Influence of Alcohol Consumption and Impact on Academic Performance

Harikrishna Prathipati

17/06/2022

## Summary

The project’s main aim involved examining the correlation between academic performance and alcohol consumption of students. Among students, alcohol has posed a significant problem associated with many negative consequences. According to Martinez et al. (2014), approximately 37.4% of students in the United States reported high alcohol consumption. Therefore, this project involved finding the association between the effects of alcohol on students’ academic performance and alcohol consumption among students. Other negative consequences of students’ alcohol consumption have also been examined in this project. From the analysis of variables for this project, it has been identified that there has been a negative correlation between the student’s GPA and the frequency of alcohol consumption. It was clear from the analysis that with the increased number of nights per week that the students consumed alcohol, there was a decrease in the grade point average for these students.

## Introduction

In most cases, it is during college that most students are allowed to make their own decisions. However, sometimes these decisions come in with alcohol consumption. In life, factors like social pressure, the easier availability of alcohol, social pressure, and independence are identified as the leading factors for alcohol consumption for students. Alcohol consumption has been identified as one of the largest parts of the culture regardless of whether the students decide to drink alcohol or not. The project has focused on the alcohol consumption impact on the student’s academic performance. In most cases, it has been identified that students tend to indulge in alcohol consumption to cope with negative emotions or stress. However, no matter why the students decide to use alcohol, alcohol consumption is a serious problem, leading to failing classes and potential dropouts.

## Literature Review

Several studies have been conducted on the association between academic performance and students' number of nights of alcohol consumption. According to Matt et al. (2012), the authors were able to identify that as the number of nights for alcohol consumption increased, there was a significant decrease in the academic performance among students. In another study by Singleton & Wolfston (2009), the authors researched alcohol consumption among students and the negative impact on their academic performance. Their research concluded that students involved in high alcohol consumption have poor sleep patterns, negatively affecting their academic performance. Pizza-Gardener (2016) researched the relationship between alcohol consumption and GPA in her study. From the research, the author identified that students with lower GPAs were involved in higher alcohol consumption than those with higher GPAs. The study by Liguori (2015) identified the correlation between student retention and alcohol consumption. From this study, it was identified that there was alcohol dependency, failure and dropouts, and academic impairment as some of the consequences associated with alcohol consumption. From the literature, there have been no statistically identified leading factors that lead to higher consumption of alcohol among students. Therefore, there is a need to develop this project. The project now aims to implement the machine learning algorithms in R to test whether there is any correlation between the students’ academic performance and the alcohol consumption among students. The project will also identify the main factors contributing to higher alcohol consumption among students.

## Theory

H1: There will be a significant relationship between frequent alcohol consumption among students and lower academic performance. H1: increased alcohol consumption contributes to a decrease in exam performance.

## Data

The data used for this project will be retrieved from <https://www.kaggle.com/uciml/student-alcohol-consumption> and it includes

## school sex age address famsize Pstatus Medu Fedu Mjob Fjob reason

## 1 GP F 18 U GT3 A 4 4 at\_home teacher course

## 2 GP F 17 U GT3 T 1 1 at\_home other course

## 3 GP F 15 U LE3 T 1 1 at\_home other other

## 4 GP F 15 U GT3 T 4 2 health services home

## 5 GP F 16 U GT3 T 3 3 other other home

## 6 GP M 16 U LE3 T 4 3 services other reputation

## guardian traveltime studytime failures schoolsup famsup paid activities

## 1 mother 2 2 0 yes no no no

## 2 father 1 2 0 no yes no no

## 3 mother 1 2 3 yes no yes no

## 4 mother 1 3 0 no yes yes yes

## 5 father 1 2 0 no yes yes no

## 6 mother 1 2 0 no yes yes yes

## nursery higher internet romantic famrel freetime goout Dalc Walc health

## 1 yes yes no no 4 3 4 1 1 3

## 2 no yes yes no 5 3 3 1 1 3

## 3 yes yes yes no 4 3 2 2 3 3

## 4 yes yes yes yes 3 2 2 1 1 5

## 5 yes yes no no 4 3 2 1 2 5

## 6 yes yes yes no 5 4 2 1 2 5

## absences G1 G2 G3

## 1 6 5 6 6

## 2 4 5 5 6

## 3 10 7 8 10

## 4 2 15 14 15

## 5 4 6 10 10

## 6 10 15 15 15

To understand data and identify the variables which the project will be working, the names of the variables have been displayed below.

