However, it also labeled many graduates as nongraduates.

Along with the upsampled logistic regression model's accuracy value, the kappa identifies how well the model is predicting. The higher the kappa value, the better the model. An accuracy value of 68% and kappa of 0.17 are fair results. The F1 value of 0.79 reflects a good performance because for the F1 score, which is the weighted average score of recall and precision, a value at one is the best performance, and at zero is the worst. The balanced accuracy measures the accuracy using an equal number of trials in each class, and the upsampled logistic regression model had a 65% balanced accuracy.

The confusion matrix results using the downsampled logistic regression model (see Table 34) had a true-positive rate of 0.68 and a false-positive rate of 0.36. This method had the higher false-positive rate of the two logistic regression models. This better performance was due to the fact that downsampling reduced the training set size significantly by removing positive observations, preventing the model from learning more about graduates, but allowing it to see a proportionally higher number of nongraduates in its training.

Table 34

Confusion Matrix Results Based on Sixth Grade Variables Used to Predict Graduation Utilizing Downsampled Logistic Regression

	Actual no	Actual yes
Predicted no	63	214
Predicted yes	36	463

The downsampled logistic regression model had an accuracy value of 68% and kappa of 0.17, which are moderately poor results and the same as the upsampled performance. Remember, the higher the kappa value, the better the model. The F1 value of 0.79 also reflects a moderately performance because it is fairly close to one, which is the best performance. The balanced accuracy measures the accuracy using an equal number of trials in each class, and the downsampled logistic regression model had a 65% balanced accuracy. This balanced accuracy value was helped by its true-negative performance of 0.63, and it was also higher than the upsampled logistic regression's true-negative performance. The upsampled logistic regression's balanced accuracy was equal to the downsampled logistic regression's balanced accuracy.

Logistic regression and LDA upsampled and downsampled data had very similar conclusions and high performance with a range of only 0.001 to 0.003 difference for each of their areas under the ROC curve. Their model accuracies were virtually identical, and they identified several of the same statistically significant variables when predicting future dropouts or late graduates.

As seen in Figures 14 and 15, the upsampled, and downsampled LDA models yield area under the ROC curve results of 0.754 and 0.736, respectively. The upsampled LDA model was thus the best-performing model overall in terms of area under the ROC curve. The downsampled model was only marginally lower than its logistic regression counterpart.

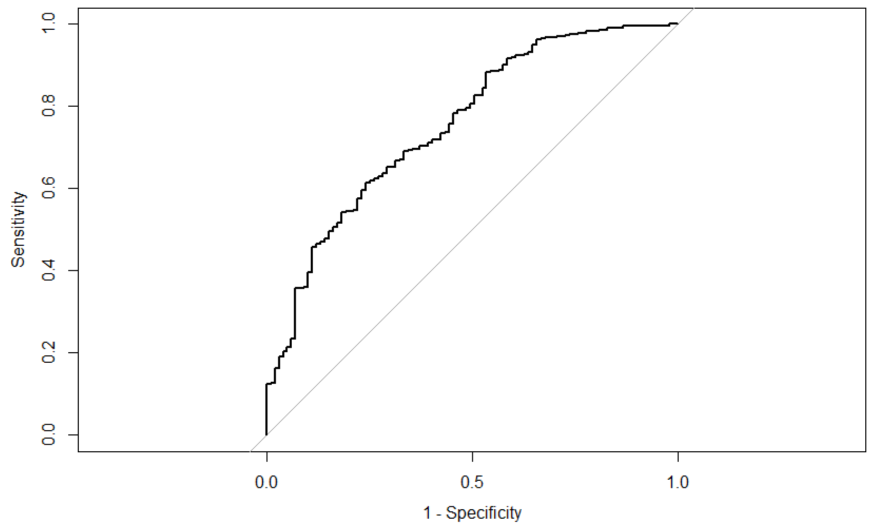

Figure 14. ROC curve results based on sixth-grade variables used to predict graduation utilizing upsampled linear discriminant analysis. Area under the curve: 0.754.

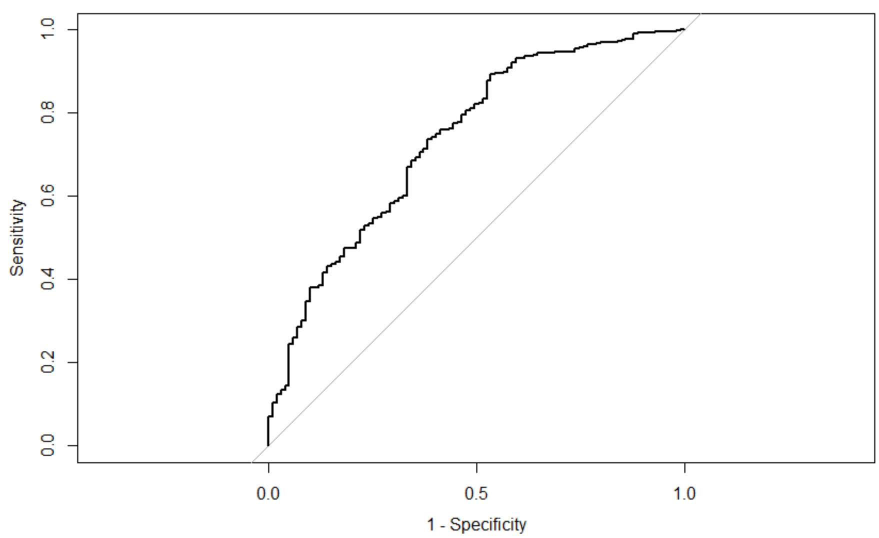

Figure 15. ROC curve results based on sixth-grade variables used to predict graduation utilizing downsampled linear discriminant analysis. Area under the curve: 0.736.

The confusion matrix results using the upsampled LDA model (see Table 35) had an accuracy of 0.69, a true-positive rate of 0.69, and a false-positive rate of 0.38. This performance was very similar to the upsampled logistic regression model.

Table 35

Confusion Matrix Results Based on Sixth Grade Variables Used to Predict Graduation Utilizing Upsampled Linear Discriminant Analysis

	Actual no	Actual yes
Predicted no	61	207
Predicted yes	38	470

The upsampled LDA model had a kappa level of 0.19, which is considered fair and similar to the upsampled logistic regression kappa value. The F1 value of 0.80 reflected a good performance and was very similar to the upsampled logistic regression performance. The balanced accuracy measures the accuracy using an equal number of trials in each class, and the upsampled LDA model had a 66% balanced accuracy. This balanced accuracy value was very similar to the upsampled logistic regression's value and was helped by its true-negative performance of 0.61.

Table 36

Confusion Matrix Results Based on Sixth-grade Variables Used to Predict Graduation Utilizing Downsampled Linear Discriminant Analysis

	Actual no	Actual yes
Predicted no	65	226
Predicted yes	34	451

The confusion matrix results using the downsampled LDA model (see Table 36) had a true-positive rate of 0.67 and a false-positive rate of 0.34. This performance was similar to the downsampled logistic regression model, revealing an important observation regarding these two models. Because the models performed quite similarly for the sixth-grade data, the practical outcomes of LDA and logistic regression could be quite similar in practice, even when the assumptions of LDA were not satisfied.

The downsampled LDA model had an accuracy value of 67% and kappa of 0.18, which are good results but not quite as good as the upsampled LDA's performance. The F1 value of 0.78 also reflects a good performance because it is close to 1, which is the best performance. The balanced accuracy measures the accuracy using an equal number of trials in each class, and the downsampled LDA model had a 66% balanced accuracy. This balanced accuracy value is the same as the upsampled LDA's value.

Figures 16 and 17 show the upsampled and downsampled QDA models yield area under the ROC curve results of 0.700 and 0.703, respectively. The QDA model performed the worst of the three models in terms of average area under the ROC curve.

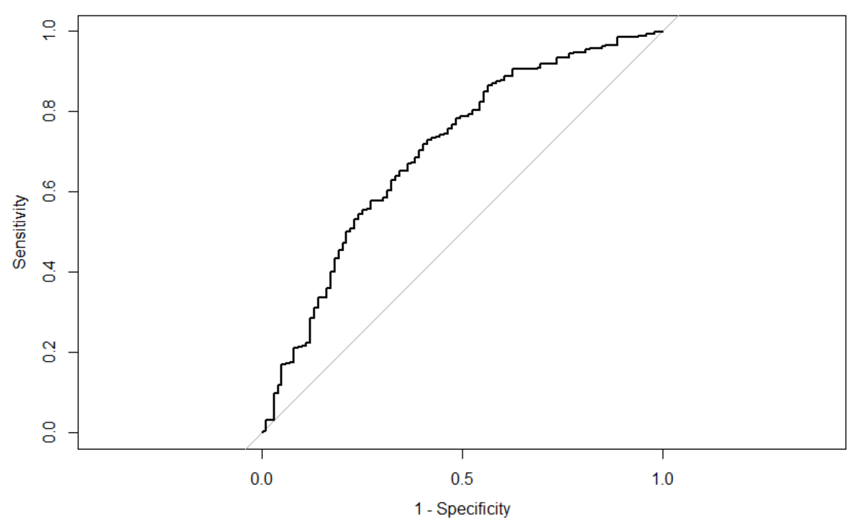

Figure 16. ROC curve results based on sixth-grade variables used to predict graduation utilizing upsampled quadratic quadratic discriminant analysis. Area under the curve: 0.700.

