USING PREDICTIVE ANALYTICS TO UNDERSTAND FACTORS AFFECTING TRANSFER STUDENT PERSISTENCE AND GRADUATION

A Dissertation
Submitted to
The Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

By
Mariya Alexandra Yanovski-Bowers
December 2019

Examining Committee Members:

Dr. Will J. Jordan, Advisory Chair, Department of Policy, Organizational and Leadership Studies
Dr. Joseph P. Ducette, Department of Psychological Studies in Education
Dr. Timothy J. Patterson, Department of Teaching and Learning
Dr. Tom Grites, External Member, Stockton University

ABSTRACT

It is the norm for institutions to report on their retention and graduation rates only for first-year student cohorts. Colleges and universities that report their first-year retention rates in the 90 % ranges often do not account for their newly-admitted transfer students. Much of the nuance in reporting retention comes from unaccounted transfer student registrations and enrollments. Reporting transfer retention is also much harder, since many transfer students do not have predictable patterns of enrollment.

This study examined factors that contribute to graduation, dropout and persistence and how they differ by race, socioeconomic class, and gender. Based on a new student questionnaire conducted in 2015, 2016, 2017 by a large research institution in the Mid-Atlantic, an exploratory statistical technique CHAID (Chi-Squared Automatic Interaction Detection) designed for a categorical dependent variable, was employed to establish the characteristics of transfer students who had a high probability to drop out after transferring to their new institution. Examining the dendrogram, one can easily classify the various "at-risk" student groups by tracing each of the terminal groups to the root of the tree.

The results of this study provide context and information for developing transfer friendly programming and interventions at both community colleges and four-year institutions. The results will be valuable to senior level staff, front-aline student support staff, faculty, and community organizations focused on helping students who seek re-enrollment after an extended academic leave period. Additionally, this study will demonstrate how modeling techniques can be used to develop predictive models for different populations, across different colleges.

ACKNOWLEDGMENTS

My parents. A kid that grew up without running water and plumbing does not usually aspire to complete a Ph.D., let alone graduate from college. Growing up three hours away from Chernobyl nuclear plant, my only dream was to grow up healthy. It is only because of my amazing parents that attending college became a reality. Your values, kindness, and perseverance, despite all the challenges that came with being an immigrant refugee in post-9/11 America helped me become the person I am today. My heart is full of love and gratitude for you. I could not have done this without you. Not a day goes by that I don't acknowledge how lucky I am to have you in my life.

Mr. Picone, it is because of you that I graduated High School and went on to college. You are the type of teacher all teachers should aspire to be. You nurtured our love for music, you built our character and taught us valuable life lessons, all in the small confines of the FLHS Music Room.

To all my colleagues and fellow students at Rutgers, FSU, Parsons, CUNY, and Temple. I have learned so much from you. Your love for your work and dedication to your students is the reason I get up for work every day. Specifically, I want to thank:

Kathleen Smith, your wisdom, charisma, and passion for helping students has forever changed how I advise. Thank you for being a mentor and a friend.

Otilia Abraham, you are an innovator and a change agent, your sense of humor and ability to always see the light at the end of the tunnel inspired me. Without you, I would have never developed an understanding and appreciation for working with online adult/transfer students and building an advising unit from the ground up.

Jim Degnan, Joe Ducette and Peter Jones. Thank you for the hours you spent teaching me data management, statistics and data analysis. *Jim*, thank you for the endless hours of mentoring and coaching me. I owe my success as a researcher to you. *Dr. Ducette*, this research would not have been possible without your willingness to help and serve on my examining committee.

Michele O'Connor, I am extremely fortunate to have had you as a mentor and a supervisor. Thank you for taking a chance on me. You have exponentially expanded my ability to navigate the landscape of higher education. The last seven years have flown faster than I can even imagine. Thank you for all that you do!

To my committee Members:

Will Jordan, thank you for always being a voice of reason, for your guidance from my first call as a Ph.D. applicant to refereeing my final defense. You always have a way to calm the nerves of your students. Thank you for not letting me give up. *Tim Patterson*, I am so lucky to have had you as one of my instructors. Your pedagogy and teaching methodology are by far one of the best that I have encountered. Thank you for teaching me how to write without fear, it only took 34 years of schooling for me to say that. *Tom Grites,* I often refer to you as my NACADA family. Meeting you at the NACADA Conference in 2009 has changed my life and given my career a much deeper purpose. Because of you, I serve as an advocate not only for my students but for my advising colleagues. Your passion for transfer students inspired me to publish on the topic and continue the work done by so many that came before me.

Rich Hetherington, Bhavesh Bambhrolia, Bernadette Mulvey, you and your teams were always willing to help. Your leadership and commitment to data transparency, integrity, and willingness to always collaborate will never be forgotten.

To my Graduate! Philadelphia colleagues, you continue to inspire me with your creativity, passion and dedication to adult students. I feel fortunate and lucky to have joined the team when I did and learn as much as I have learned. It was truly a symbiotic working relationship.

To all my students and former interns, you make me so proud. Seeing your accomplishments, hearing updates about your successes fills my heart with joy. The ripple effect of what we do spans generations and the impact can be felt by thousands of people.

Last, but not least - I am forever grateful to *my family, my amazing husband Eric Bowers, my kids Nick, Bobby, and Sophie.*

DEDICATION

"Education plays an indispensable role in the modern industrial state. This role, we believe, has two significant aspects: first, education is a major determinant of an individual's chances for economic and social success in our competitive society; second, education is a unique influence on a child's development as a citizen and his participation in political and community life.

Thus, education is the lifeline of both the individual and society."

~ 1971 Supreme Court of California – Serrano V. Priest

This study is dedicated to each one of my students, those who dedicate their lives to the pursuit of knowledge and believe that an educated society is a stronger society.

To my colleagues, who tirelessly and without hesitation, work to strengthen communities, close the achievement gaps, alleviate poverty, and increase college completion rates.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

POSITIONALITY STATEMENT

The first eight years of my professional career were spent working in advising offices. While working at CUNY in the Online Baccalaureate program, I advised a group of 900+ adult transfer students. This has shaped the way I see educational policies and programs, how I interact with students, staff and faculty. I started collecting data on my students as an advisor in 2006 and being rather skillful at Excel was able to see patterns of students' behavior based on transcripts, frequency and quality of interactions.

In 2012, I left my job at CUNY and took on a more administrative role. A role that required me to understand data and to produce freshman and transfer retention reports for the provost and the university community; the position I held when this study was conducted. One of my primary responsibilities in that role was to develop a thorough understanding of college retention and learn about risk factors of freshman, sophomore and transfer students. I have been immersed in predictive analytics, data reporting, risk modeling, and intervention design and implementation. Part of my day-to-day job responsibilities was to have access to data, make sense of it, visualize it, and explain developed reports to others. I took a chance and decided to dedicate the last seven years to expand my expertise and learn about big data, data extraction, blending, analysis, modeling and visualizations.

Data are everywhere. Companies, universities, and organizations collect data, both intentionally and unintentionally. Data doesn't always translate into information. Data files in their raw form are messy and can be manipulated, used unethically, and even illegally for the purposes of financial gain. Data governance is just as important as data

analysis. Universities need defined data standards and domain experts that can provide oversight on how data will be used, shared, and distributed.

One of the most significant institutional data challenges is the availability of longitudinal datasets by cohort and understanding transfer students and their enrollment patterns. These challenges have begun to be addressed at both the institutional level and nationally through initiatives by the Lumina foundation and the federal government. Two such examples are the establishment of the Committee on Measurement of Student Success and revisions to the Integrated Postsecondary Education Data System (IPEDS).

Good data allow for better analysis and in the long run inform policies. Even though there is an abundance of data, it usually doesn't translate into information. Higher education institutions collect data daily and at every transaction. These data in their raw and unorganized form need to be processed and organized. By developing uniform and organized data systems and designing databases that can integrate data from multiple student data systems, institutions can leverage the information by building out reports and data packages used not only for reporting, but for analysis and predictive modeling.

CHAPTER 1
INTRODUCTION

"Your life is a story of transition.
You are always leaving one chapter behind while moving on to the next."

Anonymous

A shift in the United States demographics and lower birth rates are causing a decline not only in high school graduates but in the application numbers of the entering freshman class (Grawe, 2018). According to Grawe, the post-2008 recession has not only led to a housing collapse, but a continuous and significant decrease in childbirths. In 2026, the number of college-aged children is predicted to drop by 15%, strongly impacting recruitment, admissions, faculty hiring, strategic planning, and the future of higher education for many years to come.

Figure 1.1: Total U.S. Public and Private High School Graduates (Actual and Projected) 1979 to 2032 (Hussar & Bailey, 2017)

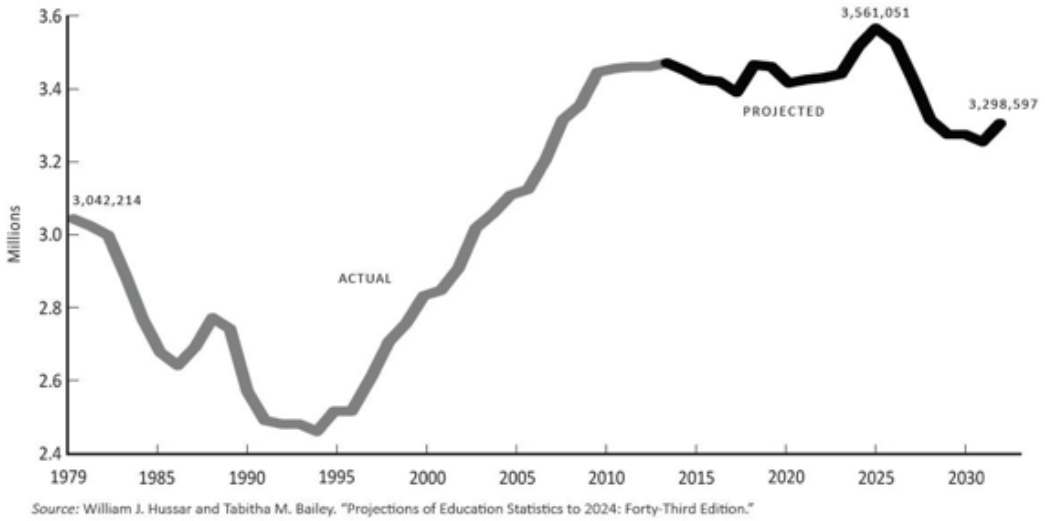
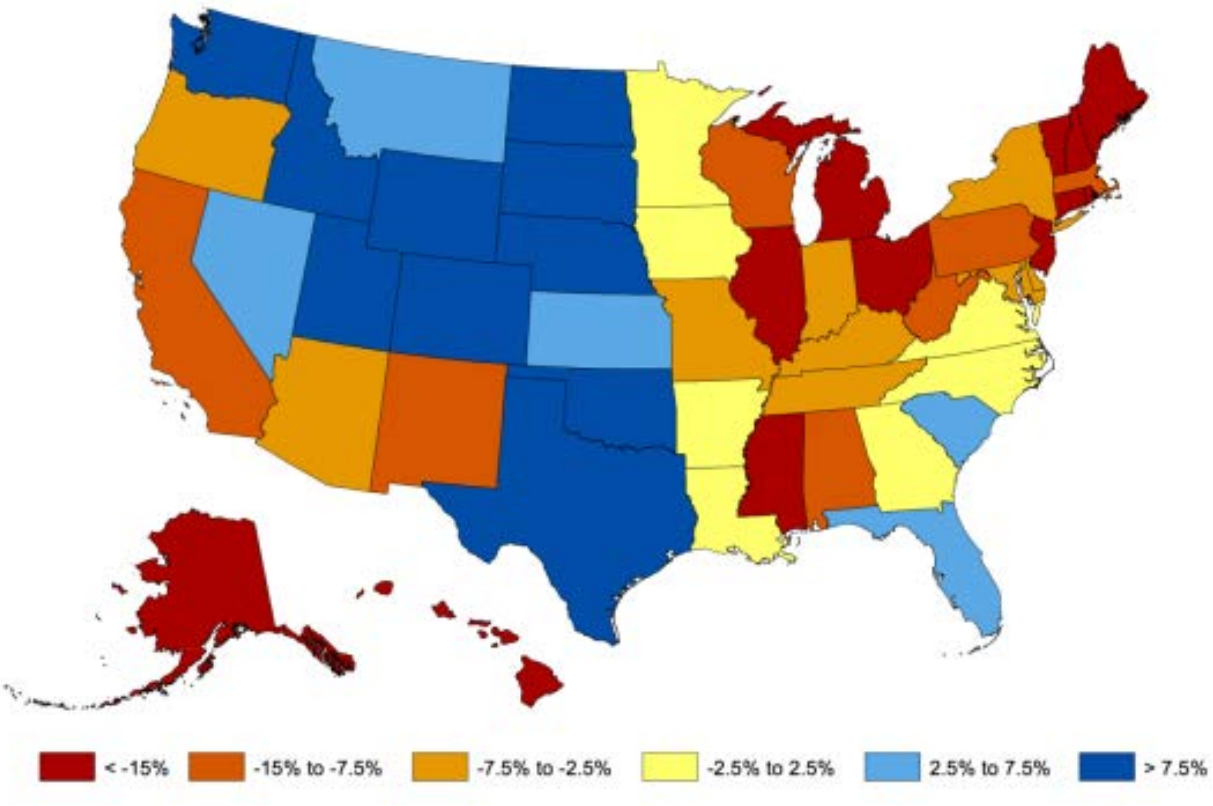


Figure 1.2: Predictions of decreases and increases in the number of first-year students enrolling at four-year colleges and universities by state (Grawe, 2018).



Note: Urban college markets and states in light and dark blue are expected to see an increase in enrollments and may diverge from national trends (Grawe, 2018)

As some colleges and universities begin to face declining enrollments of traditional freshman, many will begin to focus on transfer student recruitment, enrollment, persistence, and graduation rates in addition to building community college partnerships. The National Student Clearinghouse report – "Transfer and Mobility: A National View of Pre-Degree Student Movement in Postsecondary Institutions," published in partnership with Indiana University's Project on Academic Success, indicates that a third of students transfer before graduating. By studying transfer patterns of nearly 2.8 million students at all institutional types for five years, researchers at Indiana University found that transfer rates in the United

States fluctuated between 32.6 and 33.9% for full-and part-time students. Furthermore, during the examined five-year period, many students took breaks in between enrollments at different institutions, making them difficult to track, market to, and assess.

Every semester one out of every three new students on campus is a transfer student (Simone, 2014). Furthermore, only 41% of all transfers graduate from a four-year institution and obtain a degree; many students leave after only one semester or a year, never to return (Blackwell, 2018). Even worse, some transfer students run out of financial aid in their last year and drop out within a few credits of completing their degree. Although compared to freshman, transfer students are more experienced and have accumulated more academic credits and experiences before enrolling at a new institution, many are working full- or part-time jobs, have limited financial resources, and are less likely to seek tutoring or help with study skills (Noel-Levitz, 2011).

There are many types of transfer students. Some transfers begin their academic journey at the community colleges, others at another four-year institution; some have attended multiple schools or have taken online courses before transferring to their final college or university. Many students begin at community colleges because of proximity to home, affordability, or high school dual-enrollment programs; some take courses online and then enroll at a traditional brick and mortar institution; others complete certificate programs and years later hope to use some of those credits toward a four-year degree.

Types of Transfer Students

Education researchers have devoted considerable attention to identifying and describing the types of transfer students and their paths from one institution to the next (Joseph & Poisel, 2018), as well as to what happens to the students once they integrate into their new schools and the factors that contribute to their successes or failures, and time to graduation. Despite the research, most institutions have not improved their transfer student graduation rates and have not devoted resources to creating institution-specific risk models and interventions explicitly designed for retaining and graduating transfer students (Sissel, Hansman, & Kasworm, 2001; Tobolowsky & Cox, 2012; Walker & Okpala, 2017). Most of the literature and research uses labels to identify different types of transfer students (Laanan & Sanchez, 1996; Townsend & Dever, 1999):

- *Lateral transfers* are students who transfer from one institution to another institution of equivalence, such as transferring from one community college to another community college, or from one four-year institution to another four-year institution.
- *Upward transfers* move from two-year institutions to four-year institutions; many of them do so before completing their associate degrees.
- *Reverse transfers* transition from a four-year institution to a two-year institution, often for academic or financial reasons; these students often will transfer back to four-year universities later on.
- *Vertical transfers/Dual Admissions* students often declare their intent to transfer to a four-year institution, while still enrolled at a community college. These declarations of intent often serve as contracts between a four-year institution and the student and guarantee admission if the student reaches a specific GPA and attains an associate

degree. Dual admission students may also receive transfer scholarships from the four-year schools if they declare their intent to transfer early enough in their academic career.

- *Swirling Transfers* may move from one institution to another, accumulating credits, but not enrolling in any specific degree. Distance education, increased consumer culture, and rising costs of tuition all contribute to such student behavior. Students accumulate credits from multiple institutions and attempt to transfer them to a program that would accept them towards a degree. This is a risky proposition, but one that can pay off if students plan and research their options carefully.

Every time students switch schools, they will have to navigate new cultures, learn a new language, acronyms, policies, and protocols of their new institutions. Research from the Graduate! Network has shown that students who are not continuously enrolled or stop out for longer than 12 months are much less likely to re-enroll and graduate (Yanovski, Sheffer & Mattei, 2018).

The "Cooling Out Concept" was first coined by Burton Clark in 1960 to describe the function of junior/community college faculty in students' engagement or gradual disengagement with the pursuit of their educational goals. Articulation and dual admission agreements serve as retention tools that promote continuous enrollment and allow colleges to monitor and track students, while at the same time preventing them from getting "cold" or "cooling off"; in other words, preventing them from stopping out. Clark (1960) found that students who do stop-out often begin to show gradual disengagement with their academic life, detach from academic institutions, seek substitute achievements, and exhibit avoidance of standards as part of the "cooling out" process.

Why Students Transfer

Students transfer because of dissatisfaction with the size or quality of the institutional or geographic area, inability to pay high tuition fees, poor major fit or lack of major that suits their interests or career goals, personal circumstances, or failure to meet school academic requirements. Some are admitted to better institutions and leave to pursue more prestigious degrees. Students transfer immediately, or within months of leaving their original institution; others wait years and sometimes even decades before re-enrolling as adults. There are many reasons why stopped out students choose to continue their education at a later date. Some students seek a career change, some are mandated by their employer to pursue additional training, some are fulfilling personal goals, while others do so with the expectations of an increased earning potential. Adult transfers differ significantly from traditional-age students. Many adult students are working full-time jobs, have families, and a limited amount of time and resources required to allocate to finishing a degree.

Transfer students generally differ from traditional first-year students in age, demographics, academic and socio-economic characteristics (Mustafa & Compton, 2014). The majority of the students wait until the fall semester to begin at their new institution; however, some do transfer in the spring. About 60% of the transfer students fall in a traditional college-age group of 24 and under, with nontraditional students (over 24) comprising about 40% of the entering class (Santos, 2007). Transfer students are often first in their family to attend college, come from diverse family incomes and socio-economic groups, have different academic credentials, grade point averages, and select various academic majors. There are also significant differences across all of these factors from college to college and geographically (Mehta, Newbold, & O'Rourke, 2011; Smith, 1973).

Factors Driving Withdrawal

Transfers leave their new institutions for many reasons: some depart voluntarily, others as a result of academic failure. Voluntary withdrawal is often based on issues of student motivation, institutional fit, financial problems, misinformation, and lack of support. Transfer student motivation and goals may differ at different points of their academic careers, with some students shifting focus from academic pursuits to employment and family issues. Lack of information or misinformation about curriculum requirements, financial aid, and career options can lead to drop-out and disengagement with the new institution. Not knowing whom to ask, where to go, and what to expect from the interactions with faculty or staff can lead students to seek help from their family or friends and often become misinformed. Also, at some institutions, student-staff ratios are substantial and can lead to long wait times to get answers to questions.

Poor customer service, unsatisfactory experience with staff, faculty, or other students can lead newly admitted transfers to disengage and begin thinking about transferring elsewhere again. Lack of fit between student and institution can manifest itself in misaligned expectations. Students who struggle often underestimate the number of hours they are expected to study for face-to-face, hybrid and online classes and the time needed for reading and exams. Since many transfer students are required to select a major at the point of enrollment, selecting the wrong one can lead to poor academic performance due to lack of interest or engagement with class materials. Lack of support from family and friends and poor advising experiences can lead students to question the benefits and the long-term payoff of a college degree.

Even though most transfers have earned some college credits before transferring, many students are unaware of what it takes to be successful at their new institution. General skill deficiencies, poor academic performance, and gaps in prerequisite knowledge can manifest themselves in poor grades and the need to repeat failed courses. Lack of academic preparation, ineffective academic behaviors such as poor study habits, time management, and insufficient time devoted to studying can all lead to academic failure. University and departmental policies can also act as significant factors in students' academic success. For example, students who have previously learned to rely on tools such as calculators may find that not all math departments allow calculators for in-class quizzes or tests. Students who are used to the small classroom at their old institutions may find it difficult to adjust to the large lecture halls of their new school. They may miss timely and tailored feedback on their performance from their faculty, as they find themselves in lectures halls of 100+.

Problem Statement

As more and more emphasis is placed on retention and graduation rates by national accreditation agencies, colleges are willing to allocate time and funding to more structured and comprehensive support initiatives. Many of these initiatives are a direct result of the US Department of Education's 2020 goal to increase the number of college graduates in the United States. This goal was set by President Obama, on February 24th, 2009, during his first joint address to Congress (US Department of Education, 2011).

One of the ways colleges are addressing the issues of retention is by proactively outreaching to students identified "at-risk" with the help of predictive models. According to Narrz'ello (2014), the "at-risk" student label at colleges and universities was popularized during the 1980s. The term began to carry future orientation, describing students who were

likely to experience problems at some point in the future. By 1994, 31 states have adopted the term “at-risk” as a label in educational legislation. Frost (1994) notes that the “at-risk” definition encompasses environmental factors as well as individual characteristics. The legislation directed at at-risk students usually provides resources and programs in order to improve educational outcomes. Such programs include National Science Foundation (NSF) Grants, Pell Grants, TRIO (Upward Bound, Talent Search, and Student Support Service) Grants, McNair Scholarships, Educational Opportunity Center, et cetera.

The definition of at-risk began to shift from an arbitrarily assigned label based on common knowledge or generalizations to a category assigned as a result of a statistical model based on empirical data analysis. According to the 2015 Chronicle of Higher Education Special Report on Big Data, only 8% of public colleges and universities and 24% of private, not-for-profit universities have been using predictive analytics for more than five years (Selingo, 2015). Most of the literature on student risk-identification points to the use of regression modeling and purchasing off-the-shelf software that promises a panacea solution to institutional retention problems.

During the last fifteen years, universities began not only adapting off-the-shelf models but designing their own. With the use of statistical tools, colleges and universities build models and identify characteristics that can classify at-risk first-year students during different stages of the student lifecycle: admission, orientation, first term registration, first term mid-semester marks, and subsequent term enrollment periods. However, most institutions are still focused on developing freshman models and do not create algorithms to identify transfer students at-risk of not being retained at their new school. No existing research outlines or explains methods for developing transfer risk models and ways to assess

their effectiveness. The logistics of capturing necessary information, the right amount of historical data points, and knowing how to build predictive models for this population prevents many institutions from developing an empirical modeling approach.

Purpose of the Study

Identifying transfer students at-risk of dropping out using statistical models can lead to the development of interventions that can reduce attrition and increase persistence and graduation rates. My past experience analyzing institutional data shows that transfer students are much more likely to leave their new institution after attending for only *one or two* semesters compared to their first-time freshman classmates. There are four main reasons for conducting this study:

1. As institutions begin to grow their adult and transfer student cohorts, by marketing and enrolling more nontraditional students alongside traditional first-year students, administrators will need to gain an overall understanding of the demographics and composition of their transfer student population.
2. Departments will have to take a closer look at the services they provide and the effectiveness of these services at supporting nontraditional student populations.
3. Deans, faculty chairs, and administrative units will need to examine their policies and processes and the capacity to handle transfer student issues.
4. Lastly, this study will provide a blueprint for anyone who would like to develop an in-house risk model at their institution using software like SPSS Modeler, Tableau and Excel.

One way to view the transfer lifecycle is through the enrollment management lens: students apply, enroll, take classes and graduate. If something goes wrong in the cycle (e.g.,

the student takes the wrong class, changes his or her major) the retention algorithm is off. Statistical techniques like Configural Analysis, done with the help of IBM's SPSS Decision tree module, are used to assist in systematic post-positivist theory development and testing, as was done in this study. This method of synthesizing data uses chiefly quantitative approaches to test hypothesis and predict outcomes.

In conducting the research of this type, it is important to identify the datasets and the variables which were used in the development of the model. After the availability of the data elements is determined and the data source systems are identified, data must be extracted from the Institutional Datawarehouse or gathered from institutional functional and technical contacts responsible for each data source or database. Functional contacts are tasked with identifying data extraction methods and work technical contacts to safely transmit and store the data in accordance with the university's policies and procedures. Most institutional data are derived from the Enterprise Reporting Systems (ERPs) and auxiliary systems bought and managed by individual departments. The primary tools used to assemble the data are SQL Developer, Microsoft Excel, Tableau Prep and Desktop, SPSS Statistics and SPSS Modeler.

IBM SPSS Decision Tree model allows the researcher to determine relationships between variables and make predictions of future events. By analyzing transfer data with this method, researchers can uncover relationships between academic and psychometric variables. The decision trees themselves provide highly visual classification through dendrogram visualizations. These visualizations are useful in explaining the statistical significance levels and variable significance hierarchy. The visualizations are also useful in intervention planning and policy design discussions. SPSS decision trees allow the researcher to can gain insight into a subset of the data (such as Veteran Transfers or transfers with less

than 60 credits) by conducting further analysis on this sub-group. This way of looking at transfer student experience provides a big picture of student behaviors, without examining individual student characteristics. This type of analysis is useful when making changes to policies.

Significance of the Study

Preliminary examination of data at a large urban public university in Philadelphia indicates that close to 40% of all newly enrolled students are newly admitted transfers. Many of them are enrolling after attending community colleges or other four-year institutions, many coming from known "partner schools." However, first-year retention, graduation rates, and GPAs of these transfer students are lower compared to traditional first-year students according to my prior research as part of my full-time position. In addition, most administrators and academic departments have yet to identify and implement a comprehensive institutional best practices model that would ensure transfer student success.

In this quantitative study, I went through the process of developing a longitudinal dataset necessary for data mining, explained the process of data reduction and analyzed a longitudinal dataset comprising of three years of transfer data collected from a large urban institution in Philadelphia. I examined admissions, orientation and term data for 2015, 2016, and 2017 cohorts independently and together, and developed a predictive model for each year and one that uses combined data for all three years. In the process, I examined the role of gender, race, and socioeconomic status on retention and major/college choice. Since each student admitted to the university must complete a new student questionnaire, these records were compared to the retention metrics to ascertain correlations. Further, explanations were provided as to other factors that may affect retention. This study used Schlossberg's

Transition Theory and Tinto's theory of student departure to both position the research and guide the variable selection process.

Data-driven and informed enrollment management messaging, recruitment, and programmatic offerings can help meet the needs of transfer students. Using historical data can help build a student-centric systems approach that integrates student services with academic affairs and in the long term, improve student retention and graduation rates. Data presented in a digestible, easy to understand manner can help faculty and department chairs make course scheduling decisions, establish partnerships with feeder institutions to collaborate on curriculum alignments, and provide support to students who may not succeed due to lack of pre-requisite knowledge. In addition to providing services to transfer students, institutions will need to re-engineer many majors and academic programs to be transfer-friendly. Faculty and department chairs will need to design academic pathways that focus on critical thinking, content knowledge and skill development, while at the same time aligning curriculum content with the local economy and job market needs.

Since it is possible and most likely that multiple factors cause transfer students to withdraw without completing their degree, the ability to identify and predict who these students are and develop targeted interventions for them would be ideal for any institution with large transfer enrollment numbers. Since students at a small private institution in the middle of Kansas will probably differ significantly from students at a large public institution in the middle of Manhattan, schools must use their own empirically developed data indicators to predict when a student is "at-risk" of not being retained. These data indicators would include variables from the student lifecycle, starting with pre-enrollment data from orientation, housing application, financial aid, and placement scores and then be combined

with post-enrollment data such as mid-semester alerts, as well as social and academic experiences during the first and second semester.

Research Questions

1. What are the determinants of retention among transfer students?
    a. What are university-wide risk factors of dropping out for transfer students?
    b. Do they differ by race, socioeconomic class, or gender?
2. Do risk factors change or stay consistent from year to year?
    a. 2015, 2016, and 2017
3. What are college-specific (e.g., Education vs. Engineering) risk-factors of dropping out?
4. What are the determinants of graduation among transfer students?
    a. What are university-wide risk factors of being able to complete a degree in three years for transfer students?
    b. Do they differ by race, socioeconomic class, or gender?

Definitions of Terms Used in the Study

Academic Residency Requirements: Most institutions require that all undergraduate students complete a minimum number of credits at the school that issues their diploma. These requirements differ from school to school. For students in this study, residency rules mandate that they complete the last 45 hours of the last 60 hours of the degree. The residency requirements were revised in 2013 to increase the number of credits from 30 to 45, thus making this study possible. Previously, students with a high number of transfer credits could complete their degree in one year by taking 15 credits in the fall and 15 credits in the spring (given that all pre-requisites and degree requirements were met).

Cohort: A group of transfer students who began their education during the fall semester in a particular term. This study examines fall 2015, 2016, 2017 cohorts of newly admitted transfer students.

Freshman: First-year students at a university or college, who have completed high school graduation requirements, but have not completed any college credits after graduating from high school. Although this study mainly focuses on the experiences of transfer students, some parallel comparisons will be made to examine differences between freshman and transfer groups. Technically, first-year students may come to college with more credits than a transfer student. Freshmen may earn college credit via Dual Enrollment programs, Advanced Placement or International Baccalaureate, CLEP, as well as A-Level Examinations. Some first-year students may enter four-year institutions with 30 plus college credits. These students are still considered freshmen since all of the credits were earned before graduation from high school.

Retention: Often referred to as persistence, retention is defined as the continuous enrollment of students from one term to the next, and is often measured in academic years as opposed to semesters.

Transfer Credits: Any course credits that are recognized by the accepting institution, transfer credits are often not posted to the student's academic record at the point of admission. Students must submit final transcripts for any courses taken before admission in order to receive credit. A delay in submitting final transcripts can cause registration and scheduling difficulties, as well as inaccurate advising. Transfer credit issues will be discussed in more detail during the literature review. Students with less than 15 transfer credits are considered freshman. This definition is specific to the students selected for this study and may differ depending on the institution.

Transfer Applicant: A transfer applicant is any student who has completed at least one hour of college credit after graduating high school. Transfer students in this study will all have 15 or more college-level credits after high school graduation before enrolling at their new institution.

Retained: Retained students are registered for at least one academic credit during their second and third semesters at the institution they enrolled in as a transfer student.

Student Success: Student success is a broad concept, and most of the time it refers to successful academic and social integration on campus, as well as students' ability to secure employment after graduation. For the purpose of this study, success is defined as successful completion of the academic term for which the student enrolled with a GPA of 2.0 or better. Academic policies often place students whose GPA fell below

2.0 on academic warning or probation. Multiple semesters with a GPA below 2.0 can lead to academic dismissal.

Graduation Rates: Graduation rates serve as a key metric of accountability in higher education. They are calculated by counting the number of years a student takes to graduate from their semester of admission. For example, the institution may examine 2014 freshmen who began in the fall and calculate the number of semesters/years it took for them to graduate. Transfer cohort graduation rates are somewhat difficult to calculate, as many transfers start with a different number of credits at the point of enrollment.

Time to Degree: Similar to graduation rates, time to degree examines the number of semesters it took someone to graduate. However, instead of looking at the initially admitted cohorts, institutions calculate the time to degree by reviewing statistics about the graduating class (everyone graduating in May of 2018, for example) and calculating the amount of time it took for these students to reach graduation.

First Generation: A first-generation transfer student is a student whose parent(s)/legal guardian(s) have not completed a baccalaureate degree. Students with siblings in college are still considered first-generation. According to Bui (2002), first-generation students are more likely to fall into one or more of these groups: immigrants, nonnative speakers of English, child of a single-parent household, or emancipated minor.

Assumptions & Limitations

The impact of big data on school policies, reforms, and student learning has forever and will continue to change the landscape of the American Higher Education system. As universities begin collecting all sorts of data, many ethical questions remain unanswered. Student swipe data, location data, or course registration data all become valuable information and knowledge someone may be willing to pay for and possibly abuse. Of course, data in themselves are meaningless. Those who analyze and report data are the ones left holding the key to the future of tomorrow's reforms and policy changes. Researchers who develop predictive models and conduct student data analysis must do so with much caution. Binary variables used in the models, even though they represent human beings, do not tell the full story. For every zero or one in the dataset, there is a story, a story that often involves human struggle and years of beating the odds to get to college.

Often those in charge of the institutional data do not communicate directly with data consumers (faculty chairs, administrative unit directors, and deans). It is not uncommon for colleges and universities to have a lack of communication or siloed communication between these university groups. Students must be aware of the data that are gathered about them and have a say in what information college shares with third parties and how this information is used.

Risk models must be university-specific, based on historical student data and provide empirical evidence of their accuracy. They must be transparent to those who use the data and design interventions. The reasons a student is identified as at-risk and the factors that led to this classification should be transparent and explainable. Lastly, intervention design should not be merely data-driven; it should be data-informed. In the end, no machine will ever be

able to predict emergencies, major life events, or other life factors that can happen at any moment.

Possibility and Limits: Dark and Dirty Data

Higher education institutions are often data rich and insight poor as a result of siloed communication and data storage. Often each administrative office will manage their own Client Relationship System (CRM) or appointment system, with admissions, placement testing, residence life, financial aid and other historical data extracts stored on someone's computer. These data silos exist in isolation from the larger campus community. In 2012, when I began analyzing data in my professional role, I quickly learned that the process of developing predictive models in the academic setting is exceptionally nuanced. In order to build a predictive model, I had to have a thorough understanding of the available data sources and variables at the institution.

Data access, aggregation, understanding, and data blending are four key processes necessary for beginning the process of building a predictive model. The data extracted for this study, were provided by the institution's business intelligence team. It was mined from a data warehouse student retention package that took two years to develop by a team of university administrators, data engineers, and data architects. The student retention package is a Cognos data package that merges multiple data sources of data at the university using one primary key – student's id. The Datawarehouse was developed in 2016 and did not contain all historical snapshots that were required for this research project. The Institutional Research team and the Office of the Register were able to help fill the gaps where more data or historical data were needed. Since part of the dataset used in this study came from the Datawarehouse and the other data files used in developing the multiyear cohort file came

from excel spreadsheets, many variable names did not align from one year to the next. In order to align 2015, 2016, 2017 cohorts and stack the data together, a significant number of variable names had to be changed. Data blending was performed using Tableau-Prep software package. The availability of the historical data, access to the correct snapshots and knowing whether to pull the data from the Datawarehouse, to use live or historical data, all required thorough understanding of data governance – data availability, integrity, security, and usability.

When measuring student success, retention, or persistence, the definition of the term "at-risk" plays a significant role in the methodology and the sources of data to be collected in order to obtain accurate results. One of the biggest challenges of measuring retention for transfer students was the pattern of enrollment observed when analyzing registration patterns of this population. Transfer students were much more likely to be enrolled part-time, take one or two semesters off, and then return and continue their education. Since the models developed in this study looked to predict retention into the third semester (students second year at the institution), it became quickly apparent that students who were not registered in term three could potentially come back and continue their degree later on. Some did and even obtained a degree. Thus, the transfer model predicts persistence into the third term, and retention when using conventional definitions of these two terms. It would be challenging to predict retention without looking at multiple semesters of enrollment and collecting data spanning at least 5-10 years.

The same dilemma presented itself when trying to predict graduation. I made a conscious choice to examine whether or not transfer students from the fall 2015 cohort were able to graduate in a three-year time frame. I selected a three-year mark because students can

receive Federal Pell Grants for no more than 12 semesters (roughly six years of enrollment). Logically, if students attended community college for three years, and then transferred, after three years at their new institutions those students receiving federal aid would run out of funds. This assumption generalized transfer population persistence trends and assumed a path to graduation via a 3+3 enrollment pattern, with the first three years spent at a community college and the last three at a four-year institution. Through my preliminary research done as a part of my full-time job, this proved to be an accurate path for only a small portion of transfer students. Thus, the decision to use a 6-year graduation timeline was simply a baseline for helping me guestimate a reasonable time to degree completion.

