

College Student Success:
Using Predictive Modeling and Actionable Intelligence with a Faculty Centered
Information Portal to Improve Student Academic Performance

A Dissertation submitted
to the Graduate School
Valdosta State University

in partial fulfillment of requirements
for the degree of

DOCTOR OF PUBLIC ADMINISTRATION

in Public Administration

in the Department of Political Science
of the College of Arts and Sciences

December 2016

Andy T. Clark

MSAD: Logistics Management, Georgia College and State University, 2007
MPA, Georgia College and State University, 1999
BS, Georgia College and State University, 1998

© Copyright 2016 Andy T. Clark

All Rights Reserved

This dissertation, “College Student Success: Using Predictive Modeling and Actionable Intelligence with a Faculty Centered Information Portal to Improve Student Academic Performance,” by Andy T. Clark, is approved by:

**Dissertation
Committee
Chair**

James T. LaPlant, Ph.D.
Professor of Political Science

**Committee
Members**

Michael P. Savoie, Ed.D.
Professor of Mass Media

Robert (Sherman) Yehl, Ph.D.
Assistant Professor of Political Science

**Dean of the College
of Arts and Sciences**

Connie L. Richards, Ph. D.
Professor of English

**Dean of the
Graduate School**

James T. LaPlant, Ph.D.
Professor of Political Science

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 _____

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

Signature _____

ABSTRACT

This research project examines new models and approaches to student learning and success by concentrating on the first-year experience of beginning freshmen at Valdosta State University utilizing data from 2008-2014. With a fall freshman class ranging from 1,500 to 2,500 new students, the sample size is large enough to produce a much smaller confidence interval/sampling error, yet small enough to work with individual departments and faculty to implement and monitor the effect of changes employed through the use of predictive metrics and active intervention. The predictive metrics developed for this model use three specific indicators: (1) standardized test scores from the SAT or ACT, (2) high school grade point average and (3) where the student's high school ranks in relation to the other high schools in the state of Georgia. The purpose of this research is to develop and defend the answer in response to the research question: Can predictive modeling be used to create actionable student intelligence to improve the grades in key English and math classes resulting in higher retention rates of traditional first-year students? The findings from this research demonstrate that predictive modeling can be very effective in identifying at-risk student populations. These models provide timely insight into students' needs for additional support to be successful academically. There were five important clusters of results: (1) the pass/fail rates based upon the 1-4 rankings for high school rank, GPA, and SAT, with these data points proving to be very useful in predicting DFW rates, (2) the multivariate regression analysis also showed that these variables are statistically significant, (3) for math the difference of means test for the changes over time once the placement index was put in place improved the pass rate in math courses, (4) the analysis of financial grouping

and employment index showed that these variables also impact student success, (5) student success improved with faculty that utilized the portal vs. faculty that did not utilize the portal. This research is very closely aligned with the “Complete College America” movement.

TABLE OF CONTENTS

Chapter I: STATE OF HIGHER EDUCATION	1
Background	1
Scope of Study	3
Significance of Problem.....	4
Research Question	7
Overview of the Dissertation	8
Chapter II: A STUDY OF COLLEGE STUDENT SUCCESS	11
The Impact of Financial Aid	25
Standardized Indicators	29
Chapter III: METHODOLOGY	39
Predictive Modeling.....	39
Faculty Portal	49
Chapter IV: RESULTS	55
Introduction.....	55
Predictive Modeling.....	55
Taking Math One Step Further	72
Financial Aid Predictive Models	84
Faculty Portal, Early Alert, and Interventions	88
Chapter V: PREDICTIVE ANALYTICS AND THE FUTURE OF STUDENT SUCCESS	92
Research Question and Key Findings	92
Lessons Learned.....	101

High-Impact Practices.....	102
Additional Avenues of Research and Study	107
Cautionary Warning and Conclusion.....	109
REFERENCES	111
APPENDIX A: Institutional Review Board Exemption.....	117
APPENDIX B: Complete Data Sheet.....	119

LIST OF TABLES

Table 1: Quartiles of Student Preparation.....42

Table 2: Example of Report Card Metrics44

Table 3: Example of High School Rank44

Table 4: Additional Variables with Honors and AP45

Table 5: High School Graduation Test Results.....45

Table 6: High School Rank.....45

Table 7: Ranges for Quartiles High School GPA and SAT Math47

Table 8: How Quartiles Combine to Indicate Math Placement48

Table 9: Descriptive Statistics for Models 1 through 357

Table 10: High School Rank of 1 and Performance in English-based Courses.....58

Table 11: High School Rank of 2 and Performance in English-based Courses.....59

Table 12: High School Rank of 3 and Performance in English-based Courses.....60

Table 13: High School Rank of 4 and Performance in English-based Courses.....61

Table 14: No High School Rank (9-X-X) and Pass Rate for English-based Courses62

Table 15: High School Rank of 1 and Performance in Math-based Courses63

Table 16: High School Rank of 2 and Performance in Math-based Courses64

Table 17: High School Rank of 3 and Performance in Math-based Courses65

Table 18: High School Rank of 4 and Performance in Math-based Courses66

Table 19: High School Rank of 3 Comparing Verbal Test Scores to Math-based
Courses.....67

Table 20: Regression Model for Predictive Analysis69

Table 21: Pass Rates of All Math 1111 Courses, Fall 2005 to Fall 2014.....75

Table 22: Pass Rates of All MATH 1111 Extended Learning Students, Fall 2005 to Fall 2014	77
Table 23: Pass Rates of MATH 1111 Students without MATH 1111 Extended Learning Students, Fall 2005 to Fall 2014.....	79
Table 24: Pass Rates of Math Level One High Risk Students.....	81
Table 25: Pass Rates of Math Level One High Risk Students by Extended Learning Placement.....	83
Table 26: Retention Rate of First-time, Full-time Freshman by Financial Indicators, Fall 2007-Fall 2012 Cohorts	85
Table 27: Retention Rates by Academic and Financial Groupings	87
Table 28: Crosstabulation of Pass Rates by Faculty Page Views.....	89
Table 29: Crosstabulation of Pass Rates by Faculty Flag Set.....	90
Table 30: Retention of Fall 2012 Cohort by Portal Users	91
Table 31: Overview of QEP Project Participation, Presentations and Publications.....	103

LIST OF FIGURES

Figure 1: Mathematics Courses with Prerequisite Flow	47
Figure 2: View of Faculty Portal Login Screen	50
Figure 3: Faculty Course Roster	51
Figure 4: Early Alert Window	52
Figure 5: Pass Rates of All MATH 1111 Courses, Fall 2005-Fall 2014	74
Figure 6: Pass Rate of All MATH 1111 Extended Learning Students, Fall 2005-Fall 2014.....	76
Figure 7: Pass Rate of MATH 1111 Students without MATH 1111 Extended Learning Students	78
Figure 8: Pass Rate of Math Level One High-Risk Students.....	80
Figure 9: Pass Rate of Math Level One High-Risk Students by Extended Learning Placement.....	82
Figure 10: Retention Rates of Financial Groupings	86
Figure 11: Employment Index Retention Rates.....	87
Figure 12: Graduation Rates, Retention, Rates, and the GPAs of QEP Students.....	107

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my committee chair Dr. James LaPlant for the continuous support of my study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in the research and writing of this dissertation. I would like to also thank the other members of my committee: Dr. Michael Savoie and Dr. Sherman Yehl for their insights and questions that helped to improve this research. I would like to thank Barrie Fitzgerald in the office of Institutional Research at Valdosta State University for his help with the data used in this dissertation. I would also like to thank Dr. Joseph Weaver for his time and help with editing this research. I would like to thank Brian Haugabrook and the staff in Information Technology for the development of the Faculty Portal. To Dr. Greg Harrell and the faculty in the Department of Mathematics for their extensive work on improving student success in math courses.

DEDICATION

To my Mother and Father who always pushed me to finish the things I started. To all aspiring students who arrive in the hallowed halls of the university, that they receive the support they need to be successful. To the faculty and staff who work hard to provide and support those students in their journey.

Chapter I

STATE OF HIGHER EDUCATION

Background

Retention...Progression...Graduation. Retention...Progression...Graduation. Retention...Progression...Graduation. Throw 'Persistence' into the mix and what emerges is the mantra for those charged with enrollment management for institutions of higher learning. From its historical position of dominance as the world's leader in college completion, the United States has dropped dramatically in the 21st century. With a 10.4 college completion percentage for adults aged 25-64, the United States ranked 17th among the countries reported by the Organization for Economic Cooperation and Development for 2012 (NCES 2014, 363). Time to degree completion must also be factored into the equation. According to Complete College America, only "5% of full-time students pursuing associate degrees graduate on time" at public 2-year institutions; while the 4-year graduation rate for public institutions at non-flagship or research universities stands at 19% for full-time students (CCA 2013, 4). While the cost to those who eventually complete their degrees on a delayed basis is substantial through additional tuition, fees and lost income, the burden for those who do not complete their degrees is even worse. Often they are saddled with large student loan debt and diminished job prospects without degree in hand.

In part, these are some of the complex challenges facing institutions of higher

learning today. Because these issues affect the financial well-being of the citizens and our ability as a nation to produce an educated workforce to compete in the global market, the situation has not gone unnoticed by leaders of our state and national governments. Dissatisfaction with the status quo percolates from the chambers of the United States House and Senate and reverberates down through state and local government. Parents facing increasing costs to fund their children's education, students saddled with burdensome student loan debt that has to be repaid regardless of degree completion, employers unable to fill positions in key fields, and the precipitous fall from educational grace have thrust educational accountability into the local and national limelight. Restrictions on tuition increases to control costs to the students, mandated access to "consumer data" for prospective students to make informed educational choices, proposals to shift to performance-based funding for state colleges and universities are examples of specific attempts designed to encourage or to force colleges and universities to address the problems at the local level. The Commission on the Future of Higher Education formed in 2005 also known as the Spellings Commission brought much needed attention to the issues facing higher education, but the problems persist. A decade later, institutions of higher learning and their constituents are still grappling with the same problems regarding access, affordability, accountability and quality. If anything, the political voices have become more strident and the pressure for solutions even more pressing.

These concerns are also reflected in the mandate for Complete College America and its localized instantiations such as Complete College Georgia whose missions are to improve college completion rates. While studies and research that focus on the results

are very useful in defining the parameters of the problems, they fall short of identifying the solutions. Attempts to understand problematic graduation rates, by looking back at the causes, need to be supplemented by research done preventatively at the front end. By identifying target populations who are at-risk at the very beginning of their educational careers, it may be possible to develop and implement strategies that will enable the students to be more successful; thereby, retaining, progressing and graduating more students—the ultimate reward for persistence.

Scope of Study

This research project will examine new models and approaches to student learning and success by concentrating on the first-year experience of beginning freshmen at Valdosta State University utilizing data from 2008-2014. With a fall freshman class ranging from 1,500 to 2,500 new students, the sample size is large enough to produce a much smaller confidence interval/sampling error, yet small enough to work with individual departments and faculty to implement and monitor the effect of changes employed through the use of predictive metrics and active intervention. The predictive metrics developed for this model use three specific indicators: (1) standardized test scores from the SAT or ACT, (2) high school grade point average, and (3) where the student's high school ranks in relation to the other high schools in the state of Georgia.

The predictive metrics also known as Actionable Student Intelligence (ASI) are not ends-in-themselves but simply tools for creating intervention strategies to reach high-risk students. High-risk students are defined as a student with a high probability to struggle in academic course work or with persistence to degree completion. While certain high-risk populations are readily identifiable, others are not. It is the latter group that has

been underserved by retention efforts and in need of additional research efforts. All too often, by the time the institution becomes aware of these students' struggles, the opportunity to assist them has passed. The goal of the ASI predictive metrics model at the heart of this project is to uncover the hidden high-risk populations at a much earlier point in time, at a time that will enable the school to reach out to them when help is most useful.

Significance of Problem

The importance of projects such as this should not be underestimated since they are vital to the success of the mission of the university to educate students and to prepare them to become productive members of society. One element of the vitality of an institution is its financial well-being and the impact of a low retention rate is dramatic. In Fall 2010, Valdosta State University enrolled 2,550 first-time, full-time freshmen. By the start of the Fall 2011 term, 842 of those students were no longer at the university. The 1,708 students who returned represent a 67% retention rate, well below the retention rate (average 76%) of other comprehensive public universities of similar size. The estimated \$7 million per year in lost fees and tuition equate to over 6% of the total budget for the university in 2010. Moreover, 2010 was not an anomalous year with the retention rate holding steady at 67% retention for 2011 and 2012 as well. Even without adjusting for increases in fees and tuition, and assuming a 4-year degree completion timeline, the best case scenario shows a minimum of \$28 million in foregone revenue. This is \$28 million that the university did not have to support educational programs or service debt obligations. In college towns such as Valdosta, when the university suffers, so too does the community at large. The 842 non-returning students from 2010 are not there to

patronize local businesses and services. In addition, since tuition and fees are only a fraction of what students spend during the course of the academic year, the economic impact on the community is much greater than the \$28 million lost to the university over a 4-year span.

But money is not the sole determinant of the negative effects of low retention rates. While certainly more difficult to measure, experience indicates that losing one-third of the freshmen class each year would affect the social and psychological dynamic of those who remain. Friendships established are interrupted, if not ended. Those who remain may question their decision to stay in light of the large number not returning. Presumably, the impact on the non-returning students would be even greater. At the very least feelings of intellectual inadequacy and related emotional trauma from unsuccessful goal attainment would be experienced by many in this group.

Achieving gains in the retention percentage with its corollary impact on progression and graduation rates is a process as opposed to an event. As stated earlier, this process begins by identifying at-risk populations. Valdosta State University is not an open enrollment institution and, as a general rule, does not admit traditional freshmen students who require remediation based on standardized test scores. Valdosta State University's acceptance rate averages around 68% of applicants. Therefore, those who would be readily identifiable as at-risk students in open enrollment colleges are diffused within the freshman class as a whole at VSU.

Like many colleges and universities, VSU utilizes several key criteria in its admissions criteria whose purpose now, as it has been historically, is to assure that those students it does admit have a reasonable expectation for academic success. In Fall 2010

for traditional freshmen (those with less than 30 transferable credit hours and a high school graduation date within 5 years), VSU required those admitted to (1) have graduated from a college preparatory program, (2) achieved minimum SAT or ACT scores, and (3) earned a high school GPA sufficient to support a Freshman Index of at least 2040. The minimum scores for the SAT were 440 in Critical Reading and 410 in Mathematics (VSU Catalog 2012). For the ACT, the minimums were 18 in English and 17 in Mathematics. Prospective students who scored below the minimums were inadmissible; those who score at or above the minimums and reached the requisite Freshman Index score were exempt from remediation classes. The Freshman Index (FI) is computed by use of one of the following formulae: $FI = \text{Total English and Math SAT score} + (500 \times \text{high school GPA})$ or $FI = (42 \times \text{composite ACT score}) + 88 + (500 \times \text{high school GPA})$. Since VSU utilizes minimum standardized test scores, a prospective student with SAT scores of 440 Critical Reading and 410 Math would need a high school GPA of approximately 2.40. However, because VSU does not require a minimum high school GPA, a high standardized test score can offset a low GPA. For example, a combined SAT score of 1,400 with a GPA of only 1.28 would yield the needed 2,040 FI for admission.

Prospective college students vary widely in their preparation for college-level coursework. Unfortunately, the use of quantitative measures such as standardized tests scores and high school GPAs to gauge college readiness is not as accurate or reliable as it has been in the past. Underprepared students are clearly a major factor in the low retention rate at VSU. Valdosta State University's experience is simply a reflection of a national trend of admitting underprepared students to colleges and universities—students

who have taken college preparatory classes, and meet minimum GPA and/or standardized test scores. In 2010, a report from the California State University (CSU) system notes that a major culprit in low graduation rates is the lack of readiness for college. Further, the report shows,

The California State University (CSU), a large public university system, for many years has applied placement or readiness standards in reading, writing, and mathematics that are linked to first-year college coursework. All first-time students at all 23 CSU campuses must meet these standards, principally through performance on a common statewide placement examination. Despite system wide admissions policy that requires a college-preparatory curriculum and a grade point average in high school of B or higher, 68% of the 50,000 entering freshmen at CSU campuses require remediation in English language arts, or math, or both (NCPPE 2010, 2).

As one can imagine, the costs associated with students entering public colleges and universities without the requisite preparatory knowledge and skills are staggering to the students, to the schools, to the states and to the nation as a whole. While it is beyond the scope of this project to address the underlying problem endemic in the secondary education system, those problems can be partially mitigated by the early identification and intervention strategies for students whose lack of college preparation is masked in the admissions process.

Research Question

The purpose of this research is to develop and defend the answer in response to the research question: Can predictive modeling be used to create actionable student intelligence to improve the grades in key English and math classes resulting in higher retention rates of traditional first-year students?

The first part of the question deals with the development of a model that predicts grades of first-year students in key English and math classes. The model uses

standardized test scores as well as the rank of the high school as assigned by the state of Georgia rating system. The control group for this part of the study are traditional freshmen who enrolled at VSU from 2008-2011. This baseline population will determine the efficacy of the model in predicting gateway course grades.

The second part of the question involves the development of actionable intelligence to improve student performance in the English and math courses with the expectation that grade improvement will increase student retention. This is a dynamic exercise that involves a host of diverse actors on campus including admissions, advising, tutoring, student affairs and faculty members. The design will allow multiple entry points for the input of data with the information from the data flowing to other affected areas. Each area would then have the opportunity and ability to respond on an as-needed basis to intervene with the student as appropriate. The goal is to enable information to be collected and exchanged in such a way as to allow potential student problems to be addressed before they become unsolvable.

Overview of the Dissertation

The research project is presented in five chapters. Chapter 2 provides a comprehensive review of the existing literature that relates to statistical methods for predicting grades and a review of intervention strategies that have proven to be successful. The review is organized into three broad sections. Section I presents an overview of the use of predictive modeling in relation to retention and graduation rates. The overview contains both an historical perspective as well as contemporary studies. Section II concentrates on the subject of financial aid and its implications on retention and graduation. Section III reviews the use of standardized indicators as predictors of

grades and college success rates in key subjects such as English and mathematics.

Chapter 3 presents the development of the statistical methodology to be used to predict grades in key English and math courses. The model will assign an index score to each individual student based on the criteria outlined earlier with the scores grouped and distributed to the appropriate faculty members. The methodology also involves a detailed description of intervention strategies, the trigger points used to alert the various departments and how the results of the intervention(s) are inserted into the data flow to refine the process.

In Chapter 4 the data generated by the model and the intervention strategies that have been employed will be analyzed and reviewed over a 2-year period, 2013 and 2014. The results will be compared to course grades for students in classes whose faculty are not using the intervention strategies to determine if student success has improved at a statistically significant level. Enhancements to the predictive model and to the intervention strategies will be explained and justified. The final part of this chapter will present changes, if any, to the retention rate of the students who first matriculated in 2011.

Chapter 5 will demonstrate the broader use of predictive analytics and its pertinence to higher education in general and their relation to the public and political environment. Connections to national and state level initiatives will also be examined. The chapter will explore lessons learned from the research and also the role of high-impact practices on campuses. The chapter will also caution campuses that may be considering outsourcing analytics to proprietary companies. This chapter will conclude

with a discussion of how this project can contribute to the promotion of continued research in the discipline.

Chapter II

A STUDY OF COLLEGE STUDENT SUCCESS

For as long as students have pursued formal post-secondary degrees in the United States, invariably some percentage of the student population has failed to complete the requirements resulting in the creation of drop-out and retention statistics. However, as long as access to colleges and universities was basically restricted to the mostly affluent along with a small contingent of the poor but obviously intellectually gifted, the reasons for non-degree completion were limited in number and readily identifiable—not the kind of subject matter that intrigues academic researchers. The post-World War II era opened the floodgates to colleges and universities across the U.S. and changed the paradigm entirely. At the turn of the 20th century, a mere 4% of high school graduates went on to attend a college or university (Jacobsen 2007). By the early 21st century, a study by the U.S. Department of Education (2014) of 15,000 high school graduates showed that by the age of 27, 84% had some college education (as cited in Lauff 2014).

Consequently, going to college is no longer a rite of passage of the upper class or an entry into another world for the very smart, it is an expectation that applies to virtually every high school student in the U.S. With the massive increase in sheer numbers of those attending college by removing the financial and intellectual constraints, the problems with retention, progression and graduation have multiplied exponentially along with the attendant studies undertaken in attempts to understand and resolve the

problems. Because of the breadth and depth of the research in the field, it is most difficult to survey all the relevant literature. The ensuing literature review will focus on three broad areas: (1) predictive modeling and its relation to retention and graduation rates both historical and current, (2) financial aid and its implications on retention and graduation, and (3) use of standardized indicators such as national tests, GPA and level of high school math courses as predictors of college success as measured by grades and program completion.

Since all three sections deal with some aspect of retention and graduation, the items themselves are not categorical and there are some unavoidable overlaps. It is anticipated that the common theme in all three sections of the literature review supports the contention that by using predictive analytics, at-risk modeling, and financial indicators, it may be possible to identify future at-risk students, predict potential retention rates, modify student behavior and employ intervention strategies. The result, in theory, would be an increase in future retention and graduation rates, as well as a model to locate these at-risk students early in their careers and aid them in various academic and non-academic pathways.

Predictive Modeling

Jacobson (2007) explains that when it was first developed in 1926 by Carl Bringham, the Scholastic Aptitude Test (SAT) was intended to be a supplement to, not a replacement for, the traditional College Board entrance exams. In 1934, Harvard University became the first school to use the SAT to evaluate candidates for scholarships. By 1938 this purpose was adopted by virtually all the private and public universities in the Northeastern part of the United States. It was not until 1941 that the SAT replaced

the original College Board exams as the standard admissions test for the Northeastern schools. With the return of soldiers from World War II and the passage of the GI Bill, the use of the SAT as a tool for admissions increased exponentially as it was adopted by schools outside of the Northeast as part of their admission process. As an admissions requirement, the SAT becomes a de facto predictor of academic success since admission itself is predicated on obtaining a certain score proficiency (Jacobson 2007).

Jacobson (2007) explains in 1959 the American College Testing Program (ACT) enters the national scene. It was designed to address perceived limitations in the SAT and to appeal to a wider number of schools. As an achievement based test, the ACT was intended to demonstrate academic preparation that prospective students had received in high school. Hence, it could be used both in the admissions process as well as a placement tool to assign students to the appropriate course level. Growth for the ACT paralleled that of the SAT. As of 2012 over 3.3 million persons took one of the two tests (Jacobson 2007).

Beginning with the enrolling class of 1964 the College Board established the validity study service. This service allowed colleges to send data for an analysis of the predictive validity of college admissions measures. This method used the high school record (course grades) and the scores of the SAT verbal and SAT mathematical test to predict freshmen GPAs. These average correlations are the estimates of predictive validity for each enrolling institution (Morgan 1990). However, there are problems with drawing conclusions about these trends and predictive validity: the first has to do with colleges voluntarily participating, meaning sample size or population is not always consistent. Another potential issue lies with the characteristics of the student population

being measured. College admissions practices change, and it is not unusual for schools to increase or decrease SAT and ACT minimum score requirements as a way to regulate the size of the entering freshmen class. One way to overcome these drawbacks is to properly categorize the different types of institutions and ensure that institutions submit data each year over an extended period of time. The data in the study conducted by Goldman and Widawski suggests that smaller, private and more selective colleges see smaller changes in the predictive validity of these measures with regard to freshmen grade point average than do larger and less selective colleges and universities (Morgan 1990). According to Pomplun et al., in the post-Vietnam late 1970s the correlation between SAT scores and GPAs began to diverge. Both SAT scores and GPAs showed declines. The average SAT scores decreased substantially but while freshmen GPAs also declined, the decrease was not at the same pace as the SAT. Whether this divergence was the result of grade inflation is still undetermined. Regardless, the predictive ability of the SAT was less reliable during this period (Pomplun, Burton, Lewis 1991).