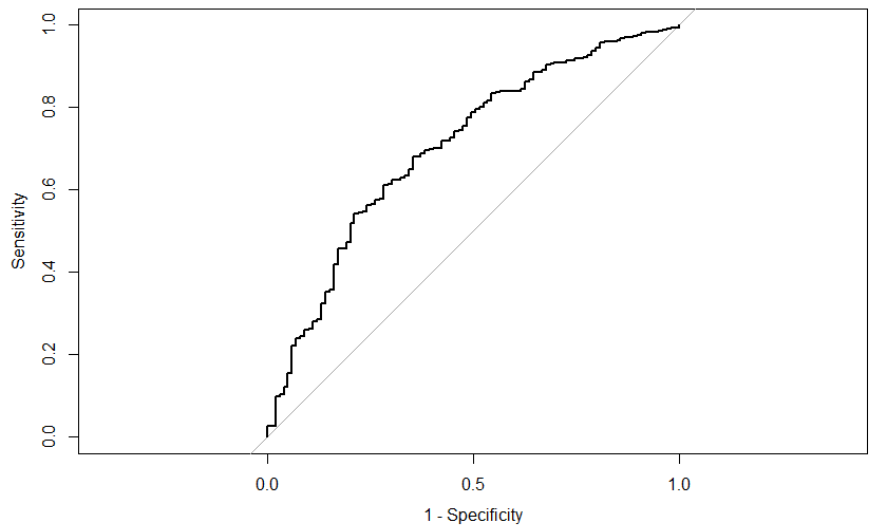

Figure 17. ROC curve results based on sixth-grade variables used to predict graduation utilizing downsampled discriminant analysis. Area under the curve: 0.703.

The confusion matrix results using the upsampled QDA model (see Table 37) had a true-positive rate of 0.88 and a false-positive rate of 0.61. Upsampling allowed the model to achieve a lower false-positive rate but also a lower true-positive rate, sacrificing recall (percentage of total relevant results correctly classified) for precision (percentage of results that are relevant).

Table 37

Confusion Matrix Results Based on Sixth-grade Variables Used to Predict Graduation Utilizing Upsampled Quadratic Discriminant Analysis

	Actual no	Actual yes
Predicted no	39	79
Predicted yes	60	598

The upsampled QDA model had an accuracy value of 82% and kappa of 0.26, which are good results and better than those of both the upsampled logistic regression and upsampled LDA values. The F1 value of 0.90 also reflects a good performance and higher than the upsampled logistic regression and upsampled LDA F1 values. The balanced accuracy measures the accuracy using an equal number of trials in each class, and the upsampled LDA model had a 64% balanced accuracy. This balanced accuracy value was lower than both the upsampled logistic regression and upsampled LDA's balanced accuracy values.

The confusion matrix results using the downsampled QDA model (see Table 38) had a true-positive rate of 0.78 and false-positive rate of 0.49. This performance was in line with the results seen for the downsampled logistic regression and LDA models, as downsampling allowed the model to achieve the lowest false-positive rate while sacrificing the most true-positive rate. Downsampling performance was a result of the equality of class sizes and the uniqueness of each observation in both classes.

Table 38

Confusion Matrix Results Based on Sixth-grade Variables Used to Predict Graduation Utilizing Downsampled Quadratic Discriminant Analysis

	Actual no	Actual yes
Predicted no	50	148
Predicted yes	49	529

The downsampled QDA model had an accuracy value of 75% and kappa of 0.20, which were good results and better than the downsampled data set for both the logistic regression and LDA models. The F1 value of 0.84 reflects a pretty good performance because it is close to 1. The balanced accuracy measures the accuracy using an equal number of trials in each class, and the downsampled QDA model had a 64% balanced accuracy. This balanced accuracy value was lower than the balanced accuracy of all of the other upsampled and downsampled models, although they all performed similarly.

Choosing the Best Model

The sixth-grade models displayed the same tradeoff between true-positive rate and false-positive rate that the ninth-grade models did, albeit with generally lower accuracy than the ninth-grade models. In terms of accuracy and true-positive rate, the upsampled QDA model is best, with an accuracy level of 0.83, a true-positive rate of 0.88, and a kappa value of 0.26. However, the high levels come at a price because the upsampled QDA also had the lowest balanced accuracy and highest false-positive rate. The model with the lowest false-positive rate was the downsampled LDA model (false-positive rate = 0.34), which showed that the downsampled LDA model performed much better in reducing false-positive predictions when it had small, equally-sized classes for its training. The downsampled LDA model also had the third-highest kappa value and the highest balanced accuracy.

In general, the prediction of graduation likelihood for sixth graders with the data available was more difficult than when using ninth graders' data. Although the sixth-grade models were not as accurate as the models that used ninth-grade variables, sixth-grade models performed with enough accuracy to allow middle school staff members to identify students in need of interventions. Those interventions can put students back on track to graduation well before the students enter high school.

Summary

When attempting to determine if one or more of the ninth-grade variables consisting of attending school at least 90% of the time, earning sufficient credits to move to the tenth grade, been suspended out of school, multiple school moves, standardized reading and math scores, failing no more than one semester of a core content course, school minority percentages, and student characteristics (ELL status, SWD status, free/reduced meal status, and male gender) were able to predict students who would not complete high school within four years, several were shown to have a statistically significant relationship. Namely, earning sufficient credits to move to the tenth grade was a significant predictor in all logistic regression and LDA models (significant at the $p = 0.001$ level in all logistic regression models). The analysis showed that if a student did not earn sufficient credits to advance to 10th grade, then the student was very unlikely to graduate from high school. The indicator variable for out-of-school suspension and the variable for multiple school moves were shown to be statistically significant predictors of graduation in the regular and upsampled logistic regression models, and they were also large determinants in the LDA models.

When deciding which statistical model was the most accurate at predicting future dropouts or late graduates utilizing ninth-grade variables, the use of the information is important. Because a main goal of the task is to identify nongraduates early in their schooling, it is wiser to choose a model with the lowest false-positive rate even if the true-positive rate suffers slightly from that choice. With that option, it is better to offer help to more students than need it than to fail to identify students who need help. This judgment leads to the conclusion that the downsampled QDA model is the best model for minimizing false-positive observations, as its false-positive rate was the lowest at 0.43, and its true-negative was the highest at 0.56. The downsampled QDA model also had the second-highest kappa value and the highest balanced accuracy. However, in terms of accuracy and true-positive rate, the downsampled LDA model is best, with an accuracy level of 0.89 and a true-positive rate of 0.98. However, the accuracy level and true-positive rate come at a price because the downsampled LDA also had the lowest kappa value, lowest balanced accuracy, and highest false-positive rate.

When determining if the sixth-grade variables consisting of failing English, failing math, attending less than 80% of the time, receiving out-of-school suspension, number of school moves, standardized reading and math scores, school minority percentages, and student characteristics (ELL status, SWD status, free/reduced meal status, race and male gender) are able to predict students who would not complete high school within four years, a few were shown to have a statistically significant relationship. Gender was a significant predictor in a few of the models, particularly logistic regression. Additionally, whether or not a student passed English was a strong predictor in the upsampled logistic regression model and the upsampled and downsampled LDA models. The analysis showed that if a student did not pass English in the sixth grade, then the student was unlikely to graduate from high school.

Two other sixth-grade variables identified in a few models as having a strong relationship with predicting high school graduation were: being suspended from school and making multiple school moves. Those two variables were identified in the upsampled data sets for both the logistic regression and QDA models. In addition, being suspended from school was identified as a strong predictor in the upsampled LDA and downsampled QDA models. Making multiple school moves in the sixth grade was also identified as a strong predictor in the downsampled logistic regression model.

Overall, variables consistently identified in most of the sixth-grade models as able to predict students who would not complete high school within four years were: (a) male gender (b) if the student was suspended from school, (c) if the student had multiple school moves, and (d) if the student did not pass English.

When deciding which statistical model was the most accurate at predicting future dropouts or late graduates utilizing sixth-grade variables, it can be concluded that the prediction of graduation likelihood for sixth graders was much more difficult than when using ninth graders' data. Because a main goal of the task is to identify nongraduates early in their schooling, it is wiser to choose a model with the lowest false-positive rate even if the true-positive rate suffers slightly from that choice. This judgment leads to the conclusion that the downsampled LDA model is the best model for minimizing false-positive observations, as its false-positive rate was the lowest at 0.34, and its true-negative was the highest at 0.65. The true-positive rate was still impressive at 0.67. The downsampled LDA model also had the third-highest kappa value and the highest balanced accuracy. However, in terms of accuracy and true-positive rate, the upsampled QDA model is best, with an accuracy level of 0.83, a true-positive rate of 0.88, and a kappa value of 0.26. However, the high levels come at a price because the upsampled QDA also had the lowest balanced accuracy and highest false-positive rate.

The results of the area under the curve and confusion matrix accuracy for both sixth grade and ninth grade also showed that it was more accurate to predict if a student would graduate within four years of high school using a student's ninth-grade data compared to a student's sixth-grade data. However, in this study, in terms of accuracy, the sixth-grade data using the downsampled LDA model outperformed the Philadelphia study of over 12,000 sixth graders (Balfanz et al., 2007). This study identified 65% of the students who would not graduate on time and had a false-positive rate of 34%, while the Philadelphia study identified 60% of the students who would not graduate on time and had a false-positive rate just under 40%.

Chapter V

SUMMARY, CONCLUSIONS, AND IMPLICATIONS

Dropping out of high school is a serious problem, not only for the individual but the school system, community, and society as well. A key strategy for economic growth is addressing the high school dropout crisis (Tucci, 2011). Improving education not only benefits the individual, but the gains compound to benefit the economy at the local, state, and national levels. Economic benefits include increased tax revenues, increased individual earnings, increased sales of homes and autos, more jobs, and economic growth, along with more spending and investment (Tucci, 2011). In addition, students who failed to graduate were more likely than their peers who graduated to be unemployed, live in poverty, receive public assistance, spend time in prison, be on death row, have poor health, and be a single parent with children who also drop out (Bridgeland et al., 2006).

The nation has focused attention on improving the graduation rates for all students, including those who have been underserved in the past or who present specific learning challenges (DePaoli et al., 2015). The increased attention from social, political, and governmental agencies has created pressure for educators to identify and intervene with students who seem likely not to graduate from high school (Bruce et al., 2011).