The National Student Clearing House defines *retention* as "continued enrollment (or degree completion) within the same higher education institution in the fall semesters of a student's first and second year" and *persistence* as "continued enrollment (or degree completion) at any higher education institution — including one different from the institution of initial enrollment — in the fall semesters of a student's first and second year" (National Student Clearinghouse p. 11). This study uses the term persistence interchangeably with the term retention, with an understanding that all transfer students come in with credits from one or more institutions. It is valuable to note that some transfer students in this sample began their academic journey at institution A (the university where the study is taking place), transferred to institutions B, C, or even D, and then came back to institution A to finish their degree. These types of students are considered to be transfer and a re-enrollee. They begin new degrees and enroll in completely different programs to what they originally applied years ago. Some of these students will seek historical curriculum exceptions. These exceptions are rare, but if granted, allow students to pursue the same program they were

enrolled in at the point of admission during their first term of enrollment. Different institutions have different rules for granting historical curriculum exceptions. For the most part, these exceptions are only granted if clerical errors have occurred, if students need only a few credits to graduate, and if the courses needed for graduation still exist and the degree is still available for graduation.

Many transfer students are not aware that the full-time academic course load is 12 to 18 credits. Students must be registered for at least 12 credits to be considered full-time. Most institutions charge the same tuition rate for credit loads between 12 and 18 credits. Community college students are often used to a per-credit tuition pay structure, and many are not used to taking more than four courses a semester. Given that many programs are structured in a way that requires students to take 15 or 16 credits each term in order to graduate in a timely manner, transfer students may need additional time to graduate because they are not completing a minimum of 30 credits a year. There is an intricate balance between working, going to school full-time, and family obligations that each student needs to manage. One of the limitations of this study is the omission of the leave of absence data. Students who file for a leave of absence, typically request one or two semesters off from their studies with the intent to return. Students may or may not return; however, they are less likely to come back if they carried a significant monetary balance on their account. Follow-up studies need to be done to understand this population and the likelihood of re-enrollment, persistence, and graduation of someone who has left.

One of the datasets (New Student Questionnaire or NSQ) administered during students' orientation to the university had been revised, with question 4 during 2017-year NSQ administration significantly deviating from the 2015 and 2016 administrations. The

final dataset of stacked data now has question 4 and question 4a as two distinct variables. Students in the 2015 and 2016 cohorts will have Q4A blank, and students in 2017 cohort will have Q4 blank. A predictive model that relies on a survey like the New Student Questionnaire will be severely limited if the survey changes from one year to the next. This change will create difficulties for anyone who would like to develop multiyear cohort files, or if predictive models use variables that no longer exist from one year to the next.

One of the most significant challenges in working with the datasets is the consistency of variable names from year-to-year and term-to-term. If student survey data are used in the analysis, numbering of questions, and changes to the format of the survey should be kept to a minimum from one academic year to another. Data need to be scanned for outliers and data entry errors. GPA of 3.15 can easily be entered in as 315 and skew the entire analysis. The person scrubbing the data needs to account for missing variables and duplicate records, variable names, level of measurement, and variable types. Variable labels and data dictionary need to be created before analysis is done.

In this study, I am built models for each academic year: 2015, 2016, 2017, and then stacked the data to run one model using all three years of data. Differences by year, college, race, and gender were conducted throughout.

To produce empirical results, the typical process of building a predictive model involves using a historical dataset that has matured (at least two years old) or stacking multiple years of data together. The one-year worth of data can be highly unbalanced - with a very small percentage of students dropping out. This is especially true when examining the retention of a smaller program or college (i.e., Undeclared or Performing Arts Students). An unbalanced dataset means that the majority of the cases for outcome retained (1) vs. not

retained (0) will be set to 1, retained. Predictive models developed using an unbalanced dataset run the risk of having lower precision and high accuracy, thus predicting that all students will be retained (1). An unbalanced dataset will bias the prediction model towards the more commonly observed outcome.

As mentioned earlier, stacking data to perform multi-cohort, multi-year analysis is arduous since some data sources and variables change significantly from one year to the next. Student survey data collected from the NSQ used in this study were difficult to stack, as not only the questions changed in content and their relative position within the set, the coding of the questions changed as well (A=1, B=2 vs A=5, B=4). Another challenge of using NSQ survey data is a significant number of student records with missing data.

Future studies that examine transfer student registration, enrollment, and graduation patterns will need to address same or similar issues. Transfer data are rarely analyzed and often getting a longitudinal dataset that tracks transfers through their academic lifecycle is challenging. The final dataset used in this study was developed by blending 23 different files and took 60 plus hours.

CHAPTER 2
THEORETICAL FRAMEWORK AND LITERATURE REVIEW

Literature Review

Students' first impression of the university typically begins with an orientation, where they meet staff, faculty, and other students for the first time. However, most orientations cater primarily to first-time freshmen, students who just graduated from high school. Many administrators report that an institution's academic culture and policies often work against transfer students, not recognizing the specific needs of this population. Tobolowsky and Cox (2012) determined that institutional policies and structures, internal and external environmental conditions, as well as personal perceptions of faculty and staff directly impact transfer student success.

The institutional response to transfer student success and issues of retention should be championed by academic and student affairs, faculty, and administrators. Success for transfer students may mean attainment of first semester grade-point averages at a new institution that exceeds or matches the GPA of the first school the students attended; thus *not* experiencing "Transfer Shock," a widely known phenomenon identified by Hills in 1965. Success may also mean making it to graduation without running out of financial aid. Furthermore, issues of transfer credit articulation and academic scheduling should be addressed holistically and before students arrive at their new institutions. Most of the recent literature on transfer students suggests that in order to aid transfer student adjustment and mitigate transfer shock, institutions must focus on both policies and practices, as well as individual and institution-specific factors that shape the transfer student experience (Glass & Harrington, 2002; Joseph & Poisel 2018; Kuh, 2003).

Academic scheduling plays a significant role in the success of any transfer student, and for the most part, is done during transfer orientation, leaving very little time for social interaction or any holistic or developmental advising. For many transfer students, orientations consist of credit re-evaluations, course prerequisite verification, and scheduling students into whatever courses are still open. Most transfer orientations last less than a day, and students often get shuffled from one office to another; typically, from advising to bursar to financial aid. Also, transfer students bring with them "a range of assumptions and expectations about college that were developed at their previous institutions" (Tobolowsky & Cox, 2012, p. 397). Students often assume policies and regulations that held true at their home institutions also hold at their new school. Advisors and faculty have a very short window during orientation to deliver crucial information that will impact student's progress throughout their degree. Assessing the needs of transfer students during orientation is a difficult task, and most advisors are limited to prescriptive advising based on the limited information they have about the student at the time of orientation (Mayhew et al., 2010).

Comprehensive transfer orientation is just as critical as the orientation offered to first-year students. The format and the delivery of transfer orientations play an integral role in how the information is retained, the level of engagement from the participants, and the impact on retention and possibly even graduation rates. Orientation formats can take many shapes – one or multiple days on campus orientations, on-campus orientations followed by a full or a shortened credit or non-credit courses that are taught online or in person, hybrid or fully online orientations that utilize technology like Skype, Google Hangout, Facebook and learning management systems (LMS). Online Orientations can be asynchronous or

synchronous, can be delivered at the point of admission, after the student deposits and commits to the institution, or right before the beginning of the semester.

Transfer Credit Evaluations

The ability to receive a clear picture of what credits transferred and what credits did not is one of the significant obstacles for transfer students. Many transfer students are misadvised and do not register for correct courses during their first semester at their new institution because of missing transcripts, forgotten AP/IB credits, outdated transfer equivalency rules or incorrect course evaluations. Universities that admit large numbers of transfer students may be guilty of significant delays in having transfer credits evaluated and posted to student's academic record. As a result, many students miss the opportunity to register on time because their credit counts are inaccurate. These delays also cause their academic transcripts to be incomplete, leading to potential errors in course planning, incorrect student standing, financial aid, and reduced loan eligibility. By improving internal processes and allocating budgets for specialized transfer admissions counselors, universities can improve the speed and accuracy of transfer credit evaluations, which would improve student record accuracy, provide students with the ability to register on time and avoid errors caused by inaccurate academic records.

Researchers used focus group format to interview recently transferred community college students about significant sources of anxiety they faced after admission. As predicted, one of the significant sources of anxiety was whether or not credits earned before admission transferred to the new institution. Students reported that it took their new institution significantly longer than expected to evaluate and post their community college credits to their account. This delay resulted in an inability for students to register for courses that had

prerequisites, register for courses they have already taken – resulting in duplicate credit, and receive incomplete advising information about what is left for them to complete in order to receive their degree. This problem was compounded with the fact that many transfer students were admitted to their new institution late into the admission cycle, and had very little time to complete administrative tasks, like setting up accounts, financial aid, e-mail, and parking. In addition to dealing with administrative hurdles, researchers found that many transfer students were also dealing with academic anxiety.

Curriculum Congruence & Articulation Agreements

Although for many transfer students academic anxiety eventually goes away, for some the issue is much larger and originates from the need for remediation in basic math, writing and science skills. There are significantly high correlations between the attainment of the baccalaureate degree and a transfer student's gender, SES, rigor of high school curriculum, self-expectations, GPA, college involvement, and math remediation (Wang, 2009). The math curriculum at the community colleges does not always match up to the math curriculum at the four-year institutions. This curricular misalignment may force students who have completed two or three levels of college math to go back and repeat courses they have already taken and push back their graduation dates (Long, 2005). This may cause them to take on more debt and limit them from taking upper-level courses that have math pre-requisites.

For the last twenty years, sometimes even earlier, colleges have been diligent in creating course articulation agreements to allow for more effortless transfer of credits. However, the articulation systems are not comprehensive, and in some instances, courses have a shelf life of five years from the time the student took them. This policy is typical for

business law courses as well as upper-division science, math, and other pre-professional courses.

In addition to the course articulation rules, four-year institutions have been working to develop articulation agreements for major electives and general education requirements. These agreements are supposed to facilitate seamless transfer from one institution to another. Successful articulation agreements are a collaboration among faculty and administrators at both sending and receiving institutions. Typical articulation agreements also called 2+2 program agreements, require students to take two years at a community college, earn the Associates Degree, transfer to a four-year institution, and complete the last two years there. In reality, once transferred, most students do not complete their degree in two years. Students must be aware of the existence of the agreement and follow the curriculum precisely as described. Any deviation from the suggested course plan or variance in the type of classes students take for suggested electives can result in additional time to degree and more debt. Even if students follow the prescribed plan as outlined in the articulation agreement, they must show proof of degree attained by submitting final transcripts. Students often forget this last step or assume that their first institution will send the transcript on their behalf. The cost of sending final transcripts varies by institution but is very small compared to the cost of having to take 15-30 extra credits that are sometimes not reflected on the earlier submitted transcripts used to make the admission decision.

A 2012 study by Montague discusses the benefits and limitations of articulation agreements, focusing on recruitment and financial benefits for the receiving institutions, particularly standardization of course credits across the institutions and smoother admissions process. Montague notes that there is a need for evaluation of transfer agreements which can

be achieved by tracking students and identifying successful agreements. Faculty and staff interactions and personal factors like relocation, family, and illness often contribute to early transfer and failure to follow articulation agreements and completion of Associate Degrees (Cejda & Kaylor, 2001).

Critical Pathways for Transfers

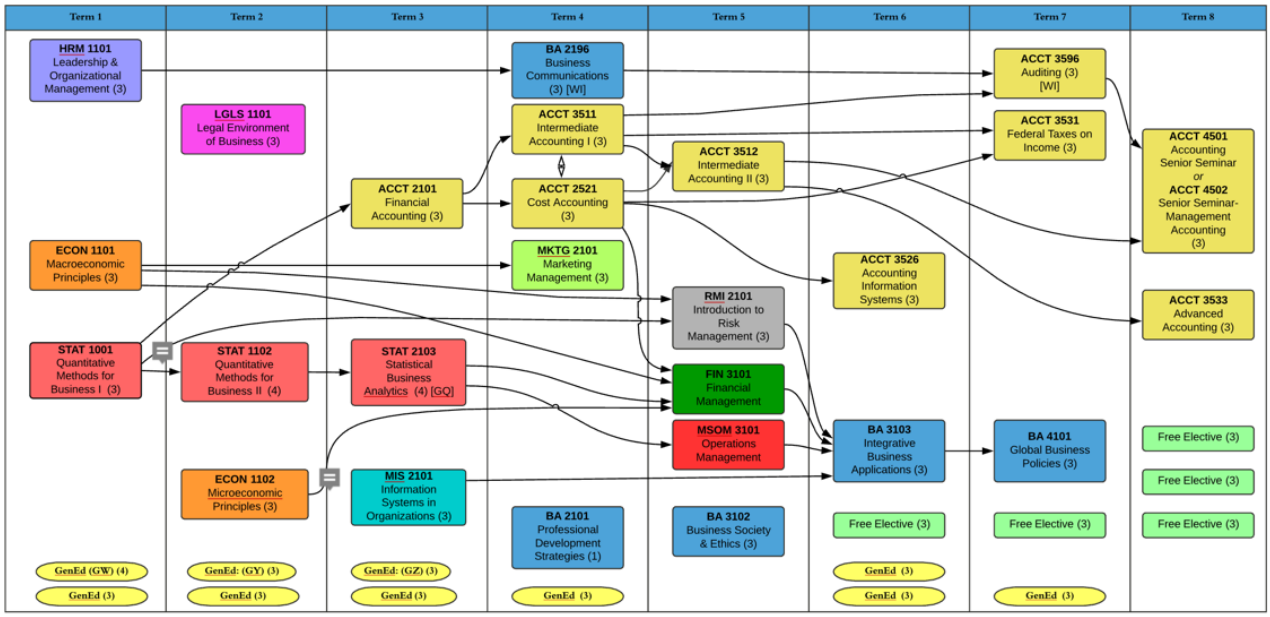
Early 2000 marked the beginning of Guided Pathways implementation at both community colleges and four-year institutions. Development of structured pathways requires participation and planning from faculty, deans, and administrators. Pathways outline and guide course selection and sequencing. The term 'critical paths or guided pathways' describes a review process that requires each school/college to examine how students move through the major. The main outcome of the review is a creation of an eight-semester plan, a pre-requisite map, and a student roadmap, which specifies milestones students must take in order to obtain a degree in a specified amount of time. Students are given much less choice in what courses they can take, but instead provided a blueprint/map with suggested curriculum and set of courses for each academic term that they are enrolled. Proper curriculum review can help identify institutional barriers to degree completion and fix bottlenecks and problem areas for students and majors. Data collected from these initiatives help provide an early indication of any student who is off track.

Departments/colleges that have completed the guided pathways program review process can identify timing for critical courses, and co-curricular requirements, as well as difficult pre-requisite structures students must complete in order to stay on track. Curricular Flowcharts are developed after a thorough review of each major is conducted, and a grid with suggested sequencing for how a student must progress through the degree is developed.

Guided pathways that are properly designed can improve time to degree completion and are extremely useful for working with transfer students and for development of articulation agreements. Curriculum flowcharts/maps visualize how students move through course sequences and identify possible challenges with pre-requisite structures, the timing for critical courses, issues with course availability, and scheduling. Sample milestones can include college and overall GPA, a total number of credits completed, specific placement test codes, and/or specific courses or grades in those courses.

Guided Pathways implementation is also useful during the review of new programs and course proposals. As a result of the Guided Pathways implementation, academic advisors can be better prepared to help students make informed decisions about taking the right courses, at the right time, and in the right sequence. This process becomes essential for transfer students and transfer counselors to understand pre-requisite sequencing and know the course equivalencies across institutions. Even more important is the increased need for transparency for this information for each school or university granting baccalaureate degrees. The following diagram is an example of the accounting pathways, with horizontal arrows symbolizing pre-requisites and vertical arrows co-requisites. Courses are color-coded to differentiate between free electives, general education, major requirements, and by department/subject.

Figure 2.1: An example of a Guided Pathways done for an Accounting Major



Guided Pathways significantly reduce the guesswork in deciding what classes a student needs to take. However, having this plan alone is not enough for students who wish to see how one institution compares to another or how their credits will transfer. In order to facilitate a smoother transfer between institutions, several states have implemented voluntary or mandated common course numbering systems. At the time of this writing, common course numbering has been mandated in fifteen states (Education Commission of the States, 2019)

In a common course numbering framework, both four-year and two-year institutions agree on a standard convention for naming and numbering courses — for example, MATH 1001 for College Algebra, and BIOL 1001 for Introduction to Biology. Thus, a student would not have to face course-by-course evaluation at their new institution, and the course from the sending institution would carry the same designation as the course at the receiving institution. Quality articulation agreements between institutions that do not follow the common numbering would include a course-by-course equivalency table along with the guided pathway alignment. In the example below, the coursework on the left represents suggested

courses student should take at their community college, and the table on the right represents the courses and their equivalencies at the receiving university. An additional table with a suggested course schedule is then developed to identify the courses the student is left to complete after he or she has been matriculated at the new institution.

Figure 2.2: An example of a Guided Pathways/Articulation agreement that represents course alignment between sending and receiving institution.

Four Year Institution			Two Year Institution	
First Semester		***Credits***	***First Semester***	
COLL 101	College Success Seminar	(1)	NON TRANSFERABLE COURSE	
COMP 110	English Composition I	3	ENG 0802	Analytical Reading & Writing
CRIJ 100	Introduction to the Administration of Criminal Justice	3	CJ 1001	Introduction to Criminal Justice
SOCI 110	Introduction to Sociology	3	SOC 1176	Introduction to Sociology
POLI 120	American State and Local Government	3	POLS 2102	American State & Local Politics
HIST 151	U.S. History: Young America	3	HIST 1101	History of US: 1600-1877
	Semester Total:	**15 (16)**		
Second Semester			***Second Semester***	
COMP 111	English Composition II	3	ENGL LXXX	English Elective
CRIJ 110	Introduction to Criminal Law	3	CJ 2501	Introduction to Criminal Law
CRIJ 160	Introduction to Juvenile Justice	3	CJ 2001	Juvenile Justice
PSYC 110	Introduction to Psychology	3	PSY 1001	Introduction to Psychology
HIST 152	U.S History: Modern America	3	HIST 1102	History of US: 1877-Present
	Semester Total:	**15**		

Appropriate Placement

Transfer students, just like first-year students, must complete appropriate placement tests to register for classes. Placement testing happens in multiple formats, online and in-person, and is administered directly by academic departments or through a third-party vendor, such as Accuplacer or Aleks (James, 2006). Students who are accurately placed into their Math and English courses form more homogenous classes, thus eliminating review and back-tracking of content done by the faculty member of those courses. Accurate placement may also prevent student withdrawal from a course late into the semester, and place only those students who are well-prepared into upper-division courses (St. Clair, 1993). University transfer policies often mention that transfer credits will trump placement; however, academic advisors and department chairs often require students who have a misalignment between their placement score and transfer credits to re-test in order to prove competency. Although

transfer students may find mandatory placement testing unreasonable, in the end, testing will facilitate success later on in their academic careers.

General Education Courses

For students who do need to take general education requirements at their new institutions, those courses must prove to be both informative and useful. Often transfer students perceive general education curriculum as a hurdle they need to overcome in order to take their major coursework. To be valuable, general education courses must blend liberal arts education with teaching of technical skills. By integrating skills desired by employers into the course assignments, students will develop a much higher appreciation for their required coursework. Liu's and Murphy's research (2013) suggests blending 21st-century skills like mobile application design, data analytics, programming and database design with liberal arts curriculum, thus converting students from technology consumers to technology creators and developers of new pedagogical strategies.

Mentoring, Social Awareness & Skills

Much has been written about the importance of social integration and academic and family stressors that contribute to attrition. Transfer students from working-class, low-income families are often the first in their family to attend college and may not have access to mentors, emotional support systems or other positive role models. Those who identify as immigrants, LatinX, Black or other minority groups are heavily represented in the transfer population (Cerezo & McWhirter, 2012). Lower expectations, lack of parental support, lack of knowledge about resources, and hesitation to seek out help from staff and faculty all play a role in transfer student retention. Research has shown that students who are invited to participate and engage with peer or faculty and staff mentors are much more adept at

developing skills necessary for facilitating social adjustment to college (Thile & Matt, 2005). Without these supports and proper mentoring students from socially or academically disadvantaged backgrounds, especially those who start at community colleges, may experience high levels of stress, social stratification in higher education, and often leave college before completing a degree. The psychological well-being of transfer students often depends on their ability to get involved in campus social activities, successful navigation of financial aid, academic and career counseling, and successful completion of academically rigorous coursework (Lee & Frank, 1990, Liu, Stevens, Wong, Yasui, & Chen, 2018).

Many transfer students reside off-campus, commute, and have family responsibilities as well as employer obligations. As a result, many students miss out on opportunities to meet with faculty, attend evening campus events, participate in student-faculty research projects, participate in study abroad, or service-learning. All of these activities correlate with environments that foster student learning and development and promote retention and student engagement (Kuh, 2008; Terenzini & Pascarella, 1998), and many of them are not accessible to transfers.

Major Selection

Many universities require transfer students to declare their major before arriving on campus. However, it is possible that some transfer students are not ready to declare or simply not aware of all possible curriculum options at their new institution. Transfer students often elect to attend community college because they are unsure about what to major in. Some students hope that by attending a local community college, they can save money, complete general education requirements, and/or prerequisites for major courses before transferring to a university. This approach can be risky, given that some academic majors/programs require

students to complete several sequences of specialized coursework that begin during freshman and sophomore years.

Furthermore, researchers determined that race and gender play a large role in the pursuit of career aspirations, specifically pursuit of science, technology, engineering, and mathematics (STEM) degrees by minorities. Since minority students are more likely to attend a community college before attending a four-year institution (Foote, 1997; Jackson, 2013), their numbers are overrepresented in the transfer populations compared to the traditional freshman. Thus, they are much less likely to receive the same supports as first-time freshman. The achievement gap translates into the opportunity gap, not only at a high-school and college level but in the employment and career choice options. Transfer students without proper career counseling early in their academic careers may end up in jobs that pay less over their lifetime and are not aware of other possible career options that may better suit them. According to a 2014 U.S Equal Employment Opportunity Commission Report, Blacks and Latinos are underrepresented in the high-tech private industry, making up only about 8% of the workforce.

Major and Career Exploration Courses for Transfer Students

Even though the majority of transfer students come to four-year institutions with a specific degree or even a career in mind, some students are undecided about their majors. Some students transfer with unrealistic expectations and change their major after the first semester (May & Casazza, 2012). Major exploration and discussion about career and life goals usually do not happen over an hour. It takes an intentional and structured intervention from faculty, career, and academic advisors to help students get on the right path. The issue of admitting and working with undeclared transfer students raises the topic of extended

orientation or transfer seminar courses. Instead of focusing on transitional issues, the primary subject of such courses becomes major-exploration and career development. Students enrolled in such courses use class time to reflect and identify personal abilities and traits that impact career choices and success. They learn about careers that best suit their interests, personality, skills, and values. In addition to self-exploration, students also develop a professional portfolio, practice interviewing, and discuss workplace ethics. Transfer students are more likely to enroll in career development courses than simply courses that focus on transitional issues. After interviewing groups of undeclared students, researchers found that students often felt panic, frustration, disappointment, and discouragement when asked about their major (Reynolds et al., 2012). Reynold's study found that involvement in the undeclared/career readiness courses helped students feel empowered, "lead more intentional lives, and come to terms with expectations about having a declared major" (p. 59). Another reason for colleges to invest resources into such courses is the finding that the earlier students declare their major, the more likely they will be able to graduate on time. Reynolds also found that students enrolled in career/life-calling classes were "just over six times more likely to earn a degree after four years" (p.59) than students who had not taken such a course, and those who did not graduate showed evidence of continuous enrollment.

Part of the reason career intervention courses may be successful is that they help students develop receptiveness to intrusive career interventions. Transition and career readiness courses allow the student to explore career options without being judged or penalized. The level of career decision-making readiness depends on the student's background, work history, identified disabilities, and prior experience with career services. (Sampson et al., 2013). Face-to-face interaction with trained career advisors allows students

to ask questions specific to their majors. However, even the most seasoned career counselors guide students to online resources, such as Focus 2, Occupational Outlook Handbook, Census, and career exploration websites like SIGI and Discover. Although extremely valuable, these resources on their own without counselor interpretation can further confuse students (Tirpak et al., 2013). It is critical that students who take computer-assisted self-assessments like Focus 2 participate in feedback sessions provided either by a trained counselor or in a classroom or office hours setting. Whiston et al. (2003) documented that Computer Assistant Career Guidance Systems are most useful when a counselor or a teacher is involved in students' career exploration process.

Current research supports the notion that transfer orientation, if delivered in a correct format, can be one of the deciding factors if students graduate, drop out, or transfer again. Orientation sessions alone are not enough to provide students with the necessary information. The best format for transfer orientation is a blended classroom, where the focus is primarily on either developmental skills or career readiness, with transitional issues interwoven into the curriculum of the class. This type of curriculum is not a traditional model and is not something many faculty members have done in their classroom. The implementation of this strategy would require academic departments to identify gateway courses and courses explicitly with high enrollment of first-semester transfer students and train faculty on transfer issues.

Transfer Seminars

Davies-Vollum and Greengrove (2010) described the impacts of the skill-focused orientation courses developed to provide transfer and non-traditional students with skills necessary to be successful in upper-division classes, build community, provide extended

orientation and encourage long term planning. Students were asked to complete assignments that taught curriculum expectations, built personal responsibility, helped them set goals, and develop learning portfolios. The course was designed and taught in a traditional classroom format at the University of Washington-Tacoma and was marketed to non-traditional students in the environmental science program. The researchers analyzed student feedback forms collected over four years for the class and found an overwhelmingly positive response to the course content and curriculum.

Transfer issues are wide-ranging, and it would be challenging to address all of them during a one-day orientation. The most successful models of extended transfer orientations have been described as upper-division seminar courses, living-learning communities in residence halls, and online orientation courses. One of the earliest mentions of transfer seminars comes from Richard Stockton College of New Jersey, where Thomas Grites, one of the founders of the National Academic Advising Association (NACADA) and the lead author on the NACADA Transfer Monologues, began developing transfer seminar courses and designing standardized transfer seminar curriculum. Grites and Farina (2012) discussed the results of the study conducted to examine the effectiveness of the Richard Stockton Transfer Seminar curriculum. The study found that students who completed the transfer course felt empowered, better prepared, more organized, less intimidated, and more involved socially. The authors also suggested that as an alternative to developing a completely separate course, institutions can identify transfer "gateway courses" in academic majors and "ask faculty to integrate transitional activities into their syllabi, without disrupting or minimizing any of the content taught in the course" (Grites & Farina, 2012, p. 15).

A large part of the transfer seminar is the process of building community among students enrolled in the class. This important aspect, in addition to having immediate access to an administrator or faculty member who can answer questions, significantly improves transfer experience and improves transfer student retention. In addition to developing comradery among newly enrolled transfer students, small (no more than 20 students) transfer courses can help students increase critical consciousness and develop an awareness of cultural congruence (Cerezo & McWhirter, 2012; Jefferson, Dougherty, Steadman, & Thomas, 2013).

Successful orientations and transfer seminars have to be theory-based and have clear learning objectives. As a result of the orientation, students should be able to make a smooth transition to their new institution. Transfer seminars should provide students with information about departments and services available at the college that will be critical to their future academic success. Students enrolled in the transfer seminars should also be able to conduct a self-assessment of their existing skills and identify any developmental needs during their first term. Transfer seminar instructors play a significant role in building a connection with transfer students, thus developing the student's connection to the university both socially and academically.

Theoretical Framework

Schlossberg's Transition Theory

This study intersects the components of Schlossberg's Transition Theory and Tinto's Theory of Student Departure in an effort to understand the transfer process and the role institutions play in transfer student success. Schlossberg (1989) stated that the transition model' allows researchers and institutional administrators to comprehend students' needs through a structured approach to predicting, measuring, and modifying reactions to change. Schlossberg's theoretical foundation is ideal for understanding how students manage transitions and adjust to college. Using this model, I am able to identify a set of variables that can further be tied to coping and support mechanisms.

According to Schlossberg (1989), four major factors influence students' ability to adjust to a new environment: situation, self, support, and strategies, which are also known as the 4 S's.

Situation:

What prompted the transfer (transfer trigger)? What are the factors that led to the change in the institution? Did the student plan to transfer, or was he or she forced to do so? Time of the transfer is important and in terms of transition: was the event considered "on-time" (i.e., student transferred right after completing an associate degree) or "off-time" (student transferred after a terrible breakup, or job loss). Timing of the transition is tied to the idea of perceived control. Does the student feel that the decision to transfer was a choice? Did they have options of transferring to multiple institutions, or did they have only one choice? If the student was well known, engaged at multiple clubs and organizations at their first institution, would he or she feel and experience the transition as a gain or a loss? If the

student is used to transferring schools at the K-16 level, how effectively would he or she cope with the implications of the current transition? Students' ability to handle stress, previous experience with transferring schools, and the ability to learn about the new environment quickly would all define how the situation surrounding transfer would impact one's adaptation to the new institution.

Self:

Personal and demographic characteristics as well as the psychological attributes all impact one's perception of the situation and ability to adapt to the new environment. Views on life, personal goals, gender, age, ethnicity, race and state of health, as well as outlook, grit, commitment and values all shape one's perception of their new institution and ability to integrate into the campus community.

Social support:

Support systems differ from student-to-student and may include family, friends, coworkers, church congregation members, significant others, teachers, mentors and even pets. The effects of these groups and their perceptions and values surrounding education may play a large role on student's ability to persist towards a degree.

Strategies, coping responses:

Students who are able to be flexible and use the multiplicity of their prior experiences to help navigate the transition are most likely to experience a sense of control and power when it comes to adapting at their new institution. By doing so, these students are able to better manager stress, modify their preexisting behaviors to better suit the new environment, and control the meaning of the narrative of what it means to be a new transfer student.

Table 2.1: Modified and Expanded - The Individual in Transition: A Detailed Look (Schlossberg, 1989)

SITUATION *The Transition*	SELF *The Student*	SUPPORT *Coping Resources*	STRATEGIES *Reactions over time for better or worse*
Type: • *Control* Anticipated Transfer vs Unanticipated Transfer • Difficulty Nonevent vs Chronic hassle *Context* • Relationship of person to transition • Setting in which transition occurs • Duration • Previous experiences with similar transition • Concurrent Stress Assessment *Impact* • Relationships • Routines • Assumptions • Roles	*Personal/Demographic Characteristics* • Socioeconomic Status • Sex role/Gender • Age and stage of life • State of health • Ethnicity/Culture *Psychological resources* • Ego development • Personality • Outlook • Commitments and values • Spirituality and Resiliency *Coping Responses* • Functions: Controlling situation, meaning or stress • Strategies: Information seeking, direct action, inhibition of action, intrapsychic behavior	*Social Support Types:* • Intimate, family unit(s), friendship networks, institutions, communities *Functions:* • affect, affirmation, aid, feedback *Options* • Actual, perceived, utilized, created *Needs:* • Acceptance • Affect & Affirmation • Self-Esteem • Aid and Honest Feedback • Love and Physical Intimacy Personal & Work • Connection Stimulation and Challenge • Role Models & Guidance • Comfort and Assistance	*What strategies is the student using that is impacting the transition?* *How effective are current strategies in helping them cope with transition?* 3 Ways of Coping: • Modifying the situation (hope and optimism) • Controlling the meaning(reframing • Managing stress after transition (selective denial) *Phases of Assimilation* • Pervasiveness • Disruptions • Integration (for better or for worse) *Appraisal* • Of transition, resources, results • Of preoccupations vs life satisfaction

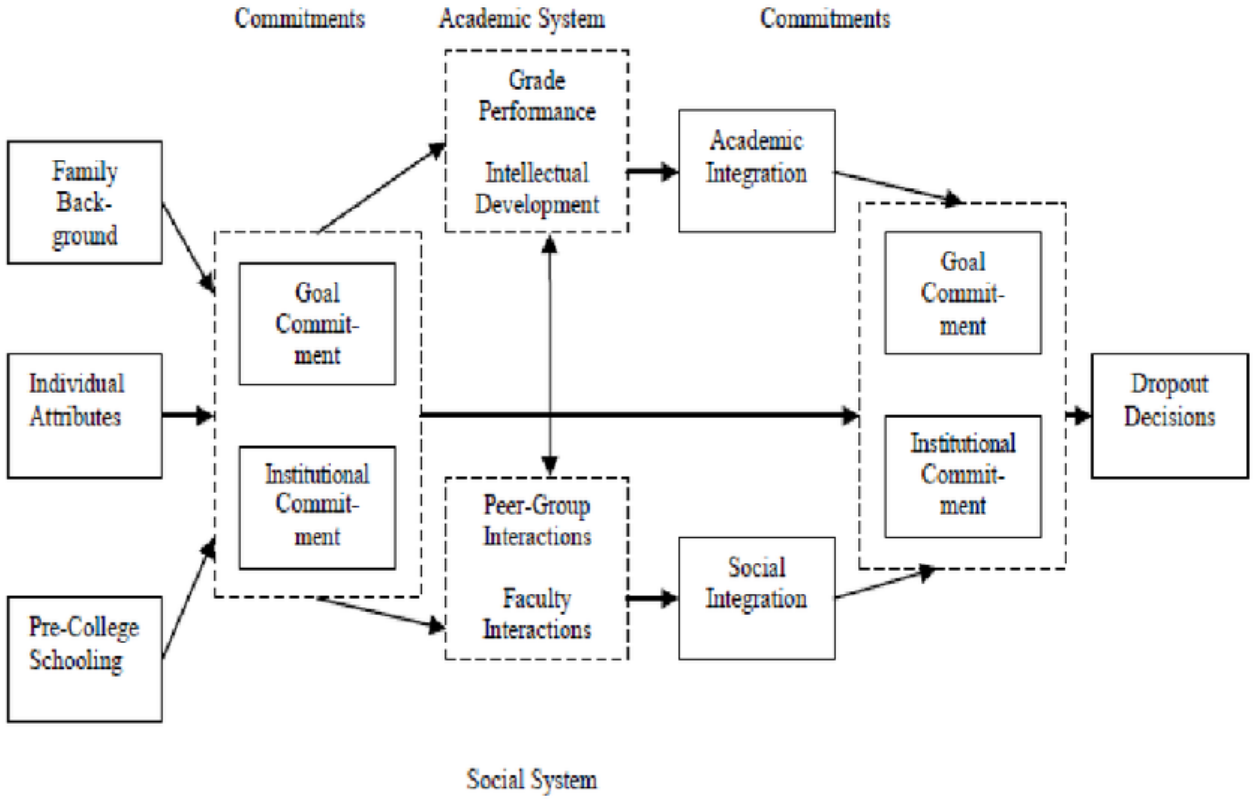
Tinto's Theory of Student Retention

When discussing college completion agenda, most higher education professionals refer to Vincent Tinto's Theory of Student Retention (1975, 1993, 1997, 2006) as a cornerstone theoretical framework required to understand student departure and attrition. Tinto's theories were integral to the publication of an extensive body of literature on retention and understanding of student's academic and social integration processes that occur within the social systems of college and universities. Tinto's (1993) theory of student departure has been one of the most influential and cited in the area of retention. Although originally developed to explain factors affecting freshman retention, it still applies when examining why transfers leave after enrolling at a new institution. Most of Tinto's theoretical foundation is an integrated framework of theories from other disciplines (sociology, psychology, economics, anthropology). Tinto uses a rigorous approach to analyzing different schemas related to student dropout rates. He applies existing theoretical frameworks to establish new areas of inquiry and became a catalyst for the development of an extensive body of research on the topic of college student retention. His publications synthesize and analyze existing theoretical knowledge, drawing from hundreds (if not thousands) of published studies on student retention.

Tinto posits that undergraduate students' continuous enrollment is influenced by their social and academic experiences, as well as their own goals, characteristics, and commitments. Both positive and negative faculty and staff interactions can impact the student's decision to stop attending; these interactions can happen both inside and outside the classroom. Students' interactions with the residence hall, student activities staff, and the campus community at large shape their reactions and feelings about their new school. Thus,

social integration and sense of belonging play a large role in one's perception of their transfer experience and decision to finish the degree.

Figure 2.3: Tinto's Student Integration Model (SIM)



Tinto draws his analysis from the Durkheim's *theory of suicide* and *theory of Cost-Benefits* analysis as they apply to education (noting that prior work on connecting Durkheim's theory to college drop-out was performed by William Spady in 1970).

