DATA DRIVEN TRANSFER STUDENTS SUPPORT ANALYSIS

A Thesis

presented to

the Faculty of California Polytechnic State University,

San Luis Obispo

In Partial Fulfillment

of the Requirements for the Degree

Master of Engineering in Industrial Engineering

by

Kathryn Steidl

June 2021

COMMITTEE MEMBERSHIP

TITLE:	Data Driven Transfer Students Support Analysis
AUTHOR:	Kathryn Steidl
DATE SUBMITTED:	June 2021
COMMITTEE CHAIR:	Dr. Lizabeth (Schlemer) Thompson, Ph.D. Professor of Industrial and Manufacturing Engineering
COMMITTEE MEMBER:	Dr. Tali Freed, Ph.D. Professor of Industrial and Manufacturing Engineering
COMMITTEE MEMBER:	Dr. Fred W. DePiero, Ph.D. Professor of Electrical Engineering

ABSTRACT

Data Driven Transfer Students Support Analysis

Kathryn Steidl

Low income, academically talented, underrepresented students within the Central Coast of California face barriers in transferring and completing their technical degree. In order to meet future work needs and improve the quality of public life, the path for transfer students needs to be more accessible. To improve access to a high-quality engineering education for local students, the ENGAGE grant (Engineering Neighbors: Gaining Access, Growing Engineers -NSF Grant numbers 1834128 and 1834154) was created. This initiative strives to support local transfer students pre-transfer, during transfer, and post transfer by providing additional academic and financial resources. Five years of Cal Poly transfer student data was collected for analysis on the factors impactful on academic success as measured by Cal Poly cumulative undergraduate degree GPA. This analysis was divided between engineering and non-engineering transfer students. Regression models were created for each subset of transfer students to identify the predictive traits of historically successful students. For engineering students, the developed model included the factors of CSU Mentor GPA (the student's application GPA), Extracurricular Activity Points (points awarded based upon the number of extracurricular activities on the application), Father's Education Code (the level of the education achieved by the student's father), Major (the major enrolled in by the student), Ethnicity Code (the ethnicity the student identified as), and the CA Resident Flag (if the student resided in California at the time of application). These factors were responsible for about 29.61% of variation within the undergraduate degree GPA. Students who had obtained a higher CSU

Mentor GPA were predicted to achieve a higher undergraduate degree GPA. Students who stem from primarily underrepresented ethnicities (such as African American/Black preference and Hispanic) and/or were first generation college students were predicted to achieve a lower undergraduate degree GPA within engineering majors. Those who were California residents were predicted more likely to succeed. For non-engineering transfer students, the factors included within the model were CSU Mentor GPA (the student's application GPA), Major (the major enrolled in by the student), Ethnicity Code (the ethnicity the student identifies as), Work Hour Range Code (the number of hours worked per week), Gender Code (the gender the student identified as), and Academic Extracurricular Leadership Points (the number of points awarded for extracurricular leadership activities). These factors were responsible for 33.88% of the variation with the undergraduate degree GPA. Students who obtained a higher CSU Mentor GPA were more likely to achieve a higher undergraduate degree GPA. Non-engineering students who identified within underrepresented ethnicities such as American Indian/Alaska Native and African American/Black Preference were predicted to achieve a lower undergraduate degree GPA. Those who engaged in six to twenty hours of work per week were predicted less likely to succeed. Based upon both models, any future initiatives in support of transfer students should consider that background of students who have historically achieved lower undergraduate degree GPAs.

Several dashboard tools utilizing the statistical program R are presented for future implementation to support the ENGAGE faculty team. These tools include a data overview, numerical variable summaries, categorical variable summaries, variable summary and plots, factor investigation, and regression model creation. These dashboards will be implemented within an interactive data sandbox that will allow users of varying data skill levels to investigate the transfer student data. Thus, through ENGAGE, further analysis of the factors that impact the success of transfer students will be possible within the data sandbox. Then, transfer student programs and resources can be directed to students who would benefit from additional support.

ACKNOWLEDGMENTS

Special thanks to Dr. Lizabeth Thompson, Dr. Fred DePiero, and Dr. Tali Freed for their aid and guidance throughout this experience. Also, thanks to my family and friends for their love and support.

TABLE OF CONTENTS

Page

LIST OF TABLES … ix

LIST OF FIGURES x

CHAPTER

1. INTRODUCTION 1

1.1 Background 2

1.2 Problem Description 5

2. LITERATURE REVIEW 7

2.1 Transfer Students Background 7

2.2 Existing Systems & Studies 9

2.3 Developing Dashboards 12

3. DATA ANALYSIS 15

3.1 Engineering Transfer Students Model 18

3.2 Non-Engineering Transfer Students Model 30

4. PROPOSED DASHBOARD TOOLS 42

5. FUTURE DIRECTIONS 56

6. CONCLUSIONS 59

REFERENCES 64

APPENDICES

A. Sample Data Variables 67

B. Major Acronym Definitions 68

C. Engineering Transfer Students Undergraduate GPA Analysis 69

D. Non-Engineering Transfer Students Undergraduate GPA Analysis 72

LIST OF TABLES

Table Page

1. Existing Programs at AHC, Cuesta, and Cal Poly in 2018 5
2. Data Variables Overview 17
3. Stepwise Alpha Selection 19
4. Engineering Model Coefficients 21
5. CSU Mentor GPA Numerical Summary 22
6. Extracurricular Activity Pts Numerical Summary 23
7. Father's Education Code 24
8. Number of Students per Major 25
9. Number of Students per Ethnicity Code 26
10. Number of Students per CA Resident Flag 28
11. Stepwise Alpha Selection 31
12. Non-Engineering Model Coefficients 33
13. CSU Mentor GPA Numerical Summary 34
14. Number of Students per Ethnicity Code 36
15. Number of Students per Work Hour Range Code 37
16. Number of Students per Gender 38
17. Number of Students per Activity EC Leadership Pts 39
18. Usability Testing User Matrix 57

LIST OF FIGURES

Figure Page

1. Fall 2020 PolyView Institutional Research 3
2. Data Cleaning Process........ 15
3. Regression Model Process........ 16
4. Transformed Model Residual Plots 18
5. Regression Model ANOVA Test 20
6. CSU Mentor GPA vs. Undergraduate GPA........ 22
7. Extracurricular Activity Points vs. Undergraduate GPA 23
8. Father's Education Code Main Effects Plot........ 24
9. Major Main Effects Plot........ 26
10. Ethnicity Code Main Effects Plot 27
11. CA Resident Flag Main Effects Plot 28
12. Transformed Model Residual Plots 30
13. Regression Model ANOVA........ 31
14. CSU Mentor GPA vs. Undergraduate GPA........ 34
15. Major Main Effects Plot........ 35
16. Ethnicity Main Effects Plot........ 37
17. Work Hour Range Code Main Effects Plot........ 38
18. Gender Main Effects Plot 39
19. Activity Extracurricular Leadership Pts Main Effects Plot 40

20. R Code Data Cleaning Process 43

21. Data Overview 44

22. Numerical Summary Tool 44

23. Categorical Summary Tool 45

24. Numerical Variable Summary and Plot 46

25. Categorical Variable Summary and Plot 47

26. Numerical Factor Investigation 49

27. Categorical Factor Investigation 51

28. Regression Model Creation 1 53

29. Regression Model Creation 2 54

30. Implementing the R dashboards 56

Chapter 1

INTRODUCTION

Completing a university degree by starting at a community college and transferring to a four-year institution is a valid path, but some four-year institutions make this path difficult. In addition, students from disadvantaged schools and first-generation college students attend community colleges at a higher rate than white or those from economically well-off families. As universities have become more and more impacted, students that start in community college are faced with increased difficulty in transferring and completing their degrees. This is represented currently in the California central coast area. Allan Hancock and Cuesta are central coast community colleges that serve largely local students. In contrast, California Polytechnic State University in San Luis Obispo does not primarily serve local central coast students including those transferring from Allan Hancock or Cuesta community college. In 2020, 64,580 students applied to Cal Poly (Cal Poly , 2020). Of these applications, 38.4% of first-time freshmen and 19.9% of transfer students were accepted. As Cal Poly has become a highly sought-after engineering university, the likelihood of local community college transfer students being accepted has decreased as the number of qualified applicants has increased. Therefore, local low-income academically talented students have had less of a chance of attending Cal Poly after completing courses at community colleges’ such as Allan Hancock and Cuesta. Improving the success of local transfer students is vital in meeting future workforce needs for technical and skilled workers specifically in STEM fields along with addressing historic institutional inequities.

1.1 Background

Unlike other universities in the CSU system, Cal Poly does not currently accommodate all the local students who want to attend due to a large influx of qualified applicants outside the area. As a result, students that attend Allan Hancock and Cuesta community college face uncertainty in transferring to Cal Poly to complete their technical degrees. Allan Hancock is a community college that primarily serves northern Santa Barbara county with campuses in Lompoc, Santa Maria, Santa Ynez, and Vandenberg Airforce base (Allan Hancock College, 2020). On average, 98% of its 11,500 students are from the local area. In its service area, less than one fourth of the population holds a college degree. Additionally, Allan Hancock is a Hispanic Serving Institution (HSI) with over 55% of students being Latinx. Cuesta is a community college that serves San Luis Obispo county with locations in San Luis Obispo, Paso Robles, and Arroyo Grande. In 2019, Cuesta enrolled over 15,000 students both online and at its locations (Cuesta College, 2020). In contrast to Allan Hancock, over 50% of its students are from outside the service area. It is also an HSI, and over 33% of its students are Latinx. California Polytechnic State University San Luis Obispo is a four-year institution part of the CSU system. Cal Poly had approximately 22,287 students in Fall 2020. The student demographics include that 54.04% of students were white while 18.33% were Latinx (Cal Poly SLO, 2020). In Fall 2020, there were 64,580 applicants. Of those applicants, 4,788 first time students were enrolled along with 1,052 transfer students. 21.7% of those transfer students were enrolled in the College of Engineering (CENG). Additionally, 37% of transfer students were Hispanic/Latino and 38% were white. Most notably, only 7.9%

of enrolled students were from the Central Coast area (San Luis Obispo, Monterey, Santa Barbara) seen in Figure 1.

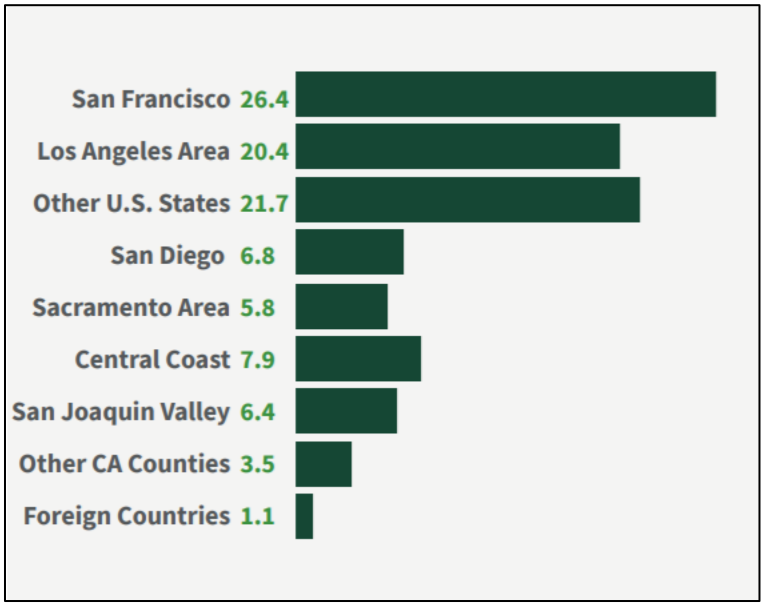


Figure 1: Fall 2020 PolyView Institutional Research

To improve the quality of public life and the ability for upward social mobility, the transfer from community college to a four-year college is imperative. It is apparent that Cal Poly does not service the local area at the same rate as both Allan Hancock and Cuesta. New enrolled students at Cal Poly were only 14% transfer students. In a study done by the Aspen Institute and Columbia University, it was discovered that lower income students were less likely to transfer to a four-year institution and complete a bachelor's degree (Jenkins & Fink, 2016). Furthermore, prior studies have also concluded that community college students from low-income backgrounds are less likely to succeed in transfer programs as compared to students from higher-income backgrounds. Therefore, to improve the quality of life in the Central Coast area it is imperative that support available to transfer students is improved and appropriate.

To enhance transfer student success, the ENGAGE grant was created (Engineering Neighbors: Gaining Access, Growing Engineers -NSF Grant numbers 1834128 and 1834154) to increase access to high quality engineering education for local students in the central coast area that would benefit from additional academic and financial support. ENGAGE is focused on supporting engineering students pre-transfer, during transfer, and post-transfer. In addition, ENGAGE hopes to create sustainable change in Cal Poly transfer practice so that Cal Poly better services the local community. ENGAGE hopes to support student development in five ways; academic, engineering transfer/career path, personal, connection, and professional.

To determine what factors impact transfer student success, transfer student application data has been retrieved from Cal Poly's records for the previous five years for analysis. Isolating the key factors and determining which types of students would benefit from further aid is vital for effectively distributing resources and tailoring programs for increased success. As a result, ENGAGE will ideally improve the retention rate and success of low-income, academically talented students within Cal Poly and the central coast.

1.2 Problem Description

According to a study completed by the Institute of Higher Education Leadership & Policy at CSU Sacramento in 2010, 70% of degree seeking students did not complete any degree after six years (Moore & Shulock, 2010). Additionally, the Latinx share of the working age population in California is projected to grow from 34% to 50% by 2040, with a share of 37.2% as of 2019 (Labor, 2020). Furthermore, only 16% of working-age Latinx adults in California have a college degree (while 50% of white adults have degrees). Considering these statistics and the fact that Latinx students are more likely to begin their education at a community college, it is essential to improve the transfer process to increase the degree completion rate. Cal Poly is working to implement new policies and practices to further support local transfer students in their desire to complete an engineering degree.

Currently, Allan Hancock, Cuesta, and Cal Poly have their own programs and resources available to transfer students. This is represented in Table 1**.** This illustrates the existing systems in place intended to aid students in extra need of support.

Table 1: Existing Programs at AHC, Cuesta, and Cal Poly in 2018

	AHC	Cuesta	Cal Poly
Specialized Programs for Target Population include:			
Mathematics, Science & Engineering Achievement (MESA) Center / Multicultural Engineering (MEP) Advising Program	x		x
LSAMP (Louis Stokes Alliance for Minority Participation in STEM) Program			x
Women's Engineering Program			x
Relevant Student Clubs such as SHPE, MESA, SHPE, NSBE, SWE	x	x	x
TRIO EOP, EOPS, SSS	x	x	x
University Transfer Center / Transfer Advisor	x	x	x
ENGR 301: Engineering Professional Success **(previous CP S-STEM)**			x
CENG Transfer Advising Program (TAP)			x
Faculty-Student Mentoring Program **(previous AHC S-STEM)**	x		x
Cultural Centers (e.g. Multicultural, Gender Equity, and Pride Centers)	x	x	x
Additional Programs include:			
Advising Center / STEM Advising Center / Engineering Advising Center	x	x	x
STEM / MATH Tutoring Lab(s)	x	x	x
Career Services / Professional Dev. Experience & Internship Support	x	x	x
Writing Center	x	x	x
Disabled Student Programs & Services	x	x	x
Veteran's Success Center/Program	x	x	x

Two student cohorts of 50 students each are participating in ENGAGE. The first started in Fall of 2019 at either Allan Hancock or Cuesta. The second cohort began in Fall of 2020. These cohorts will participate fully in ENGAGE, and therefore will be tracked and surveyed to determine their level of involvement and success. To join the ENGAGE program, students go through an application process. This process consists of an application form, personal statement, transcripts, an educational plan, a FAFSA form, and two letters of recommendation. These steps are intended to determine the eligibility of applicants in terms of academic potential and financial need. After selection, students are required to maintain eligibility by meeting specific academic goals and additionally attending ENGAGE events.

The data sample that was analyzed first consists of historical data stemming from the last five years of Cal Poly transfer students. This data includes student application information, transcript data, term data, and course data. Utilizing this data, further analysis was conducted to isolate the primary factors that impact the success of transfer students. This analysis also identified the background of students who would benefit from increased resources and support. Additionally, interactive dashboards were developed to support further analysis and research on the transfer student data.

Chapter 2

LITERATURE REVIEW

In the following sections, research is presented with a focus on the background of transfer students, existing transfer support systems and studies, and dashboard development. These sections are intended to improve comprehension of the issue holistically.

2.1 Transfer Students Background

Understanding the demographics and attributes of transfer students is essential prior to proceeding with data analysis on Cal Poly's transfer students. In a study conducted by the College Board, trends in enrollment, prices, student debt, and completion were analyzed. The authors state that community colleges are imperative in providing education for especially for low-income and first-generation college students (Ma & Baum, 2016). In 2014, 22% of the nation's community college students were Hispanic. However, in California, 43% of community college students were Hispanic. Community college students who are dependent on their family for financial support constituted 40% of the overall population. Within the dependent student community college population, 31% were from the lowest family-income quartile and 36% were first generation college students. In general, the parents of white community college students have a higher educational attainment level (Laanan, 2000). Furthermore, two-thirds of community college students worked in 2011-2012. When compared to their four-year counterparts, community college students are more likely to put in a full work week (40 hours) in addition to class preparation (Carnevale & Smith, 2018). These statistics illustrate some trends in community college students across the nation. Many community college students are from low-income families, hold jobs in addition to school, are part of underrepresented groups, and are first generation college students. These attributes can pose extra challenges for community college students hoping to complete a degree. Therefore, it is essential that future community college policies invest in understanding the backgrounds of their students to remove barriers to education and thus increase the likelihood of community college students succeeding.

Additionally, prior studies demonstrate that there is an overall lack of academic preparation (specifically in mathematics) for community college and transfer students in precollege years as compared to four-year university students (Terenizini, Lattuca, Ro, & Knight, 2014). In relation to pursuing and obtaining an engineering degree, it is essential that students are exposed to adequate mathematical preparation in precollege and community college curriculum. Weaker academic preparation can affect the level of success when students encounter more rigorous work in future college courses. In 1965, John Hills coined the term "transfer shock," relating to the drop in academic performance for transfer students when compared to their community college grades (Hills, 1965). Hills claimed that transfer students faced greater difficulty from high academic rigor and as a result may not graduate in the normal time frame. This theory has become a key motivator for studying transfer students and the transfer process as it has become vital to address why this drop in academic performance occurs.

Furthermore, the motivations behind students selecting a community college versus a four-year university need to be understood. In a study conducted at four major Texas institutions with reputable engineering programs, transfer students were surveyed for their motivations in attending a community college initially (Ogilvie & Knight, 2020). In the face of rising college costs, many students cited financial affordability as a major factor in their choice. Specifically, within the Hispanic/Latinx demographic, financial affordability was even more of a consideration. Next, nonacademic commitments including work or family were illustrated as a significant reason. Also, academic flexibility relating to how easily student's personal and education goals are being met was a significant consideration. Most notably, not obtaining admission to a four-year institution was not a significant factor in selecting community college. In an additional study completed in 1996 with a sample of 10,638, first time community college students were surveyed for their motivations in attending college. The primary reasons included obtaining a better job in the future to make more money as well as learn new things (Laanan, 2000). These primary motivating factors of students beginning in community college imply that students intending to pursue a four-year degree need a clear transfer pathway.

2.2 Existing Systems & Studies

Numerous studies have been performed across the nation in search of improving the success of transfer students. The data utilized in such studies varies from qualitative data (e.g., open-ended survey responses) to quantitative data (e.g., GPA). These studies aim to determine what factors and/or aspects of the transfer process and experience are impactful. The term transfer capital refers to the factors that are involved in a successful transition (Moser, 2020). This theory implies that factors that have a positive impact on the student transfer process contribute to the overall transfer capital. As this capital increases, the ease of transition from a community college to a four-year institution is improved. Examples of factors that could add to the transfer capital include academic counseling, faculty interaction, and a mentor relationship. The following studies are a few examples that express impactful factors and successful system designs.

At Colorado State University, the Vital Connections Transfer Program was implemented in 1993 with the intent of assisting transfer students from Colorado and Wyoming. The program primarily provided application assistance, information on transfer student events, and information on scholarships and advising. After the cohort of transfer students went through the process of transferring, they were surveyed in focus groups of eight students to determine which aspects of the program were effective and which factors were missing in improving their transfer experience. Overall, the students felt that the transfer program was successful in aiding their admission to the university, but there were a few shortcomings that needed to be addressed. Students felt that step-by-step information on the transfer process was lacking. Additionally, they felt a campus tour would have been useful in understanding the environment they would be entering. Furthermore, the transfers expressed their desire to have a peer-mentor available to them for addressing concerns and building a social network in a large university environment.

