Does a Relationship affect school attendance?

Jason Noble

School of Engineering & Technology, National University

ANA600, Fundamentals of Analytics

Matthew Vanderbilt

August 3, 2024

Abstract

To investigate the factors influencing student absences in math classes. Specifically, we aim to examine whether being in a romantic relationship is associated with the number of absences a student has. Understanding external factors can help educators and policymakers develop strategies to reduce absences and improve academic outcomes.

Does a Relationship affect school attendance?

The dataset has various attributes related to students, including personal, familial, and academic characteristics. Variation in the number of school absences among students are influenced by multiple factors. We gathered data on about 30 variables like age, sex, family size, parental status, living situation, and social activities to try and identify any that have a significant impact on math absences.

Romantic relationship stood out

Romantic relationships can have substantial influence on teenage life and can have significant impacts on various aspects of behavior and performance, including school attendance. There are several reasons to hypothesize that being in a romantic relationship might influence the number of school absences:

Emotional Impact

Relationships can bring emotional highs and lows. The stress or distraction from relationship emotions may lead to students taking time off from school to deal with stress and other issues.

Time Commitment

Romantic relationships require time and effort, which might conflict with school schedules. Students in relationships may prioritize spending time with their partners over attending classes, leading to increased absences.

Motivation and Priorities

Being in a relationship may shift a student's priorities away from academic goals to personal and social ones, leading to more absences as school becomes a lower priority.

Social Influence

Peers, including romantic partners, can significantly influence each other's behavior. If a student's partner has a different attitude toward school attendance, it might affect the student's own attendance habits.

Impact

By studying this relationship, we can gain insights into how romantic relationships might affect educational outcomes and identify potential areas for intervention to improve attendance rates.

Research Question

Does being in a romantic relationship affect the number of absences in math classes among students?

The relationship we are investigating can be expressed as:

$$\text{Number of Absences} = \text{Romantic Relationship Status} + \text{Other Factors}$$

Where:

- Number of Absences is the dependent variable
- Romantic Relationship Status is the key independent variable of interest, indicating whether a student is in a romantic relationship (Yes or No)
- Other Factors include demographic, familial, social, health, and academic variables that might also influence the number of absences

Next, we will proceed with data preparation and analysis to test this relationship.

Data Description

Frequencies of Categorical Demographic Variables

- School: GP 349 MS 46
- Sex: Female 208 Male 187
- Address: Urban 307 Rural 88
- Family Size: 4 or more 281 3 or less 114
- Parental Status: Together 354 Apart 41

Five-Number Summaries of Quantitative Demographic Variable

Variable	Min	Q1	Median	Q3	Max
Age	15	16	17	18	22

Descriptive Statistics of Outcome and Explanatory Variable

Variable	Min	Q1	Median	Q3	Max
Absences	0	0	4	8	75

- Romantic Relationship Status: No 263 Yes 132

The histogram of the Absences Variable (Chart 1) shows:

- Shape: The distribution is right-skewed, with a large number of students having few or no absences and a few students having a high number of absences
- Center: The median number of absences is 4
- Skewness: The distribution is positively skewed
- Weirdness: There are some outliers with very high numbers of absences

The visualization of absences by romantic relationship status (Chart 2) shows:

- Distribution: Students in a romantic relationship tend to have a higher median number of absences compared to those who are not in a relationship.
- Spread: There is a wider range of absences among students in a relationship, with more outliers at the higher end.

This initial visualization supports the hypothesis that being in a romantic relationship may be associated with a higher number of math absences.

Next, we will proceed with the statistical analysis to formally test the research question.

Answering the Question

Empty model

Which does not include any predictors, can be expressed as:

$$\underline{\text{Absences} = \beta_0 + \varepsilon}$$

Where:

- β_0 is the intercept, the mean number of absences
- ε is the error, variation in number of absences not explained by the model

Parameter Estimates:

- Intercept $\beta_0 = 5.709$

Interpretation of the Estimated Parameter:

- The intercept of 5.709 indicates the mean number of absences among students when no predictors are included in the model. This value is the average number of absences across all students in the sample.
- The empty model provides a baseline for understanding the average level of absences. Next, we will fit a model including the romantic relationship status to see if this variable helps explain the variation in student absences.