Two oversimplified examples can demonstrate how these theories relate to students leaving the institution:

1) *Durkheim's Theory:* Student begins his coursework and finds misalignment between his values and interests and those of the department/institution he is in. Over time, the student continues to feel lost and alone and slowly becomes detached from his environment by lack of engagement, attendance, and participation – complete social withdrawal follows, resulting in student's dropping out.
2) *Cost-Benefit's Analysis Theory:* Student spends hours on readings, assignments, and class preparation, only to fail his exams and/or papers. The amount of time spent on coursework does not demonstrate/correlate to understanding as measured by the instructor. Thus, frustration sets in, and the student feels like a failure and withdraws. In this case scenario, either the student did not understand the material, or the instructor(s) failed to provide a framework for what is required to be successful in the class. Instead of continuing the fruitless efforts, the student drops the course or completely drops out from the program.

Tinto emphasizes the importance of recognizing the variations of experiences that students undergo during their college career and the types of events at any stage (freshmen - senior) of enrollment that can lead to student departure. He draws on the research from the field of social anthropology and compares student exits from college to rites of passage in tribal societies. Tinto utilizes the work of Arnold Van Gennep and his research of the rites of

membership in tribal societies to add new perspective and new dimension to his earlier developed theory on student departure. He hopes that by utilizing Gennep's stages of separation, transition, and incorporation in the analysis of student departure, a new understanding of issues of retention will continue to develop.

Tinto acknowledges that retention has become a commodity for researchers, educators, entrepreneurs, and businesses, yet there hasn't been a substantial change in the actual student retention numbers. He argues that the next step is to help translate theory into practice, continue to shape institutional practices and policies, foster and grow faculty development programs and join efforts with other higher education professionals in order to help students graduate. Tinto's approaches to understanding student departure are holistic; his theoretical framework examines both the individual student characteristics and institutional factors that lead to the student's decision to leave the institution. This holistic view translates well to the data modeling approaches used in this study, with much of the data collected and analyzed closely aligning with Tinto's theory of student departure and drawing additional input and output variables from Schlossberg 4S' theoretical framework.

This study used historical data that have accumulated during the normal course of the student's lifecycle. The process of examining the role of theory in the data mining process is an important step and helps provide the framework for defining risk and identifying a list of data sources, variables, and tools needed for successful data reduction, analysis, and prediction. Universities are often data-rich, but insight poor, and there are thousands of variables and datasets that can be used in developing a model. By using variables that can be grounded in theory, we are better equipped to identify previously known and unknown data patterns and build predictive models that can be deployed on the current (incoming) student

data. Even though we have a set of variables collected from standardized institutional datasets, there is added value in creating new variables by combining data from multiple sources, examining variables alone and in combination, and using all five data mining techniques: classification/segmentation, associate rule/dependency modeling, anomaly/outlier detection, clustering analysis and regression.

Types of Risk Models

Risk methodology and predictive modeling have started to gain popularity in higher education and the commercial higher education space around 2010, and much has been written about developing risk assessments for first-year students. Furthermore, the majority of the risk instruments, even those designed for first-year students, have traditionally utilized *logistic regression* to understand the relationships and interactions between student and institutional characteristics. This approach heavily relied on identifying and coding risk factors that can be found in existing retention literature and institutional databases, scoring each student based on their likelihood of dropping out, dividing the population into deciles and reaching out to the highest-risk students. “Regression identifies variables with strong predictive validity across the entire student population and is less effective identifying risk factors that apply to smaller sub-groups within the population” (Jones, 2015, p. 124). This approach is best utilized when the dataset is mostly made of scale variables and has little to no missing data.

Instead of using regression, configural analysis can be utilized to identify primary, secondary, and tertiary predictors. This statistical method does not require data to fit a normal distribution, can use ordinal and nominal data as well as scale variables. Thus, configural analysis is performed by using nonparametric statistical methods to calculate feature

selection, and then using classification procedures to segment a sample into significant groups on the basis of specific theoretical and statistical conditions (Arentze et al., 2000). This technique is called CHAID and is available within the software package Statistical Package for the Social Sciences – Modeler (SPSS – Modeler). Instead of assigning scores from 0-100, and then dividing the sample into deciles, this approach divides the sample into mutually exclusive and exhaustive subgroups, based on specific conditions (Magidson, 1994).

Lastly, the most accurate risk models are those that can capture real-time data and are dynamic in their design. A majority of the commercial offerings in this space have attempted to build software based on this principle. Some of the leaders in the field of predictive analytics in higher education are Civitas, Starfish, and Advisory Board Company (EAB). Each has a slightly different methodology for calculating risk, with some predicting graduation while others focus on persistence. Some models predict outcomes at the university-wide level; others calculate scores by looking at risk in a specific college, department, or a major. By using swipe, learning management system (LMS) data, and faculty provided midsemester alert data, dynamic models can recalculate the risk score at specific intervals and provide more up-to-date risk-scores.

Two main challenges with this approach: not all faculty use learning management systems, and provide mid-semester feedback in a format that is usable or at the time the model is relying on incorporating them in the algorithm (McKenzie, 2018).These additional inputs serve as temporal variables, represent student actions that can impact their risk level and can include both academic data, as well as financial, behavioral, housing, appointment, and on-campus swipe transactions.

This study utilized CHAID/Decision Tree methodology; these techniques have not been applied to studying transfer populations until very recently, and without publication record. Some of the research has been done and presented at conferences in the form of presentations or posters; however, no similar studies have been published at this time.

Summary

Much has been written about the transfer student experience, reasons for transferring, and institutional approaches to working with this population. Most of the literature outlines best practices for new student orientation, registration, and advising, as well as strategies for math and writing remediation skills. Recent developments in higher education and the establishment of Critical Pathway Initiatives shifted the national conversation from the students to the curriculum, and the literature began to reflect this shift. In addition to critical pathways, conversations began about the establishment of a common course numbering system for different states as well as standards for the development of articulation agreements between institutions.

In addition to major selection and course selection, many transfer students need help with campus adjustment, time management, housing, and maintaining a balance between working part-time jobs, getting involved on campus and academic coursework. Transfer specific orientations and transfer seminars have been identified as key strategies for transfer student retention; however, there is limited research on their effectiveness when it comes to the impact on retention/degree completion.

CHAPTER 3
METHODOLOGY

"The possible solutions to a given problem emerge as the leaves of a tree, each node representing a point of deliberation and decision."
~ Niklaus Wirth

The purpose of this study was to identify factors that contribute to transfer attrition/drop-out/graduation, to understand if these factors differ by gender, race and socioeconomic status and by academic discipline, as well as verify if these factors stay consistent from one year to the next. This chapter examines the methodology, procedures for data extraction, data cleaning and blending, feature selection, and measures of precision, accuracy, recall, as well as other important aspects of data analysis and modeling. One of the secondary objectives of this study is the introduction of clustering technique and use of SPSS modeler in the day-to-day operations of university or college advising/retention offices trying to identify and retain high-risk transfer students.

Institutional Characteristics and Student Population

The study was conducted at a large, urban, public research university located within the Mid-Atlantic region of the United States. The study focused on transfer students, who make up one out of three newly admitted students in the fall cohort; on average, close to 2,300 new transfer students are enrolled every fall. The study focused on three cohorts of transfer students; in total, 7,498 student records were analyzed. The large majority of the newly admitted transfer students majored in psychology, accounting, finance, biology, business management, marketing, criminal justice, media studies and production, or computer science. Overall, the students in the sample were enrolled in 135 different programs of study. Many students who transfer come from local community colleges,

designated as partner institutions. Some of the students in the sample enrolled as part of the Dual Admissions Agreements and could have received scholarships in the amount up to $2,000 per year renewed for 3 years. Forty-six percent of the entire transfer class is considered the first generation; this is almost 15% higher than the freshman class at the same institution. For these students, neither parent completed a college degree.

Data Sources and Data Extraction

A majority of the data collected for this study were extracted from Ellucian Banner Enterprise Resource Planning software (ERP), and the home-grown *New Student Questionnaire* (NSQ) administered to all incoming students. Ellucian is a software company that is dedicated to higher education and provides a system of applications that automate higher education back-office functions. Formerly a partnership between Datatel Inc. and SunGard Data Systems, Ellucian is a well-known name in the higher education ERP space. Banner is a software platform that stores admissions, student demographic information, financial aid, HR, accounts receivables, registration, grade, and other data in views/tables. Banner data can be extracted using Oracle SQL Developer or an application like Toad Data Modeler.

A majority of the data in this study are derived from Banner system tables and views; 83 variables were taken from the New Student Questionnaire administered during summer orientation by the Institutional Research Office, and are not available on Banner, but can be downloaded from the institutional Data Warehouse (ODS). The NSQ is developed by the Institutional Research Team and reviewed by a committee of dedicated administrators once a year. The NSQ data come from a survey completed by newly admitted students. The NSQ survey has undergone multiple reiterations and continues to change every academic year.

The primary tools used to assemble the data were SQL Developer, Microsoft Excel, and Tableau. New (secondary/derivative) variables were created in SPSS and Excel. The SPSS files were edited (variable names and formats were changed for compatibility with SPSS coding) and assembled to form a one line per student longitudinal record containing the information needed for research purposes.

Research Design

The first research question was answered by evaluating data from all three cohorts and examining university-wide transfer risk factors by developing multiple predictive models:

A. Combined Model for 2015, 2016, 2017
    a. Pre-Enrollment Run using factors with $p < .05$, effect size $> .099$
    b. End of Term 1 Run using factors with $p < .05$, effect size $> .099$
B. Model examined with race forced as a top predictor, run using factors with $p < .05$
C. Model examined with gender forced as a top predictor, run using factors with $p < .05$
D. Model examined with Socioeconomic Class (using NSQ family self-reported family income question as first top predictor) as a top predictor, run using factors with $p < .05$.

Feature selection for each model was done using χ^2 test and tTest analysis. Effect size was calculated using Cramer V and Cohen's D measure sizes. The cut off value of .099 for the effect size was selected based on past practical experience and in consultation with other researchers. To answer the second question, I split the dataset into three parts and developed three individual models for 2015, 2016, and 2017 transfer student cohorts, and examined the

stability of the predictors from one year to the next. The process for answering the third research question is similar; however, instead of looking at the population as a whole, the model was first split by student's school/college of admission. The last model used graduation as an outcome variable and was only run on the 2015 transfer cohort.

Typically, risk models that use snapshot data should be updated multiple times a year, as the new data collected during the course of the academic year may be more valuable in predicting persistence/graduation than data collected prior to matriculation. Models based on academic data can be ran at specific intervals or on important dates.

The following events define common snapshot intervals:

1) After Placement Exams/Pre-Orientation; partial information exists on the cohort, pre-summer melt

2) During the first semester, at Add/Drop, a point in the semester when students can no longer make changes to their schedule. The cohort is defined at this point in time; students who have deposited and decided not to attend are removed from original models

3) At Mid-Semester Alerts/One week after the withdrawal date has passed (to accommodate for late/excused withdrawals)

4) At the end of the first semester, after all grades have been posted

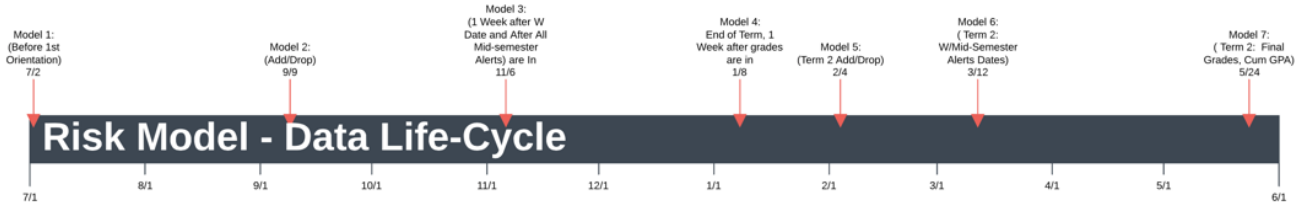
5) During the second semester, at the end of the Add/Drop

6) During the second semester, at Mid-Semester Alerts/One week after withdrawal date has passed (to accommodate for late/excused withdrawals)

7) At the end of the second semester, after all grades have been posted

The more information collected in the model the more accurate the results. This study (Question 1) focused on the event #4 in designing the predictive model.

Figure 3.1: Snapshot based intervals for running retention models on student data.



Tools/Software Packages Used

Phase I: Data Extraction: Cognos IBM, SQL
Phase II: Data Cleaning: Tableau Prep
    Part 1: Development of Cohort Based Longitudinal Datasets
    Part 2: Development of a 3 Year Stacked dataset
Phase III: Feature Selection/Data Reduction: SPSS V23:
    Part 1: Independent T-Test in SPSS Statistics
    Part II: Chi-Square Test of Independence (Crosstabs)
    Part III: Effect Size calculations
Phase IV: Modeling SPSS Modeler (Exhaustive CHAID & Auto Classifier)

Intermediate files were derived from Banner and stored as SPSS Files. New (secondary/derivative) variables were created in SPSS and Excel. The SPSS files were edited (variable names and formats are changed for compatibility with SPSS coding) and assembled to form a long record containing the information needed for research purposes.

The original data files collected included:

1) Specified Cohort File identifying two cohorts of students (Present and Past years)

2) The Longitudinal Academic History file (credit hour, transfer hours, quality points, etc.).

3) Mid Semester Alerts for each student

4) Number of "W" grades for each term, (W grade indicates student withdrew without any academic penalties)

5) Financial Aid Data: Financial Aid Eligibility, Pell Grant amounts and Estimated Family Contributions for each term (*Removed from the study due to financial aid regulations)

6) Semester history file (Special Cohorts, Housing and Living Learning Community, etc.)

Model Variable Selection

One way to view the transfer lifecycle is through the enrollment management lens: the student applies, enrolls, takes classes, and graduates. If something goes wrong in the cycle (e.g., a student takes the wrong class, changes his or her major) the retention algorithm is off. Statistical techniques like Configural Analysis done with the help of IBM's SPSS Modeler CHAID algorithm are used to assist in systematic post-positivist theory development and testing. Before conducting the analysis, it is important to identify the datasets and thus, the variables which will be used in the development of the model. After the availability of the data elements is determined, and the data source systems are identified, data must be extracted. During this process, the researcher must rely on the functional and technical contacts for each database. Functional contacts are tasked with identifying data extraction methods while technical contacts are intended to safely transmit and store the data in accordance with the university's policies and procedures.

Dynamic risk models are the best way to analyze risk and predict a student's likelihood of dropping out. Dynamic risk models would take into account student behavior and incorporate variables from Learning Management Systems, campus swipe-data building on pre-enrollment data of admissions, placement testing, housing selection, orientation and so on. When evaluated using the dynamic risk model, a student's risk-score may change from week to week. Dynamic risk models rely on a large set of cross-institutional temporal data and are currently the gold standard for risk identification. In the last ten years, companies like Civitas, Hobsons, Skyfactor, and software platforms like Veera by Rapid Insight, Teradata, and IBM have all entered into the higher education market of predictive analytics, assessments and benchmarking. Through advanced data analytics, automated data extraction

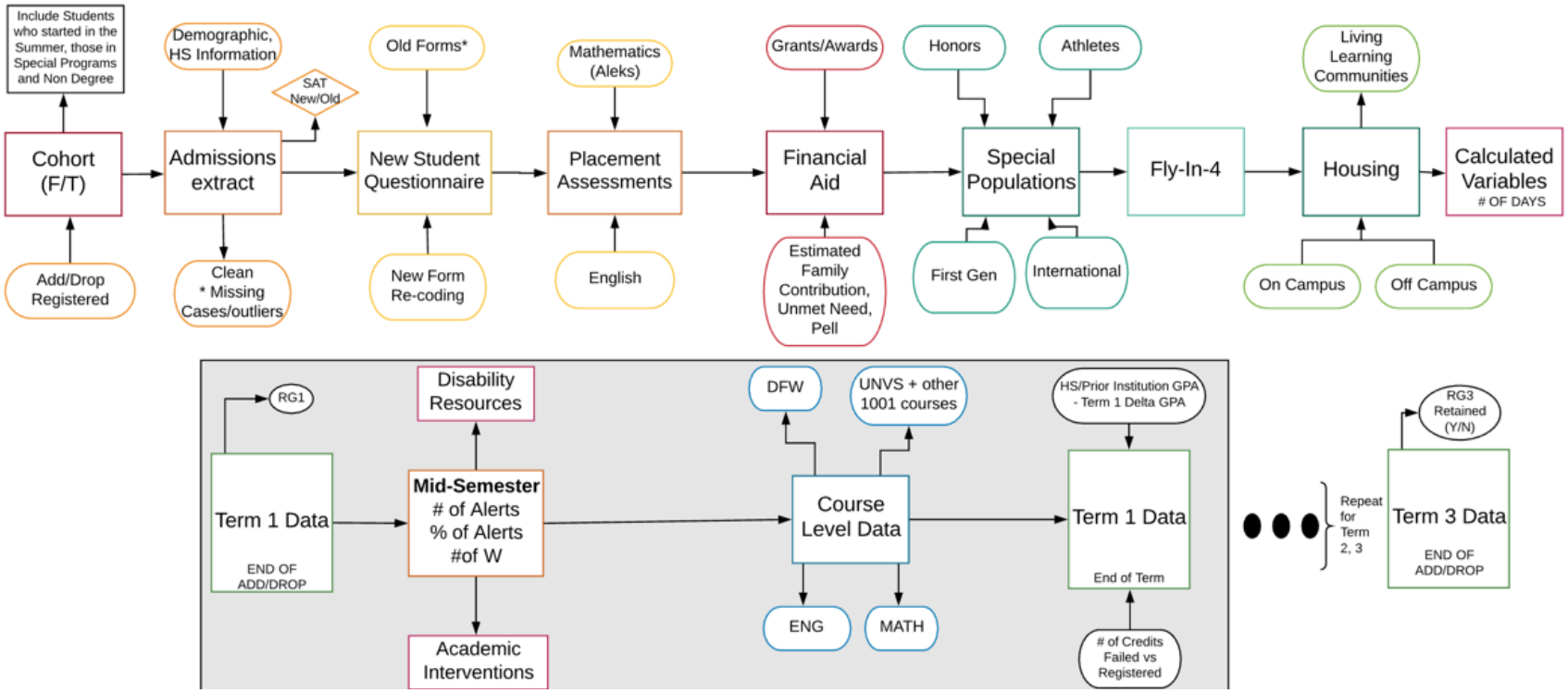
processes, and administrator dashboards, all of these companies are able to track who is at risk, when the student is first identified as at risk, and when he or she is no longer at risk.

The problem with off the shelf models is the applicability of one standard set of rules to all populations. Risk factors at a small Jesuit school in Texas will not be the same as those at a large public university in Florida or small private school in upstate New York. In order to develop the most accurate risk model, administrators need to analyze historical data at their own institution, identify risk factors that apply to their students, and build interventions that will serve their unique student populations. In addition, before developing any model, those working on developing interventions to be used with the students identified "at-risk" need to agree on the definition of at-risk and the model's outcome variable (e.g. at risk of dropping out, at risk of falling below 2.0, at risk of not graduating).

In order to build a model, the institutional research team or the business intelligence team must identify all the source systems (often siloed data used by different academic departments) for required data elements. Then, they must determine the availability of each data element and identify the functional and technical contacts for each source system. Last but not least data architects/engineers must extract and store necessary data elements, while ensuring that all security and data governance procedures and policies are adhered to.
Since the data of this study came from an institution that uses an in-house assessment that is administered to all transfer students prior to orientation called the New Student Questionnaire (NSQ), this assessment was integrated into the longitudinal student dataset. Not all questions on the NSQ are predictive of retention; however, each question was evaluated and used in the development of the final model. Many of the NSQ questions are based on the principles discussed by Tinto and Schlossberg as well as other educational theorists.

Institutional data are without boundaries, with hundreds if not thousands of datapoints available for analysis. These theories influence what variables are selected as inputs and guide the data analysis process. In addition to the NSQ, additional student life-cycle data are collected from the point the student is admitted and all the way through graduation. The diagram below breaks down the data cycle into pre-enrollment, term data, and outcome data. Pre-enrollment data consist of admissions application data, student questionnaire responses, placement assessment results, financial aid, student cohort categories, housing assignments, and additional variables that are calculated (e.g., number of days between student action like application submission, deposit, orientation, registration and the first day of the semester). Term data may include students' college, major, hours registered, completed, number and percentage of courses with midsemester alerts and withdrawals, enrollment and grades in gateway/pre-requisite courses, and grade point average.

Figure 3.2: Student Data Lifecycle



Phase I: Data Extraction

Data used in this study came from the 2015, 2016, 2017 Fall Transfer Cohorts. A cohort is a group of students who applied, was admitted, deposited, and were enrolled in courses the day after add/drop (census point). Students who were admitted for the fall, but started in the summer, took courses as a part of a special program, or as non-degree students and had a prior record of enrollment are included in the cohort. The logic to identify the initial cohort may differ depending on which office runs the data, and the time of the data pull. The data and the logic for identifying transfer cohorts for this study were provided by the Office of Institutional Research (IR). Demographic information (e.g. race, gender) is extracted along with admissions data, along with SAT, CLEP, or any other scores. Since transfer students are not required to submit high school transcripts and SAT scores, for many of them, these data items remain blank. For students who have submitted SAT Scores, it is essential to consider that these scores may not reflect the current SAT grading model. For that reason, and because a significant number of students do not submit the scores, this measure was omitted from the analysis.

After students have been admitted and paid their deposit (or had it waived), they are asked to complete a New Student Questionnaire (NSQ). Students admitted in 2015, 2016, 2017 had to answer 80+ questions that focused on identifying cohort characteristics, attitudes, intentions, and student aspirations. Students must complete the NSQ in order to register for orientation, meet with their academic advisor, and sign up for first semester courses. There is a small group of students who start the questionnaire, but do not finish; this does not prohibit them in moving forward with the enrollment process. However, the majority of the students do finish and answer all the questions asked in the survey.

After completing the NSQ students are asked to complete their placement assessments for English, Mathematics, and Foreign Language. Transfer credits will often waive many students from taking a placement assessment. Students who already completed upper level mathematics, have credits for college level algebra, pre-calculus or calculus may not need to take the proctored Math ALEKS exam. However, many are encouraged to do so, as many may have forgotten the foundational mathematics material. English placement may depend on credits earned at prior institutions, as well as the grade, with many transfer institutions accepting courses with C and some with C- or better. Students in this study did not receive credit for courses with grades below a C. Since some of the transfer students are auto placed and others have to take the assessment, transfer student data have significant gaps in the raw placement scores. Historical placement scores were not available for this study, and although very useful were omitted from the analysis.

The majority of the students in each cohort had completed their Free Application for Federal Student Aid (FAFSA), and I was able to examine the role finances play in retention and time to degree completion. Historically, it has been well documented that students with low expected family contribution (EFC), high unmet need, and those receiving Pell grants struggle financing their degrees.

Estimated Family Contribution (EFC) is a calculated measure that provides colleges and universities with a snapshot of a student's family financial strength and ability to contribute to the cost of education. EFC considers family's income, assets, and size and uses a federal formula to calculate eligibility. Students with low EFCs are eligible to receive federal and state aid, like PELL Grants. The amount of the Pell award differs depending on the EFC calculation. Not all students are eligible to receive Pell, and not all Pell recipients receive the same amount.

Colleges use EFC to determine need based financial aid awards. It is important to note that EFC alone is only part of the total equation when it comes to financing one's college cost; the unmet need is often a critical factor in one's ability to afford college costs. Unmet need is calculated by subtracting EFC and any federal aid/scholarships from the total cost of attendance, which includes tuition, fees, books and supplies, room and board, transportation, and personal expenses. EFC calculations have been criticized for providing an inaccurate portrayal of student's family financial situation, by ignoring differences in regional cost of living, parental debt, family composition, among other factors (Jones, 2011). To meet unmet need students typically borrow from private lenders and/or family members. Alternatively, some try to find employment and pay down the balances by utilizing institutional payment plans. Because of federal regulations and privacy rules regarding financial aid data, all financial variables were removed from the study. Proxy variables from NSQ and term balance were used to replace EFC, Unmet Need, and Pell Amount information. The original design for this study included financial aid information at a student level; however, due to state and federal regulations, the data points were removed from the original dataset. Depending on the office running predictive models, institutions may or may not be able to integrate EFC, Unmet Need, and Pell amounts into the algorithms. Financial aid data are usually difficult to get and it is often changing as students accept different loans, scholarships, and supply institutions with updated family information.

This study does not break down students by subpopulations (i.e., Honors, athletes, international). However, I did examine how first-generation transfer students perform compared to the rest of the population. Researchers use different criteria to define first-generation students; this study uses two categories to identify first-gen students (Hsiao 1992; Padget, Johnson, & Pascarella, 2012)

1) Students for whom neither parent attended college

2) Students for whom neither parent graduated from college.

Alternative definitions include (but not used in this study):

1) neither parents/guardians nor siblings (if any) graduated college

2) neither parents/guardians nor siblings (if any) attended college

Residency is determined by whether the student is in- or out- of- state. Out- of- state students pay higher tuition rates. Out-of-state transfer students often select to attend an institution because of its proximity to a larger city, internships, and opportunities. Other variables in the dataset include calculated variables such as *Registration_Start_DayDIF.* This variable is computed by calculating the number of days between first day of class and the student's first date of registration. Students who register for classes late in the summer often have less course options to select from and may not have a schedule or the professors they may wanted.

All of the earlier collected pre-enrollment data were merged with the registration data from terms 1, 2 and 3. Term data include registration information, such as the number of credits students registered, completed, number of transfer hours accepted by the institution prior to that term and mid-semester alerts submitted by the faculty. Students may receive midsemester alerts for poor attendance, participation, grades, a combination of the above, among other reasons. Lastly, term data include students' major, college, "w" counts (total number of courses the students have official withdrawn from) and account balances at the beginning and end of the semester.

Data Blending:

As described previously, data for this study were collected from multiple institutional sources and combined using Tableau-prep software. Tableau-prep allowed for faster and more

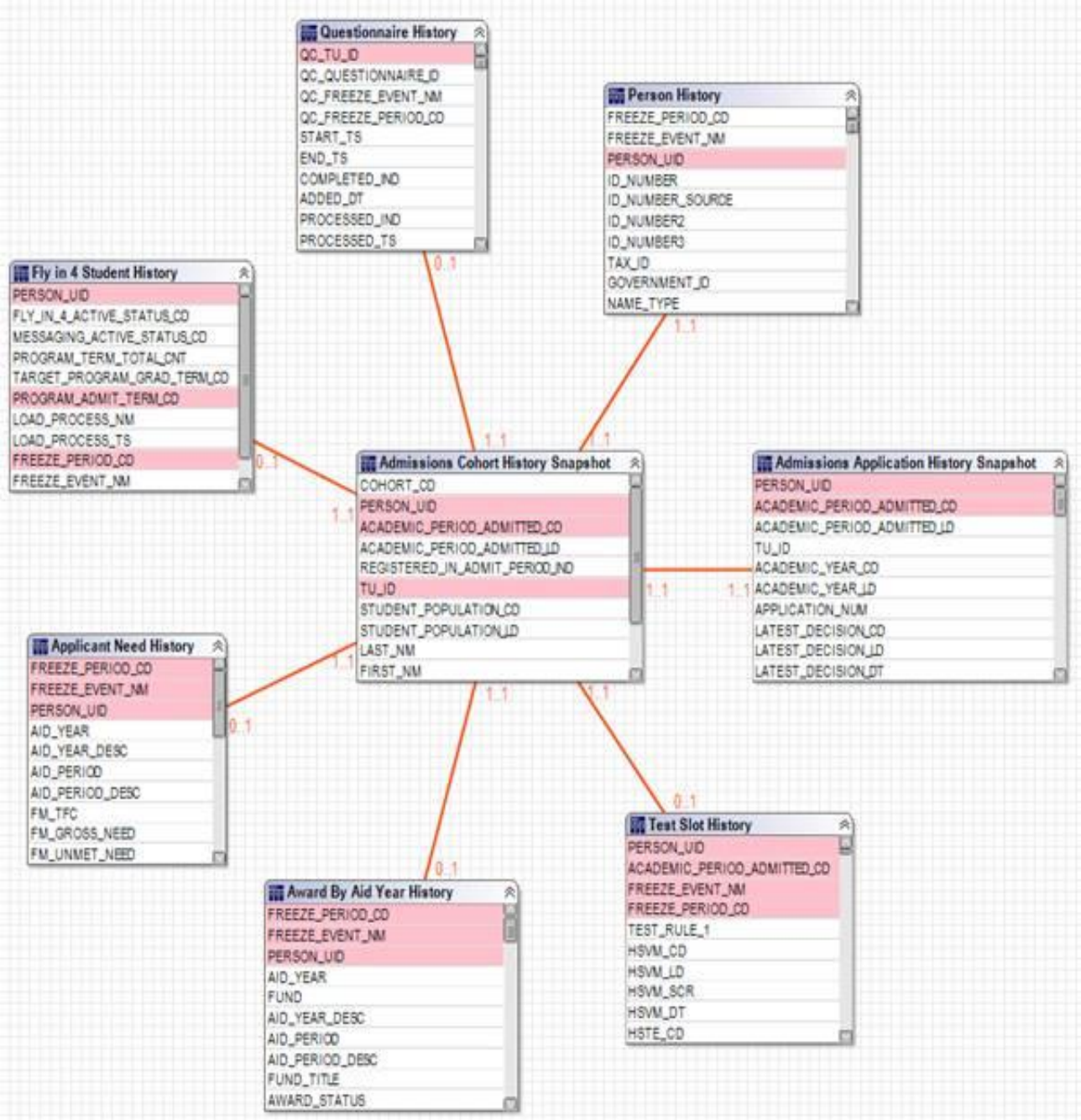
efficient data manipulation and blending. Several left-outer joins were used to merge together pre-enrollment, term, and outcome data. Multiple dataflows were created, one for each cohort, and when combined/stacked together to form multiyear dataset. Data were shaped and cleaned using the Tableau coordinated views. This process allowed for significant reduction in the time that is required to prepare data for analysis. Figure 3.3. provides a student data model and illustrates data relationships for the Cognos package used in this study.

Calculated Outcome Variables

Retained/Not Retained: The existing dataset did not include a binary variable that signified if the student was retained or not. This variable was created by taking the term level variable "Hours Registered," and dividing it by itself, which resulted in a value of one. For students who were not enrolled, the outcome variable was NULL. These two variables were recoded into Yes (1) and No (Null). This process was repeated for term one, term two and term three and named RG1, RG2, RG3 respectively. RG3 - retained/persisted into term three- is used as the main target (dependent) variable in the study.

A similar process was applied for calculating if the student graduated or not. Graduation data were collected from May, August, and December degree conferral dates. The graduation data were only collected on the 2015 fall admitted cohort. Since the cohort enrollment began in September, students who graduated any time between the conferral dates of May 2016 and August of 2018 were coded as graduates. Although a significant number of students had stopped or dropped out, many of the 2015 fall transfer admits were still enrolled in the fall 2018 and subsequent terms, extending their time to degree at the new institution to four or more years.

Figure 3.3: Student Data Model - Relationship Diagram (Source: Business Intelligence)



Missing data:

There were several student records that were incomplete, missing either NSQ data or portions of admissions data. Since this study and the models created used configural analysis

process and not regression, student records with missing data were retained. The Configural analysis CHAID algorithm treats missing values for each independent variable as a single category; for scale and ordinal independent variables that category may or may not subsequently get merged with other categories of that independent variable, depending on the growing criteria.

Phase II: Data Cleaning Process

When data were extracted, all variables with commas were recoded to remove the "," character. Since the majority of the work for the model was done using Tableau and SPSS, commas can cause major difficulties when reading or exporting datafiles by these two programs. The "Replace" command was run to substitute all commas in the file for spaces or blanks. Naming conventions were edited for some of the Cognos variables, with spaces replaced by underscores ("_"); term data were appended with _T1, _T2, T3 to help differentiate variables with the same name that repeated each semester. Blank rows were removed, and data were converted from *.xlsx to *.sav format. There were a number of records for students who were admitted but did not begin in their first semester; those students started their coursework in Term 2 (spring). These records were kept in the original dataset. Data were tested in Tableau by developing multiple pivot tables. This was done to make sure no abnormalities or coding issues happened during NSQ data conversion. NSQ answers were converted in the following fashion: "A" = 1, "B"=2, "C"=3 and so on. Script to convert some of the questions failed, resulting in "Null" values, this resulted in backtracking and cleaning up the SPSS script, and remerging of the datasets later on. Since multiple keys were used to combine the data, and anomalies were found after the dataset was deidentified, the process to rematch variables to the correct student cases took significant additional time.

Phase III: Feature Selection

The raw excel data were loaded into SPSS Statistics Data Editor, and all variables were reviewed to make sure data type, measure and data roles were set correctly. Most of the data points corresponding to the New Student Questionnaire were reset from nominal to ordinal. NSQ questions 4 and 4A were recoded to account for form differences that existed between 2015, 2016, 2017 survey versions.

The first step of developing each model was identifying features/variables that have a statistical significance in predicting the outcome of Persistence/Graduation. In order to identify key features to be included in the model, T-Test and Chi-Square analysis were performed using the SPSS Statistical Package. A Google Documents Spreadsheet was used to track statistical significance (*p-value*) and effect sizes for each association between independent and dependent variables. Only variables with p-values of <.05 were used in the building of the models.

Variables Used in the Analysis and Theoretical Alignment

Most of the variables selected for this study align with Tinto's Student Integration Model and Schlossberg's 4S transition Model and fall within one of the themes outlined in prior research. Variables used in the study are described in detail in Appendix I.

The IBM SPSS Modeler CHAID algorithm was used to determine relationships between variables and make predictions of future events (transfer drop-out). By analyzing transfer data with this method, I was able to uncover relationships between academic and psychometric variables. The decision trees provided highly visual classification through dendrogram visualizations. These visualizations were useful in explaining the statistical significance levels and variable significance hierarchy. The visualizations were also useful in the intervention planning and policy design discussions. SPSS decision trees allowed me to gain insight into a

subset of the data (such as students in specific colleges, transfers with more than 60 credits, or transfers with high unmet need) by conducting further analysis on these sub-groups. This way of looking at transfer student experience provided the big picture of student behaviors. This type of analysis can be extremely useful when proposing or making changes to institutional policies.

Final Dataset

The final deidentified dataset was provided by the institution's Institutional Research office. The study was approved by Undergraduate Studies Office, Registrar's office, and the office of the Provost. Appendix II outlines the process and the steps taken to build the predictive model and conduct the analysis in detail using the SPSS Modeler software.

Interpretation of the Model Performance: Accuracy, Precision, Recall & F1 Score

After the retention and graduation models were developed, they were tested using the confusion matrix to see how good the predictions were in determining persistence. The confusion matrix was developed for questions 1 and 4 using the Tableau software platform and focused on accuracy of prediction at pre-enrollment and at end of term in each model.

Retention Confusion Matrix

	Number of Students Predicted to be at Risk		
Actual number of Students		Prediction = Yes	Prediction = No
	Not Retained	TP = True Positive	FN = False Negative
	Retained	FP = False Positive	TN = True Negative

Graduation Confusion Matrix

	Number of Students Predicted to be at Risk		
Actual number of Students		Prediction = Yes	Prediction = No
	Did not graduate in 3 years	TP = True Positive	FN = False Negative
	Graduated in 3 years	FP = False Positive	TN = True Negative

The goal of accurately predicting retention is to minimize false negatives and false positives. In this case *false negatives* (FN) are defined by the number of students who were not retained yet were not identified as at-risk by the model. *False positives (FP)* cases represent students who were identified as at-risk but were actually retained.

True Positives (TP) represent the number of students who were accurately identified by the prediction model. True negatives (TN) represents the number of students who were accurately predicted by the model as not at risk. After the confusion matrix was developed for each model, Accuracy, Precision, Recall and F1 scores were calculated. The following parameters were used to calculate each measure and determine how good the model performed at identifying at-risk students.

