INVESTIGATING HOW THE FIRST YEAR EATS PROGRAM AFFECTS STUDENTS MENTALLY AND ACADEMICALLY

An Undergraduate Research Scholars Thesis

by

MICHAEL LEE

Submitted to the LAUNCH: Undergraduate Research office at
Texas A&M University
in partial fulfillment of requirements for the designation as an

UNDERGRADUATE RESEARCH SCHOLAR

Approved by
Faculty Research Advisors: Dr. Alan Dabney
Dr. Sumana Datta

May 2021

Major: Statistics

RESEARCH COMPLIANCE CERTIFICATION

Research activities involving the use of human subjects, vertebrate animals, and/or biohazards must be reviewed and approved by the appropriate Texas A&M University regulatory research committee (i.e., IRB, IACUC, IBC) before the activity can commence. This requirement applies to activities conducted at Texas A&M and to activities conducted at non-Texas A&M facilities or institutions. In both cases, students are responsible for working with the relevant Texas A&M research compliance program to ensure and document that all Texas A&M compliance obligations are met before the study begins.

I, Michael Lee, certify that all research compliance requirements related to this Undergraduate Research Scholars thesis have been addressed with my Research Faculty Advisors prior to the collection of any data used in this final thesis submission.

This project did not require approval from the Texas A&M University Research Compliance & Biosafety office.

TABLE OF CONTENTS

ABSTRACT

Investigating How the First Year Eats Program Affects Students Mentally and Academically

Michael Lee
Department of Statistics
Texas A&M University

Research Faculty Advisor: Dr. Alan Dabney
Department of Statistics
Texas A&M University

Research Faculty Advisor: Dr. Sumana Datta
Department of Biochemistry and Biophysics
Texas A&M University

The First Year Eats Program (FYE) provides first-year university students among select dorms across Texas A&M University with the option to participate in learning about and taking with them nutritious and delicious recipes including the ingredients to their dorm rooms. By comparing students with similar demographics in the same dorms who did not participate in FYE to students who participated, we wanted to note any significant results between the two populations in terms of their academic success measured by their GPA scores as well as the retention rate along with their mental health which was measured by two surveys: The Perceived Stress Scale (PSS) and the University Belongingness Questionnaire (UBQ). Both last academic year's batch of FYE students and this academic year's batch of FYE students have been investigated in this project. Comparing the midterm GPA scores for last year's students FYE vs. NFYE (Non-First Year Eats) have shown that FYE students showed a statistically significant

increase from their Fall 2019 Midterm GPA scores to their Spring 2020 Midterm GPA scores while NFYE students did not show any significant increase to their GPA scores within the same time. For students this year, results have shown that those who participated in FYE have shown a statistically significant and remarkable increase to their GPA scores as well as their mental health compared to those who did not participate once the effects of the program could be seen after a full semester of the program. Some future areas of interest for study regarding the program would be to further investigate other effects outside of GPA scores and mental health.

ACKNOWLEDGEMENTS

Contributors

I would like to thank my faculty advisors, Dr. Alan Dabney and Dr. Sumana Datta, for their inputs and encouragement and trusting me to help with their projects. I have learned so much from them and feel very honored to have been part of their team.

Thanks also go to my academic advisor, Alyssa Brigham, who has led me to the existence of this project in the first place while answering all questions that I had while attending Texas A&M University.

Finally, thanks to all my family, friends, and professors who have allowed my stay at Texas A&M University to be so special and always reminding me of the success that awaits with enough hard work and patience. It is through them that I can have these opportunities and pursue my academic interests.

All work conducted for the thesis was completed by the student independently.

Funding Sources

Undergraduate research was supported by an Innovations(X) grant and donations from the Association of Former Students at Texas A&M University. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the School of Innovation.

NOMENCLATURE

FYE	First Year Eats
NFYE	Non-First Year Eats
PSS	Perceived Stress Scale
UBQ	University Belongingness Questionnaire

1. INTRODUCTION

1.1 Food Insecurity and Other Issues Across US University Campuses

There have been many studies explaining how mental health affects academic performance as well as how a sense of belonging affects academic performance, both of which the First Year Eats Program hopes to address. Food insecurity has been an issue that "32.9% to 50.9% of college students in the U.S. experience" (Nikolaus et al., 2019). Unfortunately for such a major issue, it appears that not many "evidence-based programs [exist] to address food insecurity in college students" even when it has been shown that food insecurity has statistically significant negative effects on student academic performance (Van Woerden et al., 2019) and thus, the "opportunity remains to conduct intervention-specific research targeted at [reducing] food insecurity prevalence on college campuses." (Davis et al., 2020)

In regards to mental health, a sense of belonging on campus as well as having a community to rely on has been shown to contribute to a steady increase in the college retention rate for students. One study used the Social Belonging Index (SBI) and tracked the students' retention rate to show that approximately 99% of those who had a high SBI were retained while 84% of those who had a low SBI were retained (Davis et al., 2019). It is known that one of the most common causes for students to leave the school is because they do not feel like they belong at their school or environment (Pratt et al., 2017). Similar to the situation above regarding food insecurity, not much research has been done on a university level on belongingness while studies for K-12 students have been done extensively (Slaten et al., 2018). There have also been statistically significant relationships between some stress factors and academic performance where conscientiousness, agreeableness, and task focus factors have had positive correlations to

academic performance. Students who have higher levels of these positive factors have been found to be successful in terms of their academic performance (Saklofske et al., 2012).

1.2 The Purpose of The First Year Eats Program

The First Year Eats Program (FYE) has been in effect for one year and has already shown a noticeable trend amongst participants regarding their academic success in that students from disadvantaged backgrounds who participated in the program compared to those who did not were able to "close the gap" between themselves and more advantaged students in terms of GPA.

However, there are still some more factors that we would like for the program to detect and address regarding mental health and community. By allowing students to interact with and meet new people, some of whom they may befriend, while also not having to worry about food insecurity, we hope to see an increase in not only academic performance but also in mental health. It is our hope that this program will show significant positive results, expanding the program to other dormitories on TAMU, and perhaps even see this program become distributed to other campuses.

2. METHODS

The FYE Program has been designed to combat food insecurity which may be a factor in lower academic performance, lower level of sense of belonging in college, as well as a higher level of stress. Further research has shown there may be correlation present among these variables and analyzing how they affect each other is of interest for the project. Ultimately, it is important to note any statistically significant results and form careful conclusions on the impacts that the FYE Program has on its participants.

To begin with the project, it was important to decide on surveys that would best measure the mental health of both FYE and NFYE students at select dormitories in Texas A&M University. This would allow for the collection of quantifiable data regarding mental health. Next, academic success would be obtained through databases from the university to obtain GPA scores as well as retention rates for past year's students along with the current year's students. Thus, this project will use data from both the academic years 2019-2020 and 2020-2021.

2.1 Finding the Appropriate Surveys to Measure Mental Health

The very first step of the project was to propose questionnaire templates assessing student mental health and to decide which would be most appropriate to use for Texas A&M University first-year university students specifically. Once likely candidates were chosen, the questionnaires were planned to be distributed to both FYE and NFYE students at the beginning of the semester as well as the beginning of the next semester for comparison.

The two candidates that would serve to measure stress levels were either the Perceived Stress Scale (PSS) or the College Student Stress Scale (CSSS). We were able to discuss with Professor Sarah LeMire, a librarian and professor with knowledge of databases and surveys, the

positives and negatives of both surveys and which survey would be more appropriate to use in our case. After careful consideration and discussion, it was decided that the PSS would be the main scale to use to measure stress levels as the PSS has been determined to be more reliable because it has been used extensively and proven to accurately measure stress levels on a consistent level across time even up to recent years. The PSS is a survey in the form of ten questions that all commence with "How often …?" and measures student response from a Likert Scale from 1 to 5 whereas 1 would represent "Never" and 5 would represent "Very Often" (Cohen, 1983). Since questions varied whether a higher number would be a negative or positive response, we made sure to make it consistent whereas a higher number would be a positive response in any case by flipping the scale for questions that asked for negative responses.

The next measurement of interest was the emotional level of students and in particular, the sense of belonging on campus. While there were many more candidates for this topic, it was decided that the questionnaire that would be used would be the University Belongingness Questionnaire (UBQ) with the help of Professor LeMire once again. The UBQ measures how much students feel belonging to or part of the university and culture in three sections: University Affiliation, University Support and Acceptance, and Faculty and Staff Relations. Once again, the student response is measured on a Likert Scale but this time from 1 – 4 where 1 would represent "Strongly Disagree" and 5 would represent "Strongly Agree" (Slaten et al., 2018). No adjustment of scores was necessary as all questions with higher scores would result in a higher sense of belonging.

To obtain as many samples as possible, directories containing the entire FYE population as well as the NFYE students within the same Learning Communities were used to email the two surveys selected along with an incentive to take both surveys. It was important to decide the time

frame at which the datapoints would be measured as we would like to avoid bias if possible. The first set of datapoints would have to be taken before the FYE program could take effect but also not too early as to avoid initial stress students would have with moving in and preparing for classes for the first time. Thus, we decided to release both surveys two weeks into the first semester with a window of a week and a half for students to complete the surveys. Once the surveys were closed, all results were taken and organized.