Explanatory Model

Coding of X_i - The explanatory variable, romantic relationship status is coded as:

- No: Students not in a romantic relationship (baseline category)
- Yes: Students in a romantic relationship

The explanatory model can be represented as:

$$\underline{\text{Absences} = \beta_0 + \beta_1(\text{Romantic Relationship Status}) + \varepsilon}$$

Where:

- β_0 is the intercept, the mean number of absences for students not in a romantic relationship.
- β_1 is the difference in the mean number of absences between students in a romantic relationship and those not in a romantic relationship.
- ε is the error, variation in number of absences not explained by the model

Parameter Estimates:

- Intercept $\beta_0 = 4.840$, the estimated mean number of absences for students who are not in a romantic relationship.
- Romantic Relationship Status β_1: 2.599, this parameter represents the additional number of absences for students who are in a romantic relationship compared to those who are not. On average, students in a romantic relationship have 2.60 more absences than those not in a relationship.
- This difference is NOT statistically significant (PRE = 0.0235 = 2.35%)
  (Less than 5% is insignificant)

The model suggests that being in a romantic relationship is associated with a small increase in the number of school absences. On average, students in a romantic relationship tend to have about 2.60 more absences compared to those who are not in a relationship.

Comparing the models

Table 1 is an Analysis of Variance on the number of absences data broken apart by relationship status of the student. The data from running the supernova function shows:

- F-ratio = 9.47, which is statistically significant (p-value = 0.002). This indicates that the explanatory model with the romantic relationship status variable explains some variation in the number of absences compared to the empty model
- Degrees of Freedom (df): The model has 1 degree of freedom for the predictor (romantic relationship status) and 393 degrees of freedom for the residuals
- Difference in Number of Parameters: The empty model has 1 parameter (the intercept), while the explanatory model has 2 parameters (intercept and the romantic relationship status)
- Proportion Reduction in Error (PRE) = .0235, indicates that considering the romantic relationship status variable in the model reduces the error in predicting absences by approximately 2.35%. Suggesting that while the romantic variable does explain some variance in absences, its effect is relatively modest.
- Cohen's d = 0.325, calculated using the means and standard deviations from the two groups (students in a romantic relationship vs. not).

$$d = \frac{(\beta_1 - \beta_0)}{SD} \quad \begin{matrix}\text{(mean yes – mean no)} \\ \text{(sd of whole romantic dataset)}\end{matrix}$$

- Mean absences for yes 7.439 – mean no 4.840 / standard dev 8.003 = 0.325
  0.325 suggests a small to medium size effect.

This indicates that the difference in absences between students in a romantic relationship and those not in a relationship is noticeable but not too large.

Comparison Summarized

The F-ratio of 9.47 and the significant p-value indicate that the model with the romantic relationship status variable significantly improves the prediction of absences.

The PRE value of 2.35% suggests a modest reduction in error.

The Cohen's d value of 0.325 indicates a small to medium effect size.

These findings suggest that while romantic relationship status does have a statistically significant impact on student absences, the effect size is relatively small.

Conclusion

From the analysis, we learned that being in a romantic relationship is associated with a higher number of school absences among students. Specifically:

- Model Results: The explanatory model, which includes the romantic relationship variable, shows that students in a romantic relationship have, on average, 2.60 more absences than those who are not in a relationship; moderately significant.
- Effect Size: The effect size, represented by Cohen's d (approximately 0.325), indicates a small to medium effect, suggesting that the difference, while significant, is not very large in practical terms.
- PRE: The value of 2.35% suggests that the romantic relationship status accounts for a small portion of the variability in student absences.

These findings support the research question: There is an association between romantic relationship status and the number of absences. However, the effect size indicates that while significant, it explains only a small part of the absence variance.

Limitations of the Analysis

Several limitations should be considered when interpreting these findings:

1. Sample Size: The sample is 395 students, which is reasonably large, but the generalizability of the findings may be limited if not representative of the student body.

2. Measurement Error: The accuracy of the absences data and the self-reported status of romantic relationships could introduce measurement error. Students might underreport absences or inaccurately report their relationship status.

3. Sampling Bias: If the sample is not randomly selected or if it does not adequately represent different student demographics, the results may be biased. Factors like school type, geographic location, and socio-economic status could influence findings.

4. Unmeasured Confounders: Other variables not included in the model (e.g., mental health, peer influence) might also affect absences and confound the relationship between romantic relationship status and absences.

5. Timing Design: The data captures a single point in time. This limits the ability to make causal inferences about the relationship between romantic relationships and absences.

6. Potential Missing Data: If there is missing data that was not accounted for, this could bias the results. Ensuring data completeness and addressing missing values is crucial for analysis.