Now more than ever, better research and data are available for school leaders to learn both who will likely drop out and what effective interventions to implement. There are several factors as to why a student chose to drop out, and there are variables along the way that can identify who is likely to drop out (Rumberger & Lim, 2008). Early warning systems could aid dropout prevention by testing local indicators to identify accurately students who are likely to drop out and to aid schools in identifying those who need interventions (Pinkus, 2008). Accurate and early identification of students at risk for not graduating can lead to appropriate and timely interventions that increase the likelihood of students completing high school (McKee & Caldarella, 2016).

Summary of Findings

Independent variables for this study were chosen based on a thorough literature review of dropout indicators at both the ninth-grade and sixth-grade levels and, thus, were used as predictor variables in this study. The most accurate ninth-grade dropout indicators that are both influenced and not influenced by schools are: attending school at least 90% of the time, earning sufficient credits to move to the tenth grade, number of days suspended out of school, number of school moves, standardized reading and math scores, failing no more than one semester of a core course, school minority percentages, school poverty percentages, and student characteristics (ELL status, SWD status, free/reduced meal status, Black race, and male gender) (Allensworth & Easton, 2005; DePaoli et al., 2015; Kemple et al., 2013; Lee et al., 2011; Mac Iver & Mac Iver, 2010; Mac Iver & Messel, 2012; 2013; Neild, 2009; Zvoch, 2006).

The most accurate sixth-grade dropout indicators in these two categories are: failing English, failing math, attending at least 80% of the time, receiving out-of-school suspension, number of school moves, standardized reading and math scores, school minority percentages, school poverty percentages, and student characteristics (ELL status, SWD status, free/reduced meal status, Black race, and male gender; Balfanz, 2009; Balfanz et al., 2007; Jerald, 2006; Mac Iver, 2010; Rumberger, 2004; Silver et al., 2008).

The purpose of this nonexperimental, correlational study was to use longitudinal data from a mid-sized school district from two cohorts to support the creation of a dropout early warning system to predict which middle school and high school students are at risk for not graduating on time. Both ninth-grade and sixth-grade longitudinal data were used to create this dropout early warning system.

A few variables in both cohorts had missing data. Because this missing data could undermine the ability to make valid inferences for this study, the missing data were imputed. The missing data were imputed utilizing the bagged tree model within the caret package in R, a programming language for statistical computing and data visualization. For each predictor in the data, a bagged tree was created using all the other predictors in the data set.

To help with the imbalance of graduates and nongraduates, upsampling of the minority class and downsampling of the majority class of the data were performed before training the model. The end result for both upsampling and downsampling is the same number of observations from the minority and majority classes. When running each analysis, the two cohorts were used as training and testing sets. Each model could train using roughly half the available data and then test on the other half. For both the ninth-grade and sixth-grade cohorts, the 2017 variable data were used as the training data, and the 2018 variable data were used to test the accuracy of the model.

This study identified the most accurate indicators at each respective grade level and identified the most accurate statistical models that resulted in high levels of true classification and low levels of false identification for all middle schools and high schools in a mid-sized school district in Georgia.

One of the statistical procedures used in this correlational study was logistic regression. Logistic regression was used to examine the relationship among the variables to yield a maximum correlation to predict which students will drop out or not finish within four years (Ary et al., 2014). Two other alternative and widely used statistical procedures utilized in this study were linear discriminate analysis (LDA) and quadratic discriminate analysis (QDA). For the multiple variables, the models estimate the mean and variance from the data for each class. Both the LDA and QDA algorithms make predictions by estimating the probability that a new set of inputs belongs to a particular class or group. The class with the highest probability was the output class, and therefore, the prediction (James et al., 2013).

Conclusions for Ninth Grade

In this study, logistic regression was one of three statistical models used to predict whether or not a student would graduate within four years based on known student ninth-grade data. The logistic regression analysis finds the best fitting model to describe the relationship between a student's graduation status and a variety of student and school variables. This statistical model estimates the probability of certain events occurring based on some previous data. For the upsampled logistic regression prediction, the accuracy was 0.890, true-positive rate of 0.97, and false-positive rate of 0.66. The downsampled version had an accuracy of 0.881, true-positive rate of 0.94, and false-positive rate of 0.54. Another measure of the quality of the model is the Akaike Information Criterion (AIC) value. If two similar models are compared, then the model with the lower AIC is superior. The upsampled logistic regression model had an AIC level of 1545.20 and can be compared to the downsampled logistic regression model's AIC level of 272.36. Thus, the downsampled model was identified as the preferred model.

The Nagelkerke pseudo R-squared value is helpful when it is compared to another pseudo R-squared of the same type and predicting the same outcome. This value can be helpful when deciding which logistic regression model is best. A higher pseudo R-squared indicates which model does a better job of predicting the outcome (UCLA: Statistical Consulting Group., n.d.). The upsampled logistic regression model had an Nagelkerke pseudo R-squared value of 0.492 and can be compared to the downsampled logistic regression model's pseudo R-squared value of 0.518. When compared to the upsampled logistic regression model's pseudo R-squared level, the downsampled logistic regression model is again identified as the preferred model.

For the upsampled logistic regression model, out of the 13 predictor variables, eight were statistically significant: gender, attend at least 90% of days, earn sufficient credits to advance to tenth grade, failed less than two core courses, if suspended out of school, number of school moves in ninth grade, scaled score of standardized algebra test, and race, but it was shown to not be as significant as initially believed. By earning sufficient credits to advance to the tenth grade, a student increases their odds of graduating by a factor of 5.2, given all other variables are unchanged. If a student received an out-of-school suspension, the odds of a student graduating decreased by 57.9% (0.421 – 1 = -0.579), keeping other variables constant.

For the downsampled logistic regression model, three variables were statistically significant: gender, earning sufficient credits to advance to tenth grade, and the number of school moves a student makes in the ninth grade. By earning sufficient credits to advance to the tenth grade, a student increases their odds of graduating by a factor of 4.3, given that all other variables are unchanged. If a student had multiple moves in the ninth grade, the odds of a student graduating decreased by 28.9% (0.711 – 1 = -.289), keeping all other variables constant.

Like logistic regression, linear discriminant analysis (LDA) predicts a dependent variable based on the values of independent variables. Linear discriminant analysis makes predictions by estimating the probability that new inputs belong to a particular class. The coefficients of linear discriminants were the values used to predict whether or not each student would graduate within four years. The coefficients are similar to regression coefficients. The category with the highest probability was, therefore, the prediction. The higher the coefficient, the more weight it had.

The upsampled LDA model with an accuracy of 89.0%, true-positive rate of 0.97, and false-positive rate of 0.66 had results that closely aligned with those of the upsampled logistic regression models. That is, the coefficients of linear discriminants for race (0.466), gender (0.568), attend at least 90% of days (0.519), earning sufficient credits to advance to tenth grade (0.985), failing less than two core courses (0.499), and if suspended out of school (-0.550) show that those predictors were the most influential in determining if a student graduated high school, as they were the ones with the highest magnitude.

Earning sufficient credits to advance to the tenth grade was still the most influential predictor with a coefficient of 0.985, but now gender had the second-highest weighted coefficient (0.568). Another strong predictor in the upsampled LDA model was fail9 (coefficient of 0.499), suggesting that the students who fail less than two classes are much more likely to graduate. The group means confirmed this, as 77.1% of graduates failed less than two classes, and only 35.6% of nongraduates failed less than two classes in the ninth grade.

The downsampled LDA model had an accuracy score of 88.7%, true-positive rate of 0.98, and false-positive rate of 0.75, which was slightly less than the upsampled LDA model. A strong predictor in the downsampled LDA model was earning sufficient credits to advance to tenth grade, which had a coefficients of linear discriminants value of 1.064. The nongraduates had a group mean of 0.464, and the graduates had a group mean of 0.906. The group means showed that 90.6% of the graduates earned sufficient credits to advance to the 10th grade, while only 46.4% of the nongraduates did. Other predictors with high coefficients of linear discriminants are qualifying for special education (0.588), gender (0.724), and if suspended out of school (-0.762). Those predictors were the most influential in determining a student's likelihood of graduation, as they were the variables with the highest values.

Using quadratic discriminant analysis (QDA), each observation was classified in the group that had the least squared distance. Quadratic discriminant analysis was used to analyze the research data to understand the relationship between the dependent variable and different independent variables. The output contains the group means, but because the analysis involved a quadratic, rather than a linear function of the predictors, QDA does not provide the coefficient of the linear discriminant for each variable like the LDA did (Crouser, n.d.). For the upsampled QDA prediction, the accuracy was 0.874, true-positive rate was 0.94, and false-positive rate was 0.63. The downsampled version had an accuracy of 0.854, true-positive rate of 0.90, and false-positive rate of 0.43. These were both lower than the accuracies of the best-performing logistic regression and LDA models, suggesting that the decision boundary is better estimated with a linear discriminant function.

The upsampled QDA group means showed that 93.9% of the graduates earned sufficient credits to advance to the 10th grade, while only 48.3% of the nongraduates did. Similarly, only 16.2% of graduates received out of school suspension while 53.4% of nongraduates received out of school suspension, and 88.1% of graduates attended at least 90% of school days while only 49.7% nongraduates attended at least 90% of school days. These results confirm that students who earn sufficient credits to advance to the 10th grade, have fewer suspensions, and attend more school are all more likely to be on track to graduate.

The downsampled QDA group means show the drastic difference between the means of graduates and nongraduates for the variables of earning sufficient credits to advance to tenth grade, attending at least 90% of school days, and if suspended from school just as it did for the upsampled QDA model.

Overall, variables consistently identified in a majority of the ninth-grade models as able to predict students who would not complete high school within four years were: (a) if a student did not receive enough credits to advance to the tenth grade, (b) if a student did not attend school at least 90% of the time, (c) if a student was suspended from school, (d) if a student had multiple school moves in the ninth grade, and (e) male gender. Overall, race was not considered a significant predictor due to conflicting data.

