Abstract

Achieving academic excellence is a key goal for higher education students and student's success is important to educational institutions because it serves as an indictor of their performance, yet the elements that lead to exceptional performance have not yet been thoroughly examined. This study looks at the variables that affect a student's chances of receiving an AA grade. in their annual academic performance with a focus on class attendance as a key factor. Using the multiple logistic regression method, the research shows which variables within various aspects affect the students' likelihood to get an AA grade. The goal looks into and assesses how much of an impact the student's weekly study hours and attendance in class will have on their achievement of an AA grade. For this research we are given a dataset to analyze and look for different suitable variables that support the claim to our study .A combination of descriptive analysis , correlation analysis , simple ordinal logistic regression models , full model ordinal logistic regression, reduced models ordinal logistic regression model and r-squared and coefficient of determination will be employed to evaluate the relationship between different variables and to classify which variables contribute to students achieving the AA grade in their annual performance. The study's main take home message reveals that consistent class attendance and other variables are strongly linked to better academic performance and achievement of AA grade, understanding and commitment.

Research Questions

Main research question:

What factors predict whether a higher education student will achieve an 'AA' grade?

Proposed / Created research question:

1. How does class attendance increase the annual academic performance of university students to achieve AA grade?

Literature Review

Introduction

The main aim of our research is to analyze different factors that increase the student's ability or chances to obtain AA grade in their annual grades. Different factors contribute to the student's achievement. The relationship between class attendance and academic performance is an important variable in this research, with most of the research demonstrating a favorable correlation between regular attendance and higher academic achievement. The research also helps us to determine whether always attending classes is directly proportional to academic performance or if other factors also contribute. According to some academics, students can still get high marks even with irregular attendance, and attendance alone is not a reliable factor of success. This research will look at parental influence after evaluating the relationship between academic performance and attendance to address these themes. It will next look at the roles of classroom engagement and project-based learning before closing with an assessment of the consequences for education practice. Our main research question is: What factors predict whether a higher education student will achieve an 'AA' grade? With our proposed research question being: How does class attendance increase the annual academic performance of university students to achieve 'AA' grade? The goal of this research is to provide a deeper knowledge of the nature of academic performance and to identify ways of improving student achievement in educational institutions by applying this perspective to the review of existing literature.

The Effect of Attendance at Class on Academic Performance

Because it allows for active participation with peers and instructors and gives access to course materials, regular attendance is essential for promoting academic success. Numerous academic investigations have emphasized the significance of attendance in class as a determinant of academic achievement. Ancheta et al. 2021, claim that students who regularly attend class typically achieve better academically. Attending classes on a regular basis gives students the chance to interact with teachers face-to-face, get quick feedback, and ask questions that are answered, which helps them comprehend the course material better. Furthermore, Yılmaz and Sekeroglu (2020) stress the importance of in-class conversations, pointing out that these exchanges improve concept comprehension and information storage. In conclusion, there is a clear correlation between academic achievement and class attendance, with frequent attendance improving knowledge and overall academic results. Although attendance and peer interaction are important, other factors such as parental influence also play a role in students' academic outcomes

The Impact of Parental Education and Occupation on Academic Performance

Students' academic performance is greatly influenced by their parents' occupation and their level of education, with better academic results frequently correlated with greater parental education levels. Research indicates that the educational attainment of students is significantly impacted by parental education, especially that of mothers (Judith et al., 2015). Higher educated parents are more likely to recognize the value of attendance and to push their children to focus on their education (Lareau, 2011). Also, if parents respect education, their jobs may motivate their children to achieve academic success (Cheung & Chan, 2010). Parental education has an especially big impact in situations where pupils do not have access to alternative mentorship or support systems (Gonzalez, 2019). The level of support and encouragement students receive at home is often determined by the education and occupation of their parents, which plays a significant role in determining their academic achievement. The relationship between attendance and academic achievement is complex, as other factors including the number of siblings also play a part, even though parental background plays a significant influence.

The impact of the number of siblings on academic performance

A student's academic performance can be impacted by the number of siblings they have. According to research, having more than one sibling can influence academic performance by encouraging competitive and supportive discussions. This paragraph's argument is that having siblings can positively or negatively impact academic performance. According to research, siblings in larger homes frequently contribute resources and information, which supports a supportive environment for learning (Duncan & Dowsett, 2020). The academic success of every student may suffer, on the other hand, if there is more competition for parental attention and resources, which could result in less academic assistance (Baker & Gruber, 2019). Children from larger families, for example, typically perform worse academically, possibly because of less personalized parental care (Hsin & Farkas, 2009). This relationship emphasizes the importance of considering familial dynamics when examining elements that contribute to academic performance. We will examine the impact of reading specific books next since it is crucial to comprehend these dynamics to address the various aspects that contribute to academic performance.

The Influence of Reading Specific Books and Journals (Non-Scientific)

A student's academic performance might be greatly impacted by their use of non-scientific reading materials. Reading non-scientific books on a regular basis fosters critical thinking and creativity, both abilities that are essential for academic success, according to this concept. Students who engage in a variety of non-scientific books improve their vocabulary and understanding, which can lead to stronger writing and analytical skills in their academic work (Smith and Luthar, 2021). Furthermore, students who read for enjoyment develop a passion for learning and are inspired to consider many viewpoints (Nell, 2019). This enjoyment of reading not only enhances academic performance but also enriches students' overall educational experiences (Hirsch, 2016). Additionally, research by McKool and McLaughlin (2008) indicates that leisure reading positively correlates with students' motivation and engagement in school. Therefore, encouraging students to read non-scientific books can be viewed as a key strategy for enhancing their academic success, which leads us to consider the impact of additional work responsibilities on academic performance.

The Impact of Additional Work on Academic Achievement

Part-time employment and educational requirements can have a significant impact on a student's success. This thesis states that although part-time work offers financial assistance and useful experience, it may also interrupt study time and academic focus. According to research by Wang et al. (2021), students who work less hours can improve their academic performance by learning critical skills like accountability and time management. But busy job schedules can cause stress and insufficient study time, which may affect grades (Cohen & Kessel, 2018). Students who work more than 20 hours a week, for instance, frequently report feeling overloaded, which may impact their academic interest (Hirschi & Jaensch, 2015). Finding ways to assist student success requires an understanding of how to manage employment and academic obligations. Finally, our investigation of work-life balance among students improves a greater understanding of the factors affecting academic achievement.

Conclusion

This research study looked at several factors that influence students' academic achievement, namely their capacity to get an 'AA' grade. Class attendance, parental education and occupation, the number of siblings, parental status, and the effect of part-time job are some of the important topics covered. Although most studies consistently emphasise the value of consistent attendance and parental education in promoting academic performance, there are considerable contrasts in how sibling relationships and the consequences of part-time work are portrayed. The relationship between family structure and academic achievement is complicated by the fact that, whereas larger families might offer a supportive environment, they may also result in competition for parental resources. These results highlight the necessity of more investigation into the complex connections between these variables. Future research could examine the effects of family dynamics on academic performance in various educational environments, such as the number of siblings in relation to parental support. Furthermore, analysing the point at which part-time work starts to negatively impact academic achievement may offer insightful information to members of Congress and educational institutions looking to promote student achievement. Given the circumstances, this review supports the idea that academic success is complex and influenced by a wide range of related variables. Future research into the factors that predict excellent academic success will need to keep focus on the small details of these relationships. In addition to advancing our knowledge of student achievement, filling in these gaps in the literature will help develop better teaching methods and support networks

Methodology

INTRODUCTION

This study explores the different factors that predict students' annual academic performance, specifically focusing on the likelihood of achieving an AA grade. Understanding the influence of class attendance on academic success is crucial for universities aiming to improve student outcomes, as it addresses a key factor often linked to performance. In our research we have used various methods to study the relationship between student's academic performance with the main focus being on class attendance's impact. Other methodologies include statistical models like multiple logistic regression. While some research argues for the use of simpler correlation analyses, debates exist on the effectiveness of such models in capturing the complexity of student performance data. This study uses multivariate logistic regression to examine the relationship between class attendance and academic performances. The approach that was chosen allows for a more proficient study of how attendance affects the likelihood of getting an AA grade, making it suited for answering the research question. This study employes descriptive statistics to characterize data distribution, correlation analysis to analyze correlations between variables, logistic regression to predict the effect of attendance on grade outcomes, and other procedures that will be explored in the analysis. These tools will give an in-depth analysis of how class attendance affects academic performance. To guarantee diversity, the sample consists of students from various academic years and fields of study. The university's electronic attendance system provided the attendance records. Final grades and other information about academic performance were gathered from the academic database of the university. Attendance data was categorized into bins to facilitate analysis. This section will begin with a detailed explanation of the descriptive and correlation analyses, followed by both full version and reduced model of the multiple logistic regression model. The analysis will provide a clear understanding of the methods used.

DESCRIPTIVE ANALYSIS

The initial step in data analysis is descriptive analysis. To guarantee consistency and data validation, 131 undergraduate students from diverse university departments were chosen using a convenience sample technique. The dataset is ready for analysis after it has been cleaned and formatted correctly. Descriptive analysis is useful in this study because it allows us to compare the means of the variables and graphs and determine whether the data we have is normal. We may compare the distribution of class attendance to the annual grade attained with the guidance of a histogram.

