The Effects of Student Alcohol Consumption

Rayanna Harduarsingh | Sandro O'Neill

Research Problem

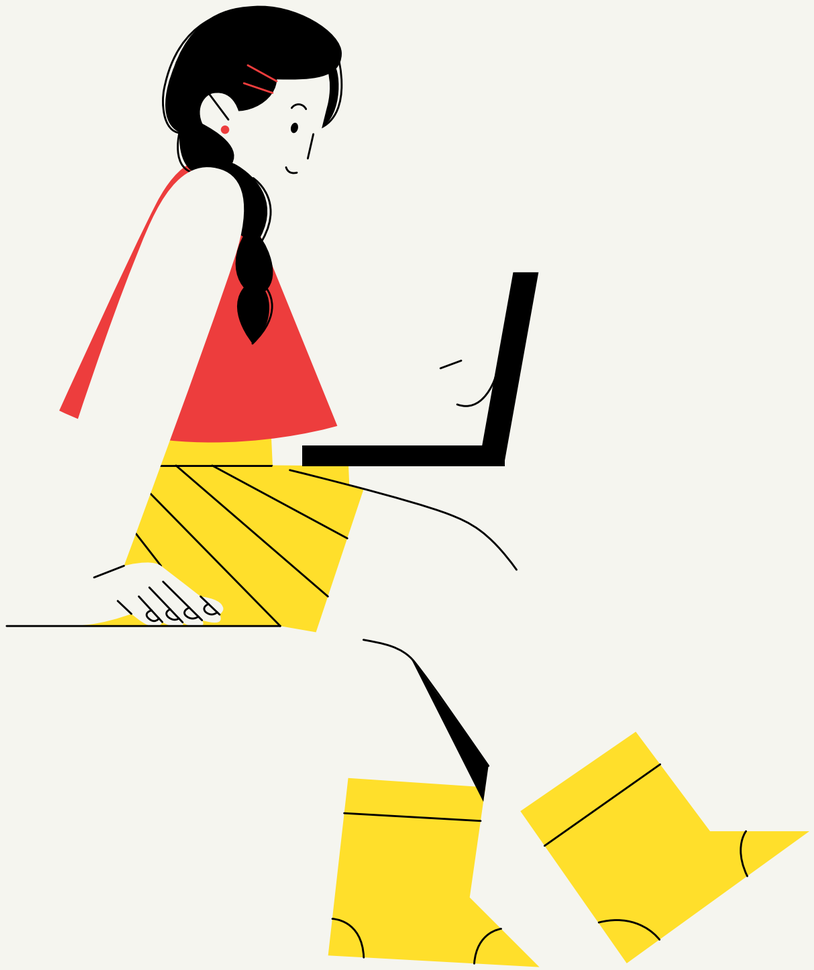
Does underage drinking affect a student's academic performance?

Underage drinking has been a serious long-term issue in the United States. About 7.05 million Americans between the ages of 12 and 20 reported current alcohol consumption. This remains a huge concern as these behaviors can lead to negative consequences and effects. One of these consequences could be the **decline in a student’s academic progress.**