This divergence may also be reflected in college grade patterns. Many of the longitudinal studies conducted using data collected by the National Center for Educational Statistics (NCES) focused in large part on grades. “Americans are fascinated by grades in education, almost as fascinated as by indicators of athletic performance in the sports section of the local newspaper or various market averages in the business section” (Adelmen 2004, 77). According to Adelmen’s research, the proportion of F grades declined slightly between the class of 1972 and the class of 1982 then rose again between 1982 and 1992. The more notable issue in changes of grading practices from the 1970s to the 1990s was the growing portion of withdrawal grades (Ws). While the W

grade is labeled as a non-penalty, there is actually a more subtle penalty reflected in the time to degree completion and the effects on financial aid. In highly selective institutions, the studies show a notable increase in the proportion of As at the expense of grades of B. At less selective institutions, there were far more grades of non-penalty or withdrawals. Withdrawals, failures and repeats are concentrated over this 30-year period of time in remedial courses, mathematics, lower division courses in the core or part of the foundation courses in the major such as history, chemistry, and introduction to business. Looking at the longitudinal studies and noting the patterns in where these D, F and W type grades are concentrated opens lines of research inquiry available for those who wish to add such institutional variables as admissions criteria to the equation (Adelmen 2004).

Not only did the student populations increase in sheer number during the latter half of the 20th century, but the student body became much more diverse, reflecting many more of the groups and sub-groups that make up the nation. These groups and sub-groups included those traditionally underrepresented in the higher education student population. Consequently, attempts to understand retention and graduation rates as well as predictive modeling moved beyond grades and standardized test scores to encompass other non-academic factors. Burgher and Davis (2013) conducted a study with the intention of using student behavior from proceeding spring semesters (2008, 2009, and 2010) in order to predict retention for the semester of spring 2011. Monitoring the individual student is difficult because there is little variation between their behaviors at a particular university. Using analyses to predict the behavior of a filtered group of students is far more effective than using it to predict behavior in individual students. Ergo, subsets of students were used in the data. The experiment consisted of two groups:

conditionally admitted and non-conditionally admitted students. Non-conditionally admitted students had a GPA over 2.5 and were automatically admitted to the university. Those with GPAs less than 2.5 were allowed admission conditionally, dependent upon a variety of other factors including ACT scores and admissions profiles. Selection bias was a strong possibility given that students with a GPA under 2.5 had to be reviewed for other criteria, subjecting the process to arbitrary outcomes.

The research determined and factored in an extensive list of possible variables. These variables included ethnicity, gender, high school GPA, ACT score, and whether or not the student was a child of an alumnus. A hybrid selection technique was used in which structural models were mixed with a statistical approach. The data model used the 2008-2010 academic years to establish the retention rates. Then, data was input for the 2011 class to generate the probability for each individual student returning to school in 2012.

There was some trouble with regards to predicting “unlikely to enroll.” Counting both conditionally and unconditionally admitted students, which totaled 2,208 students, it was only determined that 22 would be more likely not to return than to return. Overall, the predictive model was quite successful and accurate as far as retention rates. The researchers predicted an 82.1 overall retention rate, with the actual number being 83.1% (Burgher and Davis 2013, 28-31). The article closes with three reasons why it is difficult to predict behavior of individuals, which is why they used the group format. The student body itself is relatively homogeneous, especially in a situation where all students above a certain GPA are admitted. Coinciding with this, statistical analysis is based on averages, and one individual has the ability and the propensity to be farther off from the

average than a group. Lastly, the additional circumstances that occur in an individual's life make it impossible to predict with absolute certainty what factors have the most influence on a person's reasoning for continuing his or her education (Burgher and Davis 2013, 28-31).

A study by Strayhorn in 2013 examined the factors that affect college readiness for minorities at 4-year colleges and universities. Strayhorn examined high school GPA, twelfth grade National Assessment of Educational Progress (NAEP) standardized math score, and twelfth grade highest math level. The author attempts to determine if there are differences among historically underrepresented groups and their level of college readiness. Data for this study was taken from the Education Longitudinal Study and consists of a nationally representative sample of sophomores tracked during 2002 as sophomores, 2004 as seniors, and 2006 as high school graduates after 2 years from graduation. T tests were used to evaluate the differences between men and women. Women were more college ready in terms of reading standardized test scores, writing abilities, and 12th grade highest math. However, men were more college ready in terms of math standardized test scores and 12th grade math self-efficacy. All of the factors indicated that first generation college students were less college ready than continuing-generation college students. College readiness also differed significantly among races. Whites and Asians were much more college-ready than blacks, Hispanics, and Native Americans. High SES students were also deemed more college-ready than low SES students. Even among first generation college students and minority students, students who spent more time studying were often more college-ready than those who spent less time studying. Also, students who spent more time verbally communicating with parents

and friends about college were found to be generally more college-ready than those who did not have such conversations (Strayhorn 2014).

The “college for all” theme and its educational ramifications are examined in a study by Lee (2012) which demonstrated the connection between math achievement and college readiness. The goal of the study included comparing performance standards, benchmarks, and norms for college readiness, and also examining college readiness gaps between races and social subgroups. This study used three national longitudinal data sets to track math achievement levels. Data were taken from the Early Childhood Longitudinal Study-Birth Cohort (ECLS-B), the Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K), and the National Education Longitudinal Study of 1988 (NELS:88). Results found a high correlation between twelfth grade math levels and college completion. Results also indicate that there were significant math achievement level differences between 2-year versus 4-year colleges. Results also found strong gaps between racial groups and college completion. White and Asian groups were on track for 4-year college completion as early as elementary school. Hispanics were on track until Grade 3, while for African Americans the average math achievement was only good enough for 2-year college completion until primary school and went down to 2-year college entrance during middle and high school years. The researcher concludes by discussing the many limitations of the study. First, national math comparison data sets may lack validity. Second, college readiness levels were observed during different periods of data sets. Third, only college entrance and completion were observed; it may be important to understand college achievement levels as well. Lastly, the study ignored variations between institutions beyond 2-year and 4-year institutions (Lee 2012).

Lotkowski, Robbins and Noeth (2004) conducted a study exploring the correlation between academic and non-academic factors and retention. According to previous research, many of the variables that affect retention with any type of significance are non-academic in nature. Among these factors are level of commitment to obtaining a degree, level of academic self-confidence, academic skills, and level of academic and social integration into the institution (Braxton 2000; Braxton and McClendon 2002; Kennedy, Sheckley and Kehrhahn 2000; Mangold et al., 2003; O'Brien and Shedd 2001; Wyckoff 1998; as cited in Lotkowski Robbins and Noeth 2004). Socio-economic status (SES) was also measured because it has been shown to be a potential influence on retention (Hossler and Vesper 1993 as cited by Lotkowski, Robbins and Noeth 2004). Socio-economic status is determined in part by a parent's educational level and the family income both of which also have a relationship to retention.

The findings of Lotkowski et al., (2004, 41) indicated that factors such as academic-related skills, academic self-confidence, academic goals, institutional commitment, social support, certain contextual influences (financial support), and social involvement all had a positive relationship to retention. The strongest of these factors as measured by correlation coefficients were academic-related skills (.366), academic self-confidence (.359), and academic goals (.340). Academic factors, previous discussed, like high school GPA (.448) and ACT scores (.388) were also shown to have a significant relationship. Considering that the combination of non-academic and academic factors should be addressed, retention strategies reflecting these deficiencies should be deployed, such as combining tutoring and faculty-mentors and peer support (Hurd 2000; Ramirez

1997; Tinto 1997; as cited in Lotkowski, Robbins and Noeth 2004). Courses and programs that build mentoring and support groups build student involvement, motivation, and academic self-confidence (Mangold et al., 2003; Padgett and Reid 2003; as cited in Lotkowski, Robbins and Noeth 2004). Another strategy is block registration, in which a student may enroll in a program and be immersed in a specific cohort that progresses through the academic terms together. This would lead to increased social support and integration, which can lead to higher rates of persistence (Mangold et al., 2003; as cited in Lotkowski, Robbins and Noeth 2004)).

Taking a different approach, Schroeder (2013) focused on institutional processes and how these processes might affect graduation rates. He argues that undergraduate students' retention and graduation rates not only correlate with ACT and SAT scores and GPA, but also empowering institutional experiences that foster a high level of growth and success. Specifically, Schroeder (2013, 9) seeks to examine the following question: "What kinds of empowering institutional processes produce the most desirable results for the majority of students?" Schroeder discusses Kurt Lewin's theory (1936), interactionist framework, in detail. Lewis believed that human behavior is conditioned by multiple factors including those beyond the individual's control. For example, if a student is placed in a math class beyond his abilities, he does not have a high chance of performing well. The student is not the only factor that plays a role in his or her success or lack thereof; educational institutions play a role as well by guiding the student and providing an empowering experience (Lewis 1936, as cited in Schroeder 2013). Beckett (2006) found how important institutional processes are when he conducted a longitudinal study on freshman interest groups (FIGs) at the University of

Missouri-Columbia. Freshmen interest groups are in essence learning communities where students are grouped based on a common interest, such as their major. His study included over 13,000 freshmen between 1998-2001 and the FIGs had higher retention and graduation rates than non-FIGs. For example, 6-year graduation rates for FIGs were 69% and 62% for non-FIGs.

Hamrick et al. (2004) also examined the extent to which institutional characteristics affect graduation rates. This study focuses on resource allocation of different public universities. The goal of their study was to develop a statistical model that shows the relationship between resource allocation and graduation rates. The model consisted of institutional characteristics such as Carnegie type and selectivity, and resource allocations such as instructional expenditures and student affairs expenditures. The goal of the statistical model was to provide leaders of higher education with information in order to make strategic resource allocation decisions that would improve graduation rates. The data for their study came from the Integrated Postsecondary Education Data System (IPEDS) and included 513 accredited public institutions. The results of the study indicated that instructional expenditures, library expenditures, and a number of institutional classification variables were significant predictors of graduation rates (Hamrick et al., 2004, 45).

While predicting retention rates is possible and plausible, working on identifying at-risk students and working with them is another area of concern and can raise retention and graduation rates. At-risk students are found to remain at-risk throughout their college career. Also, the degree to which the student was at-risk is predictive of whether the student subsequently re-enrolls elsewhere and the type of institution at which the re-

enrollment occurs (Singell and Waddell 2010). Building on the variables described earlier by Burgher and Davis (2013), Singell and Waddell identified additional variables that affect retention rates. These variables included institutional selectivity, academic preparation, and financial aid (Singell and Stater 2006, as cited in Singell and Waddell 2010). There exists uncertainty regarding the prediction of retention and the efficacy of treatment as to whether administrative action and associated resource expenditures would yield a net benefit. This problem is becoming of more import as higher educational institutions become more conscious of asset allocation and return on investment capital.

There are two ways to view targeting at-risk students in correlation to retention rates. The first is a type I error in which treatment is foregone for all students in the class. Conversely, a type II error can be defined as applying treatment to all students, regardless of whether or not the students are classified as at-risk. The results of the research indicate, as might be hypothesized, that a balance must be struck to avoid using too much capital. There must be a way to target those who are a priori at-risk and work with them to increase retention rates without attempting treatment on everyone in the class. As stated previously, at-risk status usually continues for students throughout their collegiate life, so collecting data to analyze if a student is at-risk after their first semester allows for treatment to a large majority of people who will remain at-risk. A possible correlation that exists between expected GPA (the average of all GPAs for that class) and actual GPA for the student could assuage some of the difficulties in becoming at-risk early on. It is discovered that students who take somewhat harder classes in their initial schedules maintain a higher GPA comparative to the expected GPAs for that class (Kuh et al., 2005, 71-75). Previous literature suggests that what happens to students after they

enter college is more influential in their persistence decisions than characteristics they bring to college (Pascarella and Terenzini 2005), but because we know that at-risk students stay at-risk throughout their studies, more comprehensive advising could be beneficial. Creating a model to observe retention rates of several cohorts would be a start to experimenting with treatments. Getting enough data to predict retention rates in future first-time full-time (FTFT) cohorts, and then randomly selecting groups from those FTFT cohorts while applying a treatment program would allow a university to determine whether the treatment outweighs the financial cost of administering the plan.

Other research has focused on the implementation of dedicated counselors at schools in an attempt to aid at-risk students. Osberg (2004 as cited in Rothman 2010,1-3) determined that if a counselor, assuming the cost to the school is approximately \$40,000/year, is able to support three students who are at-risk of dropping out, will have nullified the cost incurred. As a result, a potential option for increasing or minimizing potential expenses while dealing with the retention of at-risk students includes the development of an adequate counseling system. The development of these systems and various other student aid programs is to avoid attrition at any cost. Attrition rates are detrimental to higher education institutions for a number of reasons: there is an exclusion of young, low-skilled workers from the employment ranks, the perceived misappropriation of funds from state and local governments for the investment into students that did not complete, and the lost revenue for the institution. There may be a way to alleviate attrition rates by taking a look at various aspects of the student's high school benchmarks.

A study by Kappe and van der Flier (2012) examine how personality traits can

be used to predict academic success in higher education. Students at a Netherlands college completed a survey that measured intelligence, the Big Five personality traits, motivation, and four specific personality traits. Then the following outcomes were measured: GPA, time to graduation, exam scores, skills training, team projects, internships, and a written thesis. The authors found that conscientiousness is the best predictor of all the achievement measures, even surpassing GPA and intelligence. It may be beneficial to teach about the importance of conscientiousness as early in a student's academic experience as is practical. The "conscientiousness factor" as evidenced by student behavior is also demonstrated in a study by Li et al., (2013). This study examined how math readiness and student course behavior have an impact on knowledge gain and course success. Course behavior consists of factors such as attendance, participation, and homework completion. The four variables analyzed were initial mathematics readiness, course behavior, mathematics knowledge, and course success. Participants consisted of 1,254 students at a 2-year public college in a large midwestern U.S. city. All participants were enrolled in developmental math courses. Participants were placed in the developmental courses if they did not meet COMPASS requirements. Predictor variables were mathematics readiness and student course behavior; dependent variables were mathematics knowledge and course success. Results indicate that math readiness showed strong direct effects on post-test math knowledge as well as indirect effects on course success via course behavior. Post-test math knowledge also showed strong direct effects on course success. The main implication from this study is that students can be targeted for intervention based on academic and behavioral risks before entering college. Students may even be placed into groups based on their needs, i.e.,

students lacking in algebra skills can be placed together in high school and teachers can focus on improving these skills (Li et al., 2013, 14-36).

The Impact of Financial Aid

No attempt to grapple with the issue of retention, graduation, and progression would be adequate without addressing the impact that money has on a person's ability to begin and continue a college education. Not only is the lack of financial resource a major retention factor, it is also a barrier to college matriculation. The Advisory Committee on Student Financial Assistance (ACSFA) reports that close to 50% of academically qualified low-income students failed to enroll at a 4-year college because of financial barriers (ACSFA 2002, as cited in Herzog 2008). The problem, argue Gerald and Haycock (2006), is primarily with flagship universities. It is estimated that more than 60,000 low-income students were unable to enroll in 2003 because of financial issues. These were students who were accepted academically. Furthermore, low-income students, even if they are successfully enrolled into a university, would continue to struggle. This lack of enrollment persistence or retention is attributed mainly to inadequate academic preparation and student self-reported reasons rather than financial aid (Adelman 2007 as cited in Herzog 2008).

Herzog (2008) implements a propensity score matching model that accounts for demographic, pre-college, and first-year university experience variables. A FTFT is grouped according to the type of first-year aid received; this includes grant and merit-based students against those who relied strictly on loans. These two groups are then also put together to be an "aid-treated" group and compared against students without any first-year aid, or "the untreated," based on the propensity for receiving aid. The propensity

score-matching method is estimated to remove about 90% of bias associated with observed covariates as long as it has at least five strata, or independent variables. The independent variables included were socio-demographic attributes, pre-collegiate academic preparation, college major, first-year credit load, campus residency, and timing of the university-entry test (Herzog 2008). Results were based on approximately 5,000 first-year students and measured the correlation between the likelihood of receiving aid and second-year retention both before and after taking into account first-year math experience and GPA. First-year math experience and GPA are included stepwise because they are key determinants in retention (Adelman 2004; Pascarella and Terenzini 2005; Herzog 2008).

Factoring out EFC (estimated family contribution), the study concluded that students with an average propensity to receive aid are 6% more likely to persist compared to freshman with a low chance for aid. Students who are most likely to receive aid are nearly 15% more likely to persist than the same low chance freshman (Herzog 2008, 2-3). However, this data was proven inconsequential when also factoring in academic performance (GPA and advanced math). Academic performance had much more impact regarding persistence rates whether the student was more or less likely to receive aid. Next, the influence of financial aid on retention of students from different EFC levels was observed. An EFC less than \$4,000 included low-income background students who were eligible for need-based grants such as Pell Grants and Supplemental Education Opportunity Grants. The next level consisted of an EFC between \$4,000 to \$10,000, and the students in this level received between \$800 and \$900. Lastly, students with an EFC over \$10,000 have relatively no financial need and usually only get aid such as merit-

based aid and unsubsidized loans.

There was no significant correlation between EFC and retention, and propensity for receiving aid remains uncorrelated to low income students' level of academic success. However, it was noted that a one letter grade increase in GPA was associated with a 15% rise in retention as it concerns low-income freshmen who had received aid. For students in the medium EFC range (\$4,000 to \$10,000), again there was no correlation found between likelihood of receiving aid, amount or type of aid received, and second-year persistence. Outside of financial aid factors, doing well in advanced math tends to increase the retention probability by 10% over those who only have passed a college algebra class. The FTFT in the highest EFC improved their persistence chance by 18% when receiving grant or merit-based aid. Institutional grant aid for low-income students is positively correlated with their persistence, at a rise of about 7% per \$1,000. As stated previously, retention of low EFC students was significantly associated with GPA and advanced math such that if a low EFC student took a math remediation class and failed to pass a first year math course, a 12% decline in persistence was noticed (Herzog 2008, 8-11). To recap, only middle-income students appeared to be influenced by the amount of aid they received, and everyone, regardless of EFC, exhibited a strong correlation with GPA, advanced math, and retention. There could be a variety of reasons as to why grant or merit or scholarships may have no effect on the average student. Grant aid does not entail a payback (Pell Grant, etc.) while reducing the student's cost of investment, perhaps leading to academically risky behavior. Conversely, scholarships are expected to reduce such risks as eligibility for these scholarships depend upon academic performance.

The finding that capital fails to have a significant positive impact on academic

performance has been echoed in previous literature. Stinebrickner and Stinebrickner (2004, 3-4, as cited in Herzog 2008) estimated that 85% of enrollment attrition was due to factors unrelated to financial resources. Self-financed students were less likely to experience academic failure than those with little or no personal investment, and contribute to the moral hazard associated with financial support (Val Long and Shimomura 1999 as cited in Herzog 2008). Failure to produce a positive link between financial aid and retention of low-income students may be due to other factors. Academic interventions, as stated earlier in the literature review, may make students regardless of income more academically prepared for college. The results of Herzog's analysis were based on students at a moderately selective public university with a sizable segment of first-year commuters. These and other aspects of the climate surrounding the university may differ from college to college.

Tinto (2006), a leader in the area of retention and financial indicators, noticed that although more low-income students were being accepted into 4-year institutions, they suffer from a lower graduation and persistence rate. Tinto's focus in the article concentrated on the transition from 2-year universities to 4-year universities. Among those beginning at a 2-year college, only 8% of low-income students earn a bachelor's degree within 6 years, while 25% of high-income students do (Tinto 2006, 11-12 as cited in Herzog 2008). Even within elite institutions, income matters. Students in the lowest quartile of SES while at an elite institution are less likely to graduate than students in a higher quartile by 14%. Tinto argues that in the case of admission to 2-year institutions, we must know what happens within the 2-year colleges that are more successful at producing students who are graduating successfully at 4-year colleges within a certain

timeframe.

A study by Dynarski (2000) also examined the effects of merit-based aid in the form of Georgia's HOPE Scholarship program. There was a strong enrollment response for middle and high-income youth, but relatively little benefit yield for low-income students. This subsequently widened the enrollment gap between low and high-income families. Wetzel et al., (1999) conducted a study to examine the effect of financial aid on student retention and found a similar result to that of Herzog. The empirical results suggest that net costs of attending the university (tuition minus grants) negatively correlated to the student being retained, but also that their effect on retention is small as a whole. Variables such as a student's commitment to the institution or to a college degree Herzog affirms play a bigger factor. An important finding in Singell's research regarding re-enrollment found that students with higher incomes, lower financial eligibility, and higher SAT scores were more likely to re-enroll than their counterparts. While Herzog found no relationship between type of grant and retention, Singell's research uncovered a slight positive linear relationship between types of grants/loans and retention. Grants (Pell grants, state supplemental grants) increased the probability of retention by 1.3% per \$1000, whereas subsidized loans increase retention probability by 4.3% per \$1000 (Singell 2001, 18-19). It is important to note that whether or not these increases are of significant correlational value is unclear, as the relationship could potentially be due to normal variance.

Standardized Indicators

There are many different forms of standardized indicators used to predict success in college. These are measures which can be tested and compared over time. In

addition to the well-known indicators such as the ACT and SAT for purposes of this section of the literature review, other measures including high school GPA, level of math courses taken in high school, and placement tests such as the COMPASS will be considered here. As the most quantifiable, mathematics is perhaps the area that has received the most attention. The recent decline in the U.S. in STEM (science, technology, engineering and mathematics) proficiency has been well-documented. Knowledge in these areas is crucial for the United States' ability to remain competitive in the global economy. It is imperative that students in higher education master these areas, and it is more important that they are given the necessary tools to do so. Preparation for such coursework can begin at all grade levels, especially high school.

According to recent literature, there exists a correlation between advanced math classes taken and success at the university level. Megert (2005) conducted research at a community college in New Mexico inspecting the honors scholarship retention rate when the scholarship was awarded based on completed advanced high school math classes compared to the retention rate when the scholarship was based solely based on GPA. Advanced math classes were defined as trigonometry, pre-calculus, and calculus. Results indicated that there was no significant correlation between grades earned in advanced math classes and retention, however, there was a 10% difference in retention when considering scholarships awarded based on a high GPA and the completion of an advanced math class versus scholarships based solely on a high GPA. When HighMath (having completed an advanced math class) was used in conjunction with GPA, 73% of students retained their scholarship, while only 58% retained their scholarship when GPA alone was used. Also, when MathScore (a combination of the rigor of the course, number

of courses taken, and grades received in courses above Algebra I) was combined with GPA, there was a 67% retention rate; whereas students awarded scholarships based solely on their GPA had a 48% retention rate (Megert 2005, 55-56). This research expanded upon Adelman's (1999) study, which found that completing a mathematics class in high school above the level of Algebra II doubled the odds of a student achieving a bachelor's degree. Adelman's study, however, had some limitations by restricting the records investigated to students who indicated they were seeking a degree, and by withdrawing students from the study who did not submit high school transcripts. Megert's research, being much more empirical in nature, is a solid indicator that a high GPA combined with the completion of an advanced math class increases the chances of retention (Mergert 2005).

There was a similar result found in the transition from 2-year to 4-year schools. A student's level of secondary school mathematics is a good indicator of preparedness for university mathematics, according to a study by L. J. Rylands and C. Coady (2008). Every student in the study who studied Advanced Mathematics in secondary school passed the university basic mathematics class, and 67% gained a grade higher than a pass. Also, 78% of students who had taken Intermediate Mathematics in secondary school passed, while 23% of students who had taken Elementary Mathematics passed (Rylands and Coady 2008, 741-753). Recommendations from the article include recognizing that some first-year students in mathematics have different levels of preparation and that extra assistance must be in place for students who are not as well prepared (Rylands and Coady 2008, 741-753).

Complimentary to the above design, Friedl, Pittenger, and Sherman (2012)

undertook a study examining the relationship between taking intermediate algebra at a 2-year community college and a 4-year university, and their subsequent pass rate in a college algebra course at the University of Tennessee at Chattanooga (UTC). While the average GPA in intermediate algebra for students at the 2-year college (2.86) was higher than that of the 4-year university students (2.44), the 4-year university students scored higher on the college algebra class (Friedl et al., 2012, 526-532). Undergraduates who took the intermediate algebra course at a 4-year university had a pass rate of 85% in the college algebra course at UTC, while undergraduates who took an intermediate algebra course at a 2-year community college had a pass rate of 70% (Friedl et al., 2012, 526-532).