Simple Ordinal logistic regression

Simple logistic regression is employed to evaluate the direct relationship between a single predictor variable and the outcome of interest—specifically, the likelihood of achieving an AA grade based on class attendance. This model will help in understanding the initial influence of attendance on academic performance without the complications introduced by additional predictors. The significance of the coefficient will indicate whether attendance has a statistically significant impact on the likelihood of achieving an AA grade. This initial analysis provides a baseline understanding of the relationship, which can then be expanded upon in the multiple logistic regression analyses

FULL ORDINAL LOGISTIC REGRESSION

To investigate the connection between class attendance and the probability of receiving an AA grade, a multivariate logistic regression model was created. We compared the effect of students' attendance in class to their academic performance based on this technique. The ordinal logistic technique was employed to fit the logistic regression model, and SAS procedures were used for data analysis. From the full simple ordinal logistic regression model, we will be able to create predictive equations. This model can then be used to find different predictor variables odds ratios and their significance in the model and given that they are not significant we can drop them. This will be our initial step to the reduced ordinal logistic regression model

REDUCED ORDINAL LOGISTIC REGRESSION MODEL

We can determine when the most effective logistic regression model was chosen based on particular criteria by applying selection techniques to the reduced logistic regression. This will also enable us to assess the correctness of our model when new and removed variables are incorporated into it. The selection techniques will include three options namely forward, backward and stepwise. The forward selection starts with no predictors and its algorithm will evaluate every potential predictor variable based on their p-value. The backward selection model starts with full predictors by identifying which predictors are not significant and removes them from the model. The stepwise selection technique is a combination of both forward and backward selection techniques. The above options will play a critical role in identifying the predictor variables that are highly significant for our model.

Coefficient of determination

The coefficient of determination will clarify how the student's attendance in class accounts for the entirety of the variability in the grade they received. Class attendance is significant in our model, as demonstrated by a higher R-squared or coefficient of determination, which indicates that more of the variation in the student achievement grade can be explained by it. On the other hand, a lower R-squared or coefficient of determination indicates that less of the variation in the student grade achievement of AA can be explained or was caused by class.

The explanatory power of the model is indicated by the R-squared value. A higher R-squared value suggests that the attendance rate of the students may account for a greater share of the variability in attaining an AA grade.

CONCLUSION

We will be able to determine which variables have an increased impact on the student's annual grade after running and completing our procedures. We will also be able to determine whether or not class attendance enhances the likelihood that the student will achieve an AA yearly grade. These techniques or procedures will also allow us do variance tests or ANOVA, which will help us understand more about the variables that actually effect student's annual grade accomplishment. We will also be able to see the strength of the associations between the variables. We will be able to determine which factors have a bigger influence on the students' yearly grade as well as how attending class improves the student's chances.

Initial Analysis

Introduction

This study's primary goal is to draw conclusions about the factors that influence a student's yearly achievement of an AA grade. The purpose of this study is to ascertain how a student's attendance in class affects their annual accomplishment of an AA grade. From the start of the first semester until the conclusion of the second semester in 2019, data was gathered for the full academic year. There are 33 variables in the dataset, broken down into categories such as academic, personal, and family-related. In order to draw conclusions about these variables and data, we used a variety of analytical tests and analyses, including ANOVA, but our primary method was multiple logistic regression, which only included significant variables—variables that did not significantly impact the system or the model would be eliminated. To draw more insight regarding the effects of different variables on the students annual grade, different statistical methods were used such as descriptive analysis to provide an initial summary and understanding of the dataset, simple ordinal logistic regression to assess the impact of class attendance on student’s annual grade, full model ordinal logistic regression is to analyze the combined effect of all the independent variables on the ordinal dependent variable (Grade), reduced model ordinal logistic regression to streamline the model by focusing only on the most significant variables that affect a student's likelihood of achieving an AA grade and selection techniques. Once the different analytical/ statistical techniques are employed we will be able to find the main factors that increases the chances of a student to achieve AA grade to be added in our model and which ones are unnecessary and we will exclude them from our model or as key indicators. Having outlined the overall objective of the study and the various statistical methods employed, the next step is to delve into the application of each analytical technique.

DISSCUSION AND RESULTS

Discriptive Analysis

The purpose of this descriptive analysis is to explore and summarize the key characteristics of student academic performance on achieving an AA grade. By focusing on the distribution and patterns of these variables, such as attendance, parent's education and occupation, number of siblings etc. and grades, this analysis aims to provide a clear understanding of how these factors interrelate. Through descriptive statistics, we examine central tendencies, dispersions, and the overall distribution of attendance and grade outcomes. This foundational overview not only highlights the prevalence of different attendance behaviors among students but also reveals correlations between attendance patterns and academic achievement. By visualizing these relationships, we can identify trends that may inform strategies to enhance student engagement and performance.

DISCUSSION

NORMALITY OF ATTENDANCE

In order to generate the normality and distribution graphs for the variables attendance and outcome grade that we are dealing with, descriptive analysis was used. According to attendance normalcy, there is a low variance of 0.1936 and a mean attendance of 1.259. The distribution of attendance is platykurtic due to its negative kurtosis and is right-skewed, with a skewness of 1.109 and a kurtosis of –0.78.

Looking at the frequency table results for varied attendance and grade, it can be seen that most students have been attending consistently. When attendance = 1 (always), the frequency is 97, or 74.05% of the total attendance record.

Category 2, or when Attendance = 2 (sometimes), has a frequency of 34, which represents 25.95% of the entire class—a small percentage of the class proportion—when it occurs.

There are seven students in grade 0 (fail), which is represented by frequency and accounts for only 5.34% of the entire data; in grade 7, or AA, there are seventeen students, or 12.98% of the whole student body.

Cross-tabulation frequency between attendance and grade

Attendance 1 (Attending Always)

Our cross-tabulation reveals that 5 out of the 71.43% of students with attendance of 1 are in grade 0, indicating a very weak correlation between students who consistently attend and attendance of 1. This indicates that there is little possibility of failure for students who consistently attend. According to our cross-tabulation, there are 15 students, and 88% of those with attendance 1 are in grade 7, or AA. This indicates and demonstrates the strong correlation between students who consistently attend and those who receive an AA grade. This implies and informs us that students who consistently attend class run the risk of failing.

Attendance 2 (Attending Sometimes)

Our cross-tabulation reveals that 2 and a total of 28.57% of students with attendance 2 are in grade 0, indicating a very weak correlation between students who consistently attend and attendance 1. This implies and indicates that students who attend occasionally have low failure rates. According to our cross-tabulation, there are 2 students and 11.76% of those with attendance 2 are in grade 7, or AA. This indicates and demonstrates that there is a very weak correlation between students with attendance 2 or those who attend occasionally and kids receiving an AA grade. This indicates that there are instances when attending pupils have extremely little probability of earning an AA grade.

Based on the aforementioned data, it is evident that students who consistently attend classes are more likely to pass and receive high grades than those who attend sometimes.

Grade Normality

The grade's mean is 3.2671, its skewness is 0.372, and its kurtosis is –1.204, indicating that it has a platykurtic distribution and is skewed to the right, according to grade normalcy. Five is on the third or 75% quantile, indicating that most students are receiving BB grades, while the seven is on the 100 quantiles, indicating that more students are receiving AA grades. Pupils that receive an AB grade fall into the middle, while failing students are at the bottom of the quartile. It is also evident from the normalcy histogram that a greater number of students are receiving the grades DD, DC, and AA, supporting the distribution's proper skew. The majority of students' consistent attendance indicates a connection between class attendance.

SG plot of grade by attendance

Our graph demonstrates that students who consistently attend outperform those who don't very often. The course material and interactions with the instructor may be to blame for this. Our graph indicates that students who consistently miss class have a lower average than those who attend every time.

Other circumstances and academic performance

Our grade and parental status data do indicate that married kids fare better academically than divorced students, but not better than students whose parents or one of them has passed away. According to our data, adolescents who have lost both or one of their parents typically outperform those whose parents are married or divorced. This concludes and demonstrates that having one deceased parent is a contributing factor to a student's accomplishment of AA, rather with attendance being the most important factor.

Analysis for taking notes in class and grade

According to our analysis, students who don't take any notes in class are more likely to perform worse than other students, indicating that participation and interaction with other students in the classroom have an impact on a student's potential to pass and receive an AA grade. Additionally, it demonstrates that pupils who take notes in class typically outperform those who don't. This means that while attendance is important, it is insufficient on its own. In class, students ought to engage as well as take notes. Pupils who take notes consistently do better on average and receive higher grades than those who just occasionally or never take notes.

Conclusion(Descriptive analysis)

The descriptive analysis of the data revealed important insights into the relationship between attendance and academic performance. Students who consistently attended classes were more likely to achieve AA grade compared to students who did not attend classes always or who attended sometimes. However, class attendance alone was not sufficient; additional factors such as parental status and note-taking also influenced academic success. These findings provide a strong foundation for further analysis, which can explore deeper insights into the dynamics of student achievement of AA grade in their end of year.

Coefficient of determination

The aim of this correlation analysis is to determine the strength and direction of the relationship between student attendance and academic performance, as measured by grades. By examining the correlation between these two variables, we seek to uncover whether changes in attendance are associated with corresponding changes in academic outcomes. Understanding this relationship is crucial for gaining deeper insights into the dataset, as it helps to identify whether increased attendance plays a significant role in academic success. In this study, we use correlation analysis to assess the potential linear association between attendance and grades, offering a foundation for further statistical exploration and decision-making in educational strategies.

Hypothesis Testing

H0: $p = 0$: There is no linear association between the two variables

Ha: $p \neq 0$: There is a linear association between the two variables

conclusion

The correlation between **Attendance and GRADE is p=** 0.17242

which is a weak positive correlation.

From our study and evaluation of the correlation it does show and outline that attendance has a positive relation showing that it does increase the student's ability to obtain AA grade, but the relationship is a weak one.

Simple Ordinal Logistic Regression models

Simple Ordinal Logistic Regression models are used to analyze relationships between a set of predictors and an ordinal dependent variable, where the outcomes have a natural order. The goal of this analysis is to examine how individual predictors, such as attendance or study habits, influence the probability of students achieving higher or lower academic grades. By focusing on a single predictor at a time, these models help us understand the direct impact of each factor on the ordered outcome. This approach allows for a clear interpretation of the strength and direction of these relationships, providing valuable insights into the dynamics that contribute to academic performance.