Next, at Eckard College the Quantitative Excellence in Science and Technology grant was established in 2012 to improve the transfer experience for STEM students (Wetzel & Debure). The study found three effective initiatives in supporting transfer students post transfer, especially in the crucial first semester. First, specialized mentoring was significant is making sure that transfer students understood what courses to take and the path to graduation. Second, the existence of a first-year seminar for transfer students was vital in ensuring that students were well informed and additionally had opportunity to form new connections with other transfer students. Lastly, the close monitoring of student progress by faculty to make sure transfer students are on track for success.

Another study based at the University of Massachusetts Amherst in 2001 analyzed interview data from 372 transfer students (Berger & G.D., 2003). Six different response variables were measured including academic support satisfaction, university satisfaction, social satisfaction, cumulative GPA, sense of community, and sense of academic progress. A few key takeaways were presented post data analysis. First, students that were more prepared and knowledgeable on the transfer process had a higher overall satisfaction with their university experience and even higher academic performance. This illustrates how it is imperative that clear information is available to transfer students on the transfer process and graduation requirements. Furthermore, students that engaged with the faculty and were more involved with the university seemed to have higher satisfaction rates. Additionally, the study also found that white students were more likely to receive higher grades and have higher overall satisfaction. Therefore, it was concluded that additional support would be helpful for transfer students from underrepresented groups. The term “transfer trauma” refers to the experience a transfer student has at university that has different norms or values (Bennett & Okinaka, 1990). This level of alienation at a new university can explain subsequent lack of satisfaction with the university and lower academic performance.

According to a study conducted based on data collected from a Texas institution, “educationally purposeful activities” or EPAs can be predictive of cumulative undergraduate GPA (Fauria & Fuller, 2015). After conducting a survey among both transfer and non-transfer students several factors played a role in GPA success. These factors included receiving faculty academic performance feedback, tutoring other students, participating in class discussions, and working hard. These predictive factors of cumulative undergrad GPA illustrate that steps can be taken within transfer practices to cater to the success of transfer students. Specifically, the authors state that improved faculty-training on the needs of transfer students could be highly beneficial. Additionally, faculty should be expected to challenge their students while also understanding the extra difficulties transfer students may face in a new academic environment.

An additional article looked at existing studies on factors that influence the success of transfer students, particularly Latinx students (Winterer, Froyd, Borrego, Martin, & Foster, 2020). Fifty-nine different studies were analyzed and organized to isolate key factors. The resulting factors include increasing and strengthening the level of interaction between staff and transfer students; encouraging peer student interactions via study groups or living environments; creating an inclusive cultural climate for all students; and improving the availability and quality of student advising, mentoring, or counseling services; and finally implementing further programs focused on supporting academic integration. These key findings summarized from existing studies support the notion that more steps can be taken to further support transfer students stemming from underrepresented groups.

2.3 Developing Dashboards

The ongoing analysis of transfer student data is going to be facilitated with the use of user dashboards. To design and develop effective data dashboards based upon transfer student data, it is imperative to understand dashboard creation techniques and priorities. The success of a dashboard relies heavily on the selection of the data and the selected visualizations (Janes, 2013). The data selected needs to relate to the goals of the dashboard. The visuals should allow the user to understand the meaning of the data with minimal effort. Additional considerations include understanding the needs of the user, the end goal, dashboard type, the structure of information, and minimizing the amount of information (Fard, 2020). It is necessary to develop a dashboard in a way that allows the user to understand useful information quickly that aids them in achieving their goals. Furthermore, presenting the information in an aesthetic, minimalistic, and organized manner can reduce cognitive overload on the user. Prioritizing these design considerations can aid in the process of developing the dashboards.

Once a dashboard is developed, it is highly beneficial to go through thorough user testing and review prior to being implemented. Usability testing refers to the process of ensuring that users are able to complete specific tasks within the created dashboard (Klein, 2018). In this process, users would be given a specific task to complete along with detailed instructions. Then, the user would be observed while utilizing the system, and even potentially asked to "think-aloud" (Richter Lagha, et al., 2020). This technique implies that the user speaks their thoughts as they interact with the system in order to better communicate their thoughts to the test moderator. After completing the task(s), the user can be further interviewed for their thoughts on potential revisions. Additionally, there is a popular questionnaire utilized within usability testing; the System Usability Scale (SUS). The questionnaire involves ten questions and requires the user to respond from Strongly Disagree to Strongly Agree (U.S. General Services Administration Technology Transformation Services, n.d.). A survey such as the SUS could be utilized, or a survey tailored to the system that presents the most important criteria to the system developers. Through the process of usability testing, the developed dashboard can be improved to allow the user to easily utilize the dashboard and achieve their end goals.

Chapter 3

DATA ANALYSIS

To analyze which factors are significant in the success of transfer students, a data sample spanning across five years of students was obtained from Cal Poly. The data consists of primarily application data and undergraduate metrics. For the purpose of analysis, the final (at graduation) cumulative Cal Poly undergraduate degree GPA for transfer students was selected as the response variable and definition of success. It is important to note that GPA does not fully define success for transfer students, and inherently has variation randomized across students as a result of different courses and professors. Due to lack of a superior metric, it was selected to represent transfer student success.

Prior to embarking on statistical analysis, the data required an overall data cleaning process (Figure 2). First, any columns filled with null values were removed. Next, columns with only one response level (e.g., yes for all students) were removed. Looking at the response variable of undergraduate GPA, any students that had not graduated yet were removed. Then, students with null or negative values in the remaining application fields were removed. Additionally, columns with high multicollinearity to one another such as total GPA units and total GPA grade points were removed. The remaining factors available for analysis are listed in Appendix A.

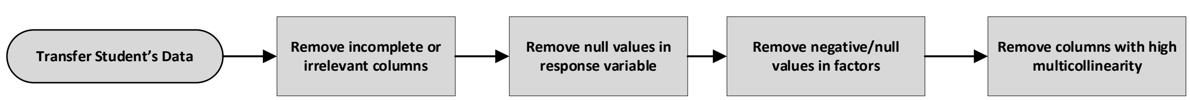


Figure 2: Data Cleaning Process

The following analysis is split into two distinct groups: engineering transfer students and non-engineering transfer students. For each sample, a regression model is explored with the significant factors on undergraduate GPA represented. In the process of creating a regression model, the following steps were followed (Figure 3).

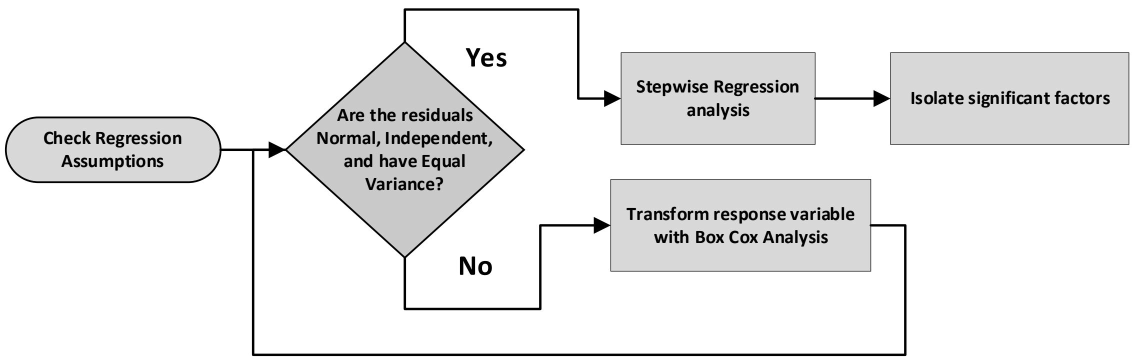


Figure 3: Regression Model Process

First, the regression assumptions were checked for both samples including normality, equal variance, and independence of residuals. Next, a Box-Cox analysis was used to determine the optimal transformation of the response variable. Then, the residual assumptions were reaffirmed prior to completing stepwise regression analysis. Finally, the regression model was developed after isolating the factors that explain a significant amount of variation within undergraduate GPA. The next step in the analysis includes analyzing each continuous factor for a relationship with undergraduate degree GPA. Each categorical factor was explored to determine which levels had a significant main effect on undergraduate GPA. Notably, the main effect was determined based on fitted means, in which the model estimates the level effect given that the design was balanced across levels. The results of this process are detailed in the following sections. Table 2 displays all the variables considered in creating the regression models.

Table 2: Data Variables Overview

Variable Name	Data Type	Brief Definition
Major Code	Categorical	Major Code (EE, IME, BUS, etc.)
CSU Mentor GPA	Continuous	Student application GPA.
Ethnicity Code	Categorical	Ethnicity as identified by student.
EOP Eligible Flag	Categorical	Educational Opportunity Program eligible (Y/N)
Activity Leadership Role Flag	Categorical	Leadership Roles (Y/N)
Extracurricular Hour Range Code	Categorical	Extracurricular Hours Range Code (0-5)
Work Hour Range Code	Categorical	Work Hours Range Code (0-6)
Work Major Related Flag	Categorical	Work Major Related (Y/N)
Gender Code	Categorical	Gender Code (M/F)
Last School Local Flag	Categorical	Local school (Y/N)
Transfer Academic Major Specified Credit Pts	Continuous	Transfer credit pts toward major.
California Resident Flag	Categorical	California Resident (Y/N)
Fathers Education Code	Categorical	Father's education code (0-7)
Mothers Education Code	Categorical	Mother's education code (0-7)
Academic Extracurricular Leadership Pts	Categorical	Leadership Pts Awarded (0, 50, 100, 250)
Academic Extracurricular Major Related Pts	Categorical	Major Related Pts Awarded (0, 10, 50, 100, 150, 250)
Academic Extracurricular Pts	Categorical	Pts awarded based on academic extracurricular activities.
Academic Work Pts	Categorical	Pts awarded based on academic work activities.
Extracurricular Activity Pts	Continuous	Pts awarded based on extracurriculars.
Transfer Academic General Ed Pts	Continuous	Transferrable general education pts.
Transfer Academic IGETC Met Flag	Categorical	Met IGETC Requirements (Y/N)

3.1 Engineering Transfer Students Model

The sample data for engineering transfer students contained 254 different transfer students who completed their undergraduate degree. Prior to analysis, the assumptions for the residuals were analyzed for normality, equal variance, and independence (see Appendix C). Although the residuals appeared to follow a normal shape, after running an Anderson-Darling test, the resulting conclusion was that the residuals do not stem from a normal population (see Appendix C). Next, a Box-Cox analysis was performed to identify a suitable transformation on the response variable (see Appendix C). The analysis did not result in a recommended transformation, but a y^2 transformation was utilized on the undergraduate GPA response variable. The resulting residuals were again analyzed for normality, equal variance, and independence (Figure 4).

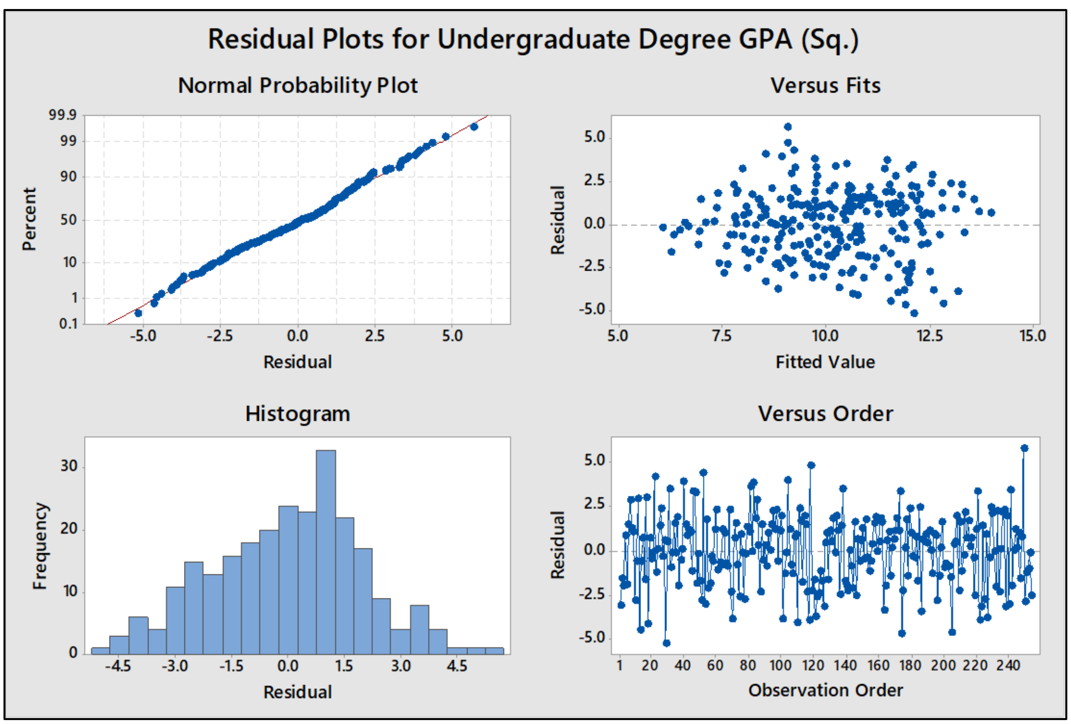


Figure 4: Transformed Model Residual Plots

An Anderson-Darling test of normality was conducted to confirm that the transformed residuals stem from a normal population (see Appendix C). Looking at the Versus Fits plot within Figure 4, there are no concerning shapes that would indicate unequal variance. Lastly, looking at the Versus Order plot, there are no concerning patterns that would indicate lack of independence. Therefore, the three key assumptions of normality, equal variance, and independence were verified prior to further analysis. Next, a regression model was created via Minitab. Stepwise analysis was utilized using a significance alpha level of 0.25 for factors to be entered and removed from the model in order to achieve a relatively high R-squared adjusted (Table 3).

Table 3: Stepwise Alpha Selection

Alpha	**R-squared (adj)**	**# Selected Variables**
0.05	26.86%	3
0.10	26.86%	3
0.15	28.74%	4
0.20	29.03%	5
0.25	29.61%	6
0.30	29.61%	6
0.35	29.84%	7
0.40	29.84%	7

The resulting significant factors can be seen in Figure 5.

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	26	624.34	24.013	5.09	0.000
CSU Mentor GPA	1	173.65	173.652	36.83	0.000
Extracurricular Activity Pts	1	33.97	33.971	7.20	0.008
Major	12	77.10	6.425	1.36	0.185
Ethnicity	5	55.29	11.057	2.34	0.042
Father_s Education Code	6	37.29	6.216	1.32	0.250
CA Resident Flag	1	10.88	10.878	2.31	0.130
Error	227	1070.36	4.715		
Total	253	1694.70			

Figure 5: Regression Model ANOVA Test

The final regression model resulted in an R-squared value of 36.84% and R-squared adjusted value of 29.61%. Thus, about 29.61% of variation within undergraduate GPA in the model stems from the factors of CSU Mentor GPA, Extracurricular Activity Points, Major, Ethnicity, Father’s Education Code, and the CA Resident Flag. The coefficients within the model are depicted in Table 4. The coefficients of each factor and its corresponding levels indicate which model terms most affect the Undergraduate Degree GPA. The highlighted green cells refer to the highest positive coefficient within a factor while the red cells refer to the lowest negative coefficient.

Table 4: Engineering Model Coefficients

Term	Coefficient
Constant	-1.52
CSU Mentor GPA	3.171
Extracurricular Activity Pts	-0.0307
Major	
BMED	0.79
CE	-0.368
CPE	0.307
CSC	0.103
EE	-0.672
ENVE	-0.206
GENE	-0.56
IE	0.67
MATE	0.454
ME	-0.985
MFGE	-0.029
SE	-1.044
Ethnicity	
ASIAN	0
BLACKPRF	-2.27
DECLINE	-0.042
HISPA	-0.305
TWOMORE	0.816
WHITE	0.844
Father's Education Code	
2	0.302
3	0.358
4	0.528
5	0.285
6	0.433
7	1.783
CA Resident Flag	
Yes	0.763
No	0

The significant continuous factors include CSU Mentor GPA and Extracurricular Activity Points. CSU Mentor GPA is the GPA from the student's application. The average CSU Mentor GPA was 3.54 across engineering transfer students (Table 5).

Table 5: CSU Mentor GPA Numerical Summary

Factor	Min	1st Q	Median	Mean	3rd Q	Max	Skewness	Standard Deviation
CSU Mentor GPA	2.65	3.31	3.56	3.54	3.78	4.00	-0.35	0.3087

The distribution of this variable can be seen in Appendix C. The relationship between CSU Mentor GPA and Undergraduate Degree GPA is illustrated in Figure 6. There is a clear positive relationship therefore indicating that a student with a higher CSU mentor GPA is more likely to achieve a higher cumulative undergraduate GPA.

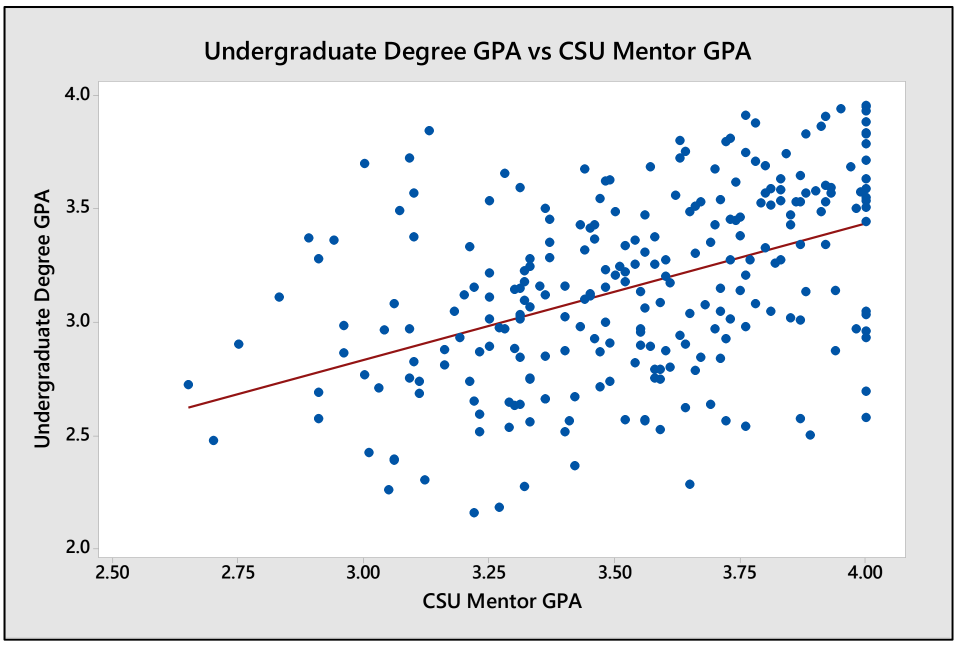


Figure 6: CSU Mentor GPA vs. Undergraduate GPA

Extracurricular Activity Points were also significant within the model. These can be defined as the level of extracurricular activity participation a student engaged in. The average number of extracurricular activity points across engineering students was 25.402 (Table 6).

Table 6: Extracurricular Activity Pts Numerical Summary

Factor	**Min**	**1st Q**	**Median**	**Mean**	**3rd Q**	**Max**	**Skewness**	**Standard Deviation**
Extracurricular Activity Pts	0	15	26	25.4	35	50	0.02	13.27

The distribution can be seen in Appendix C. The relationship between extracurricular activity points and undergraduate GPA is depicted in Figure 7. Notably, there appears to be a slight negative relationship, indicating that engineering students with fewer extracurricular activity points are more likely to obtain a higher undergraduate GPA.

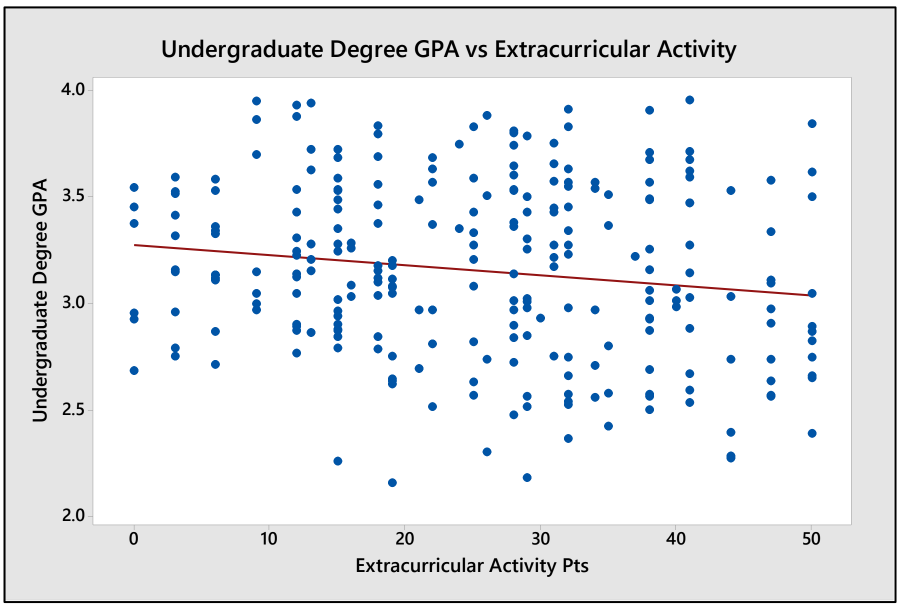


Figure 7: Extracurricular Activity Points vs. Undergraduate GPA

The significant categorical variables within the model include Father's Education level, Major code, Ethnicity code, and the CA Resident Flag. Father's education code refers to the level of education the father of the student completed. The education code, respective level of education, and number of students per level can be seen in Table 7.