The surveys would then be reused for the next semester in which at this point, the FYE program would have had time to have an effect in terms of both stress levels and university belongingness among participants if any. As usual, the timing had to be decided carefully to remove bias such as extra stress resulting from a new semester and adjusting to classes once again. The time window would be around a week and a half and once the surveys closed, and all the datapoints were organized and added to the database containing the rest of the students' demographic data including the previous semester's survey results.

2.2 Measuring Academic Performance

The second part of our desired measurement, academic performance, could be measured in terms of student success such as midterm and final grades that will also be tracked and analyzed along with the questionnaire results. A database containing all information of students including their demographics data along with their midterm and final GPAs was provided to us whereas we could then combine all data including the survey results.

Another important method of data collection was to create a new category of classifying race and ethnicity to determine underrepresented minority status. As there are an abundant number of categories within the race column and too little information on ethnicity, it was determined that combining the two categories into a new category: underrepresented minority

status was appropriate. The conditions to categorize students into this category was for students to be classified in terms of ethnicity as “Hispanic or Latino” or if they were classified as “Not Hispanic or Latino,” then their race would have to be one of the following: “Black,” “Native American or Islander,” or “Two or More Races.”

2.3 Collecting and Analyzing Last Year’s Group of Students

It was also of interest to see the academic performance of last year’s students so while waiting for the survey data to finish being collected for the current year’s students, the data of last year’s students were analyzed. Due to the worldwide pandemic, COVID-19, it was important to take this into consideration and change the method of analysis for last year’s students as the pandemic officially affected classes and scheduling in the middle of the Spring 2020 Semester. Originally, it was planned to compare the GPA of the final grades between Fall 2019 and Spring 2020 but instead, we opted to compare the GPA of the midterm grades instead because there would be no bias from the pandemic.

In addition to comparing the GPA scores, retention rates of these students into the current academic year were analyzed as well. We wanted to note any statistically significant differences between FYE participants and NFYE students regarding the ratio of students who returned to campus and were present until the end of the semester. The two definitions that we used to determine whether a student was retained or not were to consider them as retained if they were present for the 12^{th} day of classes for the semester or if their final grade for that semester was reported.

2.4 Using R for Various Data Organization and Analysis

The software R is a statistical tool that allows for both data organization and analysis. For example, R would be able to combine two datasets with a common column name. We had two

datasets where one would have both the demographics and GPA information while the other would have survey results. Both datasets had the student identification ID category to which R used this column to combine the two datasets.

R would also then be used to conduct all the main data analysis using various tools to do so including linear regression analysis, two-sample t-test, hierarchical clustering, and more for both last year's data and this year's data. A linear regression model that was used to attempt to predict GPA scores for this year's students included variables gender, underrepresented status, FYE status, First Generation status, family income level, and the two survey results. Other two linear regression models performed were to either only look at the two survey results or split the UBQ section into three categories. This would allow us to see if any initial measurements of the two surveys standing alone would be an indicator of predicting GPA scores with no other variables present. The model for the first regression model containing all factors is represented as Equation 2.1

$$y_i = \beta_{0i} + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_7 x_{7i} + \varepsilon_i \qquad (2.1)$$

where y_i, the response variable, represents the GPA score predicted, β_0 is the intercept, β_1 is the coefficient of the gender effect, β_2 is the coefficient of the underrepresented status, and so on with ε_i being the error term that results naturally from each datapoint coming from independent sampling. Each x term is unique to each datapoint and its related coefficient while simply representing an individual student's gender, status, survey score, etc.

The second linear regression if only the two main average scores of the PSS and UBQ were used and not split into three sections would be Equation 2.2

$$y_i = \beta_{0i} + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i \qquad (2.2)$$

where this time, y_i, the response variable, represents the GPA score predicted once again, β_0 is the intercept, β_1 is the coefficient of the PSS effect, β_2 is the coefficient of the UBQ effect, and ε_i being the error term.

The general idea with the usage of linear models is to approximate and determine the magnitude and significance of how each factor affects the response variable assuming all other variables are held constant. Therefore, we must observe each variable and its coefficient separately which will be explained in greater detail in the following section.

3. RESULTS

As one part of the data was being collected, another section of data would be able to be analyzed. The first completed section was the initial mental health survey results.

3.1 Mental Health Survey Results

Preliminary results of the two surveys were collected after the first two weeks of the fall semester. It was important to collect the data before the program could take effect to determine whether the two groups of students (FYE vs. NFYE) were significantly different from each other so that we can assume fair testing to avoid biases in future reruns of the surveys with these students. To do this, the data was first divided into two groups: FYE and NFYE students. Then, each section was compared to each other by the mean and standard deviation. A two-sample t-test was conducted to determine any significant differences between the two groups regarding both the sample size and data where we obtained a p-value greater than 0.05, meaning that we failed to conclude that the two groups had different means. In addition, we used R to create a histogram for both groups to compare the shape.

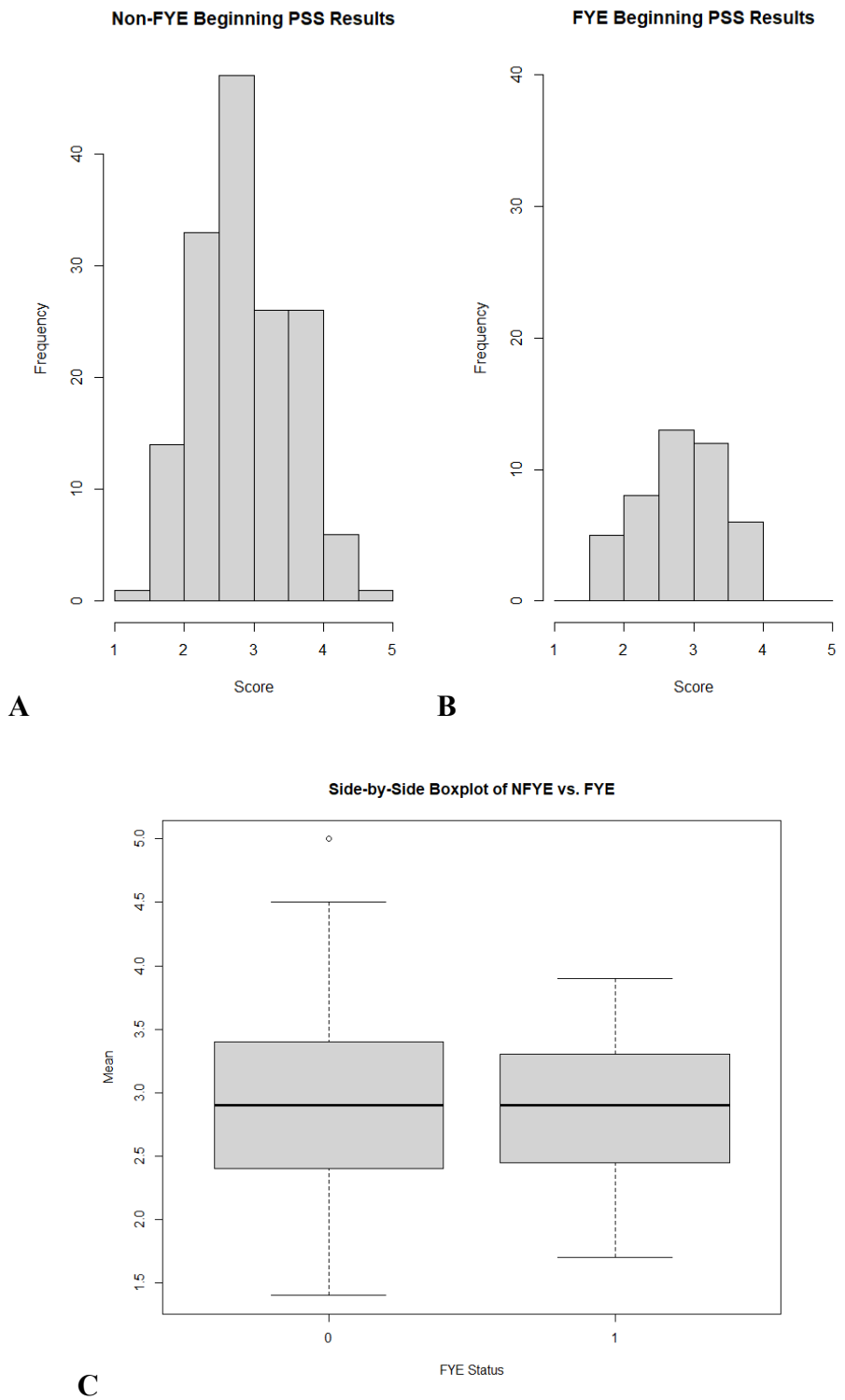


Figure 3.1: A) Histogram of NFYE PSS Results B) Histogram of FYE PSS Results C) Boxplots of Both NFYE and FYE PSS Results

Figure 3.1 A-C above shows us that the PSS results have roughly the same results and are similarly distributed. A common test that is used to test normality is the Shapiro-Wilk Test where datapoints are taken and fit into a normal distribution using the goodness-of-fit test. A p-value of less than 0.05 would indicate that the dataset strays from the normal distribution at 95% confidence. Both the FYE and NFYE results have indicated a p-value of greater than 0.05, so the assumption of normality is consistent with our data, at 95% confidence. It also shows that the two boxplots are nearly identical. Thus, with all information gathered from above, it seems that we can conclude that the two groups have a similar starting point in terms of stress levels.