A study by Bremer analyzed the outcomes of developmental English and math courses in three different states. Data was recorded on 7,898 freshmen beginning in 2009 and 2010. The authors found that many factors were related to persistence into a second term of college, such as graduation and higher overall GPA. The results of the study indicated that older students, White/non-Hispanic students, and occupational students were more likely to graduate. Math ability at the time of college entrance was a strong predictor of student success. Reading and English placements as predictors were limited to retention in the second term and/or second year (Bremer et al. 2013).

Hoyt and Sorensen (2001) examined the remedial placement rates of college students who had completed high school. Of note, the study showed that over half of students who completed college math prep classes were still placed in remedial math classes in college because of their inability to achieve a sufficient score on the placement exam. A sampling of educator sentiment showed that the high remediation rates in

college were the result of failure to take college preparatory classes, grade inflation, and lack of challenging coursework. This study was conducted in an effort to determine how high school preparation affects remedial placement rates at Utah Valley State College (UVSC). Following up on a previous study conducted at UVSC, researchers examined both English and math preparation of high school students. Students' placement in college classes was determined by their ACT and COMPASS scores. The following additional factors were measured to determine if they impacted remedial placement: preparation in high school, grades in math and English courses, gender, ethnicity, delayed entry into college, and attendance at different high schools. The data collected confirmed the researchers' hypothesis that the level of high school preparation and grades were significantly related to placement in remedial education. The researchers suggest that grade inflation was a major factor and that teachers were apparently awarding passing grades to many students who had not adequately learned the material. The authors concluded by recommending that high school teachers collaborate with higher education to improve students' preparation for college (Hoyt and Sorensen 2001).

Melguizo, Kosiewicz, Prather, and Boss (2014) examine current assessment and placement policies used to assign students to developmental math courses. So far, there are no national guidelines for what is considered developmental education and how students should be placed into developmental coursework. This has generated confusion and frustration for both students and teachers, and has also made it difficult to understand how effective placement policies truly are. The researchers conducted their case study on The Los Angeles Community College District (LASSC) because it is composed of a diverse group of 250,000 students. Both quantitative and qualitative data were gathered

through websites, transcripts, interviews and administrative documents. Overall, results found that educators lack the technical expertise to ensure that assessment and policies facilitate student success (Melguizo et al., 2014).

The lack of recognized guidelines for remediation may lend itself to problematic placement. A 2013 article by Sparks showed how remedial placements may be overused and possibly misrepresent the need for the high rates of remedial placements in U.S. colleges. According to Sparks (2013, 2), “new research findings suggest a significant portion of students who test into remedial classes don’t actually need them.” The article claims that as much as \$7 billion a year is spent on noncredit remedial classes, and many students are wasting limited financial aid on courses that may not be fully necessary to succeed in college (Sparks 2013). The recent push to eliminate remedial classes altogether or to require remedial/credit classes be taken simultaneously as co-requisites is a reflection of the cost of remediation as well as its potential to be counterproductive to retention and degree completion. A 2014 study by Hodara and Jaggars (2014) found that accelerating through remedial programs in shorter sequences led to greater access to college-level coursework and long-term success (Hodara 2014).

McGraw-Hill’s decision to launch Assessment and Learning in Knowledge Spaces (ALEKS) Placement, an open-response placement and remediation tool for higher education, is yet another strategy to address the remediation question. The primary goal of ALEKS is to improve math placement accuracy while increasing graduation rates and decreasing the need for remedial education in higher education. The combination of rising tuition rates and low college graduation rates of remedial students makes the need for successful placement techniques more necessary than ever before. ALEKS allows

students to bypass already mastered material, to reduce wasted time, and gives students their own tailored individual study plan. ALEKS is also the only placement exam that uses qualitative open-ended questions instead of the multiple-choice format. This format helps to pinpoint student's current knowledge, level of mastery, and areas that need improvement. ALEKS can also be taken online and allows students to complete assessments at their own convenience. At Kent State University, 75% of incoming freshman needed remedial math courses and ALEKS helped create a 24% increase in completion of these courses and a 10% decrease in the rate of drops, fails, and withdrawals (EDUCAUSE 2013, 1).

A study by Belfield and Crosta (2012) questions the validity of using college placement tests as predictors of college performance. Their research replicated and extended a previous study by Scott-Clayton (2012). The math and literacy tests from both the Accuplacer and COMPASS placement tests were evaluated. The results indicated that these placements do not predict college performance levels. However, there was a strong relationship between high school GPA and college GPA and the authors suggest that more weight should be given to high school GPA when determining college acceptance standards. The authors even suggest waiving college placement tests for students who have high school GPAs at or above the passing threshold. High school GPA also has a strong association with college credit accumulation. A student whose high school GPA is one grade higher will have accumulated approximately four extra credits per semester (Belfield and Crosta 2012, 22-23). Placement scores were also positively associated with credit accumulation; after three to five semesters, a student with a placement test score in the highest quartile has on average nine credits more than a

student with a placement test score in the lowest quartile (Belfield 2012, 23).

Rather than eliminating the use of placement tests, other authors advocate the use of multiple measures for level of course placement. Ngo and Kwon (2014) examine whether or not the following measures improves placement decisions in higher education: high school GPA, prior math achievement, and noncognitive measures. The authors focus on math placement because math placement tests are often inaccurate. Although there are existing studies which show that high school GPA and course completion are related to college outcomes, rarely have these measurements been used for math placement decisions. Specifically, Ngo and Kwon (2014, 25-28) attempt to answer the following: “(1) Does using multiple measures increase access to higher-level math courses, particularly for groups disproportionately impacted by remediation? (2) How do students who are placed using these additional measures into a higher-level math course perform in comparison to their peers?” High school GPA and information prior to math course taking appeared to increase access to higher level math courses and students were also more successful in those courses (Ngo and Kwon 2014, 25-28).

There are additional types of placement exams taken prior to the start of college that can predict a student’s successfulness and, in some cases, these exams will earn the student credit hours towards degree completion. From a financial perspective, taking courses in high school that may count as college credit is a way to encourage cost-efficiency, especially as tuition rates continue to grow. These exams, called Advanced Placement Exams (AP), are taken after “AP courses” are taught during the regular school year. AP courses are an all-around positive way to master material: students can begin them in high school, they may serve as college credit thus saving money, and their

benefits are extended to minorities. According to a press release, “Students who master AP courses in high school are three times more likely to graduate from college. For minority students, that multiplier is even greater: African-American and Hispanic students who succeed in AP courses are four times more likely to graduate from college” (NMSI 2014, 1). The use of AP classes as an indicator of college success was duplicated in data regarding the Chicago Public Schools. The results showed that AP math and science courses have significant effects on college outcomes. Taking AP math or science increases likelihood of enrolling in a 4-year college versus a 2-year school, and taking AP math also increases the likelihood of enrolling in a selective college. Lastly, taking either AP math or science courses related to improvement in 2-year persistence at 4-year colleges (Kelley-Kemple et al., 2011).

Koppius and Shmueli (2011) examined the importance of predictive analytics and the lack of such in information systems research. Predictive analytics includes statistical models with the intention of empirical predictions. This is different from other predictive analyses since it focuses on more than just prediction from theory. Methods for assessing the quality of those predictions in practice (predictive power) also fall under this terminology. In addition, predictive analytics plays a crucial role in theory building, theory testing, and relevance assessment (Dubin 1969; Kaplan 1964, as cited in Koppius and Shmueli 2011). The authors suggest predictive analytics prowess regarding six aspects: generating new theory, developing measures, comparing competing theories, improving existing models, assessing relevance, and assessing predictability.

To summarize, if the goal is to increase retention, progression and graduation, then models must be developed that yield empirical results, models that can be refined

and improved that will yield new empirical data that can be measured and compared over time. As the literature suggests, there is a compendium of factors that must be accounted for when considering how to improve retention. These factors include a mix of academic and nonacademic factors that have to be taken into account and integrated into a comprehensive approach. Also, institutional factors should not be overlooked. Factors like high school GPA and SAT scores, as to be expected, have a high correlation with success at a university, but the social atmosphere and context of the modern student tends to be underestimated when retention is considered. Using predictive analytics that consider all of these variables including financial aid to students from various socio-economic backgrounds, it may be possible to locate and aid future at-risk students. When these at-risk students are found, systems can be put into place such as peer support through blocked registration, freshman learning communities, and mentoring programs. Keeping in mind that at-risk students stay at-risk throughout the entirety of their collegiate life, aid or support for these students at any time would be beneficial. The literature also indicates that more research needs to be done in order to determine whether or not these counter-measures are fiscally prudent, and should be developed as part of an overarching strategic plan.

Chapter III

METHODOLOGY

Predictive Modeling

The methods of statistical analysis that will be used in this study are predictive modeling, and this research will be conducted on data provided by Valdosta State University (VSU) and it has been approved by a Protocol Exemption Report (Appendix A). Predictive modeling is an end to a means, not an end-in-itself. As such, predictive modeling is a tool designed to achieve a greater purpose, a higher goal. So, the first step in developing a predictive tool is goal definition, and is centered around what specifically needs predicting. When successful, predictive modeling accurately predicts an outcome value for a new set of observations. This is known as *prediction* for numerical data, and *classification* for a categorical outcome. If the outcome is categorical, then a separate procedure is used to rank new sets of observations according to their probability of belonging to a certain class—a process known as *ranking*.

The next step in the predictive modeling progression process is data collection and study design. Ideally, the data used for modeling and for prediction consists of the same variables and are from the same population. Predictive analytics needs to have a larger sample size than explanatory modeling or regular experimental procedures because there is a higher degree of uncertainty for predicting the behavior of individuals juxtaposed to population-level parameters. Increasing sampling size also reduces both model bias and sampling variance. Data dimension is another consideration in this phase, and is usually begun with a large number of variables dependent upon domain knowledge

and potential for new relationships. Data preparation follows from data collection, and involves missing values and data partitioning. Missing values are handled much the same as any data analysis, and uses proxy variables, dummy variables, and regression trees to counter the absent values. Data partitioning is an important part of predictive analytics, and it is divided into the training set and the holdout set. The training set is used to fit models, while the function of the holdout set is to evaluate predictive performance of a final chosen model. The holdout set is not included in the original data, but is used to measure against the predictive set from the original data. Exploratory Data Analysis (EDA) is used primarily in predictive analytics in a free-form fashion to uncover potential underlying constructs that may be less formulated. PCA, or principle components analysis, is a data reduction technique that is often used prior to the EDA to reduce sampling variance and perhaps increase predictive accuracy.

Predictive models and analytics can lead to the discovery of new constructs or relationships, and provide evidence regarding unknown patterns. They are more data driven than explanatory statistical models in that they are derived from empirical information, and provide a good source to assess practical relevance of theories. A potential problem with predictive analyses could be the extent to which they are willing to reduce sampling variance. When using the PCA and other data shrinkage methods like ridge regression and principal components regression, there is a strong sacrifice or bias for the reduction in sampling variance. The result could be an increase in the amount of method bias involved in the predictive analysis and may lead to artificial numbers. Even with their limitations, predictive models and analytics can be used effectively in the education setting to determine multiple important factors including graduation,

persistence, and retention rates.

While the two predictive models are distinct, the types of outcomes are very similar. The outcomes for mathematics-based courses are a student passes or does not pass, and the same for reading-based courses, a student passes or does not pass. This study utilized linear or ordinary least squares regression to determine the likelihood of a student passing a VSU reading-based and mathematics-based course (and its implications for retention). Logistic regression is another method that accounts for a dichotomous dependent variable outcome. Using the equation below,

$$P = \frac{e^{[\beta_0 + \beta_1(X_1) + \beta_2(X_2) + \beta_3(X_3)]}}{1 + e^{[\beta_0 + \beta_1(X_1) + \beta_2(X_2) + \beta_3(X_3)]}}$$

The primary hypothesis of this paper states: By taking into account standardized test scores X_1 , high school GPA X_2 , and the high school's performance ranking in creating an index X_3 , one can predict a student's likelihood of academic success or failure (DFW rates). This can also be taken a step further by splitting the reading and math portions of the standardized tests to create a math index and a reading index. Thus, one will be able to predict performance with regard to reading-based courses such as English and history and to mathematics-based courses exemplified by courses in math and chemistry.

The first factor in the predictive model is the SAT score, which is scored in 10 point intervals with the maximum combined score of 1600: 800 on the critical reading

(CR) and 800 on the mathematics (Math) portion of the test. This leaves a very wide range of individual scores. For the purpose of this research the SAT composite score (combined CR and Math) was broken into quartiles. (For those students who took the ACT test a conversion table was used to convert the ACT score to a comparable SAT score. An example of a conversion table can be found at <http://www.studypoint.com/ed/sat-to-act-conversion/>.) Index scores are assigned based on the following criteria as highlighted in Table 1:

Table 1: Quartiles of Student Preparation

Index Score	Quartile	Percentile
1	1	76-100 percentile
2	2	51-75 percentile
3	3	26-50 percentile
4	4	1-25 percentile

So a student in the top quartile was received an index score of 1 with regard to the standardized test score, while a student in the bottom quartile received a 4. Not only were Index Scores assigned to the composite, they can be further delineated and assessed based on the individual CR and Math portions of the SAT.

The second factor involved is the high school GPA. For comparison purposes all high school GPAs were converted to a 4-point scale based on completion of the college preparatory curriculum. This process converted all grades to a 4 point scale with 4 being the highest. (An example of a conversion table can be found at

<http://www.collegeboard.com/html/academicTracker-howtoconvert.html>.) Some high schools do not assign numerical GPAs, they simply use letter grades and sometimes they use plus and minus letter grades so the conversion table is useful in assigning values within the 4-point scale. Once all the students' GPAs were converted to a 4-point scale, they were also divided into quartiles and assigned numerical scores in the same fashion as the SAT scores with the highest GPAs in the top quartile being scored as a 1, the second quartile a 2, the third quartile a 3 and the bottom quartile a 4. This GPA score coupled with the SAT score placed every student in one of 16 categories with the students with the highest standardized test scores and GPAs coded as a 1-1 and students with the lowest GPAs and SAT scores coded as a 4-4.

Before the third factor was calculated into the equation, a contextual understanding of some important historical elements with regard to high school GPAs in Georgia was necessary. Prior to 1993 grades were distributed across a much wider distribution with the mean GPA approximately 2.6 out of a 4.0 scale. With the creation of the Georgia lottery and the Hope scholarship program in 1993, high school GPAs across the state began to rise in direct response to the Hope scholarship's requirement of a minimum 3.0 high school GPA. As a result of this grade inflation, the distribution of GPAs narrowed quite significantly. More significant is the fact that many of the high schools that perform at the lowest level with regard to the Georgia graduation exam have a higher number and percentage of students whose high school GPAs are above the 3.0 threshold (<https://gosa.georgia.gov/report-card> (2012)).

The third and final factor in the predictive model is the ranking of the 380+ public high schools into one of four categories or quartiles. The rankings in this category were

based on the report issued by the Governor’s Office of Accountability. Specifically, the metrics evaluated were Adequate Yearly Progress, Georgia High School Graduation Test and Writing Test, Advanced Placement and Honors. The tables that follow show how the high school ranking is assigned.

Table 2: Example of Report Card Metrics

CEEB	School Name	2011 AYP Status	2011 School Improvement Status	Consecutive Years In AYP	Corrective Action Status	State Directed Status
100001	Blue High School	Met	NI - Made AYP	1	NA	NA
100002	Green High School	Did Not Meet	NI	0	Yes	Yes
100003	Yellow High School	Met	Adequate	3	NA	NA
100004	Orange High School	Did Not Meet	Adequate-DNM	0	Yes	NA
100005	Red High School	Met	Distinguished	5	NA	NA
100006	Purple High School	Did Not Meet	NI	0	Yes	NA

Table 3: Example of High School Rank

CEEB	School Name	2011 AYP Status	2011 School Improvement Status	Consecutive Years In AYP	Corrective Action Status	State Directed Status	Quality points	Rank
100001	Blue High School	1	1	1	1	1	5	2
100002	Green High School	0	0	0	0	0	0	4
100003	Yellow High School	1	3	3	1	1	9	2
100004	Orange High School	0	2	0	0	1	3	3
100005	Red High School	1	4	5	1	1	12	1
100006	Purple High School	0	0	0	0	1	1	4

Table 4: Additional Variables with Honors and AP

CEEB	School Name	LAR						
		Percent Pass	Mean	Percent in Below Basic	Percent in Basic Proficiency	Percent in Advanced Proficiency	Percent in Honors	Advanced/2 + Honors
100001	Blue High School	92	245	8	35	39	18	38
100002	Green High School	60	230	40	46	14	0	7
100003	Yellow High School	85	220	15	47	24	14	26
100004	Orange High School	70	219	29	36	21	14	25
100005	Red High School	95	270	5	21	40	33	54
100006	Purple High School	83	238	17	43	34	6	23

Table 5: High School Graduation Test Results

CEEB	School Name	LAR				Math	Social Studies	Science	Writing	Total	Rank
		Percent Pass Points	Mean Points	Advanced/2 + Honors Points	Total	Total	Total	Total	Total		
100001	Blue High School	3	4	4	11	9	11	12	10	53	1
100002	Green High School	1	2	1	4	3	5	7	4	23	3
100003	Yellow High School	2	1	2	5	5	5	3	3	21	3
100004	Orange High School	1	1	2	4	4	3	3	5	19	2
100005	Red High School	4	4	4	12	10	12	9	12	55	1
100006	Purple High School	2	3	2	7	8	8	6	7	36	4

Table 6: High School Rank

CEEB	School Name	GHS&GHSWT Rank	AYP Rank	Overall Rank
100001	Blue High School	1	2	2
100002	Green High School	3	4	4
100003	Yellow High School	3	2	3
100004	Orange High School	2	3	3
100005	Red High School	1	1	1
100006	Purple High School	4	4	4

In Tables 2 to 6 LAR is the Governor’s annual report from the Georgia Office of Accountability. AYP is Annual Yearly Progress toward “No Child Left Behind.” The highest achieving high schools in the state with regard to the LAR metrics (Percent Pass, Percent below Basic, Percent Advanced Proficiency, and Percent Honors) and the AYP Metrics from “No Child Left Behind” in the Governor’s report card will be scored as a 1.

So a student from that high school was coded with a 1 (a student in the high schools from 51 to 75 percentile received a 2, 26 to 50 percentile a 3, a student in the bottom quartile received a 4). This now creates up to 64 categories that entering high school students could be divided into based on their standardized test score, high school GPA, and high school's rank. For example, a student who went to the best high school and had the best GPA coupled with a top quartile SAT score was coded with a 1-1-1, and a student from the bottom quartile in all areas was coded with a 4-4-4.

This metric will be applied to 5 years' worth of entering student data, scoring more than 17,000 students into the 64 possible categories. The method above was used to categorize the students because using the SAT score and the GPA and splitting the students into anything smaller than quartiles would produce too many cells with too low of an N; therefore, being too small to garner significance or validity.

For the at-risk metric for reading, the model used the SAT Verbal (ACT Converted) and the high school GPA to develop categories to run regression models to determine risk of failure (DFW).

With Math, the model was further refined to create a Valdosta Math Index which will be used for placement in to Math 1101, 1111, 1112, 1261, 1113, and 2261.

The tables below explain how that placement would function:

Diagram 1. Entry Level Mathematics Courses with Prerequisite Flow*

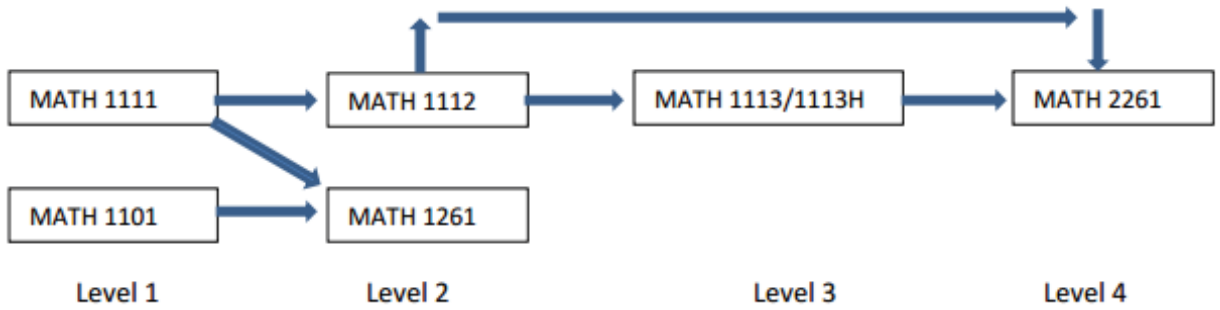


Figure 1: Mathematics Courses with Prerequisite Flow

Note: MATH 1111 is College Algebra, MATH 1101 is Mathematical Modeling, MATH 1112 is Trigonometry, MATH 1261 is Statistics, MATH 1113 is Pre-Calculus, MATH 2261 is Calculus I.

Table 7: Ranges for Quartiles High School GPA and SAT Math

Value	HS GPA	
	Low	High
1	3.41	4.00
2	3.04	3.40
3	2.69	3.03
4	0.00	2.68

Value	SAT Math	
	Low	High
1	531	800
2	486	530
3	446	485
4	0	445

Table 8: How Quartiles Combine to Indicate Math Placement

HS GPA	SAT MATH	VMI Math Level	Math Courses
1	1	Level 4	MATH 2261
1	2	Level 3	MATH 1113
1	3	Level 2	MATH 1112 or 1261
1	4	Level 1	MATH 1101 or 1111
2	1	Level 2	MATH 1112 or 1261
2	2	Level 1	MATH 1101 or 1111
2	3	Level 1	MATH 1101 or 1111
2	4	Level 1	MATH 1101 or 1111
3	1	Level 1	MATH 1101 or 1111
3	2	Level 1	MATH 1101 or 1111
3	3	Level 1	MATH 1101 or 1111
3	4	Level 1	MATH 1101 or 1111
4	1	Level 1	MATH 1101 or 1111
4	2	Level 1	MATH 1101 or 1111
4	3	Level 1	MATH 1101 or 1111
4	4	Level 1	MATH 1101 or 1111

In the above figure and tables, Figure 1 shows the courses offered at each Math Level (Math Level 1 = Math 1111 or 1101), Table 7 shows the categories for which the students receive their Valdosta Math Index, and Table 8 shows which Math the Index will place the student into. This should help with some of the issues with regard to poor success by getting the students into the correct math course from the beginning.

Historically many students were allowed to take math Level 3 because they were Science, Technology, Engineering and Math (STEM) majors even though they were only prepared for a Level 1 mathematics course.

The grades of students in each of these categories were then analyzed to see which combinations can predict better student success with regard to grades in college level courses. The hypothesis for this is that students with lower scores will perform better in college-level coursework than students with higher scores. A regression model was used along with a likelihood ratio to determine which populations of students are most at-risk of earning a DFW grade. These models drove information to faculty about students that are in their courses prior to the start of the term so that they can better prepare for those students likely to struggle. Then, at the end of the term for which the predictive modeling information is shared with the faculty and the intervention strategies have been deployed, an analysis of the grades that the students received was done using a chi-square and t test to determine if there had been significant improvement in the pass rates.

Faculty Portal

Once the predictive models have been developed, the second part of this methodology to improve student performance is creating a mechanism to deploy that information to faculty who teach and advise the students. This was accomplished by developing a faculty portal using Oracle's Application Express Software and was customized for Valdosta State University.