Accuracy score represents a ratio of correctly predicted observations to total observations. In addition to accuracy other measures were calculated due to the asymmetrical nature of the data (values of false positive and false negatives were not the same). An accuracy score of .75 represents a model that is 75% accurate, only if the data are symmetrical.

$$accuracy = \frac{TP + TN}{TP + FN + TN + FP}$$

Precision measures represent the ratio of correctly predicted at-risk observations to the total predicted positive observations. The question that this metric answers is: of all students that are labeled as retained, how many were actually retained? High precision relates to the low false positive rate.

$$precision = \frac{TP}{TP + FP}$$

Recall/Sensitivity measure represents the ratio of correctly predicted retained students to the all observations in the actual class. The question recall measure answers are: Of all the students that are truly retained, how many did we identify as not at-risk? A good recall score is above .05.

$$recall = \frac{TP}{TP + FN}$$

F1 score represents the weighted average of Precision and Recall and is used for datasets with uneven distributions. An F1 score of 1 is considered perfect precision and recall and was the measure used during this dissertation to access the validity of the models.

$$F1 = \frac{2 \times precision \times recall}{precision + recall}$$

This methodology and approach are commonly used in data science and machine learning. (Joshi 2016).

CHAPTER 4
RESULTS

"Without big data, you are blind and deaf and in the middle of a freeway."

In the previous chapter, I reviewed available institutional data sources, the tools and systems used to select, store and extract the data, explanation of the research design, and the process of building and validating a predictive model by using accuracy, precision, recall, and F1 Scores. I explained the data blending process, and went over the variables used in the research, discussed the steps of cleaning the data, addressed the issues of missing data, and developed a SPSS Modeler roadmap to facilitate better understanding of how the model can be replicated using this tool. In this chapter, I will review the descriptive findings, which identity key factors that contribute to transfer student retention/graduation and examine the relationships between these factors and the associated related factor (retention/graduation). The SPSS CHAID Modeler technique was used because it does not require the data to be normally distributed and categorizes null values into a separate category. Recall that this study aims to answer two main questions: a) What factors predict retention and graduation of transfer students? b) Do these factors differ by subpopulations? The answers to these questions can help guide policy and decisions college administrators make about strategic enrollment and spending on services needed to retain and graduate this group of students.

In Chapter 2, I reviewed previously developed models of student retention and literature by Tinto, and Schlossberg. I addressed some of best practices and institutional challenges when it comes to working with transfer students. Considering the predicted college enrollment decline of 15% over the next decade, traditional age transfer students and adult completers will become target demographics for college admissions and enrollment staff. However, if the systems

designed to support traditional freshman do not change to support the transfer population, retention and graduation rates for this group will plummet. Thus, it is important to not only understand the students in this population, but examine the services meant to serve them. In addition, policies surrounding financial aid, continuous enrollment, and credit loads will need to be reviewed given the uniqueness and patterns of the behavior of this group.

Descriptive Results

This study sheds light on the transfer population of one institution and reviews student records for enrollments between 2015, 2016, 2017. A similar study conducted at an online institution, a school or college for adult completers, Historically Black institution or right after 2008 recession would have yielded a different set of predictors and focused on a different type of students. It is important to note that enrollment trends and overall student characteristics remained consistent for the cohorts analyzed. A small decrease in international transfer students was observed in 2017, possibly due to political climate, national immigration rhetoric, travel bans, and more restrictive policies on visas. Given the scope and the complexity of the dataset, it is important to frame the study around traditional-age students, attending an urban institution, during an upward economic climate.

The following tables provide a close examination of the descriptive results necessary to understand enrollment trends, as well as the demographic, racial and socioeconomic composition of the cohorts.

Table 4.1: Total Number of Students in Each Cohort Broken Down by Gender and Persistence

Cohort	N	Did Not Persist (n)	Did Not Persist (%)	Persisted (n)	Persisted (%)
Fall 2015	2454	345	14.06%	2109	85.94%
Female	1196	166	13.88%	1030	86.12%
Male	1257	179	14.24%	1078	85.76%
Not Reported	1			1	100.00%
Fall 2016	2566	383	14.93%	2183	85.07%
Female	1308	193	14.76	1115	85.24%
Male	1257	190	15.12%	1067	84.88%
Not Reported	1	0		1	100.00%
Fall 2017	2478	398	16.06%	2080	83.94%
Female	1288	203	15.76%	1085	84.24%
Male	1188	194	16.33%	994	83.67%
Not Reported	2	1	50%	1	50%
Total	7498	1126	15.02%	6372	84.98%
Female	3792	562	14.82%	3230	85.12%
Male	3702	563	15.21%	3139	84.79%
Not Reported	4	1	25%	3	75.00%

Note: A large number of students in each cohort were originally enrolled in one of the community college partner institutions. Application, deposit and enrollment yields stayed consistent for all three years. No difference in enrollment or persistence was observed between males and females.

Large urban institutions often benefit from robust articulation agreements with partner community college schools. The current dataset does not reflect the student's prior institution of attendance, number of institutions previously attended, or degree attained prior to transfer. These data exist in the National Student Clearing House and could be brought in for further analysis. Nationally, women tend to outnumber men in higher education; this is not the pattern in this student cohort sample.

It is also important to note that the institution that is being studied has long standing relationships with local community colleges. These relationships formed on the existing

program-to-program, dual admissions, and Core-to-Core agreements help bring in a sizable transfer class each fall.

Table 4.2: Total Number of Students Broken Down by Race, Gender & Persistence into Third Semester

Gender	N	Did Not Persist (n)	Did Not Persist (%)	Persisted (n)	Persisted (%)
African American	1176	256	21.77%	920	78.23%
American Indian	6	3	50.00%	3	50%
Asian	699	73	10.44%	626	89.56%
Hispanic	593	112	18.89%	481	81.11%
International (F1 Visa)	457	60	13.13%	397	86.87%
Multiple Races	262	40	15.27%	222	84.73%
Pacific Islander	5	1	20.00%	4	80.00%
Unknown	233	43	18.45%	190	81.55%
White	4067	538	13.23%	3529	86.77%
Total	7498	1126	15.02%	6372	84.98%

Note: Differences can be observed between racial groups. Persistence rates differ between minority and white students. These rates further translate into achievement gaps visible by examining first semester GPA and graduation rates.

Table 4.3: Total Number of Students Broken Down by Race and Self-Reported Family Income

Race & Family Income	N	*% of n*	Did Not Persist (n)	Did Not Persist (%)	Persisted (n)	Persisted (%)
African American	1,176	*100.00%*	256	21.77%	920	78.23%
Did not answer	106	9.01%	30	28.30%	76	71.70%
Less than $20,000	335	28.49%	85	25.37%	250	74.63%
$20,000 to $39,999	271	23.04%	55	20.30%	216	79.70%
$40,000 to $59,999	185	15.73%	38	20.54%	147	79.46%
$60,000 to $79,999	115	9.78%	24	20.87%	91	79.13%
$80,000 to $99,999	68	5.78%	12	17.65%	56	82.35%
$100,000 to $124,999	54	4.59%	8	14.81%	46	85.19%
$125,000 to $149,999	20	1.70%	2	10.00%	18	90.00%
$150,000+	22	1.87%	2	9.09%	20	90.91%

Table 4.3: Continued

Race & Family Income	N	*% of n*	Did Not Persist (n)	Did Not Persist (%)	Persisted (n)	Persisted (%)
American Indian	6	*100.00%*	3	50.00%	3	50.00%
Less than $20,000	2	*33.33%*	2	100.00%		
$20,000 to $39,999	1	*16.67%*			1	100.00%
$40,000 to $59,999	1	*16.67%*			1	100.00%
$60,000 to $79,999	1	*16.67%*	1	100.00%		
$125,000 to $149,999	1	*16.67%*			1	100.00%
Asian	699	*100.00%*	73	10.44%	656	89.56%
Did not answer	49	*7.01%*	5	10.20%	44	89.80%
Less than $20,000	183	*26.18%*	21	11.48%	162	88.52%
$20,000 to $39,999	204	*29.18%*	23	11.27%	181	88.73%
$40,000 to $59,999	84	*12.02%*	10	11.90%	74	88.10%
$60,000 to $79,999	71	*10.16%*	5	7.04%	66	92.96%
$80,000 to $99,999	25	*3.58%*	4	16.00%	21	84.00%
$100,000 to $124,999	39	*5.58%*	2	5.13%	37	94.87%
$125,000 to $149,999	16	*2.29%*			16	100.00%
$150,000+	28	*4.01%*	3	10.71%	25	89.29%
Hispanic	593	*100.00%*	112	18.89%	481	81.11%
Did not answer	36	*6.07%*	11	30.56%	25	69.44%
Less than $20,000	135	*22.77%*	23	17.04%	112	82.96%
$20,000 to $39,999	109	*18.38%*	30	27.52%	79	72.48%
$40,000 to $59,999	81	*13.66%*	17	20.99%	64	79.01%
$60,000 to $79,999	76	*12.82%*	12	15.79%	64	84.21%
$80,000 to $99,999	60	*10.12%*	7	11.67%	53	88.33%
$100,000 to $124,999	40	*6.75%*	6	15.00%	34	85.00%
$125,000 to $149,999	23	*3.88%*	4	17.39%	19	82.61%
$150,000+	33	*5.56%*	2	6.06%	31	93.94%
International (F1 Visa)	457	*100.00%*	60	13.13%	397	86.87%
Did not answer	49	*10.72%*	16	32.65%	33	67.35%
Less than $20,000	34	*7.44%*	4	11.76%	30	88.24%
$20,000 to $39,999	80	*17.51%*	6	7.50%	74	92.50%
$40,000 to $59,999	89	*19.47%*	9	10.11%	80	89.89%
$60,000 to $79,999	58	*12.69%*	8	13.79%	50	86.21%
$80,000 to $99,999	44	*9.63%*	6	13.64%	38	86.36%
$100,000 to $124,999	39	*8.53%*	3	7.69%	36	92.31%
$125,000 to $149,999	19	*4.16%*	1	5.26%	18	94.74%
$150,000+	45	*9.85%*	7	15.56%	38	84.44%

Table 4.3: Continued

Race & Family Income	N	*% of n*	Did Not Persist (n)	Did Not Persist (%)	Persisted (n)	Persisted (%)
Multiple Races	262	*100.00%*	40	15.27%	222	84.73%
Did not answer	13	*4.96%*	2	15.38%	11	84.62%
Less than $20,000	41	*15.65%*	9	21.95%	32	78.05%
$20,000 to $39,999	45	*17.18%*	8	17.78%	37	82.22%
$40,000 to $59,999	36	*13.74%*	6	16.67%	30	83.33%
$60,000 to $79,999	43	*16.41%*	9	20.93%	34	79.07%
$80,000 to $99,999	28	*10.69%*	3	10.71%	25	89.29%
$100,000 to $124,999	22	*8.40%*	2	9.09%	20	90.91%
$125,000 to $149,999	12	*4.58%*			12	100.00%
$150,000+	2	*8.40%*	1	4.55%	21	95.45%
Pacific Islander	5	*100.00%*	1	20.00%	4	80.00%
$20,000 to $39,999	3	*60.00%*			3	100.00%
$60,000 to $79,999	2	*40.00%*	1	50.00%	1	50.00%
Unknown	233	*100.00%*	43	18.45%	190	81.55%
Did not answer	21	*9.01%*	5	23.81%	16	76.19%
Less than $20,000	41	*17.60%*	9	21.95%	32	78.05%
$20,000 to $39,999	41	*17.60%*	12	29.27%	29	70.73%
$40,000 to $59,999	43	*18.45%*	8	18.60%	35	81.40%
$60,000 to $79,999	28	*12.02%*	6	21.43%	22	78.57%
$80,000 to $99,999	20	*8.58%*	1	5.00%	19	95.00%
$100,000 to $124,999	19	*8.15%*	2	10.53%	17	89.47%
$125,000 to $149,999	7	*3.00%*			7	100.00%
$150,000+	13	*5.58%*			13	100.00%
White	4,067	*100.00%*	538	13.23%	3,529	86.77%
Did not answer	251	*6.17%*	44	17.53%	207	82.47%
Less than $20,000	344	*8.46%*	67	19.48%	277	80.52%
$20,000 to $39,999	429	*10.55%*	72	16.78%	357	83.22%
$40,000 to $59,999	550	*13.52%*	74	13.45%	476	86.55%
$60,000 to $79,999	584	*14.36%*	79	13.53%	505	86.47%
$80,000 to $99,999	546	*13.43%*	65	11.90%	481	88.10%
$100,000 to $124,999	556	*13.67%*	53	9.53%	503	90.47%
$125,000 to $149,999	286	*7.03%*	31	10.84%	255	89.16%
$150,000+	521	*12.81%*	53	10.17%	468	89.83%
All Races	7,498	*100.00%*	1,126	15.02%	6,372	84.98%
Did not answer	525	*7.00%*	113	21.52%	412	78.48%
Less than $20,000	1,115	*14.87%*	220	19.73%	895	80.27%
$20,000 to $39,999	1,183	*15.78%*	206	17.41%	977	82.59%
$40,000 to $59,999	1,069	*14.26%*	162	15.15%	907	84.85%
$60,000 to $79,999	978	*13.04%*	145	14.83%	833	85.17%
$80,000 to $99,999	791	*10.55%*	98	12.39%	693	87.61%
$100,000 to $124,999	769	*10.26%*	76	9.88%	693	90.12%
$125,000 to $149,999	384	*5.12%*	38	9.90%	346	90.10%
$150,000+	684	*9.12%*	68	9.94%	616	90.06%

Note: Differences can be observed among racial groups, especially at the lower income levels of families making less than $40,000. Differences in wealth and income distributions are also observed among groups. International students - those who are attending the institution on the F1-Visa - are coded as International, regardless of their race. There is a visible difference in persistence rates for students in the different income bands for this group, with high income international students leaving at a higher rate compared to the comparable rates for non-international students. This may be due to political climate, the timing of the study, or a possibly a student's likelihood to transfer to a higher ranked college or university like the University of Pennsylvania or Columbia University. Students who did not persist may have transferred again to enroll in another institution; thus in an essence "Swirling" between multiple institutions.

Income often serves as a proxy for the neighborhood and the high school the student may have attended in the past. Given that many lower income students may have attended underperforming or schools with significantly lower per-pupil spending, some students may not be prepared to handle college level work. Lower income students are also much more likely to have lower social capital (number of family and friend connections with college degrees, exposure to museums, number of books in the home, etc.). A significant amount of literature shows that minority students are still living in segregated neighborhoods and attend segregated schools, despite the Fair Housing Act and laws abolishing segregation. Defacto segregation-segregation by choice- continues to shape the educational landscape of not only the K-12 school districts, but also the patterns of enrollment for first time and transfer college students.

Table 4.4: Total Number of Students Broken Down by First Generation Status/Number of Hours Students Planned to Work for Money (During the School Year) and Persistence into Third Semester

First Generation & Hours Plan to work	N	% n	Did Not Persist (n)	Did Not Persist (%)	Persisted (n)	Persisted (%)
First Generation: YES	3,432	100.00%	562	16.38%	2,870	83.62%
Did Not Respond	120	3.50%	19	15.83%	101	84.17%
None	327	9.53%	42	12.84%	285	87.16%
1 to 15 hours	894	26.05%	96	10.74%	798	89.26%
16 to 20 hours	938	27.33%	154	16.42%	784	83.58%
21 to 25 hours	557	16.23%	103	18.49%	454	81.51%
More than 25 hours	596	17.37%	148	24.83%	448	75.17%
First Generation: NO	4,066	100.00%	564	13.87%	3,502	86.13%
Did Not Respond	372	9.15%	78	20.97%	294	79.03%
None	580	14.26%	58	10.00%	522	90.00%
1 to 15 hours	1,387	34.11%	139	10.02%	1,248	89.98%
16 to 20 hours	946	23.27%	139	14.69%	807	85.31%
21 to 25 hours	419	10.30%	77	18.38%	342	81.62%
More than 25 hours	362	8.90%	73	20.17%	289	79.83%
Total	7498		1126	15.02%	6372	84.98%

Note: First Generation students in this study are students who indicated that neither parent (mother, father or legal guardians) had graduated with a college degree. First generation students are much more likely to work more hours, as seen in the table above. In addition, first generation students often have family obligations and contribute financially to the household. Some first-generation students may also be immigrants. These students are often coded as citizens; however, they may struggle with English and the American education system.

The definition of first-generation student can be difficult to pin-point, as much nuance exists. The commonly used definition aligns with the earlier example, where neither parent has completed a college degree. Students who fall into this category may still have siblings who are currently enrolled or have completed college. An alternative definition is one where neither parent has attended college. There are also shades of nuance in categorizing students this way. If

parents are divorced, incarcerated, deceased or attended college oversees the definition doesn't account for these differences.

Table 4.5 below breaks down the three cohorts by race and gender and displays first semester GPA for each group. Once again differences can be easily observed among groups. Students not persisting to term three are leaving the university with GPAs lower than those who are staying. The compound effect of poor academic performance, high unmet need, and the need to work creates a situation where students decided that higher education may not be for them. Across the board, students who leave from each racial group perform worse than those who persisted.

There are significant differences in first semester GPA across groups, these differences exist across races and between retained and not retained students. The average GPAs for not retained minority students (African American and Hispanic) were significantly lower than those for white and international students. A GPA below 2.0 also often indicates that the student did not drop out but was dismissed or left on academic warning. This is alarming, as many of these students have accumulated substantial debt to pay for their education and are leaving without a credential or a diploma. Many of them will not be able to transfer to another institution because admissions criteria as most transfer programs require a GPA of 2.0 or better in previous programs.

Table 4.5: Average First Semester GPA Broken Down by Race, Gender, Third Term Persistence

Race & Gender	N	Did Not Persist (n)	Did Not Persist (%	First Semester Mean GPA	Persisted (n)	Persisted (%)	First Semester Mean GPA
Female	3792	562	14.82%	2.08	3230	85.18%	3.06
African American	691	146	21.13%	1.66	545	78.87%	2.79
American Indian	3	1	33.33%	3.83	2	66.67%	2.16
Asian	332	31	9.34%	2.48	301	90.66%	2.91
Hispanic	346	60	17.34%	1.95	286	82.66%	3.02
International (F1 Vis	232	27	11.65%	3.04	205	88.36%	3.23
Multiple Races	153	27	17.65%	2.03	126	82.35%	2.95
Pacific Islander	3	1	33.33%	3.37	2	66.67%	2.13
Unknown	107	15	14.02%	2.05	92	85.98%	3.00
White	1925	254	13.19%	2.22	1671	86.81%	3.17
Male	3702	563	15.21%	2.04	3139	84.79%	2.92
African American	484	110	22.68%	1.73	375	77.32%	2.64
American Indian	3	2	66.67%	0.00	1	33.33.%	4.00
Asian	367	42	11.44%	2.07	325	88.56%	2.76
Hispanic	247	52	21.05%	2.18	195	78.95%	2.82
International (F1 Vis	225	33	14.67%	2.38	192	85.33%	2.94
Multiple Races	109	13	11.93%	2.37	96	88.07%	2.83
Pacific Islander	2	0			2	100.00%	3.53
Unknown	126	28	22.22%	2.02	98	77.78%	2.85
White	2138	283	13.24%	2.08	1855	86.76%	3.02
Not Reported	4	1	25.00%	3.78	3	75.00%	2.93
White	4	1	25.00%	3.78	3	75.00%	2.93
Total	7498	1126	15.02%	2.06	6372	84.98	2.99

Table 4.6: Top 15 Majors with the Highest Number of Students Leaving After First Year (Ranked by Number of students who left). Number of students is often more consequential than percentages, given that they directly correspond to the tuition revenue generated or lost for each department/college.

Major/Program of Admissions	n	Did Not Persist (n)	Did Not Persist(%)	Persisted (n)	Persisted (%)
Business Management	398	62	15.58%	336	84.42%
Biology	327	55	16.82%	272	83.18%
Psychology	420	51	12.14%	369	87.86%
Accounting	378	45	11.90%	333	88.10%
Communication Studies	278	41	14.75%	237	85.25%
Finance	335	39	11.64%	296	88.36%
Computer Science	207	39	18.84%	168	81.16%
Kinesiology	235	37	15.74%	198	84.26%
Criminal Justice	246	36	14.63%	210	85.37%
Early Childhood Education	171	34	19.88%	137	80.12%
Marketing	266	33	12.41%	233	87.59%
Film and Media Arts	162	30	18.52%	132	81.48%
Undeclared-Tyler School of Art	159	29	18.24%	130	81.76%
Undeclared-Business & Management	152	22	14.47%	130	85.53%
Social Work - Undergraduate	149	22	14.77%	127	85.23%
Media Studies and Production	172	22	12.79%	150	87.21%
Undeclared-University Studies	112	19	16.96%	93	83.04%
Undeclared-Liberal Arts	122	18	14.75%	104	85.25%
Sport & Recreation Management	128	18	14.06%	110	85.94%
Mechanical Engineering	130	14	10.77%	116	89.23%
Total	4547	666	14.65%	3881	85.35%

Note: Programs displayed in the table had three-year enrollments of 100 or more students. Although these programs didn't have the highest percentage of students leaving, they had the highest number of students leaving. A total of 136 majors were represented in the sample. There were a number of majors that had enrollments of less than 100 students but lost more than 25% of those students after first year. They included: Pre-Pharmacy, Neuroscience, Biophysics, Biology, Chemistry w/Teaching, among others. It is significant to note that there were majors with enrollments with less than 10 students that lost the majority of the students at the end of the first year (Religion, Earth Science, Jewish Studies, Music Theory). These patterns are of significant concern and indicate that systematic issues may exist inside the departments that prevent certain students from succeeding.

Table 4.7:Top Majors with the Highest % of Students Leaving After First Year

		First-Generation Students				Non-First-Generation Students			
Major/Program of Admissions	n	Did Not Persist (n)	Did Not Persist (%)	Persisted (n)	Persisted (%)	Did Not Persist (n)	Did Not Persist (%)	Persisted (n)	Persisted (%)
Art History	36	3	60.00%	2	40.00%	2	15.38%	11	84.62%
Geology	16	2	50.00%	2	50.00%			4	100.00%
Anthropology	64	4	44.44%	5	55.56%	3	13.04%	20	86.96%
Geography/Urban Studies	36	3	42.86%	4	57.14%	2	18.18%	9	81.82%
Pre-Pharmacy	68	9	42.86%	12	57.14%	2	15.38%	11	84.62%
Biochemistry	136	12	41.38%	17	58.62%	4	10.26%	35	89.74%
Mathematics	66	4	36.36%	7	63.64%	1	4.55%	21	95.45%
Environmental Science	58	5	35.71%	9	64.29%	3	20.00%	12	80.00%
Neuroscience: Cell & Molecular	62	5	35.71%	9	64.29%	6	35.29%	11	64.71%
Chemistry with Teaching	6	1	33.33%	2	66.67%				
Undeclared-Engineering	16	2	33.33%	4	66.67%			2	100.00%
Philosophy	32	3	33.33%	6	66.67%	1	14.29%	6	85.71%
Undeclared-University Studies	224	13	28.89%	32	71.11%	6	8.96%	61	91.04%
Architecture Undeclared	108	8	28.57%	20	71.43%	4	15.38%	22	84.62%
Liberal Studies	70	4	26.67%	11	73.33%	4	20.00%	16	80.00%
Therapeutic Recreation	96	8	26.67%	22	73.33%	2	11.11%	16	88.89%
Undeclared-Tyler School of Art	318	13	25.49%	38	74.51%	16	14.81%	92	85.19%
Natural Sciences	12	1	25.00%	3	75.00%			2	100.00%
Sec Ed-Math Ed	14	1	25.00%	3	75.00%			3	100.00%
Art Education	22	1	25.00%	3	75.00%	1	14.29%	6	85.71%
Film and Media Arts	324	15	23.44%	49	76.56%	15	15.31%	83	84.69%
Chemistry	102	5	22.73%	17	77.27%	1	3.45%	28	96.55%
History	106	4	22.22%	14	77.78%	5	14.29%	30	85.71%
Electrical Engineering	168	10	22.22%	35	77.78%	6	15.38%	33	84.62%
Public Health	174	11	22.00%	39	78.00%	7	18.92%	30	81.08%
Biology	654	33	22.00%	117	78.00%	22	12.43%	155	87.57%
Early Childhood Education	342	22	21.78%	79	78.22%	12	17.14%	58	82.86%
Legal Studies	64	3	21.43%	11	78.57%	5	27.78%	13	72.22%
Sec Ed-Social Studies	70	3	21.43%	11	78.57%	4	19.05%	17	80.95%
Engineering	104	5	20.83%	19	79.17%	2	7.14%	26	92.86%
Criminal Justice	492	29	20.42%	113	79.58%	7	6.73%	97	93.27%
Total	4060	242	25.29%	715	74.71%	143	13.33%	930	86.67%

Notes: Many first-generation students do not persist in Science / Technology / Mathematics Pathways. In addition, students who come in undeclared also show significant declines in persistence. Programs like Pre-Pharmacy, Biochemistry, Mathematics, and Urban studies lose four out of ten students after the first year.

Table 4.8: Total Number of Students Broken Down by Number of Credit Hours Registered/Number of Hours Students Planned to Work for Money (During the School Year) and Persistence into Third Semester

Credit hours registered & Hours Planned to Work	N	% N	Did Not Persist (n)	Did Not Persist (%)	Persisted (n)	Persisted (%)
0-6	396	5.28%	133	33.59%	263	66.41%
Did Not Respond	116	1.55%	36	31.03%	80	68.97%
None	23	0.31%	5	21.74%	18	78.26%
1 to 15 hours	24	0.32%	10	41.67%	14	58.33%
16 to 20 hours	43	0.57%	17	39.53%	26	60.47%
21 to 25 hours	41	0.55%	17	41.46%	24	58.54%
More than 25 hours	149	1.99%	48	32.21%	101	67.79%
7-11	490	6.54%	130	26.53%	360	73.47%
Did Not Respond	36	0.48%	11	30.56%	25	69.44%
None	39	0.52%	10	25.64%	29	74.36%
1 to 15 hours	80	1.07%	20	25.00%	60	75.00%
16 to 20 hours	93	1.24%	28	30.11%	65	69.89%
21 to 25 hours	62	0.83%	18	29.03%	44	70.97%
More than 25 hours	180	2.40%	43	23.89%	137	76.11%
12+	6,612	88.18%	863	13.05%	5,749	86.95%
Did Not Respond	340	4.53%	50	14.71%	290	85.29%
None	845	11.27%	85	10.06%	760	89.94%
1 to 15 hours	2,177	29.03%	205	9.42%	1,972	90.58%
16 to 20 hours	1,748	23.31%	248	14.19%	1,500	85.81%
21 to 25 hours	873	11.64%	145	16.61%	728	83.39%
More than 25 hours	629	8.39%	130	20.67%	499	79.33%
Total	7498	100%	1126	15.02%	6372	84.98%

Note: Students who were registered for less than 12 credits - part time and were working to earn money at the same time had significantly lower persistence rates than students who considered

full time (12+ credits). Students who indicated that they will be working more than 20 hours a week during their first term, regardless of hours registered, had significantly lower persistence rates. Close to 90% of the transfer cohort was registered for 12+ credits during their first semester.

Research Question 1:

What are university-wide risk factors of dropping out for transfer students?

As students apply, get admitted, submit their deposit, take placement exams and submit forms new information is collected about them at each step of the way. As discussed in Chapter 2, risk models can be run at different points of time, and, of course, with only data available at the time of the run. First semester GPA and midsemester alerts are not available at the point of admission and can only be added to the model on subsequent runs. After examining three years of transfer student data, the following factors appear to be significant in predicting persistence for all students. This analysis followed extensive analysis of all variables in the dataset; only significant variables with effect size >.09 were included in the final list of factors that will be used in the modeling process.

Factors Used in The Models

Registration_Start_DayDIF: Number of Days Between Registration and First Day of Class

An independent-samples t-test was conducted to understand the difference in persistence between students who register early in the summer versus closer to the start of classes and the impact of the registration date on persistence. The results suggest, that there was a significant difference for those retained (M=43.91, SD=34.163) and those not retained (M=31.37, SD=31.38) based on their registration date and how close the registration took place to the start of the semester; t (-12.196) = 1632.94, p = 0.00, d = .38. The data were then further examined

by the CHAID model in SPSS to establish cut off points. The CHAID algorithm further subdivided the students into four groups. The first group (students who registered more than 46 days before the start of classes) second group (students who registered 22 to 46 days prior to start of classes), third group (students who registered 21 to 4 days before start of classes), fourth group (students who registered 4 days before the start of classes or during the first week of classes.

Table 4.9: Total Number of Students Broken Down by Registration (Number of Days) and Time to Start of Term

# of Days	Did Not Persist (n)	Did Not Persist (%)	Persisted (n)	Persisted (%)
>46 Days	310	10.02%	2784	89.98%
22-46 Days	187	14.18%	1132	85.82%
21 - 4 Days	404	17.42%	1915	82.58%
<4 Day – Add/Drop	225	29.37%	541	70.63%
Total	1126	15.02%	6372	84.98%

Note: Students registering for classes during August and all the way through add/drop have significantly higher attrition rates. This is most likely due to issues with course availability, missing transcripts and transfer credit re-evaluations.

New Student Questionnaire Question 2:

Number of Hours Students intend to work for money during their first semester

This is a self-reported answer collected in a survey that is required for all incoming students. This answer does not take into account if the job is on or off campus, travel distance to and from the job, and whether or not the job is a work-study position on campus.

A chi-square test of independence was performed to examine the relationship between persistence and the number of hours students intended to work for money while being enrolled. There was a statistically significant association between the two variables.

$\chi^2(4) = 110.526$, $p < .0005$; effect size was small, Cramer's V = .126.

New Student Questionnaire Question 61:

Likelihood of working full time while attending school

Similar to the previous question, NSQ question i61 asks students to indicate the likelihood they will work full time (40+ hours) while being enrolled. I02 (Hours plan to work for money per week) and I61(work full time) questions were merged in 2019 to form one question. A chi-square test of independence was performed to examine the relationship between persistence and likelihood students will work full time. There was a statistically significant association between the two variables. $\chi^2(3) = 75.656$, $p < .005$. The effect size was small (Cohen, 1988), Cramer's V = .104.

New Student Questionnaire Question 6: Self-Reported Parental Income

Students were asked to report parental income during the prior tax calendar year.

This is not an official metric used in the calculation of Estimated Family Contribution used by financial aid award distribution. A chi-square test of independence was performed to examine the relationship between persistence and parental income. There was a statistically significant association between the two variables. $\chi^2(7) = 67.147$, $p < .005$. The effect size was small (Cohen, 1988), Cramer's V = .10.

Race - RACE_CODE:
Race as reported to The Integrated Postsecondary Education Data System (IPEDS)

A chi-square test of independence was performed to examine the relationship between persistence and parental income. There was a statistically significant association between the two variables. $\chi^2(8) = 79.914$, $p < .005$. The effect was small (Cohen, 1988), Cramer's V = .10.

Term_Balance_AD_T1:
Term Balance ($ Owed to the university) at Add/Drop – Census Point

This variable also serves as a proxy for a financial aid variable of Unmet Need, a variable that was unavailable for analyses but is known to be significantly predictive of attrition. An independent-samples t-test was conducted to understand the difference between students who carried balances and those who didn't at the end of add/drop (Census) period and the impact of the balance amount on persistence. The results suggest that there was a significant difference for those retained (M=4549.5, SD=4805.92) and those not retained (M=3609.9, SD=4610.83) based on their account balance, t (6.084) = 1513.667, p = 0.00, effect size = .199. Students who carried a balance were twice as likely to stop out compared to those who didn't.

Minority students had significantly higher first term balances at the end of Add/Drop period. However, *first generation* African American students actually had lower balances (M=$768.79) compared to *non-first generation* African American Students (M=$1,039), the average balance for all African American students was $887.24. This difference may be accounted for by taking into account Pell Eligibility, scholarships for first-gen students, or other institutional aid received. Table 4.12 displays average tuition for different race groups during their first semester, white students carried significantly lower balances at the end of term one after their transfer.

Table 4.10: Average Term 1 Balance at Add/Drop (Census Point) Broken down by Student Race

Race	Average Term 1 (Census) Balance
Pacific Islander	$2,227.70
American Indian	$1,748.67
African American	$887.24
Unknown	$682.95
Hispanic	$507.78
Multi-racial	$452.95
International	$339.89
Asian	$325.04
White	$193.12

Curriculum 1_1_T1:

The major student is listed under during his or her first semester of enrollment

A chi-square test of independence was performed to examine the relationship between persistence and major choice. There was a statistically significant association between the two variables. $\chi^2(144) = 207.792$, $p < .005$. The effect size was small (Cohen, 1988), Cramer's V = .166.

Hours_Registered_T1:

Number of Credits the student is registered for during first semester, students registered for greater than 12 credits are considered to be full-time.

Full-time students pay flat rate tuition, part-time students pay per credit. An independent-samples t-test was conducted to understand the difference between students who were part-time and those who were full time and the impact of credit load on persistence. There was a small group of students who were admitted in the fall and deferred their enrollment until the spring; they did not have a record of registration for Term 1 of attendance. The results suggest that there was a significant difference for those retained (M=14.01, SD=2.652) and those not retained (M=12.68, SD=3.497) based on their credit registered count, t (-12.003) = 1315.451, p = 0.00, effect size = .42.

Count_of_W_Type_Registrations_T1:

Number of courses the student withdrew from during the first semester.

Students may withdraw from courses up to a specific point of time in the semester. Each institution sets a last date to drop a course without academic penalty; students who withdraw from classes get a grade of "W" on their transcript for that course. The "W" course grades are not used to calculate the GPA for the semester or in cumulative GPA. An independent-samples t-test was conducted to understand the difference between students who withdrew from courses and those who maintained original credit loads and the impact of withdrawal on persistence. The results suggest that there was a significant difference for those retained (M=.13, SD=.416) and those not retained (M=.43, SD=.944) based on their course withdrawal count, t (10.348) = 1164.374, p = 0.00, effect size = .411. There is a direct correlation between the number of the courses a student withdrew from and their likelihood to persist into the next year. The higher the number of 'W's the higher the attrition rate. Students withdrawing from major pre-requisites or courses that are only offered once a year may significantly delay their graduation dates. Withdrawals may also impact financial aid and the amount of money a student may need to pay out of packet for his or her tuition.

Table 4.11: Number of Students who withdrew from 1 or More Courses and Their Persistence

# of 'W's at the end of Term 1	Did Not Persist (n)	Did Not Persist (%)	Persisted (n)	Persisted (%)
0	845	12.95%	5682	87.05%
1	190	23.81%	608	76.19%
2	44	44.00%	56	56.00%
3	12	44.44%	15	55.56%
4	22	78.57%	6	21.43%
5	11	73.33%	4	26.67%
6	2	66.67%	1	33.33%

Number of Mid-Semester Alerts

Number of midsemester alerts the student received during first semester

Faculty are required to issue a mid-semester alert for students who are struggling academically in their courses. The alerts may indicate that the student is missing classes, does not participate in class discussions, missing assignments, or is struggling to comprehend the material. Not all faculty members issue mid-semester alerts; some do not structure their course to provide accurate assessment at the mid-semester alert point of the semester; and others grade on improvement or heavily weight final projects. If a faculty member fails to provide a mid-semester alert and fails the student, both the predictive model and the intervention may fail. An independent-samples t-test was conducted to understand the difference between students who received alerts and those who did not and the impact of alerts on persistence. The results suggest, that there was a significant difference for those retained (M=1.44, SD=.7) and those not retained (M=1.78, SD=.946) based on their faculty alert count; t (8.667) = 827.218, p = 0.00, effect size = .408. There is a strong relationship between the number of alerts students receive and their likelihood to persist into the next year. The higher the number of alerts, the higher the attrition rate.

Table 4.12: Number of Students who Received a Faculty Mid-Semester Alert from 1 or More Courses and Their Persistence

# Alerts	Did Not Persist (n)	Did Not Persist (%)	Persisted (n)	Persisted (%)
0	480	11.34%	3752	88.66%
1	313	15.26%	1738	84.74%
2	213	24.12%	670	75.88%
3	80	32.13%	169	67.87%
4	29	44.62%	36	55.38%
5	11	61.11%	7	38.89%
Total	1126	15.02%	6372	84.98%

Term_Undergrad_GPA_T1:

First semester term GPA

First semester GPA for transfer students is indicative of their academic performance at the new institution. This dataset does not contain students' transfer GPA. The GPA drop ("Transfer Shock" or also called "Delta GPA" is calculated by subtracting first-semester GPA the new institution from the prior institution GPA and examining how students with different drops in GPA perform over time.

An independent-samples t-test was conducted to understand the difference between students who struggled academically and those who did well in their first term. The results suggest, that there was a significant difference for those retained (M=2.99, SD=.76) and those not retained (M=2.06, SD=1.277) based on their first semester GPA, t (-23.325) = 1228.746, p = 0.00, d = .885.

