Alcohol Consumption in Schools

Hypothesis Heroes - Drew Davison, Lisa Zhang, Ellie Culman, Austin Chang

2023-11-14

Introduction and data

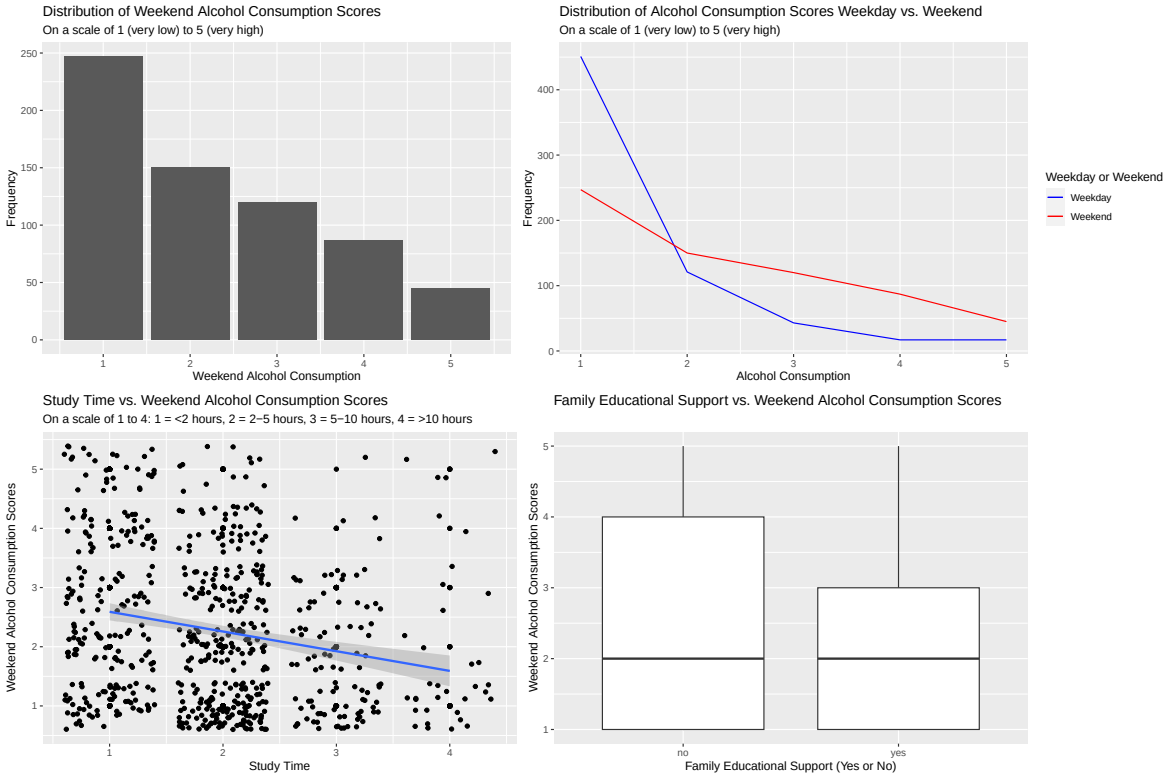
All around the world, underage drinking among students is a significant public health issue and takes a huge toll on the quality of students' lives and education. There is plenty of existing documentation on the effects of socioeconomic status on risky drinking behavior amongst college students. For example, a paper by Susan E. Collins, PhD, professor and licensed clinical psychologist, found that individuals with lower socioeconomic status as well as people of racial and ethnic minorities and homelessness experience greater alcohol-related consequences. Additionally, studies from the National Institutes of Health have found that "aspects of college life—such as unstructured time, widespread availability of alcohol, inconsistent enforcement of underage drinking laws, and limited interactions with parents and other adults" drives up rates of underage drinking, and "college students have higher binge-drinking rates and a higher incidence of driving under the influence of alcohol than their noncollege peers." Since most existing literature on underage student drinking focuses on college students, we wanted to examine the factors contributing to drinking amongst secondary school students. Our research question is as follows:

How do social indicators affect student alcohol consumption in secondary schools?

We hypothesize that factors like gender, familial status, family and school support, and other social and economic indicators will strongly influence the rates that secondary school students consume alcohol and that increased alcohol consumption is correlated with their school performance.