The ninth-grade variables identified in this study as best able to predict students who would not complete high school within four years were also identified in major studies throughout the United States. Receiving enough credits to advance to the tenth grade was also identified in a Chicago study of data from both the 2001 and 2004 freshman cohorts, which included 23,564 and 26,562 students, respectively. Students who were on track by the end of ninth grade were more than three and one-half times more likely to graduate high school in four years than off-track students. Also, 81% of the students on track at the end of their freshman year graduated from high school in four years, but only 22% of the off-track students graduated in four years. On-track students were three times as likely to graduate within five years as off-track students (Allensworth & Easton, 2005). There is a need for interventions to ensure that more ninth graders pass and earn the credits needed to stay on track to graduate because waiting until the end of ninth grade to intervene is often too late (Mac Iver, 2013).

The variable of poor attendance was identified in a study of Baltimore City Public Schools in both the 2005 (n = 6,812) and 2006 (n = 7,729) school years. This study's purpose was to determine if ninth-grade school-level factors were associated with non-graduation outcomes, including student attendance. Varying critical attendance thresholds during transition years help signal that a student is at risk of falling off the graduation path. In the Baltimore study, a student was considered chronically absent if the student missed more than 20 days. Data from two freshmen cohorts revealed that attendance played a significant role in who graduated on time and who did not. While

82% of first-time ninth graders with attendance of 95% or higher graduated on time, only 26% of those who missed more than 20 days graduated (Mac Iver & Messel, 2012; 2013).

The variable of gender, particularly being male, was identified in a Neild and Balfanz (2006) study of 130,000 students enrolled in Philadelphia schools in grades six through twelve during the 2003-2004 school year. The purpose of the study was to find the key characteristics of Philadelphia's 13,000 dropouts and those considered near-dropouts. During the 2003-2004 school year, males were considerably more likely than females to drop out but only somewhat more likely to be near-dropouts. Additional analysis in Philadelphia found that first-time freshman cohorts with expected graduation dates of 2000, 2001, 2002, 2003, 2004, and 2005 consistently graduated females at a rate of at least 10% higher than males with an almost 15% advantage from 2000 through 2003.

For the research question of which statistical model was most accurate at predicting future dropouts or late graduates utilizing ninth-grade variables, both an ROC curve and confusion matrix were used to identify the most accurate statistical model. The ROC curve is a plot of the true-positive rate (sensitivity) against the false-positive rate (specificity) at various threshold settings. The results of the area under the curve show that ninth-grade variables predict if a student will graduate within four years of high school with relatively high accuracy for all of the statistical models because the area under the curve was 0.801 and higher for every test. Of the three statistical models, the upsampled and downsampled logistic regression analyses had the highest area under the curve values at 0.842 each. Linear discriminant analysis for both the upsampled and downsampled data sets had area under the curve values of 0.841. The results of the quadratic discriminant analysis area under the curve was 0.812 for both the upsampled and downsampled data sets.

A confusion matrix is a table with two rows and two columns that reports the number of false-positives, false-negatives, true-positives, and true-negatives. The results of the confusion matrix show that ninth-grade variables predict if a student will graduate within four years of high school with relatively high accuracy for all of the statistical models because the accuracy was 0.85 and higher for all the analyses. The accuracy for both the upsampled and downsampled linear discriminant analyses was 0.89. The upsampled and downsampled linear discriminant analyses had true-positive rates of 0.97 and 0.98, respectively, and false-positive rates of 0.66 and 0.75, respectively. The upsampled logistic regression had the same accuracy of 0.89, true-positive rate of 0.97, and false-positive rate of 0.66. The downsampled logistic regression had an accuracy of .88, true-positive rate of 0.94, and a false-positive rate of 0.54. The analyses for the upsampled quadratic discriminant analysis also showed good results of 0.87 for accuracy, a 0.94 true-positive rate, and a 0.63 false-positive rate. The downsampled quadratic discriminant analysis had an accuracy level of 0.85, a true-positive rate of 0.90, and a false-positive rate of 0.43.

The balanced accuracy's best possible value is one, and the worst possible value is zero. Balanced accuracy measures the average of the proportion correct using an equal number of trials in each class. The ninth-grade balanced accuracy performed decently with values ranging from 0.62 to 0.73. However, within all of these models, performance metrics is a tradeoff between the true-positive rate and the false-positive rate. This tradeoff exists in almost all classification problems. Some models are excellent at predicting positive observations (high true-positive rate), but that is almost always associated with a low detection rate (high false-positive rate). In this study, because there was a significant class imbalance in the training data, the models all had difficulty in detecting nongraduates, as there were relatively fewer observations of nongraduates. Even when using upsampling and downsampling, the false-positive may decrease, but it still remained quite high in most of the models. One must choose a model based on the tradeoffs presented and the problem at hand.

Because a main goal is to identify nongraduates early in their schooling, then it is wiser to choose a model with the lowest false-positive rate even if the true-positive rate suffers slightly from that choice. It is a better option to offer help to more students than need it than to fail to identify students who need help. This judgment leads to the conclusion that the downsampled QDA model is the best model for minimizing false-positive observations, as its false-positive rate was the lowest at 0.43, and its true-negative was the highest at 0.57. The downsampled QDA model also had the second-highest kappa value and the highest balanced accuracy.

Conclusions for Sixth Grade

In this study, logistic regression was used to predict whether or not a student would graduate within four years based on known student sixth-grade data. For the upsampled logistic regression prediction, the accuracy was 0.679, true-positive rate was 0.71, and false-positive rate of 0.40. The downsampled version had an accuracy of 0.728, true-positive rate was 0.68, and false-positive rate of 0.36. Another measure of the quality of the model is the Akaike information criterion (AIC) value. If two similar models are compared, then the model with the lower AIC is superior. The upsampled logistic regression model had an AIC level of 1944.0 and could be compared to the downsampled logistic regression model's AIC level of 346.4. Thus, the downsampled model is identified as the preferred model.

The Nagelkerke pseudo R-squared value is helpful when it is compared to another pseudo R-squared of the same type and predicting the same outcome. A higher pseudo R-squared indicates which model does a better job of predicting the outcome (UCLA: Statistical Consulting Group., n.d.). The upsampled logistic regression model had an Nagelkerke pseudo R-squared value of 0.312 and can be compared to the downsampled logistic regression model's pseudo R-squared value of 0.366. When compared to the upsampled logistic regression model's pseudo R-squared level, the downsampled model is again identified as the preferred model.

For the upsampled logistic regression model, out of the 13 predictor variables, eight were statistically significant: receiving free or reduced meals, gender, passing English in the sixth grade, being suspended from school, moving multiple times in the sixth grade, and scaled scores on the reading and math standardized tests, and race. However, race was not supported as a significant variable in the other models. By passing English in the sixth grade, a student increases their odds of graduating by a factor of 4.7, given all other variables are unchanged. If a student received free or reduced lunch, the odds of a student graduating decreased by 70.2% (0.298 – 1 = -0.702), keeping other variables constant.

For the downsampled logistic regression model, out of the 13 predictor variables, only two were statistically significant: race and the number of school moves a student makes in the sixth grade. Again, race was not supported as a significant variable in the other models. If a student had multiple moves in the sixth grade, the odds of a student graduating decreased by 48.9% (0.511 – 1 = -0.489), keeping all other variables constant.

Linear discriminant analysis makes predictions by estimating the probability that new inputs belong to a particular class. The coefficients of linear discriminants were the values used to predict whether or not each student would graduate within four years. The category with the highest probability was selected. The upsampled LDA model had an accuracy of 0.693, true-positive rate of 0.69, and false-positive rate of 0.38 and had results that closely aligned with those of the upsampled logistic regression models. That is, the coefficients of linear discriminants—gender (0.768) and passing English in the sixth grade (0.733)—show that those predictors were the most influential in determining a student's likelihood of graduating, as they were the ones with the highest magnitude. Having a lower chance of graduating from high school was associated with the student receiving free or reduced lunch (-0.772) and being suspended out of school (-0.847).

The downsampled LDA model, which had an accuracy level of 0.670, true-positive rate of 0.67, and false-positive rate of 0.34, performed worse than the upsampled model. In this model, the most important features were passing English (1.416), race (1.22), and gender (1.02).

Using quadratic discriminant analysis (QDA), each observation was classified in the group that had the least squared distance. Quadratic discriminant analysis was used to analyze the research data to understand the relationship between the dependent variable and different independent variables. For the upsampled QDA prediction, the accuracy was 0.830, the true-positive rate was 0.88, and the false-positive rate was 0.61.

The downsampled version had an accuracy of 0.746, true-positive rate of 0.78, and false-positive rate of 0.49. For the upsampled data set, there was a big difference in group means for being suspended out of school, as the nongraduates had a group mean of 0.357, and the graduates had a group mean of 0.137. The group means showed that 13.7% of the graduates had been suspended in the sixth grade, while 35.7% of the nongraduates were suspended. These results mean that students who were suspended in the sixth grade are less likely to graduate or graduate on time than sixth-grade students who were not suspended in the sixth grade.

The downsampled QDA group means show the drastic difference between graduates and nongraduates means for the attendance and behavior variables. This difference means that sixth-grade students who missed more than 20% of school days and were suspended are less likely to graduate or graduate on time than students who missed 20% or fewer school days and were not suspended in the sixth grade.

The QDA models outperformed the upsampled and downsampled LDA models, suggesting that the decision boundary for that data might be better fit by a quadratic curve rather than a linear one. Of the three statistical models, the upsampled and downsampled QDA models had the highest accuracy at 82.0% and 74.6%, respectively. The downsampled LDA model had the lowest accuracy at 67.0%.