Sas code was ran and the following equations were obtained

GRADE 7

$y=11.929-1.1676x1$

GRADE 6

$y=11.1188-0.7621x1$

GRADE 5

$y=10.4257-0.0690x1$

GRADE 3

$y=12.8731-1.9858x1$

GRADE 2

$y=10.0972+0.4418x1$

GRADE 1

$y=10.8109+0.2877 \cdot x1$

GRADE 0

$y=-0.2675-0.0690x1$

The multiple ordinal Logistic Model (Full Model)

The multiple logistic analysis creates a model that helps predict the AA grade achievement of a student. This model will help obtain an equation for the probability of getting AA grade at the end of the year.

This model will include all the variables in the study.

Hypothesis testing

Null hypothesis:

The personal, family and educational habits are independent =0 (there is no relationship between them)

Alternative hypothesis:

The personal, family and educational habits are dependent
≠≠
0 (there is a relationship between them)

Decision rule

We will reject the null hypothesis if there is a |p| < 0.05 and accept the null hypothesis when the |p| > 0.05.

Conclusion (Hypothesis)

Student_Age: p = 0.0042, Sex: p = 0.0132, Grad_High_School_Typ: p = 0.0216,
Scholarship_Type: p = 0.0009, Additional_Mark: p = 0.0054
Salary: p = 0.0280, Transportation: p < 0.0001, Accommodation: p = 0.0002
Mother_Edu: p = 0.0358, Mothers_Occu: p = 0.0030, Fathers_Occu: p = 0.0134,
Read_Freq_Non_Scien: p = 0.0012, Seminar_Attend: p = 0.0001
Impact_Proj_Stud_Suc: p = 0.0040, Prep_Mid_Term_Exam_1: p = 0.0269
Prep_Mid_Term_Exam_2: p = 0.0205, Listening_Class: p = 0.0310
Cumal_GPA_Last_Sem: p = 0.0011, Flip_Class: p = 0.0029
The above predictor variables have p-values less than 0.05 which shows that we reject the null hypothesis therefore showing that the variables are statistically significant.

Conclusion

The logistic model reveals several factors that significantly influence students grades such as sex, transportation, scholarship type and attendance. These variables provide insights into potential areas to focus on when considering intervention for improving academic performance.

Reduced models Ordinal Logistic Regression Model

The rationale behind creating a reduced Ordinal Logistic Regression model is to simplify the analysis by focusing on the most significant predictors that influence the outcome variable. By eliminating non-essential variables, we can enhance the model's interpretability and efficiency without compromising accuracy. Some variables from the full model were rejected when examining the null hypothesis. The variables that were rejected will now be removed from the reduced model and we will apply different selection methods to be left with the most significant variables.

SELECTION TECHNIQUES

Backward selection technique

The stepwise selection option was ran to find different predictor variables that should be used for our final multiple logistic model. From the results the following predictor variables were found to be significant, Cumulative_Gpa last_sem, sex, Read_Freq_Non_Scien, Student_Age and Listening_Class.

Equations for each grade level

X1 = Cumal_GPA_Last_Sem, X2 = Sex, X3 = Read_Freq_Non_Scien

X4 = Student_Age, X5 = Listening_Class

GRADE 0

$Y = 1.8594 - 0.5473X_1 - 1.3218\,X_2 - 0.9180X_3 + 1.0423X_4 - 0.6631X_5$

GRADE 1

$Y = 4.2122 - 0.5473X_1 - 1.3218\,X_2 - 0.9180X_3 + 1.0423X_4 - 0.6631X_5$

GRADE 2

$Y = 5.0668 - 0.5473X_1 - 1.3218\,X_2 - 0.9180X_3 + 1.0423X_4 - 0.6631X_5$

GRADE 3

$Y = 5.8034 - 0.5473X_1 - 1.3218\,X_2 - 0.9180X_3 + 1.0423X_4 - 0.6631X_5$

GRADE 4

$Y = 6.2203 - 0.5473X_1 - 1.3218\,X_2 - 0.9180X_3 + 1.0423X_4 - 0.6631X_5$

GRADE 5

$Y = 6.9025 - 0.5473X_1 - 1.3218\,X_2 - 0.9180X_3 + 1.0423X_4 - 0.6631X_5$

GRADE 6

$Y = 7.6808 - 0.5473X_1 - 1.3218\,X_2 - 0.9180X_3 + 1.0423X_4 - 0.6631X_5$

Forward selection technique

The forward selection option was ran to find different predictor variables that should be used for our final multiple logistic model. From the results the following predictor variables were found to be significant, Cumulative_Gpa last_sem, sex, Read_Freq_Non_Scien, Student_Age and Listening_Class.

Equations for each grade level

X1 = Cumal_GPA_Last_Sem, X2 = Sex, X3 = Read_Freq_Non_Scien

X4 = Student_Age, X5 = Listening_Class

Equations for Each Grade Level

GRADE 0

$Y = 1.1390 + 0.1156X_1 - 0.0570X_2 - 0.00032X_3 + 0.0890X_4 + 0.4736X_5$

GRADE 1

$Y = 1.1390 + 0.1156X_1 - 0.0570X_2 - 0.00032X_3 + 0.0890X_4 + 0.8064X_5$

GRADE 2

$Y = 1.1390 + 0.1156X_1 - 0.0570X_2 - 0.00032X_3 + 0.0890X_4 + 1.1487X_5$

GRADE 3

$Y = 1.1390 + 0.1156X_1 - 0.0570X_2 - 0.00032X_3 + 0.0890X_4 + 1.4911X_5$

GRADE 4

$Y = 1.1390 + 0.1156X_1 - 0.0570X_2 - 0.00032X_3 + 0.0890X_4 + 1.8330X_5$

GRADE 5

$Y = 1.1390 + 0.1156X_1 - 0.0570X_2 - 0.00032X_3 + 0.0890X_4 + 2.1749X_5$

GRADE 6

$Y = 1.1390 + 0.1156X_1 - 0.0570X_2 - 0.00032X_3 + 0.0890X_4 + 2.5168X_5$

GRADE 7

$Y = 1.1390 + 0.1156X_1 - 0.0570X_2 - 0.00032X_3 + 0.0890X_4 + 2.8587X_5$

Stepwise selection technique

The stepwise selection option was ran to find different predictor variables that should be used for our final multiple logistic model. From the results the following predictor variables were found to be significant, Cumulative_Gpa last_sem, sex, Read_Freq_Non_Scien, Student_Age and Listening_Class.

Equations for Each Grade Level

X1 = Cumal_GPA_Last_Sem, X2 = Sex, X3 = Read_Freq_Non_Scien

X4 = Student_Age, X5 = Listening_Class

GRADE 0

$Y = 1.8594 - 0.5473X_1 - 1.3218\, X_2 - 0.9180X_3 + 1.0423X_4 - 0.6631X_5$

GRADE 1

$Y = 4.2122 - 0.5473X_1 - 1.3218\, X_2 - 0.9180X_3 + 1.0423X_4 - 0.6631X_5$

GRADE 2

$Y = 5.0668 - 0.5473X_1 - 1.3218\, X_2 - 0.9180X_3 + 1.0423X_4 - 0.6631X_5$

GRADE 3

$Y = 5.8034 - 0.5473X_1 - 1.3218\, X_2 - 0.9180X_3 + 1.0423X_4 - 0.6631X_5$

GRADE 4

$Y = 6.2203 - 0.5473X_1 - 1.3218\, X_2 - 0.9180X_3 + 1.0423X_4 - 0.6631X_5$

GRADE 5

$Y = 6.9025 - 0.5473X_1 - 1.3218\, X_2 - 0.9180X_3 + 1.0423X_4 - 0.6631X_5$

GRADE 6

$Y = 7.6808 - 0.5473X_1 - 1.3218\, X_2 - 0.9180X_3 + 1.0423X_4 - 0.6631X_5$

From the above selection techniques applied we were able to deduce significant predictor variables being cumulative gpa from the previous semester, sex, how often they read non-scientific books or journals, student age and how they listen in class. This analysis and model answers both our main and proposed research question.

CONCLUSION

Summary and conclusion

the examination of the factors in our study that influence students' success at the conclusion of the year. assessing the precise relationship between attendance and the student's capacity to receive an AA grade at the conclusion of the year. We also demonstrated and assessed additional factors in our study that affect students' performance at the conclusion of the year and their likelihood of earning an AA grade. variables include the students' marital status and attendance in class. Every element matters, and several factors influence a student's likelihood of receiving an AA on their year-end report card. For a student to receive an AA grade, attendance is a crucial component of their final grade.

Appendix

Codes

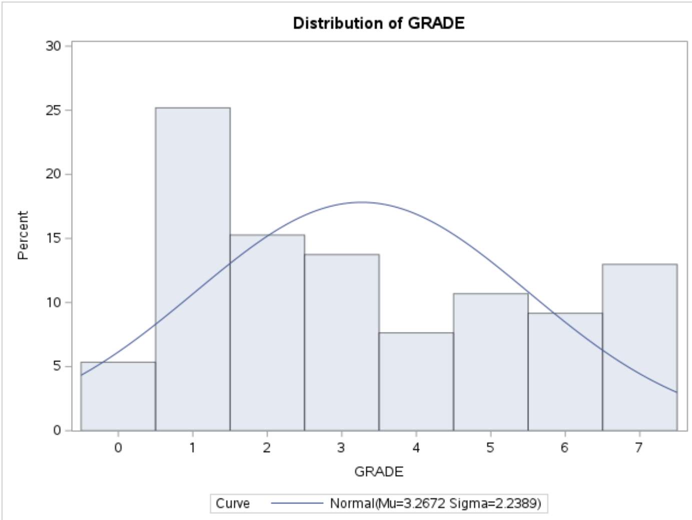
1.Descriptive statistics (Codes)

```
DATA grp6.group_6_train;

SET grp6.group_6_train

(RENAME=

  (_1=Student_Age

_2 = Sex

_3 = Grad_High_School_Type

_4 = Scholarship_Type

_5 = Additional_Mark

_6 = Artistic_Sports_Activity

_7 = Partner

_8 = Salary

_9 = Transportation

_10 = Accommodation

_11 = Mother_Edu

_12 = Father_Edu

_13 = No_Siblings

_14 =_15 = Mothers_Occu

_16 = Fathers_Occu
```

```
_17 = Weekly_Stu_Hour
_18 = Read_Freq_Non_Scien
_19 = Read_Freq_Scien
_20 = Seminar_Attend
_21 = Impact_Proj_Stud_Succ
_22 = Attendance
_23 = Prep_Mid_Term_Exam_1
_24 = Prep_Mid_Term_Exam_2
_25 = Taking_Notes_Class
_26 = Listening_Class
_27 = Discussions
_28 = Flip_Class
_29 = Cumal_GPA_Last_Sem
_30 = Expected_GPA_Grad));
DROP _dataobs_;
RUN;


ODS HTML;
PROC PRINT DATA grp6.group_6_train;
RUN;
ODS HTML CLOSE;
```

