

Project Proposal

Our group decided to analyze the data named “Student Alcohol Consumption”. This dataset contains a wide range of variables related to students' behavior and alcohol consumption. The data were obtained in a survey of students in secondary school mathematics and Portuguese language courses. It contains a lot of interesting social, gender, academic and health information about the students. As well as information about the student's family situation, why they chose this school, and the time spent in their life after school and academic performance. All of these data can effectively help us understand the relationship and impact between student alcohol consumption and other variables. You can find this database from Kaggle, which is a public database website that everyone can download.

There are multiple reasons why this dataset is interesting to us. First is that it can see the alcohol consumption habits pattern of students among different ages and diversified backgrounds. This information can help researchers or policymakers to understand the prevalence of alcohol use among students. What's more, it provides a chance to observe potential health hazards linked to student alcohol use. Researchers can use this dataset to have a general understanding between alcohol consumption and general well-being by looking at factors including health status, hazardous conduct, and academic performance. At the same time, this data set explores the causes and future academic impact of students' alcohol use from multiple perspectives and variables. For example, family factors, school location, and distribution of time outside of school. These are all issues that we would normally overlook, but are reflected in this data. We also hope to analyze the data from these two schools to predict and intervene in the future on a wider range of factors influencing student alcohol use.

First query (Yingchen Y.):

The purpose of this query was to investigate the potential impact of the frequency of going out with friends (Goout & Freetime) in their free time on students' drinking consumption (Dalc & Walc). Specifically, as the frequency of hanging out increases, does students' alcohol consumption also increase accordingly.

The result could also be used to indicate the influence of social factors in students' alcohol-related behaviors. By further exploring the underlying factors that contribute to this relationship, we can better understand the interplay between different social activities and student alcohol consumption as well.

Overall, this query can show the potential impact of interactions with friends on student alcohol consumption. If a bad relationship does exist, then prevention strategies are needed to address the negative effects of social interactions in promoting excessive drinking habits among students.

Second query (Yingchen Y.):

The purpose of this query was to explore the potential relationship between students' Internet (Internet) use and alcohol consumption (Dalc & Walc). By examining these variables, we can gain insight into the impact of the Internet on alcohol-related behaviors.

On online platforms, students have many opportunities to browse alcohol-related content, like advertisements. This frequent exposure may influence students' behaviors regarding alcohol. Thus, enhancing the consumption of alcohol. However, at the same time, we cannot ignore the positive impact that the Internet may have on alcohol consumption. For example, the Internet can show the disadvantages of excessive alcohol consumption, which may reduce students' consumption of alcohol.

Therefore, we need this query to study the relationship between alcohol consumption and Internet use. It will help develop targeted interventions to mitigate potential negative effects while using the positive aspects of online platforms to promote proper alcohol consumption.

Third query (Katherine Wei):

This query shows the relationship between alcohol consumption (both daily and weekend) and two crucial academic factors: study time and past class failures. By analyzing the relationship between weekly consumption and academic performance, we can know meaningful data about the relationship between the negative effects of alcohol and students' academic achievements.

A deep understanding of the relationship between alcohol consumption, study time, and the number of failed subjects can help teachers and parents take corresponding action to control or even prohibit children from contacting alcohol and make good strategies to guide students correctly to understand the danger of alcohol and help children away from alcohol.

Forth query (Katherine Wei):

This query analyzes the potential impact of study time and the number of failed courses on students' intentions to pursue higher education. It can be supplement information on the relationship between alcohol and study performance since the aspiration for higher education doesn't have a direct relationship with alcohol consumption but has a relationship with study time and failure. Besides, it can also show whether the number of past class failures and the allocation of spare time will affect students' inclination to take higher education.

Understanding the relationship between study time, past class failures, and the desire for higher education can guide teachers, students, and parents to have clear insight into students' self-performances in education and their willingness to take higher education. Besides, it can also help teachers and parents to make up strategies to support students' academic journeys. They can adjust the educational method based on students' demands and help them achieve their academic goals.