Unlike ninth grade, sixth grade did not have many variables that consistently had a strong relationship with predicting high school students who would not graduate on time. The variables of race and gender were significant predictors in a few of the models, particularly logistic regression. But, race was not considered a significant predictor due to conflicting data. Also, whether or not a student passed English was a strong predictor in the upsampled logistic regression model and upsampled and downsampled LDA models.

Overall, variables consistently identified in most of the sixth-grade models as able to predict students who would not complete high school within four years were: (a) the student's gender (male), (b) if the student was suspended from school, (c) if the student had multiple school moves, and (d) if the student did not pass English.

The sixth-grade variables identified in this study as best able to predict students who would not complete high school within four years were also identified in major studies throughout the United States. The variable of being male was associated with higher dropout rates using the National Longitudinal Survey of Youth (NLSY) data. Robst and de Vries (2010) used the NLSY data, which contained information about children born in 1979 to women in the 1979 NLSY cohort, including 281 boys and 289 girls. The purpose of the study by Robst and de Vries (2010) was to examine the relationship between childhood emotional and behavioral problems and the probability of graduating from high school while controlling for various sociodemographic characteristics. In 2004, graduation data were examined when the participants were between the ages of 24- and 27-years-old. There were a few gender differences in educational outcomes between boys and girls. Seventy-two percent of the boys received their high school diploma compared to 77% of girls. In the study, boys were more likely than girls to drop out of school with boys having a 14.6% dropout rate compared to an 8.6% dropout rate for girls ($X^2 = 4.7$, $p = 0.029$) (Robst & de Vries, 2010).

Being suspended from school was identified by the Balfanz et al. (2007) study of a Philadelphia cohort of approximately 13,000 students who entered the sixth grade in

September 1996. This data helped to identify flags that had both high predictive power and high yield. A flag had high predictive power if 75% or more of the sixth graders who were flagged did not graduate within one year of their expected graduation date. A flag would have a high yield if it identified 10% or more of future nongraduates. Being suspended while in the sixth grade was a flag identified in the Philadelphia study. Students with unsatisfactory behavior were 56% less likely to graduate when compared to students who did not.

Having multiple school moves is a variable identified by Wisconsin's Dropout Early Warning System (DEWS). According to the DEWS Data Brief, Wisconsin's DEWS used the R statistical software to run a state-of-the-art, machine-learning algorithm to test and combine up to 50 statistical models per grade for grades six through nine. This model predicted between 60% and 65% of future dropouts and late graduates and had a low false-positive rate (Wisconsin Department of Public Instruction, 2015). Wisconsin's DEWS did a good job of balancing the trade-off between correctly identifying likely dropouts and false-alarms by using the ROC plot. DEWS used data indicators that included school moves because of its accuracy and low false-positive rate.

Race was not strongly associated with dropping out in the Los Angeles Unified School District's (LAUSD) study. The purpose of the LAUSD seven-year longitudinal study for the class of 2005 by Silver et al. (2008) was to find the individual student's and school's characteristics at both the middle and high school levels that were associated with dropping out. That study's cohort consisted of 48,561 students who attended 163 LAUSD middle and high schools in the second-largest school district in the country. In the LAUSD study, student characteristics, including race or ethnicity, explained only 4% of the student-level variation in graduation rates.

The variable of not passing English in the sixth grade was supported by the Baltimore Education Research Consortium (2011) study. The study involved 7,887 sixth-grade students in the 2000-2001 cohort and 5,816 sixth-grade students in the 2008-2009 cohort from Baltimore City Schools. The study looked for indicators that could predict eventual dropouts with a reasonable level of certainty so that interventions could be put into place. The variable criteria included projects with at least 70% probability that a student with the indicator will not graduate, and more than 20% of the students who eventually dropped out possessed the indicator. One of the predictive sixth-grade early warning indicators for not graduating was failing English in the sixth grade. Failing a core course in the sixth grade was strongly associated with a lower likelihood of graduating from high school. Less than a third of sixth graders who failed a core course in the study eventually graduated (Baltimore Education Research Consortium, 2011).

For the research question of which statistical model was most accurate at predicting future dropouts or late graduates utilizing sixth-grade variables, the area under the curve shows that sixth-grade variables predict with relatively high accuracy for all of the statistical models. The area under the curve was 0.700 and higher for every test. The results for the area under the curve values for the upsampled and downsampled logistic regression analyses were 0.752 and 0.736, respectively. Linear discriminant analysis had an area under the curve value of 0.754 for the upsampled data set and an area under the curve value of 0.736 for the downsampled data set. The results for the area under the curve values for the upsampled and downsampled quadratic discriminant analyses were 0.700 and 0.703, respectively.

The results of the confusion matrix show that sixth-grade variables predict if a student will graduate within four years of high school with reasonable accuracy for all of the statistical models because the accuracy was 0.67 and higher for all analyses. The accuracy of the quadratic discriminant analysis showed good results of 0.83 for the upsampled data set and 0.75 for the downsampled data set. Logistic regression was close behind with an accuracy of 0.68 for both the upsampled and downsampled logistic regression. The results of the upsampled and downsampled linear discriminant analysis had an accuracy of 0.69 and 0.67, respectively.

The sixth-grade models displayed the same tradeoff between true-positive rate and false-positive rate that the ninth-grade models did, albeit with generally lower accuracy than the ninth-grade models. In terms of accuracy and true-positive rate, the upsampled QDA model is best, with an accuracy level of 0.83, a true-positive rate of 0.88, and a kappa value of 0.26. However, the high levels come at a price because the upsampled QDA also had the lowest balanced accuracy and highest false-positive rate. The model with the lowest false-positive rate was the downsampled LDA model (false-positive rate = 0.34), which showed that the downsampled LDA model performed much better in reducing false-positive predictions when it had small, equally-sized classes for its training. The downsampled LDA model also had the third-highest kappa value and the highest balanced accuracy.

In general, the prediction of graduation likelihood for sixth graders with the data available was more difficult than when using ninth graders' data. Although the sixth-grade models were not as accurate as the models that used ninth-grade variables, sixth-grade models perform with enough accuracy to allow middle school staff members to identify students in need of interventions.

Limitations

This study contributes to the current body of research available regarding indicators used to predict students at risk of dropping out. However, some limitations exist that impact the ability to generalize the results. The sample data for this study was taken exclusively from one school district; therefore, the results may not be generalized to other school districts with a different population. The variables tested in this study were identified through an extensive review of literature, but there are some variables that were omitted from this study. The omitted variables could contribute to the student's eventual high school end status. For example, the age of entering high school, level of neighborhood poverty from the Census report, and whether or not the student was at least one-year overage for his or her grade level could have been statistically significant variables.

Additionally, not all of the assumptions for the three models used were met. Assumptions for logistic regression include that the independent variables are linearly related to the log odds, and the model should have little or no multicollinearity (Kassambara, 2018b). Linear discriminant analysis assumptions require multivariate normality for each level of the grouping variable, homogeneity of the variance and covariance (homoscedasticity), and little or no multicollinearity (Brownlee, 2016). Quadratic discriminant analysis has the same assumptions as linear discriminant analysis except that each class is required to have its own covariance matrix (Kassambara, 2018a).

Additionally, both linear discriminant analysis and quadratic discriminant analysis are sensitive to outliers.

The ninth-grade data and sixth-grade data were checked for multivariate normality utilizing a multivariate Chi square Q-Q plot. The plots showed that the data were not distributed in a multivariate normal fashion because, for both grade levels, there was a heavy concentration of data in the mean regions but also heavy tails and skewness. Histograms confirmed that the data for both sixth grade and ninth grade were from a nonnormal distribution. A violation of the assumption of normality could impact the ability to trust the results and validly draw inferences about the results. Although linear discriminant analysis and quadratic discriminant analysis did not meet the assumption multivariate normality, the true accuracy of the tests was provided utilizing both the area under the curve and the confusion matrix analysis. Both models had high accuracy values that were similar to that of logistic regression, which did not have the assumption of normality. In addition, for the upsampled and downsampled data sets, the false-positive rates were often lower for linear discriminant analysis and quadratic discriminant analysis when compared to logistic regression. Therefore, not meeting the normality assumption did not significantly impact the performance of the linear discriminant and quadratic discriminant analyses when compared to logistic regression.

Coding in future studies should switch for graduates and nongraduates. Nongraduates should be coded as a one, and graduates should be coded as a zero, which is the opposite of the coding used in this study. This coding switch is necessary because the study's purpose is to identify those students who are predicted to not graduate within four years of entering high school. Therefore, nongraduates should be coded as a one instead of a zero.

Implications

Utilizing researched academic and behavioral predictors at the right time in a student's school career could help school personnel accurately identify students at risk for not graduating on time and provide the best interventions to put the students back on the path to graduation. Students often signal that they are on or off track on the path to graduation through their attendance, behavior, and course performance. An early warning system could be used at the school level as well as the district level to identify the students who are in need of guidance and interventions in order to graduate within four years of starting high school. The impact of putting students back on track to graduation could be life changing for the students and result in a better quality of life for both the students, their families, and society.

The results of this study determined if particular school-related variables and non-school-related variables identified with high accuracy and low false-positive results the students who would not graduate within four years, and if so, identify those variables. Accurately identifying potential dropouts is an important step to increasing the high school graduation rate. This study could add to the current body of knowledge about the variables that help identify students at risk for dropping out of high school and, thus, could have a positive impact on students' educational outcomes for many years to come. School personnel can be more strategic in the identification of students who need support by utilizing proven variables to accurately identify the at-risk students and provide them interventions early.

The results of this study contribute to the body of knowledge regarding the identification of variables and statistical models used to accurately identify students at risk of dropping out of high school. The results support previous research that identified particular ninth-grade and sixth-grade dropout indicators. Researchers may find the results of this study useful in designing future research to refine and further define how the use of statistical models and predictive variables can be used to increase the high school graduation rate through early identification and intervention.