Table 4.13: GPA Bands and Levels of Risk Determined by Persistence into Term 3, bands were determined using SPSS CHAID Technique

GPA Band	Risk Level	Did Not Persist (n)	Did Not Persist (%)	Persisted (n)	Persisted (%)
0.00 – 1.620	1	403	48.15%	434	51.85%
1.621 – 2.24	2	145	20.17%	574	79.83%
2.241 – 2.84	3	211	13.94%	1303	86.06%
2.841 – 3.24	4	148	10.29%	1290	89.71%
>3.24	5	219	7.32%	2771	92.68%
Total		1126	15.02%	6372	84.98%

Term_Undergrad_Hours_Completed_T1:

Number of credits student successfully completed during his/her first semester.

This measure does not indicate the number of credits that will be used towards the degree; simply that the student passed the course and had earned credit. Credits completed with C- or lower are added into this calculation; however, they may not be usable to satisfy the degree requirements.

An independent-samples t-test was conducted to understand the difference between students who were registered full time vs those who had lighter course schedules. The results suggest, that there was a significant difference for those retained (M=13.55, SD=2.93) and those not retained (M=10.9, SD=4.733) based on their hours completed during first semester, t (17.878) = 1238.428, p = 0.00, d = .673.

Table 4.14: Credits Completed and Levels of Risk Determined by Persistence into Term 3, bands were determined using SPSS CHAID Technique

Credits Completed	Risk Level	Did Not Persist (n)	Did Not Persist (%)	Persisted (n)	Persisted (%)
0 – 8	1	315	39.67%	479	60.33%
9 – 12	2	307	18.13%	1386	81.87%
12 – 14	3	238	13.09%	1580	86.91%
>14	4	266	8.33%	2927	91.67%
Total		1126	15.02%	6372	84.98%

Term_Balance_ET_T1:

Term Balance at the end of the first semester. Students with balances of greater than $200 trigger a hold on the student's account and in turn prevent future registrations.

An independent-samples t-test was conducted to understand the difference among students with different balances. The results suggest, that there was a significant difference for those retained (M=201.74, SD=1264.22) and those not retained (M=1375.99, SD=2993.74) based on their end of term account balance, $t(12.76) = 1158.08$, $p = 0.00$, $d = .51$.

Prediction Model Pre-Enrollment

Once the top predictors were identified ($p < .05$, effect size >.09) they were used to develop two predictive models for pre-enrollment and one at the end of the semester. The dataset was subdivided into smaller samples and significant groupings. The set of rules that classified students as high or low risk were presented using SQL Code. The model begins with an overall fall retention rate of 84.98%. For the 766 students who registered less than four days before classes started or during add/drop [NODE 1], the retention rate decreases to 70.6%; for the 2319 students who register in August [Node 2] or 21 days before the start of the semester the retention rate was 82%. The model subsequently searches within each of the two primary level subgroups for the next best predictor. For those students (n=283) who registered during add/drop and reported family income of less than $40,000 the retention rate is 67%. The group is further

subdivided into several sub-groups with students in different majors. At this point stopping rules prevented further separation. The configural model identified a number of risk groups based on different groups of variables. A total of 778 students were identified as at risk.

Risk Groups & Rules

Risk Group 1: [Node 7 – Chance of Dropping out – 50%; n = 78
Risk Group 2: [Node 11 – Chance of Dropping out – 46%; n = 92
Risk Group 3: [Node 24 – Chance of Dropping out – 47%; n =173
Risk Group 4: [Node 16 – Chance of Dropping out – 37%; n = 91
Risk Group 5: [Node 25 – Chance of Dropping out – 36%; n = 231
Risk Group 6: [Node 44 – Chance of Dropping out – 36%; n = 113

For the highest risk group (1) the model generated the following set of rules:

if Registration_Start_DayDIF (Number of Days Before First Day of Class) <= 4
and i06 (Self-Reported Parental Income) IS MISSING (Question was left blank)
then at Risk Node 7

For the risk group (2) the model generated the following set of rules:

if Registration_Start_DayDIF > 4
and Registration_Start_DayDIF <= 21
and ADMISSIONS_MAJOR_LD in ["Neuroscience:Cell & Molecular" "Public Health" "Architecture" "Geology" "Theory" "Health Information Management" "Chemistry with Teaching" "Exercise & Sport Science" "Biophysics" "Undeclared-Engineering" "Music Therapy" "Linguistics" "Career and Technical Education" "Geography/Urban Studies" "Composition-Music" "Environmental Enginrng" "Undeclared-Science and Technol" "Philosophy" "Pre-Pharmacy"]
then At Risk Node 11

For the risk group (3) the model generated the following set of rules:

if Registration_Start_DayDIF <= 4
and i06 <= 2 (Family Income <$40,000)

and ADMISSIONS_MAJOR_LD in ["Adult & Organizational Develop" "Public Health" "Film and Media Arts" "Biochemistry" "Biology" "Mathematics" "Legal Studies" "Undeclared-Media and Comm" "Computer Science" "Mechanical Engineering" "Early Childhood Education" "Electrical Engineering" "Information Science & Technolo" "Pre-Pharmacy" "Sport & Recreation Management" "Economics" "Neuroscience:Cell & Molecular" "Communication Studies" "Management Information Systems" "Finance" "Architecture Undeclared" "Middle Grades Education" "Undeclared-Business & Mngt" "Undeclared-University Studies" "Business Management" "Undeclared-Liberal Arts" "Chemistry" "Sec Ed-Social Studies" "Criminal Justice" "Engineering Technology" "Philosophy" "Jewish Studies" "Kinesiology"]

then at Risk Node 24

For the risk group (4) the model generated the following set of rules:

if Registration_Start_DayDIF > 21

and Registration_Start_DayDIF <= 46.

and ADMISSIONS_MAJOR_LD in [“Neuroscience:Cell & Molecular” “Horticulture” “Global Studies” “Entrprnrship & Innovation Mgt” “Architecture Undeclared” “Environmental Science” “Physics” “Legal Studies” “Political Science” “Pre-Health Info Mgt” “Art History” “Sec Ed-World Langs Ed” “Dance” “Asian Studies” “Engineering Technology”]

then at Risk Node 16

For the risk group (5) the model generated the following set of rules:

if Registration_Start_DayDIF <= 4

and i06 > 2 (Family income > 40%)

and ADMISSIONS_MAJOR_LD in ["Spch Lang Hearing Science" "Adult & Organizational Develop" "Human Develop Commnty Engagemn" "Film and Media Arts" "Accounting" "Global Studies"

"Psychology" "Anthropology" "Biochemistry" "Mathematics" "Environmental Science" "History" "Marketing" "Therapeutic Recreation" "Community Development" "Sec Ed-English Ed" "Sociology" "Mechanical Engineering" "Early Childhood Education" "Media Studies and Production" "Neuroscience:Cell & Molecular" "Communication Studies" "Liberal Studies" "Finance" "Art Education" "Civil Engineering" "Human Resource Management" "Undeclared-University Studies" "Biology with Teaching" "Physics" "Business Management" "Political Science" "Undeclared-Liberal Arts" "Chemistry" "Supply Chain Management" "Performance" "Sec Ed-Social Studies"]

then At-Risk Node 25

For the risk group (6) the model generated the following set of rules:

if Registration_Start_DayDIF > 46

and i02 > 3 or i02 IS MISSING

and ADMISSIONS_MAJOR_LD in ["Bioengineering" "Spch Lang Hearing Science" "Tourism and Hospitality Mgmt" "Neuroscience:Cell & Molecular" "Liberal Studies" "Art Education" "Religion" "Biochemistry" "Environmental Science" "Construction Mgt Tech" "Landscape Architecture" "Legal Studies" "Pre-Health Info Mgt" "Sec Ed-Math Ed" "French" "Therapeutic Recreation" "Geography/Urban Studies" "Undeclared-Science and Technol" "Performance" "Early Childhood Education" "Electrical Engineering" "Pre-Pharmacy"]

then At-Risk Node 44

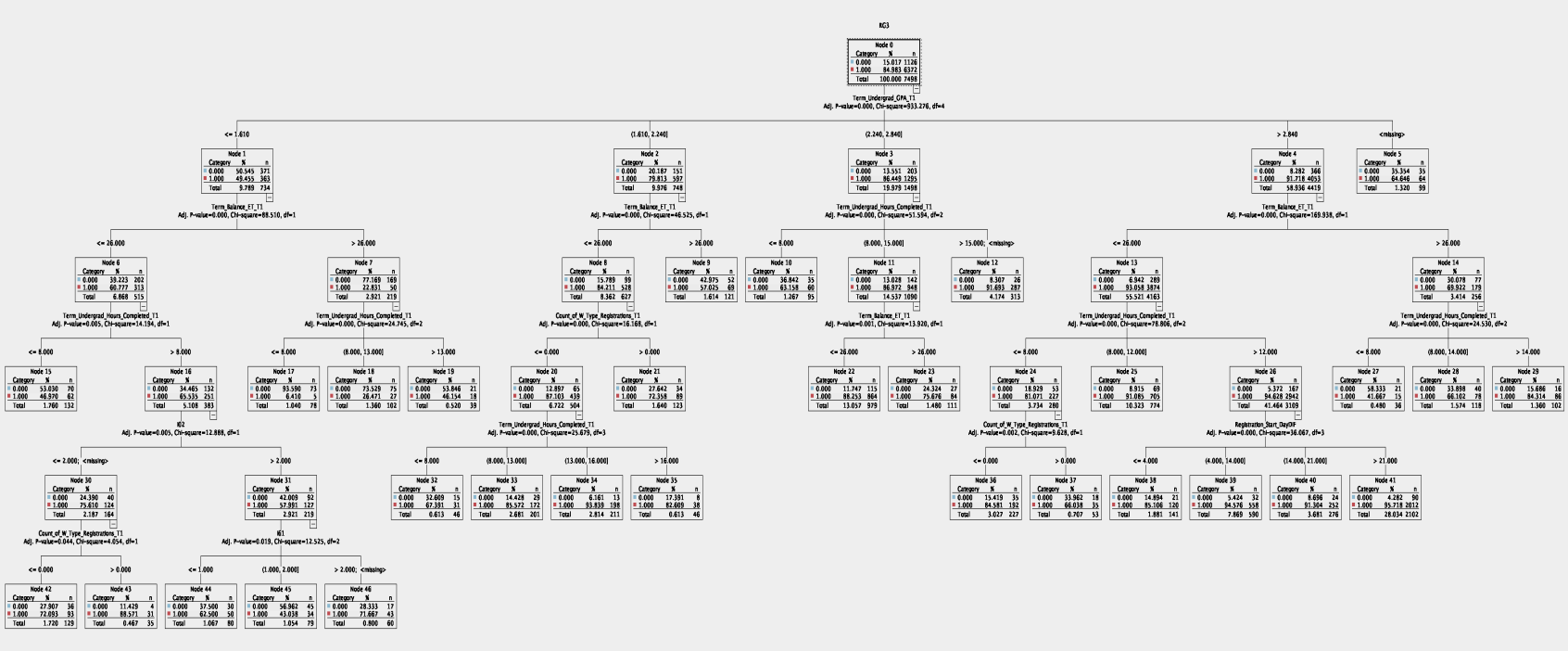
	Number of Students Predicted to be at Risk		
Actual number of Students		Prediction = Yes	Prediction = No
	Not Retained	TP = True Positive 319	FN = False Negative 807
	Retained	FP = False Positive 459	TN = True Negative 5913

Accuracy: (319+5913)/(319+807+5913+459)= .83
Precision: (319)/(319+459) = .41
Recall: (319)/(319+807) = .28
F1 = (2*.41*.28)/(.41+.28) = .33

This model is only slightly effective in predicting at-risk students; there is high misclassification of students who are identified as False Negative (Were not predicted to leave yet did not persist). One way to combat this is to decrease the at-risk pool only to groups 1 & 2 and focus on those students only. Variables like Unmet Need, EFC and Placement Scores can significantly enhance the pre-enrollment model. Those variables were not available for this study. Pre-enrollment data are also filled with Null and Missing values, as a large portion of transfer students completes enrollment requirements at the last minute.

Prediction Model Term 1 (Post Final Grades)

Figure 4.1: SPSS CHAID Decision Tree - Prediction Model Term 1 (Post Final Grades)



The second model was developed using variables that represent datapoints from the past census. The model generated 22 sub-groups with different levels of Risk. High risk subgroups were represented by Nodes 17, 18, 15, 9, 29.

Risk Group 1: [Chance of Dropping out – 92% Chance of Dropping Out; n = 78]
if Term_Undergrad_GPA_T1 <= 1.610
and Term_Balance_ET_T1 > 26
and Term_Undergrad_Hours_Completed_T1 <= 8
then At-Risk Node 17

Risk Group 2: [Chance of Dropping out – 68% Chance of Dropping Out; n = 141
if Term_Undergrad_GPA_T1 <= 1.610
and Term_Balance_ET_T1 > 26
and Term_Undergrad_Hours_Completed_T1 > 8
then At-Risk Node 18

Risk Group 3: [Chance of Dropping out – 53% Chance of Dropping Out; n = 132
if Term_Undergrad_GPA_T1 <= 1.610
and Term_Balance_ET_T1 <= 26
and Term_Undergrad_Hours_Completed_T1 <= 8
then At-Risk Node 15

Risk Group 4: [Chance of Dropping out – 43% Chance of Dropping Out; n = 121
if Term_Undergrad_GPA_T1 > 1.610
and Term_Undergrad_GPA_T1 <= 2.240
and Term_Balance_ET_T1 > 26
then At-Risk Node 9

Risk Group 5: [Chance of Dropping out – 42% Chance of Dropping Out; n = 219
if Term_Undergrad_GPA_T1 <= 1.610
and Term_Balance_ET_T1 <= 26
and Term_Undergrad_Hours_Completed_T1 > 8
and i02 > 2 (Hours working > 15)
then At-Risk Node Node 29

The following confusion matrix measures the accuracy of the predictive model:

	Number of Students Predicted to be at Risk		
Actual number of Students		Prediction = Yes	Prediction = No
	Not Retained	TP = True Positive 383	FN = False Negative 743
	Retained	FP = False Positive 308	TN = True Negative 6064

Accuracy: (383+6064)/(383+743+6064+308) = 85.9
Precision: 383/(383+308) = .55
Recall: 383/(383+743) = .34
F1: (2*.55*.34)/(.55+.34) = .42

The model has a high level of accuracy (86%); moderate precision rate (55%). The recall value is below .5 and suggests that the model is valid; however, it has significant misclassification issues exist, specifically in nodes 9 and 29.

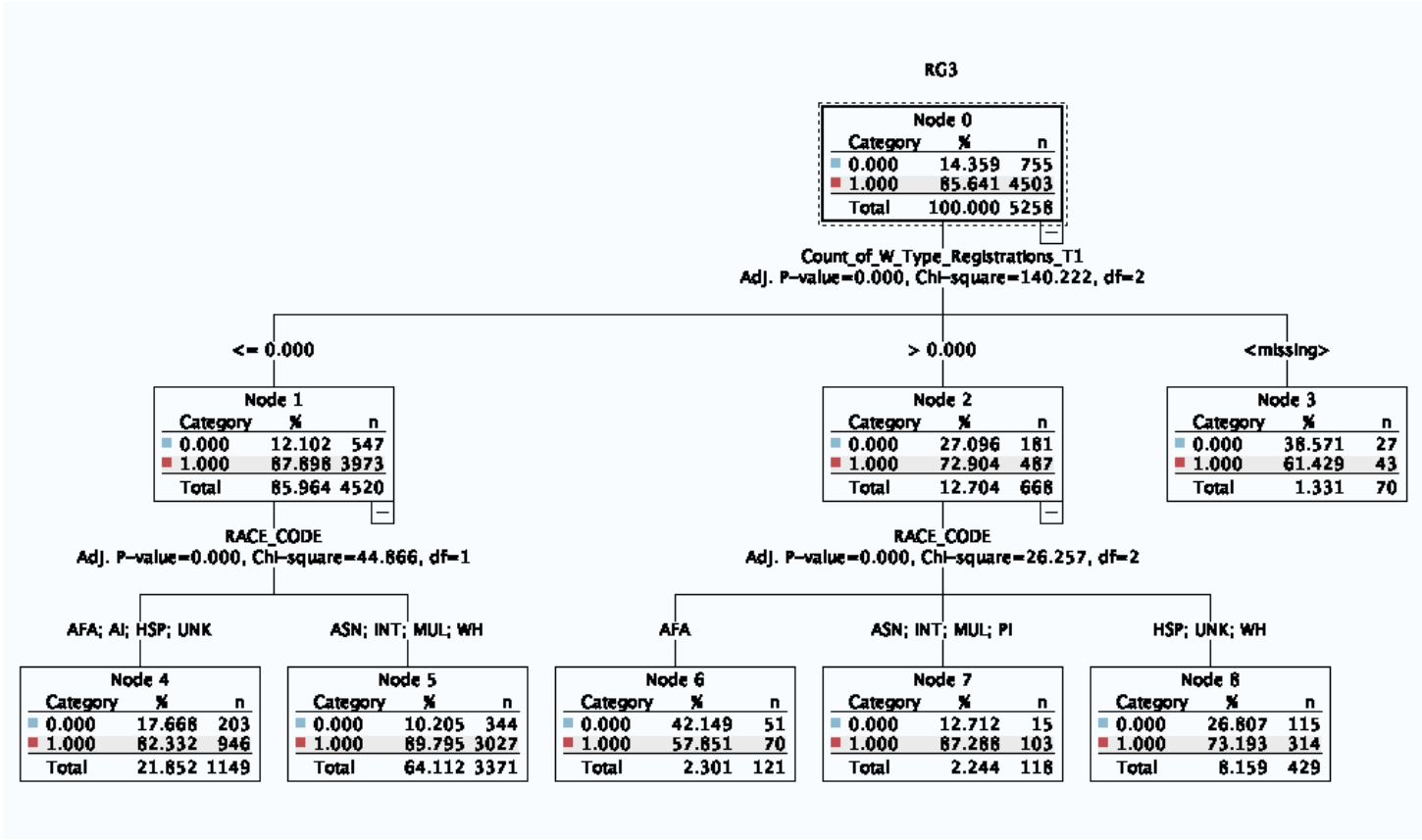
Research question 1b. Differences by race, socioeconomic class, or gender:

Differences by Race:

University-wide model picked up on the factors that were significant to the population as a whole, when examining the data and forcing in race, socioeconomic class or gender additional findings become evident. When I ran the models and forced in race as a top predictor, the results were very similar to the model described earlier. A couple of important distinctions did appear in the nodes:

1) Financial factors were much more prevalent for African American Students. Combined with poor academic performance, students who struggled financially during the first semester were much more likely to stop-out compared to their peers.
2) African American students who withdrew from courses were much more likely to stop out compared to all other students (42% of those who withdrew from 1 or more courses did not persist, compared to 12% for Asian, Multiracial, Pacific Islander, International Students, and 26% for Hispanic and White Students)

Figure 4.2: CHAID Algorithm Produced Model That Examines The Relationships Between Race, Count of Withdrawals, and Persistence into Third Semester.



3) There were also differences in the racial composition of each major. When examining majors with enrollments greater than 100 students, African Americans were much more likely to enroll in programs like Social Work (32% of the students in the major), Criminal Justice (25% of the students in the major), and Psychology (22% of the students in the major). They were least likely to enroll in majors like Accounting (7% of the students in the major), Finance (8% of the admitted students in the major) and Computer Science (8% of the admitted students in the major). Asian students were much more likely to enroll in majors like Accounting (27% of the admitted students in the major), Biology (18% of the students in the major), and Kinesiology (16% of the admitted students in the major). White students had similar patterns as Asian students. The self-segregation was obvious and happened by major.

Differences by Socioeconomic Class:

1) In predicting persistence, financial factors were much more prevalent for students from lower socioeconomic backgrounds (Family income less than $40,000). These students expressed major concerns (NSQ Question 10) about financing their education.
2) Lower income students were also more likely to work 20+ hours a week, which in turn led to lower persistence rates. Figure 4.3 further emphasizes that students who come from lower income families are much more likely to work more hours.

Figure 4.3: Heat Map that Breaks Down Students by Income and the Number of Hours They Intend to Work for Money

	Hours plan to work for money						
Estimate of Parental Income	Did Not Answer	None	1-15	16-20	21-25	>25	Grand Total
<$20,000	40	107	256	290	186	236	1,115
$20,000-$39,000	50	121	327	328	182	175	1,183
$40,000-$59,000	35	105	293	299	153	184	1,069
$60,000 - $79,000	34	110	319	281	127	107	978
$80,000 - $99,999	28	95	275	205	107	81	791
$100,000-$124,999	28	117	291	180	90	63	769
$125,000 - $149,999	18	52	157	98	41	18	384
$150,000+	15	151	288	135	56	39	684
Did Not Answer	244	49	75	68	34	55	525
Grand Total	492	907	2,281	1,884	976	958	7,498

Figure 4.4: First Semester GPA Broken Down by College of Enrollment and Estimated Parental Income.

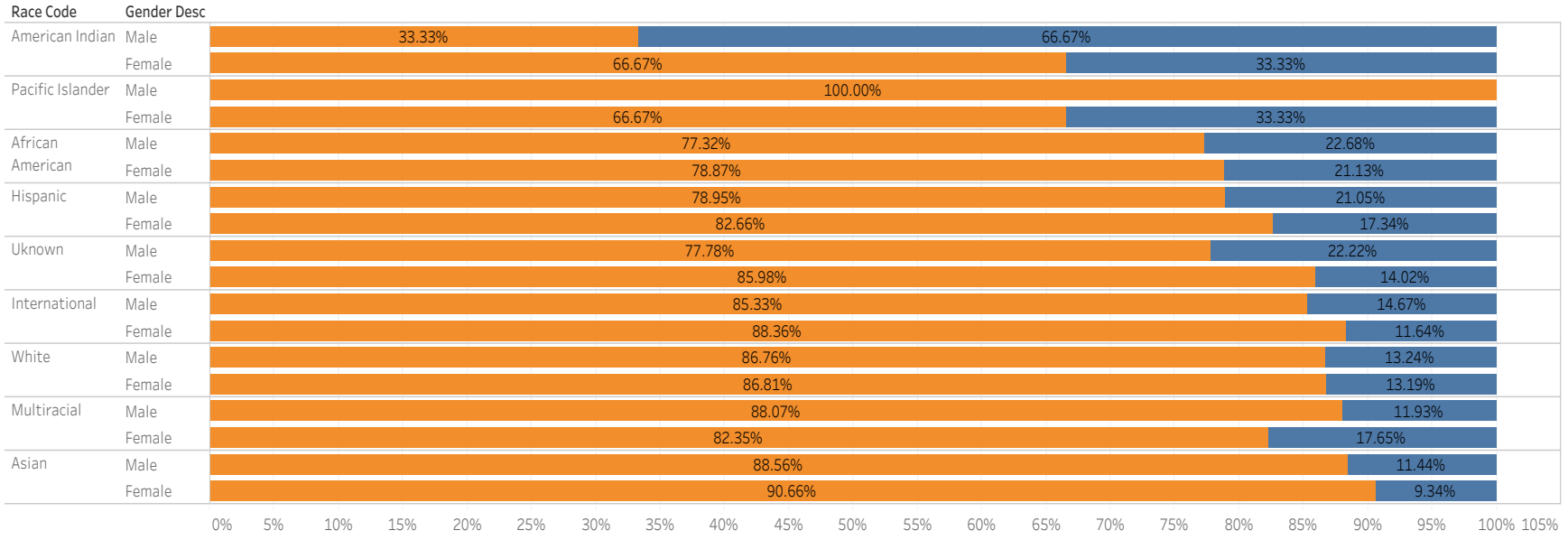
	Estimate of Parental Income							
College: Term 1	<$20,000	$20,000-$39,000	$40,000-$59,000	$60,000 - $79,000	$80,000 - $99,999	$100,000-$124,999	$125,000 - $149,999	$150,000+
Music & Dance College	2.481	2.826	3.500	2.948	3.233	3.133	3.075	3.240
Tourism & Hospitality Mngmnt	2.555	2.591	2.490	2.790	2.961	2.555	2.325	3.114
Theater Film & Media Arts	2.616	2.974	3.511	3.079	3.185	3.277	3.163	3.169
Education	2.894	3.100	3.112	3.197	3.092	3.239	2.558	3.439
Social Work	2.832	2.619	3.376	3.251	3.515	3.365	2.970	3.280
Liberal Arts	2.764	2.932	2.869	2.996	3.143	3.090	3.059	3.137
Science & Technology	2.521	2.524	2.547	2.306	2.561	2.725	2.816	2.831
Business & Mngmnt Fox School	2.621	2.765	2.879	2.750	2.834	2.897	2.928	2.921
Art Tyler School	2.831	3.065	2.696	3.068	3.322	3.149	3.214	3.081
Media & Communication	2.724	2.968	2.928	3.096	2.929	2.965	3.089	2.957
Environmental Design	3.280	3.270	3.243	4.000	3.860	3.370		3.460
College of Public Health	2.965	2.823	2.841	2.750	2.941	2.981	2.913	3.058
Sport Tourism Hospitality Mgt	3.096	2.341	2.627	2.637	2.804	2.898	2.881	3.180
University Studies	2.571	2.567	2.758	2.394	2.875	2.633	2.838	2.533
Engineering	2.717	2.566	2.479	2.614	2.796	2.954	2.793	2.639
Grand Total	**2.704**	**2.784**	**2.827**	**2.832**	**2.944**	**2.959**	**2.958**	**2.990**

Note: Lower income students majoring in Music, Dance, Tourism and Hospitality, Film, Theater and Media Arts had lower first semester GPAs compared to their more affluent peers. It is likely that these majors require perfect attendance, and penalize students who may have to work an extra shift and thus miss class. These majors may also have rejected earlier credits and required students to

begin their academic programs from the beginning without taking prior coursework into account. These factors contributed to the overall persistence in the college and not only at the university.

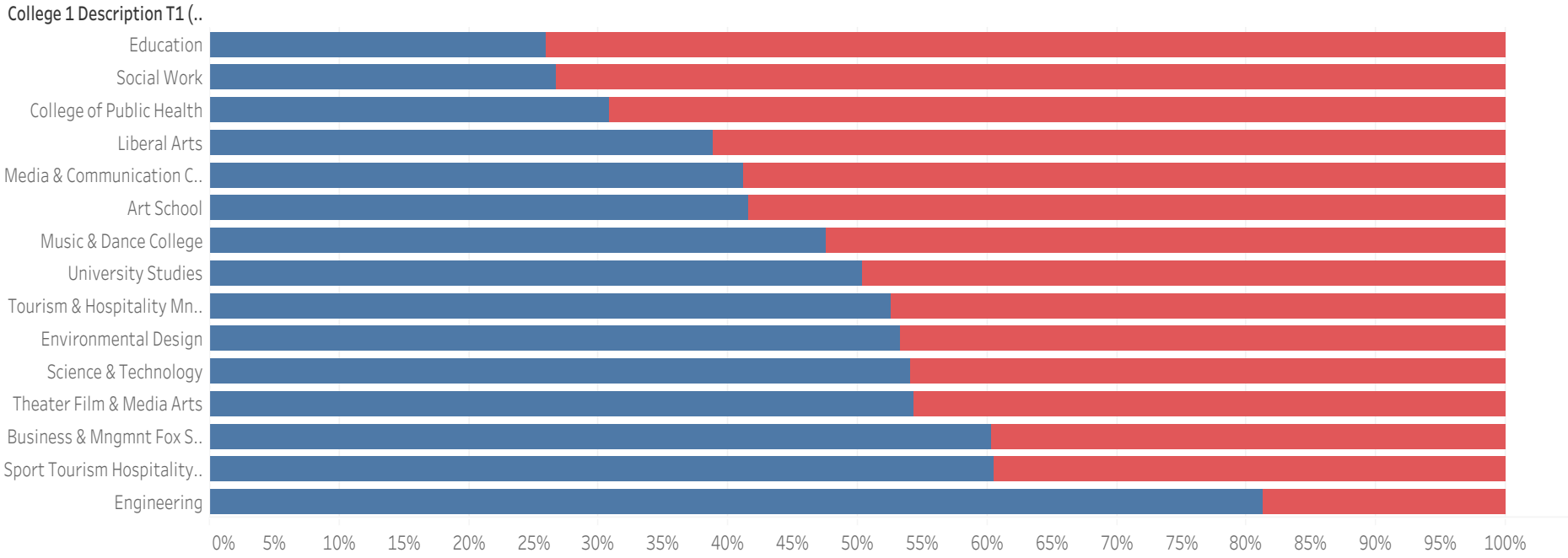
Differences by Gender and Race:

Figure 4.5: Percentage of Students That Persisted (Orange)and did not (Blue) Broken Down by Race and Gender



Note: There were no significant differences between genders when it came to persistence. The differences became visible when gender groups were further subdivided to explore the impact of intersectionality (gender, race, income, major) on persistence and retention. They were even further refined when parental income and number of hours students planned to work for money were added to the analysis. African American and American Indian males had the lowest persistence rates, while Asian females and males had the highest.

Figure 4.6: Percentage of students that were enrolled in each college broken down by Gender (Red = Female, Blue = Male)



There were a number of majors that were 100-80% female or male dominated (Figure 4.5, 4.6, 4.7, 4.8). Female dominated majors included education, dance, speech language and hearing, health professions and nursing, among others. Male dominated majors were made up of engineering, computer science, information technology, finance and supply chain management. These differences are important, as they are also reflective in the salaries most college graduates receive after graduation. Average starting salary for an educator can range around $40,000, the starting salary for someone with an engineering degree could be around $60,000.

Figure 4.7: Male Dominated Academic Programs (Red = Female, Blue = Male)

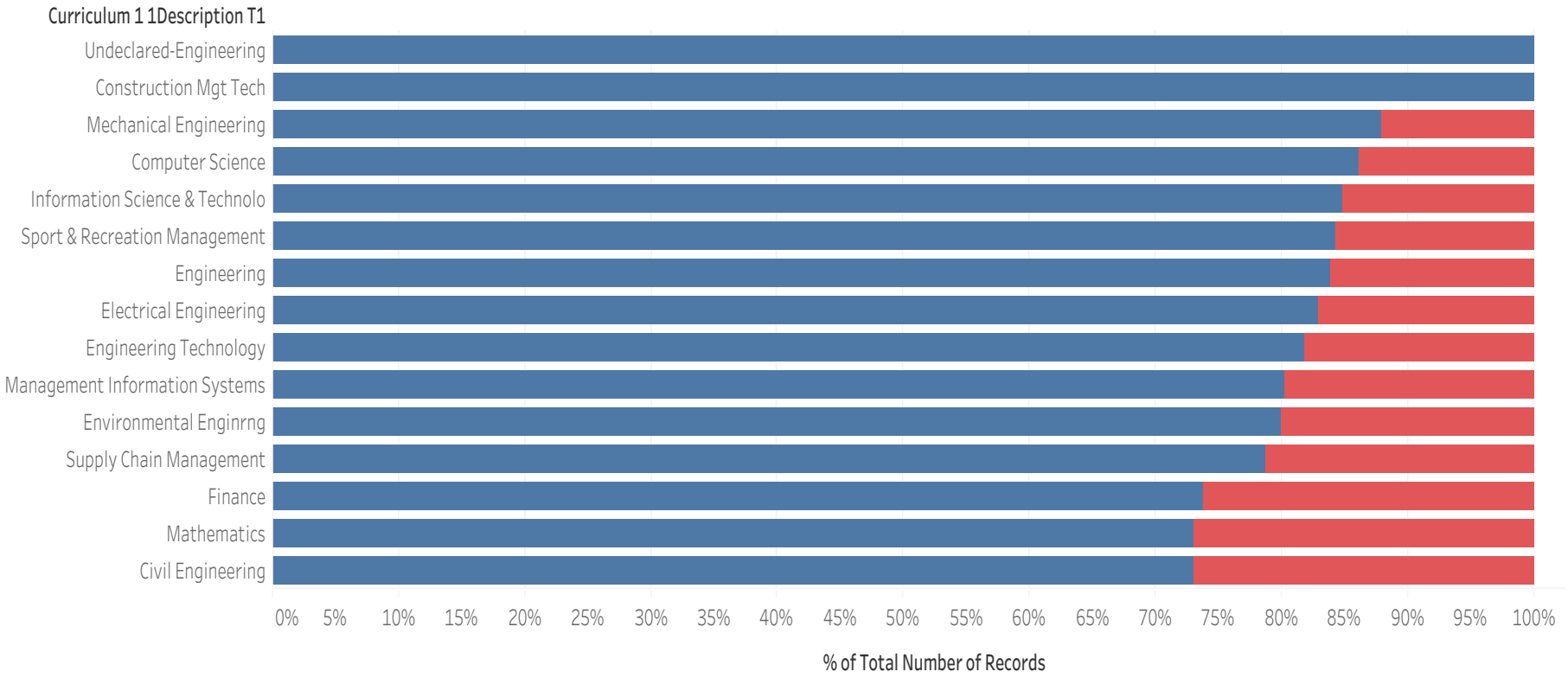
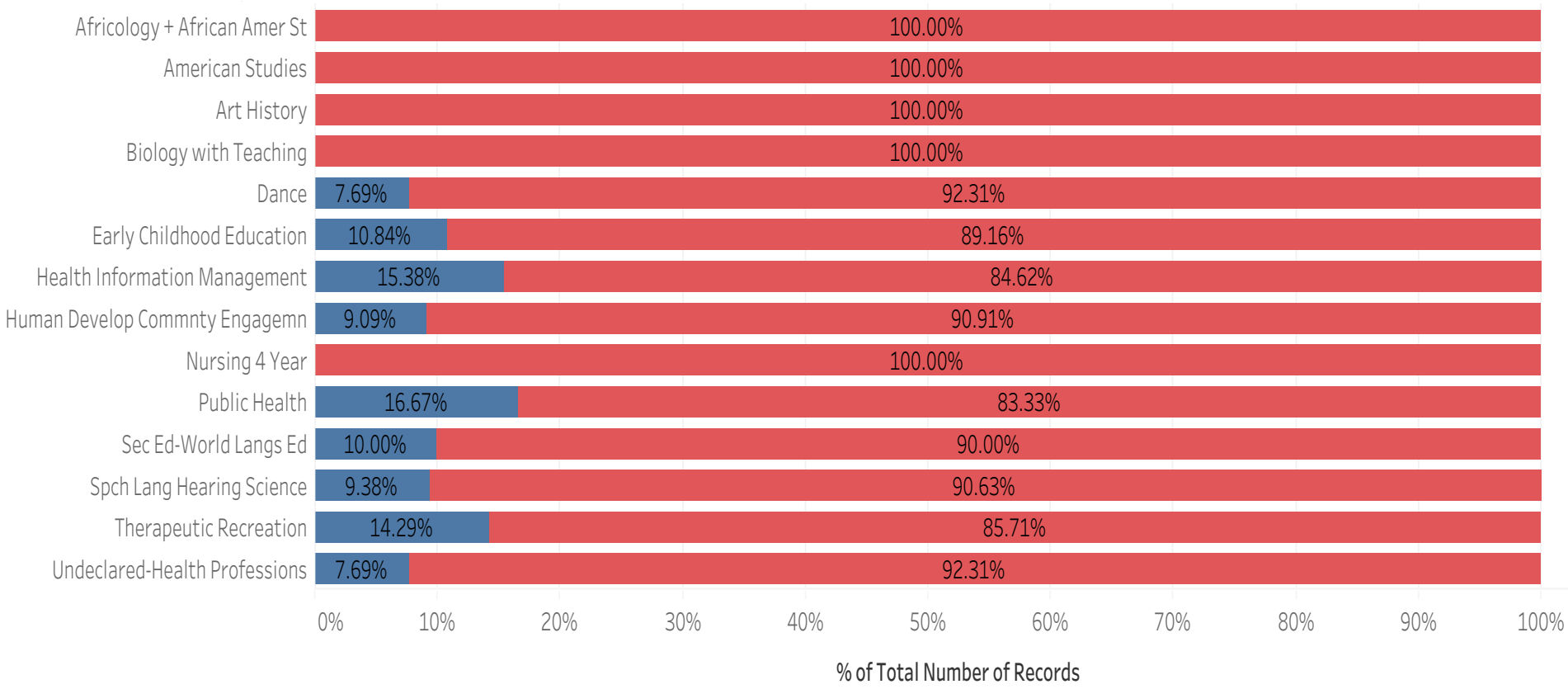


Figure 4.8: Female Dominated Academic Programs (Red = Female, Blue = Male)



Research Question 2:

Do risk factors change or stay consistent from year to year?

