

# STA304 Report

## AUTHOR

Chau Nguyen, John Luc, Adam Tadj, Sunwoo Hwang,  
Sabrina Seiler

## PUBLISHED

December 2, 2024

## 1. Abstract

University students frequently face challenges that contribute to mental health issues such as stress and anxiety. This study examines the relationship between mental health and factors such as academic workload, hours of sleep, living situation, and work-life balance. Data were collected through a Google Form survey administered to students in a third-year statistics survey & sampling course, where participants rated their agreement with statements related to various mental health indicators. The results showed a significant relationship between academic workload and stress levels. While no significant relationship was found between stress levels and hours of sleep, we observed a significant association between hours of sleep and missed social events. Additionally, no statistically significant relationship was identified between stress levels and living situation, though a marginally significant relationship was found between living situation and missed social events. Since we were not able to reach most of the conclusions that we expected, we plan to consider other factors that might influence mental health, and also incorporate attention-check questions to identify rushed responses.

## 2. Introduction

The academic workload faced by university students has increasingly become one of the critical variables for university students concerning mental health. Given a broader context of increasing academic challenges, many students experience heightened stress levels, which impact various aspects of their well-being, including sleep quality, work-life balance, and several mental health issues.

First, commonly reported mental health issues, including anxiety, depression and stress, have been found to correlate with academic pressures in highly competitive or rigorous programs where students may feel unable to meet academic and personal expectations (Beiter et al., 2014). Second, research indicates that students experiencing high academic workload tend to report poorer sleep quality and shorter sleep duration, which, in turn, can exacerbate mental health problems and academic performance challenges (Hershner & Chervin, 2014). Third, students' living environments and their ability to maintain a balanced lifestyle can also significantly impact mental health, since living in a more supportive and conducive environment positions students to better manage academic responsibilities and personal well-being (Hysing et al., 2016).

Given the above potential factors, we seek to address the following research questions in this report:

- **Research Question 1:** Is there an association between students' academic workload and the amount of mental health issues?

- **Null Hypothesis:** There is no association between academic workload and the amount of mental health issues.
- **Alternative Hypothesis:** There is a significant association between academic workload and the amount of mental health issues.
- **Research Question 2:** How does the academic workload affect the quality/length of sleep and its subsequent impact on students' mental health?
  - **Null Hypothesis:** Academic workload does not affect the quality or length of sleep, nor does it affect students' mental health.
  - **Alternative Hypothesis:** Academic workload significantly affects the quality and/or amount of sleep students get, which affects mental health.
- **Research Question 3:** Is mental health affected by students' living situations or students' ability to maintain work-life balance?
  - **Null Hypothesis:** Students' mental health is not affected by their living situation or their ability to maintain a work-life balance.
  - **Alternative Hypothesis:** Students' mental health is significantly affected by their living situation and their ability to maintain a work-life balance.

This study focuses on students enrolled in STA304H5 during Fall 2024, aiming to provide valuable insights into the effects of academic workload on their overall well-being. By exploring these relationships, the findings may inform strategies to enhance student support services and improve academic experiences.

## 3. Methodology

### 3.1. Data Collection

Between October 8th, 2024, and October 20th, 2024, the questionnaire's Google Form link was posted on the STA304H5 Piazza course discussion board. We utilized Simple Random Sampling (SRS) by using the R random generator to randomly sample 45 students from the received questionnaire responses. This method was chosen because it gives all respondents an equal chance of being included in the sample, making the sample representative. This experiment does not require methods like stratified or cluster sampling because the population in this experiment is relatively uniform and doesn't need to be divided into strata or clusters.

We considered alternative methods, such as stratified sampling or cluster sampling. However, these methods are more suitable when the characteristics of specific subpopulations need to be accounted for. Since the characteristics of students taking STA304 were expected to be similar, we decided that it would not be meaningful to distinguish between the subpopulations. Since our goal was to capture the characteristics of all students, we determined that SRS was the most appropriate method for this experiment.

On the questionnaire, we asked participants to complete 10 short-answer questions inquiring about academic workload, average hours studied weekly, amounts of experienced stress-related symptoms (stress, anxiety, concentration difficulty), hours of sleep per night, current living situation, and experienced common issues caused by academic workload (time management, social life, financial).

## 3.2. Target Population

The target population consists of students enrolled in the STA304H5 course during the Fall 2024 semester.

## 3.3. Sampling Method

We will use **Simple Random Sampling (SRS)** by combining both sections of the STA304H5 course and selecting a random sample of students from the pool.

## 3.4. Variables of interest

The following variables were identified for the research:

### 3.4.1. Academic Workload

Academic workload is defined as the number of assignments, projects, and exams and the overall demands of a course. The purpose of this study is to see if a high academic workload is associated with increased mental health problems, such as stress. Academic workload was measured through a questionnaire on a scale of 1 (very light), 2 (light), 3 (moderate), 4 (heavy) and 5 (very heavy).

### 3.4.2. Living Situation

We divided students' living arrangements into four categories: Living alone, living with family, living with roommates, and living in on-campus housing. These living arrangements are important variables that have a significant impact on students' mental health and academic stress.

### 3.4.3. Study Hours

Study time is an important indicator of students' time management and academic performance. This study analyzes the relationship between study time and mental health, exploring whether over studying leads to stress reduction or instead causes mental stress and anxiety. Study hours were measured by entering the number of hours of study per week.

### 3.4.4. Mental Health Symptoms (Stress, Anxiety, and Concentration Difficulties)

Mental health symptoms are key indicators of students' mental health and may be influenced by academic workload and living conditions. Understanding these relationships helps target interventions and support

strategies to improve students' mental well-being. These symptoms were measured through a questionnaire on a scale of 1 (Never) 2 (Sometimes) and 3 (Always).