Table 7: Father's Education Codes

Code	Level	# of Students
1	No High School	43
2	Some High School	22
3	High School Graduate	38
4	Some College	47
5	2 Year College Graduate	18
6	4 Year College Graduate	66
7	Postgraduate	20

Figure 8 illustrates the average undergraduate GPA per level of father's education. Students with a higher father's education code, specifically at the postgraduate level, were predicted to obtain a higher average undergraduate GPA. Students with a father's education level of high school graduate had the lowest fitted average undergraduate degree GPA.

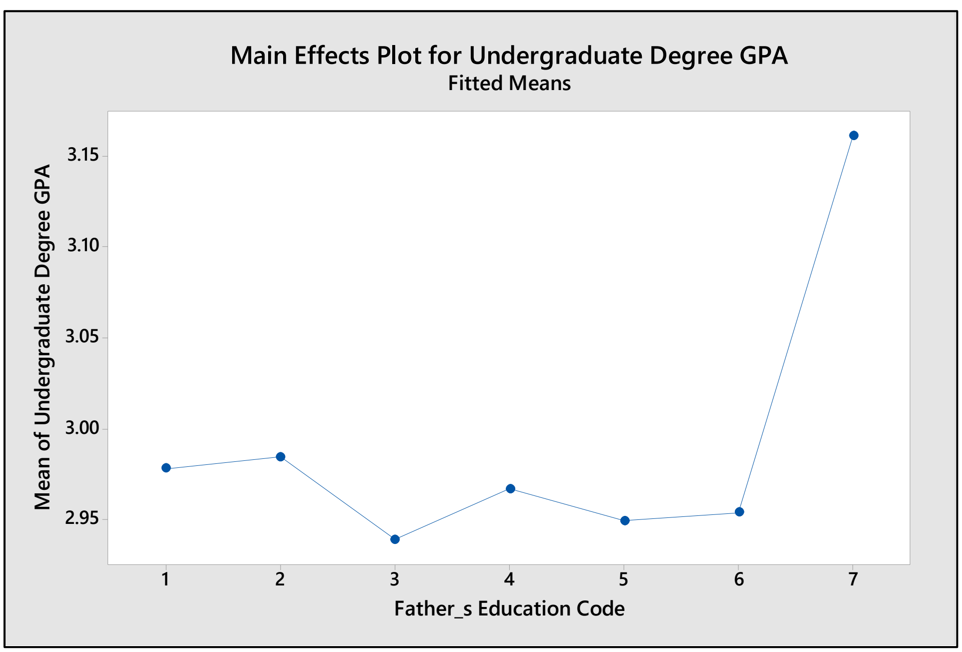


Figure 8: Father's Education Code Main Effects Plot

Major was also significant within the model. The total number of students per engineering major can be seen in Table 8. The decoded acronym of each major can be reviewed in Appendix B.

Table 8: Number of Students per Major

Major	# of Students	Major	# of Students
AERO	25	GENE	4
BMED	19	IE	4
CE	28	MATE	17
CPE	13	ME	50
CSC	30	MFGE	11
EE	30	SE	16
ENVE	7		

Looking at the main effects plot in Figure 9, the majors with the highest fitted average undergraduate GPA are Industrial Engineering and Biomedical Engineering. The majors with the lowest fitted average undergraduate GPA are Mechanical Engineering and Software Engineering. Each engineering major has different levels of rigor and course requirements, and therefore it is expected that final undergraduate GPA is significantly affected by the student’s major.

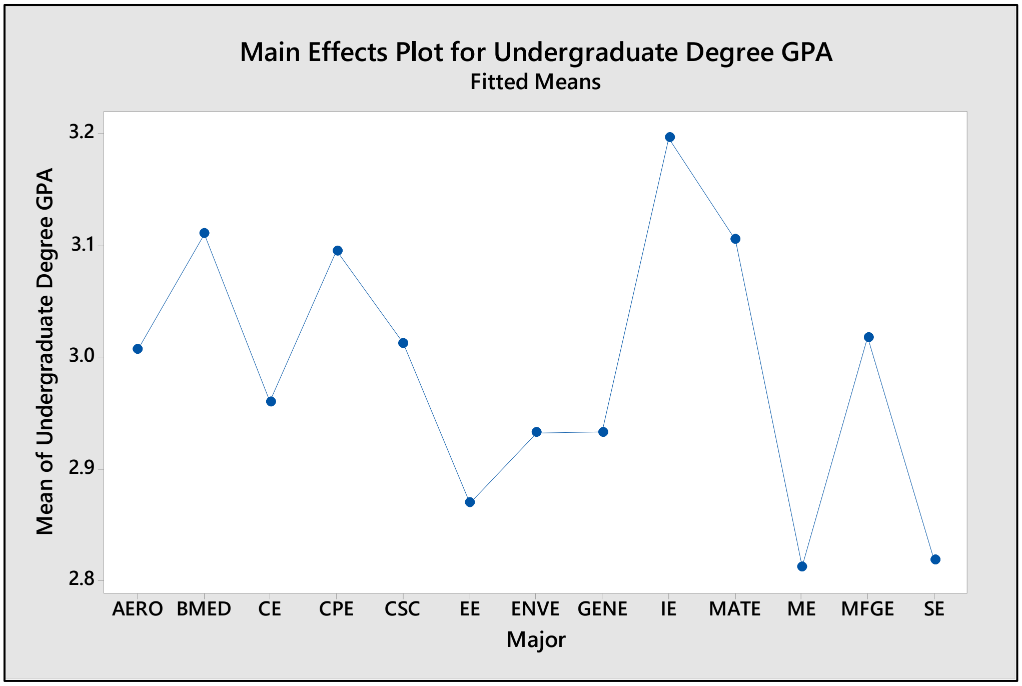


Figure 9: Major Main Effects Plot

Next, the student's Ethnicity code was significant within the model. The total number of students per ethnicity code are illustrated in Table 9.

Table 9: Number of Students per Ethnicity Code

Ethnicity Code	# of Students
ASIAN	37
BLACKPRF	2
DECLINE	8
HISPA	73
TWOMORE	20
WHITE	114

The ethnicity types considered include American Indian/Alaska Native, Asian, Black/African American Preference, Decline to state, Hispanic, Native Hawaiian/Other Pacific Islander, Two or More Ethnicities/Races, and White. The ethnicity codes of Two or More and White appear to have the highest fitted average for undergraduate GPA (Figure 10). In comparison, students who identified as Black/African American

Preference had the lowest fitted average for undergraduate GPA. However, notably there are only two individuals that identify as African American.

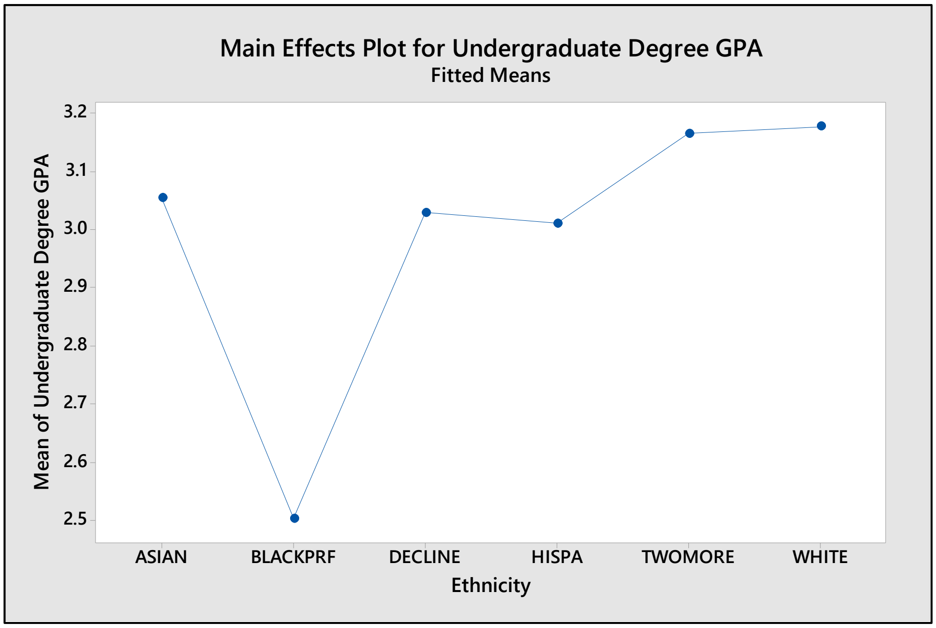


Figure 10: Ethnicity Code Main Effects Plot

The CA Resident Flag factor refers to if the student applicant resides in California. The number of students who were and were not California residents can be seen in Table 10. Notably, most engineering transfer students were California residents at the time of application.

Table 10: Number of Students per CA Resident Flag

CA Resident Flag	# of Students
Yes	232
No	22

Students who were California residents were predicted to achieve a higher undergraduate degree GPA (Figure 11).

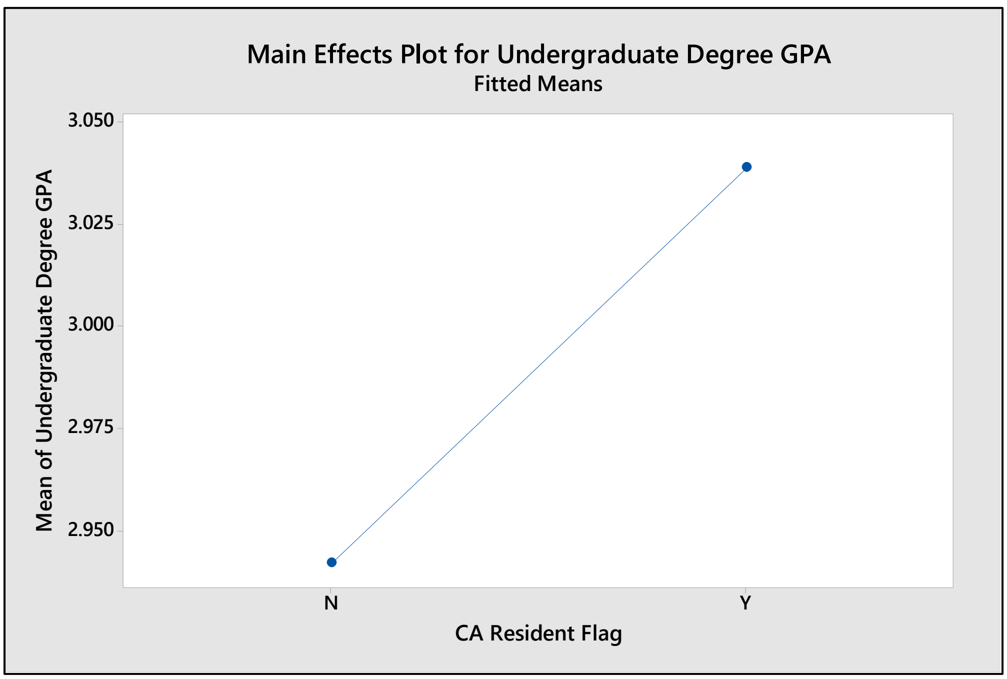


Figure 11: CA Resident Flag Main Effects Plot

In summary, the significant factors within the engineering transfer student model include:

- CSU Mentor GPA
- Extracurricular Activity Points
- Father's Education Code
- Major
- Ethnicity Code
- CA Resident Flag

For each factor, it is necessary to determine which students are more at risk of not obtaining a high undergraduate degree GPA. Engineering transfer students who had a lower CSU Mentor GPA and a higher level of Extracurricular Activity Points were predicted to achieve a lower undergraduate degree GPA overall. Students with a father's education level of high school graduate were less likely to receive a higher undergraduate degree GPA. Students in the majors of Mechanical Engineering and Software Engineering were less likely to obtain a high undergraduate degree GPA. Students who identified within the Black/African American Preference ethnicity code were least likely to complete their undergraduate degree with a high GPA. Lastly, students who did not reside within California at the time of application were less likely to obtain a higher undergraduate degree GPA. Therefore, future efforts on behalf of engineering transfer students should consider these attributes of students who have historically had less success in terms of undergraduate degree GPA.

3.2 Non-Engineering Transfer Students Model

For comparison, the sample data for non-engineering transfer students consisted of 1225 different transfer students who completed their undergraduate degree. Prior to analysis, the assumptions for the residuals were analyzed for normality, equal variance, and independence (See Appendix D). An Anderson-Darling test of normality was conducted but failed to confirm that the transformed residuals stem from a normal population (see Appendix D). Next, a Box-Cox analysis was performed to identify a suitable transformation on the response variable (see Appendix D). The analysis resulted in a recommended transformation of y^2 on the response variable. Thus, a y^2 transformation was utilized on the undergraduate GPA response variable. The resulting residuals were again analyzed for normality, equal variance, and independence (Figure 12).

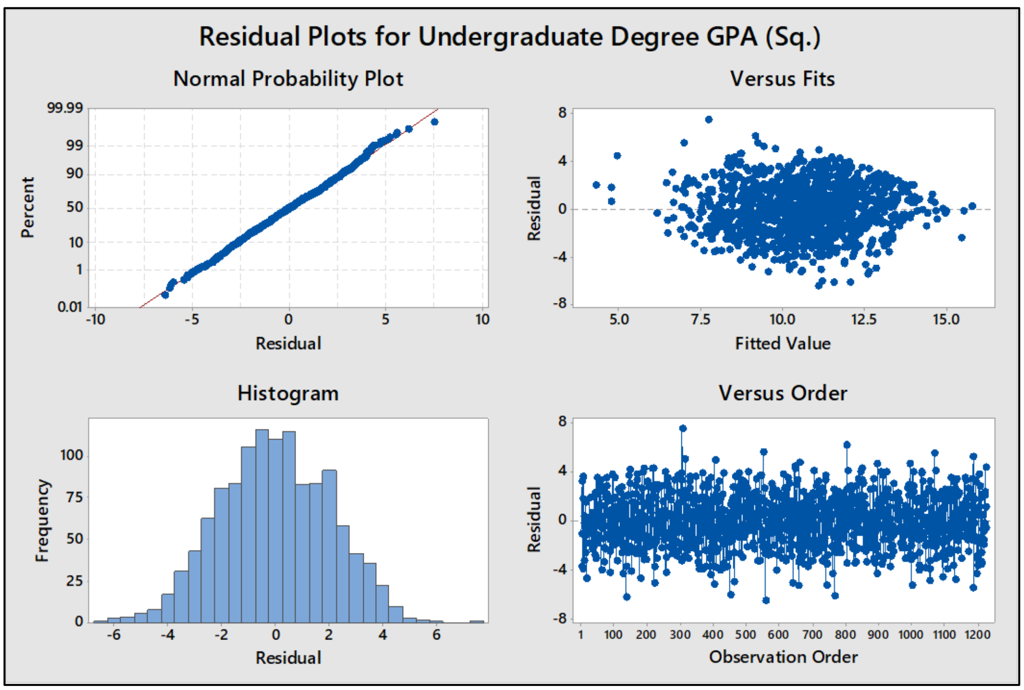


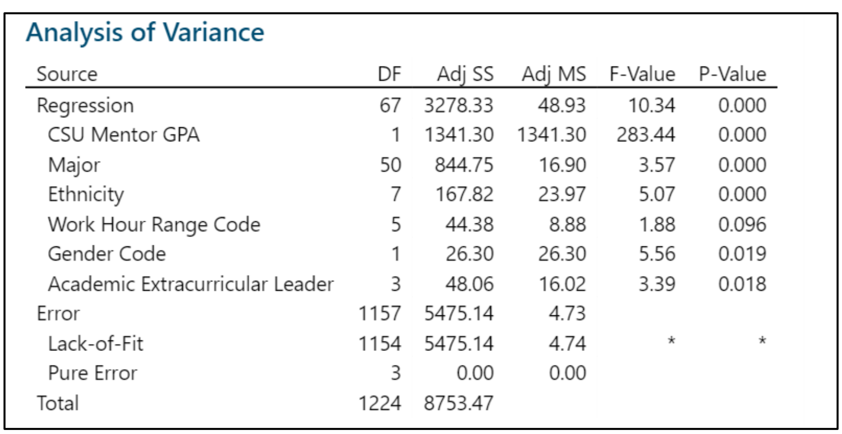
Figure 12: Transformed Model Residual Plots

Another Anderson-Darling test of normality was conducted to confirm normality of the residuals (See Appendix D). Looking at the Versus Fits plot within Figure 12, there are no concerning shapes that would indicate unequal variance. The Versus Order plot reveals no concerning patterns that would indicate lack of independence. Therefore, the three key assumptions of normality, equal variance, and independence were verified. Next, a regression model was created via Minitab. Stepwise analysis was utilized using a significance alpha level of 0.10 for factors to be entered and removed from the model to achieve a relatively high R-squared adjusted and reasonable number of variables (Table 11).

Table 11: Stepwise Alpha Selection

Alpha	**R-squared (adj)**	**# Selected Variables**
0.05	33.58%	5
0.10	33.83%	6
0.15	33.91%	7
0.20	33.91%	7
0.25	34.01%	8
0.30	34.01%	8
0.35	34.21%	11
0.40	34.25%	11

The resulting model and its factors can be seen in Figure 13.


Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	67	3278.33	48.93	10.34	0.000
CSU Mentor GPA	1	1341.30	1341.30	283.44	0.000
Major	50	844.75	16.90	3.57	0.000
Ethnicity	7	167.82	23.97	5.07	0.000
Work Hour Range Code	5	44.38	8.88	1.88	0.096
Gender Code	1	26.30	26.30	5.56	0.019
Academic Extracurricular Leader	3	48.06	16.02	3.39	0.018
Error	1157	5475.14	4.73		
Lack-of-Fit	1154	5475.14	4.74	*	*
Pure Error	3	0.00	0.00		
Total	1224	8753.47			

Figure 13: Regression Model ANOVA

The final regression model resulted in an R-squared value of 37.45% and R-squared adjusted value of 33.83%. Thus, about 33.83% of variation within undergraduate degree GPA in the model stems from factors including CSU Mentor GPA, Major, Ethnicity, Work Hour Range Code, Gender Code, and Academic Extracurricular Leadership Points. The coefficients within the model illustrate which levels of each factor were most impactful on undergraduate degree GPA (Table 12). The highlighted green cells refer to the highest positive coefficient within a factor while the red cells refer to the lowest negative coefficient.

Table 12: Non-Engineering Model Coefficients

Term	Coefficient		Term	Coefficient
Constant	-2.38		**Majors (ct.)**	
CSU Mentor GPA	3.2		MARIN	2.36
Major			MATH	-0.286
AGB	0.113		MCRO	-0.726
AGCOM	1.287		MLL	4.4
AGSC	0.587		MU	1.99
ANGEO	1.224		NUTR	1.018
ARCE	0.54		PHIL	0.881
ARCH	1.949		PHYS	0.493
ART	1.743		POLS	1.512
ASCI	1.298		PSY	1.583
ASM	2.237		REC	2.167
BCHM	-0.335		SOCIO	2.578
BIO	0.163		STAT	1.47
BRAE	1.613		TH	2.94
BUS	0.272		WVIT	0.819
CD	1.219		**Ethnicity**	
CHEM	-0.34		ASIAN	0.459
CM	1.545		BLACKPRF	0.65
COMS	-0.349		DECLINE	0.479
CRP	3.27		HISPA	0.719
DSCI	2.643		PACIF	1.32
ECON	0.522		TWOMORE	1.443
EEASC	1.93		WHITE	1.416
EESS	-0.289		**Work Hours Range Code**	
ENGL	1.654		2	0.592
ENVM	0.911		3	-0.219
ES	1.43		4	-0.071
FDSC	1.152		5	-0.101
FNR	0.869		6	0.252
GRC	1.385			-0.333
HIST	1.392		**Gender**	
IT	0.47		M	-0.333
ITP	2.282		F	0
JOUR	1.851		**Academic Extracurricular Leadership Pts**	
KINE	1.306		50	-0.729
LARC	2.661		100	-0.045
LS	1.467		250	0.615

The sole significant continuous factor was CSU Mentor GPA. This GPA stems directly from the student's application. The average CSU Mentor GPA across non-engineering transfer students was 3.47 (Table 13).

Table 13: CSU Mentor GPA Numerical Summary

Factor	Min	1st Q	Median	Mean	3rd Q	Max	Skewness	Standard Deviation
CSU Mentor GPA	1.72	3.2	3.5	3.47	3.79	4.01	-0.57	0.373

The distribution of this variable can be seen in Appendix D. The relationship between CSU Mentor GPA and undergraduate degree GPA is illustrated in Figure 14. There is a clear positive relationship, therefore indicating that a student with a higher CSU Mentor GPA is more likely to achieve a higher cumulative undergraduate GPA.