Overall the risk factors remained consistent for each cohort year (2015, 2016, and 2017). Top predictors all focused on students' ability to pay, socioeconomic status, academic preparedness and performance, and hours working for money while in school.

Predictors include in all three models are displayed in the table below:

Table 4.15: Top predictors from 2015 – variables with p<.05, effect size >.099

Predictors from 2015	*X2*	*F*	*df*	*p*	*Effect Size*
Cumulative GPA		316.625	2452	0.00	0.93
Term_Undergrad_GPA_T1		319.381	2452	0.00	0.93
Term_Undergrad_Hours_Completed_T1		189.92	2452	0.00	0.66
Number of Mid-Semester Alerts		24.673	1101	0.00	0.55
Term Balance at the End of Term		503.757	2452	0.00	0.5
Hours_Registered_T1		37.582	2452	0.00	0.43
Registration_Start_DayDIF		17.34	2452	0.00	0.42
Count_of_W_Type_Registrations_T1		200.063	2452	0.00	0.35
Balance at the start of term 1		0.208	2452	0.00	0.19
Race	39.01		8	0.00	0.126
Hours plan to work for money per week	31.847		4	0.00	0.117
Work full time	25.278		3	0.00	0.105
Estimate of parental income	24.798		7	0.00	0.105

Table 4.16: Top predictors from 2016 – variables with p<.05, effect size >.099

Predictors from 2016	*X2*	*F*	*df*	*p*	*Effect Size*
Balance at the start of term 1		8.52	497.169	0.00	0.62
Term Balance at the End of Term		592.31	373.006	0.00	0.37
Cumulative GPA		268.55	398.509	0.00	0.35
Term_Undergrad_Hours_Completed_T1		174.46	398.524	0.00	0.35
Count_of_W_Type_Registrations_T1		344.21	373.232	0.00	0.35
Term_Undergrad_GPA_T1		272.83	398.593	0.00	0.33
Number of Mid-Semester Alerts		33.54	268.154	0.00	0.33
Registration_Start_DayDIF		16.09	552.302	0.00	0.26
Curriculum1_1_T1	153.447		115	0.01	0.245
Curriculum_1_1Description_T1	153.447		115	0.01	0.245
Hours plan to work for money per week	50.034		4	0.00	0.145
Estimate of parental income	36.797		7	0.00	0.124
Race	37.23		8	0.00	0.12
College_1_Code_T1	34.027		13	0.00	0.115
College_1_Description_T1	34.027		13	0.00	0.115
Work full time	25.263		3	0.00	0.103

Table 4.17: Top predictors from 2017 – variables with p<.05, effect size >.099

Predictors from 2017	*X2*	*F*	*df*	*p*	*Effect Size*
Balance at the start of term 1		0.129	565.389	0.005	0.73
Count_of_W_Type_Registrations_T1		335.7	419.291	0.00	0.5
Cumulative GPA		303.62	444.01	0.00	0.43
Term_Undergrad_Hours_Completed_T1		206.05	449.005	0.00	0.43
Term Balance at the End of Term		303.51	427.73	0.00	0.36
Term_Undergrad_GPA_T1		311.54	443.736	0.00	0.24
Curriculum1_1_T1	184.349		129	0.001	0.237
Curriculum_1_1Description_T1	184.349		129	0.001	0.237
Registration_Start_DayDIF		19.88	591.317	0.00	0.15
Hours plan to work for money per week	43.698		4	0.00	0.138
College_1_Code_T1	46.179		13	0.00	0.137
College_1_Description_T1	46.179		13	0.00	0.137
Work full time	31.423		3	0.00	0.116
Estimate of parental income	22.353		7	0.002	0.099
Race	23.72		8	0.003	0.099
Number of Mid-Semester Alerts		19.41	317.203	0.00	0.099

Stability of the model indicates that factors rarely change and can be hard-coded into a software application to help predict those with the highest risk-rates at different points of the semester.

Research Question 3:

What are college-specific (e.g., Education vs. Engineering) risk-factors of dropping out?

Oftentimes when a large public institution has multiple colleges, the reasons for dropping out can differ for each college. Prior to conducting the analysis on each college, the dataset was filtered only for students who were originally admitted to that college. The following tables outline top predictors by college, the last chart examines the differences and similarities using one table to compare effect sizes for all top predictors broken down by college.

Table 4.18: Art School Predictors of Persistence

Top Predictors	*X2*	*F*	*df*	*p*	*Effect Size*
Mid-Semester Alerts		7.08	33.842	0.003	0.78
Cumulative GPA		78.98	57.451	0.000	0.88
Hours Registered during Term 1		0.758	300	0.002	0.47
Hours plan to work for money per week	17.210		4	0.002	0.244
Estimate of student income	11.455		4	0.022	0.199
Siblings highest level of education	19.387		5	0.002	0.258
Approximate high school average	9.859		4	0.043	0.185
Chance you will be a student leader	9.846		3	0.020	0.186
Teachers consider me a hardworker	11.298		4	0.023	0.2
Registration_Start_DayDIF		4.69	91.519	0.027	0.31
Balance at the start of term 1		1.669	304	0.001	0.48
Term Balance at the End of Term		64.68	56.905	0.008	0.49
Term_Undergrad_GPA_T1		81.41	57.298	0.00	0.86
Term_Undergrad_Hours_Completed_T1		30.58	58.697	0.00	0.66

Table 4.19: Business School Predictors of Persistence

Top Predictors	*X2*	*F*	*df*	*p*	*Effect Size*
Number of Mid-Semester Alerts		11.96	176.842	0.001	0.33
Count_of_W_Type_Registrations_T1		208.66	276.568	0	0.4
Cumulative GPA		165.82	287.344	0	0.86
Hours_Registered_T1		41	300.571	0	0.55
Hours plan to work for money per week	25.908		4	0	0.116
Estimate of parental income	20.079		7	0.005	0.102
Perception of high school preparation	21.128		4	0	0.103
Race	37.564		8	0	0.135
Registration_Start_DayDIF		31.05	384.052	0	0.47
Balance at the start of term 1		0.33	2056	0.003	0.19
Term Balance at the End of Term		305.42	276.092	0	0.54
Term_Undergrad_GPA_T1		164.93	287.286	0	0.84
Term_Undergrad_Hours_Completed_T1		136.45	290.194	0	0.72

Table 4.20: College of Public Health Predictors of Persistence

Top Predictors	*X2*	*F*	*df*	*p*	*Effect Size*
Number of Mid-Semester Alerts		2.627	261	0.003	0.420
Count_of_W_Type_Registrations_T1		109.28	92.246	0.001	0.490
Cumulative GPA		62.9	101.689	0.000	0.810
Hours_Registered_T1		27.31	104.419	0.000	0.530
Hours plan to work for money per week	29.103		4	0.000	0.236
i04A	10.464		4	0.033	0.255
Approximate high school average	12.921		4	0.012	0.156
High School grade in English	13.839		4	0.008	0.162
Work fulltime	12.470		3	0.006	0.156
Likelihood of making close friends	13.551		3	0.004	0.162
Likelihood of taking online courses	16.811		3	0.001	0.181
Likelihood of working in Philadelphia	12.838		4	0.012	0.158
Enjoy studying and reading	17.779		4	0.001	0.186
Confidence Level	12.500		4	0.014	0.156
Race	17.682		6	0.007	0.179
Registration_Start_DayDIF		2.869	549	0.005	0.330
Term Balance at the End of Term		96.12	92.905	0.001	0.480
Term_Undergrad_GPA_T1		61.62	101.631	0.000	0.820
Term_Undergrad_Hours_Completed_T1		56.39	100.495	0.000	0.800
Transfer_Hours_UG_AD_T1		2.649	549	0.045	0.210
Transfer_Hours_UG_T1		1.021	543	0.047	0.220

Table 4.21: College of Education Predictors of Persistence

Top Predictors	*X2*	*F*	*df*	*p*	*Effect Size*
Count_of_W_Type_Registrations_T1		78.32	57.04	0.016	0.450
Cumulative GPA		115.51	59.639	0	1.200
Hours_Registered_T1		8.46	71.952	0.001	0.550
Hours plan to work for money per week	21.137a		4	0	0.266
Meeting diverse students	7.503a		2	0.023	0.159
Work full time	15.893a		3	0.001	0.232
Chance of playing sports	7.847a		3	0.049	0.163
Likelihood of volunteering	15.030a		3	0.002	0.226
Likelihood of studying abroad	10.169a		3	0.017	0.186
Organization and study habits	9.849a		4	0.043	0.184
Race	20.938a		6	0.002	0.255
Term Balance at the End of Term		202.04	55.378	0.001	0.650
Term_Undergrad_GPA_T1		126.42	59.488	0	1.270
Term_Undergrad_Hours_Completed_T1		35.74	65.207	0	0.920

Table 4.22: College of Engineering Predictors of Persistence

Top Predictors	*X2*	*F*	*df*	*p*	*Effect Size*
Cumulative GPA		46.93	70.052	0.000	0.740
US Military Status	13.192a		4	0.010	0.179
Expected college average	11.989a		4	0.017	0.171
Perception of high school preparation	9.368a		4	0.053	0.152
High School grade in Natural Science	9.966a		4	0.041	0.159
Meeting diverse students	9.311a		2	0.010	0.153
Chance that you will change your major	11.280a		3	0.010	0.169
Work full time	10.388a		3	0.016	0.161
Race	21.777a		7	0.003	0.222
Registration_Start_DayDIF		4.23	98.042	0.011	0.330
Term Balance at the End of Term		62.09	66.026	0.011	0.430
Term_Undergrad_GPA_T1		46.38	70.187	0.000	0.810
Term_Undergrad_Hours_Completed_T1		46.38	70.078	0.000	0.570

Table 4.23: College of Liberal Arts Predictors of Persistence

Top Predictors	*X2*	*F*	*df*	*p*	*Effect Size*
Number of Mid-Semester Alerts		26.31	123.379	0.000	.480
Count_of_W_Type_Registrations_T1		211.04	186.212	0.000	0.450
Cumulative GPA		223.3	195.487	0.000	0.850
Hours_Registered_T1		9.42	213.091	0.000	0.320
Hours plan to work for money per week	13.146a		4	0.011	0.102
Estimate of student income	25.269a		4	0.000	0.139
Mothers highest level of education	18.364a		4	0.001	0.118
Campus Visit Importance	12.987a		3	0.005	0.100
Likelihood of working on research	12.745a		3	0.005	0.100
Likelihood of volunteering	13.958a		3	0.003	0.105
Teachers consider me a hard worker	12.733a		4	0.013	0.100
Grit Level	18.595a		4	0.001	0.121
Registration_Start_DayDIF		17.14	259.55	0.000	0.420
Balance at the start of term 1		0.002	1358	0.024	0.180
Term Balance at the End of Term		156.16	193.474	0.000	0.480
Term_Undergrad_GPA_T1		223.3	195.485	0.000	0.850
Term_Undergrad_Hours_Completed_T1		59.75	200.836	0.000	0.600
Transfer_Hours_UG_AD_T1		0.998	1358	0.031	0.170

Table 4.24: College of Media and Communication Predictors of Persistence

Top Predictors	*X2*	*F*	*df*	*p*	*Effect Size*
Number of Mid-Semester Alerts		17.97	54.37	0.004	0.550
Count_of_W_Type_Registrations_T1		144	93.643	0.001	0.470
Cumulative GPA		121.28	99.286	0.000	0.960
Hours_Registered_T1		11.13	106.602	0.010	0.320
Hours plan to work for money per week	14.776a		4	0.005	0.149
Permanent Home Location	10.324a		4	0.035	0.145
i04A	14.625a		5	0.012	0.270
US Military Status	12.595a		4	0.013	0.135
Estimate of student income	14.495a		4	0.006	0.145
Perception of high school preparation	14.388a		4	0.006	0.145
High School Grade in Natural Science	11.787a		4	0.019	0.131
“I wanted to get away” – enrollment factor	9.099a		2	0.011	0.115
“I wanted a better job” – enrollment factor	9.504a		2	0.009	0.118
Campus visit importance	10.503a		3	0.015	0.124
HS visit by college rep importance	9.321a		3	0.025	0.117
Chance you will be a student leader	10.503a		3	0.015	0.125
Race	18.979a		7	0.008	0.164
Registration_Start_DayDIF		0.015	704	0.008	0.300
Balance at the start of term 1		0.611	704	0.035	0.230
Term Balance at the End of Term		148.09	95.621	0.000	0.580
Term_Undergrad_GPA_T1		121.06	99.295	0.000	0.960
Term_Undergrad_Hours_Completed_T1		50.94	99.093	0.000	0.630

Table 4.25: College of Music and Dance Predictors of Persistence

Top Predictors	*X2*	*F*	*df*	*p*	*Effect Size*
Count_of_W_Type_Registrations_T1		6.55	68	0.005	0.480
Type of residence during first term	10.479a		3	0.015	0.364
# of Years Studied Foreign Lang. in HS	10.645a		4	0.031	0.374
Campus visit importance	9.454a		3	0.024	0.353
Reputation of specific major	9.286a		2	0.010	0.350
Chance of joining clubs	9.315a		3	0.025	0.362
Likelihood of making close friends	9.304a		2	0.010	0.362
Family encouragement	15.323a		3	0.002	0.465
Time Management	7.987a		3	0.046	0.335
Confidence Level	12.810a		4	0.012	0.425
Grit Level	11.903a		4	0.018	0.415
Organization	11.074a		3	0.011	0.395
Study habits	8.420a		3	0.038	0.347
Registration_Start_DayDIF		22.06	42.638	0.000	1.020

Table 4.26: College of Science & Technology Predictors of Persistence

Top Predictors	*X2*	*F*	*df*	*p*	*Effect Size*
Number of Mid-Semester Alerts		3.31	174.143	0.005	0.300
Count_of_W_Type_Registrations_T1		69.04	209.509	0.000	0.410
Cumulative GPA		65.65	221.558	0.000	0.760
Hours_Registered_T1		33.53	225.232	0.000	0.360
Hours plan to work for money per week	13.105a		4	0.011	0.120
Permanent home location	16.822a		4	0.002	0.164
Estimate of parental income	15.891a		7	0.026	0.132
How many yrs. you studied English in HS	17.962a		4	0.001	0.139
How many yrs. you studied Math in HS	18.063a		4	0.001	0.140
Work full time	9.694a		3	0.021	0.104
Likelihood of finding a job – perceived	11.982a		3	0.007	0.115
Race	21.856a		8	0.005	0.151
Registration_Start_DayDIF		1.319	959	0.001	0.280
Term Balance at the End of Term		142.61	198.778	0.000	0.440
Term_Undergrad_GPA_T1		68.27	221.601	0.000	0.790
Term_Undergrad_Hours_Completed_T1		91.86	214.066	0.000	0.590

Table 4.27: College of Social Work Predictors of Persistence

Top Predictors	*X2*	*F*	*df*	*p*	*Effect Size*
Cumulative GPA		32.61	23.112	0.000	1.280
Concerns about paying for college	6.940		2	0.031	0.221
University Rating (1,2,3 Choice)	6.311		2	0.043	0.210
Perception of high school preparation	10.960		4	0.027	0.278
Open house importance	9.641		3	0.022	0.261
Closeness to home importance	9.043		2	0.011	0.255
Likelihood of transferring again	10.725		3	0.013	0.278
Likelihood of taking online courses	11.915		3	0.008	0.293
Registration_Start_DayDIF		2.439	147	0.049	0.470
Term Balance at the End of Term		36.87	22.14	0.054	0.580
Term_Undergrad_GPA_T1		32.51	23.116	0.000	1.280
Term_Undergrad_Hours_Completed_T1		16.93	23.797	0.049	0.550
Transfer_Hours_UG_AD_T1		0.125	147	0.025	0.520
Transfer_Hours_UG_T1		0.7	28.342	0.013	0.620

Table 4.28: College of Sport, Tourism, Hospitality Predictors of Persistence

Top Predictors	*X2*	*F*	*df*	*p*	*Effect Size*
Number of Mid-Semester Alerts		11.14	25.127	0.018	0.690
Count_of_W_Type_Registrations_T1		81.92	30.371	0.015	0.640
Cumulative GPA		25.96	33.62	0.000	0.980
Financial Aid Compared to other schools	9.885		4	0.042	0.220
HS Grade in English	12.531		4	0.014	0.248
Grad. school prep. – enrollment factor	7.147		2	0.028	0.188
Likelihood of working in Philadelphia	14.802		4	0.005	0.271
Race	16.135		6	0.013	0.273
Term Balance at the End of Term		170.87	30.182	0.010	0.690
Term_Undergrad_GPA_T1		26.56	33.61	0.000	0.950
Term_Undergrad_Hours_Completed_T1		7.25	33.363	0.003	0.730

Table 4.29: Film & Theater Predictors of Persistence

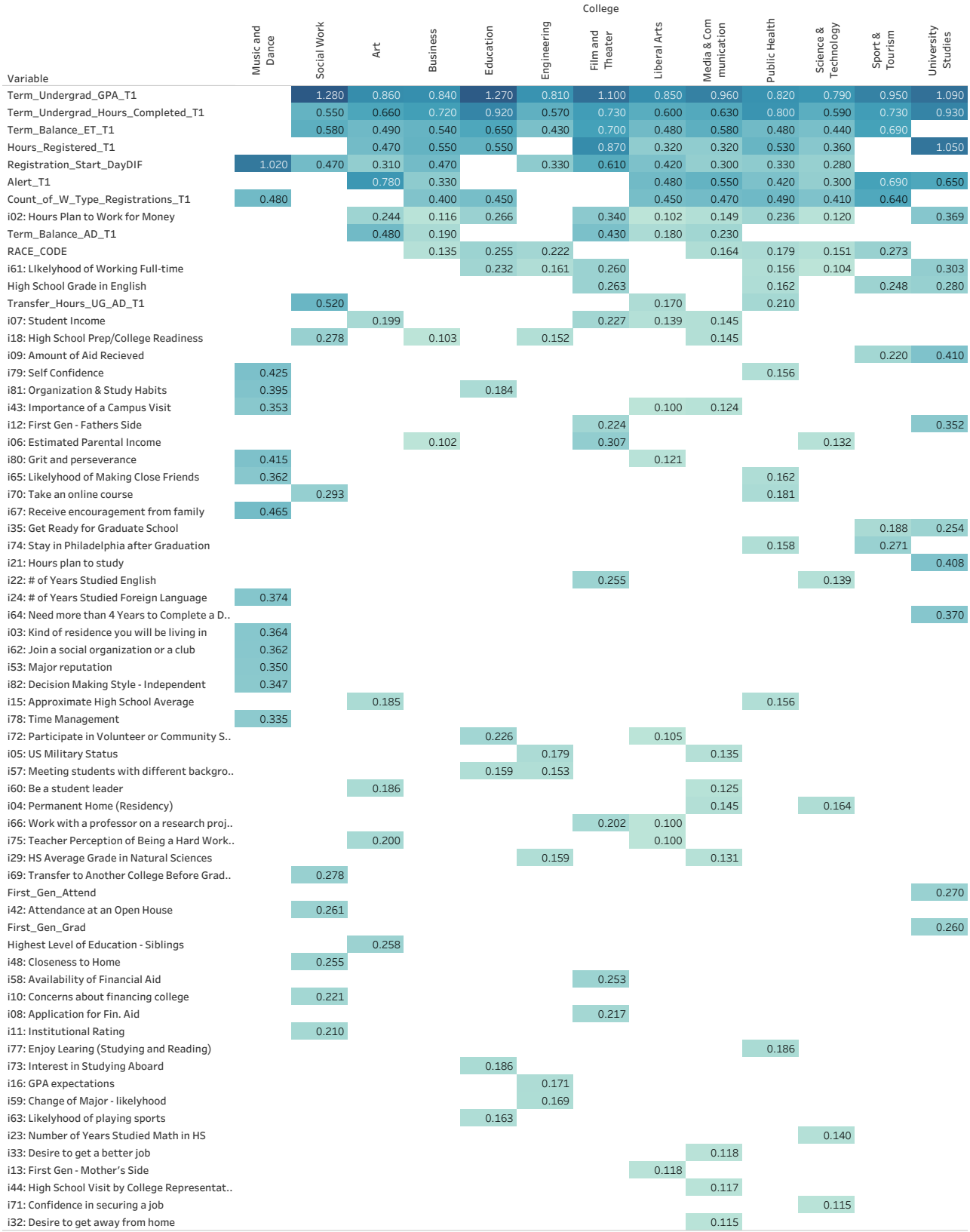
Top Predictors	*X2*	*F*	*df*	*p*	*Effect Size*
Cumulative GPA		49.44	36.948	0.000	1.120
Hours_Registered_T1		2.491	216	0.000	0.870
Hours plan to work for money per week	23.981		4	0.000	0.340
Estimate of parental income	19.357		7	0.007	0.307
Estimate of student income	10.755		4	0.029	0.227
Financial aid application status	9.937		4	0.041	0.217
Father's highest level of education	10.594		4	0.032	0.224
How many yrs. you studied English in HS	13.910		4	0.008	0.255
High school grade in English	14.682		4	0.005	0.263
Availability of aid	13.584		2	0.001	0.253
Work full time	14.372		3	0.002	0.260
Likelihood of working on research	8.723		3	0.033	0.202
Registration_Start_DayDIF		8.39	56.556	0.001	0.610
Balance at the start of term 1		2.881	217	0.012	0.430
Term Balance at the End of Term		114.24	34.572	0.005	0.700
Term_Undergrad_GPA_T1		50.71	36.916	0.000	1.100
Term_Undergrad_Hours_Completed_T1		5.83	40.662	0.001	0.730

Table 4.30: University Studies (Undeclared Students) Predictors of Persistence

Top Predictors	*X2*	*F*	*df*	*p*	*Effect Size*
Number of Mid-Semester Alerts		1.16	57	0.026	.650
First Generation: neither parent attended	8.193		1	0.004	0.270
First Generation: neither parent graduated	7.594		1	0.006	0.260
Cumulative GPA		16.26	21.124	0.001	1.080
Number of Credits Registered		1.559	110	0.000	1.050
Hours plan to work for money per week	14.427		4	0.006	0.369
i04A	10.598		3	0.014	0.502
Financial aid compared to other schools	18.009		4	0.001	0.410
Father's highest level of education	13.369		4	0.010	0.352
Hours you plan to study	18.446		4	0.001	0.408
High School Grade in English	8.562		3	0.036	0.280
i35	7.023		2	0.030	0.254
Work full time	10.017		3	0.018	0.303
Need more than 4 years to graduate	14.897		3	0.002	0.370
Term_Undergrad_GPA_T1		18.28	21.041	0.001	1.090
Term_Undergrad_Hours_Completed_T1		6.83	19.859	0.005	0.930

First Semester GPA, number of credits completed, balance at the start and end of the first term, summer registration date, number of midsemester alerts and the number of 'Ws' are the top predictors that repeat across multiple colleges. Colleges of Music and Social Work predictors were significantly different compared to other colleges. These findings are in-line with previous results that show GPA differences for students from different income levels that attend those colleges. Late registration plays a significant role in student persistence across most colleges; it does not play a role for students attending the College of Education, or those who are undeclared. These findings support the risk factors described in the literature review and focus on the following themes: 1) Financial Concerns 2) Time Management 3) Academic Difficulty & Program Fit 4) Grit/Proactiveness 5) Academic Preparedness 6) Responsibilities Beyond School/ Working to many hours 7) Campus Involvement/Leadership Roles/Research.

Figure 4.9: Top Predictors for Each College and Their Corresponding Effect Size



Variable	College												
	Music and Dance	Social Work	Art	Business	Education	Engineering	Film and Theater	Liberal Arts	Media & Communication	Public Health	Science & Technology	Sport & Tourism	University Studies
Term_Undergrad_GPA_T1		1.280	0.860	0.840	1.270	0.810	1.100	0.850	0.960	0.820	0.790	0.950	1.090
Term_Undergrad_Hours_Completed_T1		0.550	0.660	0.720	0.920	0.570	0.730	0.600	0.630	0.800	0.590	0.730	0.930
Term_Balance_ET_T1		0.580	0.490	0.540	0.650	0.430	0.700	0.480	0.580	0.480	0.440	0.690	
Hours_Registered_T1			0.470	0.550	0.550		0.870	0.320	0.320	0.530	0.360		1.050
Registration_Start_DayDIF	1.020	0.470	0.310	0.470		0.330	0.610	0.420	0.300	0.330	0.280		
Alert_T1			0.780	0.330				0.480	0.550	0.420	0.300	0.690	0.650
Count_of_W_Type_Registrations_T1	0.480			0.400	0.450			0.450	0.470	0.490	0.410	0.640	
i02: Hours Plan to Work for Money			0.244	0.116	0.266		0.340	0.102	0.149	0.236	0.120		0.369
Term_Balance_AD_T1			0.480	0.190			0.430	0.180	0.230				
RACE_CODE				0.135	0.255	0.222			0.164	0.179	0.151	0.273	
i61: LIkelyhood of Working Full-time					0.232	0.161	0.260			0.156	0.104		0.303
High School Grade in English							0.263			0.162		0.248	0.280
Transfer_Hours_UG_AD_T1		0.520						0.170		0.210			
i07: Student Income			0.199				0.227	0.139	0.145				
i18: High School Prep/College Readiness		0.278		0.103		0.152			0.145				
i09: Amount of Aid Recieved												0.220	0.410
i79: Self Confidence	0.425									0.156			
i81: Organization & Study Habits	0.395				0.184								
i43: Importance of a Campus Visit	0.353							0.100	0.124				
i12: First Gen - Fathers Side							0.224						0.352
i06: Estimated Parental Income				0.102			0.307				0.132		
i80: Grit and perseverance	0.415							0.121					
i65: Likelyhood of Making Close Friends	0.362									0.162			
i70: Take an online course		0.293								0.181			
i67: Receive encouragement from family	0.465												
i35: Get Ready for Graduate School												0.188	0.254
i74: Stay in Philadelphia after Graduation										0.158		0.271	
i21: Hours plan to study													0.408
i22: # of Years Studied English							0.255				0.139		
i24: # of Years Studied Foreign Language	0.374												
i64: Need more than 4 Years to Complete a D..													0.370
i03: Kind of residence you will be living in	0.364												
i62: Join a social organization or a club	0.362												
i53: Major reputation	0.350												
i82: Decision Making Style - Independent	0.347												
i15: Approximate High School Average			0.185							0.156			
i78: Time Management	0.335												
i72: Participate in Volunteer or Community S..					0.226			0.105					
i05: US Military Status						0.179			0.135				
i57: Meeting students with different backgro..					0.159	0.153							
i60: Be a student leader			0.186						0.125				
i04: Permanent Home (Residency)									0.145		0.164		
i66: Work with a professor on a research proj..							0.202	0.100					
i75: Teacher Perception of Being a Hard Work..			0.200					0.100					
i29: HS Average Grade in Natural Sciences						0.159			0.131				
i69: Transfer to Another College Before Grad..		0.278											
First_Gen_Attend													0.270
i42: Attendance at an Open House		0.261											
First_Gen_Grad													0.260
Highest Level of Education - Siblings			0.258										
i48: Closeness to Home		0.255											
i58: Availability of Financial Aid							0.253						
i10: Concerns about financing college		0.221											
i08: Application for Fin. Aid							0.217						
i11: Institutional Rating		0.210											
i77: Enjoy Learing (Studying and Reading)										0.186			
i73: Interest in Studying Aboard					0.186								
i16: GPA expectations						0.171							
i59: Change of Major - likelyhood						0.169							
i63: Likelyhood of playing sports					0.163								
i23: Number of Years Studied Math in HS											0.140		
i33: Desire to get a better job									0.118				
i13: First Gen - Mother's Side								0.118					
i44: High School Visit by College Representat..									0.117				
i71: Confidence in securing a job											0.115		
i32: Desire to get away from home									0.115				

Research Question 4:

What are the determinants of graduation among transfer students?

Analysis conducted for this question examined factors that predicted student' ability to graduate in three years after the point of admission. Even though many transfer students come in with credits, for many of these students some of these credits are not usable towards their new degree. It is also important to note that students who do not persist cannot graduate; thus, factors that contribute to persistence will also be the same factors that contribute to graduation.

Table 4.31: Top predictors – variables with p<.05, effect size >.099

Predictors from 2015 Cohort	*X2*	*F*	*df*	*p*	*Effect Size*
Term_Undergrad_GPA_T1	321.29		2142.34	0.00	0.87
Transfer_Hours_UG_T1	62.54		2438.92	0.00	0.69
Number of Mid-Semester Alerts	128.07		1079.35	0.00	0.57
Term_Undergrad_Hours_Completed_T1	175.7		2267.27	0.00	0.51
Registration_Start_DayDIF		30.9	2272.22	0.00	0.47
Hours_Registered_T1	103.34		2379.01	0.00	0.36
ADMISSIONS_MAJOR_LD	272.198		112	0.00	0.33
Count_of_W_Type_Registrations_T1	232.82		2037.69	0.00	0.33
Term Balance at the End of Term	228.53		1778.61	0.00	0.32
ADMISSIONS_COLLEGE_LD	136.162		13	0.00	0.24
Balance at the start of term 1		5.1	2361.84	0.00	0.22
Race	61.848		8	0.00	0.16
Need more than four years to graduate	35.615		3	0.00	0.13
High School Grade in English	28.496		4	0.00	0.11
Chance of changing your major field of study	26.917		3	0.00	0.11
Siblings highest level of education	27.595		5	0.00	0.11
Estimate of parental income	24.551		7	0.00	0.10

Transfer Hours_UG was one of the variables present in the graduation model and not in the retention model. This study did not examine the number of hours students had prior to matriculation, only the number of hours that were accepted by their new institution. It is also important to note that not all accepted hours can be used towards a degree; many transfer students find themselves with excess credits. These credits are counted towards the degree but

not towards the academic program declared by the student. Overall, transfer students who graduated in three years were coming in with slightly higher number of credits than those who did not. It is likely that these students were also coming in with AA degrees completed and were waived from doing general education requirements, thus producing a faster route to graduation. Students who were coming in as undeclared had significantly fewer transfer credits than other students.

Table 4.32: Average Number of Transfer Credits for Students Graduating in Three Years, Compared to Those Who Did Not Graduate in The Specified Timeframe (broken down by college)

Predictors from 2015 Cohort	*Did not Graduate*	*Graduated*
College – Term 1	*Average # of Transfer Credits*	
Art School	44.09	60.89
Business & Management	42.65	59.42
College of Public Health	43.74	56.31
Education	49.31	60.5
Engineering	52.79	74.88
Environmental Design	49.41	50.25
Liberal Arts	42.32	55.48
Media and Communication	36.24	49.53
Music & Dance College	42.06	60.88
Science & Technology	46.66	59.87
Social Work	52.95	55.94
Theater Film & Media Arts	41.53	51.84
Tourism & Hospitality Management	38.6	52.85
University Studies - Undeclared	25.4	35.81
Total	43.63	56.86

New Student Questionnaire Answers

Many students were aware that the degree will take them longer than four years, as seen by their answers on question 64 of the new student questionnaire. High School English grades were also predictive of the likelihood to graduate faster. This is not surprising as a significant amount of literature points to students' earlier levels of literacy predicting college graduation. Students who indicated that they have a high likelihood of changing their majors also took longer to graduate. Students who had no siblings or had siblings with college degrees or experience had higher graduation rates than students whose siblings had no prior college experience. This finding points to families' lack of experience with higher education, as well as issues families may have funding college when both children are enrolled.

College of Enrollment

There are significant differences in three-year graduation rates among different colleges. Music and dance students, undeclared students, science and engineering students had much lower 3-year graduation rates compared to students who were enrolling in Communication, Social Work, Film or Liberal Arts. The curriculum structure of these programs can be further examined to see how major complexity plays a role in the student's path to graduation.

Table 4.33: 2015 Transfer Cohort - Three Year Graduation Rates By College

Predictors from 2015 Cohort	*Graduated*	*Did Not Graduate*
Media and Communication	63.75%	36.26%
Social Work	62.50%	37.50%
Theater Film & Media Arts	56.16%	43.84%
Liberal Arts	54.95%	45.05%
Business & Management	43.98%	56.02%
College of Public Health	43.00%	57.00%
Education	42.06%	57.94%
Art School	36.36%	63.64%
Tourism & Hospitality Management	34.62%	65.38%
Engineering	31.91%	68.09%
Science & Technology	28.14%	71.86%
Environmental Design	26.67%	73.33%
University Studies - Undeclared	21.62%	78.38%
Music & Dance College	19.05%	80.95%
Total	44.78%	55.22%

Special Note: There were 15 students in the dataset who did not persist into term 3; however, they did graduate in a three-year time span from the point of their enrollment. This further highlights that transfer students are much more likely to stop out and return at some later time to complete their degree.

Table 4.34: Data Abnormality (Did not Persist, but Graduated)

3Rd Semester Enrollment (2015 Cohort)	*3 Year Grad*	
	Did Not Graduate	*Graduate*
Did Not Persist	330	15
Persisted	1025	1084

CHAPTER 5
DISCUSSION

Free the data, turn on the data, share the data

~ Geoffrey Moore

During the last fifty years, numerous campus services, activities and programs have been created to assist students identified as "at-risk". Most often these programs serve first-generation, first-year, first-time in college students and center on college preparedness, financial literacy, academic preparedness and institutional engagement, climate and fit. Some colleges and universities include wellness and psychological issues, provide orientations to both students and parents; others offer one credit seminar courses that address issues mentioned above, as well as teach students about accountability, expectations, professionalism and work ethic. Very few of these efforts address the needs of transfer students, and even fewer are effective at increasing graduation rates for this population. Many campus programs are executed without strategic planning and assessment, and do not have a direct impact on the retention and graduation rates. Individualized approaches, intrusive advising, academic coaching, and additional funding to cover tuition and living expenses all have been shown to lift retention rates. These services are labor-intensive, and if provided to the entire student population can create a significant drain on institutional resources. Targeted interventions driven by reporting and predictive analytics can help identify students in need of additional help and guide them at the appropriate time.

Transfer shock is not a new concept, but it is one that is difficult to validate without proper data. Transfer shock can also manifest differently for community college students compared to students who are switching between four-year institutions. First-generation,

minority students, and students from low-income families may feel out of place on a new campus after transferring. These students may be able to afford tuition but may struggle with living expenses, book costs, and additional unforeseen college expenses. Unexpected expenses like orientation costs, meal plan, entertainment and extracurricular costs, health insurance, parking, and major-specific costs (course lab fees, or equipment rental fees) may trigger the feeling of isolation and inadequacy. These feelings can manifest into what other researchers have defined as an "Imposter Syndrome" and lead to self-doubt, poor academic performance, and eventually withdrawal. Although these feelings may be completely unjustified from the academic standing perspective, the element of social support and integration is critical to college success. Many transfer students will need to build new social support and friendship networks and explore ways to get involved on campus.

Continuous enrollment plays a large role in the student success and degree completion. Students who take time off, pursue their degree part- time and take less than six credits each semester can encounter many difficulties upon transferring to their new institution. Re-enrollment tasks like transcript requests from prior institutions and negotiations with loan providers (both for deferments and possible defaults) can hinder future academic success and stand in the way of degree completion. Transition to college and adjustment to upper-division courses can be especially difficult for students who already struggle with managing living expenses and may be working 20+ hours to support themselves and their families. It is absolutely critical for transfer students to have a clear picture of what credits will be accepted by their new institution and what degree requirements are left for them to complete. For some students this process can be easily facilitated by following an articulation agreement; for others, academic advising and

admission staff have to do due diligence in articulating all prior coursework. In addition to transfer credit equivalencies, students must also be aware of prior learning assessment and credit for experience policies at their new institutions.

Chapter five provides guidance and implications for practice to university and college administrators on how to design affective transfer programs using earlier discussed findings. The recommendations include best practices that have been tested at multiple institutions and align with findings from this study. Implications for policy and future research are also discussed, highlighting the need to continue focusing on the transfer student population both at the institutional level and nationally.