### 3.4.5. Work-life balance Factors (Financials, Time Management)

Two key indicators of how well a student is adjusting to an independent life in university is their financial status and ability to manage time properly. Therefore, we believe that investing these factors would help to shed light on whether students' work-life balance are being impacted by academic workload. These are measured through a questionnaire of how often students feel these factors are of concern as a result of their workload, on a scale of 1-5 (Never, Rarely, Sometimes, Often, Very Often).

## 3.5. Sample Size

In our research, we applied Simple Random Sampling (SRS) on 200 students. To determine the sample size using the bound of error as 0.29, we used the following formula to find the magnitude of the confidence interval:

$$D = \frac{B^2}{4}$$

Where  $B = 0.29$ , giving:

$$D = \frac{(0.29)^2}{4} = 0.021025$$

Using the standard deviation ( $\sigma$ ) of the academic workload variable, we computed the required sample size using the formula:

$$n = \frac{N \times \sigma^2}{(N - 1) \times D + \sigma^2}$$

Where:

$N = 200$  (total population size)  $B = 0.29$  (bound of error)  $D = B^2/4 = 0.021025$   $\sigma$  is approximated as (range of academic workload)/4 range = maximum academic workload - minimum academic workload

To randomize our sampling process, we used R's random sampling function `sample()` with a fixed seed (`set.seed(1)`) to randomly select participants from the full dataset based on the calculated sample size  $n$ .

Thus, below is the initial setup for all of our statistical tests.

## 4. Analysis

### Section 1. Kruskal-Wallis Testing

#### Research Question 1: Academic Workload and Stress Levels

RQ1 questions whether academic workload contributes to a student’s stress levels. To answer this question, we attempted using ANOVA (Analysis of Variance) to test the difference in average stress levels between students with different academic workloads. However, not all our assumptions were satisfied.

Shapiro-Wilk Test for Normality was used to test normality for all our groups; however, for all groups, we received a *p-value* below 0.05. It is also impossible for our groups to have >30 observations, and thus we failed the normality assumption for ANOVA. The only assumption we could meet was that the samples between the groups were randomly sampled, and responses were independent from each other. This allowed us to test using the non-parametric version of ANOVA, called the Kruskal-Wallis test, where normality and equal variances aren’t assumed.

We obtained:

- Kruskal-Wallis Chi-squared = 6.4099,
- *df* = 3,
- *p-value* = 0.0932.

Since the calculated *p-value* of 0.0932 is larger than the 0.05 significance level, we fail to reject the null hypothesis. This indicates that the difference between the average stress levels across different academic workloads is not statistically significant enough to suggest that different academic workloads have a meaningful impact on student mental health.

Note that if we look at the mean stress levels for each rating of academic workloads:

Workload Rating	Mean Stress Level
1	1.0000
2	2.2857
3	2.6154
4	2.7000

While the mean stress level appears to be higher for students who report heavier academic workloads, the mean difference are not significant enough to suggest that higher workload indicates higher stress levels.

## Research Question 2: Sleep Quality and Stress Levels

Regarding sleep quality against average stress levels, we first ran a Shapiro-Wilk test to see if the data followed a normal distribution. When we ran the Shapiro-Wilk test, we got a *p-value* of 0.02971, which is less than 0.05, showing that hours of sleep don’t follow a normal distribution. This violates the assumptions of ANOVA analysis of hours of sleep, so we decided to use the Kruskal-Wallis test.

We obtained:

- *p-value* = 0.1115,

Since the  $p$ -value is greater than 0.05, we can't conclude that there is a significant difference between hours of sleep and stress levels. Therefore, we can't reject the null hypothesis that sleep duration doesn't affect stress levels.

We also examined the mean stress level across different sleep categories:

Sleep Category	Mean Stress Level
Less than 6 hours	2.00
6-8 hours	2.67
More than 8 hours	2.22

While the Kruskal-Wallis test shows no significant difference between stress and sleep duration, averaging the stress levels across hours of sleep shows that the group with 6-8 hours of sleep has a higher stress level than the other two groups. However, the differences are not dramatically different, so we can't draw conclusions based on simple averages. It's possible that we didn't find statistical significance because of the small sample size.

### Research Question 3: Social Events and Stress Levels

For RQ3, we examined the relationship between average stress levels and the frequency of missed social events of the sampled students. This relationship would imply a possible effect on students' mental health due to varying work-life balances.

Similar to the other two tests, we ran the Shapiro-Wilk test of normality, obtaining a  $p$ -value of 0.01873, again failing the normality assumption for ANOVA testing. Therefore, we proceeded with the Kruskal-Wallis test.

We obtained:

- Kruskal-Wallis Chi-squared = 2.6958,
- $df = 4$ ,
- $p$ -value = 0.0932.

Since the calculated  $p$ -value of 0.0932 is larger than the 0.05 significance level, we fail to reject the null hypothesis. This shows that the difference between the average stress levels across different frequencies of missing social events is not significant enough to suggest that work-life balance has a meaningful impact on student mental health.

Note that if we look at the mean stress levels for each rating of work-life balance:

Social Events Missed	Mean Stress Level
1	2.00
2	2.60

Social Events Missed	Mean Stress Level
3	2.70
4	2.44
5	2.50

While the mean stress level appears to be higher for some categories of students who report missing social events, the mean difference are not significant enough to suggest that missing social events indicates higher stress levels.

## 4.2. Chi-square Test of Independence

Research Question 1 questions if there is a relationship between the academic workload of a student and their mental health. Research question 3 questions if there is a relationship between students’ living situation or ability to maintain work-life balance. Our assumption that the data is randomly sampled holds given our method of sampling outlined above.