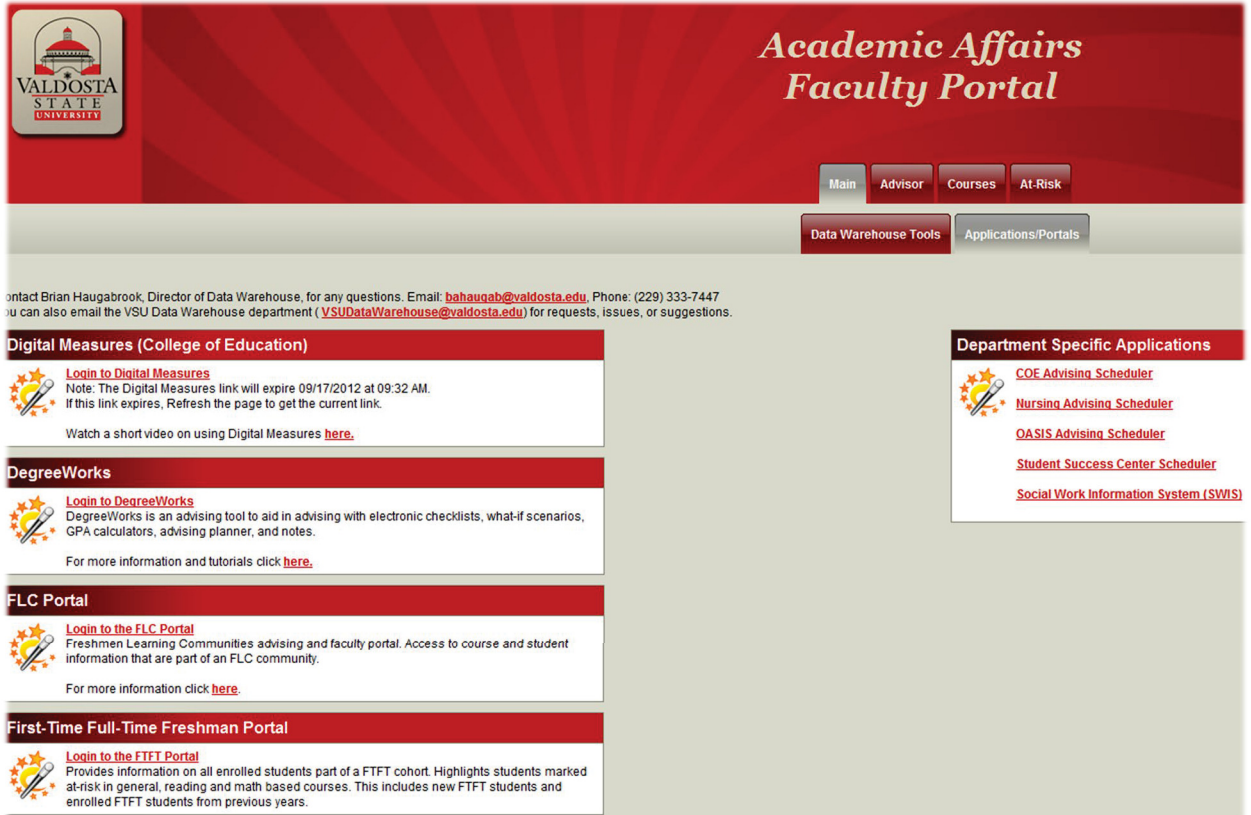


Figure 2: View of Faculty Portal Login Screen

The Faculty Portal contains the metrics produced by the predictive models in an easy-to-use roster format that will also provide information from multiple enterprise systems used on campus. Pictures of students were provided in the portal roster that comes from BlackBoard Transact, general student information such as name, email, ID, phone number were provided from Banner (see Figure 3). The portal also had integration to Degree Works (a degree audit and mapping solution) that provided faculty with even richer information about the students and where they were in progress toward their degree.

Back to Course List

Course: ACCT 2101 A - Principles of Accounting I

Rows: 100 Go

At Risk General = "Yes"
 At Risk Reading = "Yes"
 At Risk Math = "Yes"

1 - 84 of 84




Student Photo	Last Name	First Name	Middle Name	Attendance/Course Progress	Email	At Risk General	At Risk Math	At Risk Reading	DegreeWorks
	Doe	J	Tanner	Attendance/Course Progress Flags	bahaugab@valdosta.edu	Yes	Yes	No	DegreeWorks
	Doe	J	Jahkeem	Attendance/Course Progress Flags	bahaugab@valdosta.edu	No	No	No	DegreeWorks
	Doe	J	Gamar	Attendance/Course Progress Flags	bahaugab@valdosta.edu	No	No	No	DegreeWorks

Figure 3: Faculty Course Roster

The most important feature of this Faculty Portal was the early alert capability so that at any time from the start of the class faculty had the ability to flag a student for attendance or academic progress (see Figure 4). These flags drove automated communications to different professional staff based on who the student was, making the intervention very individualized. For example, a student who was an athlete, living in residence, who is flagged through the early alert system for attendance issues had automated communications go out to athletics, housing, and to the student's advisor. These three areas communicated with the student to find out why attendance was problematic and then put that information back into the system, thus closing the loop and informing the faculty member of the issues.

Return To List

Edit Student Attendance

Course ACCT 2101 - Principles of Accounting I

Edit Name Jane Doe

Edit Status Level 0: No Absences

Comments

0 of 2000

Update Attendance

ATTENDANCE

Attendance in good standing

Figure 4: Early Alert Window

This presented the development of the statistical methodology used to predict grades in key English (Reading Based) and math courses. The model assigned an index score to each individual student based on the criteria outlined earlier with the scores grouped and distributed to the appropriate faculty members through the faculty portal. The methodology also involved a detailed description of intervention strategies, the trigger points used to alert the various departments and how the results of the intervention(s) were inserted into the data flow to refine the process.

Intervention strategies observed for this study were Student Success Center Intervention, Freshmen Experience Intervention, and Professional Advisor Intervention. There were other strategies that were being developed and deployed, but these were applied equally to students based on the setting of the attendance or academic progress flags in the faculty portal.

The Student Success Intervention occurs when a faculty member flags a student for academic progress in the faculty portal. Once the flag is set an automated communication is sent to the Student Success Center with the student's information and any comments that the faculty member may have written in regard to the flag. The Student Success Center will then reach out to the student, first by email and if the student does not respond to the email or come to the Student Success Center within two business days, the student will then receive a phone call to encourage them to come in for tutoring on the course they are struggling with. The tutoring center will then keep records of the students who attend tutoring (number of sessions and hours). All students who have academic flags will enter this workflow and intervention strategy. However, the system will also deploy additional workflows and intervention strategies based on the student's attributes; for example, if the student is an athlete an automated communication will go to both the Student Success Center and the Athletic Department.

A second intervention, Professional Advisor Intervention, was evaluated in regard to attendance. If a faculty member set an attendance flag, then an automated communication was sent to the student's advisor. For professional advisors this triggered a workflow that had them reach out to the student through multiple channels of communication. The advisor first attempted contact through email. If the student had not responded in 48 hours, they followed up with phone calls and text messaging. When the advisor made contact, they engaged the student in a series of questions: How are you doing? It has come to our attention that you have been absent from Class X. Are you having any issues with Class X? The advisor explained the importance of attendance and attendance policies. Then the advisors fed the information from that conversation

back to the faculty member closing the loop. As with the academic flags, there were other interventions that were deployed based on the student's attributes; for example, if a student lived in housing an additional intervention strategy was deployed and that work flow will send a resident assistant to check on the student with a "warm bed check" by knocking on the student's door and making sure the student was not sick and having a conversation with the person about class attendance.

The effectiveness of these intervention strategies was evaluated using the Chi-square and *t*-test statistics to determine if the interventions had a significant effect on the performance of the students who received them versus the students who did not. To control for bias, only students in the middle 50 percentile were evaluated based on faculty flagging them and receiving the intervention. Faculty portal users were compared to non-portal users in the analysis to ascertain if there is a significant improvement in the success of students when compared to historical data.

Summary

Predictive modeling combines the science of statistics with the art of analysis. The intermediate goal is to blend data generation with human intervention with student success as the final end. The methodology described in this chapter has been designed to achieve both the intermediate and ultimate goals. The results of the predictive models and the implementation of early alerts through the faculty portal will be examined in the subsequent chapter. It is anticipated that the intervention strategies described earlier will translate into improved course grades as a measure of student success.

Chapter IV

RESULTS

Introduction

The purpose of this research is to develop and defend the answer in response to the research question: Can predictive modeling be used to create actionable student intelligence to improve the grades in key English and math classes resulting in higher retention rates of traditional first-year students? This research project will examine new models and approaches to student learning and success by concentrating on the first-year experience of beginning freshmen at Valdosta State University utilizing data from 2008-2014. With a fall freshman class ranging from 1,500 to 2,500 new students, the sample size is large enough to produce a much smaller confidence interval/sampling error, yet small enough to work with individual departments and faculty to implement and monitor the effect of changes employed through the use of predictive metrics and active intervention. The predictive metrics developed for this model use three specific indicators: (1) standardized test scores from the SAT or ACT, (2) high school grade point average and (3) where the student's high school ranks in relation to the other high schools in the state of Georgia.

Predictive Modeling

In analyzing the data, the formula for scoring the students into cohorts based on the three criteria (SAT score, high school GPA, and high school rank) was applied to 5 years of entering freshmen student data. This yields a total population of students being

analyzed at 17,600 which should be a large enough population to garner validity and accept or deny the null hypothesis. The data on the students were analyzed in three ways: At-Risk-General, At-Risk-Reading, and At-Risk Math as independent variables and grades (grades are defined as A,B,C = Pass and D,F,W = Fail) as the dependent variable. The three models were:

1. Using the student's SAT Verbal + high school GPA + high school rank, this study attempted to find significant predictability of grades within reading-based courses with a focus primarily on DFW rates. (At-Risk-English)
2. Using the student's SAT Math + high school GPA + high school rank, this study attempted to find significant predictability of grades within math-based courses with a focus primarily on DFW rates. (At-Risk-Math)
3. Using the student's total SAT scores + high school GPA + high school rank, this study attempted to find significant predictability of college grades. (At-Risk-General)

Table 9: Descriptive Statistics for Models 1 through 3

Descriptive Statistics			
At-Risk-English	Mean	Std. Deviation	N
Pass DFW Revised	0.22	0.417	31189
HSRank	2.56	1.1086	31189
HSGPA	2.52	1.1103	31189
SATV	2.576	1.0896	31189
At-Risk-Math	Mean	Std. Deviation	N
Pass DFW Revised	0.37	0.483	12438
HSRank	2.506	1.1376	12438
HSGPA	2.452	1.1063	12438
SATM	2.473	1.1344	12438
At-Risk-General	Mean	Std. Deviation	N
Pass DFW Revised	0.24	0.427	56357
HSRank	2.526	1.107	56357
HSGPA	2.507	1.1114	56357
SATT	2.583	1.1799	56357

In Table 9 the descriptive statistics show the mean for each variable, for Pass DFW Revised that represents the failure rate for the course (DFW grades are coded with a 1 and passing grades are coded zero). For example in the At-Risk-General 0.24 shows that of the 56,357 grades 24% were DFWs. Also with regard to the other variables (high school rank, high school GPA, and SAT) they are all represented in quartiles with a minimum of 1 and maximum of 4. The value of 9 for no high school rank was removed from the regression analysis (Blank-1-1 or Blank-4-4).

Tables 10 to 19 illustrate the grade performance, retention from fall to fall, and the credit hours earned for each of the student indices, 1-1-1 through 4-4-4 (9-1-1 through 9-4-4 for which the 9 represents no high school rank for example, out-of-state high school) for each of the three models listed above. The tables also break out the analysis by Open

(courses that are not English or math-based ex. music), English and math-based courses, and an analysis of the total. In studying these results it was clear that the variables that were analyzed were very predictive when looking at students' likelihood to pass or fail courses. The tables also show the percent retained to the following fall semester. Within the tables, the shading of green represent students with minimal risk for failure and the shading of red represents very high risk of failure. In the following tables the results demonstrate how these variable impact English-based courses.

Table 10: High School Rank of 1 and Performance in English-based Courses

HS Rank, HS GPA, SAT	Retained to Following	English										
	% Retained	Pass DFW Revised						Out of 10 Students		Credit Hours		
		Pass	DFW	Total	% Passed	% DFW	Mean GPA	Pass	DFW	Hours Enrolled	Hours Earned	% Earned
Total	68.5%	24,220	6,969	31,189	77.7%	22.3%	2.60	7	3	90,421	77,468	85.7%
Total	68.5%	24,220	6,969	31,189	77.7%	22.3%	2.60	7	3	90,421	77,468	85.7%
Total	68.5%	24,220	6,969	31,189	77.7%	22.3%	2.60	7	3	90,421	77,468	85.7%
1,1,1	83.2%	382	20	402	95.0%	5.0%	3.49	9	1	1,165	1,121	96.2%
1,1,2	73.1%	275	26	301	91.4%	8.6%	3.22	9	1	869	813	93.6%
1,1,3	70.8%	231	23	254	90.9%	9.1%	3.16	9	1	732	689	94.1%
1,1,4	77.4%	74	8	82	90.2%	9.8%	3.15	9	1	239	224	93.7%
1,2,1	73.1%	388	45	433	89.6%	10.4%	3.08	8	2	1,260	1,182	93.8%
1,2,2	74.4%	425	63	488	87.1%	12.9%	2.99	8	2	1,397	1,261	90.3%
1,2,3	71.5%	391	48	439	89.1%	10.9%	2.92	8	2	1,270	1,190	93.7%
1,2,4	65.6%	310	66	376	82.4%	17.6%	2.73	8	2	1,068	955	89.4%
1,3,1	62.3%	280	74	354	79.1%	20.9%	2.71	7	3	1,023	879	85.9%
1,3,2	70.1%	355	90	445	79.8%	20.2%	2.57	7	3	1,288	1,139	88.4%
1,3,3	65.3%	440	108	548	80.3%	19.7%	2.63	8	2	1,574	1,382	87.8%
1,3,4	65.6%	319	125	444	71.8%	28.2%	2.31	7	3	1,306	1,093	83.7%
1,4,1	54.6%	216	95	311	69.5%	30.5%	2.24	6	4	900	718	79.8%
1,4,2	57.8%	344	172	516	66.7%	33.3%	2.14	6	4	1,499	1,164	77.7%
1,4,3	57.4%	492	278	770	63.9%	36.1%	2.04	6	4	2,224	1,720	77.3%
1,4,4	60.4%	382	226	608	62.8%	37.2%	2.08	6	4	1,747	1,356	77.6%
Rank 1	66.5%	5,304	1,467	6,771	78.3%	21.7%	2.72	7	3	19,561	16,886	86.3%

In Table 10 it is clear that students who attend the most rigorous high schools in the state of Georgia perform well in English-based courses when factoring in the SAT Verbal (average ACT English and Reading) with the high school rank and high school GPA. Another observation is the pass percentage starts at the top with the 1-1-1 students

at 95% and gradually falls to 62.8% with the 1-4-4 students. Only two categories (1-4-3 and 1-4-4) fell below the 65th percentage with regard to pass rates and none fell below 60%. High school rank of 1 had an overall pass rate of 78.3%. With regard to retention rates, students whose high school GPA fell into category 4 had the lowest retention rates regardless of the SAT scores with all at 60.4% or below.

Table 11: High School Rank of 2 and Performance in English-based Courses

HS Rank, HS GPA, SAT	Retained to Following % Retained	English										
		Pass DFW Revised						Out of 10 Students		Credit Hours		
		Pass	DFW	Total	% Passed	% DFW	Mean GPA	Pass	DFW	Hours Enrolled	Hours Earned	% Earned
Total	68.5%	24,220	6,969	31,189	77.7%	22.3%	2.60	7	3	90,421	77,468	85.7%
Total	68.5%	24,220	6,969	31,189	77.7%	22.3%	2.60	7	3	90,421	77,468	85.7%
Total	68.5%	24,220	6,969	31,189	77.7%	22.3%	2.60	7	3	90,421	77,468	85.7%
Rank 1	66.5%	5,304	1,467	6,771	78.3%	21.7%	2.72	7	3	19,561	16,886	86.3%
2,1,1	80.8%	761	52	813	93.6%	6.4%	3.45	9	1	2,383	2,281	95.7%
2,1,2	85.4%	540	35	575	93.9%	6.1%	3.29	9	1	1,662	1,585	95.4%
2,1,3	78.4%	441	38	479	92.1%	7.9%	3.18	9	1	1,376	1,298	94.3%
2,1,4	78.1%	182	22	204	89.2%	10.8%	2.95	8	2	597	554	92.8%
2,2,1	71.7%	542	82	624	86.9%	13.1%	3.01	8	2	1,827	1,658	90.7%
2,2,2	76.2%	514	69	583	88.2%	11.8%	2.93	8	2	1,699	1,583	93.2%
2,2,3	81.8%	477	97	574	83.1%	16.9%	2.73	8	2	1,668	1,494	89.6%
2,2,4	71.1%	476	88	564	84.4%	15.6%	2.72	8	2	1,619	1,459	90.1%
2,3,1	63.7%	287	87	374	76.7%	23.3%	2.54	7	3	1,097	900	82.0%
2,3,2	66.8%	495	125	620	79.8%	20.2%	2.61	7	3	1,804	1,557	86.3%
2,3,3	69.8%	545	162	707	77.1%	22.9%	2.47	7	3	2,048	1,771	86.5%
2,3,4	68.3%	421	150	571	73.7%	26.3%	2.36	7	3	1,663	1,416	85.1%
2,4,1	56.6%	220	102	322	68.3%	31.7%	2.25	6	4	948	745	78.6%
2,4,2	57.7%	321	155	476	67.4%	32.6%	2.05	6	4	1,376	1,132	82.3%
2,4,3	59.5%	511	262	773	66.1%	33.9%	2.10	6	4	2,215	1,783	80.5%
2,4,4	58.6%	427	292	719	59.4%	40.6%	1.98	5	5	2,103	1,530	72.8%
Rank 2	70.4%	7,160	1,818	8,978	79.8%	20.2%	2.66	7	3	26,085	22,746	87.2%

In Table 11 it is still clear that students who attend the more rigorous high schools in the state of Georgia perform well in English-based courses when factoring in the SAT Verbal (average ACT English and Reading) with the high school rank and high school GPA. Another observation is the pass percentage starts at the top with the 2-1-1 students at 93.6% and gradually falls to 59.4% with the 2-4-4 students who would be considered high risk for failure. High school rank 2 had an overall pass rate of 79.8%.

With regard to retention rates, students whose high school GPA fell into category 4 again had the lowest retention rates regardless of the SAT scores with all below 60%.

Table 12: High School Rank of 3 and Performance in English-based Courses

HS Rank, HS GPA, SAT	Retained to Following % Retained	English										
		Pass DFW Revised						Out of 10 Students		Credit Hours		
		Pass	DFW	Total	% Passed	% DFW	Mean GPA	Pass	DFW	Hours Enrolled	Hours Earned	% Earned
Total	68.5%	24,220	6,969	31,189	77.7%	22.3%	2.60	7	3	90,421	77,468	85.7%
Total	68.5%	24,220	6,969	31,189	77.7%	22.3%	2.60	7	3	90,421	77,468	85.7%
Total	68.5%	24,220	6,969	31,189	77.7%	22.3%	2.60	7	3	90,421	77,468	85.7%
Rank 1	66.5%	5,304	1,467	6,771	78.3%	21.7%	2.72	7	3	19,561	16,886	86.3%
Rank 2	70.4%	7,160	1,818	8,978	79.8%	20.2%	2.66	7	3	26,085	22,746	87.2%
3,1,1	78.2%	511	40	551	92.7%	7.3%	3.35	9	1	1,608	1,529	95.1%
3,1,2	77.4%	378	44	422	89.6%	10.4%	3.11	8	2	1,220	1,135	93.0%
3,1,3	79.6%	401	64	465	86.2%	13.8%	3.02	8	2	1,346	1,219	90.6%
3,1,4	81.2%	286	34	320	89.4%	10.6%	2.96	8	2	917	865	94.3%
3,2,1	65.5%	255	67	322	79.2%	20.8%	2.70	7	3	947	824	87.0%
3,2,2	71.4%	343	68	411	83.5%	16.5%	2.64	8	2	1,180	1,072	90.8%
3,2,3	71.9%	470	122	592	79.4%	20.6%	2.60	7	3	1,720	1,503	87.4%
3,2,4	69.5%	398	109	507	78.5%	21.5%	2.52	7	3	1,452	1,267	87.3%
3,3,1	52.8%	132	69	201	65.7%	34.3%	2.13	6	4	592	452	76.4%
3,3,2	66.0%	210	87	297	70.7%	29.3%	2.29	7	3	870	704	80.9%
3,3,3	60.0%	365	187	552	66.1%	33.9%	2.20	6	4	1,586	1,231	77.6%
3,3,4	63.6%	387	178	565	68.5%	31.5%	2.15	6	4	1,628	1,324	81.3%
3,4,1	51.3%	74	38	112	66.1%	33.9%	2.18	6	4	323	241	74.6%
3,4,2	45.5%	98	113	211	46.4%	53.6%	1.54	4	6	611	384	62.8%
3,4,3	54.3%	230	167	397	57.9%	42.1%	1.91	5	5	1,150	828	72.0%
3,4,4	60.6%	297	205	502	59.2%	40.8%	1.83	5	5	1,458	1,079	74.0%
Rank 3	67.5%	4,835	1,592	6,427	75.2%	24.8%	2.44	7	3	18,608	15,657	84.1%

In Table 12 the results reveal that students who attend the less rigorous high schools in the state of Georgia perform well in English-based courses when factoring in the SAT Verbal (average ACT English and Reading) with the high school rank and high school GPA. Another observation is the way the pass percentage starts at the top with the 3-1-1 students at 92.7% and begins to fall more sharply to as low as 46.4% with the 3-4-2 students and for three groups this would be considered high risk for failure. High school rank 3 had an overall pass rate of 75.2%. With regard to retention rates, students whose high school GPA fell into category 4 again had the lowest retention rates regardless of the

SAT scores with all at 60.6% or below, with 3-4-2 as low as 45.5%.

Table 13: High School Rank of 4 and Performance in English-based Courses

HS Rank, HS GPA, SAT	Retained to Following % Retained	English										
		Pass DFW Revised						Out of 10 Students		Credit Hours		
		Pass	DFW	Total	% Passed	% DFW	Mean GPA	Pass	DFW	Hours Enrolled	Hours Earned	% Earned
Total	68.5%	24,220	6,969	31,189	77.7%	22.3%	2.60	7	3	90,421	77,468	85.7%
Total	68.5%	24,220	6,969	31,189	77.7%	22.3%	2.60	7	3	90,421	77,468	85.7%
Total	68.5%	24,220	6,969	31,189	77.7%	22.3%	2.60	7	3	90,421	77,468	85.7%
Rank 1	66.5%	5,304	1,467	6,771	78.3%	21.7%	2.72	7	3	19,561	16,886	86.3%
Rank 2	70.4%	7,160	1,818	8,978	79.8%	20.2%	2.66	7	3	26,085	22,746	87.2%
Rank 3	67.5%	4,835	1,592	6,427	75.2%	24.8%	2.44	7	3	18,608	15,657	84.1%
4,1,1	78.6%	422	31	453	93.2%	6.8%	3.35	9	1	1,316	1,245	94.6%
4,1,2	76.9%	319	33	352	90.6%	9.4%	3.13	9	1	1,024	976	95.3%
4,1,3	79.7%	363	42	405	89.6%	10.4%	2.98	8	2	1,175	1,105	94.0%
4,1,4	87.7%	266	32	298	89.3%	10.7%	2.97	8	2	864	803	92.9%
4,2,1	67.9%	172	53	225	76.4%	23.6%	2.76	7	3	645	525	81.4%
4,2,2	70.7%	270	72	342	78.9%	21.1%	2.60	7	3	990	849	85.8%
4,2,3	67.9%	359	83	442	81.2%	18.8%	2.63	8	2	1,288	1,143	88.7%
4,2,4	73.7%	340	107	447	76.1%	23.9%	2.45	7	3	1,302	1,119	85.9%
4,3,1	70.2%	94	47	141	66.7%	33.3%	2.33	6	4	406	313	77.1%
4,3,2	65.6%	167	72	239	69.9%	30.1%	2.19	6	4	691	526	76.1%
4,3,3	55.8%	236	113	349	67.6%	32.4%	2.14	6	4	1,018	826	81.1%
4,3,4	65.3%	268	131	399	67.2%	32.8%	2.10	6	4	1,166	937	80.4%
4,4,1	56.0%	34	31	65	52.3%	47.7%	1.86	5	5	191	121	63.4%
4,4,2	55.1%	83	57	140	59.3%	40.7%	1.79	5	5	415	310	74.7%
4,4,3	59.8%	138	125	263	52.5%	47.5%	1.70	5	5	762	546	71.7%
4,4,4	50.0%	200	154	354	56.5%	43.5%	1.85	5	5	1,021	700	68.6%
Rank 4	69.1%	3,731	1,183	4,914	75.9%	24.1%	2.43	7	3	14,274	12,044	84.4%

In Table 13 a very similar pattern emerged with regard to the bottom quartile high schools in the state of Georgia with regard to performance in English-based courses when factoring in the SAT Verbal (average ACT English and Reading) with the high school rank and high school GPA. Another observation was in the way the pass percentage started at the top with the 4-1-1 students at 93.2% and gradually fell to a low of 52.3% with the 4-4-1 students. Both of these tables (Table 12 and Table 13) show that the GPA and standardized test scores have more impact on the students' performance with regard to English than does the high school rank. These categories (4-4-4) allow one to begin to see the predictability of the model. The overall pass rates for each high

school rank were very similar (1 = 78.37%, 2 = 79.87%, 3 = 75.27% and 4 = 75.97%) and consistently the lowest GPAs represented the lowest retention and pass rates.