7. Contextual Factors: The cultural, social, and educational context in which the data was collected might influence the results. These factors should be considered when generalizing the findings to other settings.

Summary

The analysis indicates that being in a romantic relationship is significantly associated with a higher number of school absences among students. However, the effect size is small, and several limitations must be considered. Future research could address these limitations by using larger, more representative samples, longitudinal designs, and incorporating additional relevant variables to provide a more comprehensive understanding of the factors influencing student absences.

Tables

Table 1

Analysis of Variance Table (Type III SS)
Model: absences ~ romantic

Data analyzed	SS	df	MS	F	PRE	p
Model (error reduced)	593.711	1	593.711	9.469	.0235	.0022
Error (from model)	24641.808	393	62.702			
Total (empty model)	25235.519	394	64.050			

Note: Discussed on page 9

Chart 1

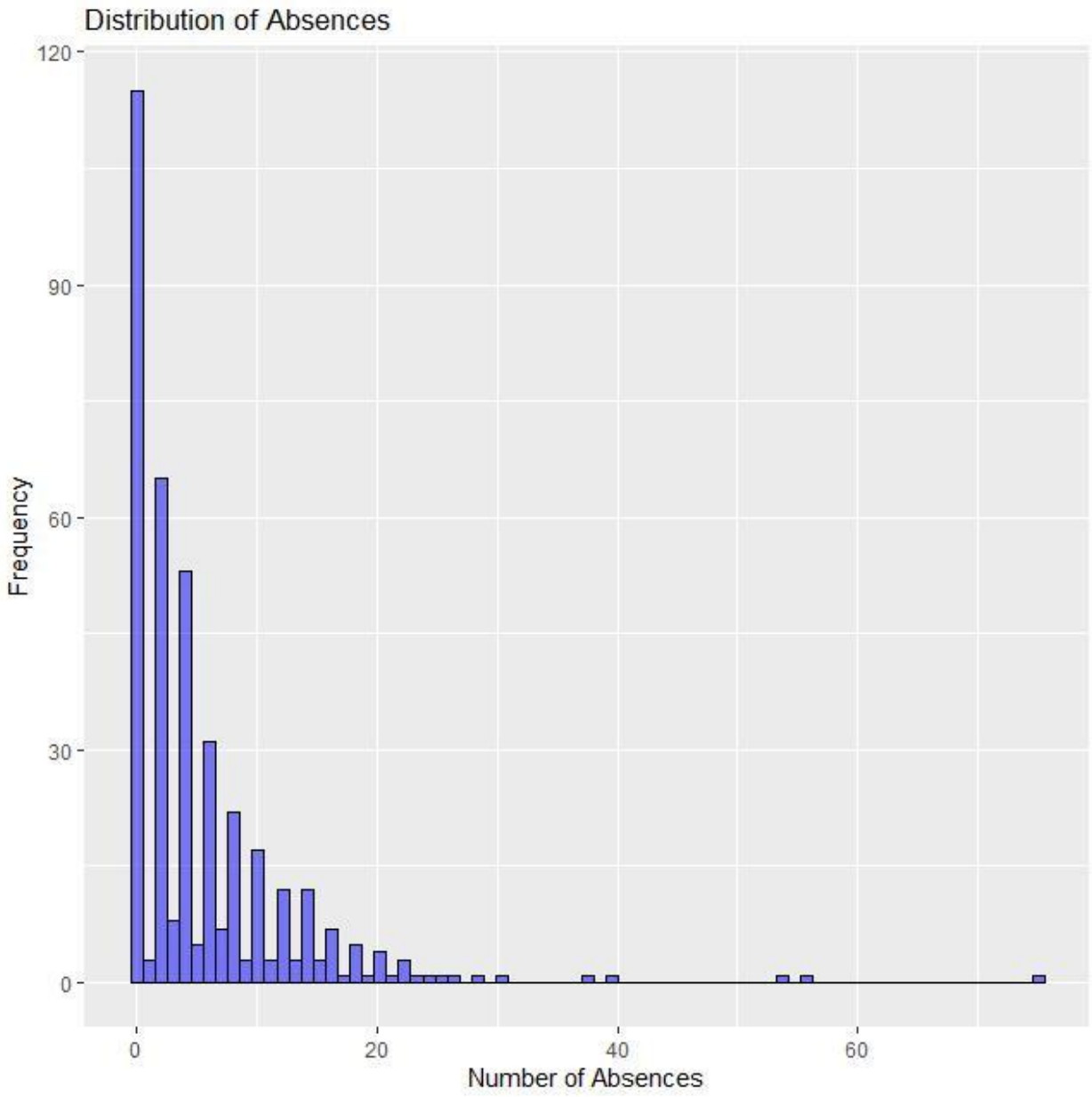
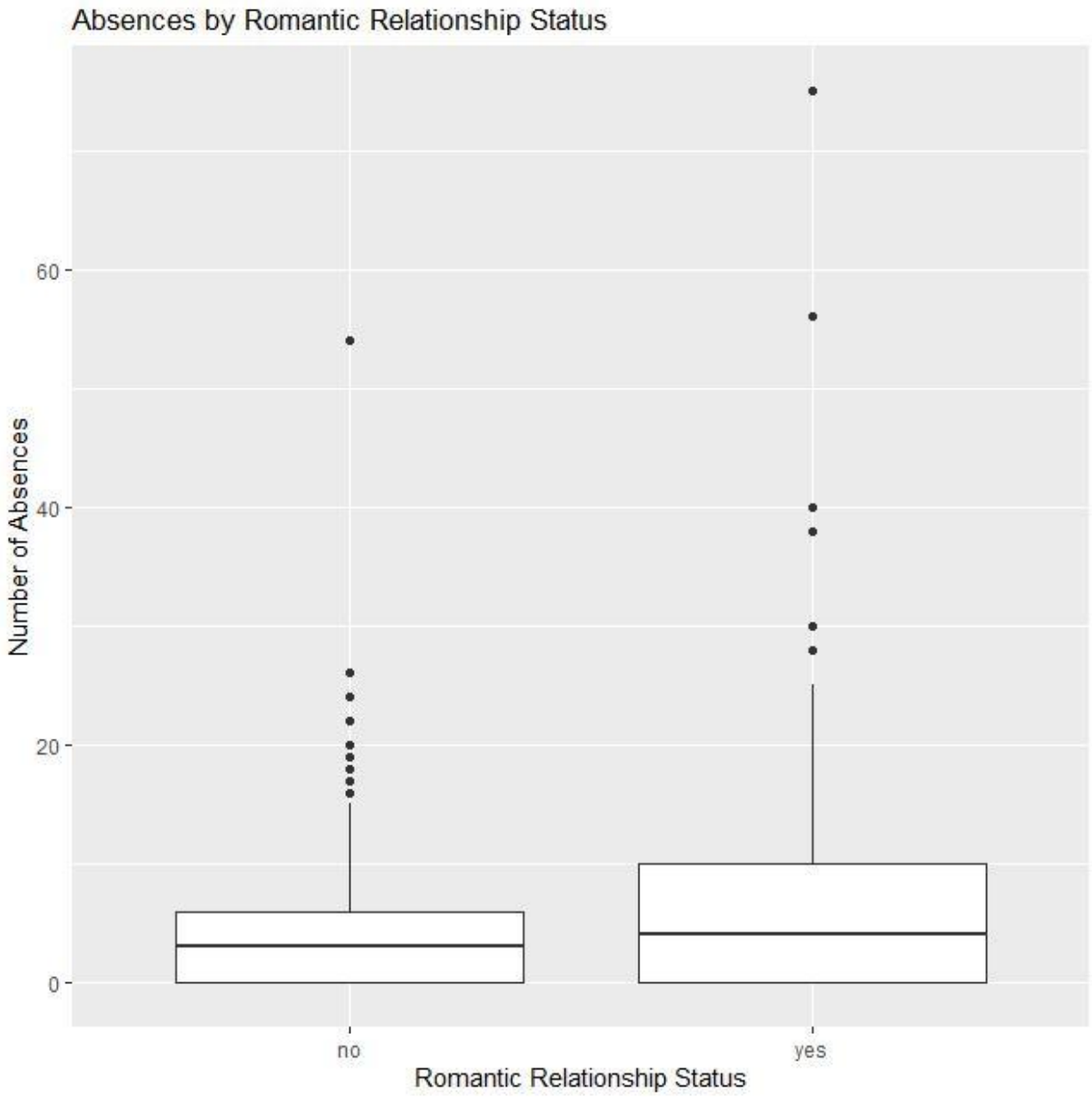

Distribution of Absences
Frequency
120
90
60
30
0
0
20
40
60
Number of Absences

Chart 2


Absences by Romantic Relationship Status
Number of Absences
60
40
20
0
no
yes
Romantic Relationship Status