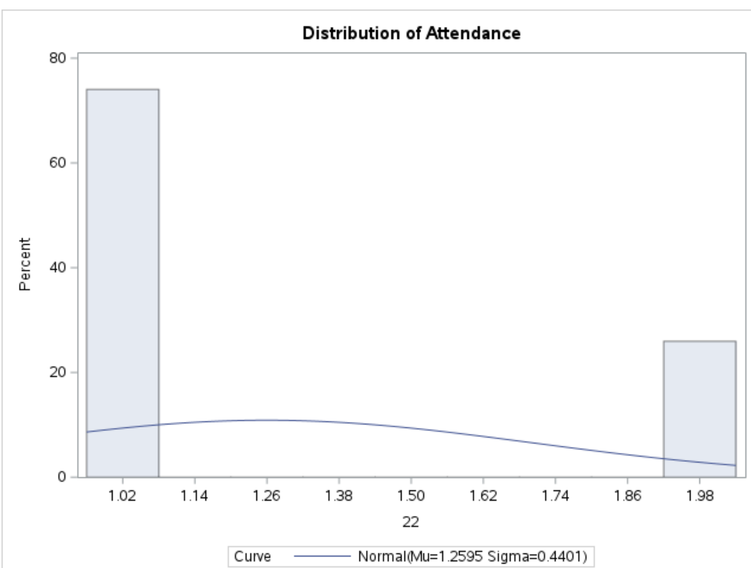
Distribution of grade

```
proc univariate data=grp6.group_6_train normal;
   var Grade;
   histogram Grade / normal;
   probplot Grade / normal(mu=est sigma=est);
   inset mean std skewness kurtosis / position=ne;
run;
```



Normality of Attendance

```
proc sgplot data=grp6.group_6_train;
  vbox grade / category=parental_Stat ;
  xaxis label="parental status";
  yaxis label="grade";
run;
```



Attendance

Cross tabulation

```
proc freq data=grp6.group_6_train;
        tables attendance*grade/ missing;
run;
```

The FREQ Procedure

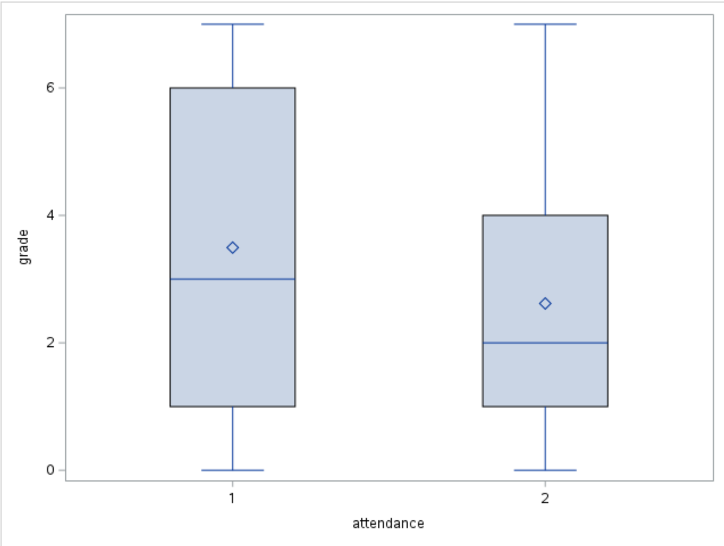
Frequency Percent Row Pct Col Pct

Table of Attendance by GRADE

	GRADE								
Attendance(22)	0	1	2	3	4	5	6	7	Total
1	5 3.82 5.15 71.43	21 16.03 21.65 63.64	12 9.16 12.37 60.00	17 12.98 17.53 94.44	7 5.34 7.22 70.00	10 7.63 10.31 71.43	10 7.63 10.31 83.33	15 11.45 15.46 88.24	97 74.05
2	2 1.53 5.88 28.57	12 9.16 35.29 36.36	8 6.11 23.53 40.00	1 0.76 2.94 5.56	3 2.29 8.82 30.00	4 3.05 11.76 28.57	2 1.53 5.88 16.67	2 1.53 5.88 11.76	34 25.95
Total	7 5.34	33 25.19	20 15.27	18 13.74	10 7.63	14 10.69	12 9.16	17 12.98	131 100.00

SG Plot for grade and attendance

```
ods graphics on;

proc sgplot data=grp6.group_6_train;

  scatter x=Attendance y=Grade / markerattrs=(symbol=circlefilled);

  xaxis label="Attendance";

  yaxis label="Eat Graduation";

run;
```



2.Correlation Analysis

ods html;

proc corr data=grp6.group_6_train PLOTS (ONLY) = SCATTER;

var Expected_GPA_Grad Attendance grade ;

run;

ods html close

Pearson Correlation Coefficients, N = 131 Prob > |r| under H0: Rho=0			
	Expected_GPA_Grad	Attendance	GRADE
Expected_GPA_Grad 30	1.00000	-0.20240 0.0204	0.27483 0.0015
Attendance 22	-0.20240 0.0204	1.00000	-0.17242 0.0489
GRADE	0.27483 0.0015	-0.17242 0.0489	1.00000

3. SIMPLE REGRESSION MODEL

ods html;

```
proc logistic data=grp6.group_6_train descending;
class grade (param=ref ref='4');
model grade = attendance / link=glogit expb;
run;
ods html close;
```

Analysis of Maximum Likelihood Estimates							
Parameter	GRADE	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp(Est)
Intercept	7	1	1.9297	1.2928	2.2280	0.1355	6.888
Intercept	6	1	1.1188	1.3434	0.6936	0.4049	3.061
Intercept	5	1	0.4257	1.2469	0.1165	0.7328	1.531
Intercept	3	1	2.8731	1.4629	3.8575	0.0495	17.692
Intercept	2	1	0.0972	1.1675	0.0069	0.9337	1.102
Intercept	1	1	0.8109	1.0856	0.5580	0.4551	2.250
Intercept	0	1	-0.2675	1.4848	0.0325	0.8570	0.765
Attendance	7	1	-1.1676	1.0212	1.3073	0.2529	0.311
Attendance	6	1	-0.7621	1.0374	0.5397	0.4625	0.467
Attendance	5	1	-0.0690	0.9090	0.0058	0.9395	0.933
Attendance	3	1	-1.9858	1.2389	2.5692	0.1090	0.137
Attendance	2	1	0.4418	0.8274	0.2852	0.5933	1.556
Attendance	1	1	0.2877	0.7792	0.1363	0.7120	1.333
Attendance	0	1	-0.0690	1.0845	0.0040	0.9493	0.933

Figure 3.1

4. Multiple Logistic Regression Model (Full Model)

```
PROC LOGISTIC DATA=grp6.group_6_train DESCENDING;
  CLASS Attendance (PARAM=REF)
        Sex (PARAM=REF)
        Grad_High_School_Type (PARAM=REF)
        Scholarship_Type (PARAM=REF)
        Artistic_Sports_Activity (PARAM=REF)
        Partner (PARAM=REF)
        Salary (PARAM=REF)
        Transportation (PARAM=REF)
        Accommodation (PARAM=REF)
```

```
      Mother_Edu (PARAM=REF)
      Father_Edu (PARAM=REF)
      No_Siblings (PARAM=REF)
      Parental_Stat (PARAM=REF)
      Mothers_Occu (PARAM=REF)
      Fathers_Occu (PARAM=REF)
      Read_Freq_Non_Scien (PARAM=REF)
      Read_Freq_Scien (PARAM=REF)
      Seminar_Attend (PARAM=REF)
      Impact_Proj_Stud_Succ (PARAM=REF)
      Prep_Mid_Term_Exam_1 (PARAM=REF)
      Prep_Mid_Term_Exam_2 (PARAM=REF)
      Listening_Class (PARAM=REF)
      Discussions (PARAM=REF)
      Flip_Class (PARAM=REF)
      attendance (PARAM=REF);

MODEL grade(event='7') =
      Student_Age
      Sex
      Grad_High_School_Type
      Scholarship_Type
      Additional_Mark
      Artistic_Sports_Activity
      Partner
      Salary
      Transportation
```

```
        Accommodation
        Mother_Edu
        Father_Edu
        No_Siblings
        Parental_Stat
        Mothers_Occu
        Fathers_Occu
        Weekly_Stu_Hour
        Read_Freq_Non_Scien
        Read_Freq_Scien
        Seminar_Attend
        Impact_Proj_Stud_Succ
        Prep_Mid_Term_Exam_1
        Prep_Mid_Term_Exam_2
        Taking_Notes_Class
        Listening_Class
        Discussions
        Flip_Class
        Cumal_GPA_Last_Sem
        attendance
        expected_gpa_grad;
RUN;
```

6.Multiple Logistic Regression Model (Reduced Model);

SELECTION TECHNIQUES BACKWARD

```
proc logistic data=grp6.group_6_train;
  model grade(event='7') =
     Student_Age
    Sex
    Grad_High_School_Type
    Scholarship_Type
    Additional_Mark

    Salary
    Transportation
    Accommodation
    Mother_Edu
    No_Siblings
    Fathers_Occu
    Weekly_Stu_Hour
    Read_Freq_Non_Scien
    Seminar_Attend
    Impact_Proj_Stud_Succ
    Prep_Mid_Term_Exam_1
    Prep_Mid_Term_Exam_2
    Listening_Class
      Flip_Class
    Cumal_GPA_Last_Sem

    / selection=backward slstay=0.05;
run;
```