Objective

What we want to achieve

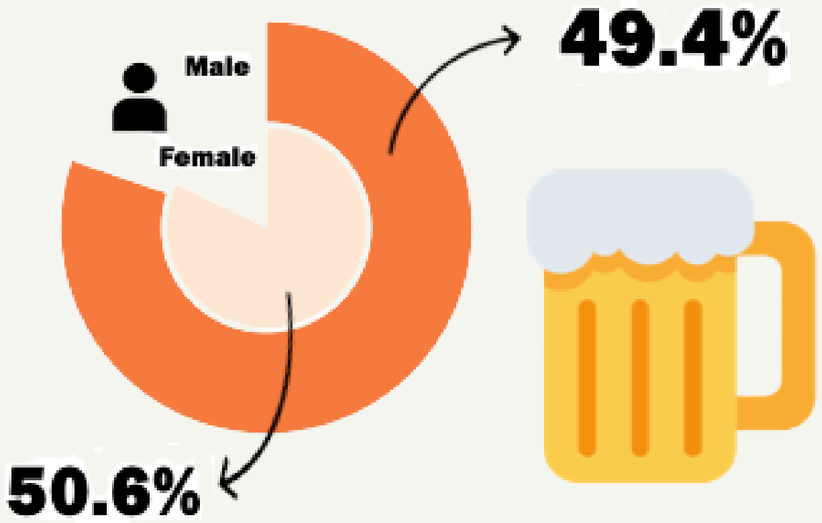
The purpose of this study is to explore the relationship between a student's alcohol consumption and demographic factors, along with their current grades and attendance to determine **if alcohol has a negative effect on their academic performance or no effect at all.**

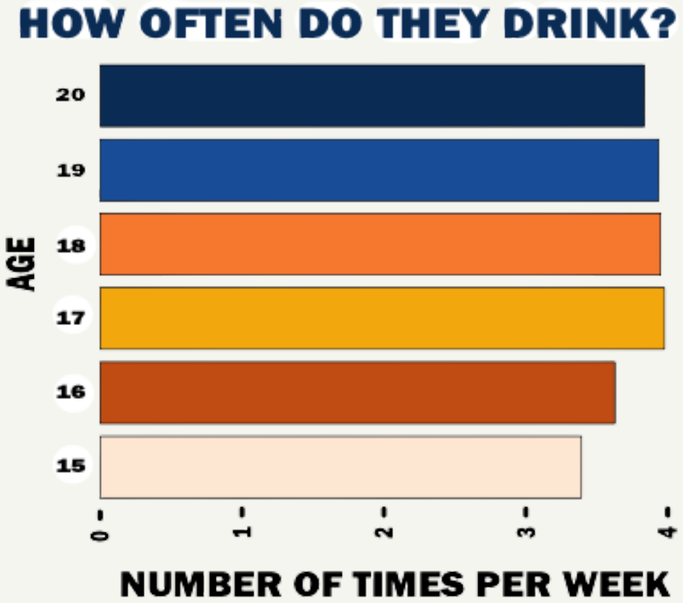


Data

Content

The University of California obtained data from a survey of students in a Portuguese language course in secondary school. It contains 33 variables and 649 observations.





Methods

1.

Regression Analysis

Explore possible connections or correlations between variables to determine positive or negative relationships.

2.

Clustering

Using k-means clustering we could segment the students based on several metrics. The results of the analysis could be helpful to the school for targeting efforts for student outreach.

3.

Classification

Using several prediction methods such as Naive Bayes, Support Vector Machine, Decision Tree, Random Tree to classify accurate relationships.

4.

Association Rule Mining

Finding common variables with students with high and low alcohol consumption metrics.

Results

Logistic Regression

Dependent Variable: Alcohol Consumption (Range 2-10)

p ≈ 0

- Increases the more a student goes out (0.59) **[1,5]**
- Increases if student is male (1.082) **[0,1]**
- Decreased by quality of family felationships (-0.326) **[1,5]**
- Increases with absences (0.05) **[0,32]**
- Health status increases (0.12) **[1,5]**
- If school choice reason is other increases (0.52) **[0,1]**
- Increasing study time decreases (0.195) **[1,5]**
- Family size less than or equal to 3 increases (0.327) **[0,1]**
- Attending nursery school decreases (-0.3567) **[0,1]**
- Mothers job in services increases (0.632) **[0,1]**
- Age increases (0.1) **[15,22]**