To investigate whether there exists a relationship between these variables, we use the Chi-squared test of independence with  $\alpha = 0.05$  as the significance level among a permutation of the variables of interests (academic workload and mental health symptoms) and (living situations, work-life balance). In total, 12 Chi-squared tests were performed, with the results and conclusion as follows:

### Chi-squared Test between Academic Workload and Stress

$\chi^2 = 25.936, df = 6, p\text{-value} = 0.0002288$

Since the calculated  $p\text{-value}$  was 0.0002288 and is less than  $\alpha = 0.05$ , we reject the null hypothesis. There is enough statistical eviende that there is a significant association between academic workload and the stress levels of a student.

### Chi-squared Test between Academic Workload and Anxiety

$\chi^2 = 14.113, df=6, p\text{-value} =0.0284$

Since the calculated  $p\text{-value}$  was 0.0284 and is less than  $\alpha = 0.05$ , we reject the null hypothesis. There is enough statistical evidence to indicate that a significant association between the academic workload and the anxiety levels of a student exists.

### Chi-squared Test between Academic Workload and Concentration

$\chi^2 = 6.5, df=6, p\text{-value} =0.3696$

Since the calculated  $p\text{-value}$  was 0.3696 and is greater than  $\alpha = 0.05$ , we fail to reject the null hypothesis. There isn’t enough statistical evidence to indicate that the academic workload and the concentration levels of a student are related.

**Chi-squared Test between Living Situation and Stress**

$$\chi^2 = 4.9936, df=6, p\text{-value} = 0.5446$$

Since the calculated  $p\text{-value}$  was 0.5446 and is greater than  $\alpha = 0.05$ , we fail to reject the null hypothesis. There isn't enough statistical evidence that the living situation and the stress levels of a student are related.

**Chi-squared Test between Living Situation and Anxiety**

$$\chi^2 = 6.6042, df=6, p\text{-value} = 0.359$$

Since the calculated  $p\text{-value}$  was 0.359 and is greater than  $\alpha = 0.05$ , we fail to reject the null hypothesis. There isn't enough statistical evidence that the living situation and the anxiety levels of a student are related.

**Chi-squared Test between Living Situation and Concentration**

$$\chi^2 = 2.8773, df=6, p\text{-value} = 0.8241$$

Since the calculated  $p\text{-value}$  was 0.8241 and is greater than  $\alpha = 0.05$ , we fail to reject the null hypothesis. There isn't enough statistical evidence that the living situation and the concentration levels of a student are related.

**Chi-squared Test between Time Management and Stress**

$$\chi^2 = 10.894, df=8, p\text{-value} = 0.2078$$

Since the calculated  $p\text{-value}$  was 0.2078 and is greater than  $\alpha = 0.05$ , we fail to reject the null hypothesis. There isn't enough statistical evidence that the time management skills and the stress levels of a student are related.

**Chi-squared Test between Time Management and Anxiety**

$$\chi^2 = 6.5567, df=8, p\text{-value} = 0.5851$$

Since the calculated  $p\text{-value}$  was 0.5851 and is greater than  $\alpha = 0.05$ , we fail to reject the null hypothesis. There isn't enough statistical evidence that the time management skills and the anxiety levels of a student are related.

**Chi-squared Test between Time Management and Concentration**

$$\chi^2 = 9.3321, df=8, p\text{-value} = 0.3151$$

Since the calculated  $p\text{-value}$  was 0.3151 and is greater than  $\alpha = 0.05$ , we fail to reject the null hypothesis. There isn't enough statistical evidence that the time management skills and the concentration levels of a student are related.

**Chi-squared Test between Financials and Stress**

$$\chi^2 = 9.3321, df=8, p\text{-value} = 0.7223$$



Since the calculated  $p$ -value was 0.3151 and is greater than  $\alpha = 0.05$ , we fail to reject the null hypothesis. There isn't enough statistical evidence that the financial situation and the stress levels of a student are related.

### Chi-squared Test between Financials and Anxiety

$$\chi^2 = 5.3134, df=8, p\text{-value} = 0.7236$$

Since the calculated  $p$ -value was 0.7236 and is greater than  $\alpha = 0.05$ , we fail to reject the null hypothesis. There isn't enough statistical evidence that the financial situation and the anxiety levels of a student are related.

### Chi-squared Test between Financials and Concentration

$$\chi^2 = 6.1471, df=8, p\text{-value} = 0.6308$$

Since the calculated  $p$ -value was 0.6308 and is greater than  $\alpha = 0.05$ , we fail to reject the null hypothesis. There isn't enough statistical evidence that the financial situation and the concentration levels of a student are related.