Conclusion

Because of the negative consequences caused by dropping out of high school, it is imperative that schools take steps to ensure all students graduate. Not only do students' futures depend on it, but society's as well. Districts can take the first step in solving the dropout problem by identifying likely high school dropouts using a dropout early warning system. An early warning and multi-tiered response system is essential to ending the dropout crisis in schools, districts, and the nation. Creating an accurate dropout early warning system involves finding the best combination of indicators considering both true-positive and false-positive results and identifying when it is best to begin identifying the students. The variables and models identified in this study add to the literature and help schools identify students early, address students' needs, and put them on track to graduate on time by utilizing effective interventions. The social and economic contributions of these young people cannot be underestimated.

The variables identified in this study can easily be used by school staff to identify those students who are at risk of not graduating high school within four years. In the ninth grade, variables such as checking credits at the end of the school year, school attendance, school suspensions, gender, and number of schools attended in the ninth grade can all easily be pulled from the school's student database. The same is true for sixth-grade variables identified in this study. Those sixth-grade variables include school suspensions, gender, number of schools attended in the sixth grade, and if the student did not pass English.

Those variables in certain models can predict with relatively high true-positive rate and low false-positive rate. It is wiser to choose a model with the lowest false-positive rate even if the true-positive rate suffers slightly from that choice. It is a better option to offer help to more students than need it than to fail to identify students who need help. It is this option that leads to the conclusion that for ninth grade, the downsampled QDA model is the best model for minimizing false-positive rate which was the lowest at 0.43 and its true-negative value which was the highest at 0.56. It also had an accuracy level of 0.85 and a true-positive rate of 0.90. For the sixth grade, the downsampled LDA model had the lowest false-positive rate at 0.34 and accuracy and true-positive rate of 0.67. This study's sixth-grade accuracy was higher than all of the sixth-grade studies in the literature reviews mentioned.

Although school staff can more accurately identify students who will not graduate on time when students are in the ninth grade, it is important to also utilize sixth grade to intervene and help put students back on the path to graduation earlier. The reduction in accuracy is negligible when compared to the benefits of helping identified students get back on the path to graduation three years sooner.

Creating an accurate dropout early warning system is a critical component to combatting the dropout problem. This study identified both middle school and high school students likely to drop out or not graduate within four years of entering high school. Identification of those students allows schools to provide interventions to get them on track for on-time graduation. By accurately identifying the students most at risk for not graduating on time, schools could use their limited resources effectively. Dropout prevention initiatives are expensive and should not be wasted on students who do not require an intervention. In addition, not providing an intervention to future dropouts who need it is a disservice because they could have been saved.

Another benefit of the use of a dropout early warning system throughout a district and potentially the state could be to help increase the graduation rate and, thus, improve the lives of thousands of people. This improvement could come in the form of higher earned income, less crime, higher employment rates, and even better quality of life for future generations. The most compelling reason to focus on early identification is that the benefits can be far-reaching for society and make a better quality of life in general. After all, education is central to the well-being of society.

REFERENCES

Alexander, K., Entwisle, D., & Horsey, C. (1997). From first grade forward: Early foundations of high school dropout. *Sociology of Education*, *70*(2), 87–107. Retrieved from https://www.learnatnorth.org/lchaffee/PSY100/Journal Articles/Alexander et al 1997.pdf

Allensworth, E. (2013). The use of ninth-grade early warning indicators to improve Chicago schools. *Journal of Education for Students Placed at Risk*, *18*(1), 68–83. https://doi.org/10.1080/10824669.2013.745181

Allensworth, E., & Easton, J. (2005). The on-track indicator as a predictor of high school graduation. *Consortium on Chicago School Research at the University of Chicago*. Retrieved from http://www.aaronjmeyer.com/storage/OnTrackIndicator.pdf

Allensworth, E., & Easton, J. (2007). What matters for staying on-track and graduating in Chicago Public High Schools. *Consortium on Chicago School Research*. Retrieved from http://files.eric.ed.gov/fulltext/ED498350.pdf

Allensworth, E., Gwynne, J., Moore, P., & de la Torre, M. (2014). Looking forward to high school and college: Middle grade indicators of readiness in Chicago public schools. *The University of Chicago Consortium on Chicago School Research*, (November), 1–138. Retrieved from http://ccsr.uchicago.edu/publications/looking-forward-high-school-and-college-middle-grade-indicators-readiness-chicago

Alliance for Excellent Education, Everyone Graduates Center, Civic Enterprises, A. P. A. (2015). *Progress is no accident: Why ESEA can't backtrack on high school graduation rates*. Retrieved from http://all4ed.org/wp-content/uploads/2015/11/NoAccident.pdf

Alliance for Excellent Education. (2011). *The high cost of high school dropouts: What the nation pays for inadequate high schools*. Retrieved from http://www.all4ed.org/files/HighCost.pdf

Amos, J. (2008). Dropouts, diplomas, and dollars: U.S. high schools and the nation's economy. *Alliance for Excellent Education*. Retrieved from http://all4ed.org/wp-content/uploads/2008/08/Econ2008.pdf

Ary, D., Jacobs, L., Sorensen, C., & Walker, D. (2014). *Introduction to Research in Education* (9th ed.). Belmont, CA: Wadsworth.

Balfanz, R. (2009). Putting middle grades students on the graduation path: A policy and practice brief. *National Middle School Association*. Retrieved from file:///Users/mbienkowski/Dropbox/zPapers/Library.papers3/Articles/2009/Balfanz/2009 Balfanz.pdf%5Cnpapers3://publication/uuid/E7AC62A6-F5F0-4EBD-81C3-565AE9C35E8A

Balfanz, R. (2011). Back on track to graduate. *Educational Leadership*, *68*(7). Retrieved from http://content.ebscohost.com.cucproxy.cuchicago.edu/ContentServer.asp?T=P&P=AN&K=59832717&S=R&D=f5h&EbscoContent=dGJyMMTo50Sep7I4zdnyOLCmr0mep7VSsK64TLGWxWXS&ContentCustomer=dGJyMPGuskm3qLFRuePfgeyx44Dt6fIA%5Cnhttp://cucproxy.cuchicago.edu/login?url=http:

Balfanz, R. (2016). Missing school matters. *Kappan Magazine*, (October), 8–13. https://doi.org/10.1177/0031721716671898

Balfanz, R., Bridgeland, J., Bruce, M., & Fox, J. (2012). *Building a Grad Nation*. Retrieved from https://files.eric.ed.gov/fulltext/ED530320.pdf

Balfanz, R., Byrnes, V., & Fox, J. (2014). Sent home and put off-track: The antecedents, disproportionalities, and consequences of being suspended in the ninth grade. *Journal of Applied Research on Children: Informing Policy for Children at Risk*, *5*(2), 1–19. Retrieved from http://digitalcommons.library.tmc.edu/childrenatrisk%5Cnhttp://digitalcommons.library.tmc.edu/childrenatrisk

Balfanz, R., & Fox, J. (2011). Early warning systems: Foundational research and lessons from the field. In *National Governors Association*. Retrieved from https://docplayer.net/15154851-Early-warning-systems-foundational-research-and-lessons-from-the-field.html

Balfanz, R., Herzog, L., & Mac Iver, D. (2007). Preventing student disengagement and keeping students on the graduation path in urban middle-grades schools: Early identification and effective interventions. *Educational Psychologist*, *42*(4), 223–235. https://doi.org/10.1080/00461520701621079

Baltimore Education Research Consortium. (2011). *Destination graduation: Sixth grade early warning indicators for Baltimore City schools*. Retrieved from http://www.baltimore-berc.org/pdfs/SixthGradeEWIFullReport.pdf

Bartishevich, C., Fogg, N., Motroni, J., Palma, S., Pond, N., & Harrington, P. (2003). *The hidden crisis in the high school dropout problems of young adults in the U.S.: Recent trends in overall school dropout rates and gender differences in dropout behavior*. Retrieved from https://files.eric.ed.gov/fulltext/ED479286.pdf

Bornsheuer, J., Polonyi, M., Andrews, M., Fore, B., & Onwuegbuzie, A. (2010). The relationship between ninth-grade retention and on-time graduation in a southeast Texas high school. *Journal of At-Risk Issues*, *16*(2), 9–16. Retrieved from http://eric.ed.gov/?id=EJ960072

Bowers, A., Sprott, R., & Taff, S. (2013). Do we know who will drop out?: A review of the predictors of dropping out of high school: Precision, sensitivity, and specificity. *The High School Journal*, *96*(2), 77–100. https://doi.org/10.1353/hsj.2013.0000

Boyle, T. (2019). Dealing with Imbalanced Data. Retrieved from https://towardsdata science.com/methods-for-dealing-with-imbalanced-data-5b761be45a18

Bridgeland, J., Dilulio, J., & Morison, K. (2006). The silent epidemic: Perspectives of high school dropouts. *Civic Enterprises*, (March), 1–35. Retrieved from http://eric. ed.gov/?id=ED513444

Brownlee, J. (2016). Linear Discriminant Analysis for Machine Learning. Retrieved from https://machinelearningmastery.com/linear-discriminant-analysis-for-machine-learning/

Bruce, M., Bridgeland, J., Fox, J., & Balfanz, R. (2011). *On track for success: The use of early warning indicator and intervention systems to build a grad nation*. Retrieved from http://files.eric.ed.gov/fulltext/ED526421.pdf

Brundage, A. (2014). *The use of early warning systems to promote success for all students*. Retrieved from http://www.fldoe.org/core/fileparse.php/5423/urlt/ews.pdf

Bureau of Labor Statistics, U. S. D. of L. (2019). Median weekly earnings $606 for high school dropouts, $1,559 for advanced degree holders. Retrieved August 25, 2020, from https://www.bls.gov/opub/ted/2019/median-weekly-earnings-606-for-high-school-dropouts-1559-for-advanced-degree-holders.htm

Burke, A. (2015). Early identification of high school graduation outcomes in Oregon Leadership Network schools. *Regional Educational Laboratory Northwest*, (April), 34. Retrieved from http://131.211.208.19/login?auth=eng&url=http://ovidsp.ovid.