Table 14: No High School Rank (9-X-X) and Pass Rate for English-based Courses

HS Rank, HS GPA, SAT	Retained to Following % Retained	English										
		Pass DFW Revised						Out of 10 Students		Credit Hours		
		Pass	DFW	Total	% Passed	% DFW	Mean GPA	Pass	DFW	Hours Enrolled	Hours Earned	% Earned
Total	68.5%	24,220	6,969	31,189	77.7%	22.3%	2.60	7	3	90,421	77,468	85.7%
Total	68.5%	24,220	6,969	31,189	77.7%	22.3%	2.60	7	3	90,421	77,468	85.7%
Total	68.5%	24,220	6,969	31,189	77.7%	22.3%	2.60	7	3	90,421	77,468	85.7%
Rank 1	66.5%	5,304	1,467	6,771	78.3%	21.7%	2.72	7	3	19,561	16,886	86.3%
Rank 2	70.4%	7,160	1,818	8,978	79.8%	20.2%	2.66	7	3	26,085	22,746	87.2%
Rank 3	67.5%	4,835	1,592	6,427	75.2%	24.8%	2.44	7	3	18,608	15,657	84.1%
Rank 4	69.1%	3,731	1,183	4,914	75.9%	24.1%	2.43	7	3	14,274	12,044	84.4%
9,1,1	84.7%	386	31	417	92.6%	7.4%	3.37	9	1	1,195	1,136	95.1%
9,1,2	83.3%	231	30	261	88.5%	11.5%	3.11	8	2	758	704	92.9%
9,1,3	79.4%	153	24	177	86.4%	13.6%	3.09	8	2	511	461	90.2%
9,1,4	71.4%	134	25	159	84.3%	15.7%	2.87	8	2	458	410	89.5%
9,2,1	69.9%	167	47	214	78.0%	22.0%	2.65	7	3	608	512	84.2%
9,2,2	68.9%	170	35	205	82.9%	17.1%	2.86	8	2	588	504	85.7%
9,2,3	73.3%	169	34	203	83.3%	16.7%	2.71	8	2	595	535	89.9%
9,2,4	88.0%	112	18	130	86.2%	13.8%	2.74	8	2	378	352	93.1%
9,3,1	70.6%	170	40	210	81.0%	19.0%	2.79	8	2	599	508	84.8%
9,3,2	68.7%	212	64	276	76.8%	23.2%	2.48	7	3	815	684	83.9%
9,3,3	56.0%	196	49	245	80.0%	20.0%	2.49	8	2	720	631	87.6%
9,3,4	72.6%	153	48	201	76.1%	23.9%	2.45	7	3	574	489	85.2%
9,4,1	54.5%	202	78	280	72.1%	27.9%	2.45	7	3	816	669	82.0%
9,4,2	52.0%	182	104	286	63.6%	36.4%	2.12	6	4	838	628	74.9%
9,4,3	57.5%	294	108	402	73.1%	26.9%	2.21	7	3	1,158	994	85.8%
9,4,4	63.0%	259	174	433	59.8%	40.2%	2.15	5	5	1,282	918	71.6%
Rank 9	68.4%	3,190	909	4,099	77.8%	22.2%	2.66	7	3	11,893	10,135	85.2%

In Table 14 the results were only calculated on high school GPA and the SAT verbal (average ACT English and Reading). The value of 9 was used because the student went to an out-of-state or private high school. Again it was observed that the high school GPA category 4 students have the lowest retention rates. Also, the pass rate for the no high school rank group was in line with the other categories at 77.8%. Only two of the categories 9-4-2 and 9-4-4 had pass rates below 65%.

In the following tables the results demonstrate how these variables impact math-based courses.

Table 15: High School Rank of 1 and Performance in Math-based Courses

HS Rank, HS GPA, SAT	Retained to Following % Retained	Math										
		Pass DFW Revised						Out of 10 Students		Credit Hours		
		Pass	DFW	Total	% Passed	% DFW	Mean GPA	Pass	DFW	Hours Enrolled	Hours Earned	% Earned
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
Rank 1	66.5%	1,639	905	2,544	64.4%	35.6%	2.30	6	4	7,497	5,746	76.6%
Rank 2	70.4%	2,311	1,182	3,493	66.2%	33.8%	2.21	6	4	10,173	7,975	78.4%
Rank 3	67.5%	1,540	1,079	2,619	58.8%	41.2%	1.90	5	5	7,630	5,505	72.1%
Rank 4	69.1%	1,287	835	2,122	60.7%	39.3%	1.90	6	4	6,175	4,554	73.7%
Rank 9	68.4%	1,055	605	1,660	63.6%	36.4%	2.27	6	4	4,882	3,654	74.8%
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
1,1,1	80.3%	193	12	205	94.1%	5.9%	3.37	9	1	596	574	96.3%
1,1,2	66.0%	100	10	110	90.9%	9.1%	3.25	9	1	329	302	91.8%
1,1,3	77.6%	57	10	67	85.1%	14.9%	3.03	8	2	194	173	89.2%
1,1,4	87.9%	28	13	41	68.3%	31.7%	2.14	6	4	115	87	75.7%
1,2,1	72.0%	204	51	255	80.0%	20.0%	2.81	8	2	736	641	87.1%
1,2,2	73.8%	122	49	171	71.3%	28.7%	2.48	7	3	501	417	83.2%
1,2,3	68.1%	109	43	152	71.7%	28.3%	2.30	7	3	443	372	84.0%
1,2,4	71.1%	77	29	106	72.6%	27.4%	2.28	7	3	307	265	86.3%
1,3,1	68.2%	129	50	179	72.1%	27.9%	2.27	7	3	525	411	78.3%
1,3,2	67.1%	132	48	180	73.3%	26.7%	2.31	7	3	530	441	83.2%
1,3,3	63.2%	88	91	179	49.2%	50.8%	1.66	4	6	537	375	69.8%
1,3,4	65.8%	81	74	155	52.3%	47.7%	1.77	5	5	473	313	66.2%
1,4,1	61.1%	60	79	139	43.2%	56.8%	1.56	4	6	404	269	66.6%
1,4,2	53.6%	107	113	220	48.6%	51.4%	1.65	4	6	653	414	63.4%
1,4,3	58.5%	85	128	213	39.9%	60.1%	1.41	3	7	637	365	57.3%
1,4,4	60.3%	67	105	172	39.0%	61.0%	1.36	3	7	517	327	63.2%
Rank 1	66.5%	1,639	905	2,544	64.4%	35.6%	2.23	6	4	7,497	5,746	76.6%

Table 15 looks at performance in math, taking into account the high school rank, high school GPA, and the math portions of the standardized test (SAT and ACT). With regard to math and top quartile high schools in the state of Georgia, there was a very similar pattern to what was observed with English; however, there were more categories showing risk for math, with the six bottom categories from 1-3-3 to 1-4-4 all showing far greater risk of not passing the courses. The 1-4-4 students had a 39% likelihood of passing math-based courses. All observed less than 65% retention rates for students

whose high school GPA fell into the 4 category. The average pass rate for the top ranked high schools in math (64.4%) were more than 10% lower than what was observed with English.

Table 16: High School Rank of 2 and Performance in Math-based Courses

HS Rank, HS GPA, SAT	Retained to Following % Retained	Math										
		Pass DFW Revised						Out of 10 Students		Credit Hours		
		Pass	DFW	Total	% Passed	% DFW	Mean GPA	Pass	DFW	Hours Enrolled	Hours Earned	% Earned
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
Rank 1	66.5%	1,639	905	2,544	64.4%	35.6%	2.23	6	4	7,497	5,746	76.6%
2,1,1	83.4%	368	45	413	89.1%	10.9%	3.20	8	2	1,163	1,079	92.8%
2,1,2	81.5%	204	31	235	86.8%	13.2%	3.01	8	2	685	635	92.7%
2,1,3	77.0%	104	23	127	81.9%	18.1%	2.80	8	2	365	330	90.4%
2,1,4	77.0%	75	20	95	78.9%	21.1%	2.58	7	3	263	231	87.8%
2,2,1	73.9%	241	63	304	79.3%	20.7%	2.77	7	3	876	758	86.5%
2,2,2	78.6%	190	58	248	76.6%	23.4%	2.65	7	3	725	629	86.8%
2,2,3	75.8%	151	67	218	69.3%	30.7%	2.27	6	4	643	521	81.0%
2,2,4	71.8%	126	55	181	69.6%	30.4%	2.06	6	4	542	452	83.4%
2,3,1	64.3%	124	60	184	67.4%	32.6%	2.28	6	4	532	419	78.8%
2,3,2	65.7%	124	90	214	57.9%	42.1%	1.89	5	5	629	466	74.1%
2,3,3	66.2%	116	87	203	57.1%	42.9%	1.86	5	5	598	451	75.4%
2,3,4	72.2%	124	129	253	49.0%	51.0%	1.54	4	6	754	519	68.8%
2,4,1	53.9%	72	62	134	53.7%	46.3%	1.76	5	5	390	279	71.5%
2,4,2	59.1%	88	99	187	47.1%	52.9%	1.66	4	6	543	349	64.3%
2,4,3	59.3%	98	119	217	45.2%	54.8%	1.45	4	6	640	395	61.7%
2,4,4	59.1%	106	174	280	37.9%	62.1%	1.27	3	7	825	462	56.0%
Rank 2	70.4%	2,311	1,182	3,493	66.2%	33.8%	2.19	6	4	10,173	7,975	78.4%

In Table 16 with regard to the high school rank of 2 in the state of Georgia, the risk of failing math was strong in 7 of the 16 categories shown as very high risk highlighted in red. The 2-4-4 students post a 37.9% likelihood of passing math-based courses. The retention pattern with regard to the high school GPA category of 4 continues to be a concern with all below 60%.

Table 17: High School Rank of 3 and Performance in Math-based Courses

HS Rank, HS GPA, SAT	Retained to Following % Retained	Math										
		Pass DFW Revised						Out of 10 Students		Credit Hours		
		Pass	DFW	Total	% Passed	% DFW	Mean GPA	Pass	DFW	Hours Enrolled	Hours Earned	% Earned
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
Rank 1	66.5%	1,639	905	2,544	64.4%	35.6%	2.23	6	4	7,497	5,746	76.6%
Rank 2	70.4%	2,311	1,182	3,493	66.2%	33.8%	2.19	6	4	10,173	7,975	78.4%
3,1,1	75.4%	232	30	262	88.5%	11.5%	3.08	8	2	761	700	92.0%
3,1,2	82.1%	178	42	220	80.9%	19.1%	2.83	8	2	620	541	87.3%
3,1,3	80.0%	128	42	170	75.3%	24.7%	2.51	7	3	501	426	85.0%
3,1,4	79.3%	74	20	94	78.7%	21.3%	2.49	7	3	274	238	86.9%
3,2,1	69.2%	100	47	147	68.0%	32.0%	2.30	6	4	422	327	77.5%
3,2,2	73.1%	119	79	198	60.1%	39.9%	2.03	6	4	575	431	75.0%
3,2,3	66.7%	135	76	211	64.0%	36.0%	1.98	6	4	604	459	76.0%
3,2,4	71.1%	102	94	196	52.0%	48.0%	1.69	5	5	570	406	71.2%
3,3,1	53.0%	50	37	87	57.5%	42.5%	1.71	5	5	253	188	74.3%
3,3,2	55.8%	78	69	147	53.1%	46.9%	1.70	5	5	427	294	68.9%
3,3,3	63.2%	83	92	175	47.4%	52.6%	1.63	4	6	520	347	66.7%
3,3,4	65.7%	108	138	246	43.9%	56.1%	1.45	4	6	732	454	62.0%
3,4,1	48.7%	21	29	50	42.0%	58.0%	1.32	4	6	140	77	55.0%
3,4,2	52.8%	24	54	78	30.8%	69.2%	1.21	3	7	233	127	54.5%
3,4,3	55.7%	52	96	148	35.1%	64.9%	1.23	3	7	436	235	53.9%
3,4,4	56.8%	56	134	190	29.5%	70.5%	0.99	2	8	562	255	45.4%
Rank 3	67.5%	1,540	1,079	2,619	58.8%	41.2%	1.88	5	5	7,630	5,505	72.1%

In Table 17 with regard to the high school rank of 3 in the state of Georgia, the risk of failing math was strong in 9 of the 16 categories shown as very high risk highlighted in red. The 3-4-4 students post a 29.5% likelihood of passing math-based courses, of 190 who attempted math only 56 passed with a grade of C or better. The retention pattern with regard to the high school GPA category of 4 continues to be a concern with all below 60%, even the higher SAT math score in the 3-4-1 group did not seem to impact retention as this group posted a 48.7% rate which is the lowest in this table.

Table 18: High School Rank of 4 and Performance in Math-based Courses

HS Rank, HS GPA, SAT	Retained to Following % Retained	Math										
		Pass DFW Revised						Out of 10 Students		Credit Hours		
		Pass	DFW	Total	% Passed	% DFW	Mean GPA	Pass	DFW	Hours Enrolled	Hours Earned	% Earned
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
Rank 1	66.5%	1,639	905	2,544	64.4%	35.6%	2.30	6	4	7,497	5,746	76.6%
Rank 2	70.4%	2,311	1,182	3,493	66.2%	33.8%	2.21	6	4	10,173	7,975	78.4%
Rank 3	67.5%	1,540	1,079	2,619	58.8%	41.2%	1.90	5	5	7,630	5,505	72.1%
Rank 4	69.1%	1,287	835	2,122	60.7%	39.3%	1.90	6	4	6,175	4,554	73.7%
Rank 9	68.4%	1,055	605	1,660	63.6%	36.4%	2.27	6	4	4,882	3,654	74.8%
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
Rank 1	66.5%	1,639	905	2,544	64.4%	35.6%	2.23	6	4	7,497	5,746	76.6%
Rank 2	70.4%	2,311	1,182	3,493	66.2%	33.8%	2.19	6	4	10,173	7,975	78.4%
Rank 3	67.5%	1,540	1,079	2,619	58.8%	41.2%	1.88	5	5	7,630	5,505	72.1%
4,1,1	79.7%	242	33	275	88.0%	12.0%	3.10	8	2	772	710	92.0%
4,1,2	85.0%	161	26	187	86.1%	13.9%	3.00	8	2	541	498	92.1%
4,1,3	74.4%	108	30	138	78.3%	21.7%	2.48	7	3	398	333	83.7%
4,1,4	81.1%	80	32	112	71.4%	28.6%	2.46	7	3	340	281	82.6%
4,2,1	80.2%	82	23	105	78.1%	21.9%	2.52	7	3	299	258	86.3%
4,2,2	65.5%	90	43	133	67.7%	32.3%	2.23	6	4	377	279	74.0%
4,2,3	71.5%	89	66	155	57.4%	42.6%	1.94	5	5	459	337	73.4%
4,2,4	67.9%	123	93	216	56.9%	43.1%	1.76	5	5	643	477	74.2%
4,3,1	55.3%	36	30	66	54.5%	45.5%	1.96	5	5	190	137	72.1%
4,3,2	71.0%	37	42	79	46.8%	53.2%	1.69	4	6	229	156	68.1%
4,3,3	65.4%	67	91	158	42.4%	57.6%	1.37	4	6	457	299	65.4%
4,3,4	59.9%	68	126	194	35.1%	64.9%	1.29	3	7	575	313	54.4%
4,4,1	47.4%	11	9	20	55.0%	45.0%	1.67	5	5	58	37	63.8%
4,4,2	65.8%	18	26	44	40.9%	59.1%	1.58	4	6	127	68	53.5%
4,4,3	61.4%	24	62	86	27.9%	72.1%	1.10	2	8	252	133	52.8%
4,4,4	48.7%	51	103	154	33.1%	66.9%	1.08	3	7	458	238	52.0%
Rank 4	69.1%	1,287	835	2,122	60.7%	39.3%	1.95	6	4	6,175	4,554	73.7%

In Table 18 with regard to the bottom quartile high schools in the state of Georgia, the risk of failing increased as we move down the categories, and 10 of the 16 categories show very high risk as highlighted in red. The 4-4-3 students posted an abysmal 27.9% likelihood of passing math-based courses. The overall pass rate for this group is 60.7% which was 15% lower than the comparable English pass rates. This further supported the idea that the use of predictive modeling can be leveraged to flag students as they enter the institution as at-risk for math-based courses.

Table 19: High School Rank of 3 Comparing Verbal Test Scores to Math-based

Courses

HS Rank, HS GPA, SAT	Retained to Following % Retained	Math										
		Pass DFW Revised						Out of 10 Students		Credit Hours		
		Pass	DFW	Total	% Passed	% DFW	Mean GPA	Pass	DFW	Hours Enrolled	Hours Earned	% Earned
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
Total	68.5%	7,832	4,606	12,438	63.0%	37.0%	2.17	6	4	36,357	27,434	75.5%
Rank 1	66.5%	1,639	905	2,544	64.4%	35.6%	2.30	6	4	7,497	5,746	76.6%
Rank 2	70.4%	2,311	1,182	3,493	66.2%	33.8%	2.21	6	4	10,173	7,975	78.4%
3,1,1	78.2%	184	36	220	83.6%	16.4%	2.91	8	2	636	568	89.3%
3,1,2	77.4%	155	32	187	82.9%	17.1%	2.87	8	2	533	471	88.4%
3,1,3	79.6%	150	36	186	80.6%	19.4%	2.73	8	2	548	474	86.5%
3,1,4	81.2%	123	30	153	80.4%	19.6%	2.67	8	2	439	392	89.3%
3,2,1	65.5%	68	51	119	57.1%	42.9%	2.04	5	5	345	219	63.5%
3,2,2	71.4%	100	68	168	59.5%	40.5%	1.96	5	5	488	382	78.3%
3,2,3	71.9%	179	81	260	68.8%	31.2%	2.15	6	4	759	590	77.7%
3,2,4	69.5%	109	96	205	53.2%	46.8%	1.77	5	5	579	432	74.6%
3,3,1	52.8%	38	38	76	50.0%	50.0%	1.58	5	5	228	146	64.0%
3,3,2	66.0%	59	63	122	48.4%	51.6%	1.62	4	6	356	238	66.9%
3,3,3	60.0%	102	121	223	45.7%	54.3%	1.54	4	6	651	412	63.3%
3,3,4	63.6%	120	114	234	51.3%	48.7%	1.63	5	5	697	487	69.9%
3,4,1	51.3%	17	19	36	47.2%	52.8%	1.68	4	6	108	57	52.8%
3,4,2	45.5%	23	71	94	24.5%	75.5%	0.99	2	8	271	128	47.2%
3,4,3	54.3%	42	103	145	29.0%	71.0%	1.06	2	8	434	225	51.8%
3,4,4	60.6%	71	120	191	37.2%	62.8%	1.20	3	7	558	284	50.9%
Rank 3	67.5%	1,540	1,079	2,619	58.8%	41.2%	1.90	5	5	7,630	5,505	72.1%

In Table 19 the results were very interesting for the SAT given that there were more groups that show high-risk (non-success) with regard to math-based courses. All SAT verbal ranks of 1 and 2 were at-risk with the exception of 3-1-1 and 3-1-2. In Table 19 all but five categories show high-risk for non-success in math-based courses when using the SAT verbal compared to the all but seven categories when looking at Table 17 showing the same high school rank and high school GPA but using the math standardized test scores. Note the drop in pass rates from the top high school GPA to the second tier was about 20%. Also it was noted that the categories of high school GPA 1 showed no risk within both high school rank 3 and high school rank 4.

For the next stage of this study, all 17,600 students were analyzed by the three models for General, English and Math using the scores from all categories (1-1-1 to 4-4-4 including blank for no high school rank in 9-1-1 to 9-4-4) independently through a multivariate regression analysis. The mean scores for the students were .24, .22 and .37 which represents the percentage of grades that were failing for each model. A student with a .33 failed one out of three classes. For the purpose of this study, analysis one was referred to as “At-risk General.” All 17,600 students’ grades were individually converted to a range of 0 to 1 to calculate the dependent variable for a total of 52,329 grades in introductory coursework. The dependent variable in this analysis was failing grades (as defined by D, F and W grades = 1 with passing grades being = 0) and the independent variables were high school rank, high school GPA and standardized test scores (all of which were quartiled). Using ordinary least squares regression, the General model looks at all grades while the English model only looked at grades in English and reading courses, and the Math model looked at math and science grades. In Table 20, the unstandardized regression coefficients showed that for every unit increase in high school GPA (student moves from the second to the third quartile) the DFW rates increased by 9% in the General at-risk model as well as the English at-risk model. For every unit increase in the high school GPA quartile, the DFW rate in the Math at-risk model increased by 13%. As we move from an SAT Math quartile of 1 to 2 or 3 to 4, the DFW rate in the Math at-risk model increased by 5.5%. The results are summarized in Table 20.

Table 20: Regression Model for Predictive Analysis

Independent Variables	General At-risk Model	English At-risk Model	Math At-risk Model
HS Rank	.002** (.001) t=2.692	.002** (.001) t=2.61	.002 (.002) t=1.169
HS GPA	.092** (.002) t=56.359	.090** (.002) t=42.703	.130** (.004) t=34.333
SAT Total	.015** (.002) t=9.811		
SAT English		.017** (.002) t=7.847	
SAT Math			.055** (.004) t=15.004
F Scores	1307.247**	713.766**	597.626**
Adjusted r ²	.065	.064	.126
N	56357	31189	12438

Cell entries are unstandardized regression coefficients, data in parentheses show the standard error. *p < .05 and **p < .01

From the results one can see that the codes that were calculated for the students “At-risk General” were highly predictive of college GPA at the end of the students’ first year. If you look at the ANOVA it was clear that with 3 degrees of freedom, F value of 1307.247 and a significance of .000, that at least one of these factors produced a highly predictive result and the model was accepted based on the above significance and predictability of the dependent variable DFW rates. Also an explained variance of 6.5%

for both the general and English model and an explained variance of 12.6% for the math model were seen in the adjusted r^2 . This left a large amount of the variance unexplained. It was also important to note the dominance of the high school GPA in the model, while the other variables show significance, it is clear the high school GPA is the most powerful predictor. This was revealed by the standardized regression coefficients which were largest for high school GPA (ranging from .24 to .30 across the three models) although not depicted in Table 20. In Appendix B, the predictability of each coded category is shown. This data was used to assign future students codes that will provide faculty and academic support staff information about students and their likelihood to struggle with classwork. Also the t scores further validate this model as shown in Table 20.

For the purposes of delineation the next two data analyses in Table 20 split student grades into English-based courses (history, English, sociology, political science, etc.) for the English At-risk model and math-based courses (math, physics, chemistry, etc.) for the Math At-risk model. In the analysis of student performance with regard to grades in reading-based courses, the coding metric for At-risk Reading represent the independent variable and grades in English-based courses the dependent variable. The N for this analysis is 17,600 students and 31,189 grades in introductory reading-based courses. The goal of this analysis was to provide predictive metrics for grades in reading-based courses.

In reviewing the regression model in Table 20, the at-risk English metric proved to be a significant predictor of success and/or failure in college introductory reading courses. This can be done by analyzing the ANOVA with 3 degrees of freedom

and an F value of 713.766 with a significance of .000. The unstandardized coefficients also further validate these findings with *t* scores that also showed satisfied significance. The coefficients in Table 20 for the model were .002 for high school rank, .090 for high school GPA and .017 for the SAT English. For example, a student with a 4-4-4 At-risk Reading codes will have a likelihood of failure of 43% (an average GPA 1.85) with regard to reading-based courses. This information will be used to assign students codes if it is highly likely they will struggle with regard to reading-based courses.