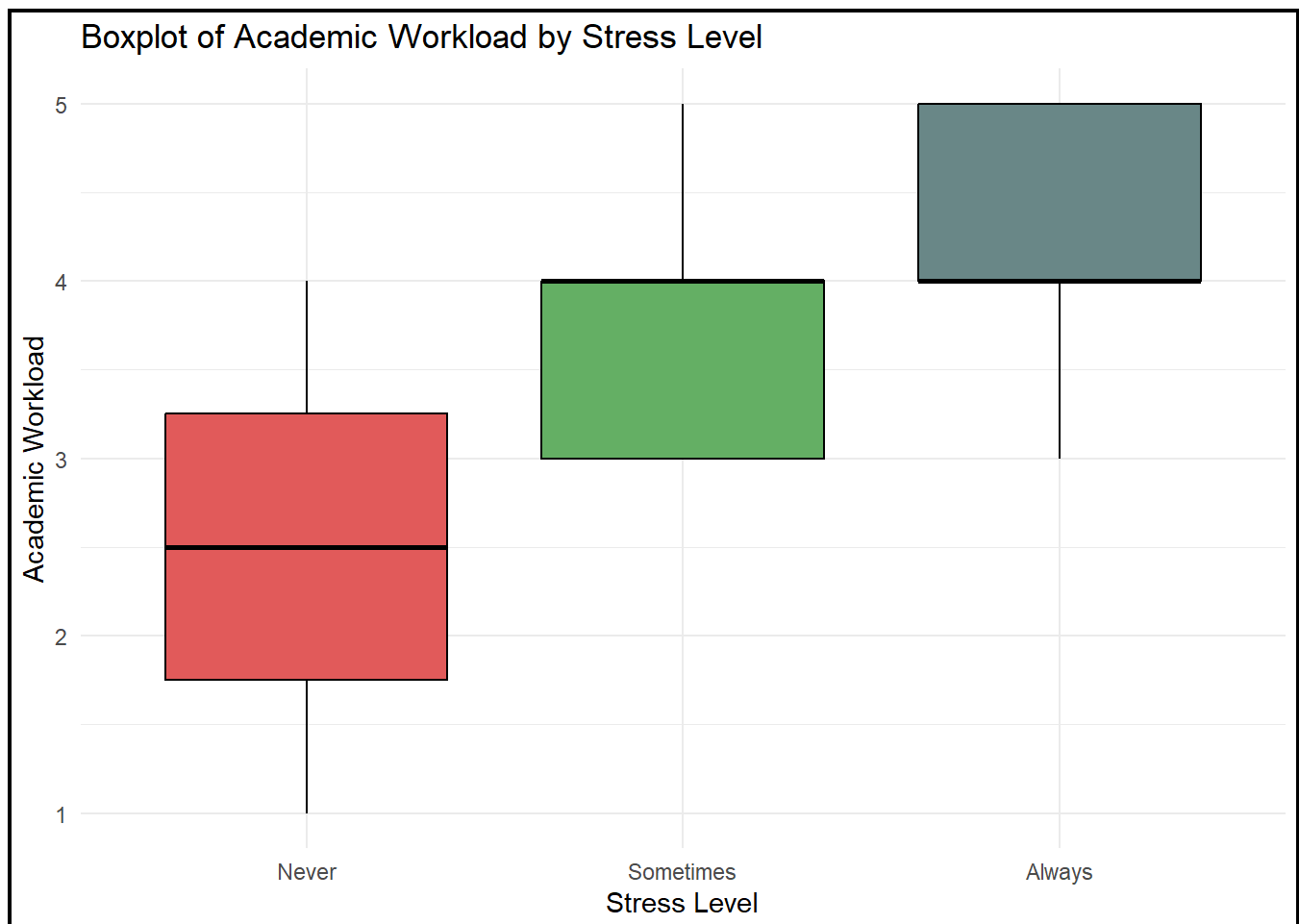
## 4.3. Simple and Multiple Linear Regressions

This part of the analysis focuses on applying linear regression to address the three research questions outlined in the introduction. We will explore whether there is a relationship between students' academic workload and the prevalence of mental health issues; between academic workload, the amount of sleep students get, and its impact on mental health; and finally, between mental health, students' living situations, and their ability to maintain a work-life balance. All of these relationships were found to be statistically significant.

### 4.3.1. Research Question 1

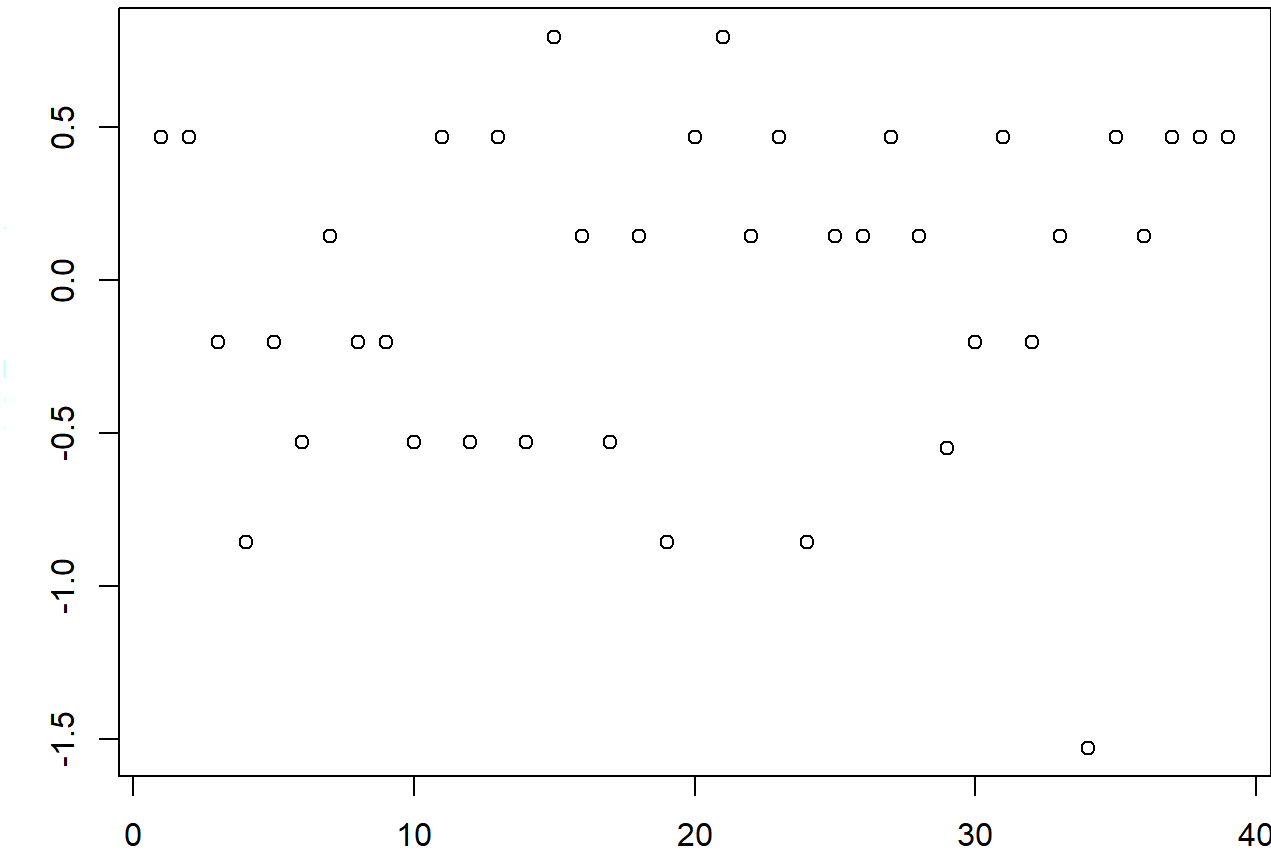
#### Relevant Graphs & Tables

Research Question 1 is concerned with determining whether there is a relationship between students' academic workload and the prevalence of mental health issues in students. To answer this question, we start by plotting a boxplot of academic workload by stress level. This plot suggests that higher stress levels are associated with a higher academic workload. On average, students who "Always" or "Sometimes" feel stressed tend to have a higher workload than those who "Never" feel stressed.