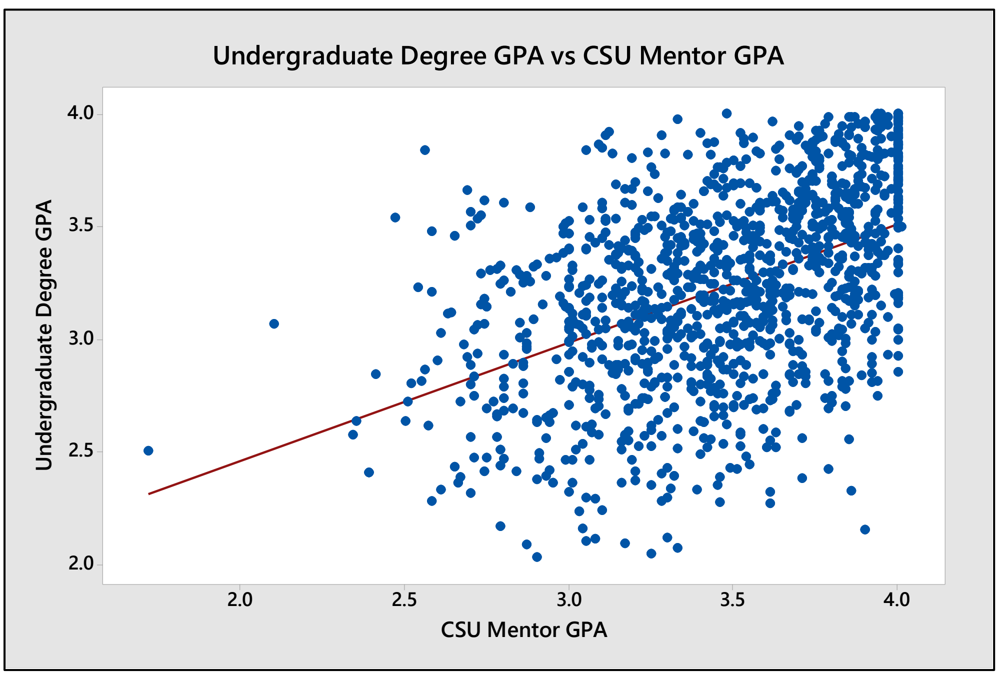


Figure 14: CSU Mentor GPA vs. Undergraduate Degree GPA

The significant categorical variables include Major, Ethnicity code, Work Hour Range Code, Gender code, and Academic Extracurricular Leadership Points. The decoded acronym of each major can be reviewed in Appendix B. The distribution of students across majors can be seen in Appendix D. The majors resulting in the highest fitted average of undergraduate degree GPA include Statistics (STAT) and Modern Languages and Literatures (MLL) (Figure 15). The majors resulting in the lowest fitted average include Environmental Earth and Soil Sciences (EESS) and Microbiology (MCRO). All different majors have different levels of difficulty; thus, it is reasonable that the undergraduate degree GPA of a non-engineering students is significantly affected by the student's major.

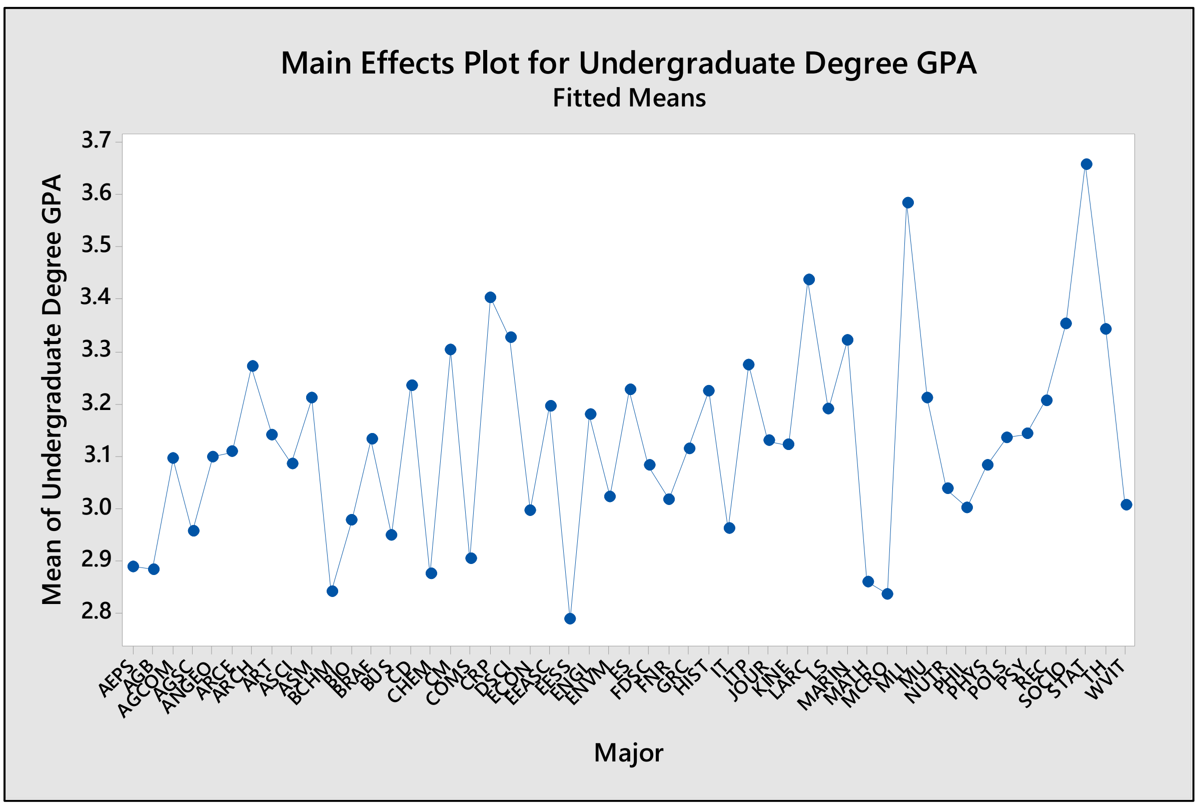


Figure 15: Major Main Effects Plot

The ethnicity code of non-engineering transfer students was also significant within the model. The ethnicity types considered include American Indian/Alaska Native, Asian, Black/African American Preference, Decline to state, Hispanic, Native Hawaiian/Other Pacific Islander, Two or More Ethnicities/Races, and White. The number of students within each ethnicity code can be seen in Table 14.

Table 14: Number of Students per Ethnicity Code

Ethnicity Code	**# of Students**
AMIND	9
ASIAN	93
BLACKPRF	9
DECLINE	31
HISPA	319
PACIF	2
TWOMORE	106
WHITE	656

The ethnicity codes with the highest number of students were White and Hispanic. Looking at the main effects plot in Figure 16, non-engineering transfer students who identified as White and Two or More had the highest predicted average for undergraduate degree GPA. The lowest fitted predicted averages stemmed from American Indian/Alaska Native and Black/African American preference ethnicity codes. Like the engineering transfer student model, those who identified within the White or Two or more ethnicity codes were predicted more likely to achieve a higher undergraduate degree GPA.

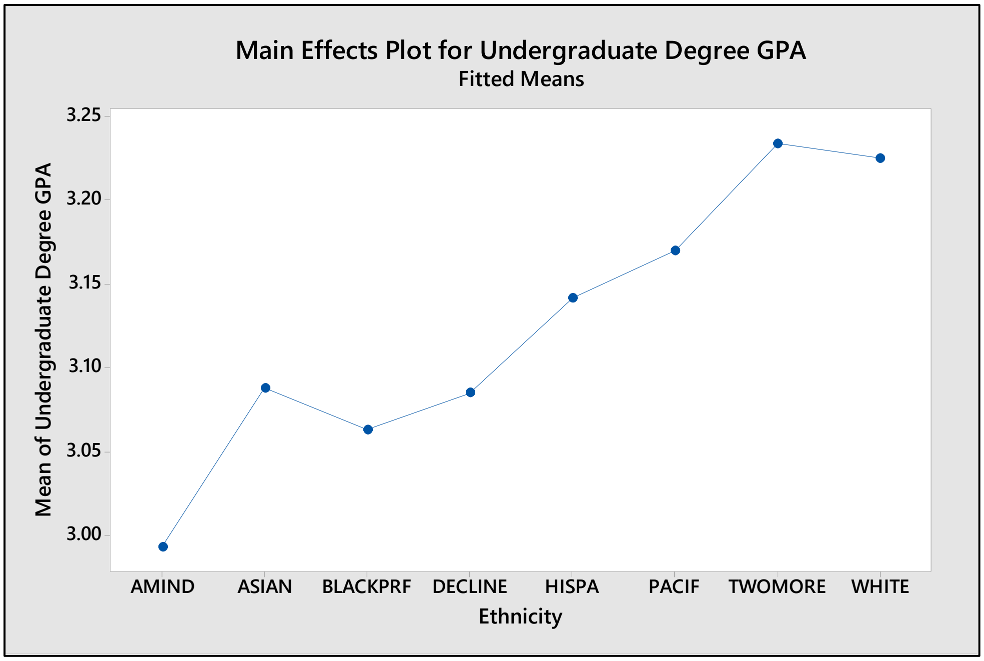


Figure 16: Ethnicity Main Effects Plot

The Work Hour Range Code was also significant within the model. The distribution of students across each level along with a code definition can be seen in Table 15. This variable can be defined as the number of a hours a student worked while attending their prior school.

Table 15: Number of Students per Work Hour Range Code

Work Hour Range Code	Hours	# of Students
1	0	77
2	1-5	34
3	6-10	97
4	11-15	159
5	16-20	277
6	21+	581

Students who worked between one and five hours per week were predicted to achieve the highest undergraduate degree GPA (Figure 17). Those who worked between six to ten, eleven to fifteen, and sixteen to twenty hours were predicted to achieve a lower undergraduate degree GPA.

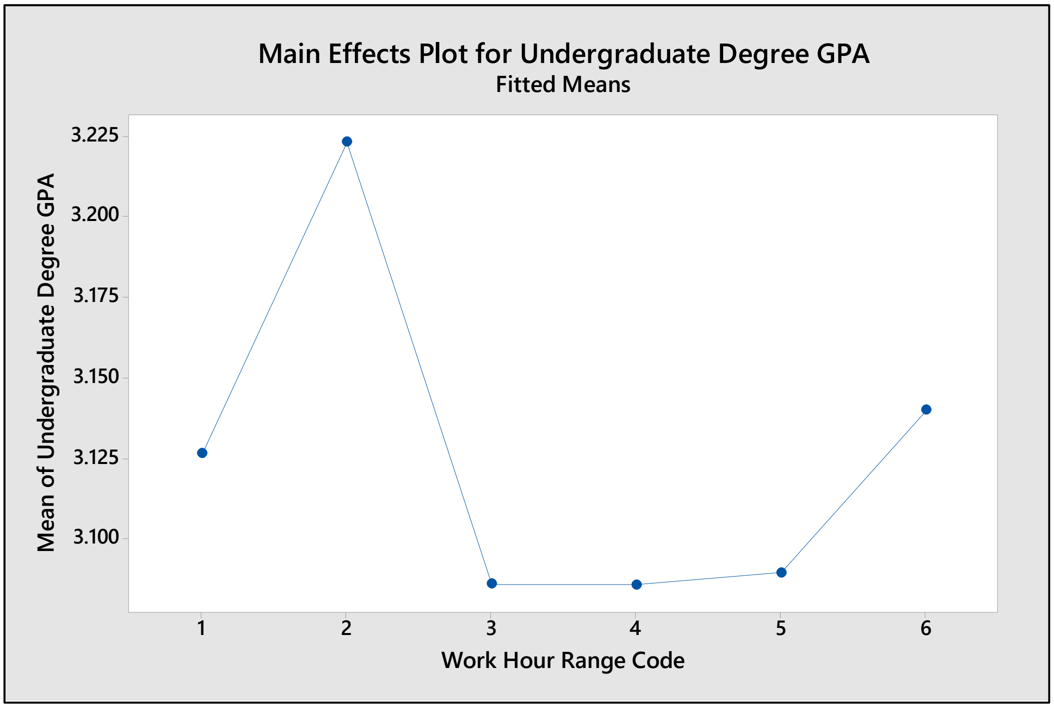


Figure 17: Work Hour Range Code Main Effects Plot

Gender was also significant within the model among non-engineering transfer students. The distribution of students can be seen in Table 16.

Table 16: Number of Students per Gender

Gender Code	# of Students
Female	679
Male	546

Female non-engineering transfer students had a higher fitted average undergraduate degree GPA than male students (Figure 18). Thus, a female non-engineering transfer student is predicted to obtain a slightly higher undergraduate degree GPA.

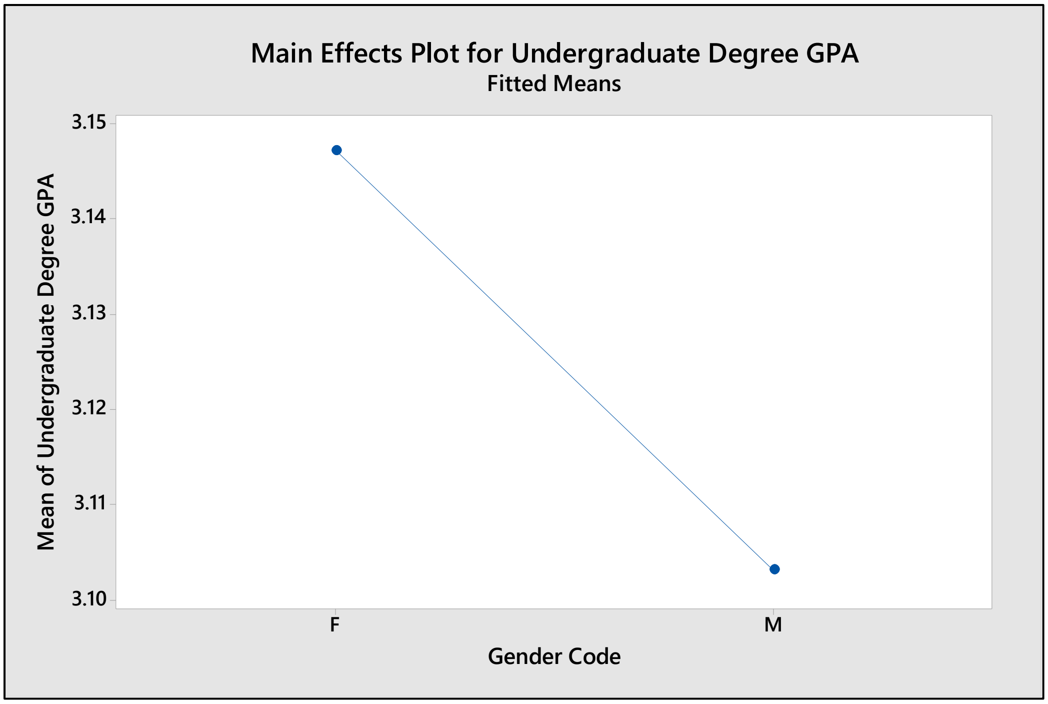


Figure 18: Gender Main Effects Plot

Next, Activity Extracurricular Leadership Points were also significant within the model. This variable can be defined as the number of points awarded based upon extracurricular leadership activities on the student's application. The distribution of non-engineering students can be seen in Table 17.

Table 17: Number of Students per Activity EC Leadership Pts

Academic EC Leadership Pts	# of Students
0	553
50	210
100	424
250	38

Students who received the highest level of 250 points were predicted to achieve a higher undergraduate degree GPA (Figure 19). Therefore, a student with a higher level of extracurricular leadership activity may be more successful.

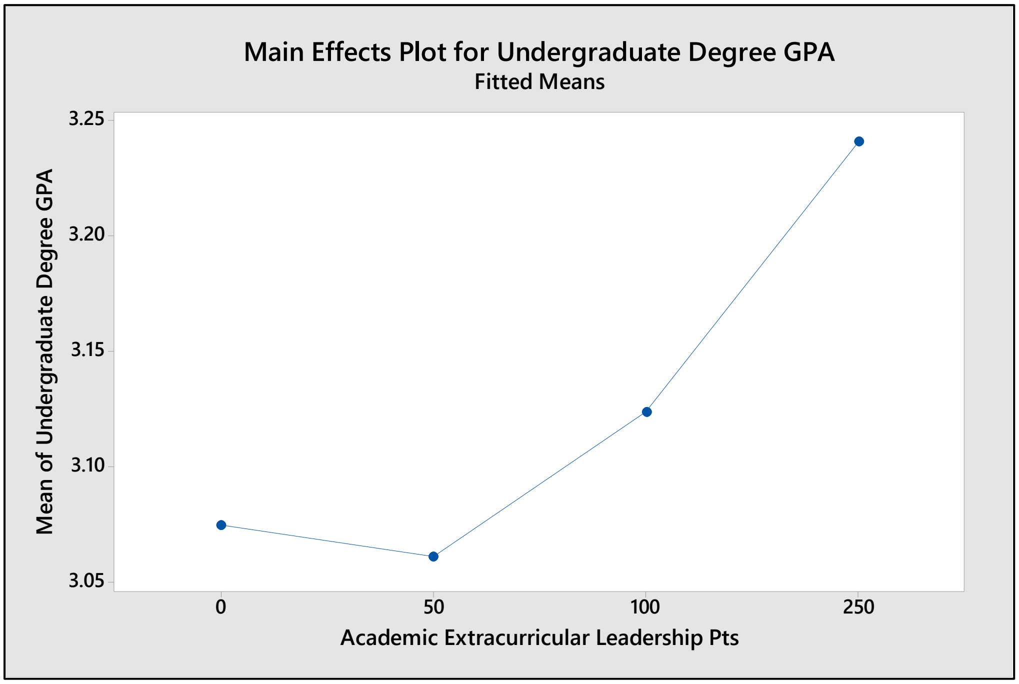


Figure 19: Activity Extracurricular Leadership Pts Main Effects Plot

In summary, the significant factors within the non-engineering transfer student model include:

- CSU Mentor GPA
- Major
- Ethnicity Code
- Work Hour Range Code
- Gender Code
- Activity Extracurricular Leadership Pts

Non-engineering transfer students who had a lower CSU Mentor GPA were predicted less likely to achieve a higher undergraduate degree GPA. Students within the Environmental Earth and Social Sciences and Microbiology majors were least likely to receive high undergraduate degree GPAs. Students who identified within the American

Indian/Alaska Native and African American/Black Preference ethnicity codes had the lowest fitted average undergraduate degree GPA. Students who worked between six and twenty hours generally were predicted to achieve a lower undergraduate degree GPA. Male non-engineering transfer students were predicted less likely to obtain a higher undergraduate degree GPA. Lastly, students who received fewer Activity Extracurricular Leadership Points were predicted less likely to succeed.

In comparison to engineering transfer student model, a few significant trends can be observed. Both models illustrate that students who identify within the White ethnicity code and Two or More are more likely to obtain a higher undergraduate degree GPA. In general, transfer students who obtained a higher CSU Mentor GPA were more likely to succeed. Notably, in the engineering transfer student's model, Father's Education level was significant with postgraduate resulting in the highest predicted undergraduate degree GPA. In contrast, Father's Education level was not significant in the non-engineering transfer student's model. In both models, different majors resulted in different predicted success levels due to differing rigor and courses.

Chapter 4

PROPOSED DASHBOARD TOOLS

In addition to identifying significant factors impacting the success of transfer students, dashboard tools utilizing the statistical program R were created to allow those involved in ENGAGE research to further analyze transfer students. Key considerations in designing the dashboards included understanding the end user, the goal of each dashboard, and presenting the information in a succinct and impactful manner. The purpose of creating accessible dashboards for analyzing transfer students' data is to make data analysis and statistical techniques available to researchers with varying data background levels. This will allow further research and analysis on future transfer students that will aid in identifying how to create sustainable change within Cal Poly's transfer student practice. The following dashboards (Figures 20–29) are proposed for eventual implementation within the ENGAGE research data sandbox. These dashboards include:

- Data Overview
- Numerical & Categorical Variable Summaries
- Variable Summary & Plot
- Factor Investigation
- Regression Model Creation

Each dashboard has a defined functionality and allows the user to analyze and view the data in different ways. The information included within each dashboard is relevant to its innate purpose and is organized in a hierarchal order. Therefore, the user can understand the key information quickly and easily. Users with very minimal statistical understanding could utilize every dashboard apart from the regression model creation tool. The regression model tool requires the user to understand how to interpret regression model outputs, including significant coefficients and R-squared values. The code utilized to clean the raw dataset involved several key steps employed across each dashboard (Figure 20). First, the desired columns were selected and appropriately renamed. Then, any null responses such as "-1" entries were replaced with NA values. Next, each field was converted to either a numerical field or a factor field. This was determined by the number of levels within each variable. Therefore, a field that was continuous was numerical, and a field that contained select categories was a factor. Then, the data frame was separated into all transfer students, engineering transfer students, and non-engineering transfer students. Lastly, depending on the user selected inputs of each dashboard, any null student records were removed. This process within the code is consistent across all the proposed dashboards.