p < 0.05

K-means clustering

Age

17.08

Alcohol Consumption

4.46

Absences

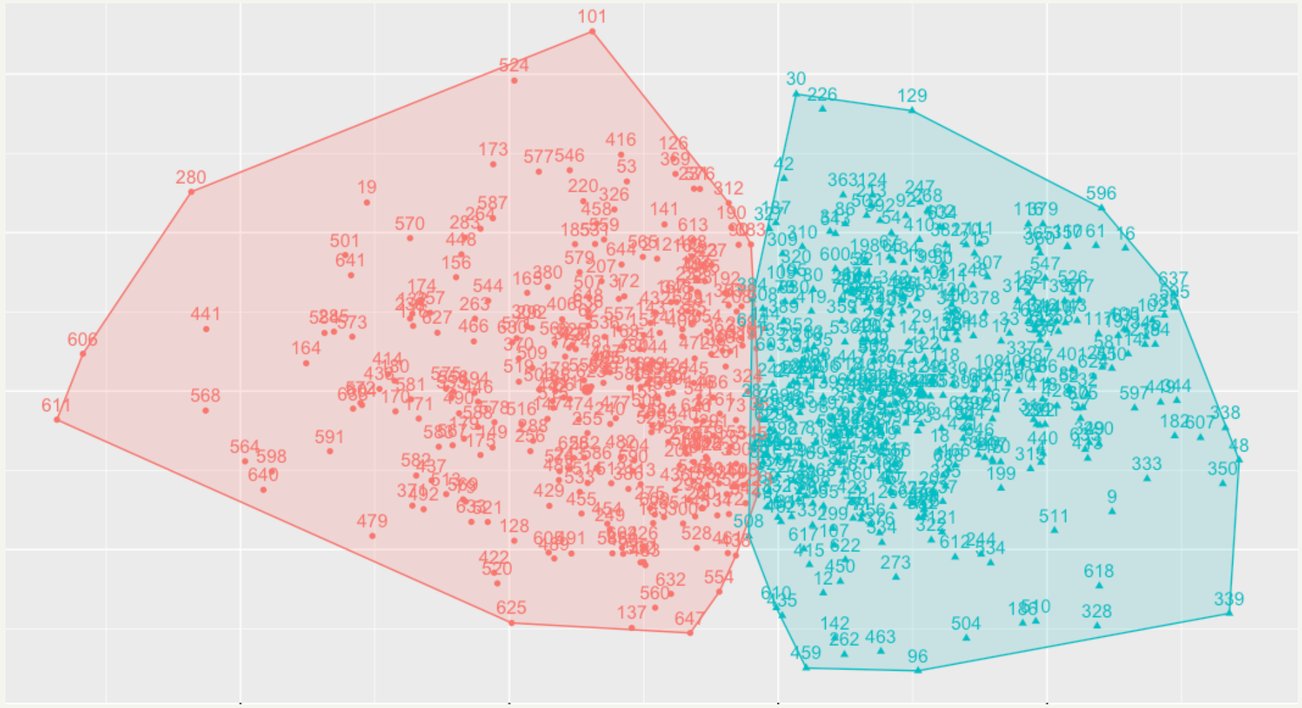
4.77

Failures

0.5

Grade

G1: 9.1
G2: 9.2
G3: 9.3



Age

16.5

Alcohol Consumption

3.3

Absences

2.83

Failures

0.02

Grade

G1: 13.1
G2: 13.4
G3: 13.8

Association Rules

- Above average alcohol consumption associates with males, high absences, non-parental guardians, and choosing school due to proximity.

- Low absences associates with low alcohol consumption, no romantic relationship, not paying for extra tutoring, and the mother having less than a primary education.

- Zero alcohol consumption associates with low going out, wanting to achieve higher education, attending nursery school, and receiving extra educational support.

- High absences associates with at least one class failure, non-parental guardian, having internet access at home, and scoring within 15% of the average.