This data is from a Kaggle public dataset, originally sourced from the UC Irvine Machine Learning Repository. The data consists of information collected in 2008 on secondary school students from two schools in Portugal: Gabriel Pereira and Mousinho da Silveira. There are 649 observations, each one being a student, and 33 variables which cover a range of characteristics about each student's family, education, social situation, alcohol consumption, and grades in their Portuguese language class. Some key variables we will be examining are schoolsup, which is whether or not their school provides them extra academic support, famsup, which is whether or not their family provides them extra academic support, Mjob and Fjob, which is

the categories that their parents' jobs fall under, freetime and studytime, which is the amount of free time after school and amount of time spent studying on a scale of 1-5, number of absences, and amount of class failures, on a scale of 1-4. Our response variable is Walc, which is the students' average weekend alcohol consumption, on a scale of 1-5, with 5 being very high.



Methodology

```
Linear Regression Model Specification (regression)

Computational engine: lm
```

Model 1

term	estimate	std.error	statistic	p.value
(Intercept)	2.947	0.550	5.360	0.000
freetime	-0.065	0.124	-0.528	0.598
studytime	-0.497	0.207	-2.406	0.017
absences	0.050	0.013	3.721	0.000
failures	-0.299	0.320	-0.937	0.349
sex_M	0.658	0.122	5.396	0.000
schoolsup_yes	0.211	0.196	1.078	0.282
famsup_yes	-0.804	0.539	-1.493	0.136
Mjob_health	0.388	0.459	0.844	0.399
Mjob_other	-0.172	0.228	-0.753	0.452
Mjob_services	0.233	0.275	0.846	0.398
Mjob_teacher	-0.062	0.364	-0.172	0.864
Fjob_health	-0.645	0.836	-0.772	0.441
Fjob_other	-0.273	0.383	-0.712	0.477
Fjob_services	-0.258	0.403	-0.640	0.523
Fjob_teacher	-0.597	0.603	-0.990	0.323
freetime_x_studytime	0.076	0.065	1.168	0.244
failures_x_studytime	0.245	0.186	1.320	0.188
failures_x_absences	-0.013	0.048	-0.278	0.781
failures_x_studytime_x_absences	-0.006	0.031	-0.192	0.848
famsup_yes_x_Mjob_health	-1.055	1.341	-0.787	0.432
famsup_yes_x_Mjob_other	-0.124	0.609	-0.204	0.838
famsup_yes_x_Mjob_services	-0.586	0.761	-0.770	0.442
famsup_yes_x_Mjob_teacher	0.173	1.307	0.133	0.894
famsup_yes_x_Fjob_health	-0.064	1.505	-0.042	0.966
famsup_yes_x_Fjob_other	0.809	0.569	1.421	0.156
famsup_yes_x_Fjob_services	0.932	0.645	1.445	0.149
famsup_yes_x_Fjob_teacher	0.473	1.415	0.334	0.738
famsup_yes_x_Mjob_health_x_Fjob_health	1.137	1.807	0.629	0.530
famsup_yes_x_Mjob_health_x_Fjob_other	0.817	1.330	0.614	0.539
famsup_yes_x_Mjob_health_x_Fjob_services	0.479	1.382	0.346	0.729
famsup_yes_x_Mjob_other_x_Fjob_health	1.583	1.787	0.886	0.376
famsup_yes_x_Mjob_other_x_Fjob_other	-0.016	0.618	-0.026	0.980
famsup_yes_x_Mjob_other_x_Fjob_services	0.156	0.711	0.220	0.826
famsup_yes_x_Mjob_other_x_Fjob_teacher	-0.150	1.514	-0.099	0.921
famsup_yes_x_Mjob_services_x_Fjob_health	1.059	1.521	0.696	0.487
famsup_yes_x_Mjob_services_x_Fjob_other	0.233	0.777	0.299	0.765
famsup_yes_x_Mjob_services_x_Fjob_services	0.544	0.820	0.663	0.507
famsup_yes_x_Mjob_services_x_Fjob_teacher	0.341	1.542	0.221	0.825
famsup_yes_x_Mjob_teacher_x_Fjob_health	1.750	1.842	0.950	0.343
famsup_yes_x_Mjob_teacher_x_Fjob_other	0.137	1.319	0.104	0.917
famsup_yes_x_Mjob_teacher_x_Fjob_services	-0.595	1.385	-0.430	0.668
famsup_yes_x_Mjob_teacher_x_Fjob_teacher	NA	NA	NA	NA

! Important

Before you submit, make sure your code chunks are turned off with `echo: false` and there are no warnings or messages with `warning: false` and `message: false` in the YAML.