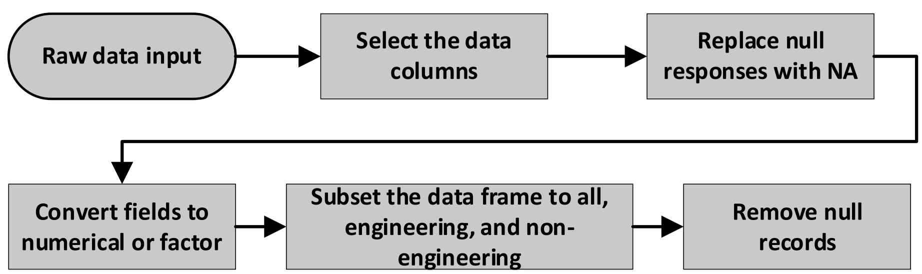


Figure 20: R Code Data Cleaning Process

The first tool allows the user to view, filter, sort, and export the data (Figure 21). It is imperative that the user can have an overview of the data prior to embarking on any further analysis.

Data Overview

Copy CSV Excel

Search:

Index	Major	CSU Mentor GPA	Academic Score	Ethnicity	EOP Flag	Activity Leadership Flag	Extracurricular Hour Range Code	Work Hour Range Code	Work Major Related Flag	Gender Code
			A			All	All			
1	BUS	4	4728	ASIAN	N	Y	3	5	N	M
2	ASCI	3.39	4363	PACIF	N	Y	2	6	Y	F
3	BUS	3.03	4605	HISPA	N	N	5	6	Y	M
4	AGSC	3.17	4147	HISPA	N	N	4	5	Y	M
5	MLL	3.43	4202	TWOMORE	N	N	3	6	Y	F
6	EE	3.52	4812	WHITE	N	Y	6	5	Y	M

Figure 21: Data Overview

Next, the user can view the overall variable summaries for the entire data, or the data separated by engineering and non-engineering (Figures 24-25). This includes summaries for both numerical and categorical variables.

Numerical Summary

CSU Mentor GPA	Academic Score	Transfer Academic Calculated GPA	Transfer Academic Major Specific Credit Pts	Academic Rank	Extracurricular Activity Pts	Transfer Academic General Ed Pts	Transfer Academic GPA Grade Pts	Transfer Academic GPA Pts
Min. :1.720	Min. :2594	Min. :2.480	Min. : 0	Min. : 1.00	Min. : 0.0	Min. : 115	Min. : 47.0	Min. : 934
1st Qu.:3.230	1st Qu.:4271	1st Qu.:3.220	1st Qu.: 935	1st Qu.: 5.00	1st Qu.: 60.0	1st Qu.: 923	1st Qu.:137.0	1st Qu.:1474
Median :3.510	Median :4493	Median :3.530	Median :1250	Median : 14.00	Median :180.0	Median :1250	Median :180.0	Median :1740
Mean :3.477	Mean :4469	Mean :3.483	Mean :1398	Mean : 41.23	Mean :184.6	Mean :1130	Mean :188.1	Mean :1756
3rd Qu.:3.780	3rd Qu.:4702	3rd Qu.:3.800	3rd Qu.:1750	3rd Qu.: 34.00	3rd Qu.:250.0	3rd Qu.:1500	3rd Qu.:228.0	3rd Qu.:1970
Max. :4.010	Max. :5000	Max. :4.000	Max. :2500	Max. :1082.00	Max. :750.0	Max. :1750	Max. :488.0	Max. :2750

Figure 22: Numerical Summary Tool

The numerical summaries include the minimum, maximum, median, mean, 1st quartile, and 3rd quartile. An example of this tool can be seen in Figure 22.

The categorical summary table includes the total number of students per each level within the category. An example of this tool can be seen in Figure 23.

Categorical Summary

Major	Ethnicity	EOP Flag	Activity Leadership Flag	College Code	Extracurricular Hour Range Code	Work Hour Range Code	Work Major Related Flag	Gender Code	Last School Local Flag	CA Resident Flag	Father's Education Code	Mother's Education Code	Academic Extracurricular Leadership Pts	Academic Extracurricular Major Related Pts	Academic Extracurricular Pts	Academic Work Pts	IGETC Met Flag
BUS :270	AMIND : 9	N:1446	N:649	10:313	1:203	1:113	N:631	F:727	N:791	N: 81	1:172	1:171	0 :666	0 :645	0 :237	100 :333	N:817
KINE : 71	ASIAN :130	Y: 37	Y:834	20: 48	2:402	2: 57	Y:852	M:756	Y:692	Y:1402	2:112	2: 99	10 :142	10 :121	60 :175	0 :149	Y:666
PSY : 58	BLACKPRF: 11			40:310	3:309	3:129					3:278	3:263	50 :212	50 :396	30 :136	80 :148	
AGB : 53	DECLINE : 39			48:348	4:181	4:198					4:282	4:242	100:425	100:259	20 :129	150 :147	
CD : 53	HISPA :392			52:254	5:120	5:330					5:132	5: 97	250: 38	150: 4	150 :118	60 :104	
NUTR : 51	PACIF : 2			76:210	6:268	6:656					6:357	6:416		250: 58	3 : 83	15 :100	
ME : 50	TWOMORE :126										7:140	7:195			90 : 72	120 : 74	
SOCIO : 47	WHITE :774										8: 10				100 : 69	25 : 64	
BIO : 46															40 : 68	12 : 52	
JOUR : 33															6 : 67	40 : 52	
ECON : 31															120 : 58	20 : 51	
REC : 31															25 : 49	9 : 39	
ASCI : 30															15 : 48	90 : 39	
CSC : 30															5 : 46	6 : 32	
(Other):629															(Other):128	(Other): 99	

Figure 23: Categorical Summary Tool

The next dashboard allows the user to investigate a specific variable and view the appropriate summaries, plots, and statistical tests. First, the data can be filtered by all transfer students, engineering transfer students, or non-engineering transfer students. Then, the user can select the variable they would like to investigate. The tool then summarizes the variable depending on whether it is numerical or categorical. If the variable is numerical, the plots developed are a histogram and a boxplot. These plots allow the user to visualize the distribution of the variable and additionally identify any concerning outliers. The user can also enter a hypothesized value for a two-sided, greater than, or less than hypothesis test. This allows the user to hypothesize the likelihood of a value within a certain variable. If the variable distribution is normal, a one-sided t-test will be performed. If it is not normal, a one-sided non-parametric Wilcoxon test will be utilized. A numerical variable example using CSU Mentor GPA can be seen in Figure 24.

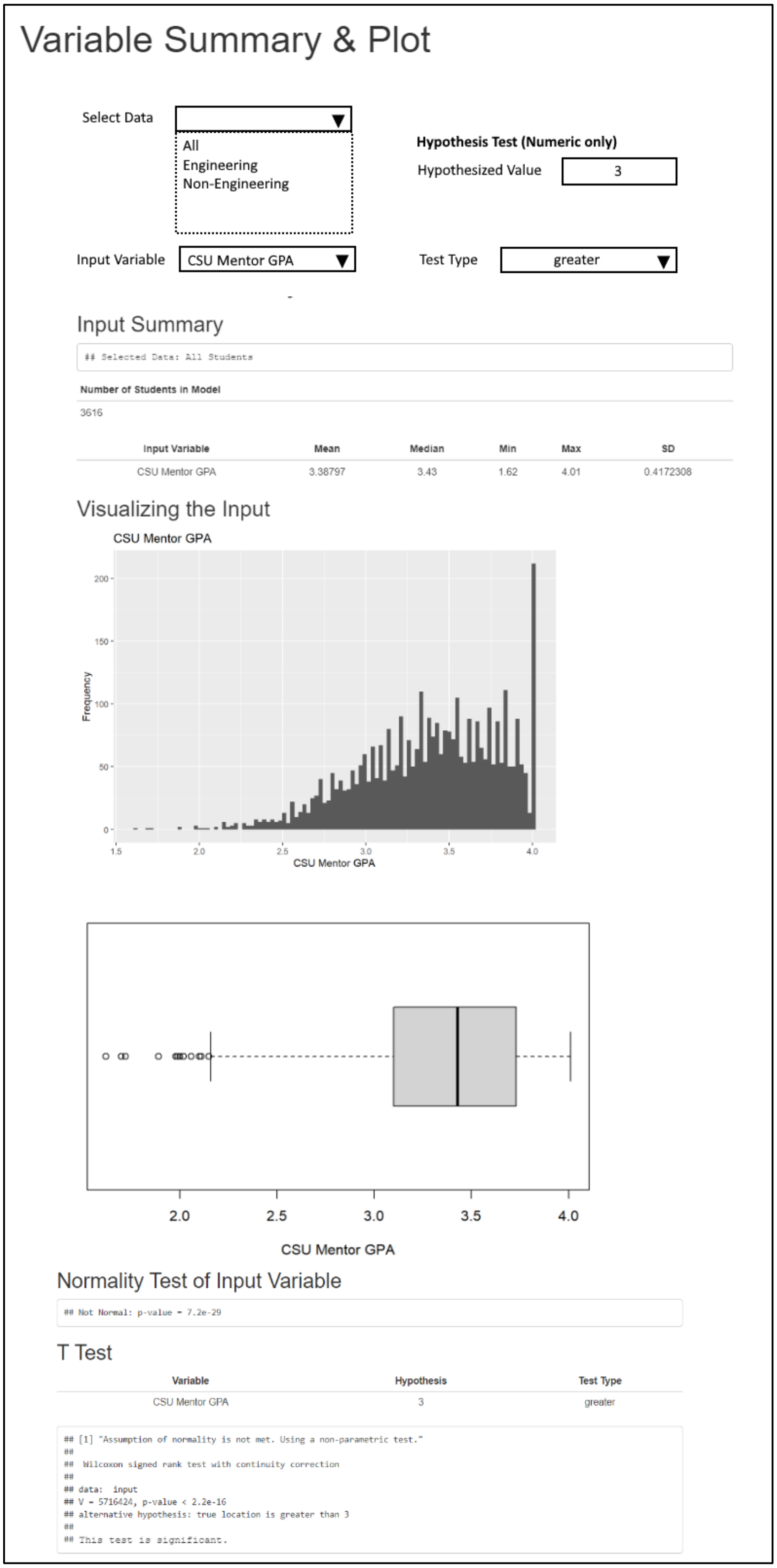
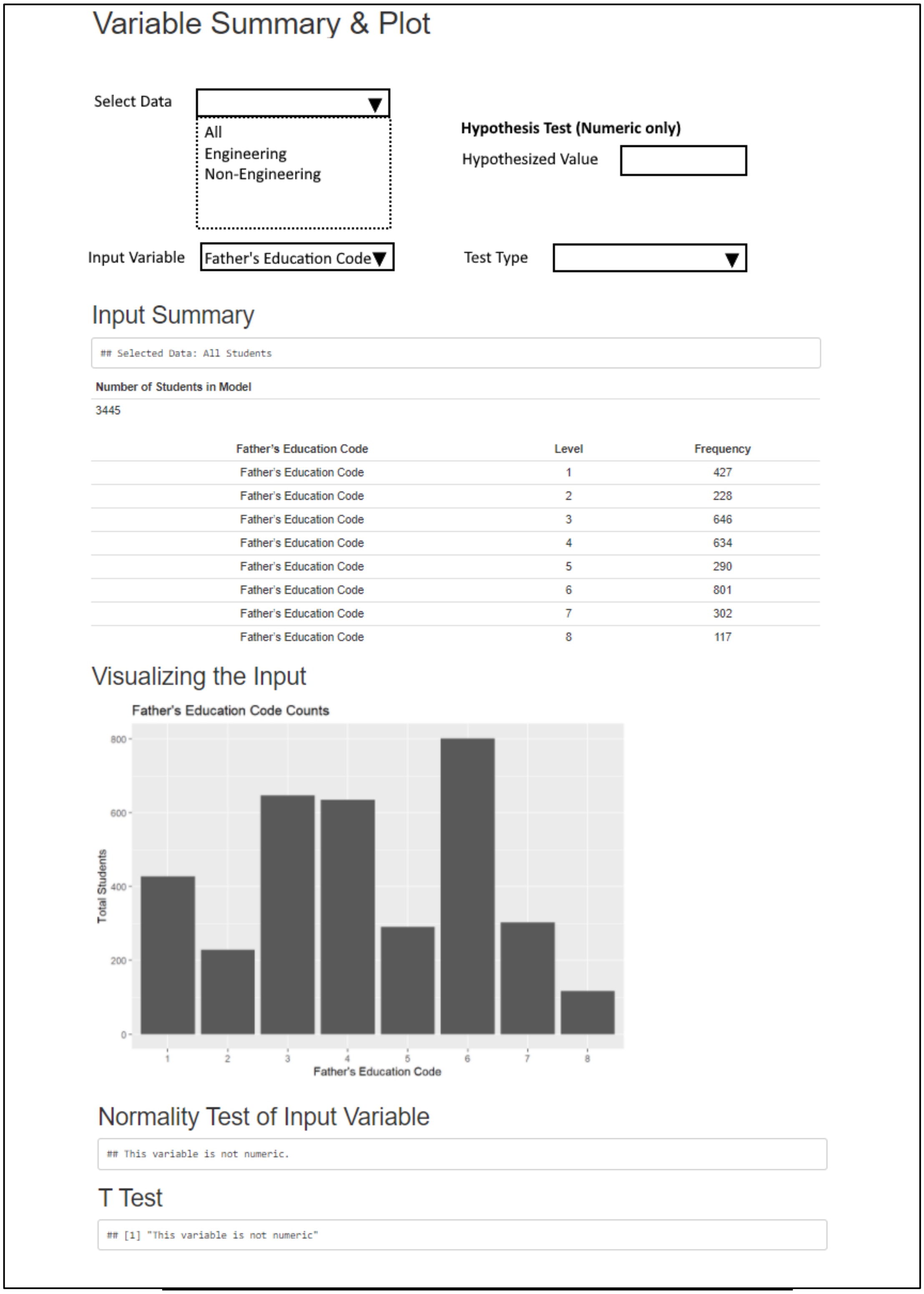


Figure 24: Numerical Variable Summary and Plot

A categorical example utilizing the input of Father's Education Code can be seen in Figure 25.



Father's Education Code	Level	Frequency
Father's Education Code	1	427
Father's Education Code	2	228
Father's Education Code	3	646
Father's Education Code	4	634
Father's Education Code	5	290
Father's Education Code	6	801
Father's Education Code	7	302
Father's Education Code	8	117

Figure 25: Categorical Variable Summary and Plot

The categorical variable summary tool allows the user to select an input variable and filter the data by all transfer students, engineering transfer students, or non-engineering students. The dashboard returns the total number of students per level of the selected variable. Then, a bar chart is populated to illustrate the distribution of students across levels. This tool allows the user to quickly view and understand a categorical variable.

Additionally, a factor investigation dashboard tool was created. This tool aims to illustrate the relationship between an input variable and a response variable. Different outputs exist depending on the type of input variable (numerical versus categorical). For example, if a user wanted to look at the impact of CSU Mentor GPA on the undergraduate degree GPA of transfer students. The user would select these variables, and additionally provide a hypothesized value for CSU Mentor GPA and the type of hypothesis test. Then, the dashboard tool would return the total number of students within the model and summary statistics. Additionally, it would return a scatterplot illustrating the relationship between the input and response variable. Also, a boxplot would be developed to illustrate the overall distribution of the input variable and to identify if there are extensive outliers. A response variable histogram would also be produced to understand the distribution of undergraduate degree GPA. Finally, the dashboard would return an appropriate hypothesis test depending on the normality of the input variable. If the input variable is normal, a one-sided t-test will be performed utilizing the user’s hypothesis inputs. If the input variable is not normal, a one-sided non-parametric Wilcoxon test will be utilized. The result clarifies if the test is significant. An example of a numerical factor investigation can be seen in Figure 26.

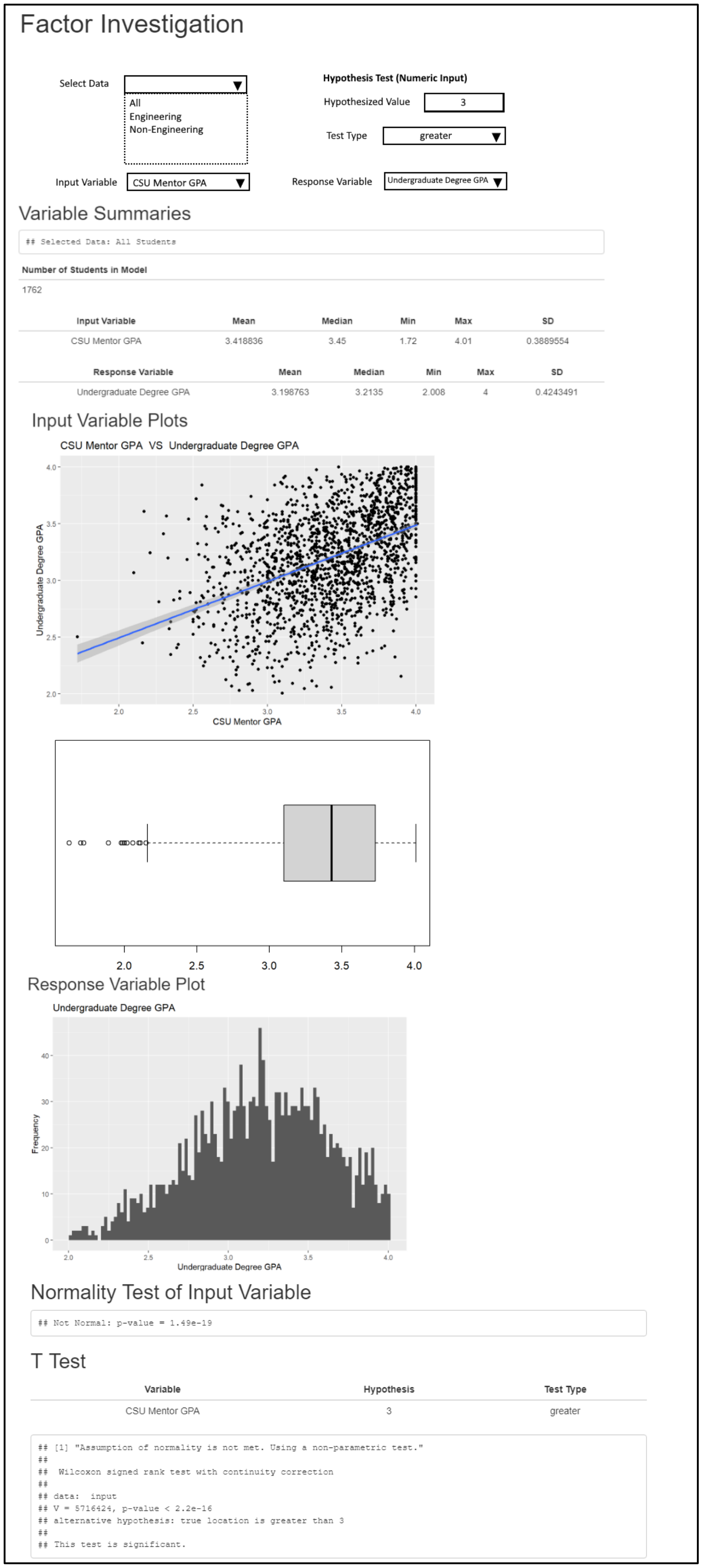


Figure 26: Numerical Factor Investigation

Furthermore, in the case of a categorical variable, the user would select an input variable such as Father's Education Code with the response variable of undergraduate degree GPA. The dashboard would then return the number of students within the model as well as the number of students per level of the categorical variable. Then, a bar chart of the input variable is presented to illustrate the distribution of students across the variable. Next, a main effects plot is developed to demonstrate the individual effect of each level of the categorical variable on the response variable. Lastly, the response variable is plotted with a histogram to depict the overall distribution. An example of Father's Education Code versus undergraduate degree GPA can be seen in Figure 27.