Finally, this study analyzed the At-risk Math model. The N for this analysis is 17,600 students and 12,438 grades (some students actually stop out of college before recording a math grade; this is why the N is lower than the total population). Those without a math grade were excluded as Null. This model was significantly predictive of students' grades in introductory math courses with a significant F value of 597.626. While the high school rank did not show statistical significance in the model, the high school GPA and the standardized test scores were both very significant. With regard to math grades, this research found that the Math at-risk model is significantly predictive of grades in introductory math-based courses. It can also be seen in the coefficients in Table 20 with high school GPA at .130 and SAT math at .055, with *t* scores that further validate the model. In Appendix B the spreadsheets show that math was the greatest challenge for the institution as students are more likely to struggle with math. For example, a student with a 4-4-4 for At-risk Math has a 67% likelihood of failure in introductory math courses (a GPA in those courses of 1.08). This compares to the 1-1-1 coded students with a 94.1% pass rate in introductory math (with an average GPA in introductory math courses of 3.37). Forty-two percent of the entering freshmen were at-risk of getting a

D,F,W in introductory math courses.

Given the potential relationship between high school GPA, high school rank and standardized test scores, there was a concern that the data may have some problems with multicollinearity. In order to test for multicollinearity a Variance Inflation Factor (VIF) was run on each of the independent variables to make sure that the scores were below 5. With independent variables such as high school GPA and high school ranking and standardized test scores, a test was conducted for whether or not the independent variables were strongly correlated with each other. All VIF scores were below 5 ranging from 1.000 to 1.090.

Given the findings in Table 20 that have been generated through this analysis of 5 years of student data, it is clear that the metrics that have been created work to show which students are at a higher likelihood to struggle based on the At-risk-variables (At-risk General, At-risk Reading, At-risk Math). This analysis proves that given the data gathered on students, the model predicted success or failure with some certainty.

Taking Math One Step Further: Analysis of MATH 1111 Courses from Fall 2005 to Fall 2014

Earned final grades of first-time, full-time freshmen (cohort students) who enrolled in a MATH 1111 at Valdosta State University (VSU) from Fall 2005 to Fall 2014 were collected. In determining the pass rates, grades of A, B, and C were classified as passing, while grades of D, F, and W were classified as not passing. Cohort students who earned a grade of NR were excluded from the analysis. Overall a total of 2 grades were excluded. In order to determine the placement of students, admission standards (high school grade point average and standardized test scores) were analyzed. All ACT

Math scores were converted to an SAT score. If a student had both an SAT Math and ACT Math, the higher of the two were taken. The scores were then given an index based on the math placement index values. If the index indicated that the student is at a high risk of not passing MATH 1111, a modified index, $\text{ModificationIndex} = (\text{HS GPA} * 500) = \text{SATmath}$, was created to determine which students needed to go into the MATH 1111 Extended Learning.

All MATH 1111 Courses

Figure 5 displays a line graph of the pass rate for MATH 1111 courses from Fall 2005 to Fall 2014, while Table 21 displays the pass rates and number of students. The pass rate has increased from 0.676 in 2005 to 0.747 in 2014, which is a 0.071 percentage point increase. With the math placement index in the second year of operation, meaning 2013 was the first year for the placement, the pass rate increased an additional 0.048 points from 2013. In order to determine whether the pass rate from 2013 to 2014 was significant, a chi-square test of independence was conducted. The relation between two terms was significantly different, $\chi^2(1, n = 1767) = 4.610, p = 0.032$. A t test was also run to determine significance and the $t = 2.286$ is significant at the $p < .025$ level in accordance with the t distribution. This means that the 2014 pass rate of MATH 1111 courses are significantly higher than the 2013 pass rate.

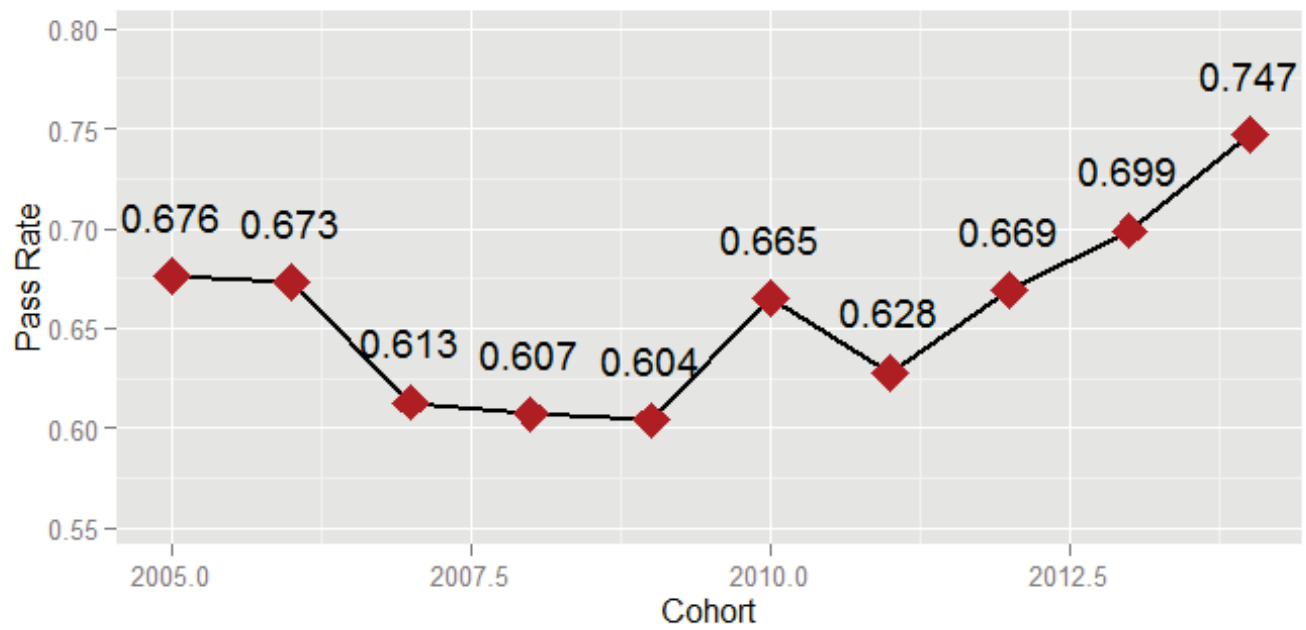


Figure 5: Pass Rates of All MATH 1111 Courses, Fall 2005-Fall 2014

Table 21: Pass Rates of All Math 1111 Courses, Fall 2005-Fall 2014

Term	Pass Rate	Number of Students
2005	0.676	782
2006	0.673	1057
2007	0.613	1110
2008	0.607	1199
2009	0.604	1441
2010	0.665	1465
2011	0.628	1397
2012	0.669	1195
2013	0.699	1001
2014	0.747	766

All MATH 1111 Extended Learning Students

After analyzing the predictive models for students from Fall 2005 to Fall 2014, the Math Department worked to place students for Fall 2014 into the equivalent MATH 1111 Extended Learning sections (which meet 5 days per week). Table 21 compares their performance (pass/fail rates) to the previous Fall semesters. The results are displayed in Figure 6 and Table 22. The pass rate from 2006-2012 was very flat and it has increased from 0.528 for the 2005 cohort to 0.69 for the 2014 cohort. This is an increase of 0.162 percentage points. Since the math placement is in its second year and the extended learning is in the first year, a pass rate increase of 0.173 was experienced. In

order to determine if the pass rate of the pilot group of the extended learning math students in 2014 was significant from like students from 2013, a chi-square test for independence was conducted. The relation between the two terms was significantly different, $\chi^2(1, n = 287) = 8.250, p = 0.004$. A t test was also run to determine significance and the t score of 3.035 is very significant at the $p < .005$ level in accordance with the t distribution. This means that the 2014 pass rate for the extended learning students are significantly higher than the 2013 pass rate.

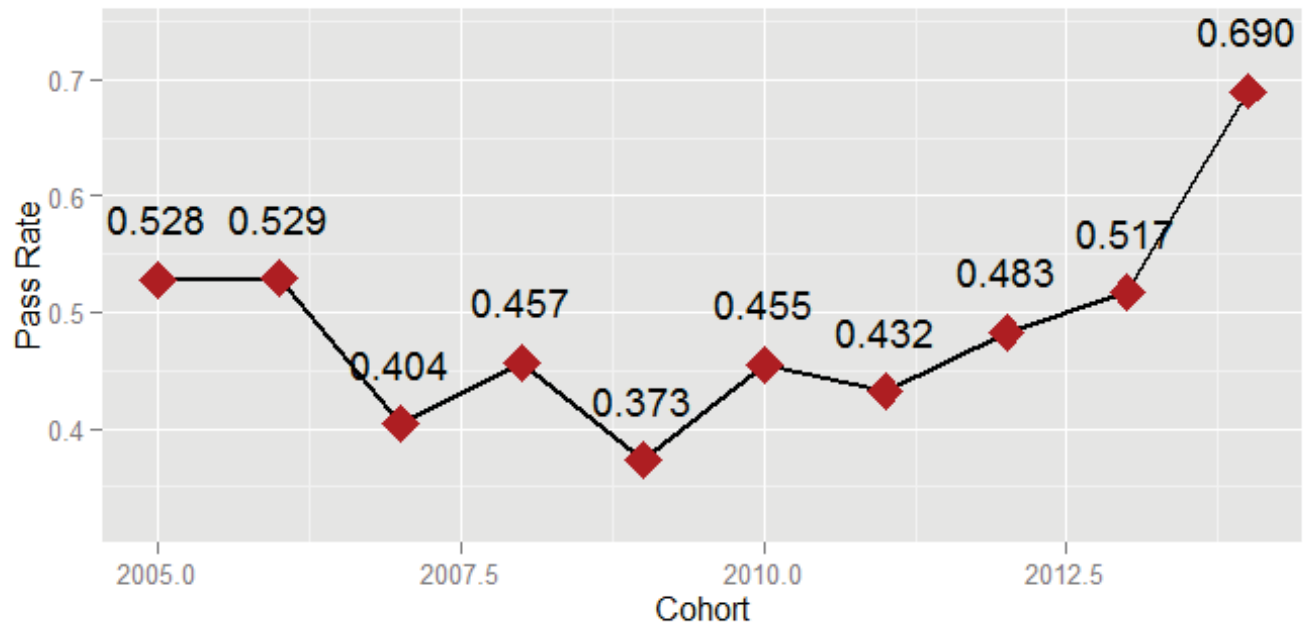


Figure 6: Pass Rate of All MATH 1111 Extended Learning Students, Fall 2005-Fall 2014

Table 22: Pass Rates of All MATH 1111 Extended Learning Students, Fall 2005-Fall 2014

Term	Pass Rate	Number
2005	0.528	142
2006	0.529	187
2007	0.404	161
2008	0.457	243
2009	0.373	252
2010	0.455	242
2011	0.432	260
2012	0.483	174
2013	0.517	145
2014	0.690	142

All MATH 1111 without Extended Learning Students

Figure 6 and Table 22 show the pass rates of the MATH 1111 students with the students who would have and were enrolled in a MATH 1111 Extended Learning removed from the analysis. The pass rate from the 2005 cohort to the 2014 cohort has experienced an increase of 0.051 points. When examining the pass rate difference from the two years of operation for the placement, the difference is 0.03 points (from the 2013 cohort to the 2014). A Chi-square test for independence indicated that the pass rate between the 2013 to the 2014 cohorts was not significantly different, $\chi^2(1, n = 1480) =$

1.490, $p = 0.222$. A t test was also run to determine significance and the t score of 1.304 is not significant at the $p < .05$ level in accordance with the t distribution.

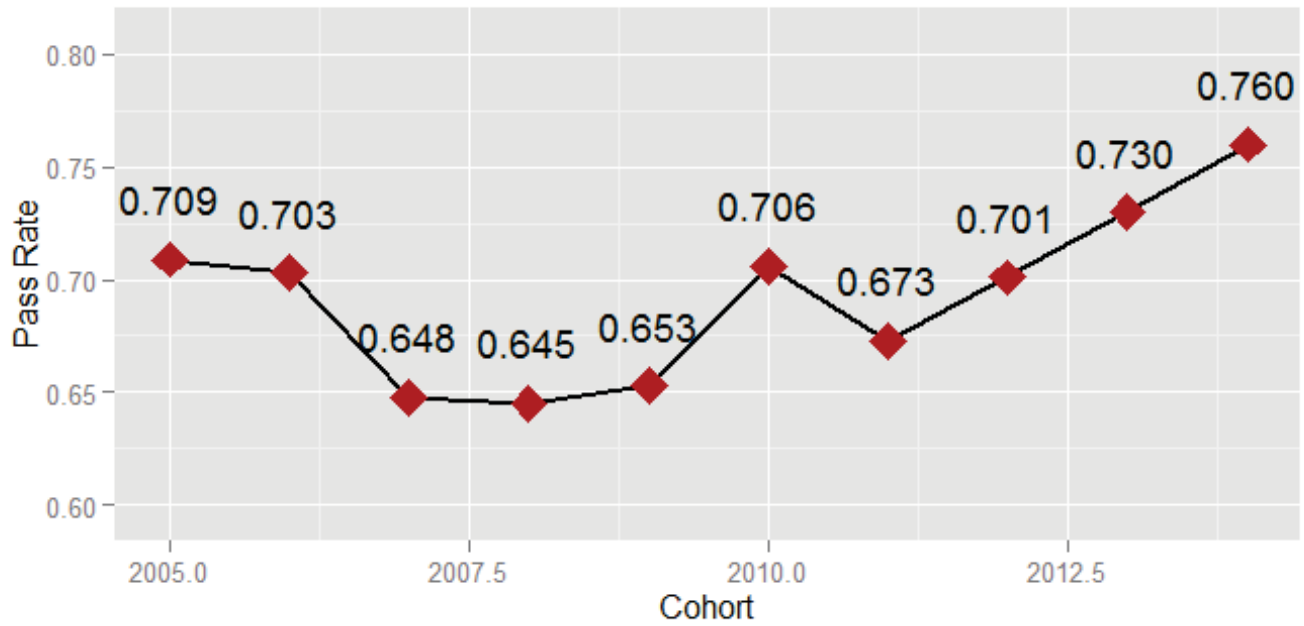


Figure 7: Pass Rate of MATH 1111 Students without MATH 1111 Extended Learning Students

Table 23: Pass Rates of MATH 1111 Students without MATH 1111 Extended Learning Students, Fall 2005-Fall 2014

Term	Pass Rate	Number
2005	0.709	640
2006	0.703	870
2007	0.648	949
2008	0.645	956
2009	0.653	1189
2010	0.706	1223
2011	0.673	1137
2012	0.701	1021
2013	0.730	856
2014	0.760	624

All Math Level One High-Risk Students

Figure 8 and Table 24 display the pass rates of the Math Level One (students who have less than a 60% chance of success) high-risk students. The pass rate for the 2014 cohort was 0.632, while the 2005 cohort's pass rate was 0.516. When examining the pass rate between the 2013 and 2014 cohorts, there was a 0.093 points increase. A Chi-square test for independence revealed that difference between the 2013 and 2014 pass rates was significantly different, $\chi^2(1, n = 743) = 6.111, p = 0.013$. A *t* test was also run to

determine significance and the t score of 2.58 is very significant at the $p < .005$ level in accordance with the t distribution.

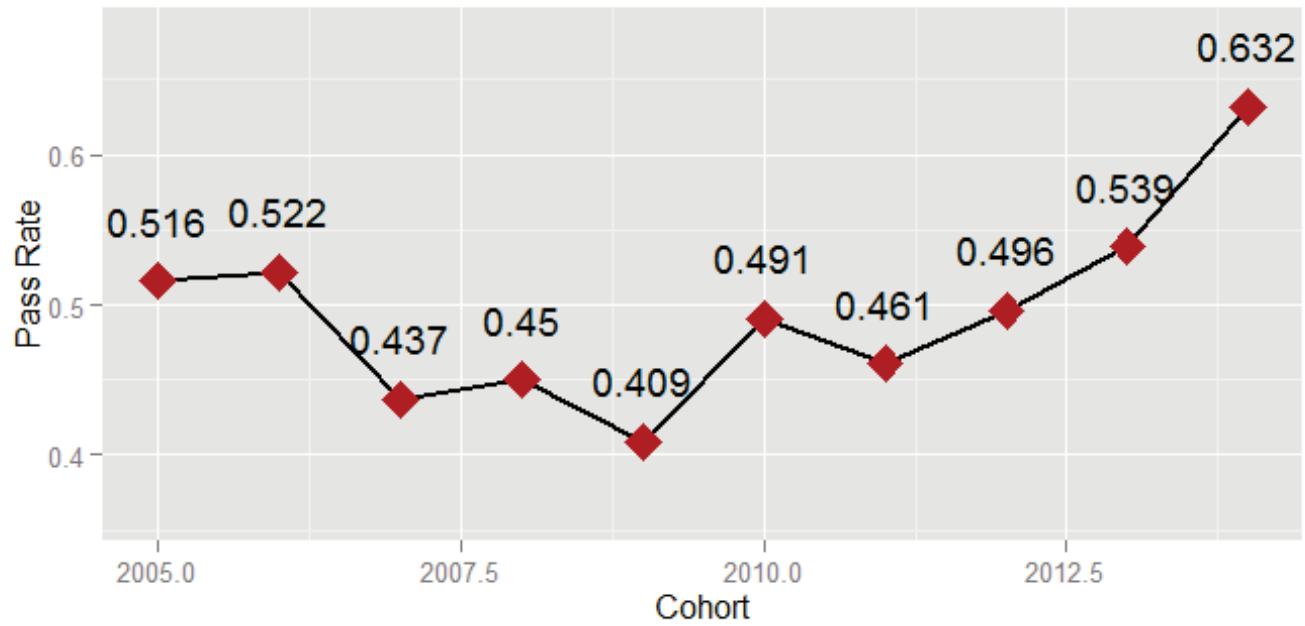


Figure 8 : Pass Rate of Math Level One High-Risk Students

Table 24: Pass Rates of Math Level One High-Risk Students

Term	Pass Rate	Number
2005	0.516	349
2006	0.522	504
2007	0.437	515
2008	0.450	576
2009	0.409	707
2010	0.491	684
2011	0.461	691
2012	0.496	522
2013	0.539	425
2014	0.632	318

All Math Level One High-Risk Students by Extended Learning Placement

Prior to the 2014 cohort, the pass rate of students who would have been placed into extended learning (the middle 50% of the high risk students) was lower than the students in the bottom and top 25% of math level one except for the 2005, 2006, and 2008 cohorts. When examining the 2014 cohort, the students who were in extended learning had a pass rate of 0.688, while students who were not placed into extended learning had a pass rate of 0.589. This is a differences of 0.099 points. A Chi-square test indicated that there was no significant difference in the 2014 pass rates of the two groups, $\chi^2(1, n = 318) = 2.912, p = 0.088$.

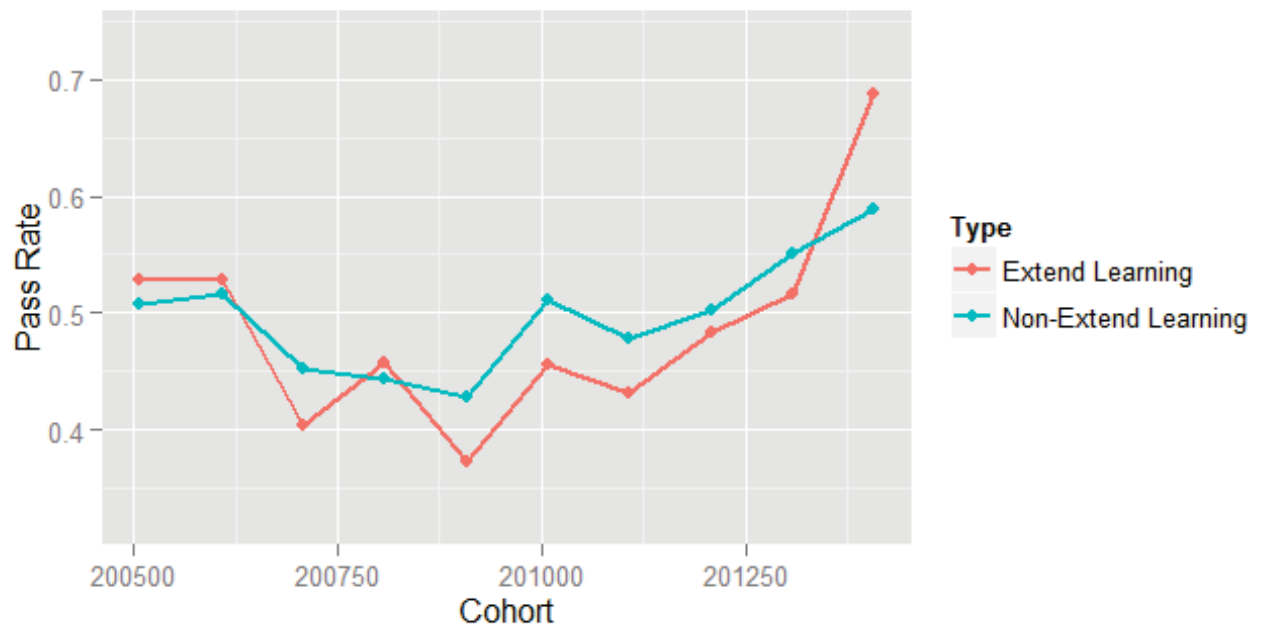


Figure 9: Pass Rate of Math Level One High-Risk Students by Extended Learning Placement

Table 25: Pass Rates of Math Level One High-Risk Students by Extended Learning Placement

Term	EL Pass Rate	EL Number	No EL Pass Rate	No EL Number
2005	0.528	142	0.507	207
2006	0.529	187	0.517	317
2007	0.404	161	0.452	354
2008	0.457	243	0.444	333
2009	0.373	252	0.429	455
2010	0.455	242	0.511	442
2011	0.432	260	0.478	431
2012	0.483	174	0.503	348
2013	0.517	145	0.550	280
2014	0.688	138	0.589	180

The Math Department is a great example of how predictive modeling can be leveraged to bring better data and information to the table when discussing student success. The data from these models drove a conversation that changed the way Math 1111 was delivered and taught to different levels of prepared students. The improvement of student learning and grades is clear from the results, students are doing better in foundational math courses.

Financial Aid Predictive Models

Financial Indicators

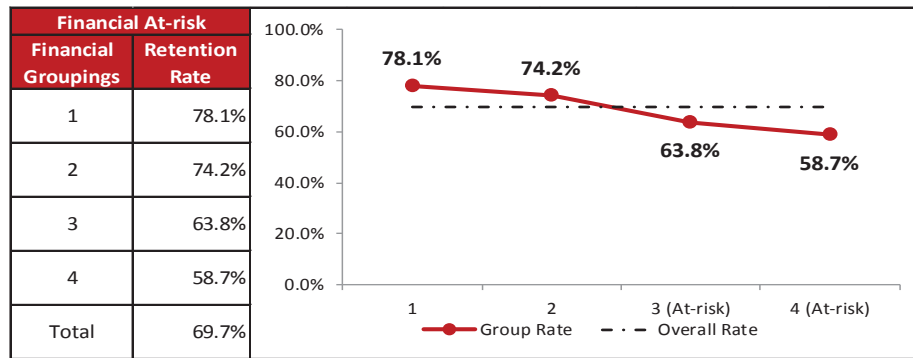
As the cost of higher education has increased over time, another piece of data that can be used in a predictive model is the financial aid data. Table 26 displays an analysis of retention rates based on the financial indicators of first-time, full-time (FTFT) freshman cohorts from Fall 2007 to Fall 2012. These financial indicators include: the HOPE Scholarship, which is an academic performance based scholarship in the state of Georgia; Pell Grant which is a needs based aid program that is based on income level; and parents' contribution is the amount of money that parents should be paying toward a child/dependents post-secondary education (the calculation considers income and dependents in the household). The findings when combining these variables and analyzing retention are that FTFT freshman who received the HOPE Scholarship, did not receive any Pell Grants, and had an expected parents' contribution at or above the average dollar amount had the highest retention rate (79.7%). At 55.7% retention rates, students who do not receive the HOPE scholarship and receive below average in both the Pell Grant dollars and the expected parents' contribution have the lowest retention rate. Financial groupings were applied by using quartiles of the retention rates.