Implications for Practice

There is much to consider when it comes to creating a successful transfer-friendly campus. Universities that are beginning to focus on the transfer population and have yet to adopt strategies to support this student group can model their services on best practices that are working at other universities and colleges. One example of such services is the creation of the Transfer Center, a one-stop shop for transfer students. The administrative staff in the Transfer Center would help transfer students navigate curricular pathways and make a seamless transition to the institution, provide holistic and developmental advising and plan thoughtful transfer open houses, orientation and welcome week events. Administrative staff working with transfer students must understand the concerns of today's transfers, be knowledgeable in how to handle students experiencing "transfer shock" and be proactive in addressing any issues that are present at the institution (e.g., old articulation agreements or equivalency rules, mis-aligned curriculum sequences or inflexible departmental culture when it comes to transfer credit evaluations). Most importantly, the center staff must be attuned to

the demographic differences and the needs of the African American and Minority students who are often first in their family to attend college. The center must provide a combination of programming that supports both the transfer population and minority students. In this study alone, 1 out of 5 African American students and Hispanic students did not persist into their second year (see Table 4.2 and Table 4.7). In addition to supporting minority students transfer counselors must understand the role family income and socioeconomic status plays in one's academic journey. One out of five students who come from families making less than $40,000 (see Table 4.3) do not persist into their second year. First-generation students who work more than 25 hours a week are also considered at high-risk of dropping out. Many of these students choose to go to work instead of class in order to survive (see Table 4.6). Institutions must become aware of students struggling with food and housing insecurities and provide appropriate refers to support these students. Faculty should set clear expectations for attendance and participation, academic advisors and student success coaches must address time management, priorities and academic expectations with all new and incoming students.

Every single model completed as a part of this study pointed to debt/balance at start of term as the most predictive factor in trying to predict persistence. In addition to being knowledgeable about academic requirements Transfer Center personnel can assist students with the Free Application for Federal Student Aid (FAFSA), institutional and external scholarships, help resolve holds and answer questions about student bills and ways to secure work-study positions on campus. This combination of academic advising/financial aid approach can significantly improve the transfer student experience during the most critical time of transition.

Senior leadership working on transfer issues may need to conduct assessments and surveys to gauge campus climate. In collaboration with Business Intelligence (BI) and Institutional Research (IR) teams, senior leaders can develop a clear picture of the transfer population by creating dashboards using institutional and Clearinghouse student data and implement intervention programs to support transfer students who are "at-risk" of dropping out. In addition to the BI and IR teams, it is essential for senior leadership to raise awareness of transfer issues with the faculty, especially if the institution is relying on transfer enrollments for revenue generation. By starting a dialog with faculty on this topic, and developing strong partnerships with feeder institutions, colleges can focus both on the enrollment and the retention and success of new transfers.

By examining policies surrounding transfer credit, life-experience, residency requirements, and training faculty and staff to serve as transfer allies, any institution can ensure that the campus culture is not only transfer-receptive but transfer-supportive. All offices, regardless of what stage of the student lifecycle they interact with the students, can develop a successful transfer action plan, implement transfer welcome programs and create social media initiatives to improve their approach to working with this population significantly.

Professional organizations like National Academic Advising Association (NACADA), Consortium for Student Retention Data Exchange (CSRDE) and National Association of Colleges and Employers (NACE), as well as many others provide guidance, professional development and consulting services to institutions that wish to improve services for specific populations or develop risk models and interventions that will streamline retention and graduation initiatives. Conferences organized by these organizations provide

administrators with an opportunity to learn about best practices and network with a colleague from other institutions. College administrators must also understand the complexities of the K-12 system, the issues of school funding, gerrymandering and school choice in order to draw connections between higher education and K-12 issues, and how finances (both personal and institutional) shape educational experience. Administrators must also be familiar with tools like the Student Achievement Measure (SAM) and the Beginning College Survey of Student Engagement (BCSSE).

SAM provides a way to track student progress and degree completion by tracking students who enroll in multiple institutions, thus an excellent resource for tracking transfer students. SAM allows for tracking an increasingly mobile population and getting a complete picture of enrollment for students who may have left the institution to pursue a degree at another university, thus couldn't be really counted as dropouts. This tool is useful for tracking students in certificate, associate, and bachelor's degree programs.

Students who complete the BCSSE Survey are given an in-depth psychometric report that can be revised with the academic advisors or a coach and develop a plan that would address areas of concern. Survey questions focus on students' expectations for the academic year and examine prior academic experiences, reasons for leaving the prior institution, learning style and strategies, and anticipated stressors for the upcoming academic term. The data collected from BCSSEE can be used to develop risk instruments necessary to predict attrition or failure to succeed.

To gain a better understanding of the transfer data, institutions can enlist the help of the Business Intelligence team to create powerful dashboards using Cognos, Power BI, Tableau or other data visualization tools. These dashboards can help senior administration

gain insight at a college level and provide a big picture of student migration from the feeder schools (schools with a high number of students in transfer). At a college level, dashboards can help visualize enrollments by major, specific department, course, or a student level. Data visualization tools can also be used to calculate enrollment yields, articulation agreement effectiveness, retention and graduation statistics by special populations, DFW reports used to build supplemental instruction and faculty development programming and track individual student success.

Programs That Can Improve Retention

Minority Male Mentoring (African American, Hispanic, American Indian)

This study highlighted differences between races when it comes to persistence and academic performance (see Table 4.7). Minority males represented one of the largest groups of stop outs and were identified to be at the highest risk for dropout. Retention numbers for African American and Hispanic males are significantly lower than white students. In order to help this vulnerable group of students, institutions should increase their commitment to effectively integrate minority students into the campus culture, provide academic and social support, and develop Minority Male Mentoring programs. The success of these programs lies in staff and faculty's willingness to serve as mentors, change and re-evaluation of institutional policies, targeted programming, and institutional commitment to funding these initiatives.

A successful mentoring program aims to refine further and develop the mentee's, abilities, skills, and understanding of the new environment. In order for these mentoring programs to become successful, they must be designed with supporting, warm, and nurturing atmosphere and with significant involvement of African American minority faculty and staff.

They must be funded to be sustainable and must be embraced by the entire institution, not just individual departments or diversity champions on campus. Mentoring can strengthen self-esteem, increase career and professional development opportunities, improve academic and social skills, and most importantly increase one's likelihood of obtaining a college degree.

Summer Bridge/High School and Middle School Programs

There are significant differences in persistence for students in majors like Pre-pharmacy, biochemistry, engineering and mathematics (see Table 4.9). Early Exposure to Science/Mathematics/Engineering at Technology is crucial for success in college level courses. Summer bridge programs in STEM, faculty partnerships across all levels in K-16, and scaffolding of academic supports like supplemental instruction and tutoring at all levels of academic curriculum. Affordable summer bridge programs may save STEM students an additional year of tuition if executed properly and developed with student strengths and school and academic district deficiencies in mind.

Articulation Agreements

The data in this study shows that students who transfer in with 60+ credits, compared to those who come in with the average of 45 are much more likely to graduate in three-year period. Articulation agreements can provide a mechanism to make sure students take the right credits at their first institution and have those credits count towards the degree after they transfer. Development of articulation agreements can serve as an initial step in creating partnerships between a four-year institution and two-year institution. Well-executed, centrally-managed articulation agreements can provide a pathway to a better curriculum planning and updates process. As departments and faculty update and plan their course

sequences, those departments that invite transfer applications will need to make sure that alignment exists between their course content and content of courses for peer institutions. Articulation agreements can focus on general education requirements or define program to program rules. They can also help align workflows for course petition approvals and facilitate institutional transfer policy reviews. Articulation agreements can also serve as recruitment tools and a way to strategically market academic programs to prospective students and their counselors. Specifically, for programs that have low female or minority enrollment numbers. For STEM students, articulation agreements can pave a path to degree completion within a reasonable timeline and outline all major pre-requisites and barriers to completion.

Admissions

Admissions and strategic enrollment staff spend a considerable amount of time with new and prospective students. The transfer credit evaluation process for all submitted transcripts often falls on the admissions team. This process can be expedited with the help of a transfer equivalency tool, like the one developed by companies like Transferology or CourseLeaf. In addition to transparency in course equivalencies, institutions must display the full cost of tuition that includes housing, books, meal plans, and special course fees for each academic program. Admissions counselors must be comfortable in discussing program costs with prospective students during Transfer Welcome days and other admissions events. If possible, transcripts should be evaluated early in the admissions cycle, including courses in progress on the student's records. Fast processing of transfer credits and prior learning assessments will facilitate better advising relationships and decrease the chances of errors or misinformation during advisor-student conversations. If at all possible, admissions offices

should stop admitting students two weeks before the start of the semester. Data shows that students admitted late into the summer do not persist at the same rate as their peers (see Table 4.11). Furthermore, one out of three students who are admitted and enroll part-time while working 20 plus ours a week do not return for their second year (Table 4.10). This group makes up about 4% of the transfer population and can be targeted directly to ensure they do not loose academic momentum after first semester and feel supported by the administration and faculty.

Last two weeks of summer should be used to run special programs (Summer Bootcamps) for transfer students admitted with academic deficiencies or low placement scores. In addition, students enrolled in Dual Admissions programs, who are in good academic standing, should register early in the summer for their fall courses, to make sure of proper placement and course availability. Ideally these summer programs would target students admitted to high drop-out majors (See Table 4.8). Community College students who intend to transfer should be advised to create and maintain an administrative binder that contains relevant documents. The contents of the binder should contain the following documents:

- Syllabi for all Courses taken prior to matriculation
- Copies of transcripts from all institutions attended w/Degree Conferred (including high school)
- Description of degree program students have completed at the Community College: AA/AAS/AS obtained from the bulletin from the term of admission
- Score reports for AP/CLEP/Prior Learning Credits received (if applies)
- Descriptions of any study abroad experiences and records of internships (if applies)
- IB diploma (if applies)
- Prior placement test results or any waivers received

- Current registration (Summer/Spring if registered)
- Financial Aid and Scholarship Documentation and letters
- Copy of Articulation Agreement, if one was followed
- Copy of transfer credit evaluations received
- Copy of the admission documents and college welcome letters

Orientation Programming

Transfer Orientation should consist of a full day on campus and be coordinated by Student and Academic Affairs. At the point of orientation, transfer students should meet with financial aid counselors and academic advisors to discuss tuition, payment plans, and course curriculum. Ideally, students should be pre-registered for courses and their transcripts evaluated before orientation; advisors should make a minor adjustment to student schedules and focus more on explaining curriculum sequencing and the elements that make up the degree. Each transfer students should be assigned to a designated advisor, who works closely with the admissions team to facilitate a smooth transition to the university and may also be knowledgeable about additional services like Pre-Law, Pre-Med or graduate school requirements.

This study specifically highlighted the issues of persistence for first-generation students working 20+ hours a week. First-generation students, students with high unmet need, and those who plan to work more than 20 hours a week should be matched with academic success coaches who will help with time management, prioritization and serve as mentors during the student's academic journey. During welcome week, transfer students should be invited to an on campus part-time job fair to help those who need a new job closer to campus identify suitable employment options. Additional Welcome week events may

include transfer t-shirt swap, networking dinners, parent orientations, off-campus housing tours, and representative presentations.

In addition to a face- to-face orientation, transfer students should have an option to complete an online program as well as proctored online testing. Online orientation and any assessments completed as a part of the onboarding process should be transfer specific and vetted for that population. Institutions should have separate social media campaigns specifically designed for transfer students, and regular e-newsletters that answer common transfer questions often asked during the planned on-campus events.

Transfer Seminars

As described in chapter 2, transfer seminars serve a vital role in a successful transfer student transition to the university. These seminars can introduce students to the resources and opportunities at their new university while integrating principles of career development into the curriculum. Students who complete transfer seminars will be much more comfortable in identifying and using campus and community resources. The structure and the curriculum of the seminars help students to connect with the University offices and student organizations. Students enrolled in the seminars are guided by their instructors and as a result:

Network and build relationships with other transfer students, campus leaders, and community partners

· Learn about academic opportunities and resources available at their new institution

· develop an understanding of how to create an academic plan and prepare for the registration process for subsequent semesters

· Identify values, interests, skills, and strengths, as well as transferable skills as they relate to varying career communities

· Explore and identify more information about desired career paths and occupations

· Understand the process for obtaining an internship, studying abroad, applying to graduate school, and the job search process.

Seminars increase student engagement and help promote retention and student satisfaction. Transfer seminars can also be topic or population-specific (i.e., First Gen, STEM, Undeclared, Pre-Health, International, Graduate School Prep)

Transfer Academic Advising

One of the most important aspects of a transfer transition is a smooth handoff and transition of students from admission to advising. Cohort based advising models are optimal for the best student experience. Advisors should track each cohort and monitor the performance of students using a learning management system or an external database/client relationship management system (CRM). CRMs are critical for document storage, tracking student notes, and maintaining student files. Learning management systems can also serve as a depository of internal policy and planning documents that can be easily shared with multiple departments and individuals.

Designated transfer advisors should receive additional training and meet together to discuss transfer student issues and cases, review policies, and plan interventions for students struggling to meet college expectations. Advisors should assist students with the transition to college, provide timely responses to student questions, and serve as advocates for students and the institution. Advising directors should also plan and execute programming designated to assist part-time transfer students, review policies surrounding transfer articulation, and

work with department chairs to make changes in curriculum arrays that do not negatively impact existing articulation agreements. Advising offices should formulate a Transfer Student Success Accountability Framework and develop Transfer Advising syllabus to assess student programming and interventions. Materials are available using Council for the Advancement of Standards in Higher Education (CAS) to support advising teams. Advising standards should be reviewed at least on a bi-annual basis.

Risk-Based Intervention Design

Academic Advising should be supported and based on strategic planning and data-driven decision making. Robust and college-specific risk-tools should be developed to identify populations of students that are struggling or have the highest risk of dropping out. Campus administrators should allocate resources to support these small pockets of at-risk students and guide and support them using specifically trained support personnel. Students should not be discouraged from pursuing majors like Engineering or Nursing simply because of their prior academic preparation. On the contrary, support systems, like academic and life coaching, should be mandated for those who are likely to take more than four years to complete their degrees.

Risk models should be validated using at least three to five years of student data; they should be designed with actionable data points that lead to specialized interventions, which in turn could be assessed to see their effectiveness. Advisors should be able to answer the following questions at any point of the semester:

- Who is at risk?
- Why are these students at risk?

- What interventions and communication strategies are being implemented to support at-risk students?

Academic advisors should be cross-trained as student success coaches, and receive training on tools like LASSI, FOCUS 2, Life-Wheel, DISC, and other assessments. Academic coaches or advisors should engage, influence, and inspire transfer students to achievable goals, meet deadlines, and develop parallel plans. Advising intervention design should be centered around deadlines, strategic communication, and assessment and utilize Project Management framework, Cognitive Behavior Therapy techniques, and multiple theories of student development. These tools are most effective when introduced early in the student's academic career (such as a first-year seminar course) and then again to students who have two or more unsatisfactory mid-semester alerts or those who withdrew from 2 or more courses in their first semester (See Table 4.13 and Table 4.14). Institutions may be able to finetune their outreach and identify attrition factors that are unique to their institutions. Colleges and department may also want to finetune prediction algorithms for their own populations, and design an intervention strategy using those, more specialized models.

Attrition Loop

For years, families have relied on the Free Application for Federal Student Aid (FAFSA), to help determine the amount of grants, work-study, and loans a student can receive. The application uses parental income and the number of dependents in college. Completely driven by the information from W-2 forms and 1040 Tax Forms, it is often a complete puzzle to average working families and a cause for much frustration and stress during the college admission process. Many families find themselves too rich for financial aid, with the Estimated Family Contribution (EFC) (a number calculated by FAFSA), being

completely unrealistic. A FAFSA application takes into consideration the family's income, savings, assets, and more, but completely fails to look at debts, cost of living, any medical or other expenses, parent's own student debt or the need to save for retirement.

For many students, admission to their dream institution comes with a hefty price tag, also called the cost of attendance by the financial aid counselors. That cost is lowered by federal and institutional aid and public and private scholarships. After all aid is exhausted students are left with unmet need, a balance that needs to be paid off before the start of term. Many students opt into payments plans to pay down that balance and seek to pursue employment on top of their studies.

Data shows that one out of four first-generation transfer students from low-income families who work more than 20 hours a week will not make it past their first year at the new institution. These students do not engage with the university, participate in extracurriculars, or join student organizations; they go from parking lot to the classroom to their job off-campus. Although many students manage academics with part-time jobs without any problems, some can experience significant struggles as a result of being overwhelmed or poor time management. Students who experience difficulty need financial aid and academic coaching.

Academic Coaching/ Peer Mentor Coaching

Academic coaches are best used to guide intrusive interventions and work with high-risk students to promote retention and student success. Coaches may work with transfer students at any point in their academic career; however, they are most effective during the first several months of the semester. Coaching relationships can be ongoing or short-term and can be initiated by an alert or by the student him/herself.

Sample Alerts:

- Account Balance Holds Alerts – Many students enroll in classes even without having a solid plan of how they will afford their tuition. Often those without a solid financial plan aimed at paying off tuition bills will see balance holds on their accounts.
- Attendance (on-line and on-ground) alert – A student has not attended class. This one is difficult to implement since many instructors do not track attendance for their courses.
- Mid-Semester Grade Alerts – Mid-Semester Grades are excellent predictors of how students are doing in their courses during week seven or eight of the term. Typically, mid-semester warnings alert academic advisors about students' attendance, grades, missing assignments and/or participation. Often Mid-Semester alerts serve as a red flag that either student or an advisor need to act to help student with a course. (Institutions should mandate faculty to submit alerts for all grade levels)
- GPA Alerts – The GPA alert identifies students where their GPA is below a certain threshold, often used to identify students going on warning or probation or those struggling in their major
- Holds – Student with holds are often blocked from registering for classes.
- Missing Registration Alerts – Students who have not registered for an upcoming term but also have not petitioned for graduation
- Wrong Registration Alerts – wrong registration alerts notify the advisor if the student has registered for wrong courses

Both professional and peer mentor coaches can engage "at-risk" students in discussions about affordability, academic preparedness, college adjustment, time management, mental

health, life commitments, and work/school balance. Coaches can motivate their mentees, introduce learning tools and strategies, discuss barriers to attendance, recommend child-care providers (if needed), refer students to food pantries, or help apply for government assistance. Coaches can help students work better on teams, learn communication skills and discuss ways to network and develop bonds with peers in the classroom. Coaches and their mentees should stay in frequent contact through the first and second semester and help students deal with difficulties and tragedies while also celebrating major academic milestones. Coaches should be professionally trained and certified using the International Coaching Federation standards.

New Faculty Training and Adjunct Training

New faculty members must understand the role curriculum complexity plays on transfer student success. Faculty should understand the role their course plays in the overarching structure of a degree and whether or not it is mandatory for both first-year and transfer students.

Curriculum complexity index refers to

- Number of Free Elective Credits in a major
- Number of courses sequences with more than five pre-requisites
- The ratio of Major Courses to Major Electives (Flexibility Ratio)
- Number Of Fall/Spring Only Offerings

Both full time and adjunct faculty should be able to make campus referrals to a student who experiences difficulty and assist first-generation college students with campus adjustment. Most importantly, faculty should be aware if they have transfer students in their class. Department chairs should create culturally responsive practices across all classrooms

and increase equity and access when it comes to diversity hiring. Faculty who teach courses with high failure or withdrawal rates should collaborate with advisors to present workshops with strategies on how to best pass their courses. Faculty should also provide timely midsemester alerts and work with advisors or student success coaches to help struggling students.

Implications for Policy

For many years the focus of national conversation has been on access and equity, recent news and congress conversations have shifted to student loan debt. What is even more frightening is the situation for students who accumulate student debt, but do not complete a degree. This is the reality for many transfer students who are first-generation, working their way through school and attempting to complete college credits at night or online. For many transfer students, the basic needs of shelter and food take precedence over studying and homework. The dream of becoming a medical professional or a distinguished musician is often disrupted by the realities of day to day life, and the financial pressures of basic survival. For a long time, a college degree has been a path to a middle class, a way to income stability and benefits, today that path is blocked and the toll to pass is unaffordable for many students from lower-income brackets.

As states and federal institutions review higher education budgets, significant focus needs to be on college affordability. Federal Pell Grants are no longer enough, at least for students who are pursuing a four-year degree at a state institution. Pell is a small fraction of a total cost and often does not even cover full tuition charges and fees. Pell can be seen as a down payment on a house one still cannot afford. Unmet need or account balance at the start of the first semester is often a key indicator of persistence. Students who decide to pay down

their balances by working, risk academic failure if a proper time management plan is not implemented during the first semester. As of right now, there are no caps on how much students can borrow from private lenders even though such caps exist on home and other loans. Federal policies must address issues of student debt, predatory lending and put caps on how much students can borrow.

State policies should focus on programs like TRIO and ASAP, and policy like statewide curricular pathways and course numbering. The role of community colleges is becoming much more important in the educational landscape, as many students turn to two years institutions with the hope of saving money or to develop relevant skills for the workplace. It is critical for states and institutional leaders to facilitate relationships and agreements between two-year colleges and four-year public institutions to help streamline transfer processes and enhance transfer programs.

For admissions and recruitment officers, the goals of equity and access will be more challenging to achieve with time, as low-income students will be priced out of higher education. Those who do decide to enroll will face significant challenges without additional scaffolding of federal, state, and institutional supports. Institutions must develop departmental and admissions policy with the "Attrition Loop" in mind. They must mandate financial entrance interviews at the point of admissions and deposit.

Deans and department chairs must carefully plan out curriculum and consider transfer students in the process. When planning electives and major requirements, department heads who seek higher transfer enrollments need to make sure that the degree has a sufficient number of free elective credits that can be used to bring in transfer coursework. Each department should also engage in self-study and focus specifically on the transfer population.

By reexamining curriculum, understanding DFW Transfer Rates and reviewing course equivalencies regularly, faculty can enhance the transfer student experience from within the classroom.

Implications for Future Research

This study focused on the big picture of transfer persistence at a large urban institution in an effort to understand major reasons that lead to drop-out. As with any large dataset, the research conducted focused on major patterns and global themes that existed in the data. Further research on this population will help to understand the nuances and individual stories behind the numbers. Many of the students identified as being "at-risk of dropping out" actually persist and go on to graduate; their stories are extremely valuable as they demonstrate the traits necessary for perseverance and success. With many students in the sample coming from Community College Feeder schools further research is needed to understand the effectiveness of the articulation agreements and dual admissions contracts, and what roles do both play in time to degree completion.

Additional research is needed to understand New Types of Transfer Students (DOATH)

1. Students who enroll via Dual Enrollment and Dual Admissions Pathways *(D)*
2. Online students who transition between face to face, hybrid and face distance learning *(O)*
3. Adults (24+) who return after a significant break from their education and their reasons for finishing the degree *(A)*
4. Traditional college age transfers (17-24) with attendance record at multiple institutions (swirlers) *(T)*
5. Homeschooled High School students who take college level courses in order to satisfy graduation requirements and then use those courses to obtain a baccalaureate degree *(H)*

Follow-up studies with a qualitative focus can integrate focus groups or individual interviews with narrative analyses of the social media student posts (Facebook, Instagram, Twitter) on the university accounts to gain an in depth understanding of the language and the issues that students bring up at different point of their student lifecycle. Interviews with student support personnel both in student and academic affairs can provide an understanding of the experiences of students who were coded as high risk, sought out help, were retained and went on to graduate. Student/Advisor satisfaction survey can shed light on systematic issues that need to be addressed and policies that come into question during individual appointments.

Additional quantitative research can be performed to understand the role of curriculum, course sequencing and scheduling on persistence, role of mid-semester alerts, and the impact of stacked credentialing [[[Certificates]Associate degrees]Bachelor's Degree]. Individual risk models can also be developed for special populations (i.e. honors, athletes, veterans, international students), and students who are undeclared or those who change their major multiple times. Any additional research conducted should be widely shared both via academic journals and using Linked-In, YouTube, and Blogs and disseminated at conferences and webinars run by The National Institute for the Study of Transfer Students.

Research can be shared by publishing quick and easy to understand visuals, like the ones below, using various social media platforms. Research that is only shared via academic journals may not be readily available to academic practitioners or easily understood in the format that it is normally presented. Research infographics and one-page summaries are much simpler to share and put into practice.

Figure 5.1: Example LinkedIn Image 1: Handout - The Attrition Loop

Figure 5.2: Image - Top Predictor of Attrition (Source College Board)

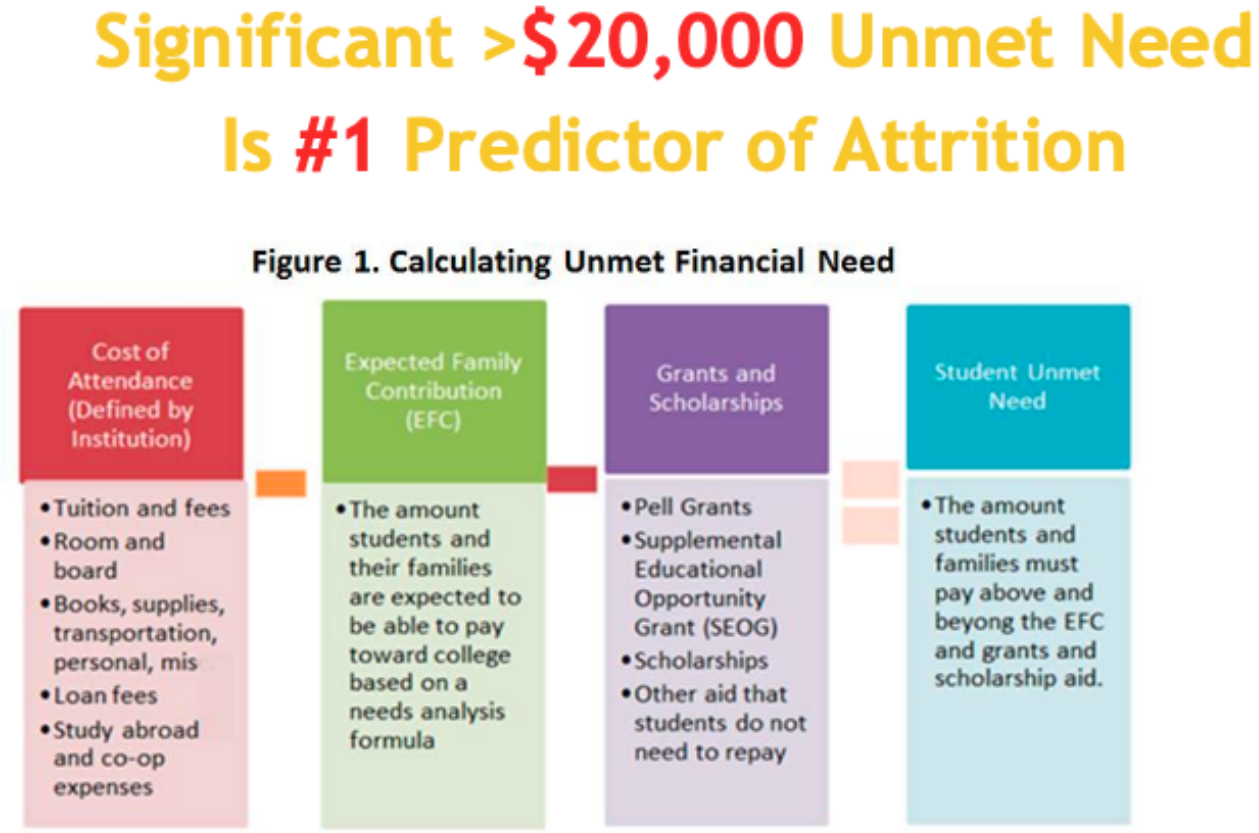


Figure 5.3: Flyer - Financial Aid Interventions

Interventions for Students with High Unmet Need and Financial Concerns

Variables Used to Predict

- EFC | Unmet Need | Offer Amount | Pell Eligibility Y/N | Pell Amount
- Work-study Available || Hired (Y/N)
- Job on campus (Y/N)
- Concerns with financing education
- Parental Income (Self Reported)
- Personal Income
- Balance at the start/end of Term 1

College/University Level Recommendations

Textbook Borrowing (Library/Student Success Center)
Open Source Textbooks/Faculty Grants to Develop open source materials

Pell Matching Grants
Institutional Grants/Scholarships

Part Time Job Fairs during Fall/Spring Terms
Off Campus Work-study Opportunities

Community College Partnerships (Program to program and Core-to-Core Agreements)
Educating of High School Guidance Counselors about pathways:
Dual Admissions/Enrollment/Community College - University Agreements
Parent and student education about loan repayment cost per month after graduation in Community college pathway vs direct enroll four year pathway (in High School)

First Semester Budgeting Workshops
Technology Budgeting Tools (YNAB, I-Grad)

Wraparound Services helping students file for government assistance
Pantry
Professional Wardrobe Closets
Daycare on campus
Laptop Rentals

Policy Recommendations:

- investing in public institutions and need-based aid programs
- improving calculations in the Free Application for Federal Student Aid (FAFSA) process
- modernizing financial aid and economic security policies

Scholarship/Private Funding

- Increased scholarships for low income students
- Increased investment from corporate sponsors in curriculum development and scholarship funding

Figure 5.4: Word Map - Issues Impacting Transfer Students

Conclusion

Transfer students have long been considered a forgotten student subgroup, a special population that needs both institutional support and financial investment. Until recently, there has been limited reporting and data produced about this group. As both researchers and administrators begin to emphasize the importance of not only enrolling this group, but also helping them persist and graduate more and more best practices will begin to emerge. Institutions will need to educate incoming transfer students on various issues ranging from policies to curriculum composition, and furthermore review the structure of their programs to make sure there are enough free electives built in to accommodate any unsatisfied pre-requisites and brought in transfer credits. Faculty and department chairs need to play an active role in curriculum review and make sure that any program/major changes align with institutional mission and goals to accommodate growing transfer population.

REFERENCES

Arentze T.A., Hofman F., Mourik H., Timmermans H.J.P. & Wets, G. (2000) *Using decision tree induction systems for modeling space-time behavior*. *Geographical Analysis,* 32(4) pp.330–350.

Blackwell, M. (2018). *The transfer experience at the institutional, state, and national level for student success: Pre-transfer, transfer, post-transfer, and reverse transfer.* Presentation at Valencia College, Orlando, FL.

Bui, K. V. T. (2002). First-generation college students at a four-year university: Background characteristics, reasons for pursuing higher education, and first-year experiences. *College Student Journal*, *36*(1), 3-12.

Cejda, B. D., & Kaylor, A. J. (2001). Early transfer: A case study of traditional-aged community college students. *Community College Journal of Research & Practice*, *25*(8), 621-638.

Cerezo, A., & McWhirter, B. (2012). A brief intervention designed to improve social awareness and skills to improve Latino college student retention. *College Student Journal*, *46*(4), 867-879.

Chin-Newman, C. S., & Shaw, S. T. (2013). The Anxiety of Change: How New Transfer Students Overcome Challenges. *Journal of College Admission*, (221).

Clark, B. R. (1960). *The" cooling-out" function in higher education.* American journal of Sociology, 65(6), 569-576.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Davies-Vollum, K. S., & Greengrove, C. (2010). Developing a" Gateway" Course to Prepare

Nontraditional Students for Success in Upper-Division Science Courses. *Journal of College Science Teaching*, *40*(1), 28.

Education Commission of the States (2019). *Transfer and Articulation Statewide Common Course Numbering June 2018*. Retrieved from http://ecs.force.com/mbdata/MBquest3RTA?Rep=TR1802

Foote, E. (1997). Community Colleges: General Information and Resources. ERIC Digest.

Frost, L. E. (1994). At-Risk Statutes: Defining Deviance and Suppressing Difference in the Public Schools. *JL & Educ.*, *23*, 123.

Glass, J. C. & Harrington, A. R. (2002). Academic performance of community college transfer students and" native" students at a large state university. *Community College Journal of Research and Practice*, *26*(5), 415-430.

Grawe, N. D. (2018). *Demographics and the demand for higher education*. JHU Press.

Grites, T. J., & Farina, A. (2012). Enhancing transfer student success: The transfer seminar. *Teacher-Scholar: The Journal of the State Comprehensive University*, *4*(1), 2.

Hills, J. R. (1965). Transfer shock: The academic performance of the junior college transfer. *The Journal of Experimental Education*, *33*(3), 201-215.

Hossler, D., Shapiro, D., Dundar, A., Ziskin, M., Chen, J., Zerquera, D., & Torres, V. (2012). Transfer and mobility: A national view of pre-degree student movement in postsecondary institutions. *Signature report*, *2*.

Hsiao, K.P. (1992). First-generation college students. ERIC Digest, pp. 1-5.

Hussar, W. J., & Bailey, T. M. (2017). Projections of Education Statistics to 2025. NCES 2017-019. *National Center for Education Statistics*.

Jackson, D. L. (2013). A balancing act: Impacting and initiating the success of African

American female community college transfer students in STEM into the HBCU environment. *The Journal of Negro Education*, *82*(3), 255-271.

James, C. L. (2006). ACCUPLACER OnLine: Accurate Placement Tool for Developmental Programs? Journal Of Developmental Education, 30(2), 2-8.

Jefferson, G. D., Dougherty, F. C., Steadman, S., & Thomas, T. G. (2013). Transfer Student Retention. In *Proc. ASEE Annu. Conf. Expo* (pp. 23-1266).

Jones, D. A. (2011, December 29). Reforming Federal Student Aid. *Chronicle.com*, Retrieved January 5, 2019, from https://www.chronicle.com/blogs/brainstorm/reforming-federal-student-aid-part-3/9188

Jones, P. R. (2015). Changing the Culture: Using Risk/Need Principles to Move Student Retention Practice From Passive-Reactive to Strategic Intervention.

Joseph, S., & Poisel, M. A. (Eds.). (2018). *Building Transfer Student Pathways for College and Career Success*. Stylus Publishing, LLC.

Joshi, R. (2016, September 9). Accuracy, Precision, Recall & F1 Score: Interpretation of Performance Measures [Web log post]. Retrieved from https://blog.exsilio.com/all/accuracy-precision-recall-f1-score-interpretation-of-performance-measures/

Kuh, G. D. (2003). What we're learning about student engagement from NSSE: Benchmarks for effective educational practices. *Change: The Magazine of Higher Learning*, *35*(2), 24-32.

Kuh, G. D. (2008). Excerpt from high-impact educational practices: What they are, who has access to them, and why they matter. *Association of American Colleges and*

Universities.

Laanan, F. S., & Sanchez, J. R. (1996). New ways of conceptualizing transfer rate definitions. *New directions for community colleges*, *1996*(96), 35-43.

Lee, V., & Frank, K. (1990). Students' Characteristics that Facilitate the Transfer from Two-Year to Four-Year Colleges. *Sociology of Education, 63*(3), 178-193. doi:10.2307/2112836

Liu, C. H., Stevens, C., Wong, S. H., Yasui, M., & Chen, J. A. (2018). The prevalence and predictors of mental health diagnoses and suicide among US college students: Implications for addressing disparities in service use. *Depression and anxiety*.

Liu, X. M., & Murphy, D. (2013). Using mobile apps to entice general education students into technology fields. *Information Systems Education Journal*, *11*(1), 25.

Long, B. T. (2005). The remediation debate: Are we serving the needs of underprepared college students. *National CrossTalk*, *13*(4), 11-12.

Magidson, J. (1994) *The CHAID approach to segmentation modeling: Chi-squared automatic interaction detection*. In R.P. Bagozzi (Ed.) Advanced Methods of Marketing Research. Blackwell, Cambridge, Massachusetts, pp 118–160.

May, R. W., & Casazza S. P. (2012). *Academic Major as a Perceived Stress Indicator: Extending Stress Management Intervention.* College Student Journal, 46(2), 264-273

Mayhew, M. J., Vanderlinden, K., & Kim, E. K. (2010). A multi-level assessment of the impact of orientation programs on student learning. *Research in Higher Education*, *51*(4), 320-345.

McKenzie, L. (2018, September 11). Academics question system for measuring academic performance, flagging potential problems. Retrieved from

https://www.insidehighered.com/news/2018/09/11/academics-question-system-measuring-academic-performance-flagging-potential-problems

Mehta, S. S., Newbold, J. J., & O'Rourke, M. A. (2011). Why do first-generation students fail?. *College Student Journal*, *45*(1), 20-36.

Montague, N. R. (2012). Articulation agreements: No credits left behind. *Issues in Accounting Education*, *27*(1), 281-298.

Mustafa, S. & Compton, P. (2014). *A Comparison of Graduation Rates: Transfers and Continuing Students at 4-year University Main Campuses*. University System of Ohio Institutions.

Narrz'ello, Gary. "At-risk students." *Education and sociology: An encyclopedia* (2014): 49.

National Student Clearinghouse: *First-Year Persistence and Retention. Snapshot Report*, Summer 2018 (pp. 11-12, Rep.). (2018). Herndon, VA:. (ERIC Document Reproduction Service No. ED591975)

Noel-Levitz (2011). *Transfer Students-the Attitudes of Second-Year College Students. Research Report*. Doi: https://bit.ly/2VKrFY1

Reynolds, P. J., Gross, J. K., Millard, B., & Pattengale, J. (2010). *Using longitudinal mixed-methods research to look at undeclared students.* New Directions For Institutional Research, 201053-66.