## Necessary Assumptions

To confirm this association, we implement a simple linear regression model, where the dependent variable is `stress_numeric` (the stress variable, converted to numeric values), and the independent variable is `academic_workload`. To implement this model, we check the following assumptions using the following plot of the residuals:



- 1. Independence of errors: Since the plot has no noticeable pattern, the independence assumption is satisfied.
- 2. Normality of errors: Since our sample size is sufficiently large, by the Central Limit Theorem the sample mean (and by extension, the sampling distribution of residuals) approximately follows a Normal distribution.
- 3. Homogeneity of variances amongst errors: Since the residuals plot has a "horizontal band appearance", this suggests that variance of the residuals are the same for all values of the independent variable academic\_workload.

Since all the assumptions are met, we proceed with the regression analysis.

### Computation and Statistical Test Output

The regression analysis produced the following results:

Statistic	Estimate	SE	t-value	p-value
Intercept ( $\beta_0$ )	1.22389	0.40942	2.989	0.00495
academic_workload ( $\beta_1$ )	0.32655	0.09945	3.284	0.00224

The residual standard error is 0.5353 with 37 degrees of freedom. The  $R^2$  value is 0.2256, and the F-statistic is  $F(1, 37) = 10.78$ , with a  $p$ -value of 0.002.

The equation of the regression line is given by:

Stress = 1.22389 + 0.32655 · Academic Workload

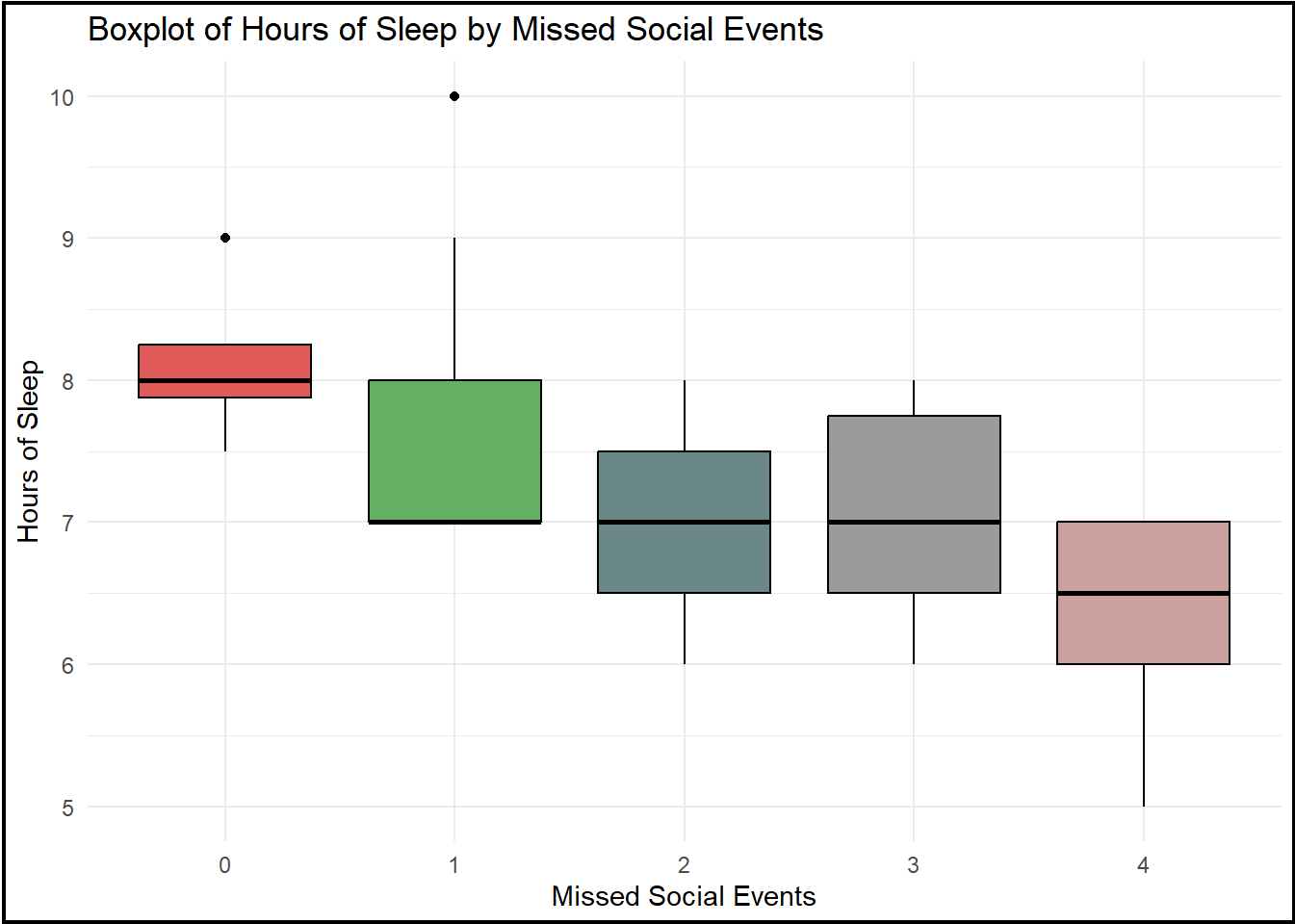
Interpretation

The small  $p - value = 0.00224$  indicates that academic workload is a significant predictor of stress. Therefore, we can conclude that there is a significant relationship between students' academic workload and their stress levels.