Table 26: Retention Rate of First-time, Full-time Freshman by Financial Indicators, Fall 2007-Fall 2012 Cohorts

Financial Indicators				
HOPE Scholarship?	Pell Grant? Average \$1,705	Parents' Contribution? Average \$15,838	Retention Rate	Financial Groupings
No	At or Above Average	No	57.5%	4
No	At or Above Average	Below Average	62.7%	3
No	Below Average	Below Average	55.7%	4
No	No	No	64.0%	3
No	No	Below Average	60.8%	4
No	No	At or Above Average	64.5%	3
Yes	At or Above Average	No	75.1%	2
Yes	At or Above Average	Below Average	76.5%	1
Yes	Below Average	Below Average	67.9%	2
Yes	No	No	74.9%	2
Yes	No	Below Average	77.5%	1
Yes	No	At or Above Average	79.7%	1
Total			69.7%	

Source: VSU SRA Analysis of Data Warehouse, 2013.

When the variables are grouped by the retention rates into 4 categories, Figure 10 shows the retention rates of the newly formed groups. Group 1 (78.1%) and Group 2 (74.2%) retention rates were above the overall institutional retention rate of 69.7%. Group 3 (63.8%) and Group 4 (58.7%) retention rates were below the overall institutional retention rates. First-time full-time freshmen who are in Groups 3 and 4 are at-risk of being retained at VSU for financial reasons. Carrying the analysis a step further we can incorporate the academic data mentioned in Chapter 3 to create a better sense of how finances play a role in success.



Source: VSU SRA Analysis of Data Warehouse, 2013.

Figure 10: Retention Rates of Financial Groupings

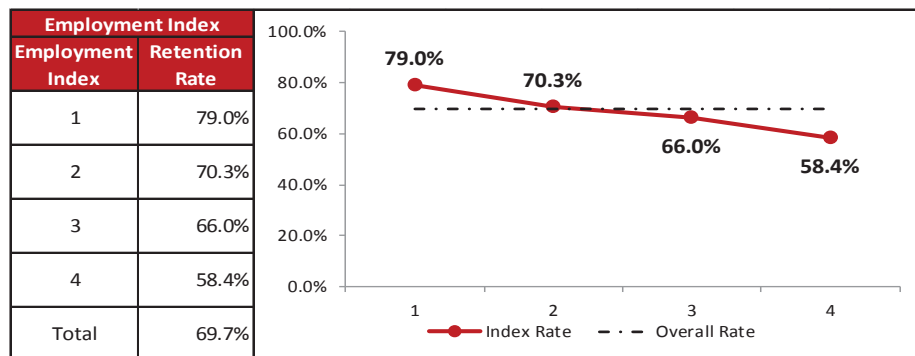
In Table 27 the findings show the retention rates of the academic and financial at-risk quartile groups. FTFT students in Academic and Financial Group 1 have the highest retention rates (84.4%). Also it can be noted that the students in Financial Group 1 continue to retain better than the institutional average; for example, Academic 2 and Financial 1 (77.1%) and Academic 3 and Financial 1 (72.6%). This shows that finances and socioeconomic background have an effect on retention. Academic and Financial Group 4 FTFT students have the lowest rate (55.2%). This finding gives way to the creation of an Employment Index which is applied by utilizing quartiles to identify students that would benefit from campus employment based on the averages of the Academic and Financial groupings. The employment index was a field created with the intent to start an on-campus job program for entering freshmen that had both financial and academic risk indicators. This would have been used as an intervention strategy for students that fell into these categories; however, due to budget cuts at the institution the program was not funded.

Table 27: Retention Rates by Academic and Financial Groupings

Groupings		Retention Rate	Employment Index
Academic	Financial		
1	1	84.4%	1
1	2	79.4%	1
1	3	72.6%	2
1	4	63.0%	3
2	1	77.1%	1
2	2	74.8%	1
2	3	68.6%	2
2	4	65.4%	3
3	1	72.6%	2
3	2	67.6%	2
3	3	66.4%	3
3	4	60.3%	4
4	1	58.3%	4
4	2	66.1%	3
4	3	59.9%	4
4	4	55.2%	4
Total		69.7%	

Source: VSU SRA Analysis of Data Warehouse, 2013.

In Figure 11 the findings show the retention rates of the Employment Index. Groups 1 and 2 on average are above the overall institutional retention rate. Groups 3 and 4 are below the overall retention rate and are most likely to benefit from an on-campus job.



Source: VSU SRA Analysis of Data Warehouse, 2013.

Figure 11: Employment Index Retention Rates

Faculty Portal, Early Alert, and Interventions

With the predictive analytics developed to provide information about the entering students to faculty, a portal, called Valdosta State University Faculty portal, was launched in August 2012. The faculty portal provides faculty with information on students who are enrolled in their courses and it also provides information on faculty advisees. This information includes the following: a picture of the student, student contact information, the at-risk variables, integration with Degree Works, and a link to a faculty reporting form for attendance and academic progress. This allows the institution to provide critical student success metrics to faculty rapidly, and this information is available to faculty as soon as registration opens for a given term. When a faculty member reports a student for attendance or academic progress, the technology running behind these forms sends automated communications to staff members who are charged with providing interventions to help the students succeed. The timely reporting of the students and deployment of the interventions should lead to enhanced student success with regard to pass rates in their courses.

When the portal with the predictive analytics was first implemented at the university, there were three expected or desired results that would stem from the implementation:

1. Faculty who used the portal to inform students about their progress would have higher pass rates than those who did not use the portal.
2. Students who were flagged at-risk academically, either through predictive analytics or by the faculty, would have improved pass rates.
3. As a result of improved pass rates, the cohort retention would increase.

Over the course of the academic year 2012-2013, the data collected by the portal was analyzed, especially focusing on the faculty who had a high number of at-risk students enrolled in their courses. Largely, this was the Department of Mathematics and Computer Sciences. Table 28 shows the crosstabulation of the pass rates by faculty views. The threshold was set at least 100 views for improvement to occur. Pass rates of faculty who had 100 views or more had a 6.3% higher pass rate than those who had less than 100 views. In order to determine if the increased pass rates were statistically significant, a chi-square test for independence was conducted. The relation was significantly different, $\chi^2(1, N = 7,475) = 28.097, p < .001$. The size effect, Cramer's V, indicates a weak relationship, .061. This means that students who had a faculty who had at least 100 views in the portal are more likely to have higher pass rates than students who had a faculty who had less than 100 views.

Table 28: Crosstabulation of Pass Rates by Faculty Page Views

Views		DFW	Pass	Total
Less than 100 views	Number	1,795	3,444	5,239
	Percent	34.3%	65.7%	100.0%
100 views or more	Number	626	1,610	2,236
	Percent	28.0%	72.0%	100.0%
Total	Number	2,421	5,054	7,475
	Percent	32.4%	67.6%	100.0%

$$X^2 = 28.097***$$

*** p < .001

Additionally, the flag set by a faculty member would potentially show faculty's intentions of helping a student to succeed in the course. The threshold was set at a

minimum of five flags. Table 29 shows the crosstabulation of pass rates by the faculty who set at least a minimum of five flags. Of the faculty who set at least five flags, the pass rate is 10.2% higher than the pass rates of the faculty who set fewer than five flags. In order to determine if the increase in pass rates was statistically significant, a Chi-square test for independence was conducted. The relation was significantly different, $\chi^2(1, N = 7,475) = 50.078, p < .001$. This means that faculty who set at least five flags in the portal are more likely to have higher pass rates than the faculty who had set less than five flags.

Table 29: Crosstabulation of Pass Rates by Faculty Flag Set

Flags		DFW	Pass	Total
Less Than Five Flags	Number	2,114	4,080	6,194
	Percent	34.1%	65.9%	100.0%
Five Flags or More	Number	307	974	1,281
	Percent	24.0%	76.0%	100.0%
Total	Number	2,421	5,054	7,475
	Percent	32.4%	67.6%	100.0%

$$X^2 = 50.070***$$

*** $p < .001$

Due to the success in the increase of passing grades, the retention rate within the Fall 2012 cohort increased if they had a faculty member who used the portal. Table 30 shows the crosstabulation of the retention rates of the cohort students. Students who had a faculty who utilized the portal had a 4.9% higher retention rate than those who did not have a faculty that utilized the portal. A Chi-square test for independence was conducted to determine the significance of the relationship. The relationship was found to be

statistically significant, $\chi^2(1, N = 1,880) = 4.776, p = 0.029$. This means that students who had a faculty member utilize the portal are more likely to be retained than those who did not have a faculty member utilize the portal.

Table 30: Retention of Fall 2012 Cohort by Portal Users

Portal Usage		Not Retained	Retained	Total
Non-Portal Users	Number	216	416	632
	Percent	34.2%	65.8%	100.0%
Portal Users	Number	365	883	1,248
	Percent	29.2%	70.8%	100.0%
Total	Number	581	1,299	1,880
	Percent	30.9%	69.1%	100.0%

$$X^2 = 4.776^*$$

*** $p < .05$

The findings that this research has provided present a strong case for the use of predictive modeling in higher education. The results of the predictive modeling and the use of intervention strategies for the data are very clear, especially with regard to the Math Department in this study. These types of analysis, data collection and deployment of information can lead to better informed decisions that impact student success. While there was not an implemented intervention strategy for the use of the financial data, it is clear that socioeconomic and finances play a significant role in student success. The intervention strategies that were deployed and measured show improvements in student success, however, more analysis and improvements to the strategies need to be implemented.

Chapter V

PREDICTIVE ANALYTICS AND THE FUTURE OF STUDENT SUCCESS

Higher education continues to face challenges on numerous fronts as it deals with a student population whose needs are fundamentally different from those of previous generations. Students, faculty and staff, administrators, accreditors, public officials and private employers are all affected in some way by these challenges. Whether it is the exploding student debt problem, increased demands for accountability, use of performance-based funding models, each is in some way related to student success as defined by classroom performance, retention, and timeliness of degree completion. These issues are directly reflected in the mandate for Complete College America and its local instantiations such as Complete College Georgia whose missions are to improve college completion rates while preparing an educated population to enter into the workforce of the 21st century.

Research Question and Key Findings

The primary emphasis of this project has been focused on a single aspect of the overarching issue, namely the relationship between improved student performance and retention. More specifically, the research question asked: Can predictive modeling be used to create actionable student intelligence to improve the grades in key English and math classes resulting in higher retention rates of traditional first-year students?

The first part of the question dealt with the development of a model that predicted grades of first-year students in key English and math classes. The predictive

metrics developed for these models used three specific indicators: (1) standardized test scores from the SAT or ACT, (2) high school grade point average and (3) where the student's high school ranks in relation to the other high schools in the state of Georgia. The sample group for this part of the study was traditional freshmen who enrolled at Valdosta State University from 2008-2011. This baseline population was used to determine the efficacy of the model in predicting gateway course grades for years 2012 to 2014.

The second part of the question involved the development of actionable intelligence to improve student performance in the English and math courses with the expectation that grade improvement will increase student retention. This was (and is) a dynamic exercise that involved a host of diverse actors on campus including admissions, advising, tutoring, student affairs and faculty members. The design allowed multiple entry points for the input of data with the information from the data flowing to other affected areas. Each area then had the opportunity and ability to respond on an as-needed basis to intervene with the student as appropriate. The goal was to enable information to be collected and exchanged in such a way as to allow potential student problems to be addressed before they became unsolvable.

Chapter 2 provided a comprehensive review of the existing literature that relates to statistical methods for predicting grades and a review of intervention strategies that have proven to be successful. The review was organized into three broad sections. Section I presented an overview of the use of predictive modeling in relation to retention and graduation rates. The overview contained both an historical perspective as well as contemporary studies. As noted in the research of Pomplun et al., in the post-Vietnam

late 1970s the correlation between SAT scores and GPAs began to diverge. Both SAT scores and GPAs showed declines. The average SAT scores decreased substantially, but while freshmen GPAs also declined, the decrease was not at the same pace as the SAT. Whether this divergence was the result of grade inflation is still undetermined. Regardless, the predictive ability of the SAT was less reliable during this period. (Pomplun, Burton, Lewis 1991). Section II concentrated on the subject of financial aid and its implications on retention and graduation. As demonstrated in the research of Tinto (2006), a leader in the area of retention and financial indicators, although more low-income students were being accepted into 4-year institutions, they suffer from a lower graduation and persistence rate. Tinto's focus in the article concentrated on the transition from 2-year universities to 4-year universities. Among those beginning at a 2-year college, only 8% of low-income students earn a bachelor's degree within 6 years, while 25% of high-income students do (Tinto 2006, 11-12). Section III reviewed the use of standardized indicators as predictors of grades and college success rates in key subjects such as English and mathematics. Megert's (2005) research at a community college in New Mexico inspected the honors scholarship retention rate when the scholarship was awarded based on completed advanced high school math classes compared to the retention rate when the scholarship was based solely based on GPA. Advanced math classes were defined as trigonometry, pre-calculus, and calculus. Results indicated that there was no significant correlation between grades earned in advanced math classes and retention, however, there was a 10% difference in retention when considering scholarships awarded based on a high GPA and the completion of an advanced math class versus scholarships based solely on a high GPA. When HighMath

(having completed an advanced math class) was used in conjunction with GPA, 73% of students retained their scholarship, while only 58% retained their scholarship when GPA alone was used. Also, when MathScore (a combination of the rigor of the course, number of courses taken, and grades received in courses above Algebra I) was combined with GPA, there was a 67% retention rate; whereas students awarded scholarships based solely on their GPA had a 48% retention rate (Megert 2005, 55-56).

While this literature review accounts for and validates that this type of research has been conducted it is still somewhat different than what this research project proposes. While studies and research that focus on the results are very useful in defining the parameters of the problems, they fall short of identifying the solutions. Attempts to understand problematic graduation rates by looking back at the causes, need to be supplemented by research done preventatively at the front end. By identifying target populations who are at-risk at the very beginning of their educational careers, it may be possible to develop and implement strategies that will enable the students to be more successful; thereby, retaining, progressing and graduating more students—the ultimate reward for persistence.

In summary of Chapter 3 the development of the statistical methodology to be used to predict grades in key English and math courses was presented. The model assigned an index score to each individual student based on the criteria outlined earlier with the scores grouped and distributed to the appropriate faculty members. The methodology also involved a detailed description of intervention strategies, the trigger points used to alert the various departments and how the results of the intervention(s) were inserted into the data flow to refine the process.

In Chapter 4 there were five important clusters of results: (1) the pass/fail rates as highlighted in green and red in those initial tables based upon the 1-4 rankings for high school rank, GPA, and SAT, (2) the multivariate regression analysis, (3) the difference of means test for the changes over time once the placement index was put in place, (4) the analysis of the financial groupings and employment index, and (5) the chi-square and Cramer's V test for portal users vs. non-users.

In summarizing the first important cluster there were 9 tables that analyzed grade performance, retention from fall to fall, and the credit hours earned for each of the student indices, 1-1-1 through 4-4-4 (including 9-1-1 through 9-4-4 for which the 9 represents no high school rank such as out-of-state high school) for each of the three models listed above. The tables also broke out the analysis by Open (courses that are not English or Math based such as Music), English and math-based courses, and an analysis of the total. In studying these results it is clear that the variables that have been analyzed are very predictive when looking at students' likelihood to pass or fail courses. The tables also show the percent retained to the following fall semester. Within the tables, the shading of green represents students with minimal risk for failure and the shading of red represents very high risk of failure. The findings also interestingly reveal that all SAT verbal ranks of 1 and 2 show high risk for math-based courses with the exception of 3-1-1 and 3-1-2. In Table 19 all but five categories show high-risk for non-success in math-based courses when using the SAT verbal compared to the all but seven categories when looking at Table 17 which shows the same high school rank and high school GPA but uses the math standardized test scores. The drop in pass rates from the top high school GPA to the second tier is about 20%. Also it is noted that the high school GPA of 1 show

no risk within both high school rank 3 and rank 4.

The second important set of results were the regression models and from the results in Table 20 the hypothesis that these factors can produce a highly predictive model can be accepted based on the ANOVA, statistical significance of the regression coefficients, and predictability of the dependent variable DFW rate. The three independent variables account for 6.5% of the variance in the DFW rates in the general and English models while they account for 12.6% of the variance in the DFW rate for the math model as seen in the adjusted r^2 . This leaves a large amount of the variance unexplained. It is also important to note the dominance of the high school GPA in the model, while the other variables show significance, it is clear the high school GPA is the most powerful predictor. In Appendix B, the predictability of each coded category is shown. This data will be used to assign future students codes that will provide faculty and academic support staff information about students and their likelihood to struggle with classwork. Also the t scores further validate this model as shown in Table 20.

The third important set of results, through a difference of means test, examines the impact of the math placement index. This model was based on earned final grades of first-time, full-time freshmen (cohort students) who enrolled in a MATH 1111 at Valdosta State University (VSU) from Fall 2005 to Fall 2014. In determining the pass rates, grades of A, B, and C were classified as passing, while grades of D, F, and W were classified as not passing. Cohort students who earned a grade of NR were excluded from the analysis. Overall a total of 2 grades were excluded. In order to determine the placement of students, admission standards (high school grade point average and standardized test scores) were analyzed. All ACT math scores were converted to an SAT

score. If a student had both an SAT math and ACT math, the higher of the two was taken. The scores were then given an index based on the math placement index values. If the index indicated that a student is at a high risk of not passing MATH 1111, a modified index, $\text{ModifiedIndex} = (\text{HS GPA} * 500) + \text{SATmath}$, was created to determine which students needed to go into the MATH 1111 Extended Learning.

After analyzing the predictive models for students from Fall 2005 to Fall 2014, the Math Department worked to place students for Fall 2014 into the equivalent MATH 1111 Extended Learning sections (which meet 5 days per week). Table 21 compares their performance (pass/fail rates) to the previous Fall semesters. The results are displayed in Figure 6 and Table 21. The pass rate from 2006-2012 was very flat and it has increased from 0.528 for the 2005 cohort to the 0.69 for 2014 cohort. This is an increase of 0.162 points. Since the math placement was in its second year and the extended learning was in the first year, a pass rate increase of 0.173 was experienced. In order to determine if the pass rate of the pilot group of the extended learning math students in 2014 was significant from like students from 2013, a Chi-square test for independence was conducted. The relation between the two terms was significantly different, $\chi^2(1, n = 287) = 8.250, p = 0.004$. A t test was also run to determine significance and the t score of 3.035 is very significant at the $p < .005$ level in accordance with the t distribution. This means that the 2014 pass rate for the extended learning students is significantly higher than the 2013 pass rate. It was also observed if students were not in the Math 1111 extended classes and were enrolled in the regular Math 1111 courses there was not a significant difference in their performance.

The fourth important cluster of results comes from the analysis of the financial

groupings as seen in Table 25, an analysis of retention rates based on the financial indicators of first-time, full-time (FTFT) freshman cohorts from Fall 2007 to Fall 2012. These financial indicators include: the HOPE Scholarship, which is an academic performance based scholarship in the state of Georgia; Pell Grant which is a need-based aid program that is based on income level; and parents' contribution is the amount of money that parents should be paying toward a child/dependents post-secondary education (the calculation considers income and dependents in the household). The findings, when combining these variables and analyzing retention, reveal that FTFT freshman who received the HOPE Scholarship, did not receive any Pell Grants, and had an expected parents' contribution at or above the average dollar amount had the highest retention rate (79.7%). With a retention rate of 55.7%, students who do not receive the HOPE scholarship and receive below average in both the Pell Grant dollars and the expected parents' contribution have the lowest retention rate. Financial groupings were applied by using quartiles of the retention rates.

The fifth important cluster of results comes from the analysis of the portal users versus the non-portal users. Student success, defined as the increase of passing grades and the retention rate within the Fall 2012 cohort, increased if they had a faculty member who used the portal. Table 30 showed the crosstabulation of the retention rates of the cohort students. Students who had a faculty member who utilized the portal had a 4.9% higher retention rate than those who did not have a faculty member that utilized the portal. A Chi-square test for independence was conducted to determine the significance of the relationship. The relationship was found to be statistically significant, $\chi^2(1, N = 1,880) = 4.776, p = 0.029$. This means that students who had a faculty member utilize the

portal are more likely to be retained than those who did not have a faculty member utilize the portal.

The data generated by the model and the intervention strategies that have been employed were analyzed and reviewed over a 2-year period, 2013 and 2014. The results were compared to course grades for students in classes whose faculty are not using the intervention strategies which revealed that student success has improved at a statically significant level. Enhancements to the predictive model and to the intervention strategies occurred with regard to the Math Department. The financial aid model proves that student's ability to afford higher education plays an important role in their success. In addition to the findings that conclusively demonstrated the link between improved student performance through active intervention strategies and student retention, there are some very important corollary insights that can be drawn from this project. Chief among these is the various ways that different constituent groups consume data. For example, the Professional Staff Advisors used the information not only for intervention purposes but also for advising students on their schedules. The Department of Housing and Residence Life used the data to partner and do programming with the Student Success Center.

The best example of that is with the Math Department. They not only used the data to better understand student success in math course but also engaged in enhancement of the model. They also used the data to change the way they offered math. This adoption and use of the data is exactly what this research intended—to make data actionable.

There is no arguing that models can be created for specific majors, on-line students, and even student engagement could be modeled. This could be done through

collection of additional data on students participating in different types of activities (the data could be collected utilizing the student ID card swipe). This type of data not only can further enhance the models that were built in this project, but it will also give much needed information to Student Affairs and other service and support areas about which students are actively engaged in the wide range of activities on campus. One thing that was clear in the research findings of this project was that even though a group of students are at-risk many of them are still going to succeed, this additional data and analysis can help to future subdivide the at-risk populations and create better more predictive models. However, the models and data will not change the success of the students, it is how the data is used that will improve the likelihood a student can be successful.

Lessons Learned

The most surprising data in the findings was the dominance of the high school GPA in the regression models and in the clustering. Even in the cases where the students had high SAT scores if the GPA was low it was the better predictor of student likelihood of not being successful. Considering that Georgia has some grade inflation due to the HOPE Scholarship Program, high school GPA is still a very powerful indicator when measuring student success.

The regression models showed high levels of statistical significance; however, they also had low levels of explained variance across the three models. This could be better understood if there were ways to incorporate other data or variables into the analysis. For example, if you had the high school grades by course rather than the cumulative high school GPA the models could then be reconfigured and possibly this would better address some of the variance. Combining the financial data into the

regression models could also possibly account for some of the variance. This was not possible in this research as the financial aid data spanned a slightly different time period than that of the student success models. If the data could be merged at the student level then the analysis could be run.

Another powerful lesson learned came from analyzing the faculty that used the portal and looking at how the pass rate improved when compared to those in previous terms. Less than 40% of faculty used the portal and set flags on students. This presents a problem with the adoption of the technology and how to incorporate student success into faculty responsibilities. If more faculty used the tools, the research shows that student success would increase.

High-Impact Practices

When you can couple these models with high-impact practices like those mentioned in the work of George Kuh (2008) such as freshmen learning communities, service learning and undergraduate research an institution can see increases in retention, progression and graduation. During the time this research was conducted at Valdosta State University (VSU 2016), the university community was engaged in a Quality Enhancement Plan (QEP) that centered on undergraduate research projects. These QEP projects demonstrate how successful these high-impact practices can be. The university conducted 13 projects in the first 2 years of the QEP as summarized in Table 31.