Sampson Jr., J. P., McClain, M., Musch, E., & Reardon, R. C. (2013). *Variables Affecting Readiness to Benefit From Career Interventions*. Career Development Quarterly, 61(2), 98-109.

Santos Laanan, F. (2007). Studying transfer students: Part II: Dimensions of transfer students' adjustment. *Community College Journal of Research and Practice*, *31*(1), 37-59.

Schlossberg, N. K. (1989). *Improving higher education environments for adults: Responsive programs and services from entry to departure*. Jossey-Bass Inc., Publishers, 350 Sansome St., San Francisco, CA 94104-1310.

Selingo, J. J. "Student success: Building a culture for retention and completion on college campuses." *The Chronicle of Higher Education* (2015).

Simone, S. A. (2014). *Transferability of postsecondary credit following student transfer or coenrollment*. US Department of Education, National Center for Educational Statistics, Institute of Education Sciences.

Sissel, P. A., Hansman, C. A., & Kasworm, C. E. (2001). The politics of neglect: Adult learners in higher education. *New Directions for Adult and Continuing Education*, *2001*(91), 17-28.

Smith, R. R. (1973). The Academic Achievement and Social Mobility of Community College Transfer and Native Students in a Professional Education Program.

St. Clair, K. L. (1993). *Community college transfer effectiveness: Rethinking enhancement efforts.* Community College Review,21(2), 14.

Terenzini, P. T., & Pascarella, E. T. (1998). Studying college students in the 21st century: Meeting new challenges. *The review of higher education*, *21*(2), 151-165.

Thile, E. L., & Matt, G. E. (1995). The ethnic mentor undergraduate program: A brief description and preliminary findings. *Journal of Multicultural Counseling and Development*, *23*(2), 116-126.

Tinto, V. (1975). Dropouts from higher education: a theoretical synthesis of recent research. Review of Educational Research, 45, 89-125.

Tinto, V. (1988). Stages of student departure: Reflections on the longitudinal character of

student leaving. The Journal of Higher Education, 438-455.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). Chicago: University of Chicago Press.

Tinto, V. (1997). Classrooms as communities. The Journal of Higher Education, 68(6), 599.

Tinto, V. (2006). Research and practice of student retention: what next? Journal of College Student Retention: Research, Theory and Practice, 8(1), 1-19.

Tirpak, D. M., & Schlosser, L. Z. (2013). *Evaluating FOCUS-2's Effectiveness in Enhancing First-Year College Students' Social Cognitive Career Development*. Career Development Quarterly, 61(2), 110-123.

Tobolowsky, B. F., & Cox, B. E. (2012). Rationalizing neglect: An institutional response to transfer students. *The Journal of Higher Education*, *83*(3), 389-410.

Townsend, B. K., & Dever, J. T. (1999). What do we know about reverse transfer students? *New directions for community colleges*, *1999*(106), 5-14.

U.S. Equal Employment Opportunity Commission. (2014) Special Report: Diversity in High Tech. https://www.eeoc.gov/eeoc/statistics/reports/hightech/index.cfm. Accessed November 2018.

US Department of Education. (2011). Meeting the nation's 2020 goal: State targets for increasing the number and percentage of college graduates with degrees. Washington, DC.

Walker, K. Y., & Okpala, C. (2017). Exploring community college students' transfer experiences and perceptions and what they believe administration can do to improve their experiences. *The Journal of Continuing Higher Education*, *65*(1), 35-44.

Walker, K. Y., & Okpala, C. (2017). Exploring community college students' transfer

experiences and perceptions and what they believe administration can do to improve their experiences. *The Journal of Continuing Higher Education*, *65*(1), 35-44.

Wang, X. (2009). Baccalaureate attainment and college persistence of community college transfer students at four-year institutions. *Research in Higher Education*, *50*(6), 570-588.

Whiston, S. C, Brecheisen, B. K., & Stephens, J. (2003). *Does treatment modality affect career counseling effectiveness.* Journal of Vocational Behavior, 62, 390-41

Wirth, N. (1978). Program Development by Stepwise Refinement. In Programming Methodology (pp. 321-334). Springer, New York, NY.

Yanovski, A., Sheffer, H., Mattei, A. (2018). *Research Summary: A Closer Look at The Graduate! Philadelphia Program.* Data that Moves Us Project. Presentation at LearnX Conference, presented by Graduate Network, Kansas City, Missouri.

Appendix A: Variables Used in the Study

Person_KEY	Unique student/key identifier
COHORT_CODE	Fall cohort of admission
GENDER_DESC	Gender code description (Individual Attributes)
RACE_CODE	Calculated Race code issued for IPEDS reporting
First_GEN_Grad	Student is considered first generation based on the NSQ Questions about parental education. For students to be considered first-gen neither father nor mother completed a college degree.
First_GEN_Attend	Student is considered first generation based on the NSQ Questions about parental education. For students to be considered first-gen in this category neither father nor mother attempted a college degree.
NSQ Questions: i1 – i82	New Student Questionnaire questions 1-82
RESIDENCY_LD	Long description for the residency code, indicates if the student is considered in or out of state for tuition calculation purposes. Students from out of stated are charged higher tuition rates
ADMISSIONS_COLLEGE_LD	College at the point of admissions
ADMISSIONS_MAJOR_LD	Major at the point of admissions

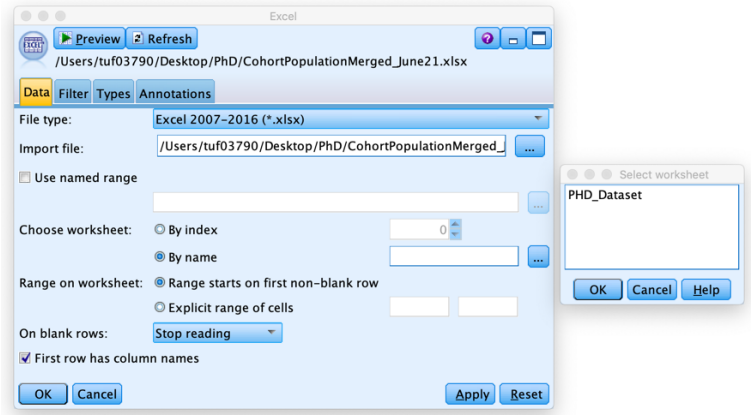
FIRST_REGISTERED_DT	Date of the first registration
Registration_Start_DayDIF	Number of days between first day of classes and original registration day
Term_Code_1	First Semester Code (Fall 2015, Fall 2016, Fall 2017)
Transfer_Hours_UG_AD_T1	Number of transfer hours accepted by the institution
Term_Balance_AD_T1	Account balance at the end of add/drop during student's first semester
College_1_Description_T1	College at the point of Add/Drop for Term 1
Curriculum_1_1 Description_T1	Major at the point of Add/Drop for Term 1
RG1	Registration indicator (Y/N) for Term 1
Hours_Registered_T1	Number of credits registered for Term 1
Term_Undergrad_Hours_Completed_T1	Number of credits completed for Term 1
Term_Undergrad_GPA_T1	Term 1 GPA
Count_of_W_Type_Registrations_T1	Number of "W" – Courses withdrawn – Term 1
Number of Mid-Semester Alerts	Number of Midsemester Alerts for Term 1 (Total number of alerts across all courses) Types of Alerts (Attendance, Participation, Low Grades, Missing Assignments, etc.)
Term_Balance_ET_T1	Account balance at the end of first semester, balances greater $200 may prevent student registration

Term_Code_2	Second semester code (Spring 2016, Spring 2017, Spring 2018)
Term_Balance_AD_T2	Account balance at the end of add/drop during student's second semester
College_1_Description_T2	College at the point of Add/Drop for Term 2
Curriculum 1_1 Description_T2	Major at the point of Add/Drop for Term 2
RG2	Registration indicator (Y/N) for Term 2
Hours_Registered_T2	Number of credits registered for Term 2
Term_Undergrad_Hours_Completed_T2	Number of credits completed for Term 2
Term_Undergrad_GPA_T2	Term 2 GPA
GPA_UG_T2	Overall GPA
Count_of_W_Type_Registrations_T2	Number of "W" – Courses withdrawn – Term 2
Term_Balance_ET_T2	Account balance at the end of second semester, balances greater $200 may prevent student registration
Term_Code_T3	Third semester code (Fall 2016, Fall 2017, Fall 2018)
Term_Balance_AD_T3	Account balance at the end of add/drop during student's second semester
College_1_ Description_T3	College at the point of Add/Drop for Term 3
Curriculum 1_1 Description_T3	Major at the point of Add/Drop for Term 3
RG3 (Target Variable)	Registration indicator (Y/N) for Term 3
Hours_Registered_T3	Number of credits registered for Term 2

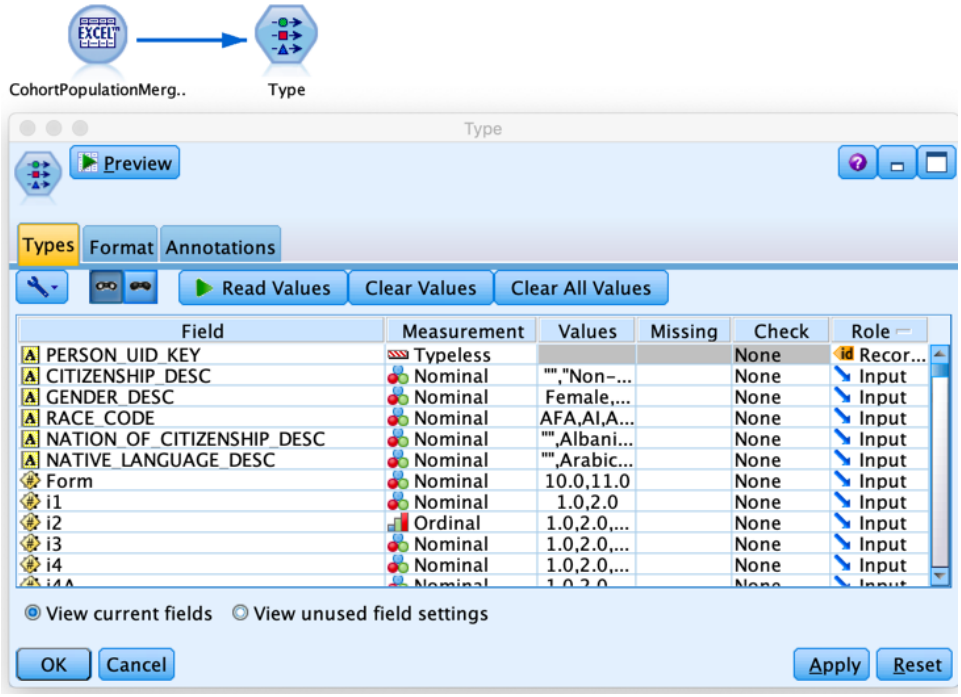
Term_Undergrad_Hours_Completed_T3	Number of credits completed for Term 2
Term_Undergrad_GPA_T3	Term 2 GPA
GPA_UG_T3	Overall GPA
Cumulative_Credits UG_T3	Total number of credits by the end of Term 3
Count_of_W_Type Registrations_T3	Number of "W" – Courses withdrawn – Term 3

Appendix B: Building of the Predictive Models

Data Import - Excel file was brought in for analysis in Modeler; the data were read in with correct the worksheet identified by program.



Data Type Settings - Measurement levels were set for each variable to correct values (Nominal, ordinal, scale), variable roles were defined as either record id, input, or target. The majority of the NSQ variables were recoded from scale to ordinal.

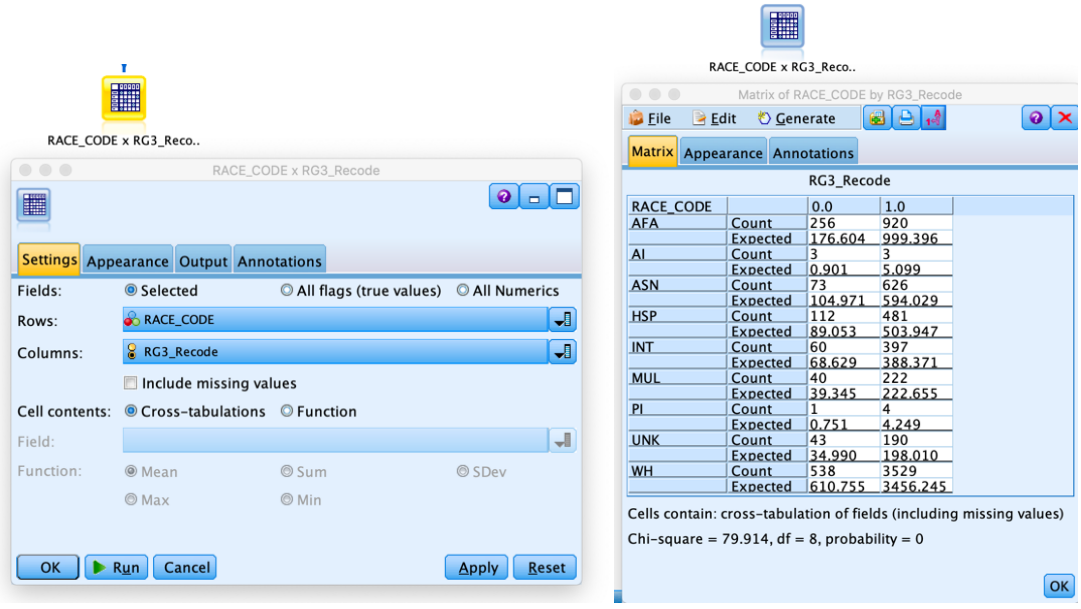


Feature Selection:

In order to identify variables that were used in the modeling process, two tests were performed for all independent variables: chi-square tests for association and independent sample t-Tests. Chi-square was used to determine whether there was an association between persistence and nominal variables, and independent sample t-tests were used to determine whether scale variables (like unmet need, Term 1 GPA), differed between those who persisted and who did not. Only variables that were statistically significant at $p < .05$ were used in the final predictive model.

Part 1

Feature Selection - Matrix/Chi-Square



The matrix browser displays cross-tabulated data between selected input variables (only nominal and ordinal variables) and whether the student was retained (Variable: RG3,

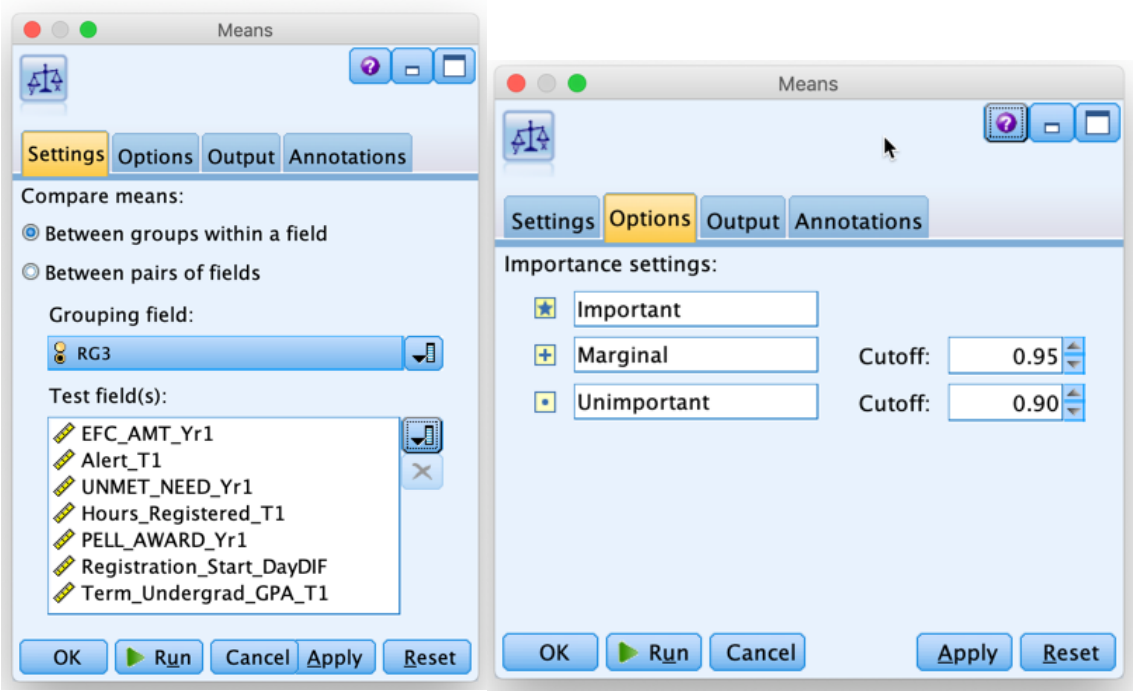
Measurement set to FLAG, and Role set to target). Records with missing values were omitted from the analysis.

Matrix Node serves as a cross-tabulation and calculates Chi-square statistics. The Pearson chi-square statistic is used to determine if the variable will be used in the final model. The Probability (*p*-value) obtained for each input variable indicates the probability that two fields are unrelated based on observed and expected counts. Probability score were recorded for each variable analyzed.

Part 2:

The independent-samples t-test was used to determine if a difference exists between students who were retained and those who were not retained on continuous independent variables (i.e. unmet need, Pell amount, number of mid-semester alerts, etc.).

Feature Selection - Means/Independent Samples T-Test:



A data dictionary was created where p-values, t-statistics, degrees of freedom (df), n, and means were tracked for each variable and the effect size was calculated using excel.

The predictive models were built using the CHAID algorithm. CHAID stands for Chi Square Automatic Interaction Detection and performs multilevel splits on the dataset to build a classification tree. In the tree layout there are parent and child nodes; the first node (the parent node) is split into multiple child-nodes. CHAID analysis is well suited for large datasets, like the transfer student lifecycle-data, and allows the researcher to observe the segmentation within the dataset. CHAID accounts for misclassification costs and provides an estimate of prediction error risk in the output.

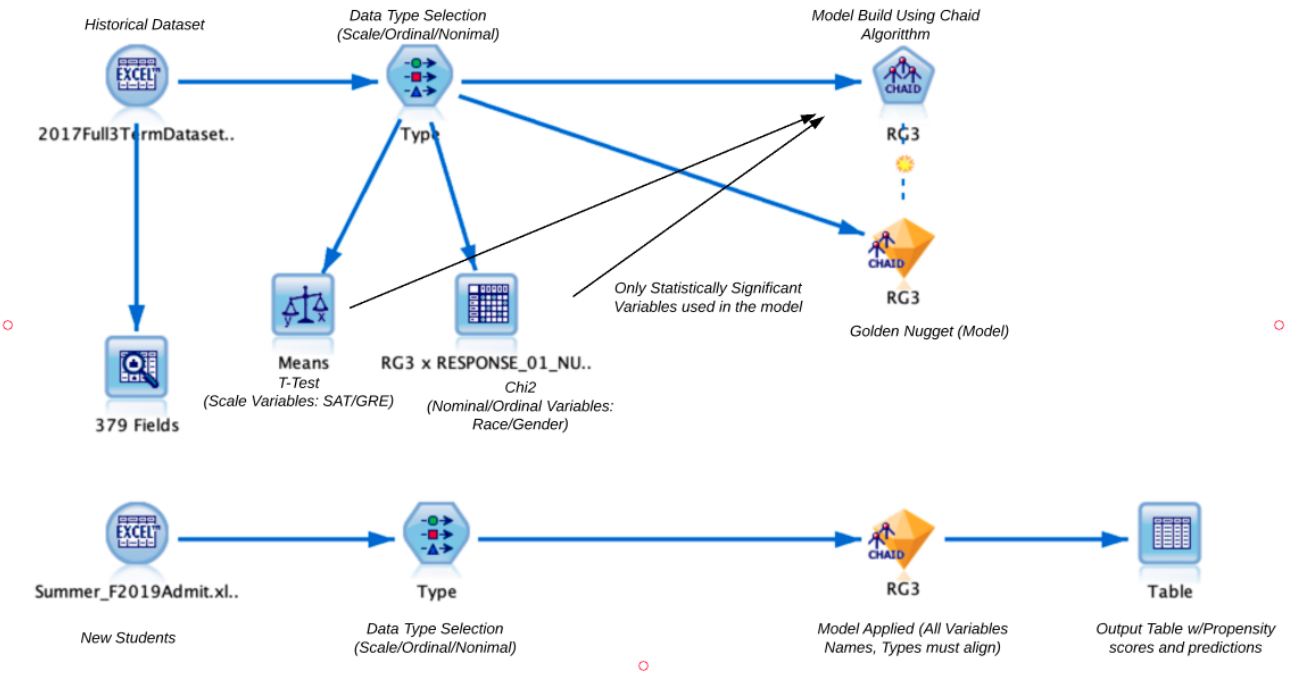
10-fold Cross Validation (C.V) is used to safeguard against overfitting the CHAID Tree. 10-Fold Cross validation is a technique where the dataset is broken up into test set (to evaluate the model) and training set (to train the model); the process is repeated 10 times. Each of the training subsamples is used for validation once, and each fold allows for a separate model learning experience. Cross validation allows all data points to be used for both testing and training the model, instead of simply dividing the original data into the training and testing subsets. After running the Cross-validation experiment, the average of the individual testing for each data fold provides more accurate model prediction results.

The stopping rules for CHAID analysis define the minimum end parameter, which in turn controls which nodes should be split further, and which should be terminal nodes (Child nodes). Based on past experience, I used absolute an value of 70 to define the minimum number of records in a parent branch and 30 for minimum number records in a child branch. Maximum Tree Depth by default was left at 5, which indicated that the root node could only be split

recursively that many times and controlled for tree complexity. The probability for splitting and merging used the Bonferroni method to adjust for significant values and was set to 0.05.

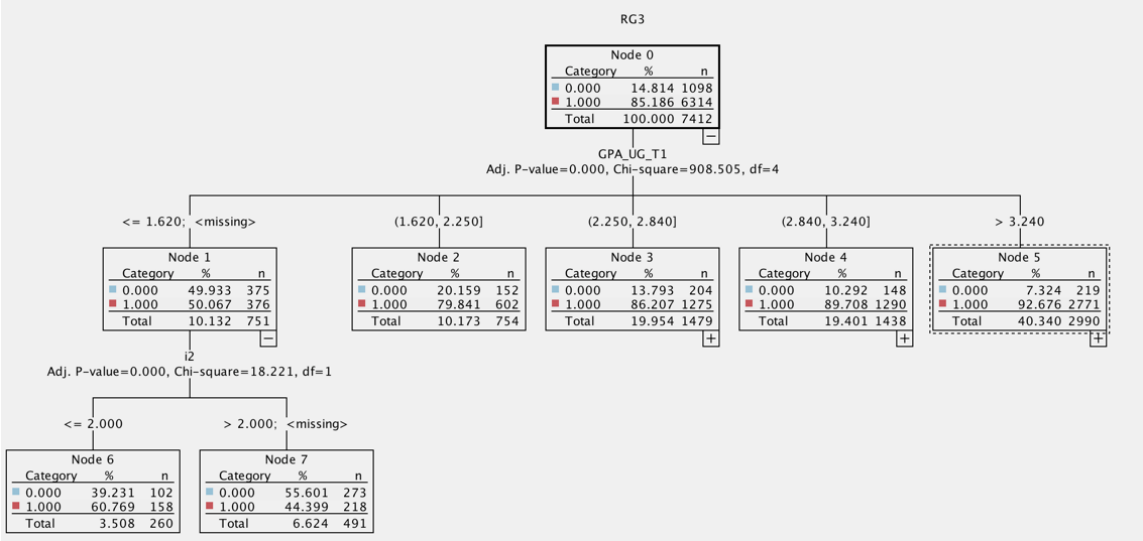
This determined which categories would be split or merged together and governed if the split was found to be statistically significant. Splits are made until no more significant splits are found, while following the roles of maximum tree depth. A Bonferroni adjustment is used to control the type 1 error rate.

Data Flow Overview: SPSS Modeler set-up needed to develop the models



In day-to-day operations of a business unit, the next logical step would be to apply the developed model to the incoming class to calculate predictive risk scores for new students based on the information gathered from historical data. This study did not follow this pattern; however, the findings were shared with the administrators at the institution, and the predictive model was deployed on the incoming transfer cohort.

Figure 0.1 illustrates SPSS Decision Tree output explaining first semester retention of Transfer Students with main predictor variable being Term 1 GPA and secondary predictor #of Hours Students plan to work for money.



The output for the CHAID algorithm provides a dendrogram visualization; a tree structure that is very useful in understanding the relationships of predictors with persistence. Examining the dendrogram, one can classify the various at-risk groups by tracing each of the child (terminal) nodes to the root of the tree. In the model above, the baseline persistence for the population of 7412 was 85.8% with 14.8% of 7412 was 85.8% with 14.8% of the population stopping out by the third semester. The attrition for terminal Node 7 was 44.4% with 55.6% of the group defined by the node (n = 491) stopping out by term three. This subgroup was stopping out at 3.75 times the baseline of the entire group and is considered as high risk. Students in node 7 indicated that they planned to work more than 15 hours a week and had a first semester GPA of 1.62 or below. This simplified version of the model helped highlight a population that would have needed an immediate intervention at the point of add/drop during second semester. Students in nodes with 2.5 times the base rate were considered high-risk.

Appendix C: NSQ

1. Is English your native language?
    a. Yes
    b. No

2. During the school year, on the average, how many hours do you plan to work (for money) per week?
    a. None
    b. 1 to 15 hours
    c. 16 to 20 hours
    d. 21 to 25 hours
    e. More than 25 hours

3. In what kind of residence will you be living during your first semester?
    a. University-owned housing (including residence halls)
    b. Home of parents/relatives
    c. Your own home or apartment
    d. With other family members
    e. Other

4. Where is your permanent home?
    a. City of Philadelphia
    b. Suburban Philadelphia
    c. Pennsylvania, other than suburban Philadelphia
    d. United States, other than Pennsylvania
    e. Other country

5. What is your U.S. Military status?
    a. No military service
    b. Active military service
    c. Veteran
    d. Reserves or ROTC
    e. Other

6. What is your best estimate of the total income of your PARENTAL FAMILY during the past year?

a. Less than $20,000
b. $20,000 to $39,999
c. $40,000 to $59,999
d. $60,000 to $79,999
e. $80,000 to $99,999
f. $100,000 to $124,999
g. $125,000 to $149,999
h. $150,000+

7. What is your best estimate of YOUR OWN total income during the past year?

a. Less than $2,000
b. $2,000 to $3,999
c. $4,000 to $5,999
d. $6,000 to $7,999
e. $8,000 or more

8. Did you apply for financial aid for this year?

a. Yes, and I received an aid package
b. Yes, but I did not receive aid
c. Yes, but I have yet to hear about an aid package
d. No, but I intend to apply
e. No, I will not need financial aid to attend

9. How did the amount of financial aid you received from UNIVERSITY compare to other schools to which you were admitted?

a. Higher than most
b. About the same
c. Lower than most
d. Applied for aid only at UNIVERSITY
e. Did not apply for financial aid

10. Do you have any concern about your ability to finance your college education?

a. None (I am confident that I will have sufficient funds)
b. Some concern (but I will probably have enough funds)
c. Major concern (not sure I will have enough funds to complete college)

11. What was your rating of UNIVERSITY at the time you applied for admission?

a. UNIVERSITY was my first choice
b. UNIVERSITY was my second choice
c. UNIVERSITY was my third or lower choice

12. What is the highest level of formal education completed by your father or guardian 1?
    a. Did not graduate from high school
    b. Graduated from high school
    c. Some college education
    d. Graduated from college (a bachelor's degree)
    e. Postgraduate or professional degree

13. What is the highest level of formal education completed by your mother or guardian 2?
    a. Did not graduate from high school
    b. Graduated from high school
    c. Some college education
    d. Graduated from college (a bachelor's degree)
    e. Postgraduate or professional degree

14. What is the highest level of formal education completed by any of your college age (18 years of age or older) siblings?
    a. Did not graduate from high school
    b. Graduated from high school
    c. Some college education, but did not graduate
    d. Currently enrolled in college
    e. Graduated from college (a bachelor's degree or higher)
    f. I do not have any college age siblings

15. What was your approximate high school average?
    a. A
    b. B+
    c. B
    d. B-
    e. C+ or lower

16. What scholastic average do you expect to obtain in college?
    a. A
    b. B+
    c. B
    d. B-
    e. C+ or lower

17. Scholastically, where did you rank in your high school graduating class?
    a. Top 10%
    b. Top 20%
    c. Top 30%
    d. Top 50%
    e. Not among top 50%

18. In general, how well do you feel that your high school prepared you to do college work?
    a. Very well
    b. Fairly well
    c. Uncertain
    d. Poorly
    e. Very poorly

19. Have you decided on an academic major?
    a. Yes
    b. No

20. Do you consider yourself to be a person who has a disability?
    a. Yes
    b. No
    c. Uncertain

21. During the school year, on the average, how many hours do you plan to study per week?
    a. None
    b. 1 to 15 hours
    c. 16 to 20 hours
    d. 21 to 25 hours
    e. More than 25 hours

During high school (Grades 9-12), how many years did you study each of the following subjects?

22. English
    a. None
    b. One
    c. Two
    d. Three
    e. Four

23. Mathematics
    a. None
    b. One
    c. Two
    d. Three
    e. Four

24. Foreign Language
    a. None
    b. One
    c. Two
    d. Three
    e. Four

25. Natural Sciences
    a. None
    b. One
    c. Two
    d. Three
    e. Four

During high school (grades 9-12) on average what was your grade in the following subjects?
26. English
    a. A
    b. B+
    c. B
    d. B-
    e. C+ or lower

27. Mathematics
    a. A
    b. B+
    c. B
    d. B-
    e. C+ or lower

28. Foreign Language
    a. A
    b. B+
    c. B
    d. B-
    e. C+ or lower

29. Natural Sciences
    a. A
    b. B+
    c. B
    d. B-
    e. C+ or lower

How important were the following in your decision to go to college?

30. I wanted to get a general education.
    a. Very important
    b. Somewhat important
    c. Not important

31. My family wanted me to go.

a. Very important
b. Somewhat important
c. Not important

32. I wanted to get away from home.
    a. Very important
    b. Somewhat important
    c. Not important

33. I wanted to be able to get a better job.
    a. Very important
    b. Somewhat important
    c. Not important

34. I wanted to learn more about things that interest me.
    a. Very important
    b. Somewhat important
    c. Not important

35. I wanted to prepare myself for graduate or professional school.
    a. Very important
    b. Somewhat important
    c. Not important

36. It seemed like a good thing to do as a transition to work.
    a. Very important
    b. Somewhat important
    c. Not important

How important were the following in selecting the UNIVERSITY

37. Personal call from a student
    a. Very important
    b. Somewhat important
    c. Not important
    d. Does not apply / did not attend

38. Social media channels (Facebook, Twitter, Instagram, etc.)
    a. Very important
    b. Somewhat important
    c. Not important
    d. Does not apply / did not attend

39. Brochures or mailings
    a. Very important
    b. Somewhat important
    c. Not important
    d. Does not apply / did not attend

40. E-mail communication
    a. Very important
    b. Somewhat important
    c. Not important
    d. Does not apply / did not attend

41. Web site
    a. Very important
    b. Somewhat important
    c. Not important
    d. Does not apply / did not attend
42. Open house or reception
    a. Very important
    b. Somewhat important
    c. Not important
    d. Does not apply / did not attend

43. Regular campus visit
    a. Very important
    b. Somewhat important
    c. Not important
    d. Does not apply / did not attend

44. High school visit by representative or college fair
    a. Very important
    b. Somewhat important
    c. Not important
    d. Does not apply / did not attend

45. Affordable tuition
    a. Very important positive factor
    b. Somewhat important positive factor
    c. Not a positive factor

46. Student body size
    a. Very important positive factor
    b. Somewhat important positive factor
    c. Not a positive factor

47. Social atmosphere
    a. Very important positive factor
    b. Somewhat important positive factor
    c. Not a positive factor

48. Closeness to home
    a. Very important positive factor
    b. Somewhat important positive factor
    c. Not a positive factor

49. Urban location
    a. Very important positive factor
    b. Somewhat important positive factor
    c. Not a positive factor

50. Variety of academic programs available
    a. Very important positive factor
    b. Somewhat important positive factor
    c. Not a positive factor

51. Variety of co-curricular programs available
    a. Very important positive factor
    b. Somewhat important positive factor
    c. Not a positive factor

52. Reputation of UNIVERSITY
    a. Very important positive factor
    b. Somewhat important positive factor
    c. Not a positive factor

53. Reputation of your specific major
    a. Very important positive factor
    b. Somewhat important positive factor
    c. Not a positive factor

54. Advice and experience of family
    a. Very important positive factor
    b. Somewhat important positive factor
    c. Not a positive factor

55. Advice and experience of friends
    a. Very important positive factor
    b. Somewhat important positive factor
    c. Not a positive factor

56. Meeting students with backgrounds and interests similar to yours
    a. Very important positive factor
    b. Somewhat important positive factor
    c. Not a positive factor

57. Meeting students with backgrounds and interests different from yours
    a. Very important positive factor
    b. Somewhat important positive factor
    c. Not a positive factor

58 Availability of financial aid
    a. Very important positive factor
    b. Somewhat important positive factor
    c. Not a positive factor

What is the chance you will do the following:

59. Change your major field of study
    a. Very good chance
    b. Some chance
    c. Very little
    d. No chance

60. Be a student leader
    a. Very good chance
    b. Some chance
    c. Very little chance
    d. No chance

61. Work full time while attending college
    a. Very good chance
    b. Some chance
    c. Very little chance
    d. No chance

62. Join a social organization or club
    a. Very good chance
    b. Some chance
    c. Very little chance
    d. No chance

63. Play varsity / intercollegiate athletics
    a. Very good chance
    b. Some chance
    c. Very little chance
    d. No chance

64. Need more than four years to complete degree requirements
    a. Very good chance
    b. Some chance
    c. Very little chance
    d. No chance

65. Make close friends
    a. Very good chance
    b. Some chance
    c. Very little chance
    d. No chance

66. Work with a professor on a research project
    a. Very good chance
    b. Some chance
    c. Very little chance
    d. No chance

67. Receive encouragement from family while you're in college
    a. Very good chance
    b. Some chance
    c. Very little chance
    d. No chance

68. Get tutoring help in specific courses
    a. Very good chance
    b. Some chance
    c. Very little chance
    d. No chance

69. Transfer to another college before graduating
    a. Very good chance
    b. Some chance
    c. Very little chance
    d. No chance

70. Take an on-line course
    a. Very good chance
    b. Some chance
    c. Very little chance
    d. No chance

71. Find a job after college in your chosen field of study
    a. Very good chance
    b. Some chance
    c. Very little chance
    d. No chance

72. Participate in volunteer or community service work
    a. Very good chance
    b. Some chance
    c. Very little chance
    d. No chance

73. Study abroad/away
    a. Very good chance
    b. Some chance
    c. Very little chance
    d. No chance

74. I want to live and work in Philadelphia after graduation.
    a. Definitely agree
    b. Somewhat agree
    c. Neither agree or disagree
    d. Somewhat disagree
    e. Definitely disagree

75. Most of my teachers considered me one of the harder workers in their class.
    a. Definitely agree
    b. Somewhat agree
    c. Neither agree or disagree
    d. Somewhat disagree
    e. Definitely disagree

76. I find it difficult to keep a plan of action in my school work.
    a. Definitely agree
    b. Somewhat agree
    c. Neither agree or disagree
    d. Somewhat disagree
    e. Definitely disagree

77. I enjoy studying and reading about things on which I am working.
    a. Definitely agree
    b. Somewhat agree 2075
    c. Neither agree or disagree
    d. Somewhat disagree
    e. Definitely disagree

78. I know how to manage my time well.
    a. Definitely agree
    b. Somewhat agree
    c. Neither agree or disagree
    d. Somewhat disagree
    e. Definitely disagree

79. I am self-confident.
    a. Definitely agree
    b. Somewhat agree
    c. Neither agree or disagree
    d. Somewhat disagree
    e. Definitely disagree

80. My plans have frequently seemed so full of difficulties that I have had to give them up.
    a. Definitely agree
    b. Somewhat agree
    c. Neither agree or disagree
    d. Somewhat disagree
    e. Definitely disagree

81. I am organized and have good study habits.
    a. Definitely agree
    b. Somewhat agree
    c. Neither agree or disagree
    d. Somewhat disagree
    e. Definitely disagree

82. I prefer to be independent of others in deciding what I want to do.
    a. Definitely agree
    b. Somewhat agree
    c. Neither agree or disagree
    d. Somewhat disagree
    e. Definitely disagree