4.3.2. Research Question 2

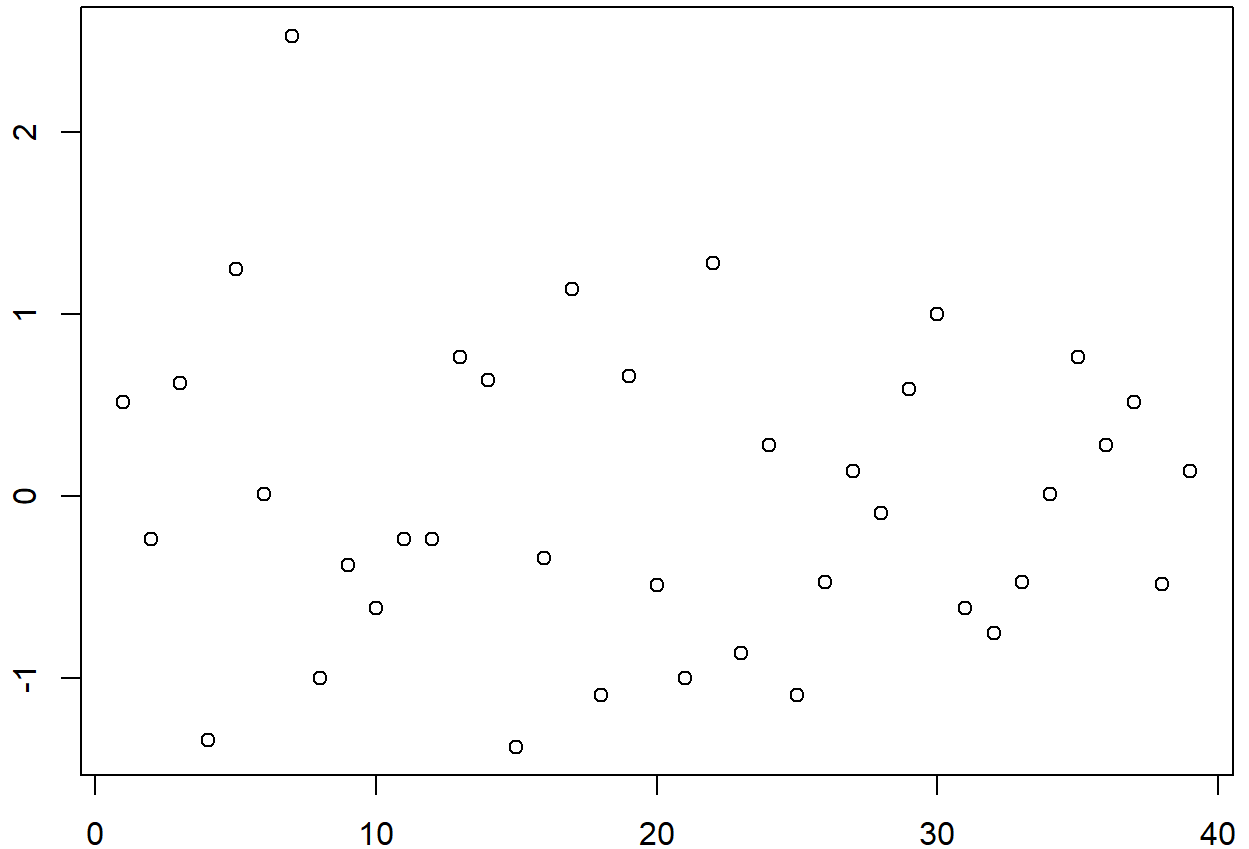
Relevant Graphs & Tables

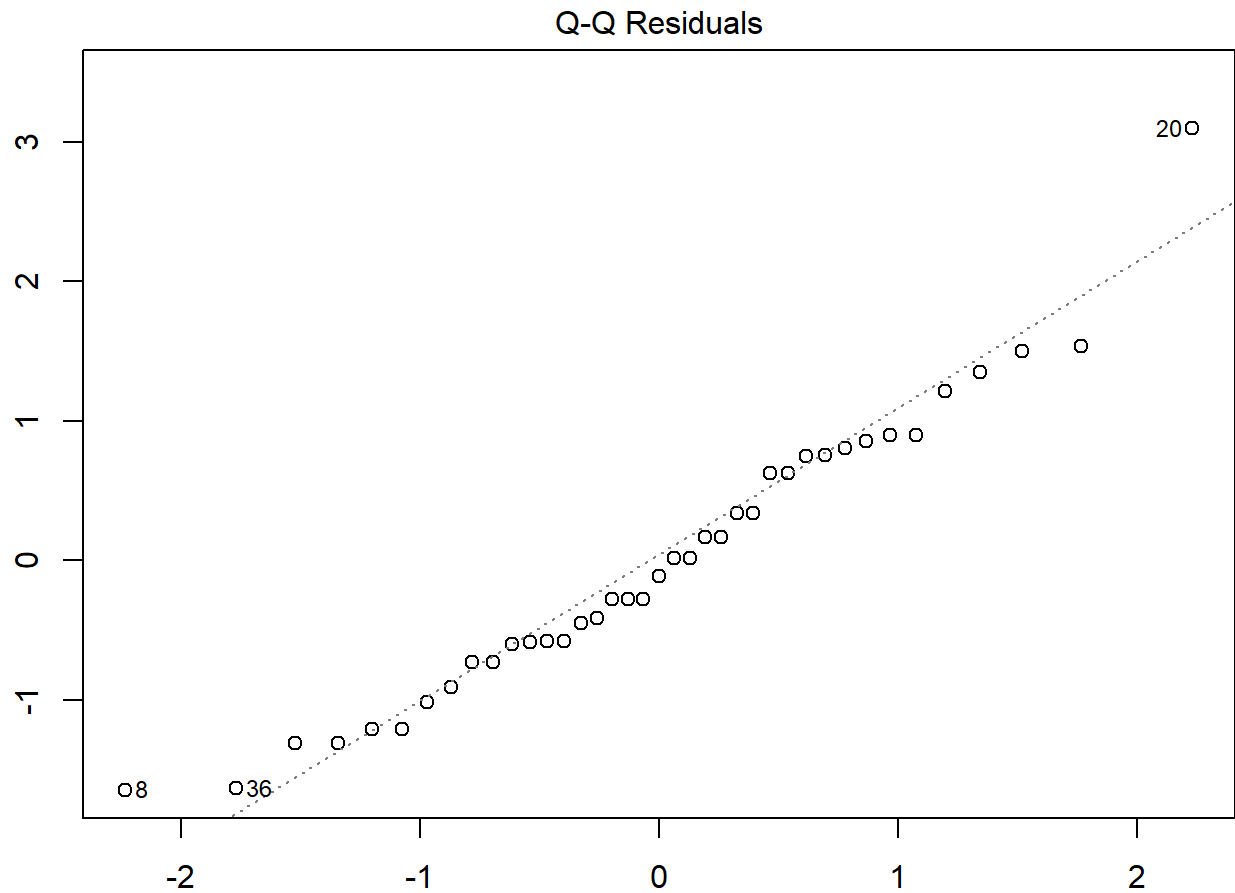
Research Question 2 investigates whether there is a relationship between students' academic workload, the amount of sleep they get, and impact on mental health. To investigate the effect number of missed social events, we plot a boxplot of hours of sleep by the number of missed social events. This plot suggests students who miss more social events tend to have fewer hours of sleep on average.



## Necessary Assumptions

To assess this relationship, we use a multiple linear regression model with hours of sleep as the dependent variable, and academic workload and missed social events as the independent variables. Before interpreting this model, we check the following assumptions:





- Independence of errors: The residual plot shown above indicates no clear patterns, which suggests independence.
- Normality of errors: Since the sample size is large, the Central Limit Theorem implies that the residuals should approximate a normal distribution. Also, a QQ-plot confirms approximate normality.