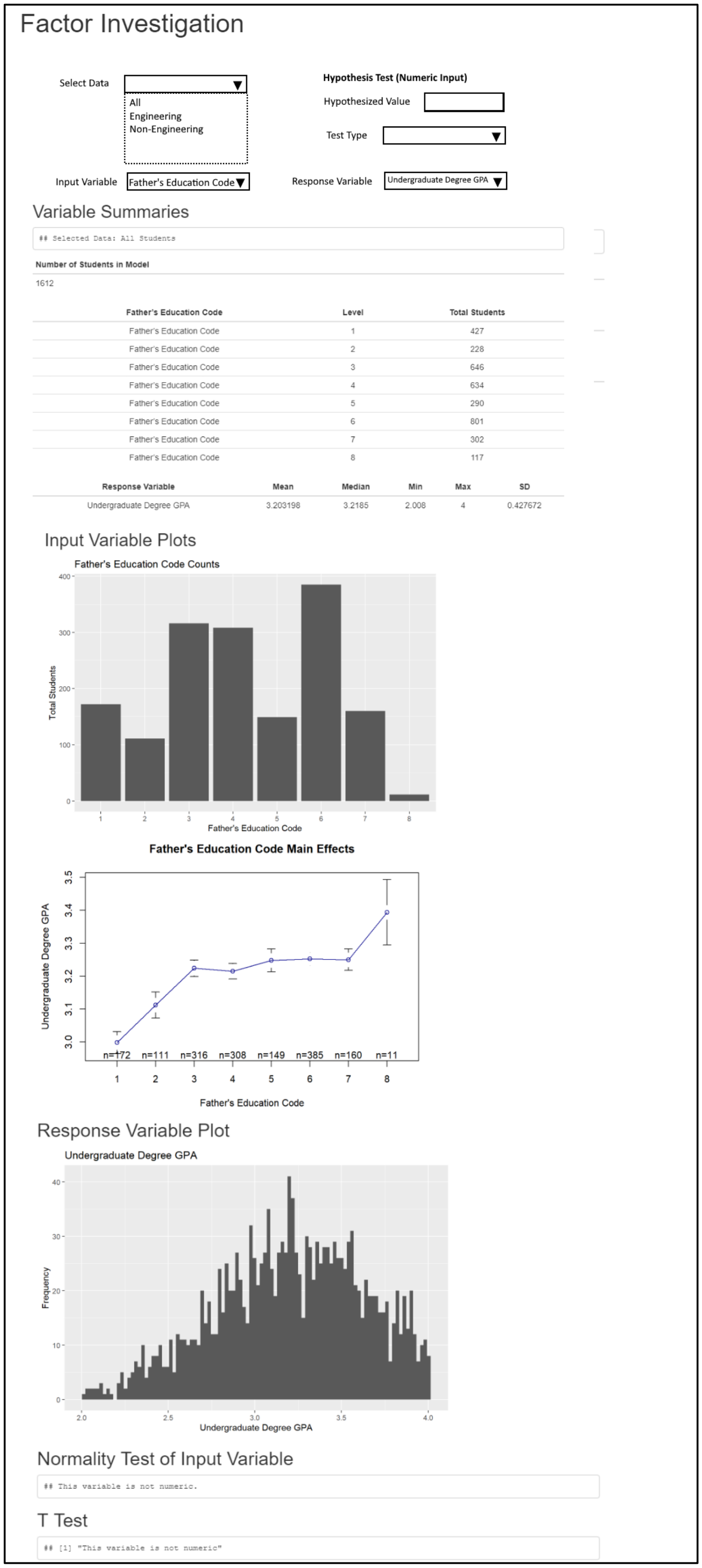
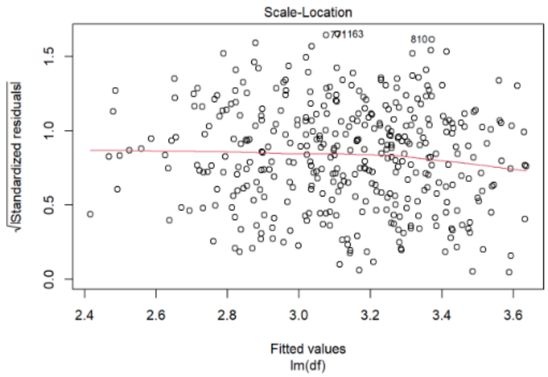
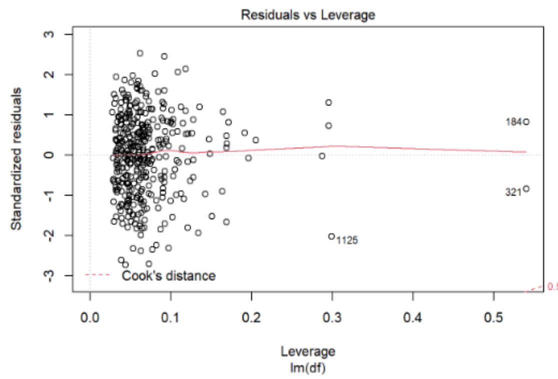


Figure 27: Categorical Factor Investigation

A dashboard tool was also developed to create a regression model. The user would enter the desired input variables into a text box, select a response variable, and filter the data by all transfer students, engineering transfer students, or non-engineering transfer students. The purpose of creating a regression model is to allow the user to investigate which variables are significant on the response variable of undergraduate degree GPA. Prior to creating the regression model, the tool validates the normality of the residuals. If the residuals are not normal, the appropriate transformation is performed on the response variable to achieve normality of the residuals if possible. Then, the residual plots are developed for the user to analyze for model adequacy. These residual plots include a residuals versus fitted plot, normal probability plot, standardized residuals plots, and a Cook's distance plot. These plots can be utilized to determine if the residuals meet the assumptions of equal variance, independence, and normality. The Cook's distance plot can be used to determine if there are any significant outliers. Finally, a linear regression model is fitted with the selected variables. The coefficients of each term are presented along with the overall significance of the model and R-squared values. The R-squared values illustrate the level of variation caused by the variables contained within the model. This can be useful for the user to determine which variables and their levels have an impact on the undergraduate degree GPA of transfer students. An example of creating a regression model with the tool can be seen in Figure 28 and Figure 29. The variables selected for this example include Major, Ethnicity, CSU Mentor GPA, Father's Education Code, and Extracurricular Activity Points. The resulting model is significant, and results in a R-squared adjusted value of 0.324. Therefore, about 32.4% of variation within the response variable of undergraduate degree GPA is explained by the selected variables.

Figure 28: Regression Model Creation 1





Fitting the Model

```
##
## Call:
## lm(formula = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.91804 -0.22965  0.02087  0.23686  0.84384
##
## Coefficients:
##                                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     1.446424   0.320653   4.511 8.78e-06 ***
## MajorBMED                       0.082369   0.086019   0.958   0.3389
## MajorCE                        -0.031127   0.077100  -0.404   0.6867
## MajorCPE                        0.015659   0.099589   0.157   0.8751
## MajorCSC                        0.016028   0.077115   0.208   0.8355
## MajorEE                        -0.158624   0.078260  -2.027   0.0434 *
## MajorENVE                       0.052774   0.116104   0.455   0.6497
## MajorGENE                      -0.224838   0.145810  -1.542   0.1240
## MajorIE                         0.185771   0.116653   1.593   0.1122
## MajorMATE                       0.112824   0.092893   1.215   0.2253
## MajorME                        -0.143852   0.070923  -2.028   0.0433 *
## MajorMFGE                       0.009442   0.114533   0.082   0.9343
## MajorSE                        -0.092489   0.096761  -0.956   0.3398
## EthnicityASIAN                 -0.190291   0.254095  -0.749   0.4544
## EthnicityBLACKPRF              -0.401731   0.303159  -1.325   0.1860
## EthnicityDECLINE               -0.164863   0.272070  -0.606   0.5449
## EthnicityHISPA                 -0.268784   0.254194  -1.057   0.2910
## EthnicityTWOMORE               -0.096820   0.260170  -0.372   0.7100
## EthnicityWHITE                 -0.063156   0.252471  -0.250   0.8026
## `CSU Mentor GPA`                0.536030   0.057954   9.249  < 2e-16 ***
## `Father's Education Code`2      0.065747   0.072533   0.906   0.3653
## `Father's Education Code`3      0.014726   0.069056   0.213   0.8313
## `Father's Education Code`4      0.079637   0.068332   1.165   0.2446
## `Father's Education Code`5      0.042413   0.083688   0.507   0.6126
## `Father's Education Code`6      0.039247   0.065438   0.600   0.5490
## `Father's Education Code`7      0.213157   0.082842   2.573   0.0105 *
## `Extracurricular Activity Pts` -0.002287   0.001311  -1.745   0.0819 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3437 on 356 degrees of freedom
## Multiple R-squared:   0.37,  Adjusted R-squared:  0.324
## F-statistic: 8.041 on 26 and 356 DF,  p-value: < 2.2e-16
```

Figure 29: Regression Model Creation 2

The proposed dashboard tools above incorporate key data and statistical analysis techniques that will allow users to investigate and understand which factors significantly impact the success of transfer students. This is critical for furthering research on how to support transfer students pre-transfer, during transfer, and post-transfer. Implementing these tools will allow the ENGAGE initiative to create enduring changes within Cal Poly to better support and service engineering transfer students. In addition, these dashboards can be expanded as data is added or new analysis methods are desired.

Chapter 5
FUTURE DIRECTIONS

The existing R code has not yet been integrated within the data sandbox system. Therefore, it is necessary to consider that implementing the developed R dashboards within a data sandbox will require several key steps. First, the existing R code will be integrated within the framework of the desired dashboards. These dashboards will likely include the developed dashboards presented in the previous section. To reiterate, these include the dashboards of a data overview, numerical and categorical variable summary, variable summary and plot, factor investigation, and regression model creation. The framework for each of these dashboards will need to incorporate the relevant existing R code. Then, each dashboard can be briefly tested to ensure overall functionality. Once the code is incorporated within the framework, usability testing can be performed utilizing volunteers already involved within ENGAGE research. These users can trial each dashboard, then provide feedback and revision recommendations for further improvements. Each trial should instruct the user to complete a specific task within the respective dashboard while being observed by a test moderator. This step can be repeated until all involved parties approve the dashboards for use. These steps are illustrated in Figure 30.

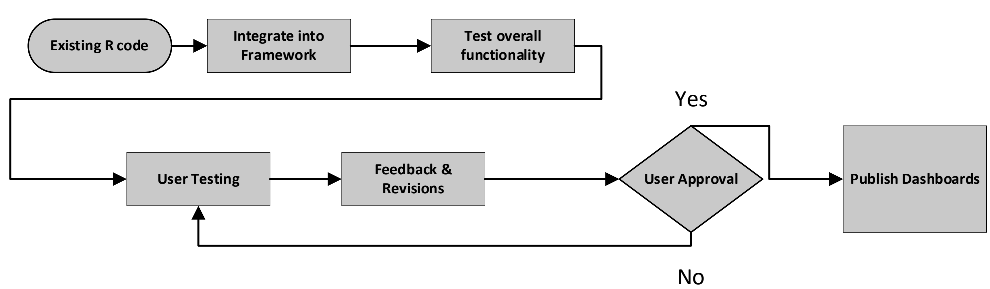


Figure 30: Implementing the R dashboards

Looking specifically at the usability testing, it is essential to consider the major goals of the overall data sandbox. The users should rate the dashboards on several key categories after each trial. The following matrix represents possible categories for user ratings and feedback (Table 18).

Table 18: Usability Testing User Matrix

Category	Rate (1-10)	Additional Comments
Ease of Navigation		
Achieves Purpose		
Ease of Use		
Overall Satisfaction		

Ease of navigation refers to how easily the user can navigate within each dashboard. For example, how easily a user can filter the data, select the desired variables, and view the appropriate output. Achieves purpose relates to how well each dashboard achieves its initial intent. For example, if the dashboard's goal is to illustrate key variable summaries and plots, which summaries and plots are provided and do they successfully aid the user. Next, ease of use refers to how easily the user can understand the functionalities of the dashboards. This specifically is intended to rate the ease of use and/or understanding for any type of user. For example, can a user with little data or statistical analysis background complete basic analysis with ease and understanding. This category is key is making data analysis techniques available to users of varying experience levels. Finally, the overall satisfaction category refers to how satisfied the user is with the dashboards. This could relate to the overall aesthetics or layout of the dashboards. It could also relate to how the user feels after utilizing the dashboards in terms of completing meaningful data analysis. Overall, it is essential that once the existing R code is integrated within the data sandbox, future users are involved in usability testing to ensure that the completed dashboards achieve their intended purpose.

Once the dashboards are implemented and revised, they will likely be available to more users for further research into the success of transfer students. The accessibility of the presented dashboards will allow all types of users to investigate which factors relating to transfer students impact success. Therefore, further discoveries will be possible to make in terms of determining which tools and practices can be implemented to improve the support for transfer students pre-transfer, during transfer, and post-transfer. Furthermore, with the support of the existing dashboards, more dashboards could be developed to achieve additional goals. More transfer student data and types of data could be added to the data sandbox to further the number of factors considered on the success of transfer students. In the future, qualitative data survey data stemming from the two ENGAGE cohorts will be available. This survey data could be added to the data sandbox for further analysis on what factors benefit engineering transfer students. Thus, the dashboards proposed for implementation are the first steps in furthering the research on the success of transfer students in an effort to support these students more effectively in the future.

Chapter 6

CONCLUSIONS

Local low-income academically talented students face barriers in transferring to Cal Poly to complete their technical degrees. To improve the quality of life in the Central Coast area, it is vital that support is available for transfer students. Through the ENGAGE initiative, further research will continue to analyze which factors are impactful on the success of transfer students. The current data analyzed included five years of transfer students separated by engineering and non-engineering majors. The metric for success selected was the Cal Poly cumulative undergraduate degree GPA. The factors analyzed stem from primarily application data, and can be seen in Appendix A.

The developed model based upon 254 engineering transfer students included the factors of CSU Mentor GPA, Extracurricular Activity Points, Major, Ethnicity, Father's Education Code, and the CA Resident Flag. Together, these factors had a significant effect on the cumulative undergraduate degree GPA of engineering transfer students. Key takeaways include that transfer students with a lower CSU Mentor GPA were predicted to achieve a lower undergraduate degree GPA. Students with a father who completed a postgraduate degree were most likely to achieve a higher undergraduate degree GPA while those with fathers who graduated high school were predicted to achieve the lowest undergraduate degree GPA. Students who identified within White and Two or More ethnicities were more likely to achieve a higher undergraduate degree GPA. Students who identified as Black/African American preference and Hispanic were least likely to obtain a high undergraduate degree GPA. Those who resided within California at the time of application were predicted more likely to succeed. Through this model, it is apparent students who stem from primarily underrepresented ethnicities and/or are first generation college students are predicted to achieve a lower undergraduate degree GPA.

The model based upon 1225 non-engineering transfer students included the predictive factors of CSU Mentor GPA, Major, Ethnicity, Work Hour Range Code, Gender Code, and Academic Extracurricular Leadership Points. These factors each influenced the undergraduate degree GPA of non-engineering transfer students. Students with a lower CSU Mentor GPA were less likely to achieve a higher undergraduate degree GPA. Students who identified within the White and Two or more ethnicities were most likely to obtain a higher undergraduate degree GPA. Students who identified within the American Indian/Alaska Native and Black/African American preference were predicted least likely to receive a higher undergraduate degree GPA. Male non-engineering transfer students were predicted to achieve a lower undergraduate degree GPA. Those with a higher level of Academic Extracurricular Leadership Points as a result of their extracurricular activities were more likely to succeed.

Both developed models stemming from the transfer student data illustrated several factors that had an impact on the Cal Poly cumulative undergraduate degree GPA. Across both engineering and non-engineering transfer students, those who identified within the White or Two or More ethnicities ultimately were predicted more likely to obtain a higher undergraduate degree GPA. Those who identified within underrepresented groups such as Black/African American preference, Hispanic, and American Indian/Alaska Native were less likely to achieve a high undergraduate degree GPA. Students who had a higher CSU Mentor GPA were more likely to succeed in their undergraduate degree

GPA. Thus, any further initiatives should consider these factors in terms of impacting success as defined by undergraduate degree GPA.

The current support programs offered independently by Allan Hancock, Cuesta, and Cal Poly include advising, mentoring, clubs, career services, a writing center, and tutoring. Based upon existing transfer support systems discussed in the literature review, several program recommendations are included. To the benefit of current potential transfer students, step by step transfer information is already readily available. Also, campus tours are explicitly offered to allow potential transfer students to understand the environment they would be entering. These programs provide opportunities and support for transfer students to receive additional aid pre-transfer, during transfer, and post transfer. One potential improvement in improving the overall transfer experience is offering a first quarter seminar for transfer students to build their social network and have immediate support and resources. This would potentially further the available transfer capital; the factors impacting the success of students transferring.

Lastly, several dashboard tools developed through the coding program R were presented for eventual implementation within a data sandbox available to users of varying experience levels. These dashboard tools will allow the user to investigate transfer student data to further the research on the factors that most impact the success of transfer students. It will allow the user to discover trends across transfer students and view the distribution of students across key variables. This will allow researchers of all backgrounds to investigate and understand transfer students with key data analysis and statistical techniques. Through the implementation of these dashboards, further research can be performed to continue the ENGAGE initiative of implementing support for engineering transfer students pre-transfer, during transfer, and post-transfer. These efforts will allow low-income academically talented students to receive support that will ease barriers for completing vital technical degrees.

This material is based upon work supported by the National Science Foundation under Grant No. (NSF Grant numbers 1834128 and 1834154)

REFERENCES

1. Allan Hancock College. (2020, April 22). *All About Allan Hancock College*. Retrieved from Allan Hancock College: https://www.hancockcollege.edu/about/index.php

2. Bennett, C., & Okinaka, A. (1990). Factors Related to Persistence Among Asian, Black, Hispanic, and White Undergraduates at a Predominantly White University: Comparison Between First- and Fourth-Year Cohorts. *Urban Review*, 33-60.

3. Berger, J., & G.D., M. (2003). Assessing the Transition of Transfer Students from Community Colleges to a University. *NASPA Journal*, 1-23. doi:10.2202/1949-6605.1277

4. Cal Poly . (2020). *Fall 2020 Census Admissions Info Brief*. Retrieved from https://content-calpoly-edu.s3.amazonaws.com/ir/1/images/admissions-2020_51279695%20%2814%29.pdf

5. Cal Poly SLO. (2020). *Poly View / Fall 2020*. Retrieved from https://content-calpoly-edu.s3.amazonaws.com/ir/1/images/PolyView%20combind.pdf

6. Carnevale, A., & Smith, N. (2018). *Balancing Work and Learning: IMPLICATIONS FOR LOW INCOME STUDENTS.* Georgetown University Center on Education and the Workforce.

7. Cuesta College. (2020). *Cuesta College Fast Facts*. Retrieved from Cuesta College: https://www.cuesta.edu/about/info/collegefacts.html

8. Fard, A. (2020, February 10). *9 best practices when designing a dashboard*. Retrieved from UX Collective: https://uxdesign.cc/9-best-practices-when-designing-a-dashboard-ae79fe182b68

9. Fauria, R., & Fuller, M. (2015). Transfer student success: educationally purposeful activities predictive of undergraduate GPA. *Research & Practice in Assessment*, 39+. Retrieved from https://link.gale.com/apps/doc/A419764341/AONE?u=calpolyw_csu&sid=AONE&xid=75094d1e

10. Hills, J. (1965). Transfer Shock -- The Academic Performance of the Junior College Transfer. *The Journal of Experimental Education*, 210-215. doi:10.1080/00220973.1965.11010875

11. Janes, A. &. (2013). Effective dashboard design. *Cutter IT Journal. 26*, 17-24.

12. Jenkins, D., & Fink, J. (2016). *Tracking Transfer: New Measures of Institutional and State Effectivesness in Helping Community College Students Attain Bachelor's Degrees.* : Community College Research Center, Teachers College, Colombia University.

13. Klein, L. (2018, July 23). *Is Your Embedded Dashboard Useful and Usable?* Retrieved from Logi Analytics: https://www.logianalytics.com/designing-dashboards/is-your-dashboard-useful-and-usable/

14. Laanan, F. (2000). Community College Students' Career and Education Goals. *NEW DIRECTIONS FOR COMMUNITY COLLEGES,*, 19. Retrieved from https://doi-org.ezproxy.lib.calpoly.edu/10.1002/cc.11202

15. Labor, U. S. (2020, October 13). *TED: The Economics Daily*. Retrieved from U.S. Bureau of Labor Statistics: https://www.bls.gov/opub/ted/2020/hispanics-made-up-44-point-7-percent-of-the-labor-force-in-new-mexico-in-2019.htm

16. Ma, J., & Baum, s. (2016). *Trends in Community Colleges: Enrollment, Prices, Student Debt, and Completion.* College Board Research.

17. Moore, C., & Shulock, N. (2010). *Divided We Fail: Improving Completion and Closing Racial Gaps in California's Community Colleges.* CSU Sacramento, Institute for Higher Education Leadership & Policy.

18. Moser, K. (2020). *Redefining transfer student success: Transfer capital and the Laanan-transfer students' questionnaire (L-TSQ) revisitied.* Graduate Theses and Dissertations. Retrieved from https://lib.dr.iastate.edu/etd/12414

19. Ogilvie, A., & Knight, D. (2020). Engineering Transfer Students' Reasons for Starting at Another Institution and Variation Across Subpopulations. *Journal of Hispanic Higher Education*, 69-83. doi:10.1177/1538192718772659

20. Richter Lagha, R., Burningham, Z., Sauer, B. C., Leng, J., Peters, C., Huynh, T., . . . Kramer, B. J. (2020). Usability Testing a Potentially Inappropriate Medication Dashboard: A Core Component of the Dashboard Development Process. *Applies clinical informatics*, 528-534.