## [1] "school" "sex" "age" "address" "famsize"

## [6] "Pstatus" "Medu" "Fedu" "Mjob" "Fjob"

## [11] "reason" "guardian" "traveltime" "studytime" "failures"

## [16] "schoolsup" "famsup" "paid" "activities" "nursery"

## [21] "higher" "internet" "romantic" "famrel" "freetime"

## [26] "goout" "Dalc" "Walc" "health" "absences"

## [31] "G1" "G2" "G3"

Using dim()function, the size of the data was displayed.

dim(student\_data)

## [1] 395 33

The dataset is made up of 395 records and 33 total variables.

colSums(is.na(student\_data))

## school sex age address famsize Pstatus Medu

## 0 0 0 0 0 0 0

## Fedu Mjob Fjob reason guardian traveltime studytime

## 0 0 0 0 0 0 0

## failures schoolsup famsup paid activities nursery higher

## 0 0 0 0 0 0 0

## internet romantic famrel freetime goout Dalc Walc

## 0 0 0 0 0 0 0

## health absences G1 G2 G3

## 0 0 0 0 0

There are no missing values in the column.

## Methodology

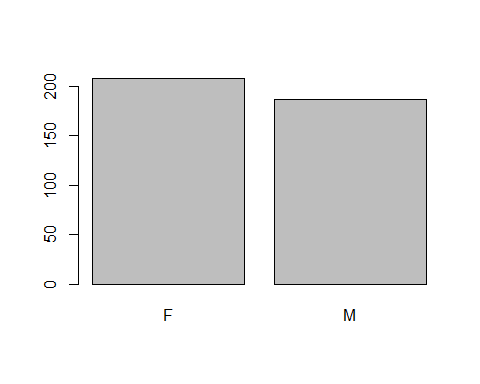
Having identified that the data is clean, the next step was to work with the student\_data and gain insights from the data regarding alcohol consumption among students, the reasons for alcohol consumption, and the effects of alcohol consumption on their academic performance.

First, the project checks if alcohol consumption for the student is influenced by the student's social-economic status and biological factors.

## student\_gender

## F M

## 208 187



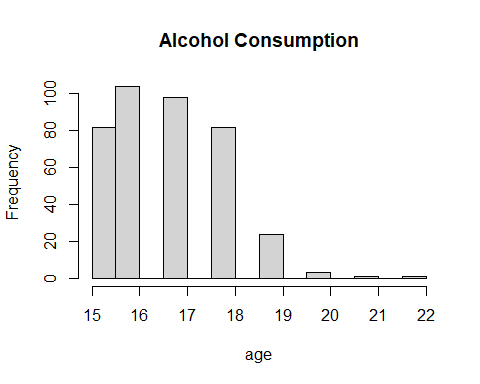
It is clear that there are more female students in the data than male students.

studentage = student\_data$age

# convert studentage variable to a numeric vector

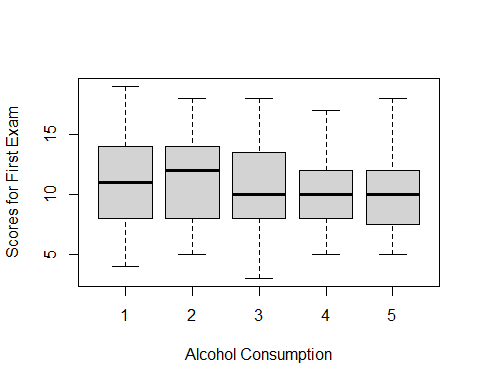
age\_numeric = as.numeric(studentage)

Once the age variable was converted to a numeric vector, a histogram was created to illustrate the alcohol consumption among students by age.

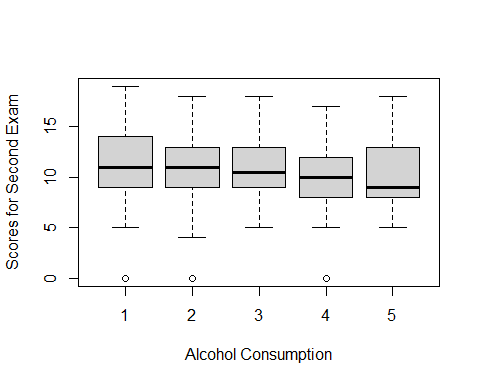


## [1] 16.6962

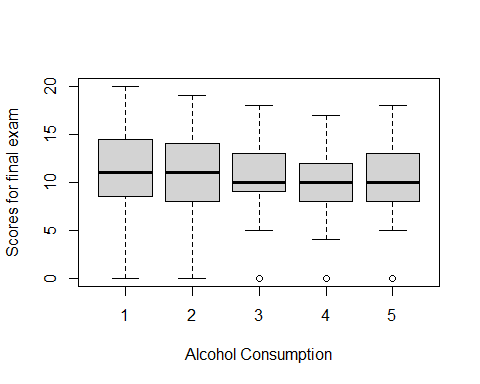
Next, to identify any correlation between exam performance and level of alcohol consumption for students, the project involved creating several boxplots. The first boxplot includes the distribution of scores for the first exam and level of alcohol consumption.