com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=eric3&AN=ED556119;%5Cnhttp://sfx.library.uu.nl/utrecht?sid=OVID:ericdb&id=pmid:&id=doi:&issn=&isbn=&volume=&issue=&spage=&pages=&date=2015&title=

Christle, C., Jolivette, K., & Nelson, C. (2007). School characteristics related to high school dropout rates. *Remedial and Special Education*, *28*(6), 325–339. https://doi.org/10.1177/07419325070280060201

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences. In *Routledge Academic* (2nd ed.). https://doi.org/10.2307/2529115

Cohen, J., & Smerdon, B. (2009). Tightening the dropout tourniquet: Easing the transition from middle to high school. *Preventing School Failure: Alternative Education for Children and Youth*, *53*(3), 177–184. https://doi.org/10.3200/PSFL.53.3.177-184

Conduct and interpret a point-biserial correlation. (n.d.). Retrieved from http://www.statisticssolutions.com/point-biserial-correlation/

Crouser, R. (n.d.). LDA and QDA in Python. Retrieved from http://www.science.smith.edu/~jcrouser/SDS293/labs/lab5-py.html

Davis, M., Herzog, L., & Legters, N. (2013). Organizing schools to address early warning indicators (EWIs): Common practices and challenges. *Journal of Education for Students Placed at Risk*, *18*(1), 84–100. https://doi.org/10.1080/10824669.2013.745210

DePaoli, J., Balfanz, R., & Bridgeland, J. (2016). *Building a grad nation 2016: Progress and challenge in raising high school graduation rates*. Retrieved from https://files.eric.ed.gov/fulltext/ED572731.pdf

DePaoli, J., Fox, J., Ingram, E., Maushard, M., Bridgeland, J., & Balfanz, R. (2015). *Building a grad nation: Progress and challenge in ending the high school dropout epidemic*. Retrieved from https://files.eric.ed.gov/fulltext/ED530320.pdf

Dynarski, M., & Gleason, P. (2002). How can we help? What we have learned from recent federal dropout prevention evaluations. *Journal of Education for Students Placed at Risk*, *7*(1), 43–69. https://doi.org/10.1207/S15327671ESPR0701

Epstein, J., & Mac Iver, D. (1990). *Education in the middle grades: Overview of national practices and trends*. Columbus, Ohio: National Middle School Association.

Finn, J. (1989). Withdrawing from school. *Review of Educational Research*, *59*(2), 117–142. https://doi.org/10.3102/00346543059002117

Ford, S., Alukonis, D., Byrd-Bennett, B., Cohn, C., Dickson, S., Easton, J., … Smith, C. (2005a). Mathematics: The nation's report card. *National Assessment of Educational Progress*. Retrieved from https://www.nationsreportcard.gov/reading_math_2005/s0025.aspx?subtab_id=Tab_2&tab_id=tab2&printver=#chart

Ford, S., Alukonis, D., Byrd-Bennett, B., Cohn, C., Dickson, S., Easton, J., . . . Smith, C. (2005b). Reading: The nation's report card. *National Assessment of Educational Progress*. Retrieved from https://www.nationsreportcard.gov/reading_math_2005/s0015.aspx?subtab_id=Tab_2&tab_id=tab2&printver=#chart

Fortmann-Roe, S. (2012). Accurately measuring model prediction error. Retrieved from http://scott.fortmann-roe.com/docs/MeasuringError.html

Frick, M. (2018). *Georgia's graduation rate rises to all-time high*. Retrieved from http://www.gadoe.org/External-Affairs-and-Policy/communications/Pages/PressReleaseDetails.aspx?PressView=default&pid=644

Georgia Department of Education. (n.d.). College and Career Ready Performance Index. Retrieved from http://www.gadoe.org/CCRPI/Pages/default.aspx

Gwynne, J., Lesnick, J., Hart, H., & Allensworth, E. (2009). What matters for staying on-track and graduating in Chicago public high schools: A focus on students with disabilities. *Consortium on Chicago School Research at the University of Chicago*, (December). Retrieved from https://files.eric.ed.gov/fulltext/ED507419.pdf

Hammond, C., Linton, D., Smink, J., & Drew, S. (2007). *Dropout risk factors and exemplary programs: A technical report*. Retrieved from https://files.eric. ed.gov/fulltext/ED497057.pdf

Heppen, J., & Therriault, S. (2008). *Developing early warning systems to identify potential high school dropouts*. Retrieved from http://files.eric.ed.gov/fulltext/ ED521558.pdf

James, G., Witen, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning with Applications in R. *ScienceDirect,* 64(9-12), 856-875. https://doi.org/ 10.1016/j.peva. 2007.06.006

Jerald, C. (2006). Identifying potential dropouts: Key lessons for building an early warning data system. *Achieve, Inc.* Retrieved from http://www.achieve.org/ files/IdentifyingPotentialDropouts.pdf

Jerald, C. (2007). Keeping kids in school: What research tells us about preventing dropouts. *Center for Public Education*, 1–18. Retrieved from http://www. centerforpubliceducation.org/Main-Menu/Staffingstudents/Keeping-kids-in-school-At-a-glance/Keeping-kids-in-school-Preventing-dropouts.html

Jimerson, S., Anderson, G., & Whipple, A. (2002). *Winning the battle and losing the* *war: Examining the relation between grade retention and dropping out of high school*. *39*(4), 441–457. Retrieved from http://www.seg.guanajuato.gob.mx/Ceducativa/CartillaB/6antologia/antecedentes/pdf/35.- WINNING THE BATTLE AND LOSING THE WAR EXAMINING THE RELATION BETWEEN GRADE RETENTION AND DROPPING OUT OF HIGH SCHOOL.pdf

Jobs for the Future. (2014). *Early warning indicators and segmentation analysis: A technical guide on data studies that inform dropout prevention and recovery*. Retrieved from http://www.jff.org/sites/default/files/publications/materials/early warningindicators.pdf

Jones, C. (2016). *District high schools see gains in graduation rates: Howard High exceeds state graduation rate*. Retrieved from https://www.bcsdk12.net/cms/lib01/GA01000598/Centricity/domain/1/press releases/2016 Graduation Rates 110116.pdf

Kassambara, A. (2018a). Discriminant Analysis Essentials in R. Retrieved from http://www.sthda.com/english/articles/36-classification-methods-essentials/146-discriminant-analysis-essentials-in-r/

Kassambara, A. (2018b). Logistic Regression Assumptions and Diagnostics in R. Retrieved from http://www.sthda.com/english/articles/36-classification-methods-essentials/148-logistic-regression-assumptions-and-diagnostics-in-r/

Kemple, J. J., Segeritz, M. D., & Stephenson, N. (2013). Building On-Track Indicators for High School Graduation and College Readiness: Evidence from New York City. *Journal of Education for Students Placed at Risk*, *18*(1), 7–28. https://doi.org/10.1080/10824669.2013.747945

Kennelly, L., & Monrad, M. (2007). Approaches to dropout prevention: Heeding early warning signs with appropriate interventions. *American Institutes for Research*, 2014–2015. https://doi.org/10.1037/e538292012-001

Knowles, J. (2014). Of needles and haystacks: Building an accurate statewide dropout early warning system in Wisconsin. *Journal of Educational Data Mining*, *7*(3), 1–52. https://doi.org/doi:10.6084/m9.figshare.1142580

Knowles, J., & White, D. (2013). *Wisconsin dropout early warning system (DEWS)*. Retrieved from https://dpi.wi.gov/sites/default/files/imce/dews/pdf/DEWS Action Guide 2015.pdf

Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modeling*. New York: Springer. Retrieved from https://doi.org/10.1007/978-1-4614-6849-3

Laird, J., Kienzl, G., DeBell, M., & Chapman, C. (2007). Dropout rates in the United States: 2005. *U.S. Department of Education*. Retrieved from http://files.eric.ed.gov/fulltext/ED497226.pdf

Lee, T., Cornell, D., Gregory, A., & Fan, X. (2011). High suspension schools and dropout rates for black and white students. *Education and Treatment of Children*, *34*(2), 167–192. https://doi.org/10.1353/etc.2011.0014

Leech, N., Barrett, K., & Morgan, G. (2008). *SPSS for intermediate statistics: Use and interpretation* (3rd ed.). New York: Taylor & Francis Group, LLC.

Lemon, J. C., & Watson, J. C. (2011). An investigation of the relationship among wellness, perceived stress, mattering, and at-risk status for dropping out of high school. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, *71*(8-A), 2766. Retrieved from http://proxy.lib.sfu.ca/login?url=http://

search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2011-99030-371&site=ehost-live

Levin, H., Belfield, C., & Rouse, C. (2007). *The costs and benefits of an excellent education for all of America's children*. Retrieved from http://www.literacy cooperative.org/documents/Thecostsandbenefitsofanexcellentedforamerchildren.pdf

Linear & quadratic discriminant analysis. (n.d.). Retrieved from http://uc-r.github.io /discriminant_analysis

Mac Iver, M. (2010). Gradual disengagement: A portrait of the 2008-09 dropouts in the Baltimore City Schools. *Baltimore Education Research Consortium*, 1–16. Retrieved from http://acy.hs-cluster-1.net/upimages/Gradual_Disengagement.pdf

Mac Iver, M. (2013). Early warning indicators of high school outcomes. *Journal of Education for Students Placed at Risk (JESPAR)*, *18*(1), 1–6. https://doi.org/ 10.1080/10824669.2013.745375

Mac Iver, M., & Mac Iver, D. (2009). Beyond the indicators: An integrated school level approach to dropout prevention. *The George Washington University Center for Equity and Excellence in Education*. Retrieved from http://maec.ceee.gwu.edu