Fifth query (Bonnie):

This query examines the relationship between first period grade (G1), second period grade (G2), and final grade (G3), the variables of workday alcohol consumption (Dalc) , and weekend alcohol consumption (Walc). By analyzing these variables together, we can gain insights into how alcohol consumption affects students' academic performance in Math or Portuguese over time.

The findings from this query can provide valuable information on the potential impact of alcohol consumption on students' academic achievements. If the analysis reveals that higher levels of alcohol consumption are associated with lower grades in the first period, second period, and final exams, it suggests that alcohol consumption might hinder a student's ability to perform well in their academic subjects.

Understanding this relationship can help educators, parents, and policymakers develop targeted interventions and support systems to address the potential negative consequences of alcohol consumption on students' academic progress. It could also emphasize the importance of promoting responsible drinking habits and creating awareness of the potential impact of alcohol on academic performance in specific subjects like Math or Portuguese.

Sixth query (Ziyu Lin):

This query will explore the influence of family factors on students' alcohol use. Family factors are often an important component in shaping students' character. The amount of alcohol abuse among students will be explored by looking at variables such as the general geographic location of the family, the size of the family, whether the parents live together, and the quality of the family relationship to find out if there is a correlation.

For the data, we will pick school (column 1), sex (column 2), and age (column 3). These three will be the location of the basic information about the student. The selection of address (column 4), famsize (column 5), pstatus (column 6), and famrel (column 24) will be used to study the effect of these factors on alcohol abuse in the future. While Dalc (column 27) and Walc (column 28) were counted by the amount of alcohol abuse among students on weekdays and days off. Therefore, this query is necessary for the overall analysis.

Seventh query (Bonnie):

This query examines the relationship between parent's cohabitation status (pstatues), sex, age, mother's education (Medu), father's education (Fedu), mother's job (Malc), father's job (Dalc), and the variables of workday alcohol consumption and weekend alcohol consumption. The aim is to investigate whether parents' jobs, education, or guardian status are related to alcohol consumption.

By analyzing these variables together, we can gain insights into how various factors such as parents' cohabitation status, sex, age, education, and occupation may be associated with alcohol consumption habits.

The findings from this query can provide valuable information on the potential relationship between parental factors and alcohol consumption patterns. For example, if the analysis reveals that certain parental occupations or education levels are associated with higher alcohol consumption, it suggests that there might be a correlation between these factors and alcohol habits.

Eighth query (Ziyu Lin):

We will explore the analysis of school choice throughout the chart through the creation of this query. This database analyzes the amount of alcohol used by students at two schools, and why students of different genders choose this school may also be a reason for their alcohol use.

The choice of columns in this query is also relatively simple: school (variable 1), sex (variable 3) and reason (variable 11). The "reason" is the reason for choosing this school. These three variables were used to create the queries to derive the most important factors for choosing a school and to help in future analysis.

Ninth query (Katherine Guo):

This query will explore the relationship between extra educational support and the student who wants to take higher education. It will analyze whether students who aspire to pursue higher education receive additional educational support from their family or school. It will also compare the academic performance(grades) of students who express a desire to pursue higher education versus those who do not.

Columns picked for this query are (variable 16) Schoolsoup, (variables 17) famsup, (variable 18) paid, (variable 21) higher, and (variable G3). These variables can represent the key influencing factors of students who want to take higher education.

Tenth query (Katherine Guo):

This query aims to analyze the relationship between health conditions and the levels of alcohol consumption among students considering different age groups and sexes. It provides an opportunity to investigate potential differences in health outcomes of students based on these variables. We can gain insights that whether there is a difference in health status between male and female students. How does health status vary across different age groups of students? The age is from 15 to 22, we can group by one stage every 2 years to see the overall trends of the affection of alcohol consumption among the different groups.

Columns for this query are (variable2) sex, (variable3) age, (variable27) Dalc, (variable28) Walc, and (variable29) health, higher alcohol consumption will lead to bad health conditions. The result can be used to guide policies and programs that promote better lifestyles and lessen the possible negative health effects of alcohol intake, especially for particular age groups and genders.

(wc: 1699)