21. Terenizini, P., Lattuca, L., Ro, H., & Knight, D. (2014). *America's Overlooked Engineers: Community Colleges and Diversity in Undergraduate Education.* Retrieved from http://hdl.handle.net/2027.42/107460

22. U.S. General Services Administration Technology Transformation Services. (n.d.). *System Usability Scale (SUS)*. Retrieved from usability.gov: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html

23. Wetzel, L., & Debure, K. (n.d.). The Role of Faculty in Fostering STEM Transfer Student Success. *Journal of College Science Teaching*, 42+. Retrieved from https://link.gale.com/apps/doc/A529490339/AONE?u=calpolyw_csu&sid=AONE&xid=467e4bae

24. Winterer, E., Froyd, J., Borrego, M., Martin, J., & Foster, M. (2020). Factors influencing the academic success of Latinx students matriculating at 2-year and transferring to 4-year US institutions-implications for STEM majors: a systematic review of the literature. *International Journal of STEM Education*. Retrieved from https://link.gale.com/apps/doc/A628786347/AONE?u=calpolyw_csu&sid=AONE&xid=de91e8b2

APPENDICES

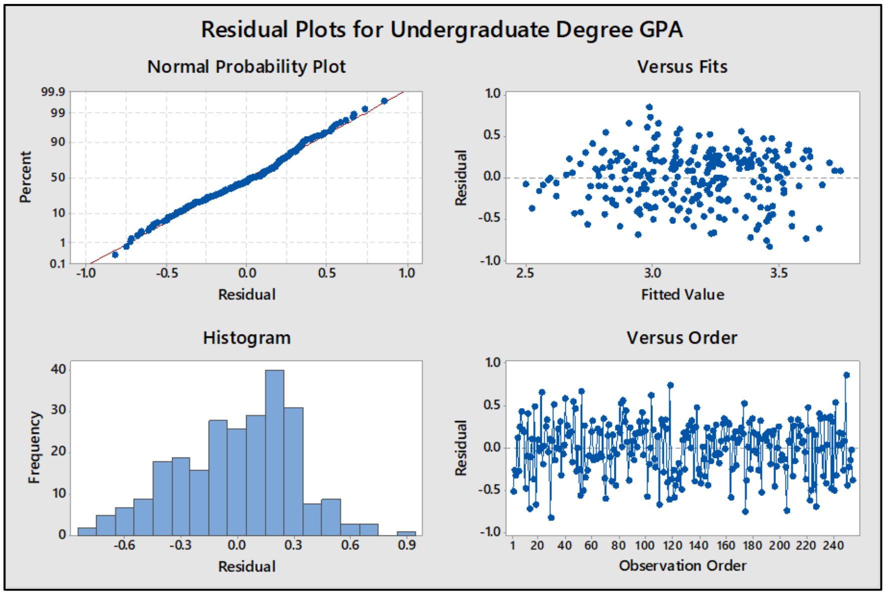
A. Sample Data Variables

Variable Name	Data Type	Brief Definition
Major Code	Categorical	Major Code (EE, IME, BUS, etc.)
CSU Mentor GPA	Continuous	Student application GPA.
Ethnicity Code	Categorical	Ethnicity as identified by student.
EOP Eligible Flag	Categorical	Educational Opportunity Program eligible (Y/N)
Activity Leadership Role Flag	Categorical	Leadership Roles (Y/N)
Extracurricular Hour Range Code	Categorical	Extracurricular Hours Range Code (0-5)
Work Hour Range Code	Categorical	Work Hours Range Code (0-6)
Work Major Related Flag	Categorical	Work Major Related (Y/N)
Gender Code	Categorical	Gender Code (M/F)
Last School Local Flag	Categorical	Local school (Y/N)
Transfer Academic Major Specified Credit Pts	Continuous	Transfer credit pts toward major.
California Resident Flag	Categorical	California Resident (Y/N)
Fathers Education Code	Categorical	Father's education code (0-7)
Mothers Education Code	Categorical	Mother's education code (0-7)
Academic Extracurricular Leadership Pts	Categorical	Leadership Pts Awarded (0, 50, 100, 250)
Academic Extracurricular Major Related Pts	Categorical	Major Related Pts Awarded (0, 10, 50, 100, 150, 250)
Academic Extracurricular Pts	Categorical	Pts awarded based on academic extracurricular activities.
Academic Work Pts	Categorical	Pts awarded based on academic work activities.
Extracurricular Activity Pts	Continuous	Pts awarded based on extracurriculars.
Transfer Academic General Ed Pts	Continuous	Transferrable general education pts.
Transfer Academic IGETC Met Flag	Categorical	Met IGETC Requirements (Y/N)

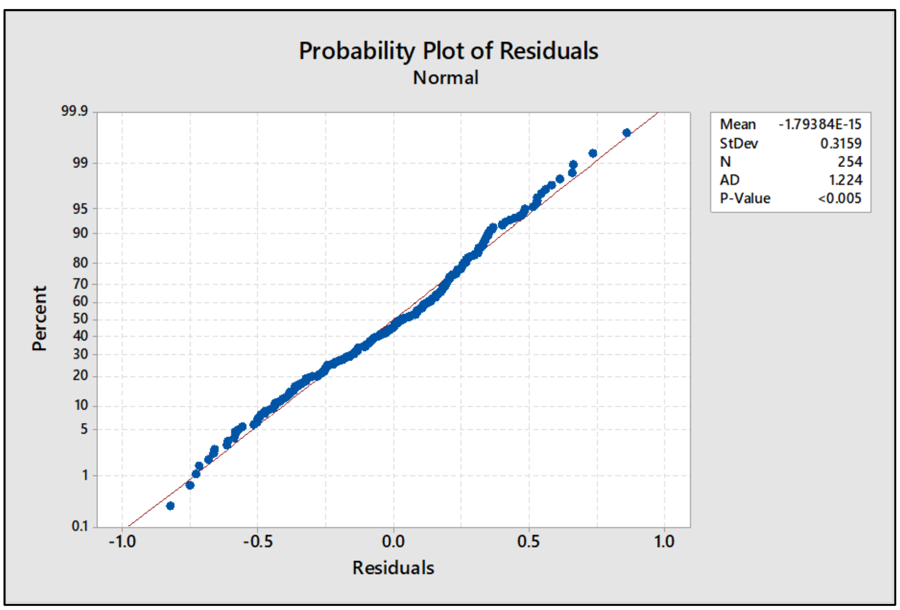
B. Major Acronym Definitions

Majors	
AEPS - Agricultural and Environmental Plant Sciences	FNR - Forestry and Natural Resources
AERO - Aerospace Engineering	GENE - General Engineering
AGB - Agricultural Business	GRC - Graphic Communication
AGCOM - Agricultural Communication	HIST - History
AGSC - Agricultural Science	IE - Industrial Engineering
ARCE - Architectural Engineering	ITP - Industrial Technology and Packaging
ARCH - Architecture	JOUR - Journalism
ART - Art and Design	ITP - Industrial Technology and Packaging
ASCI - Animal Science	JOUR - Journalism
ASM - Agricultural Systems Management	KINE - Kinesiology
BCHM - Biochemistry	LARC - Landscape Architecture
BIO - Biological Sciences	LS - Liberal Studies
BMED - Biomedical Engineering	MATE - Materials Engineering
BUS - Business Administration	MATH - Mathematics
CD - Child Development	MCRO - Microbiology
CE - Civil Engineering	ME - Mechanical Engineering
CM - Construction Management	MFGE - Manufacturing Engineering
COMS - Communication Studies	MLL - Modern Languages and Literatures
CPE - Computer Engineering	MU - Music
CRP - City and Regional Planning	NUTR - Nutrition
CSC - Computer Science	PHIL - Philosophy
DSCI - Dairy Science	PHYS-Physics BS
ECON - Economics	POLS - Political Science
EE - Electrical Engineering	PSY - Psychology
EESS – Environmental Earth & Soil Sciences	REC - Recreation, Parks, and Tourism Administration
ENGL - English	SE - Software Engineering
ENVE - Environmental Engineering	SOCIO - Sociology
ENVM - Environmental Management and Protection	TH - Theatre Arts
ES - Comparative Ethnic Studies	WVIT - Wine and Viticulture

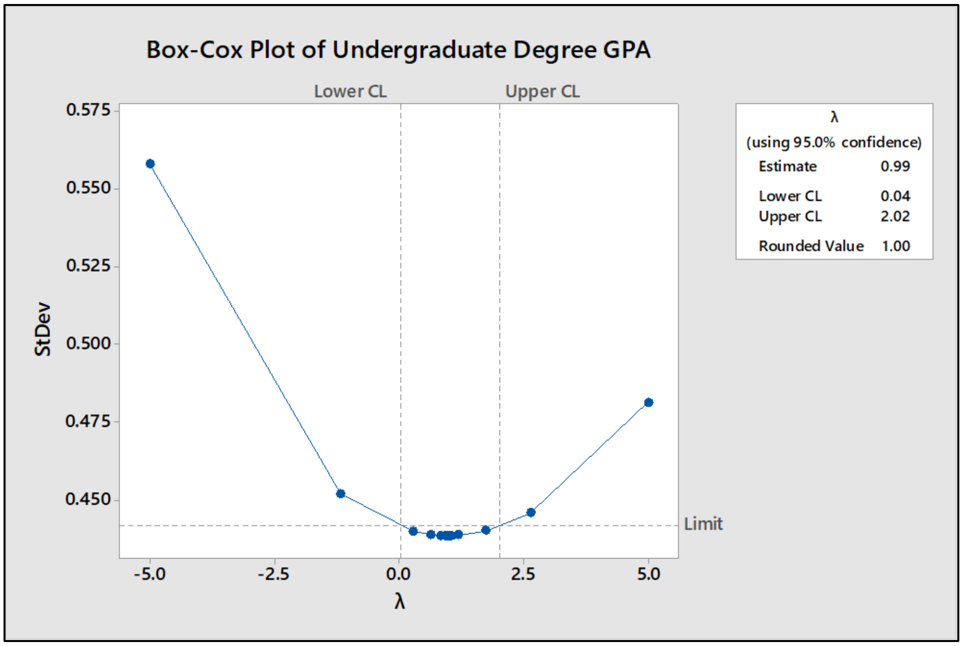
C. Engineering Transfer Students Undergraduate GPA Analysis



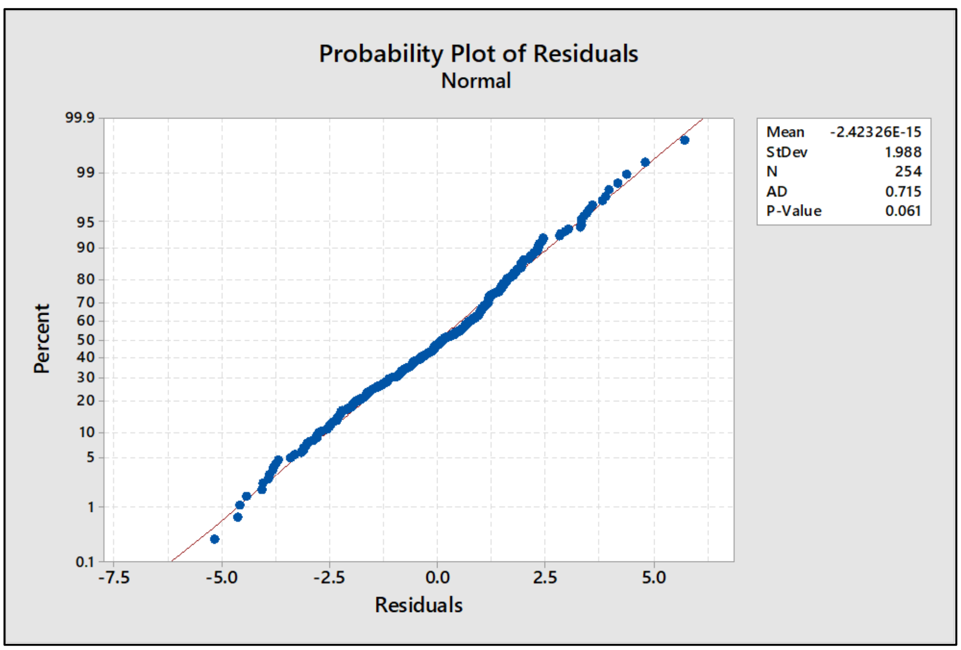
Undergraduate GPA Residual Plots



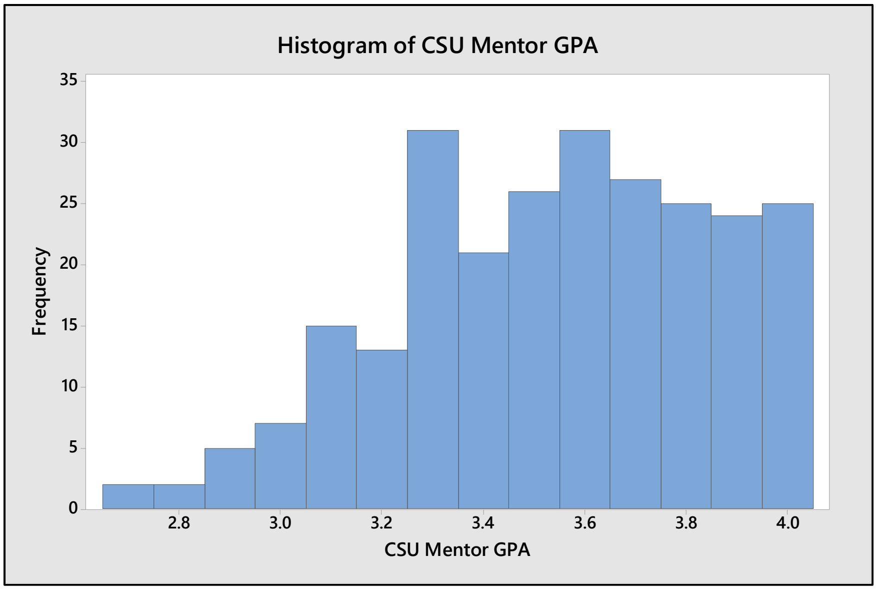
Undergraduate GPA Anderson-Darling Normality Test



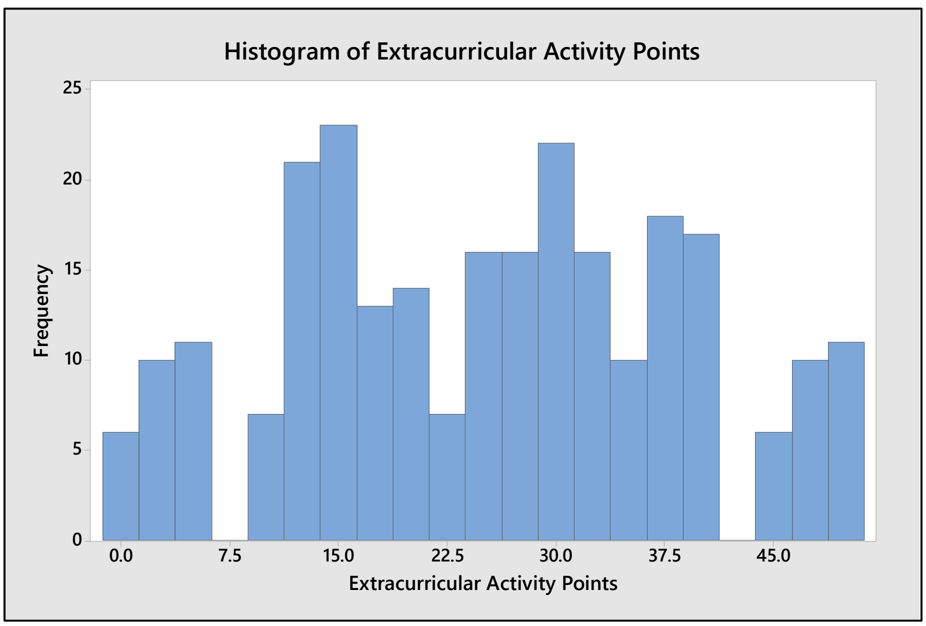
Box Cox Transformation on Undergraduate GPA



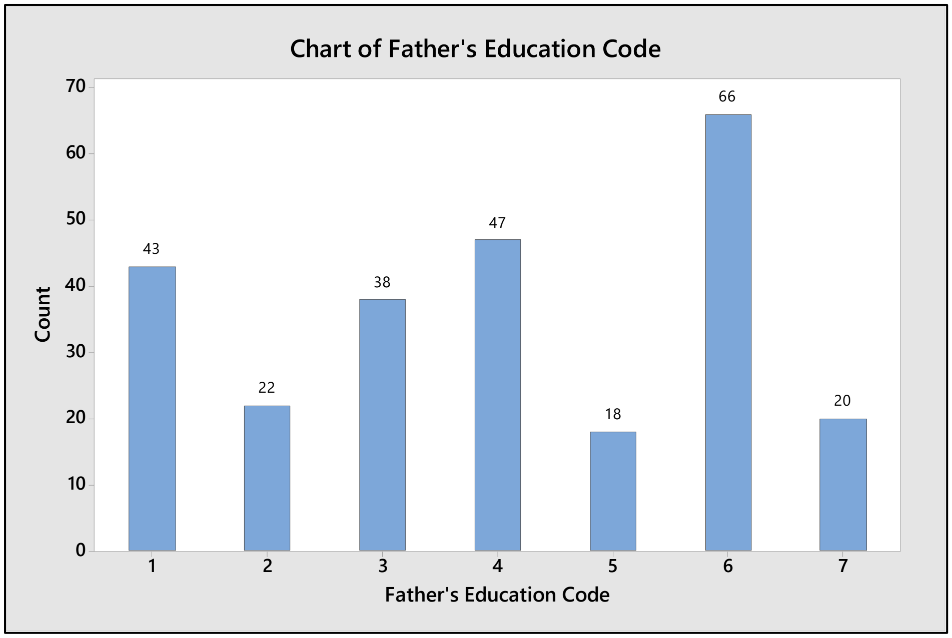
Undergraduate GPA Squared Anderson-Darling Normality Test



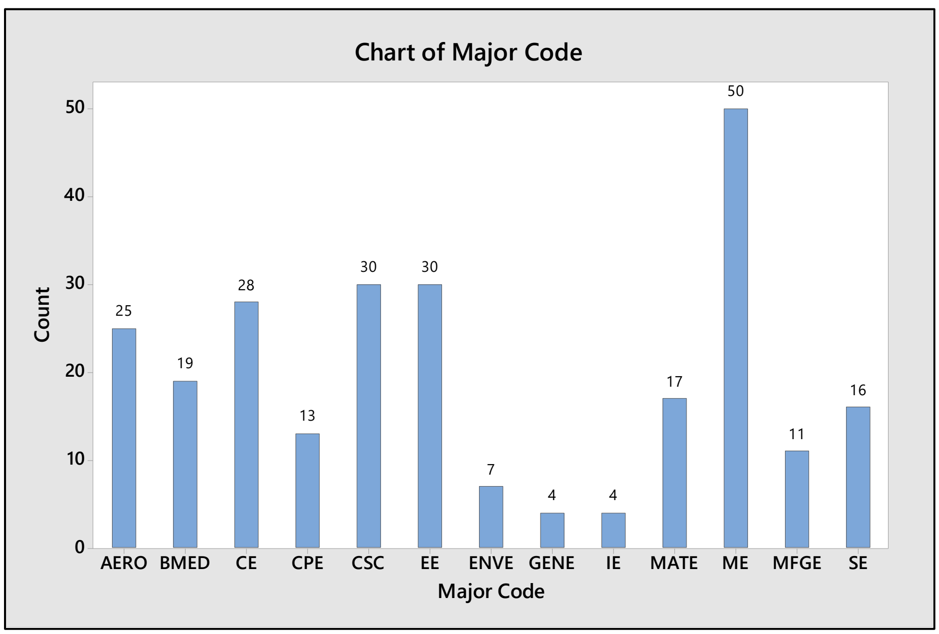
CSU Mentor GPA Histogram



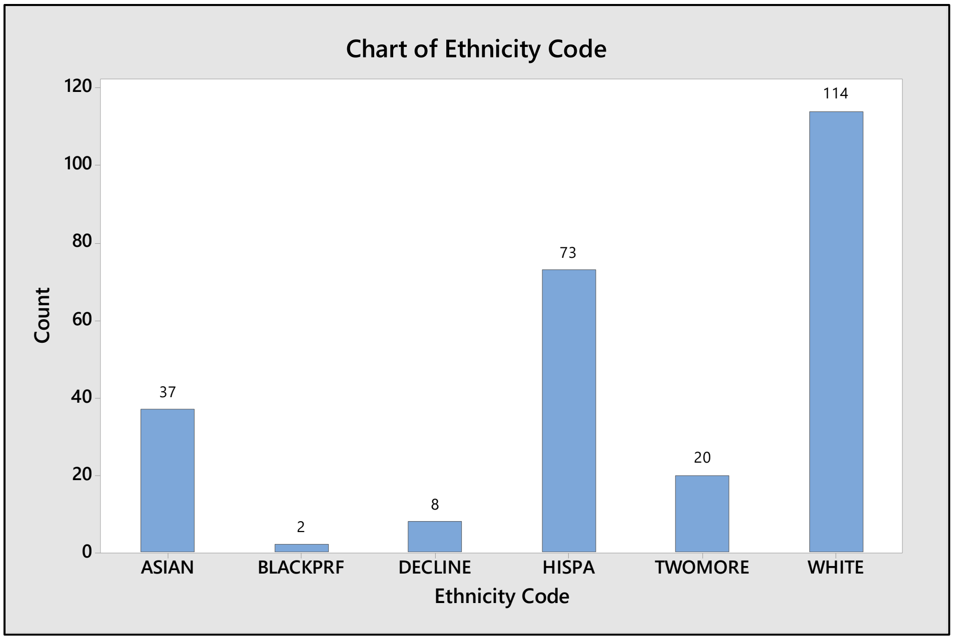
Extracurricular Activity Points Histogram