The same process has been repeated for the UBQ results and the histograms for the two groups are shown below:

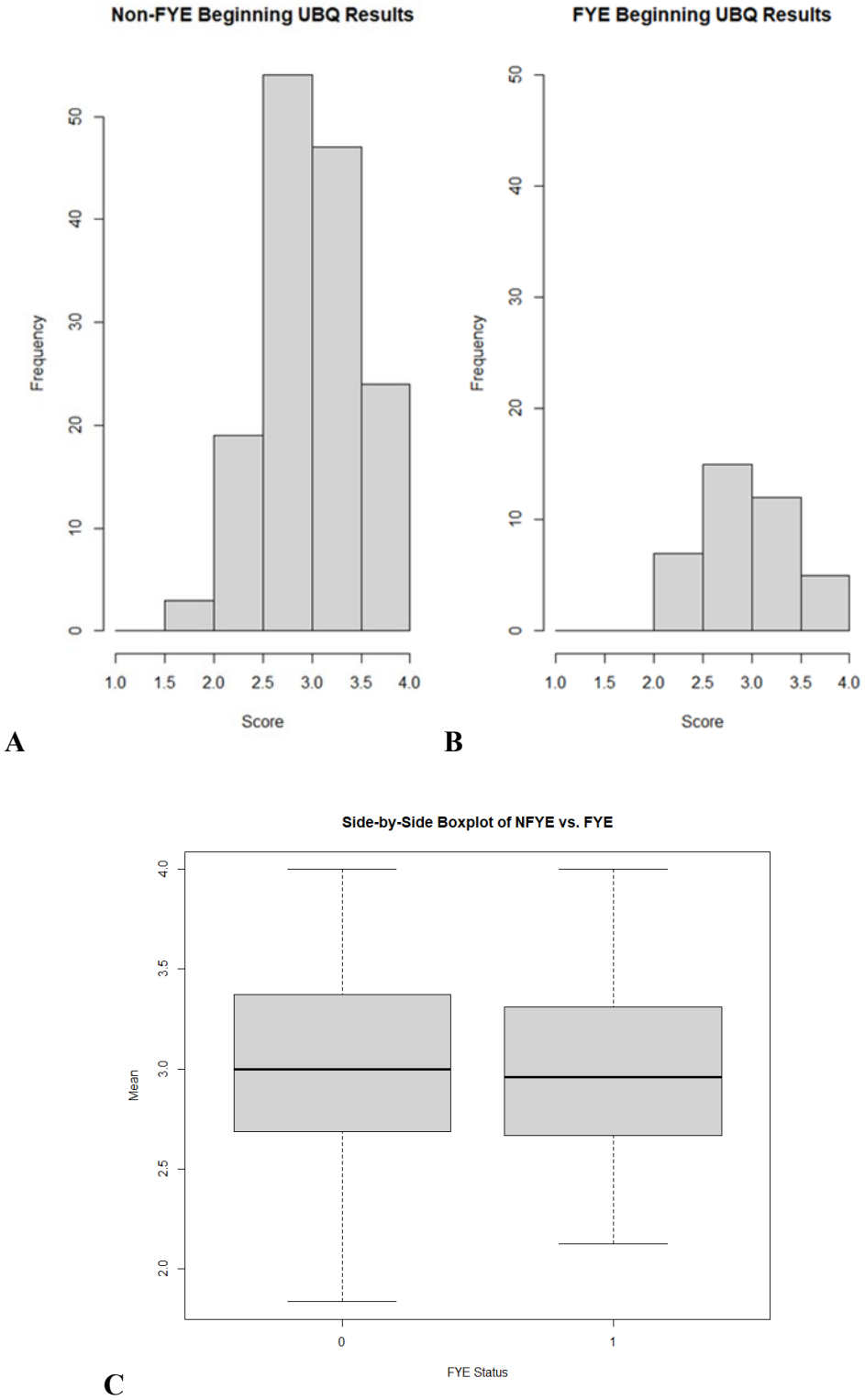
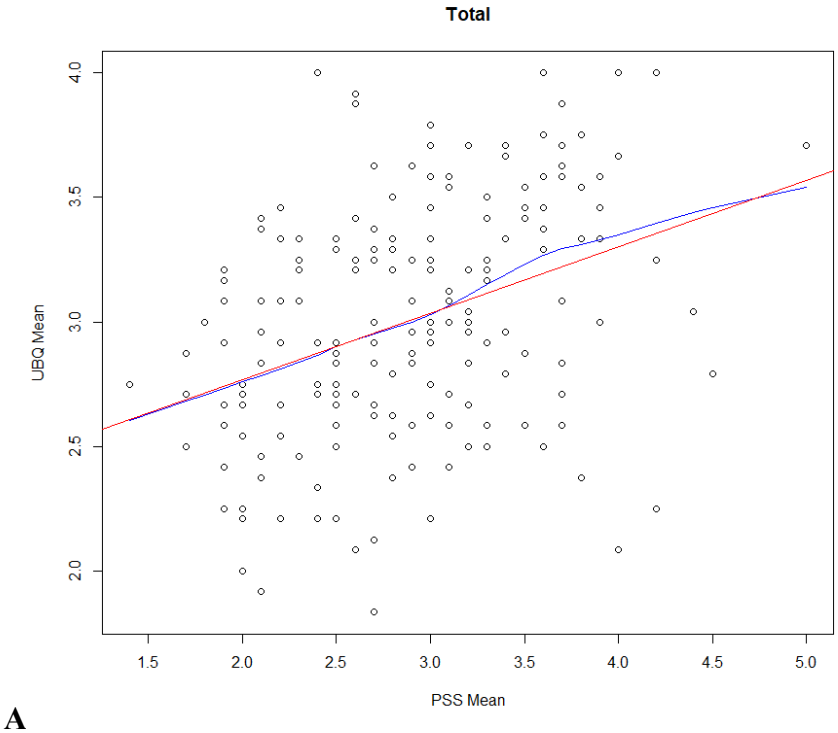


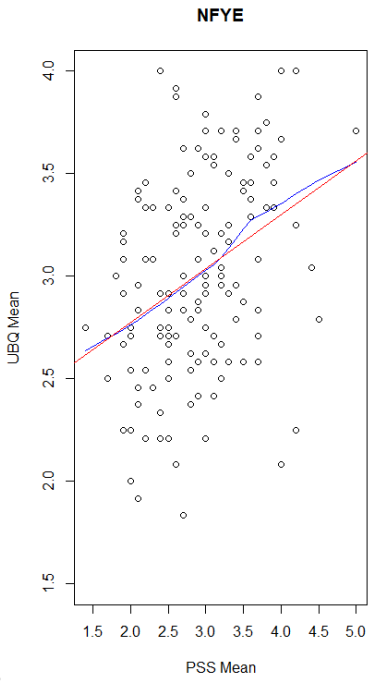
Figure 3.2: A) Histogram of NFYE UBQ Results B) Histogram of FYE UBQ Results C) Boxplots of Both NFYE and FYE UBQ Results

Figure 3.2 A-C once again shows us that the shapes are similar along with the boxplots. Tests of normality and significant difference in means between the two groups have also shown the same results as the PSS survey so we can likely conclude the two groups do not differ preliminarily in the sense of belonging on campus.

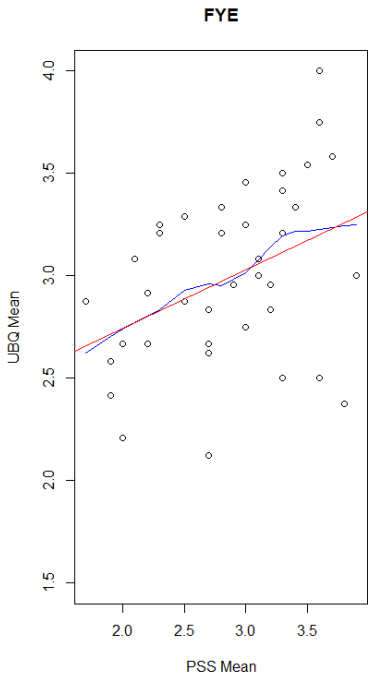
The next set of observations are to decide if there is any significant correlation between the two survey results. In other words, we want to answer if students scoring “better” on stress scales had higher scores for belongingness and vice-versa. The graph below is the result:



A



B



C

Figure 3.3: A) Correlation Graph of Combined NFYE and FYE B) Correlation Graph of NFYE C) Correlation Graph of FYE

We can see from Figure 3.3 A-C that there is a positive albeit rather weak trend between the two scores. The correlation coefficient value, R^2, is 0.372 which is slightly weak but still a noticeable trend. We also divide this into the two groups again to see if the trend remains.

Table 3.1 confirms the trend, and we can observe that because of the smaller sample size for FYE, the 95% Confidence Interval is much wider than the other two making for a slightly unreliable trend.