Table 31: Overview of QEP Project Participation, Presentations and Publications

QEP Round 1				
Project Name (Department)	# of Students	# of Presentations	Manuscripts in Progress	# of Publications
Cutting Edge Cancer Research (Chemistry)	48	15	3	12
Summer Archival Field Experience in History (History)	5	12	1	1
Investigating Social Inequalities of Hispanic Immigrants through the U.S.-Mexico Borderland Experience (Women's and Gender Studies)	15	3	0	0
Preparing Scholars of Tomorrow to Effectively Analyze Language Sample Data for Parent-Child Turn Taking (Communication Sciences and Disorders)	68	12	2	0
Evidence-Based Practice Strategies for Nursing and Health Care (Nursing)	52	2	0	0
Discovering Unrealized Generational Differences in Kitchen Design Preferences Between Next Generation Interior Designers and Current Resident-Users (Art)	8	7	0	0
Total	196	51	6	13

QEP Round 2				
Project Name (Department)	# of Students	# of Presentations	Manuscripts in Progress	# of Publications
Towards the Internationalization of the Language Curriculum (Modern and Classical Languages)	14	20	5	1
Developing a New Group of Medicinal Agents (Chemistry)	103	30	1	18
Improving Mechanical Test Methods in Biomaterials and Engineering (Engineering Studies)	8	7	0	0
Studying Human Impacts on Water Quality in Lakes and Rivers of Lowndes County, GA (Biology)	13	10	1	1
The Design of Future Wise Cars (Math and Computer Science)	34	9	0	2
Engaging Students in Understanding Academic Cultures (English and Honors College)	9	7	0	0
Research on Climate Change Action Plans for Small Southern Cities (Geosciences)	14	4	1	0
Total	195	87	8	22

In the first round of QEP projects, the Chemistry project involved producing economical and effective medicinal agents from the sea. The students developed/synthesized five new cancer drugs that were accepted for testing by the National Cancer Institute. The project utilized a “pharmaceutical aquaculture” approach to produce bryostatin, an effective but very expensive cancer and Alzheimer’s drug. Students were part of a patent application for this new group of drugs. The students also developed a new tuberculosis drug that was accepted for testing by the Infectious

Diseases section at the National Institutes of Health. The History project was a summer research experience in which students conducted archival research at the U.S. Army Heritage and Education Center, which is the Army's main archive on America's oldest military base. Students learned how to navigate a major national archive, identify and request materials, and gather information from those materials for a culminating research paper and presentation on a topic pertaining to military history. The Women's and Gender Studies project also involved an intensive summer research experience in which undergraduates were immersed in the border culture of the El Paso-Ciudad Juarez region. Students were placed in local agencies and non-profits, which serve the diverse Hispanic population of the border region, to engage in qualitative research via ethnographic interviews and participant observations. Students investigated how gender, race, ethnicity, class, and nationality status affect Hispanics residing in the borderland region as well as how globalization impacts migration, immigration, and poverty. For the Communication Sciences and Disorders project, undergraduate students worked in pairs to collect a language sample of parent-child interaction during daily routines. The students utilized innovative recording technology known as LENA (Language Environment Analysis System) to collect and analyze the data. Students investigated adult word count (the total number of adult words the child hears), conversational turns (the total number of conversational interactions the child engages in with an adult), and child vocalizations (continuous speech spoken by a child wearing a Digital Language Processor). In the fifth QEP project, Nursing students learned research concepts and developed skills supporting evidence-based nursing practice. Working with a major hospital system comprised of six acute care hospitals, the students investigated end-of-

life care communication and the level of moral distress. In the final project of the first round, students from the Interior Design program in the Department of Art explored the differences between kitchen layouts in assisted living facilities designed by interior design students and the layouts designed by residents of assisted living facilities. Students analyzed and coded past studio projects, interviewed project participants, and built models for participants.

There were seven projects in the second round of the QEP. The first project from Modern and Classical Languages incorporated discipline-based inquiry into a study abroad program to Cadiz, Spain. Students in the Spanish for Professionals program and the Foreign Language Education program analyzed the health care and education systems in the United States and Spain through service-learning as well as discipline-based inquiry projects. In the Chemistry project, students analyzed the efficiency and effectiveness of drug delivery for cancer drugs and antibiotics. Students synthesized new pharmaceutical compounds and investigated the copper (II) ion as a delivery platform. For Engineering Studies, the students engaged in two discipline-based inquiry projects: mechanical testing of dental ceramics and an analysis of hydrogen diffusion and corrosion damage in pipeline steel. For the first project, undergraduates designed and developed suitable mechanical tests and computational techniques for determining the strength of dental porcelains. Students also developed three-dimensional models of various dental restorations and analyzed their performance under simulated biting forces using three dimensional printers and finite element stress analysis. In the second Engineering project, students investigated the impact of material microstructure on hydrogen diffusion and corrosion damage in pipeline steel. The Biology students

explored human impacts on the water quality of lakes and rivers in Lowndes County, Georgia, by monitoring and describing the presence of nutrients (carbon, nitrogen, and phosphorous) and various metals. Undergraduate students utilized and applied multiple sampling and limnological techniques, as well as laboratory analyses, to determine which human and environmental stressors have the greatest impact over the health of the system and the biota occupying the lakes and rivers. Students in Mathematics and Computer Science applied their computer and communication knowledge to the essential daily task of driving. Undergraduates investigated the use of Radio Frequency Identification tags, readers, and short-range wireless communicators to create a “social network of cars” and use “wisdom of crowds” to provide a low-cost and yet more accurate and reliable solution for driving assistance so that nearby cars can exchange road and traffic information, issue warnings in case of danger, and achieve safe and autonomous driving. In the project “Engaging Students in Understanding Academic Cultures,” first-year students from the Honors College researched their experiences in classrooms to determine what their classes value, ascertain how to learn effectively in those courses, and recommend improvements in undergraduate education. Through an ENGL 1102 Composition II course, the students examined and discussed the academic cultures in their Biology, Psychology, Honors Seminar, Theatre, Computer Science, Music, Sociology, and Chemistry classes. For the final QEP project in the second round, students in the Geosciences collected, assembled, and analyzed data from other cities on their climate change action plans, and then proposed a plan for the City of Valdosta, Georgia. Students constructed GIS (Geographic Information System) maps of climate change and related policy in different cities.

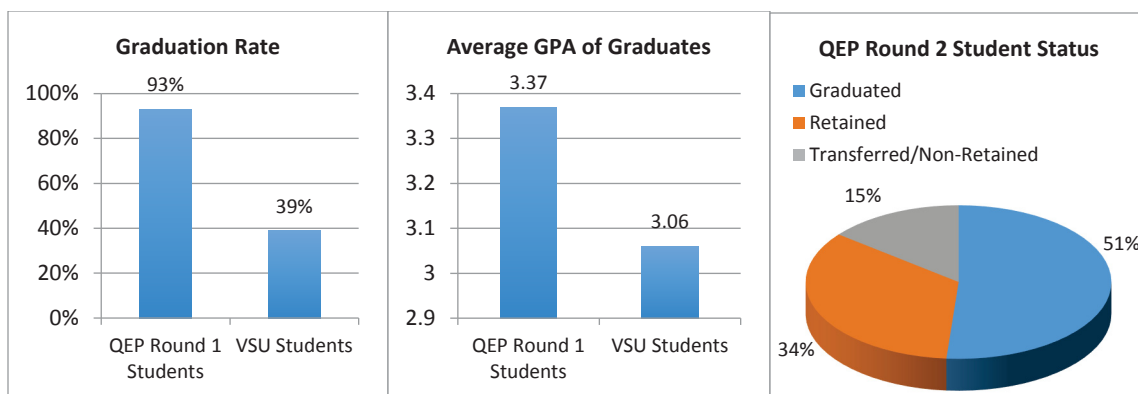


Figure 12: Graduation Rates, Retention Rates, and GPAs of QEP Students

As seen in Figure 12, students who participated in these projects showed much higher graduation and retention rates as well as the GPAs than that of other students, with a 93% graduation rate of QEP students in the first round of projects compared to a 40% institutional graduation rate. Furthermore, QEP students from that first round graduated with an overall GPA of 3.37 compared to 3.06 for the overall student body. As noted at the outset, discipline-based inquiry was selected as the theme of the QEP because of the powerful impact undergraduate research has on graduation and retention rates. Among the students in the second round of projects, 85% have graduated or have been retained at Valdosta State University.

VSU will be pursuing a Title III Strengthening Institutions grant from the Department of Education in 2017 to improve undergraduate student retention which will marry the high-impact practices from the QEP along with the predictive analytics that this project has brought into practice through the student success portal (see Faculty Portal from Chapter 3).

Additional Avenues of Research and Study

The intervention strategies that were analyzed in this research project were effective but there is plenty of room for enhancement of the existing initiatives and

creation of new more impactful strategies. The need for better data collection on the strategies was also noted as the results were analyzed. It would be helpful if a standard survey instrument were used when the advisors are deploying the intervention strategies with the students that have been flagged for academic progress. This would not only collect better data on the intervention but also gather qualitative information on the student's motivation.

A student portal which could provide students with single sign-on to the services they use most on campus (Web registration, LMS, Email, etc.) could also be utilized to channel information to students. This information could be tailored with resources to help them with classes they are enrolled in that they show a level of risk for. For example, if a student was enrolled in college algebra and was at risk of doing poorly they could receive through ads in the portal when tutoring and other helpful resources were available. All of the information in the portal (logins and activity) could be recorded and analyzed to measure effectiveness.

Another intervention strategy that is needed centers on the cost of higher education and students' ability to finance their education. It was clear from the Financial Model and Jobs index from Chapter 4 that a campus program that coupled academic preparation, financial planning and job skills and packaged these skills into a job that could pay these students a stipend around mid-semester could have a significant impact on the retention, progression and graduation of these students. Financial factors and a student's ability to maintain eligibility is one of the primary reasons students are not retained.

Cautionary Warning and Conclusion

There is also a need to caution institutions, a warning about black box analytics and how companies are introducing these solutions in higher education. Institutional Research professionals partnered with Information Technology professionals should be providing solutions like these in house. These black box predictive models (while they may be great models) erode one of the key tenets of the Institutional Research (IR) profession, that of researcher. Higher education and the American Association of Institutional Research (AIR) should be working to develop IR professionals and equip them with the skills to build and produce a variety of predictive analytics and learning models to aid institutions in better understanding their unique student populations. However, when institutions outsource these functions to a non-transparent third party solution they become less likely and less empowered to mature with regard to analytics, business intelligence and the basic principles of research methods. In the last 5 years I have seen more universities spend and waste scarce financial resources on a myriad of these products with little return, or no return, but what is usually observed is even more costs. Outsourcing a key principle (research methods), which all graduate and doctoral students should be well versed in, is a slippery slope for higher education. The future of higher education should be empowerment, transparency and a community of developing and sharing innovative solutions, models, and analytics. The IR profession is now beginning to engage in this conversation and move in this direction and we need to support practices to further that cultural development and NOT pursue an outsourced subscription cloud model.

It is also important to reiterate that the best measure of student performance in college level work is their performance in high school work. If you were a good student

in high school and made good grades there is a high likelihood that you will perform well in college or university level work. This can be seen in the way the some schools have placed more emphasis on the high school GPA and other criteria (such as engagement). One university that has made the standardized test scores optional is George Washington University, as stated by Laurie Koehler, senior associate provost for enrollment of George Washington University, “We hope the test-optional policy sends a message to prospective students that if you are smart, hard-working and have challenged yourself in a demanding high school curriculum, there could be a place for you here” (CBS News 2015). This can also be seen at other schools that place a higher weighted value on the high school GPA, over the standardized test scores in the selection criteria for admission.

In closing there is a great need for American public education to better understand and teach the students at all levels. The concept of blended classrooms where students of varying levels of understanding are mixed is not working. As seen with the math students in this study, different students need different approaches. Some need one hour of math while others need two. It is not one size fits all. Higher education has the technology, the data and the human and intellectual capital to do this and put the United States back in the top five in the world with regard to college completion.

REFERENCES

- Adelman, C. 2004. *Principal Indicators of Student Academic Histories in Postsecondary Education, 1972-2000*. Washington DC: US Department of Education.
- Beckett, A. 2006. *Relationship between Participation in a Residentially-Based Freshman Interest Group and Degree Attainment*. Columbia, MO. University of Missouri Press.
- Belfield, C., Crosta, P., & Columbia University, C. C. 2012. "Predicting Success in College: The Importance of Placement Tests and High School Transcripts." CCRC Working Paper No. 42. Columbia University: Community College Research Center
- Bremer, C. D., Center, B. A., Opsal, C. L., Medhanie, A., Jang, Y. J., & Geise, A. C. 2013. "Outcome Trajectories of Developmental Students in Community Colleges." *Community College Review* 41(2): 154-175.
- Burgher, K., & Davis, M. 2013 "Predictive Analytics for Student Retention: Group vs. Individual Behavior." *College and University* 88(4): 63-65.
- CBS News. 2014. "George Washington is Latest School to not Require SAT, ACT Scores." 28 July.
- Complete College America. 2011. *Time is the Enemy*. September 2011. Accessed on May 25, 2014. http://completecollege.org/docs/Time_Is_the_Enemy.pdf
- Dynarski, S. 2000. "HOPE for Whom? Financial Aid for the Middle Class and its Impact on College Attendance." *National Tax Journal* 53(3): 629-61.
- EDUCAUSE. 2013. "McGraw-Hill Education Launches First Adaptive Math Placement Platform at 2013 EDUCAUSE Annual Conference." *PR Newswire US*.

- Ewell, P. T. 1984. *Conducting Student Retention Studies*. New York: National Center for Higher Education Management Systems and College Board.
- Friedl, J., Pittenger, J. D., & Sherman, M. 2012. "Grading Standards and Student Performance in Community College and University Courses." *College Student Journal* 14: 526-532.
- Gerald, D., & Haycock, K. 2006. *Engines of Inequality: Diminishing Equity in the Nation's Premier Public Universities*. Washington DC: Education Trust.
- Hamrick, F., Schuh, J., & Shelley, M. 2004. "Predicting Higher Education Graduation Rates from Institutional Characteristics and Resources Allocation." *Education Policy Analysis Archives* 12(19) 1-23.
- Herzog, S. 2008. *Estimating the Influence of Financial Aid on Student Retention: A Discrete Choice Propensity Score-matching Model*. University of Nevada: CRDA STATLAB.
- Hodara, M. S. 2014. "An Examination of the Impact of Accelerating Community College Students' Progression through Developmental Education." *Journal of Higher Education* 85(2): 246-276.
- Hoyt, J. E., & Sorensen, C. T. 2001. "High School Preparation, Placement Testing, and College Remediation." *Journal of Developmental Education* 25(2): 26-34
- Jacobson, E. 2007. "A (Mostly) Brief History of the SAT and ACT Tests." Accessed September 27, 2014. <http://www.erikthered.com/tutor/sat-act-history.html>

- Kappe, R., & van der Flier, H. 2012. "Predicting Academic Success in Higher Education: What's More Important than Being Smart?" *European Journal of Psychology of Education* 27(4): 605-619.
- Kelley-Kemple, T., Proger, A., Roderick, M., & Society for Research on Educational Effectiveness. 2011. *Engaging High School Students in Advanced Math and Science Courses for Success in College: Is Advanced Placement the Answer?* Evanston, IL: Society for Research on Educational Effectiveness.
- Koppius, O., & Shmueli, G. 2011. "Predictive Analytics in Information Systems Research." *MIS Quarterly* 35(3): 553-572.
- Kuh, G. D., Kinzie, J., Schuh, J. H., & Whitt, E. J. 2005. *Student Success in College*. California: Jossey-Bass.
- Lauff, E., & Ingels, S. 2014. *Education Longitudinal Study of 2002 (ELS:2002): A First Look at 2002 High School Sophomores 10 Years Later*. U.S. Department of Education NCES.
- Lee, J. 2012. "College for All: Gaps between Desirable and Actual P-12 Math Achievement Trajectories for College Readiness." *Educational Researcher* 41(2): 43. doi:10.3102/0013189X11432746
- Li, K., Zelenka, R., Buonaguidi, L., Beckman, R., Casillas, A., Crouse, J., & Robbins, S. 2013. "Readiness, Behavior, and Foundational Mathematics Course Success." *Journal of Developmental Education* 37(1): 14-36.
- Lotkowski, V., Robbins, S., & Noeth, R., 2004. *The Role of Academic and Non-Academic Factors in Improving College Retention*. ACT Policy Report. Iowa City, IA: ACT Inc.

- Megert, A. D. 2005. "The Relationship between High School Math Courses, High School GPA, and Retention of Honors Scholarships." *NASFAA Journal of Student Financial Aid* 35(2): 51-60.
- Melguizo, T., Kosiewicz, H., Prather, G., & Bos, J. 2014. "How Are Community College Students Assessed and Placed in Developmental Math? Grounding Our Understanding in Reality." *Journal of Higher Education* 85(5): 691-722.
- Morgan, R. 1990. *Predictive Validity within Categorizations of College Students*. New Jersey: Educational Testing Services.
- National Center for Education Statistics (NCES). 2014. "Education Longitudinal Study of 2002 (ELS:2002): A First Look at 2002 High School Sophomores 10 Years Later First Look." NCES 2014-363. U.S. Department of Education. January 2014. Accessed May 23, 2014. <https://nces.ed.gov/surveys/els2002/>
- National Center for Public Policy and Higher Education. 2010. "Beyond the Rhetoric Improving College Readiness through Coherent State Policy." The National Center for Public Policy and Higher Education. Accessed May 24, 2014. http://www.highereducation.org/reports/college_readiness/gap.shtml
- National Math + Science Initiative (NMSI), 2014. "National Math + Science Initiative Boosts College Readiness for More Than 13,000 Students Nationwide, Based on 2013-14 Advanced Placement* Exam Results." National Math + Science Blog. Accessed March 26, 2015. <https://www.nms.org/Blog/>
- Ngo, F., & Kwon, W. W. 2014. "Using Multiple Measures to Make Math placement decisions: Implications for access and success in community colleges." *Research in Higher Education*, 55(2): 442. doi:10.1007/s11162-014-9352-9.

- Pascarella, E. & Terenzini, P. 2005. *How College Affects Students*. San Francisco, CA: Jossey-Bass Publishing.
- Pomplum, M., Burton, N., & Lewis, C. 1991. *An Exploration of the Stability of Freshmen GPA, 1978-1985*. Princeton, NJ: Educational Testing Service.
- PR Newswire. 2014. (October 27). "Dallas ISD High Schools Partnering with the National Math + Science Initiative Outperform Peer High Schools on Advanced Placement* Exams in 2013-14." National Math and Science Initiative, PR Newswire US. Accessed March 27, 2015.
<http://www.prnewswire.com/news-releases/486305785.html>
- Rothman, R. 2010. *Mission: A Drop in Dropouts*. Washington DC: ASCA School Counselor.
- Rylands, L. J. & Coady, C. 2008. "Performance of Students with Weak Mathematics in First year Mathematics and Science." *International Journal of Mathematical Education in Science and Technology* 40(6): 741-753.
- Schroeder, C. 2013. "Reframing Retention Strategy: A Focus on Process." *New Directions for Higher Education* 161: 39-47.
- Shmueli, G., & Koppius, O. 2011. "Predictive Analytics in Information Systems Research," *MIS Quarterly* 35(3): 553-572.
- Singell, L. D. 2001. *Come and Stay a While: Does Financial Aid Effect Enrollment and Retention at a Large Public University?* New York: Cornell Higher Education Research Institute.

- Singell, L. D., & Waddell G. R. 2010. "Modeling Retention at a Large Public University: Can At-risk Students Be Identified Early Enough to Treat?" *Research Higher Education* 51: 546-572.
- Sparks, S. D. 2013. "Remedial Placements Found to Be Overused." *Education Week* 32(2): 1-21.
- Strayhorn, T. L. 2014. "Modeling the Determinants of College Readiness for Historically Underrepresented Students at 4-Year Colleges and Universities: A National Investigation." *American Behavioral Scientist* 58(8): 972.
doi:10.1177/0002764213515230
- Valdosta State University (VSU). 2016. "Quality Enhancement Plan Impact Report." Submitted to the Southern Association of Colleges and Schools Commission on Colleges for Reaccreditation.
- Valdosta State University. 2014. *Undergraduate Catalog: Academic Catalog 2014-2015*.
https://www.valdosta.edu/academics/catalog/1415/ugrad/documents/vsu_undergraduate_catalog_2014-15.pdf
- Wetzel, J., O'Toole, D. & Peterson, S. 1999. "Factors Affecting Student Retention Probabilities: A Case Study." *Journal of Economics and Finance* 23(1): 45-55.

APPENDIX A

Institutional Review Board Exemption



**Institutional Review Board (IRB)
for the Protection of Human Research Participants**

PROTOCOL EXEMPTION REPORT

PROTOCOL NUMBER: 03349-2016

INVESTIGATOR: Andy Clark

PROJECT TITLE: College Student Success

INSTITUTIONAL REVIEW BOARD DETERMINATION:

This research protocol is **exempt** from Institutional Review Board oversight under Exemption Category(ies) 4. You may begin your study 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/SUGGESTIONS:

Although not a requirement for exemption, the following suggestions are offered by the IRB Administrator to enhance the protection of participants and/or strengthen the research proposal:

N/A

- 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 W. Olphie *3/28/16*
Elizabeth W. Olphie, IRB Administrator Date

*Thank you for submitting an IRB application.
Please direct questions to irb@valdosta.edu or 229-259-5045.*

Revised: 12.13.12

Appendix B

Complete Data Sheet

Panel	62.50	20.208	6.000	31.800	77.76	22.36	2.40	7	3	96.242	77.688	82.792	7.282	1.238	6.000	9.406	2.27	6	4	36.529	22.834	42.55	13.204	56.529	7.626	2.63	7	3	152.723	327.897	838.8
11.1	62.50	20.208	6.000	31.800	77.76	22.36	2.40	7	3	96.242	77.688	82.792	7.282	1.238	6.000	9.406	2.27	6	4	36.529	22.834	42.55	13.204	56.529	7.626	2.63	7	3	152.723	327.897	838.8
11.2	62.50	20.208	6.000	31.800	77.76	22.36	2.40	7	3	96.242	77.688	82.792	7.282	1.238	6.000	9.406	2.27	6	4	36.529	22.834	42.55	13.204	56.529	7.626	2.63	7	3	152.723	327.897	838.8
11.3	62.50	20.208	6.000	31.800	77.76	22.36	2.40	7	3	96.242	77.688	82.792	7.282	1.238	6.000	9.406	2.27	6	4	36.529	22.834	42.55	13.204	56.529	7.626	2.63	7	3	152.723	327.897	838.8
11.4	62.50	20.208	6.000	31.800	77.76	22.36	2.40	7	3	96.242	77.688	82.792	7.282	1.238	6.000	9.406	2.27	6	4	36.529	22.834	42.55	13.204	56.529	7.626	2.63	7	3	152.723	327.897	838.8
11.5	62.50	20.208	6.000	31.800	77.76	22.36	2.40	7	3	96.242	77.688	82.792	7.282	1.238	6.000	9.406	2.27	6	4	36.529	22.834	42.55	13.204	56.529	7.626	2.63	7	3	152.723	327.897	838.8
11.6	62.50	20.208	6.000	31.800	77.76	22.36	2.40	7	3	96.242	77.688	82.792	7.282	1.238	6.000	9.406	2.27	6	4	36.529	22.834	42.55	13.204	56.529	7.626	2.63	7	3	152.723	327.897	838.8
11.7	62.50	20.208	6.000	31.800	77.76	22.36	2.40	7	3	96.242	77.688	82.792	7.282	1.238	6.000	9.406	2.27	6	4	36.529	22.834	42.55	13.204	56.529	7.626	2.63	7	3	152.723	327.897	838.8
11.8	62.50	20.208	6.000	31.800	77.76	22.36	2.40	7	3	96.242	77.688	82.792	7.282	1.238	6.000	9.406	2.27	6	4	36.529	22.834	42.55	13.204	56.529	7.626	2.63	7	3	152.723	327.897	838.8
11.9	62.50	20.208	6.000	31.800	77.76	22.36	2.40	7	3	96.242	77.688	82.792	7.282	1.238	6.000	9.406	2.27	6	4	36.529	22.834	42.55	13.204	56.529	7.626	2.63	7	3	152.723	327.897	838.8
12.0	62.50	20.208	6.000	31.800	77.76	22.36	2.40	7	3	96.242	77.688	82.792	7.282	1.238	6.000	9.406	2.27	6	4	36.529	22.834	42.55	13.204	56.529	7.626	2.63	7	3	152.723	327.897	838.8