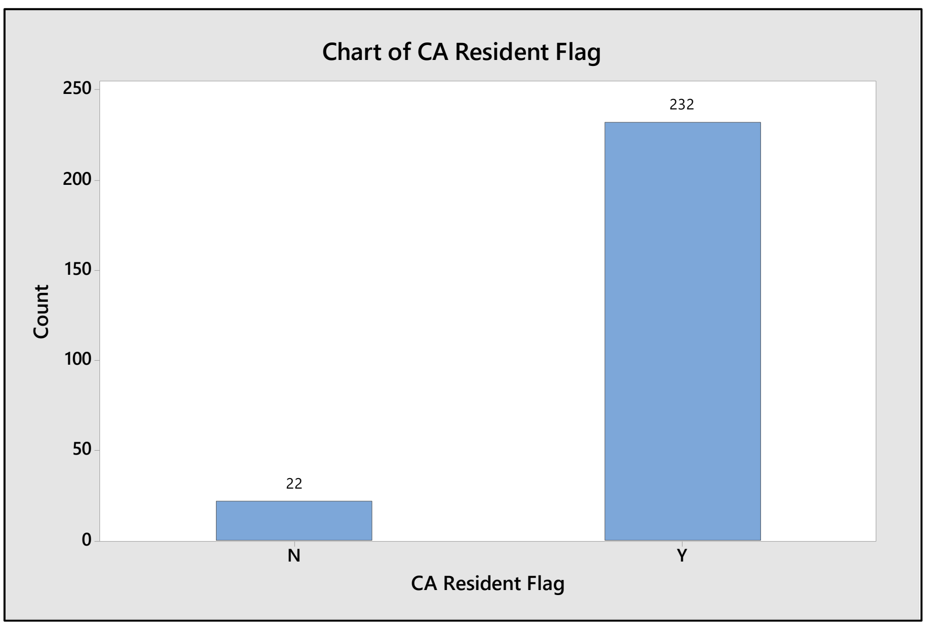
Father's Education Code Counts



Major Code Counts

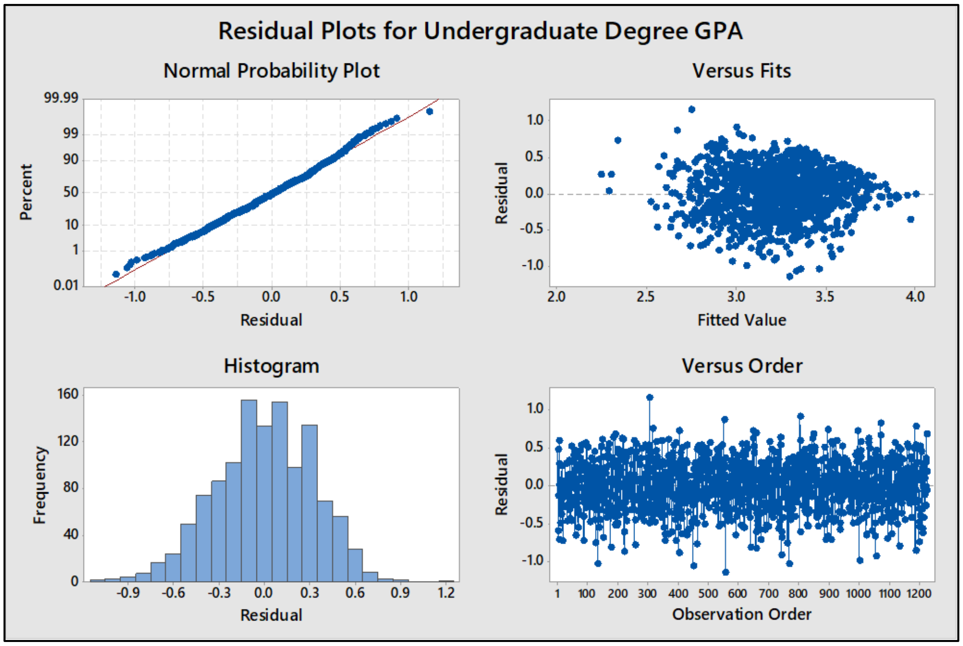


Ethnicity Code Counts

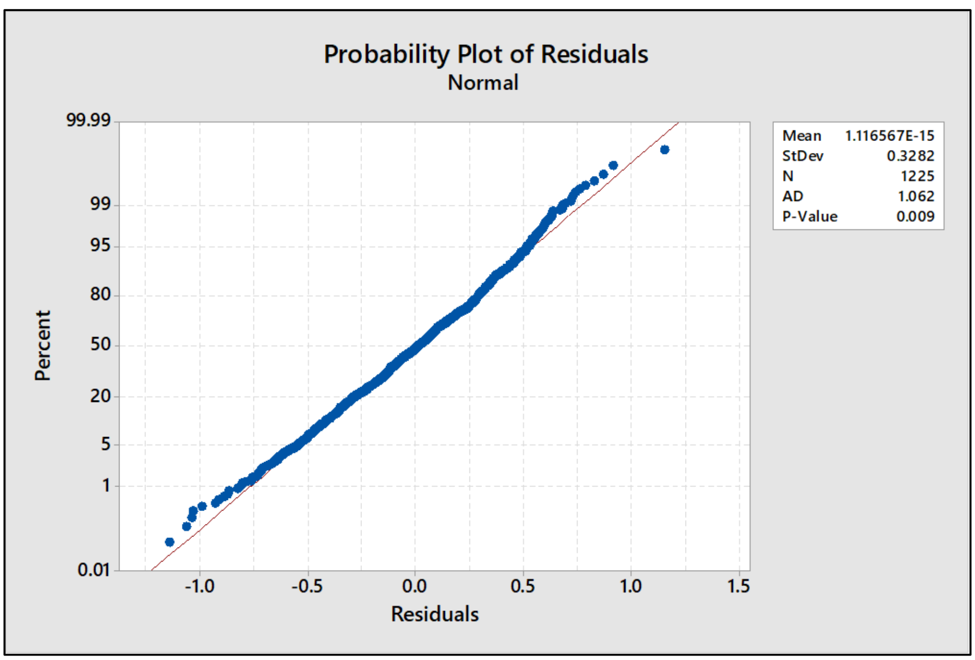


CA Resident Flag Counts

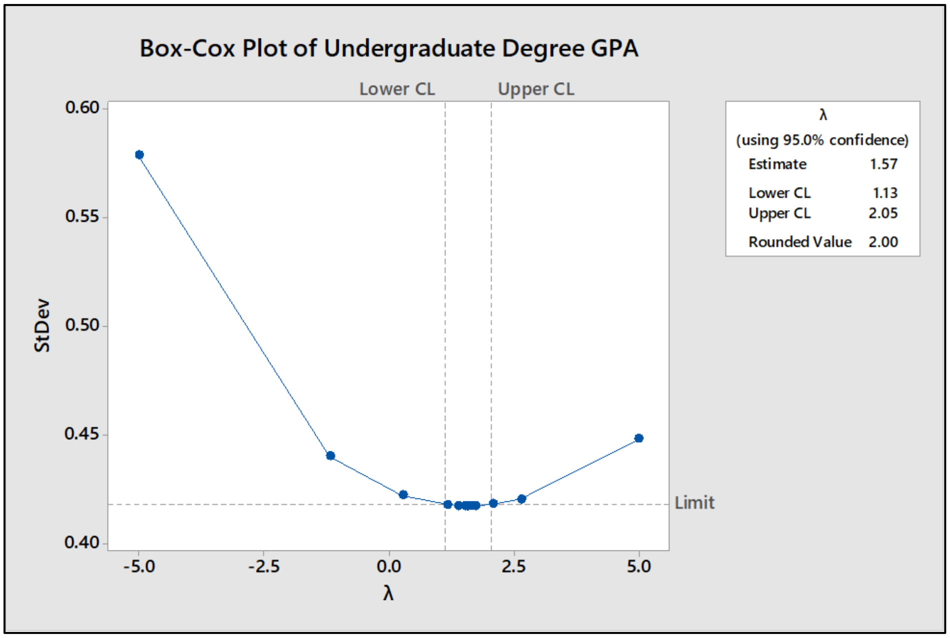
D. Non-Engineering Transfer Students Undergraduate GPA Analysis



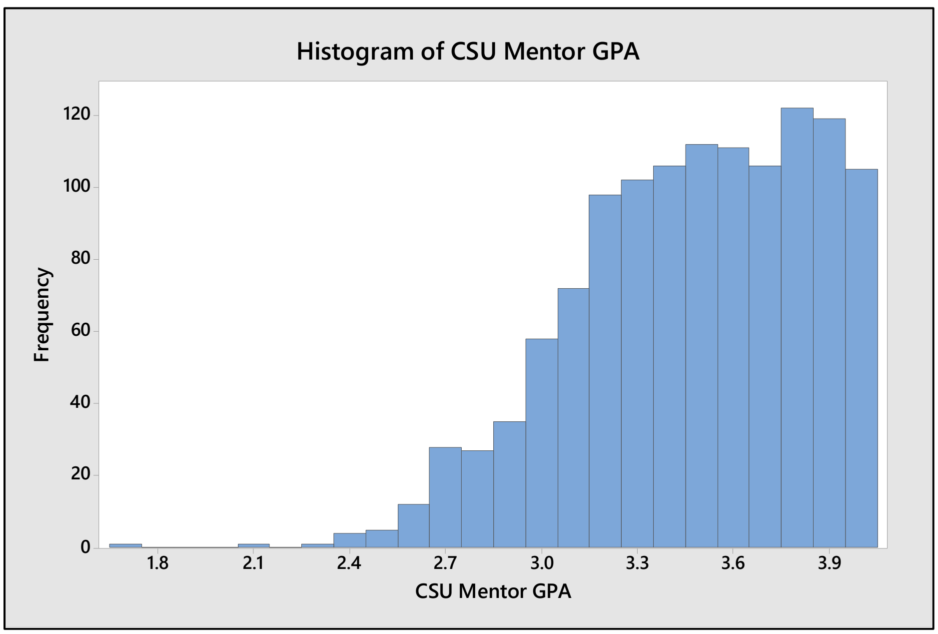
Undergraduate GPA Residuals Plots



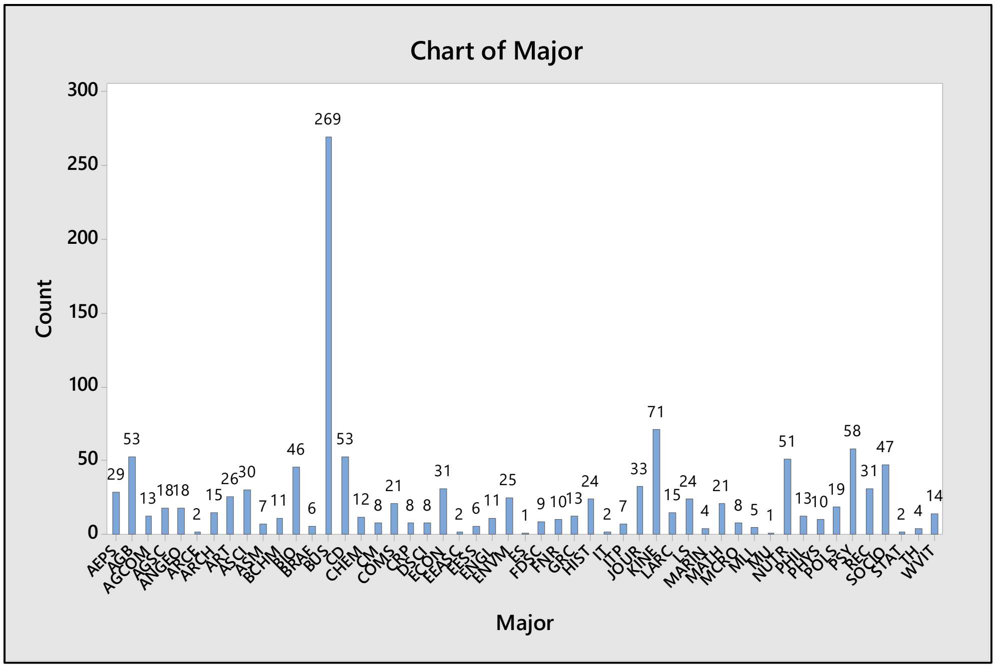
Undergraduate GPA Anderson-Darling Normality Test



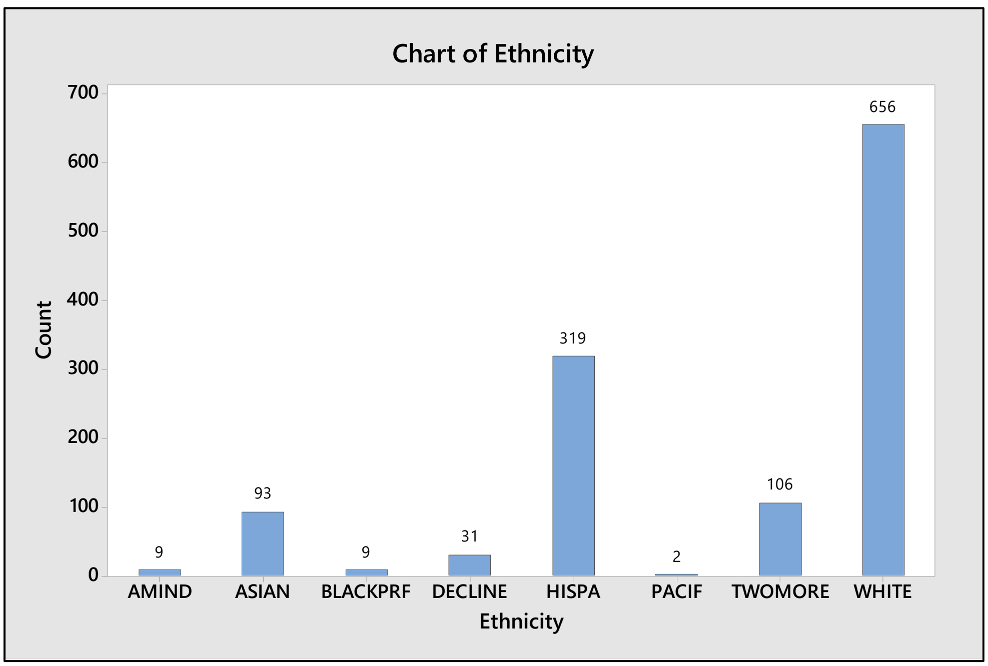
Box Cox Transformation on Undergraduate Degree GPA



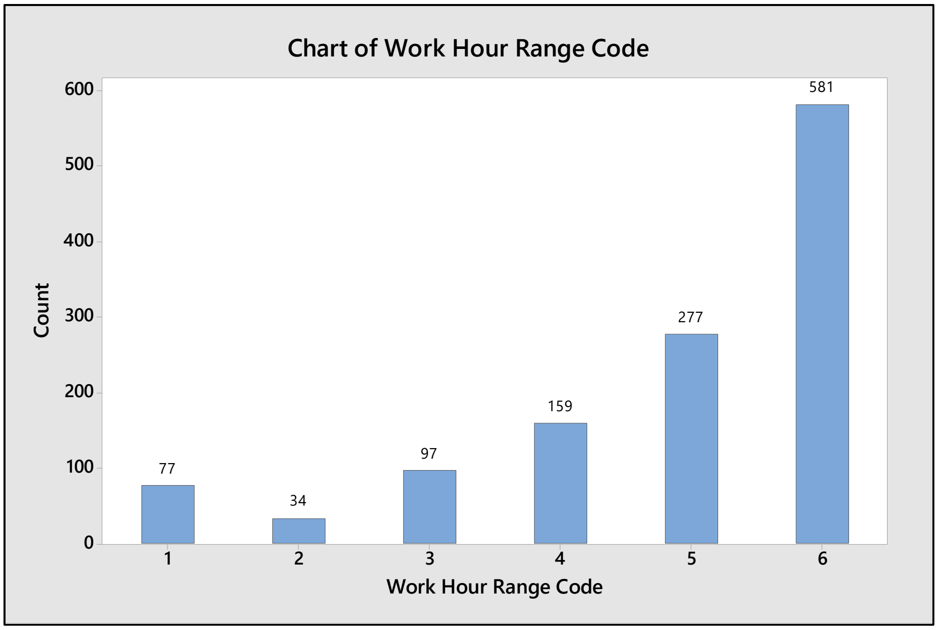
CSU Mentor GPA Histogram



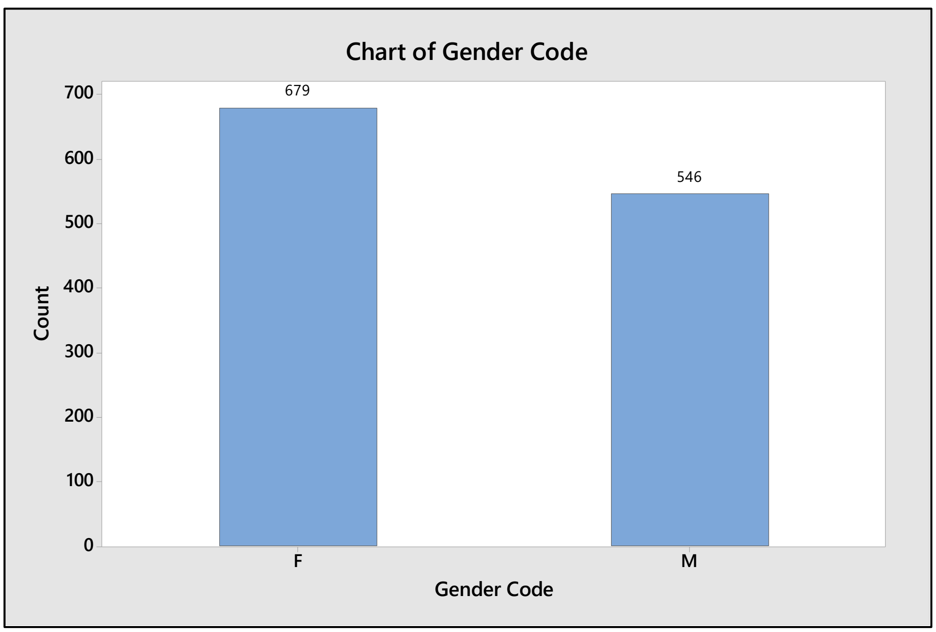
Major Code Counts



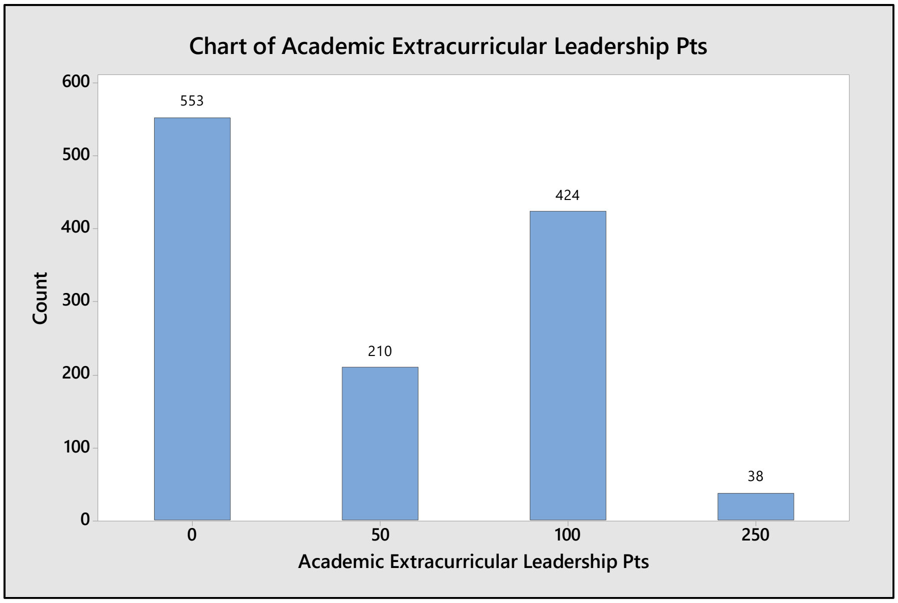
Ethnicity Code Counts



Work Hours Range Code Counts



Gender Code Counts



Academic Extracurricular Leadership Pts Counts