- Homogeneity of variances: The residuals plot exhibits a roughly horizontal band, indicating constant variance across predicted values.
- No multi-collinearity: We calculated the VIF of each independent variable, and found that  $VIF(\text{academic\_workload}) = 1.076844$ , and  $VIF(\text{missed\_social\_events}) = 1.076844$ . Since they are both less than 5, this indicates there is little correlation between them.

With all assumptions adequately met, we proceed with the regression analysis.

## Computation and Statistical Test Output

The regression analysis produced the following results:

Statistic	Estimate	SE	t-value	p-value
Intercept ( $b_0$ )	8.5521	0.6661	12.838	$5.3 \times 10^{-15}$
academic_workload ( $b_1$ )	-0.1412	0.1666	-0.848	0.40208

Statistic	Estimate	SE	t-value	p-value
missed_social_events ( $b_2$ )	-0.3763	0.1194	-3.151	0.00327

The residual standard error is 0.864 with 36 degrees of freedom. The  $R^2$  value is 0.2653, the adjusted  $R^2$  is 0.2245, and the F-statistic is  $F(2, 36) = 6.5$ , with a  $p$ -value of 0.003889.

The equation of the regression line is given by:

Hours of Sleep = 8.5521 − 0.1412 · Academic Workload − 0.3763 · Missed Social Events