Classification

Defaults

- 70% training
- 30% testing
- workday & weekend alcohol consumption combined into single column (2-10 scale)
- going out, free time, absences, & study time

Naïve Bayes

37.9%

SVM

39.5%

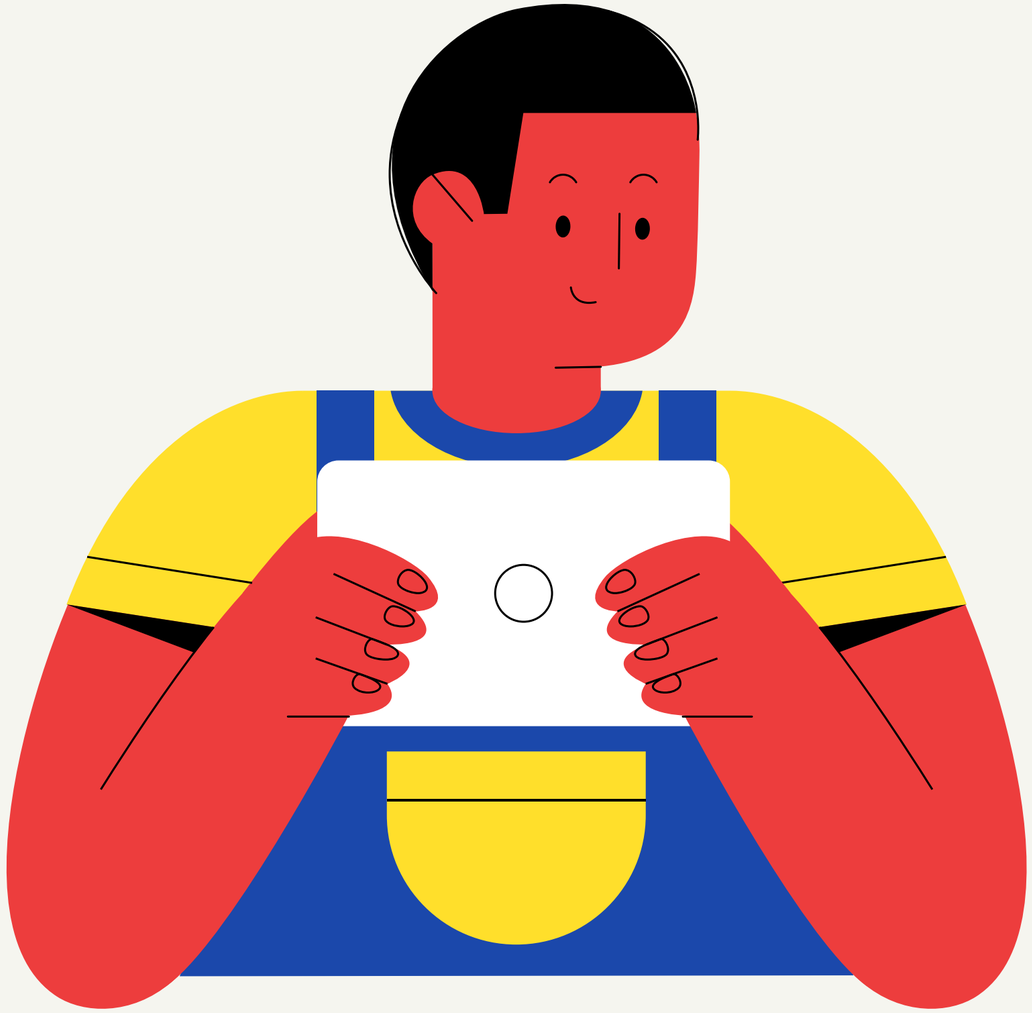
Decision Tree

34%

Random Forest

32%

Conclusion & Thoughts



Conclusion

- Regression & Association Rule Mining yielded the most insightful analysis
- Underage drinking affects students, but not dramatically based off of this data set.
    - Not much of a correlation or relationship with grades.
    - More associated with higher absences, low study time, and failures.

Thoughts

- Some data observations were repetitive leading to small accuracies for classification
- Bigger data set could have provided a more promising analysis