Summary of Backward Elimination						
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq	Variable Label
1	No_Siblings	1	19	0.0427	0.8364	13
2	Scholarship_Type	1	18	0.1492	0.6993	4
3	Salary	1	17	0.1486	0.6999	8
4	Prep_Mid_Term_Exam_1	1	16	0.2315	0.6304	23
5	Mother_Edu	1	15	0.3242	0.5691	11
6	Weekly_Stu_Hour	1	14	0.5763	0.4478	17
7	Impact_Proj_Stud_Suc	1	13	1.1518	0.2832	21
8	Seminar_Attend	1	12	1.9710	0.1603	20
9	Prep_Mid_Term_Exam_2	1	11	2.1388	0.1436	24
10	Flip_Class	1	10	2.2455	0.1340	28
11	Grad_High_School_Typ	1	9	2.4427	0.1181	3
12	Accommodation	1	8	2.8927	0.0890	10
13	Fathers_Occu	1	7	2.4981	0.1140	16
14	Transportation	1	6	3.4461	0.0634	9
15	Additional_Mark	1	5	3.0087	0.0828	5

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	0	1	1.8594	1.0705	3.0172	0.0824
Intercept	1	1	4.2122	1.0694	15.5138	<.0001
Intercept	2	1	5.0668	1.0955	21.3905	<.0001
Intercept	3	1	5.8034	1.1222	26.7452	<.0001
Intercept	4	1	6.2203	1.1376	29.9002	<.0001
Intercept	5	1	6.9025	1.1630	35.2249	<.0001
Intercept	6	1	7.6808	1.1925	41.4879	<.0001
Student_Age		1	1.0423	0.2832	13.5503	0.0002
Sex		1	-1.3218	0.3519	14.1049	0.0002
Read_Freq_Non_Scien		1	-0.9180	0.2919	9.8879	0.0017
Listening_Class		1	-0.6631	0.2464	7.2403	0.0071
Cumal_GPA_Last_Sem		1	-0.5473	0.1353	16.3687	<.0001

SELECTION TECHNIQUES FORWARD;

```
proc logistic data=grp6.group_6_train;
   model grade(event='7') =
      Student_Age
      Sex
      Grad_High_School_Type
      Scholarship_Type
      Additional_Mark
      Salary
      Transportation
      Accommodation
      Mother_Edu
      No_Siblings
      Fathers_Occu
      Weekly_Stu_Hour
      Read_Freq_Non_Scien
      Seminar_Attend
      Impact_Proj_Stud_Succ
```

```
        Prep_Mid_Term_Exam_1
        Prep_Mid_Term_Exam_2
        Listening_Class
        Flip_Class
        Cumal_GPA_Last_Sem
        / selection=forward slentry=0.05;
run;
```

Summary of Forward Selection						
Step	Effect Entered	DF	Number In	Score Chi-Square	Pr > ChiSq	Variable Label
1	Cumal_GPA_Last_Sem	1	1	15.3623	<.0001	29
2	Sex	1	2	8.3848	0.0038	2
3	Read_Freq_Non_Scien	1	3	6.6723	0.0098	18
4	Student_Age	1	4	8.1949	0.0042	1
5	Listening_Class	1	5	7.3945	0.0065	26

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	0	1	1.8594	1.0705	3.0172	0.0824
Intercept	1	1	4.2122	1.0694	15.5138	<.0001
Intercept	2	1	5.0668	1.0955	21.3905	<.0001
Intercept	3	1	5.8034	1.1222	26.7452	<.0001
Intercept	4	1	6.2203	1.1376	29.9002	<.0001
Intercept	5	1	6.9025	1.1630	35.2249	<.0001
Intercept	6	1	7.6808	1.1925	41.4879	<.0001
Student_Age		1	1.0423	0.2832	13.5503	0.0002
Sex		1	-1.3218	0.3519	14.1049	0.0002
Read_Freq_Non_Scien		1	-0.9180	0.2919	9.8879	0.0017
Listening_Class		1	-0.6631	0.2464	7.2403	0.0071
Cumal_GPA_Last_Sem		1	-0.5473	0.1353	16.3687	<.0001

SELECTION TECHNIQUE STEPWISE;

```
proc logistic data=grp6.group_6_train;
   model grade(event='7') =
      Student_Age Sex Grad_High_School_Type  Scholarship_Type Additional_Mark
        Salary Transportation Accommodation  Mother_Edu No_Siblings Fathers_Occu
        Weekly_Stu_Hour Read_Freq_Non_Scien Seminar_Attend
Impact_Proj_Stud_Succ
     Prep_Mid_Term_Exam_1 Prep_Mid_Term_Exam_2 Listening_Class Flip_Class
```

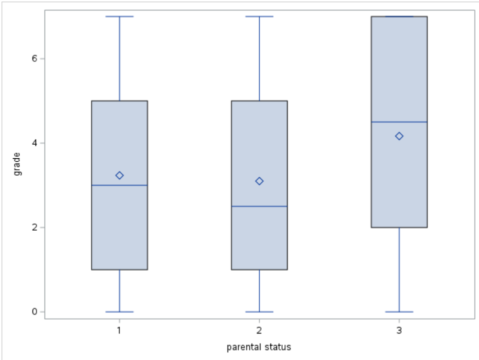
```
        Cumal_GPA_Last_Sem
        / selection=stepwise slstay=0.05;
run;
```

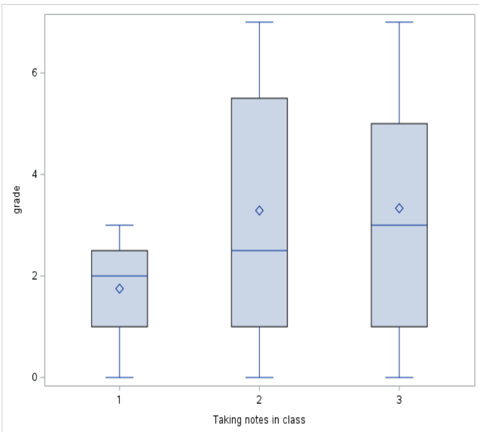
Summary of Stepwise Selection							
	Effect						
Step	Entered	Removed	DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
1	Cumal_GPA_Last_Sem		1	1	15.3623		<.0001
2	Sex		1	2	8.3848		0.0038
3	Read_Freq_Non_Scien		1	3	6.6723		0.0098
4	Student_Age		1	4	8.1949		0.0042
5	Listening_Class		1	5	7.3945		0.0065

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	0	1	1.8594	1.0705	3.0172	0.0824
Intercept	1	1	4.2122	1.0694	15.5138	<.0001
Intercept	2	1	5.0668	1.0955	21.3905	<.0001
Intercept	3	1	5.8034	1.1222	26.7452	<.0001
Intercept	4	1	6.2203	1.1376	29.9002	<.0001
Intercept	5	1	6.9025	1.1630	35.2249	<.0001
Intercept	6	1	7.6808	1.1925	41.4879	<.0001
Student_Age		1	1.0423	0.2832	13.5503	0.0002
Sex		1	-1.3218	0.3519	14.1049	0.0002
Read_Freq_Non_Scien		1	-0.9180	0.2919	9.8879	0.0017
Listening_Class		1	-0.6631	0.2464	7.2403	0.0071
Cumal_GPA_Last_Sem		1	-0.5473	0.1353	16.3687	<.0001

Graphs

Sg plot of parental status





2. Correlation

3. Simple logistic

ods html;

proc logistic data=grp6.group_6_train descending;

class grade (param=ref ref='4');

```
model grade = attendance / link=glogit expb;

run;

ods html close;
```

Analysis of Maximum Likelihood Estimates							
Parameter	GRADE	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp(Est)
Intercept	7	1	1.9297	1.2928	2.2280	0.1355	6.888
Intercept	6	1	1.1188	1.3434	0.6936	0.4049	3.061
Intercept	5	1	0.4257	1.2469	0.1165	0.7328	1.531
Intercept	3	1	2.8731	1.4629	3.8575	0.0495	17.692
Intercept	2	1	0.0972	1.1675	0.0069	0.9337	1.102
Intercept	1	1	0.8109	1.0856	0.5580	0.4551	2.250
Intercept	0	1	-0.2675	1.4848	0.0325	0.8570	0.765
Attendance	7	1	-1.1676	1.0212	1.3073	0.2529	0.311
Attendance	6	1	-0.7621	1.0374	0.5397	0.4625	0.467
Attendance	5	1	-0.0690	0.9090	0.0058	0.9395	0.933
Attendance	3	1	-1.9858	1.2389	2.5692	0.1090	0.137
Attendance	2	1	0.4418	0.8274	0.2852	0.5933	1.556
Attendance	1	1	0.2877	0.7792	0.1363	0.7120	1.333
Attendance	0	1	-0.0690	1.0845	0.0040	0.9493	0.933