Table 3.1: Correlation Coefficient and Their 95% Confidence Interval

Category	Correlation Coefficient	95% C.I.
Total	0.372	(0.240, 0.490)
NFYE	0.365	(0.215, 0.499)
FYE	0.401	(0.093, 0.639)

3.2 Regression Modeling for GPA

Returning to the regression modeling mentioned in the last section, we use R to fit a predictive model with GPA being the response variable. The full model for the first regression model is as follows:

Table 3.2: Regression Model Factors and Their Effects

Factor	Factor Effect	p-value
Intercept	3.275	<0.0001
Gender	-0.074	0.6754
Underrepresented	-0.811	<0.0001
FYE Status	0.310	0.0569
First Generation Status	-0.500	0.0337
Income Level	-0.074	0.2622
PSS	0.074	0.5152
UBQ	0.108	0.4781

The interpretation of Table 3.2 is that the factor effects mention the rough magnitude of effect that each factor has on predicting the GPA for the midterm. The p-value represents how significant the factor is, so a lower p-value would represent a more significant factor. The results show us that there are three statistically significant factors with 90% confidence: Underrepresented, FYE Status, and First Generation Status.

According to a study done collecting data from 3,245 university students, food insecurity and GPA scores were also affected by ethnicity, Pell Grant-eligibility, and first-generation status, so our results seem to be relatively consistent to this study as well (Camelo et al., 2019). The results were also telling in that after only a month and a half of the program, FYE students already seemed to have a higher GPA score compared to NFYE students. On the other hand, it seems that the two survey results were not significant factors in predicting GPA scores.

3.3 Analyzing Last Year's Students

The first component of analyzing results from last year's students is to compare the GPA changes between the Fall 2019 to Spring 2020 semesters for both FYE and NFYE students as illustrated below:

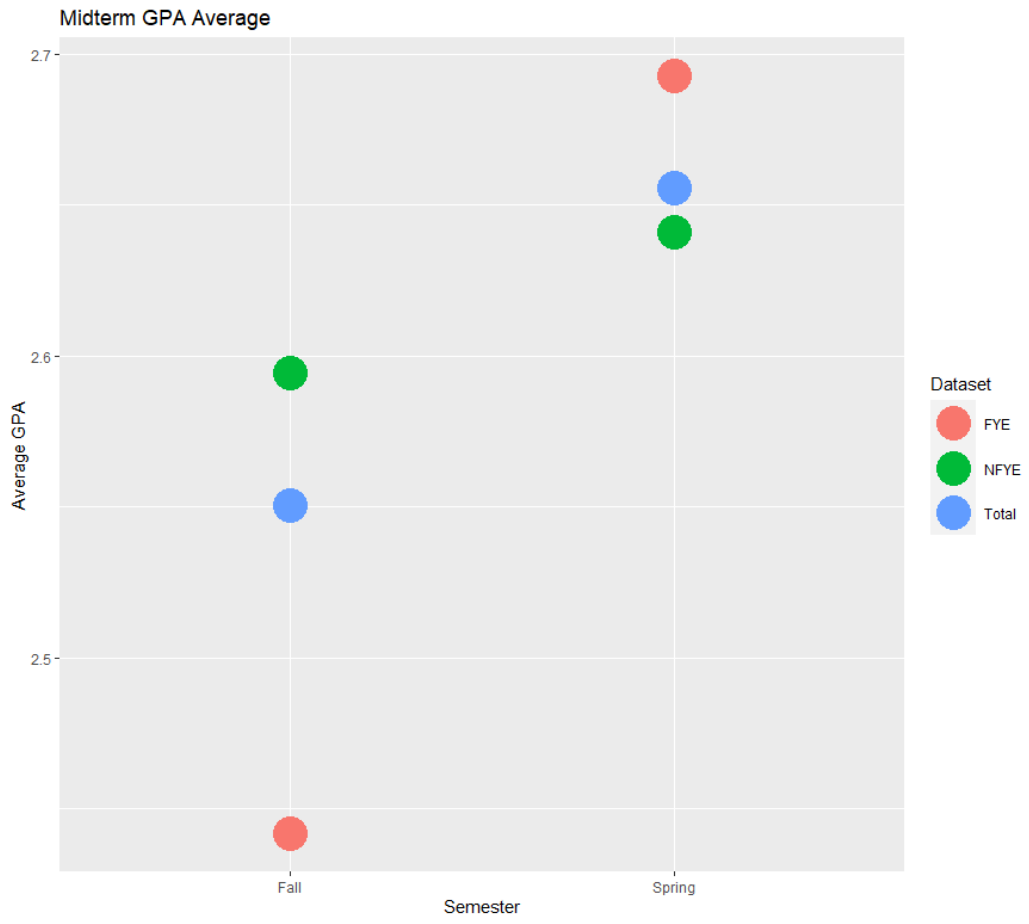


Figure 3.4: Midterm GPA Averages for FYE and NFYE

As Figure 3.4 above shows, it seems that originally, FYE students had a lower performance compared to NFYE in the Fall but not only managed to catch up but even perform better than NFYE students in the Spring. However, this graph does not tell the full story and further analysis is required so we performed a population means comparison and, in that test, we get a result that indicates this observation not being statistically significant. This again reminds

us that what meets the eye is not necessarily what is happening when statistical methods are involved to scientifically prove significance.

However, when comparing GPA performance by only looking at the changes in FYE, it nets a statistically significant result in which there is a noticeable GPA increase for FYE students between the fall and spring semesters but the same could not be said for NFYE students. Thus, in this regard, the program has had a statistically significant effect in the academic success of its students.

The next set of analysis involves stratification and any potential patterns that can be noticed by splitting the data into groups. The first stratification was with gender and whether any significant patterns could be seen by dividing and analyzing students by gender. Quick test runs have shown that stratifying by gender have netted no significant results and regression models from last year's study has shown no significant factor effects from gender which agrees with the result obtained.

Another stratification group was by college and if certain colleges were shown to have better GPA performance:

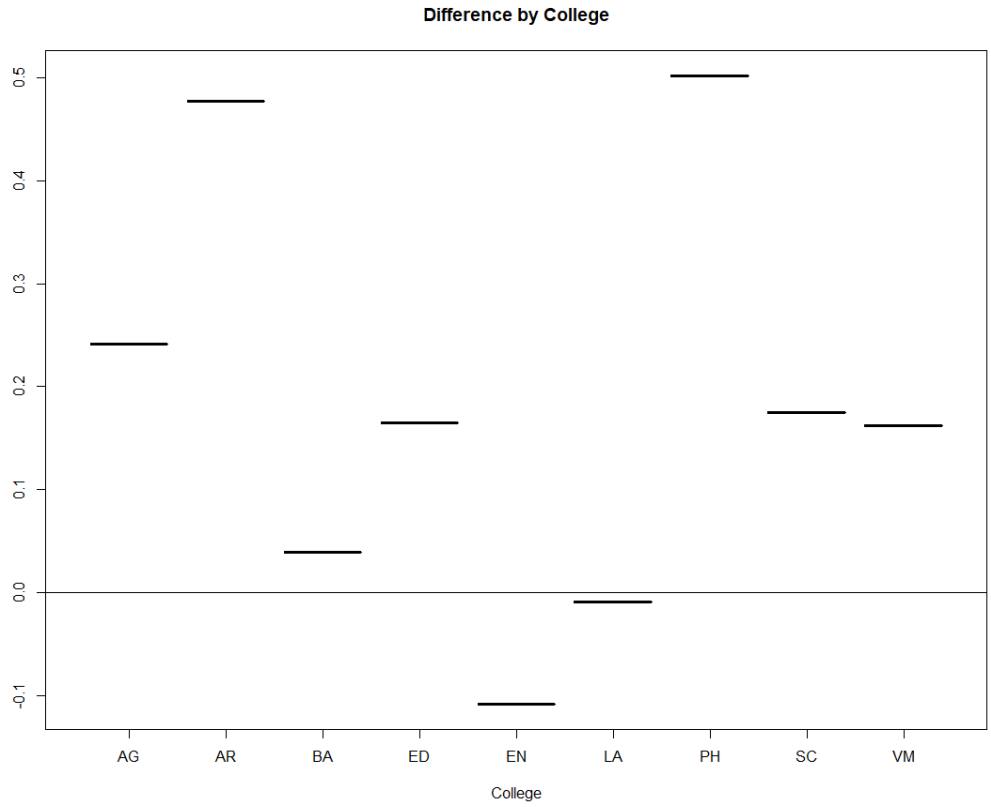


Figure 3.5: Stratifying by College for GPA Difference

The datapoints represented by lines in Figure 3.5 indicate the result of Spring 2020 minus the Fall 2019 GPA scores. Thus, everything above the line at 0 indicates an increase to their GPA in the Spring compared to the Fall. Error bars were not represented as differences between colleges were not significant for the most part and provides no extra information but only serves to clutter the chart.

For a more in-depth analysis, analysis of variance (ANOVA), was performed for the colleges and their differences and a Tukey's Honest Significant Difference (HSD) Test was conducted to find significant mean differences between any pair of colleges. Two pairs of differences that were significant at the 90% level according to this test were Engineering and Agriculture as well as Engineering and Veterinary Medicine. When doing individual two-means

comparison test for Engineering and other departments, it seemed that engineering students significantly differed in their means compared to the other departments. This is also consistent with the results gained from last year where Engineering was one of the significant factors of determining GPA scores for last year's students.