Mac Iver, M., & Mac Iver, D. (2010). *Keeping on track in ninth grade and beyond: Baltimore's ninth graders in 2007-08*. Retrieved from http://new.every1graduates. org/wp-content/uploads/2012/03/Keeping_On_Track_in_Ninth_Grade_and_Beyond .pdf

Mac Iver, M., & Messel, M. (2012). Predicting high school outcomes in the Baltimore City Public Schools. *The Senior Urban Education Research Fellowship Series*, *VII*(Summer 2012). Retrieved from http://baltimore-berc.org/pdfs/PredictingHigh

SchoolOutcomes.pdf

Mac Iver, M., & Messel, M. (2013). The ABCs of keeping on track to graduation: Research findings from Baltimore. *Journal of Education for Students Placed at Risk*, *18*(1), 50–67. https://doi.org/10.1080/10824669.2013.745207

McCallumore, K. M., & Sparapani, E. F. (2010). The importance of the ninth grade on high school graduation rates and student success in high school. *Education*, *130*(3), 447–456. Retrieved from https://login.proxy.library.msstate.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ917128&site=eds-live%5Cnhttp://www.projectinnovation.biz/education_2006.html

McIntosh, K., Flannery, K., Sugai, G., Braun, D., & Cochrane, K. (2008). Relationships between academics and problem behavior in the transition from middle school to high school. *Journal of Positive Behavior Interventions*, *10*(4), 243–255. https://doi.org/10.1177/1098300708318961

McKee, M., & Caldarella, P. (2016). Middle school predictors of high school performance: A case study of dropout risk indicators. *Education*, p. 515. Retrieved from http://eds.a.ebscohost.com/eds/pdfviewer/pdfviewer?vid=1&sid=44c5b4e5-47a6-4f39-aca7-b38c2d219f86%40sessionmgr4006

National Education Association. (2008). *English language learners face unique challenges*. Retrieved from http://www.nea.org/assets/docs/HE/ELL_Policy_Brief_Fall_08_(2).pdf

Neild, R. (2009). Falling off track during the transition to high school: What we know and what can be done. *Future of Children*, *19*(1), 53–76. https://doi.org/10.1353 foc.0.0020

Neild, R., & Balfanz, R. (2006). Unfulfilled promise: The dimensions and characteristics of Philadelphia's dropout crisis, 2000–2005. *Finance and Development*, *40*(1), 1–49. Retrieved from http://www.schoolturnaroundsupport.org/sites/default/files/resources /Neild_Balfanz_06.pdf

Neild, R., Balfanz, R., & Herzog, L. (2007). An early warning system. *Educational Leadership*, *65*(2), 28–33. Retrieved from http://new.every1graduates.org/wp-content/uploads/2012/03/Early_Warning_System_Neild_Balfanz_Herzog.pdf

Neild, R., Stoner-Eby, S., & Furstenberg, F. (2008). Connecting entrance and departure: The transition to ninth grade and high school dropout. *Education and Urban Society*, *40*(5), 543–569. https://doi.org/10.1177/0013124508316438

Pinkus, L. (2008). Using early-warning data to improve graduation rates: Closing cracks in the education system. *Washington, DC: Alliance for Excellent Education*, (August). Retrieved from http://beta.fresnounified.org/gradtf/Shared Documents/Using Early Warning Data to Improve Graduation Rates, Closing Cracks in the Education System.pdf

Prabhakaran, S. (2017). Logistic Regression – A Complete Tutorial With Examples in R. Retrieved from https://www. machinelearningplus.com/machine-learning/logistic-regression-tutorial-examples-r/

Public high school graduation rates. (2018). Retrieved from https://nces.ed.gov/programs /coe/indicator _coi.asp

Reschly, A., & Christenson, S. (2006). Prediction of dropout among students with mild disabilities: A case for the inclusion of student engagement variables. *Remedial and Special Education*, *27*(5), 276–292. https://doi.org/10.1177/07419325060270050301

Robst, J., & de Vries, J. (2010). Childhood behavioral problems and dropping out of school. *Eastern Economic Journal*, *36*(4), 523–538. https://doi.org/10.1057/eej.2009.40

Roderick, M. (1993). *The path to dropping out*. Westport, CT: Auburn House.

Rumberger, R. (2004). Why students drop out of school. In *Dropouts in America: Confronting the graduation rate crisis* (pp. 131–155). Cambridge: Harvard Education Press.

Rumberger, R., & Larson, K. (2007). Student mobility and the increased risk of high school dropout. *American Journal of Education*, *107*(1), 1–35. https://doi.org/10.1086/444201

Rumberger, R., & Lim, S. (2008). *Why students drop out of school: A review of 25 years of research*. Retrieved from http://www.cdrp.ucsb.edu/pubs_reports.htm

Schoeneberger, J. A. (2012). Longitudinal attendance patterns: Developing high school dropouts. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 85(1),* 7-14. Retrieved from https://doi.org/10.1080/00098655.2011.603766

Sebring, P., Bryk, A., Roderick, M., Camburn, E., Luppescu, S., Thum, Y., Smith, B., & Kahne, J. (1996). *Charting reform in Chicago: The students speak*. Retrieved from https://consortium.uchicago.edu/sites/default/files/publications/ChartingReformInChicago_TheStudentsSpeak.pdf

Silver, D., Saunders, M., & Zarate, E. (2008). *What factors predict high school graduation in the Los Angeles Unified School District*. Retrieved from www.lmri.ucsb.edu/dropouts

Sparks, E., Johnson, J., & Akos, P. (2010). Dropouts: Finding the needles in the haystack.

Educational Leadership, *67*(5), 46–49. Retrieved from http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=47791473&site=ehost-live&scope=site

Stanard, R. (2003). High school graduation rates in the United States: Implications for the counseling profession. *Journal of Counseling and Development*, *81*, 217–221. Retrieved from http://eds.a.ebscohost.com/eds/pdfviewer/pdfviewer?vid=1&sid=ef60e59e-abc1-4b6a-a2b4-30a829cf52d2%40sessionmgr4009

Stark, P., Noel, A., & McFarland, J. (2015). Trends in high school dropout and completion rates in the United States: 1972–2012. *National Center for Education Statistics*, 1–12. Retrieved from https://files.eric.ed.gov/fulltext/ED557576.pdf

Stuit, D., O'Cummings, M., Norbury, H., Heppen, J., Dhillon, S., Lindsay, J., & Zhu, B. (2016). *Identifying early warning indicators in three Ohio school districts*. Retrieved from http://ies.ed.gov/ncee/edlabs

Suh, S., & Suh, J. (2004). *Risk factors and levels of risk for high school dropouts*, *10(3)*, 297–306. Retrieved from https://www.jstor.org/stable/42732523?readnow=1&googleloggedin=true&seq=1#metadata_info_tab_contents

Sum, A., Khatiwada, I., & McLaughlin, J. (2009). *The consequences of dropping out of high school: Joblessness and jailing for high school dropouts and the high cost for taxpayers*. Retrieved from https://www.prisonpolicy.org/scans/The_Consequences_of_Dropping_Out_of_High_School.pdf

The Governor's Office of Student Achievement: Downloadable data. (2018). Retrieved from https://gosa.georgia.gov/downloadable-data

The High School Pipeline. (2006). *Education Week*, *25*(41S), 16. Retrieved from https://

www.edweek.org/ew/articles/2006/06/22/41s_patterns-pipeline.h25.html?sec=freshman

Tucci, T. (2011). *Education and the economy: Boosting state and local economies by improving high school graduation rates*. Retrieved from https://34.231.97.227/wp-content/uploads/EdEconomy_seb_leb.pdf

U.S. Department of Education. (2016a). Graduation rates by state. *Education Week*, *35*(33). Retrieved from https://www.edweek.org/ew/dc/2016/map-graduation-rates-by-state-demographics.html

U.S. Department of Education. (2016b). *Issue brief: Early warning systems*. Retrieved from http://ies.ed.gov/ncee/edlabs/projects/ews.asp.

UCLA: Statistical Consulting Group. (n.d.). Introduction to SAS. Retrieved May 1, 2020, from https://stats.idre.ucla.edu/other/mult-pkg/faq/general/faq-what-are-pseudo-r-squareds/

Wisconsin Department of Public Instruction. (2015). *DEWS data brief*. Retrieved from https://dpi.wi.gov/sites/default/files/imce/dews/pdf/DEWS Data Brief 2015.pdf

Zvoch, K. (2006). Freshman year dropouts: Interactions between student and school characteristics and student dropout status. *Journal of Education for Students Placed at Risk*, *11*(1), 97–117. https://doi.org/10.1207/s15327671espr1101

APPENDIX A:

Institutional Review Board Protocol Exemption Report

Institutional Review Board (IRB)

For the Protection of Human Research Participants

Protocol Number:	03751-2018	**Investigator:**	Dana Morris
		Supervising Faculty:	Dr. Lantry Brockmeier
PROJECT TITLE:	*Using Sixth Grade and Ninth Grade Indicators within an Early Warning System to Predict Students at Risk for Dropping Out.*		

INSTITUTIONAL REVIEW BOARD DETERMINATION:

This research protocol is **Exempt** from Institutional Review Board (IRB) oversight under Exemption **Category 4**. Your research study may begin immediately. If the nature of the research project changes such that exemption criteria may no longer apply, please consult with the IRB Administrator (irb@valdosta.edu) before continuing your research.

ADDITIONAL COMMENTS:

- *Upon completion of the research study all data (transcripts, data lists, etc.) must be securely maintained (locked file cabinet, password protected computer, etc.) for a minimum of 3 years and only accessible by the researcher.*

☒ ***If this box is checked, please submit any documents you revise to the IRB Administrator at irb@valdosta.edu to ensure an updated record of your exemption.***

Elizabeth Ann Olphie 12.05.2018
Elizabeth Ann Olphie, IRB Administrator

Thank you for submitting an IRB application.
Please direct questions to irb@valdosta.edu or 229-253-2947.

Revised: 06.02.16