4. Multiple logistic (Full Model)

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	7	1	-14.7552	6.5251	5.1135	0.0237
Intercept	6	1	-13.4670	6.5104	4.2789	0.0386
Intercept	5	1	-12.3289	6.4980	3.5999	0.0578
Intercept	4	1	-11.6775	6.4911	3.2364	0.0720
Intercept	3	1	-10.5660	6.4797	2.6590	0.1030
Intercept	2	1	-9.2841	6.4699	2.0592	0.1513
Intercept	1	1	-5.4776	6.4117	0.7299	0.3929
Student_Age		1	-1.5575	0.5442	8.1916	0.0042
Sex	1	1	-1.4464	0.5838	6.1381	0.0132
Grad_High_School_Typ	1	1	-2.2666	0.9805	5.3442	0.0208
Grad_High_School_Typ	2	1	-0.0138	0.6846	0.0004	0.9839
Scholarship_Type	1	1	0.5041	2.6164	0.0371	0.8472
Scholarship_Type	2	1	7.9474	2.1915	13.1518	0.0003
Scholarship_Type	3	1	1.9197	0.7750	6.1362	0.0132
Scholarship_Type	4	1	3.1693	0.8889	12.7110	0.0004
Additional_Mark		1	1.8260	0.6558	7.7517	0.0054
Artistic_Sports_Acti	1	1	-1.0380	0.5934	3.0596	0.0803
Partner	1	1	-0.8541	0.5349	2.5493	0.1103
Salary	1	1	2.9209	1.4950	3.8175	0.0507
Salary	2	1	1.3215	1.4989	0.7774	0.3780
Salary	3	1	0.8040	1.6400	0.2403	0.6240
Salary	4	1	1.4211	2.1955	0.4190	0.5174
Transportation	1	1	4.6518	0.9528	23.8381	<.0001
Transportation	2	1	5.0198	1.2973	14.9726	0.0001
Transportation	3	1	2.1482	2.7592	0.6062	0.4362

Accommodation	1	1	-2.5841	2.7597	0.8768	0.3491
Accommodation	2	1	0.0957	2.7336	0.0012	0.9721
Accommodation	3	1	-0.3150	2.8622	0.0121	0.9124
Mother_Edu	1	1	-3.9080	2.3882	2.6777	0.1018
Mother_Edu	2	1	-2.7962	2.5440	1.2081	0.2717
Mother_Edu	3	1	-2.8960	2.4218	1.4299	0.2318
Mother_Edu	4	1	-2.4364	2.4172	1.0159	0.3135
Mother_Edu	5	1	-10.0639	3.5418	8.0738	0.0045
Father_Edu	1	1	-1.3317	2.5536	0.2720	0.6020
Father_Edu	2	1	-1.8541	2.5823	0.5155	0.4728
Father_Edu	3	1	-3.2671	2.5861	1.5960	0.2065
Father_Edu	4	1	-2.2828	2.5781	0.7840	0.3759
Father_Edu	5	1	-5.1724	2.9591	3.0554	0.0805
No_Siblings	1	1	0.1270	0.9424	0.0182	0.8928
No_Siblings	2	1	-1.6917	0.8791	3.7026	0.0543
No_Siblings	3	1	-0.7949	0.8012	0.9844	0.3211
No_Siblings	4	1	-2.0599	0.8940	5.3088	0.0212
Parental_Stat	1	1	1.5284	1.6538	0.8541	0.3554
Parental_Stat	2	1	1.9329	1.8442	1.0985	0.2946
Mothers_Occu	1	1	8.9904	2.4639	13.3140	0.0003
Mothers_Occu	2	1	5.6954	2.0815	7.4869	0.0062
Mothers_Occu	3	1	4.3171	2.0574	4.4028	0.0359
Mothers_Occu	4	1	5.7254	2.1522	7.0771	0.0078
Fathers_Occu	1	1	1.7148	0.8681	3.9019	0.0482
Fathers_Occu	2	1	0.8261	1.1251	0.5390	0.4628
Fathers_Occu	3	1	-0.3173	0.9237	0.1180	0.7312
Fathers_Occu	4	1	-0.6986	0.8921	0.6133	0.4336
Weekly_Stu_Hour		1	-0.1651	0.3144	0.2759	0.5994

Read_Freq_Non_Scien	1	1	-2.8636	1.0631	7.2553	0.0071
Read_Freq_Non_Scien	2	1	-2.8151	0.7711	13.3290	0.0003
Read_Freq_Scien	1	1	0.9769	0.9661	1.0223	0.3120
Read_Freq_Scien	2	1	-0.2408	0.6575	0.1342	0.7142
Seminar_Attend	1	1	2.9872	0.7805	14.6490	0.0001
Impact_Proj_Stud_Suc	1	1	1.2196	0.8109	2.2617	0.1326
Impact_Proj_Stud_Suc	2	1	5.9951	1.8041	11.0428	0.0009
Prep_Mid_Term_Exam_1	1	1	-2.5780	1.2902	3.9926	0.0457
Prep_Mid_Term_Exam_1	2	1	-0.9625	1.2992	0.5488	0.4588
Prep_Mid_Term_Exam_2	1	1	0.8886	2.8283	0.0987	0.7534
Prep_Mid_Term_Exam_2	2	1	3.2362	2.9185	1.2295	0.2675
Taking_Notes_Class		1	0.6271	0.4688	1.7892	0.1810
Listening_Class	1	1	-1.4053	0.7801	3.2453	0.0716
Listening_Class	2	1	-1.7336	0.6617	6.8637	0.0088
Discussions	1	1	1.1670	1.0940	1.1379	0.2861
Discussions	2	1	0.0490	0.4957	0.0098	0.9213
Flip_Class	1	1	1.0506	0.7281	2.0820	0.1490
Flip_Class	2	1	2.2280	0.6867	10.5268	0.0012
Cumal_GPA_Last_Sem		1	0.8953	0.2753	10.5784	0.0011
Attendance	1	1	-0.8485	0.5611	2.2866	0.1305
Expected_GPA_Grad		1	0.2203	0.3767	0.3420	0.5587

2	6	12
3	5	14
4	4	10
5	3	18
6	2	20
7	1	33
8	0	7

Probabilities modeled are cumulated over the lower Ordered Values.