Next, a look into clusters produced by R using hierarchical clustering was done to find any other previously unsuspected clusters that may be present in factors or combination of factors that could have been overlooked initially from doing exploratory analysis. Hierarchical clustering utilizes any continuous numeric factors and attempts to group them into a specified number of clusters based on how similar their collective means are to each other. In other words, it simply tries to conjoin seemingly unrelated points together to hopefully find a trend when looking at each group and their patterns. First, it is important to determine the recommended number of clusters by using a scree plot to view the differences between each cluster or how impactful splitting the data into an arbitrary number of clusters would be using this chart below:

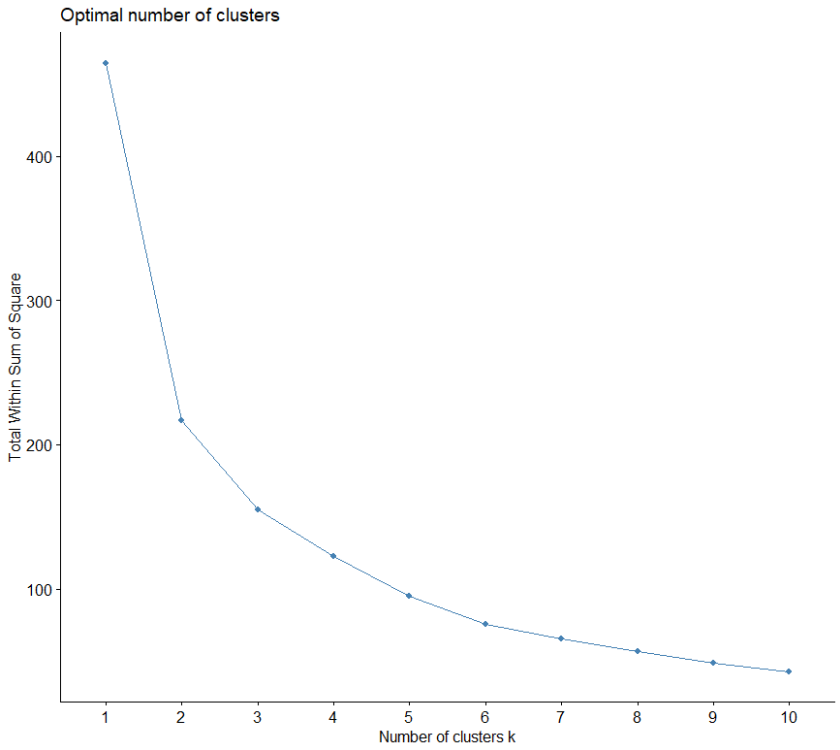
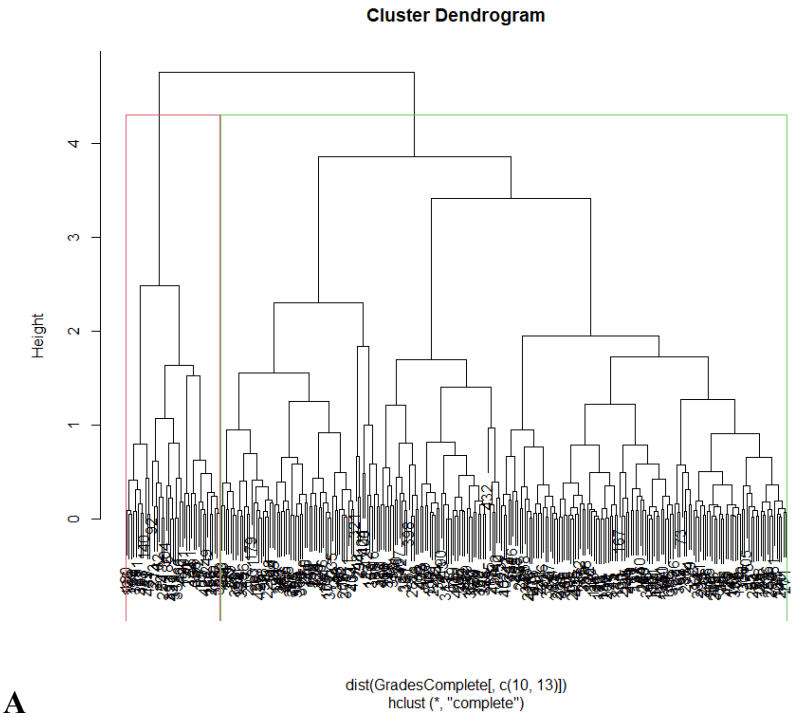


Figure 3.6: Determining Optimal Number of Clusters

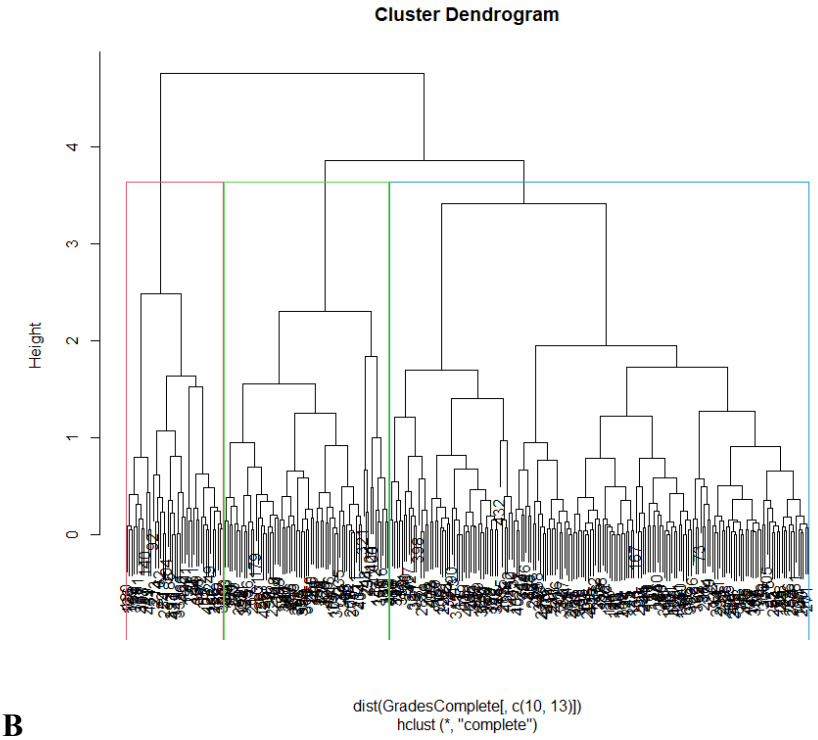
It was determined from Figure 3.6 that after two or three clusters, the effectiveness of dividing clusters into even greater values would be unnecessary and the Principle of Parsimony was taken into account where "simple is better."

This allowed us to decide the number of clusters to be either two or three and run the program in accordance. There are three different main methods of linkages in how R will divide the clusters. However, one of the three, "single," has shown to be obsolete for this data as it fails to divide data into three clusters. Thus, the default, "complete," as well as the other option, "average," were pursued. Further hierarchical clustering analysis was conducted using only the default linkage method, "complete," for a more general interpretation as well as because it is the recommended method for analysis. However, it would also be interesting to investigate further

with the “average” method later. Asking the program to visually split the data into a dendrogram, we receive this:



A



B

Figure 3.7: A) Dendrogram of Two Groups B) Dendrogram of Three Groups

Since every data point was split into different groups in Figure 3.7 A-B, it is rather difficult to see which specific groups contain which student grades. Therefore, we plotted the individual groups containing the students to see a more clear division:

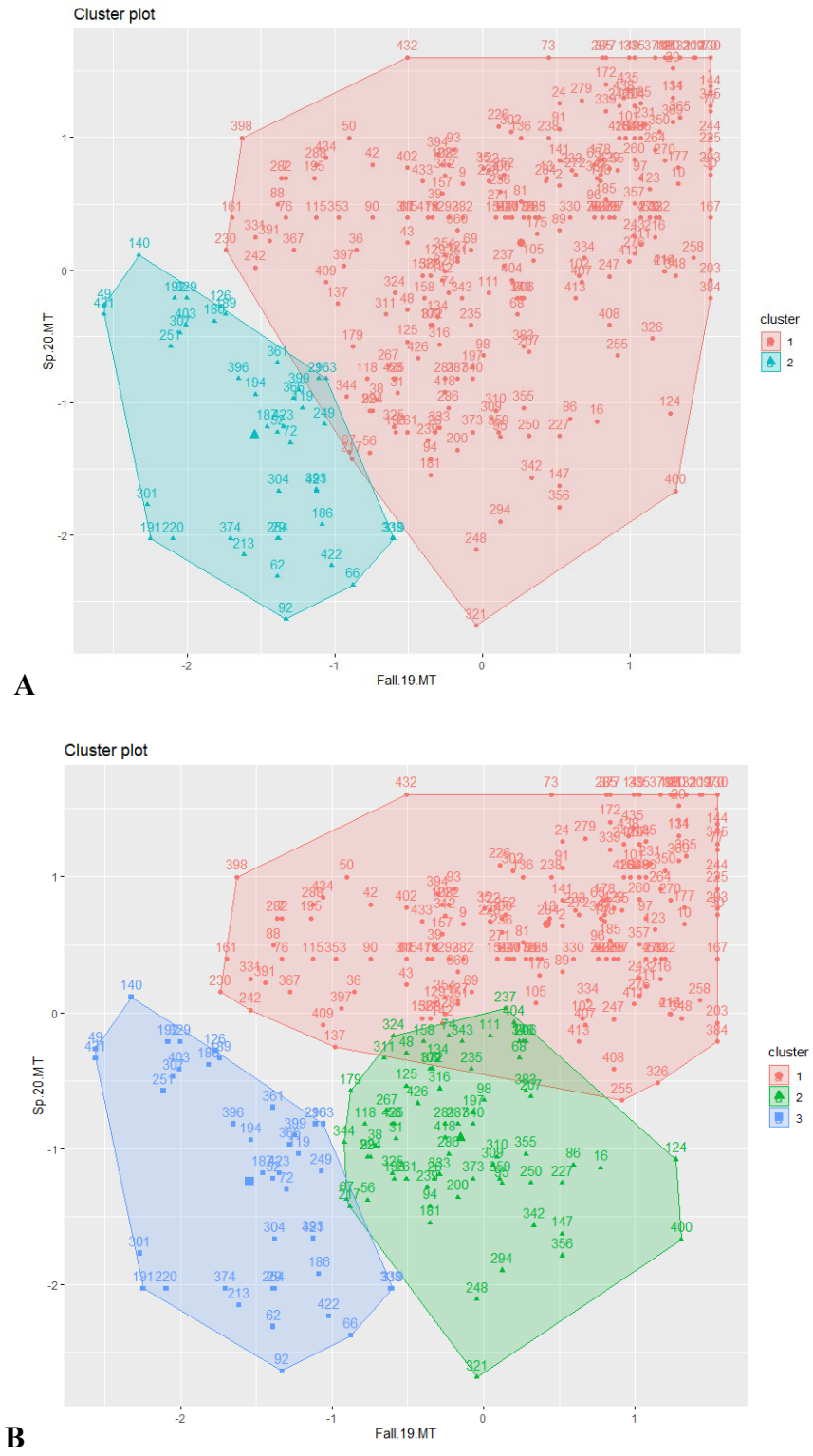
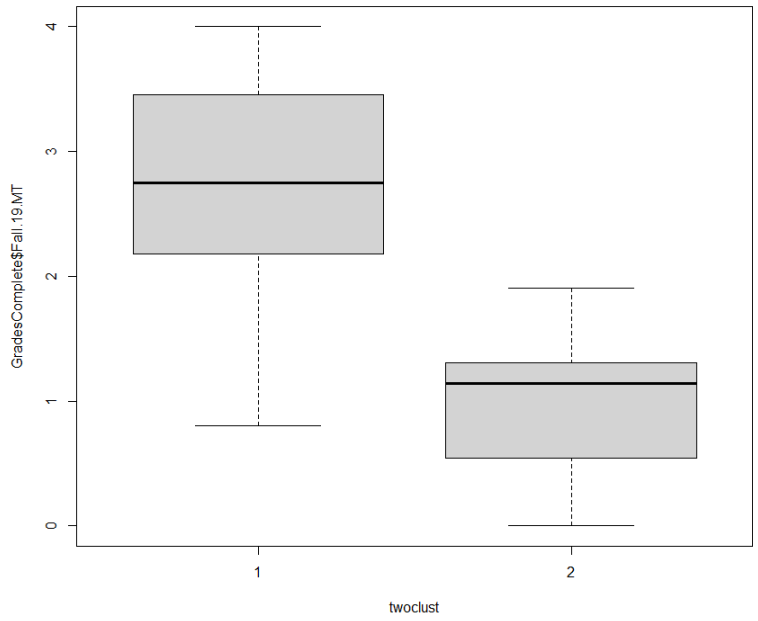


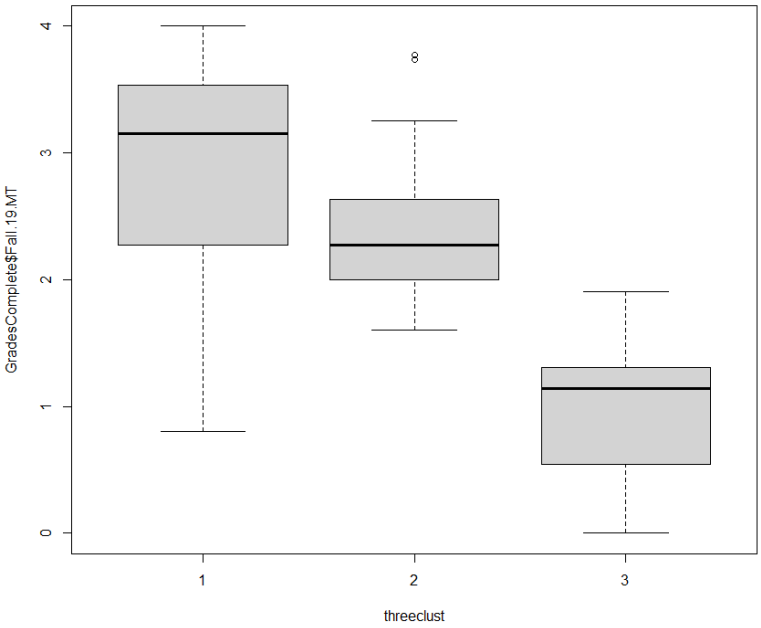
Figure 3.8: A) Cluster Plot of Two Groups B) Cluster Plot of Three Groups

As we can see from Figure 3.8 A-B, it seems that the larger of the two groups has been split into two itself when requesting for an additional split while the second, smaller group retains its shape and size.

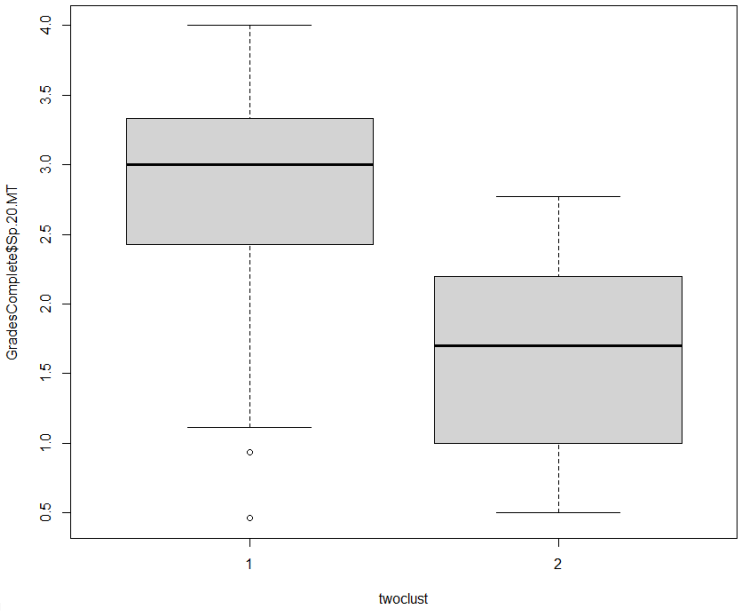
Since the only two continuous numeric variables available were the two midterm grades, the program determined clusters based on how low or high as well as how much individual grades changed between the semesters and grouped them together based on how similar these values were to each other:


GradesComplete$Fall.19.MT
0
1
2
3
4
1
2
twoclust

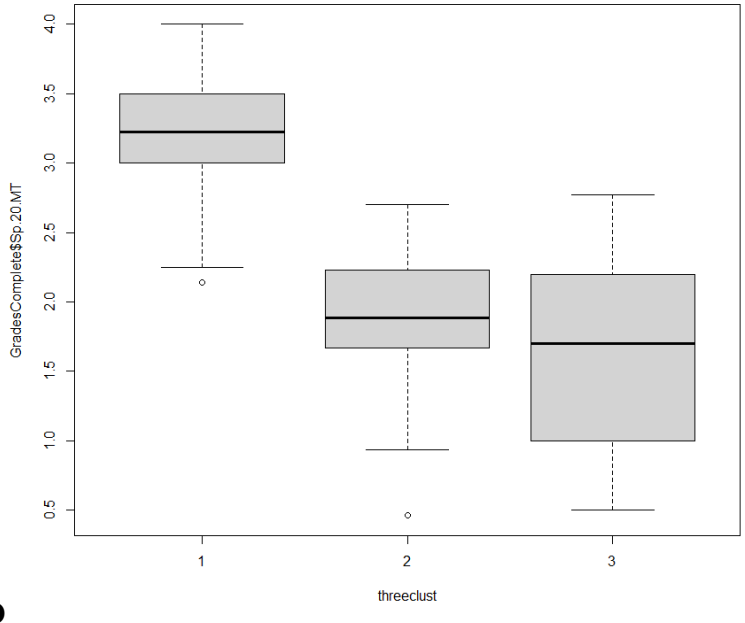
A


GradesComplete$Fall.19.MT
0
1
2
3
4
8
1
2
3
threeclust

B



C



D

Figure 3.9: A) Fall 2019 Midterm Two Clusters B) Fall 2019 Midterm Three Clusters C) Spring 2020 Midterm Two Clusters D) Spring 2020 Midterm Three Clusters

We can see that higher grades were seen in the former clusters vs the latter from observing both sets of graphs in Figure 3.9 A-D.