The second boxplot includes the distribution of scores for the second exam and the level of alcohol consumption.

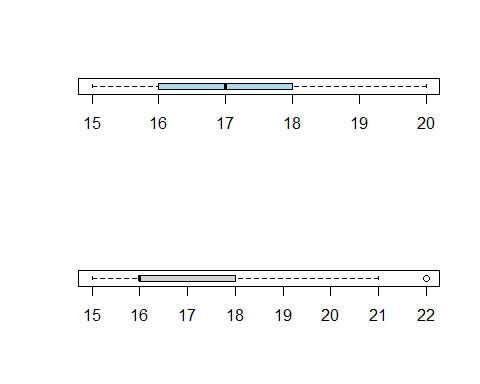


The third boxplot includes the distribution of the scores for the final exam and the level of alcohol consumption for the students.



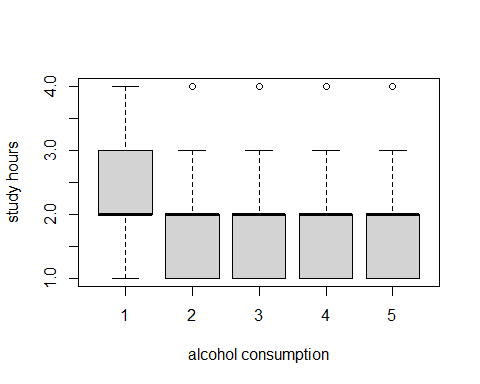
## Results

After analyzing the significant variables and identifying the major research questions for the project in the dataset, several insights were gained.

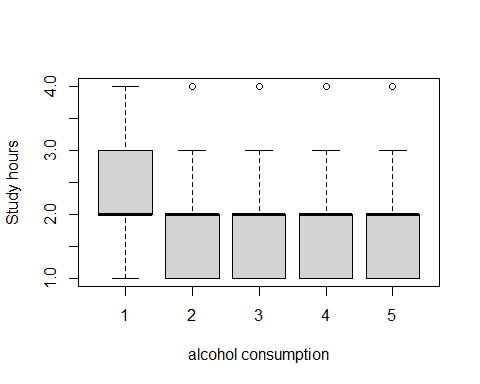


The results for age and alcohol consumption illustrated that alcohol consumption is more among students of age between 16-17. In addition, gender has been identified not to have a significant role in the influence of alcohol consumption among students.

In regards to studying time and alcohol consumption, results show that there is a very strong negative correlation. For example, there is the same value of 2 hours of study for the 25th percentile, the median, and alcohol consumption of level 1, which indicates that 25% of students study for 2 hours and 50% of the students always study more than 2 hours.



Regarding the correlation between exam performance and alcohol consumption, it has been identified that an increase in alcohol consumption has a negative correlation with exam performance. By assumption, students who drink less have more time to study, and their exam performance may be more affected by the study time. This assumption is supported by the boxplot below, where it is clear that an increase in alcohol consumption significantly decreases the study time.



## Conclusion

The theory that increased alcohol consumption contributes to a decrease in exam performance was true. There is some degree of correlation between small levels of alcohol consumption with improved exam scores. There is a decrease in exam scores with a high level of alcohol consumption. High levels of alcohol consumption among students significantly harm their ability to perform.

# References

Matt, R., Shannon, C. & Oliver, H. (2012). Does alcohol consumption affect academic performance? Retrieved from <http://www.shannclark.weebly.com/uploads/2/5/1/4/25149157/the_effect_of_alcohol_consumption_on_academic_performance.pdf>

Singleton, R. A., & Wolfson, A. R. (2009). Alcohol consumption, sleep, and academic performance among students. Journal of studies on alcohol and drugs, 70(3), 355-363.

Piazza-Gardner, A. K., Barry, A. E., & Merianos, A. L. (2016). Assessing drinking and academic performance among a nationally representative sample of students. Journal of Drug Issues, 46(4), 347-353.