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

5 Anova

The ANOVA Procedure

Class Level Information		
Class	**Levels**	**Values**
Attendance	2	1 2

Number of Observations Read	131
Number of Observations Used	131

The ANOVA Procedure

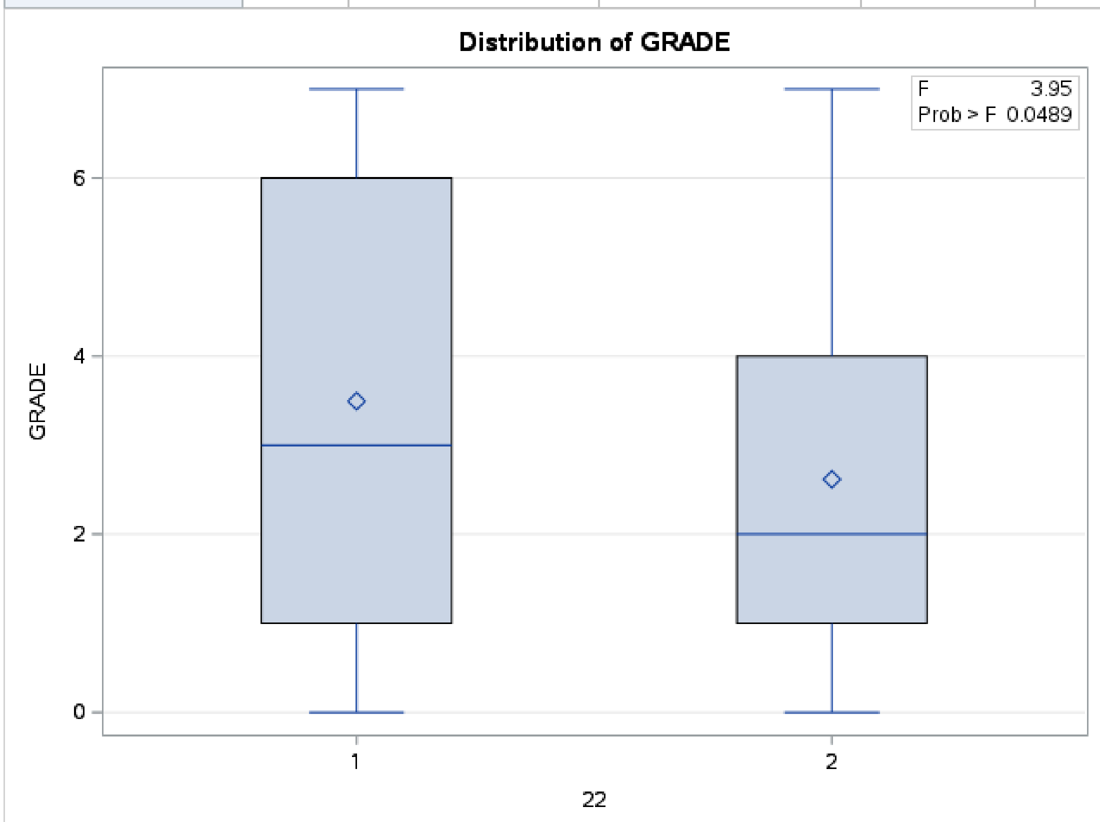
Dependent Variable: GRADE

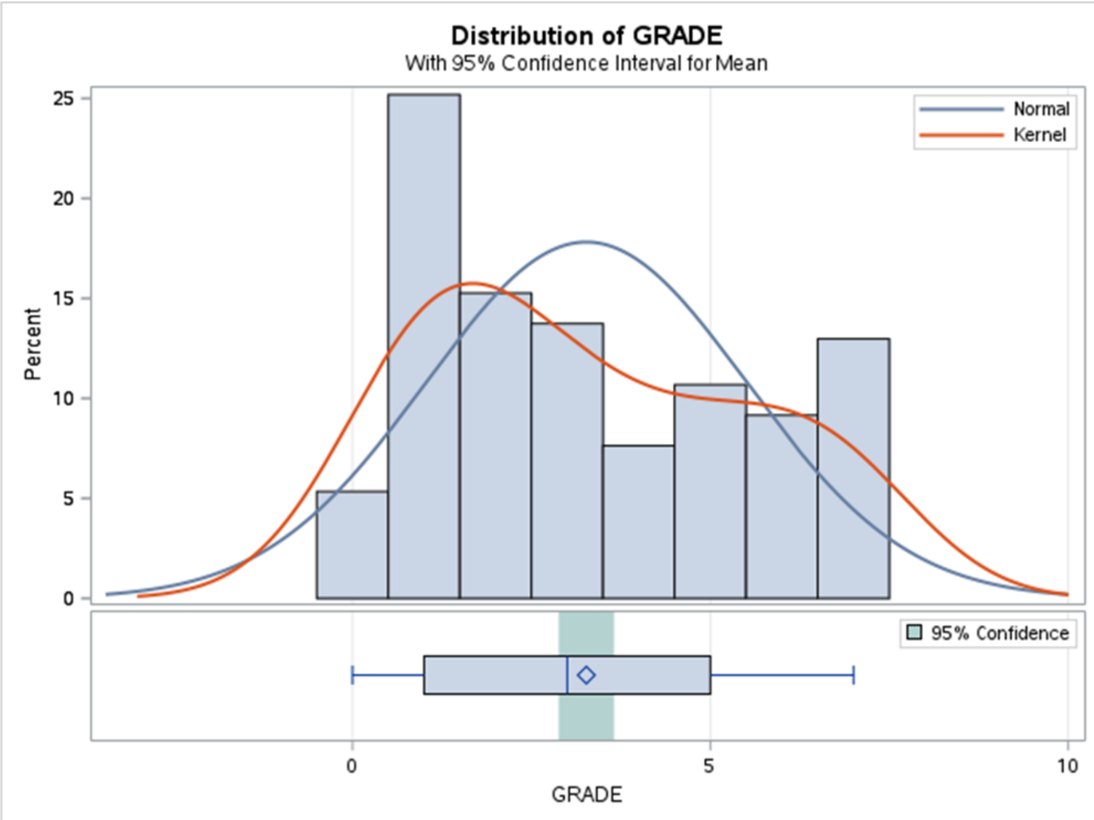
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	19.3720205	19.3720205	3.95	0.0489

Error	129	632.2768344	4.9013708		
Corrected Total	130	651.6488550			

R-Square	Coeff Var	Root MSE	GRADE Mean
0.029728	67.76201	2.213904	3.267176

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Attendance	1	19.37202052	19.37202052	3.95	0.0489





7. R-Square and Reduced Multiple Model

The LOGISTIC Procedure

Model Information	
Data Set	WORK.GROUP_6
Response Variable	GRADE
Number of Response Levels	8
Model	cumulative logit
Optimization Technique	Fisher's scoring

Number of Observations Read	131

Number of Observations Used	131

Response Profile

Ordered Value	GRADE	Total Frequency
1	7	17
2	6	12
3	5	14
4	4	10
5	3	18
6	2	20
7	1	33
8	0	7

Probabilities modeled are cumulated over the lower Ordered Values.

Backward Elimination Procedure

Step 0. The following effects were entered:

Intercept_7 Intercept_6 Intercept_5 Intercept_4 Intercept_3 Intercept_2 Intercept_1 Student_Age Sex Grad_High_School_Typ Additional_Mark Transportation Accommodation Fathers_Occu Weekly_Stu_Hour Read_Freq_Non_Scien Seminar_Attend Impact_Proj_Stud_Suc Prep_Mid_Term_Exam_2 Listening_Class Flip_Class Cumal_GPA_Last_Sem

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Score Test for the Proportional Odds Assumption

Chi-Square	DF	Pr > ChiSq
214.8225	90	<.0001

Model Fit Statistics		
Criterion	**Intercept Only**	**Intercept and Covariates**
AIC	533.494	493.637
SC	553.621	556.892
-2 Log L	519.494	449.637

R-Square	0.4133	**Max-rescaled R-Square**	0.4213

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	69.8571	15	<.0001
Score	51.9256	15	<.0001
Wald	58.9202	15	<.0001

Note: No (additional) effects met the 0.5 significance level for removal from the model.

Analysis of Maximum Likelihood Estimates							
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**	**Exp(Est)**

Intercept	**7**	1	-10.2481	2.0224	25.6771	<.0001	0.000
Intercept	**6**	1	-9.3499	1.9919	22.0335	<.0001	0.000
Intercept	**5**	1	-8.5636	1.9660	18.9744	<.0001	0.000
Intercept	**4**	1	-8.0902	1.9511	17.1933	<.0001	0.000
Intercept	**3**	1	-7.2679	1.9273	14.2211	0.0002	0.001
Intercept	**2**	1	-6.3609	1.9076	11.1191	0.0009	0.002
Intercept	**1**	1	-3.8515	1.9011	4.1044	0.0428	0.021
Student_Age		1	-1.0848	0.3227	11.2984	0.0008	0.338
Sex		1	1.2763	0.3829	11.1099	0.0009	3.583
Grad_High_School_Typ		1	0.6199	0.3195	3.7655	0.0523	1.859
Additional_Mark		1	1.0784	0.3925	7.5488	0.0060	2.940
Transportation		1	-0.4786	0.1740	7.5629	0.0060	0.620
Accommodation		1	0.4339	0.2411	3.2378	0.0720	1.543

Fathers_Occu		1	-0.2227	0.1319	2.8495	0.0914	0.800
Weekly_Stu_Hour		1	-0.1465	0.1930	0.5763	0.4478	0.864
Read_Freq_Non_Scien		1	1.3350	0.3288	16.4896	<.0001	3.800
Seminar_Attend		1	-0.4618	0.4495	1.0552	0.3043	0.630
Impact_Proj_Stud_Suc		1	-0.3164	0.2835	1.2456	0.2644	0.729
Prep_Mid_Term_Exam_2		1	0.7620	0.4335	3.0903	0.0788	2.143
Listening_Class		1	0.7677	0.2612	8.6422	0.0033	2.155
Flip_Class		1	-0.3261	0.2110	2.3881	0.1223	0.722
Cumal_GPA_Last_Sem		1	0.5385	0.1416	14.4751	0.0001	1.714

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	75.3	**Somers' D**	0.510

Percent Discordant	24.4	**Gamma**	0.511
Percent Tied	0.3	**Tau-a**	0.436
Pairs	7285	**c**	0.755

Odds Ratio Estimates and Wald Confidence Intervals

Effect	Unit	Estimate	95% Confidence Limits	
Student_Age	1.0000	0.338	0.180	0.636
Sex	1.0000	3.583	1.692	7.590
Grad_High_School_Typ	1.0000	1.859	0.994	3.477
Additional_Mark	1.0000	2.940	1.362	6.346
Transportation	1.0000	0.620	0.441	0.872
Accommodation	1.0000	1.543	0.962	2.476
Fathers_Occu	1.0000	0.800	0.618	1.037
Weekly_Stu_Hour	1.0000	0.864	0.592	1.261
Read_Freq_Non_Scien	1.0000	3.800	1.995	7.238
Seminar_Attend	1.0000	0.630	0.261	1.521
Impact_Proj_Stud_Suc	1.0000	0.729	0.418	1.270
Prep_Mid_Term_Exam_2	1.0000	2.143	0.916	5.011

Listening_Class	1.0000	2.155	1.292	3.595
Flip_Class	1.0000	0.722	0.477	1.091
Cumal_GPA_Last_Sem	1.0000	1.714	1.298	2.261

The LOGISTIC Procedure

Model Information	
Data Set	WORK.GROUP_6
Response Variable	GRADE
Number of Response Levels	8
Model	cumulative logit
Optimization Technique	Fisher's scoring

Number of Observations Read	131
Number of Observations Used	131

Response Profile		
Ordered Value	**GRADE**	**Total Frequency**
1	0	7

2	1	33
3	2	20
4	3	18
5	4	10
6	5	14
7	6	12
8	7	17

Probabilities modeled are cumulated over the lower Ordered Values.

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Score Test for the Proportional Odds Assumption		
Chi-Square	**DF**	**Pr > ChiSq**
214.8225	90	<.0001

Model Fit Statistics		
Criterion	**Intercept Only**	**Intercept and Covariates**
AIC	533.494	493.637
SC	553.621	556.892
-2 Log L	519.494	449.637

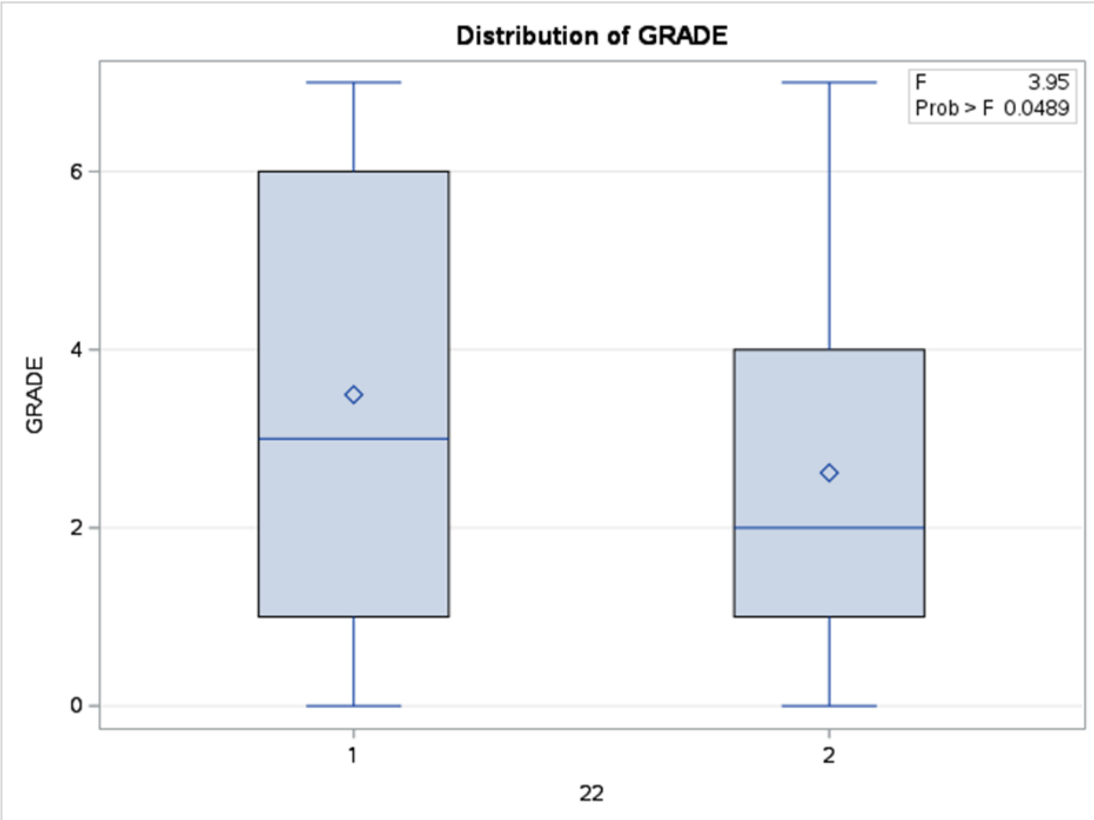
R-Square	0.4133	Max-rescaled R-Square	0.4213