It is still unclear what kinds of trends or patterns are hidden. However, R allows for clear data viewing using tabulation so we can explore other variables using this clustering data. Individual data points were labeled with their respective cluster number and the new row columns were added onto the original dataset containing every student's demographic information. Tabulating by Family Income Status gave us the following results:

Table 3.3: Tabulating by Family Income Level

		Two-Cluster Groups			**Three-Cluster Groups**			
		1	**2**	**Total**	**1**	**2**	**3**	**Total**
Family Income Level	**0**	107	19	*126*	73	36	19	*126*
	1	36	9	*45*	29	6	9	*45*
	2	46	7	*53*	34	11	7	*53*
	3	44	4	*48*	33	11	4	*48*
	4	12	2	*14*	7	5	2	*144*
	Total	*245*	*41*		*176*	*69*	*41*	

We can see that overall, there were no noticeable differences in each group compared to another group when viewing Table 3.3. We cannot find any patterns worthy of note looking at the clusters in terms of family income level. A noticeable difference for example would be if we had a significant ratio of one family income level in a cluster but a completely different ratio noticed in another cluster.

Next, the college that the student is attending was considered and was also tabulated into the following:

Table 3.4: Tabulating by College

		Two-Cluster Groups			Three-Cluster Groups			
		1	**2**	**Total**	**1**	**2**	**3**	**Total**
College	**AG**	61	12	*73*	37	24	12	*73*
	AR	4	0	*4*	3	1	0	*4*
	BA	3	0	*3*	3	1	0	*4*
	ED	**12**	**2**	*12*	9	1	3	*13*
	EN	**50**	**1**	*51*	**41**	**10**	**0**	*51*
	GS	1	0	*1*	1	0	0	*1*
	LA	**14**	**1**	*15*	11	3	1	*15*
	PH	13	3	*16*	9	4	3	*16*
	SC	25	6	*31*	21	4	6	*31*
	VM	62	16	*78*	41	21	16	*78*
	Total	*245*	*41*		*176*	*69*	*41*	

This time in Table 3.4, there seems to be a more noticeable difference in Engineering and to a lesser extent in Education and Liberal Arts within the clusters. This is interesting to note because again, it was determined that last year, Engineering was an important statistically significant factor in predicting GPA scores of these same students, so this helps confirm that being in the College of Engineering is significant in determining grade in the case that we assume all other factors are disregarded.

Similar tabulation processes were performed for gender and first-generation status, but no significant differences were found. While performing hierarchical performance did not yield many new details to add to our analysis, it was important to use as it will give us further support or give reason to not support information we had before.

Next, a brief analysis in Principal Components was done. The purpose of Principal Components Analysis (PCA) is to reduce the number of continuous numeric factors into one and thus reduce unnecessary complexity in our model. Since once again, we only have access to two such factors, we would only be reducing two variable coefficients to one, but this will become

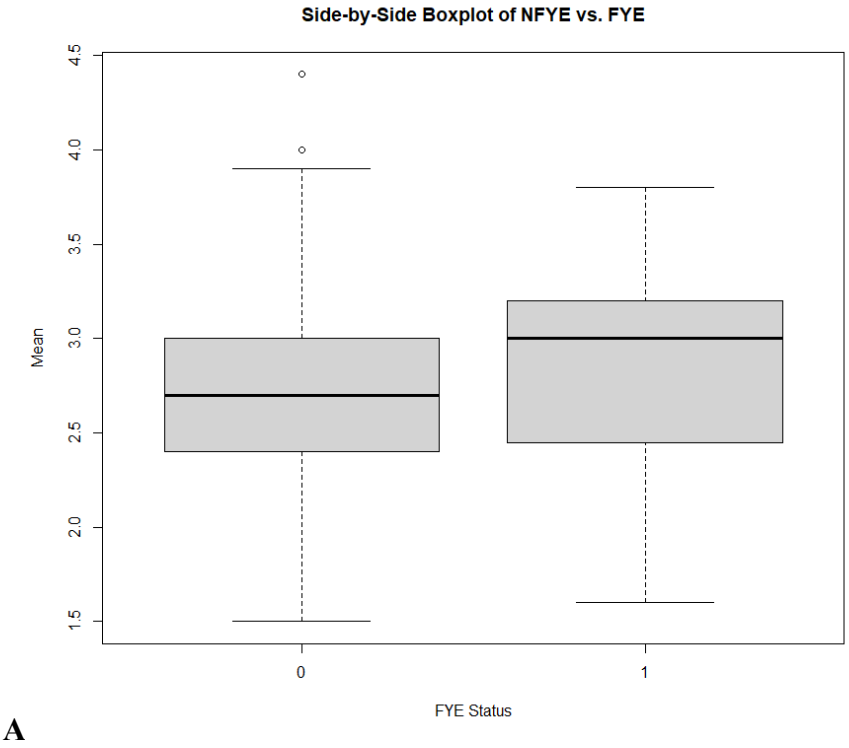
significantly more useful when other factors are added in during a future instance. The results have shown that the fall midterm portion has given 0.806 while the spring midterm portion has given 0.592. These numbers simply represent the magnitude of each factor, but another important value is the proportion of variance that each component gives or in other words, how much weight each portion has. The proportion of variance explained using the first column of data above (PCA1) is 77.5% while the second column which is not provided (PCA2) is 22.5%. Thus, PCA1 has much more weight than PCA2, so that we can use these coefficients when utilizing them for regression modeling.

Finally, retention rates were observed as these students have now been able to enter and complete their first semester after they have finished participation with the FYE Program. Two separate analyses were made with the first considering retention to be whether the student has attended their 12^{th} day of the new semester wherein out of the 122 FYE students, 119 have and out of the 108 NFYE students, 103 have. At first glance, it may seem that FYE students are clearly being retained better but running the Fischer Exact Test that helps determine any significant differences between proportions by considering both ratio and sample size for greatest power in reporting a conclusion has shown a result with a p-value of 0.4795 which is nowhere near significant unfortunately.

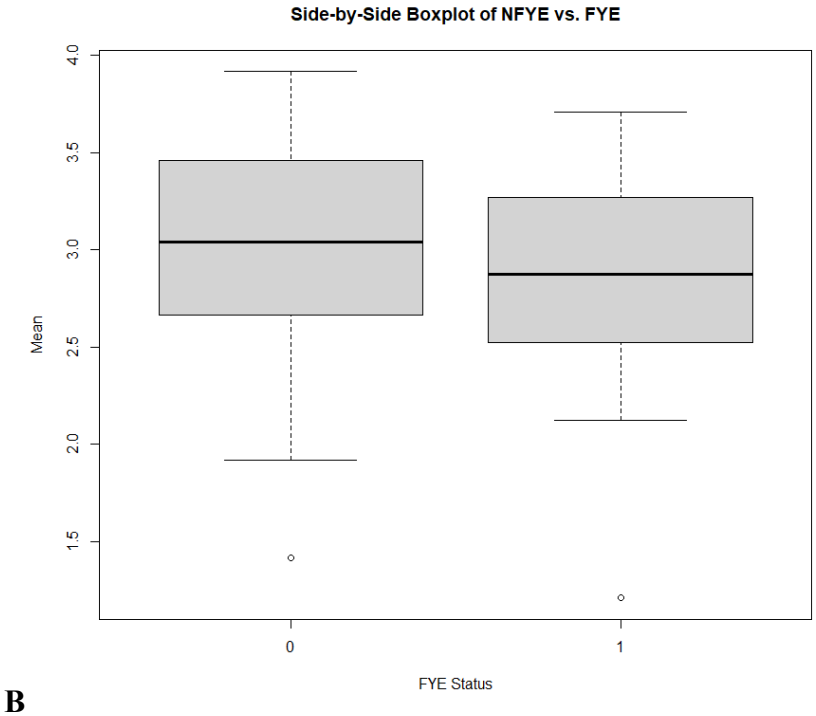
However, using the second definition of retention where Fall Final semester grades must be reported in order to be considered retained has shown different results wherein 120 out of 122 FYE students were considered to be retained and 101 out of 108 NFYE students were retained. This time, the p-value is 0.0871 which is significant with 90% confidence in concluding that FYE students have been more retained than NFYE students.

3.4 Return to the Mental Health Surveys

Returning to the mental health surveys sent out at the beginning of this year's fall semester, we sent out these same surveys again and results were collected for the current year's students. More analysis was done in terms of both how mental health changed over time and how if any the stress scales may be correlated to predicting GPA for this year's students:



A



B

Figure 3.10: A) Boxplot of New PSS Results B) Boxplot of New UBQ Results

Unfortunately, it seems that like the boxplots above in Figure 3.10 A-B suggest, there seems to be no significant differences between FYE and NFYE for both tests for the second round.

A series of two-sample t-tests were conducted to detect any differences of any pattern including FYE vs. NFYE in PSS and UBQ as well as differences between Fall vs. Spring for both groups for both surveys and all have suggested no significant differences. Some suggestions for such results were that both stress and belongingness can very easily become more negative as time goes by as students may have more knowledge of the pressure and having to overcome challenges that they may have noticed during their first semester. The additional stress of having to raise or maintain GPA may be another source of worry that students would not have had for their very first semester of classes. Another suggestion was that the situation of the global pandemic and its continued presence on campus not allowing for more opportunities to bond and reduce stress when most experience is limited to a computer screen with little to no human-to-human interactions could negatively impact students or not cause any changes from the beginning. This unfortunately cannot be investigated any further until the pandemic ceases.