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	69.8571	15	<.0001
Score	51.9256	15	<.0001
Wald	58.9202	15	<.0001

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	75.3	Somers' D	0.510
Percent Discordant	24.4	Gamma	0.511
Percent Tied	0.3	Tau-a	0.436
Pairs	7285	c	0.755

Odds Ratio Estimates and Wald Confidence Intervals				
Effect	Unit	Estimate	95% Confidence Limits	
Student_Age	1.0000	2.959	1.572	5.570
Sex	1.0000	0.279	0.132	0.591
Grad_High_School_Typ	1.0000	0.538	0.288	1.006

Additional_Mark	1.0000	0.340	0.158	0.734
Transportation	1.0000	1.614	1.147	2.270
Accommodation	1.0000	0.648	0.404	1.039
Fathers_Occu	1.0000	1.249	0.965	1.618
Weekly_Stu_Hour	1.0000	1.158	0.793	1.690
Read_Freq_Non_Scien	1.0000	0.263	0.138	0.501
Seminar_Attend	1.0000	1.587	0.658	3.830
Impact_Proj_Stud_Suc	1.0000	1.372	0.787	2.392
Prep_Mid_Term_Exam_2	1.0000	0.467	0.200	1.092
Listening_Class	1.0000	0.464	0.278	0.774
Flip_Class	1.0000	1.386	0.916	2.095
Cumal_GPA_Last_Sem	1.0000	0.584	0.442	0.770



Tables

Variables included in the model

Timeline

1st Submission
• **Initial Date: 22 July 2024**
• **First Draft of the Research Proposal:**
o **Date: 24 July 2024**
o **Contributors: Arehone Matodzi, Busi Manzini, Lulamile Manzini, Mthetheleli Simphiwe Mbeje, Msindiseni Mathebula**
o **Notes: Completed on 24 July 2024.**
• **Plagiarism Check on Turnitin:**
o **Date: 24 July 2024**
o **Conducted by: Arehone Matodzi**
o **Notes: Completed on 24 July 2024.**
• **First Draft Feedback Consultation:**
o **Date: 26 July 2024**
o **Participants: Arehone Matodzi, Busi Manzini, Lulamile Manzini, Mthetheleli Simphiwe Mbeje, Msindiseni Mathebula**
o **Notes: Met to discuss feedback received from Mr. Matthew Wayne Valentine.**
• **Second Draft:**
o **Date: 27 July 2024**
o **Contributors: Arehone Matodzi, Busi Manzini, Lulamile Manzini, Mthetheleli Simphiwe Mbeje, Msindiseni Mathebula**
o **Notes: Completed on 27 July 2024.**
• **Submission of Second Draft:**
o **Date: 29 July 2024**
o **Submitted by: Arehone Matodzi**
o **Notes: Submitted on 29 July 2024.**
• **Consultation with Mr. Matthew:**

o Date: 29 July 2024
o Participants: Arehone Matodzi, Busi Manzini, Lulamile Manzini, Mthetheleli Simphiwe Mbeje, Msindiseni Mathebula
o Notes: Met to discuss feedback from Mr. Matthew Wayne Valentine.
• Final Draft of Research Abstract:
o Date: 30 July 2024
o Contributors:
♣ Introduction: Lulamile Manzini
♣ Body: Busi Manzini, Mthetheleli Simphiwe Mbeje
♣ Conclusion: Arehone Matodzi, Msindiseni Mathebula
o Notes: Completed on 30 July 2024.
• Due Date for Submission: 02 August 2024
o Final Submission: Completed by Busi Manzini.
2nd Submission
• Initial Date: 02 August 2024
• Due Date: 16 August 2024
• Drafting the Second Submission:
o Start Date: 03 August 2024
o Responsible Party: Busi Manzini
o Notes: Begin drafting the submission.
• Feedback Consultation:
o Date: 10 August 2024
o Participants: Arehone Matodzi, Busi Manzini, Lulamile Manzini, Mthetheleli Simphiwe Mbeje, Msindiseni Mathebula
o Notes: Schedule a meeting with Mr. Matthew for feedback.

• **Submission of Second Submission:**
o **Due Date: 16 August 2024**
o **Submitted by: Manzini Lulamile**
o **Notes: Finalize and submit the document.**
3rd Submission
• **Initial Date: 16 August 2024**
• **Due Date: 30 August 2024**
• **Drafting the Third Submission:**
o **Start Date: 17 August 2024**
o **Responsible Party: Arehone Matodzi, Busi Manzini, Lulamile Manzini, Mthetheleli Simphiwe Mbeje, Msindiseni Mathebula**
o **Notes: Begin drafting the submission.**
• **Feedback Consultation:**
o **Date: 24 August 2024**
o **Participants: All team members**
o **Notes: Schedule a meeting with Mr. Matthew for feedback.**
• **Submission of Third Submission:**
o **Due Date: 30 August 2024**
o **Submitted by: Mathebula Msindiseni**
o **Notes: Finalize and submit the document.**
4th Submission
• **Initial Date: 30 August 2024**
• **Due Date: 20 September 2024**

• **Drafting the Fourth Submission:**
o **Start Date: 31 August 2024**
o **Responsible Party: [Name to be determined]**
o **Notes: Begin drafting the submission.**
• **Feedback Consultation:**
o **Date: 10 September 2024**
o **Participants: All team members**
o **Notes: Schedule a meeting with Mr. Matthew for feedback.**
• **Submission of Fourth Submission:**
o **Due Date: 20 September 2024**
o **Submitted by: Arehone Matodzi**
o **Notes: Finalize and submit the document.**
Final Submission
• **Initial Date: 20 September 2024**
• **Due Date: 07 October 2024**
• **Drafting the Final Submission:**
o **Start Date: 21 September 2024**
o **Responsible Party: [Name to be determined]**
o **Notes: Begin drafting the final submission.**
• **Feedback Consultation:**
o **Date: 30 September 2024**
o **Participants: All team members**
o **Notes: Schedule a meeting with Mr. Matthew for feedback.**
• **Submission of Final Paper:**
o **Due Date: 07 October 2024**

o **Submitted by: Mthetheleli Simphiwe**
o **Notes: Finalize and submit the document.**

References

1. Ancheta, D. M., Reyes, P. M., & Sison, C. P. (2021). The impact of class attendance on academic performance: A study among university students. *Journal of Educational Research*, *114*(4), 301-310.
2. Baker, D. P., & Gruber, K. J. (2019). Siblings and academic achievement: An exploration of how birth order influences educational outcomes. *Journal of Family Issues*, *40*(5), 610-635.
3. Cheung, C. S., & Chan, A. C. (2010). The role of parental involvement in children's academic performance: A study of Chinese families in Hong Kong. *Educational Psychology*, *30*(3), 273-290.
4. Cohen, R., & Kessel, M. (2018). Balancing work and academics: A study of the impact of part-time employment on college students' academic performance. *Journal of College Student Development*, *59*(1), 23-36. https://doi.org/10.1353/csd.2018.0003
5. Duncan, G. J., & Dowsett, C. J. (2020). Family resources and children's academic achievement: The importance of siblings. *Child Development*, *91*(1), e99-e113. https://doi.org/10.1111/cdev.13238
6. Gonzalez, J. (2019). Parental influences on academic achievement in high school students: A longitudinal study. *Journal of Educational Psychology*, *111*(2), 203-216. https://doi.org/10.1037/edu0000289
7. Hirschi, A., & Jaensch, V. (2015). The impact of work on academic performance: A study of the work-life balance of students. *Journal of Vocational Behavior*, *88*, 101-110. https://doi.org/10.1016/j.jvb.2015.02.001
8. Hirsch, E. D. (2016). *Why knowledge matters: Rescuing our children from failed educational theories*. Yale University Press.
9. Hsin, A., & Farkas, G. (2009). Cultural capital and the academic achievement of siblings: The role of family dynamics in educational outcomes. *Sociology of Education*, *82*(4), 363-383. https://doi.org/10.1177/003804070908200404
10. Judith, L., Smith, A. B., & Brown, C. D. (2015). Parental education and academic achievement among adolescents: The role of family dynamics. *Educational Studies*, *41*(1), 45-62.

11. Lareau, A. (2011). *Unequal childhoods: Class, race, and family life*. University of California Press.
12. McKool, S., & McLaughlin, T. (2008). The relationship between student engagement and academic achievement. *Literacy Research and Instruction*, 47(1), 1-16.
13. Nell, V. (2019). *Lost in a Book: The Psychology of Reading for Pleasure*. Yale University Press.
14. Smith, L., & Luthar, S. S. (2021). *The impact of reading habits on academic performance*. *Educational Psychology*, 41(1), 20-34. Wang, Y., Zhang, L., & Ma, Y. (2021). The effects of part-time work on college students' academic performance: The mediating role of time management skills. *Educational Psychology*, *41*(5), 657-671.
15. Yılmaz, H., & Sekeroglu, R. (2020). The role of classroom discussions in enhancing students' understanding of concepts. *International Journal of Instruction*, *13*(2), 245-258. https://doi.org/10.29333/iji.2020.13216a