3.5 GPA Changes for Current Students

Finally, it was of interest to compare academic performance between FYE and NFYE students of the current year. We have seen earlier looking at the midterm grades for the fall semester that there was already a significant difference favoring the FYE students compared to the NFYE students, but we also wanted to see if this trend would hold for the fall final grades. Three two-sample t-tests were performed once again to find significant differences with the results summarized in the table below:

Table 3.5: Results of two-sample t-tests Comparing FYE and NFYE Students

Comparison of Interest	**p-value**
Final FYE vs. NFYE	0.1859
FYE Midterm vs. Final	0.0220
NFYE Midterm vs. Final	<0.0001

We can see from the results in Table 3.5 that while the final grades between the FYE and NFYE students were no longer significant, FYE students still on average had higher GPA scores compared to NFYE students and both groups have improved from their midterm scores.

A possible explanation for this situation is that there may be an unmeasured "cap" or "ceiling" of how much students can improve their grades regardless of whether they were a part of the FYE program or not. Perhaps there may exist a diminishing returns situation where initial improvements in grades are quick to occur, but further improvements provided by the program may return smaller increases compared to before. Finally, there may be other unaccounted factors that we simply do not have access to for either group such as outside tutoring, seeking academic help, etc. that may affect GPA scores. These however serve as great suggestions for further research in this topic.

4. CONCLUSION

Throughout this long journey, there have been so many new observations, some fruitful while others not quite, that have given basis to the solutions for our questions at the beginning of the year. It seems overall that the program has certainly played an effect as denoted by significant changes and characteristics when comparing FYE and NFYE students in various methods.

4.1 Discussion of Results

Beginning from the results with the survey data for the current year's students, it seems that there were simply too many external factors and too few sample sizes to give proper conclusions on how the surveys played an effect in determining academic success or whether the FYE Program has been significant in affecting stress or belongingness. As both the first round and the second round of the surveys have shown no changes between FYE or NFYE students or total changes between the two semesters, not much could be concluded in this regard.

For a better explanation regarding the low sample size, while we initially had plenty of students respond nearing 200 students for each survey, numbers were heavily reduced for the second round of surveying whereas only about half responded. In addition, there were many new students that have answered the second round of surveys meaning that not the same group of students who answered the first round of surveys were necessarily present for the second round. At the end, when only observing those who answered both surveys both times, the sample size reduced to 56 which is considerably lower than what we would have ideally preferred.

As this was the first year to implement these surveys, it can be improved significantly in later years, and perhaps we may find more concrete results with a better system to survey the students so that we can maximize the sample size.

As mentioned before, unseen factors and variables may additionally hinder the conclusions that can be drawn from the survey results as the pandemic remains an ever-looming presence as well as traditional external factors such as pressure to do well in classes or the stress or worry of adjusting to a new semester after a month-long break.

Even though the survey results have not offered too much insight, it was beneficial to first having determined the two surveys that should be used as well as being able to have experience of how to distribute and collect data for future studies. However, I do believe there is room for improvement in both finding or adjusting the types of surveys needed and the logistics of conducting the surveys.

In terms of academic success in both academic years and their GPA scores and retention rates, it seems that the FYE Program had a more visible and statistically significance in this area. First, the retention rates of last year's students into this year has shown that there were significantly more students retained who were part of the FYE Program last year compared to NFYE students.

What is even more encouraging is that this is by purely observing two factors: FYE status and retention, while considering nothing else which means that other factors do not come into effect when conducting the Fisher Exact Test to find any significant differences between the two proportions of retention. In other words, it is unlikely that being a participant in the FYE Program increases retention is simply coincidence. The program has indeed given a positive effect in student retention rate whether this may be the result of former FYE participants feeling

more comfortable and knowledgeable in avoiding food insecurity while also eating nutritiously to support themselves while taking classes or otherwise.

GPA scores have also shown great prospective for FYE students in that the program has given quick results in around just a month and a half to distinguish the FYE students statistically significantly in terms of GPA where FYE students had higher midterm GPAs. While it is true that NFYE students have managed to catch up to no longer be significantly different from FYE students in Final GPA scores, FYE students have continued to show significant improvements with their GPA, retaining the positive effects of the program. It can be argued that NFYE students have also shown increases to their GPA scores, but this can be attributed once again to other unknown factors that were not measured or simply the effect of diminishing returns as explained before.

Looking for covariates using methods such as hierarchical clustering and tabulation was also a great way to uncover uncertainties or further prove the observations that were present before. This will hopefully benefit the direction of the FYE Program and what it needs to target for the program to show even more benefits for a larger portion of the participants with their respective needs and tendencies and being able to meet these standards. Hierarchical clustering has for example shown that even with the limited data that we have, it was noticeable that students in the College of Engineering tended to have higher GPA scores across the board while other colleges may have a harder time in keeping up.

4.2 Suggestions for Future Research

A suggestion that I would like to suggest for future research on the program and its effects is to explore more variables and find methods of obtaining such data that we did not have for either year of the program. While variables such as demographics were readily available, data

such as how much hours a student worked or surveys that could detect stress levels occurring specifically from certain factors such as classwork or food insecurity rather than a general stress detector may be interesting to look up further on.

Variables that were suggested for analysis that needed to be continuous and numeric included standardized test scores to determine whether one's pre-university academic performance could be a factor in academic success in college or how much time a student studies on a given week. I believe finding a method to retrieve these data and then including them for regression analysis or a separate analysis could uncover even more information. Additional information that I would recommend to try to obtain are any data that would allow to look deeper into preliminary academic background or level of students in the FYE program to find whether these students were more likely to have had higher academic success before joining the program. Some of these factors may be SAT scores or high school GPA scores. This will hopefully answer any disparities if any before the program takes into effect.

Some suggestions I had resulting from not having enough time or data that would be great candidates for research next time would be first having some categories in the demographics dataset for the current year's students to be complete and be used for regression such as the college as this year's dataset did not have it available. While I do not expect too much to change regarding the regression and significant factors, having more factors that may affect GPA scores are always great to look at and include. Also, it would be nice to perform a logistic regression with retention as the response variable to predict retention similar to how GPA scores were predicted.

Finally, proposing more suggestions for the program itself by looking at the history of the program and conducting surveys to the participants themselves would be another great way to improve the program and have even further impacts on its participants.

4.3 Parting Words

The FYE Program and its leaders and creators since the inception have been constantly finding ways to improve and have shown passion in bettering the lives of many students in their first year of university life. Nearly up to half of all university students have experienced food insecurity at one point, and the program has done so much not only to help improve student grades and numbers but by also providing a humanitarian solution for students who may not know where their next meal may come from and thus letting them know that there are always people who are willing to help them through difficult situations. I hope that the program will continue and keep improving to affect as many lives as possible as the program is such an invaluable resource for students, especially those transitioning into a new phase of their life away from their security at home and supporting themselves for the first time.

REFERENCES

Camelo, K., & Elliott, M. (2019). Food insecurity and academic achievement among college students at a public university in the United States, Journal of College Student Development, 60(3), 307-318, DOI: 10.1353/csd.2019.0028

Cohen, Sheldon, et al. "A Global Measure of Perceived Stress." Journal of Health and Social Behavior, vol. 24, no. 4, Dec. 1983, pp. 385–396. EBSCOhost, doi:10.2307/2136404.

Davis, G. M., Hanzsek-Brill, M. B., Petzold, M. C., Robinson, D. H. (2019). Students' Sense of Belonging: The development of a predictive retention model, Journal of the Scholarship of Teaching and Learning, 19(1), 117-127, DOI: 10.14434/josotl.v19i1.26787

Davis, H., Sisson, S. B., & Clifton, S. (2020). A call for evidence to support food security interventions on college campuses, Journal of American College Health, DOI: 10.1080/07448481.2019.170589

Nikolaus, C. J., Ellison, B., Nickols-Richardson, S. M. (2019). Are estimates of food insecurity among college students accurate? Comparison of assessment protocols, PLoS ONE, 14(4), e0215161, DOI: 10.1371/journal.pone.0215161

Pratt, I. S., Harwood, H. B., Cavazos, J. T., Ditzfield, C. P. (2017). Should I stay or should I go? Retention in first-generation college students, Journal of College Student Retention: Research, Theory & Practice, 0(0), 1-14, DOI: 10.1177/1521025117690868

Saklofske, D. H., Austin, E. J., Mastoras, S. M., Beaton, L., Osborne, S. E. (2012). Relationships of personality, affect, emotional intelligence and coping with student stress and academic stress: Different patterns of association for stress and success, Learning and Individual Differences, 22(2), 251-257, DOI: 10.1016/j.lindif.2011.02.010

Slaten, C. D., Elison, Z. M., Deemer, E. D., Hughes. H. A., Shemwell, D. A. (2018). The Development and Validation of the University Belonging Questionnaire, The Journal of Experimental Education, 86:4, 633-651, DOI: 10.1080/00220973.2017.1339009

Van Woerden, I, Hruschka, D., Bruening, M. (2019). Food insecurity negatively impacts academic performance, J Public Affairs, 19, e1864, DOI: 10.1002/pa.1864