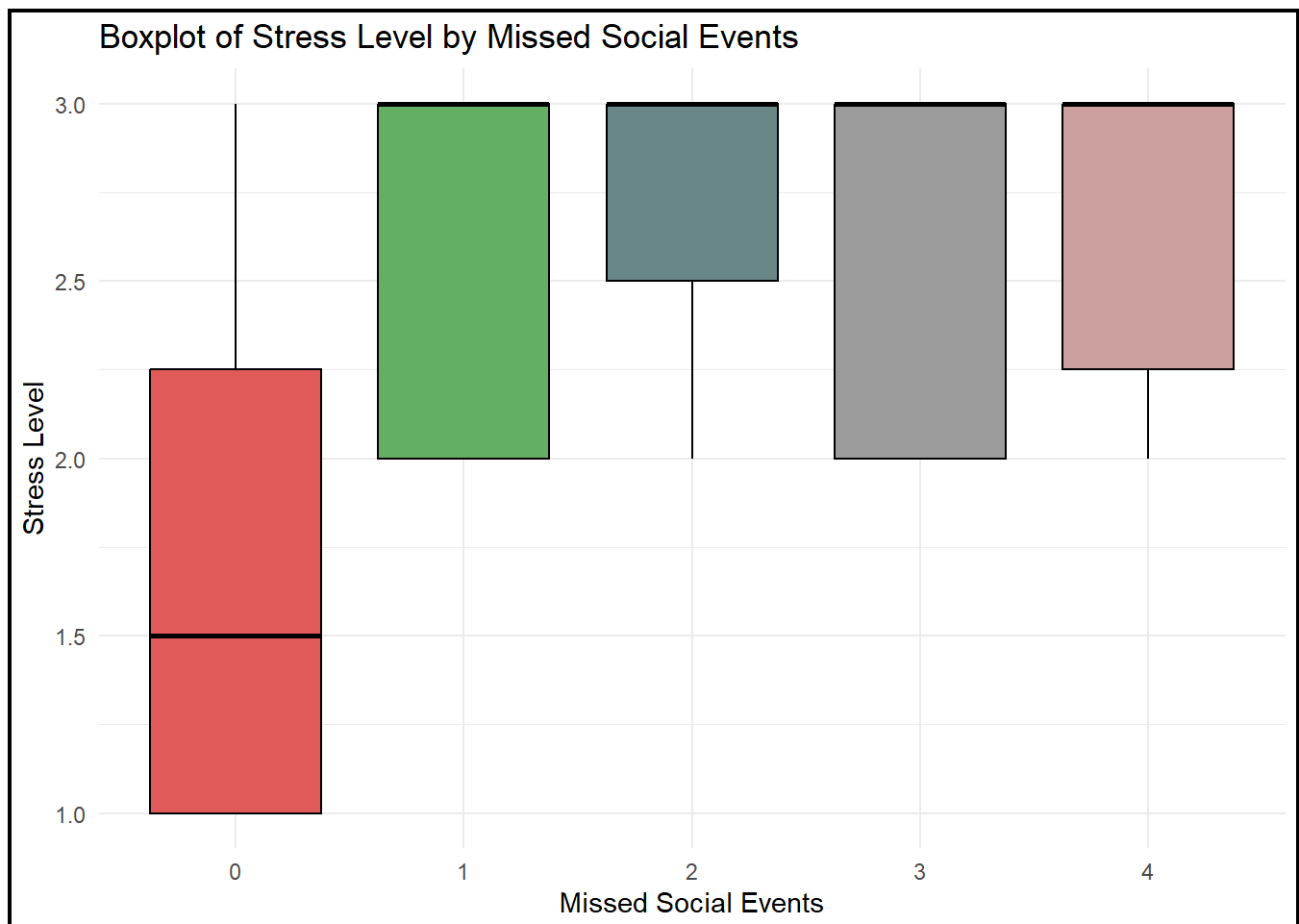
d. Interpretation

The output of the regression analysis shows that the missed social events have a statistically significant negative association with hours of sleep ( $p = 0.00327$ ). However, academic workload does not have a significant relationship with hours of sleep ( $p = 0.40208$ ).

4.3.3. Research Question 3

Relevant Graphs & Tables

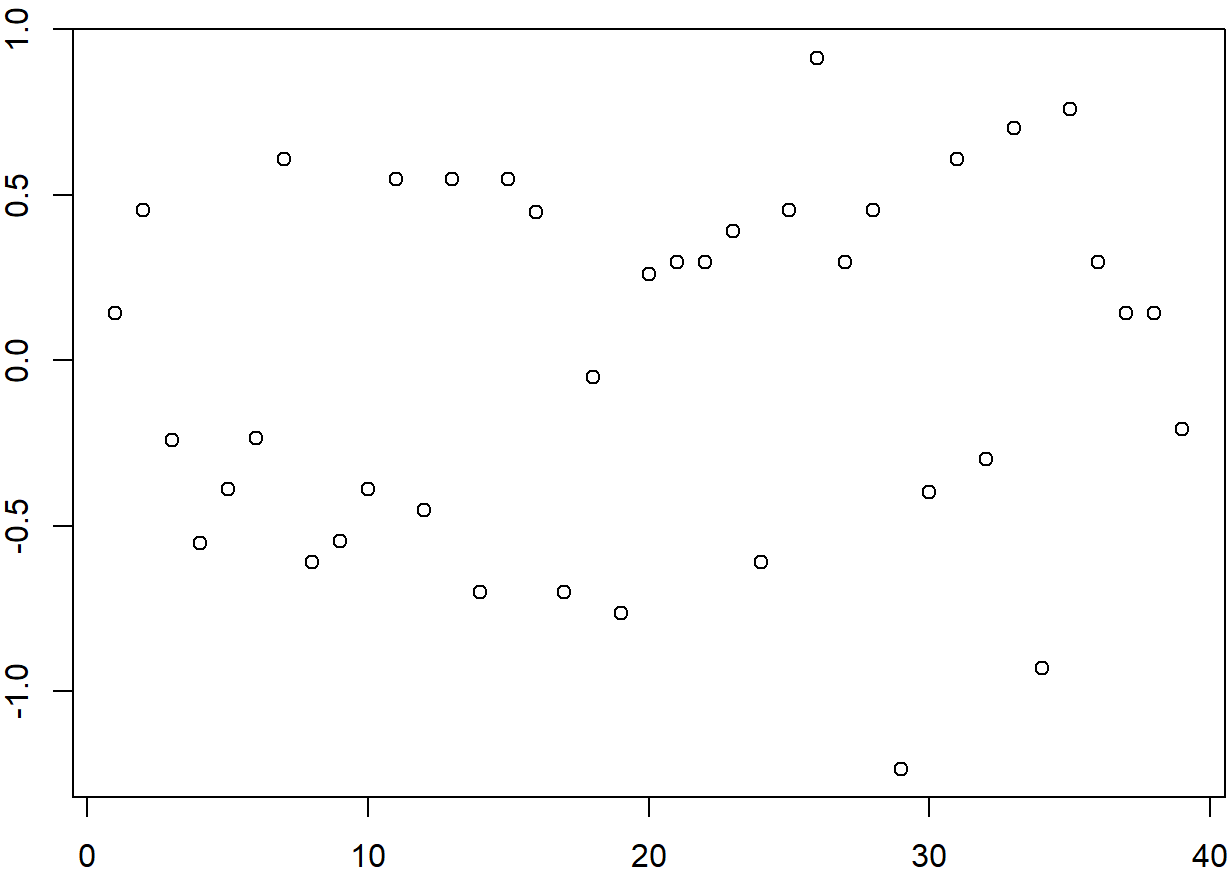
Research Question 3 examines whether mental health, represented by stress levels, is affected by students' living situations and their ability to maintain a work-life balance. To explore this relationship, we first plot a boxplot of stress level by missed social events, with students grouped by stress categories ("Always," "Never," and "Sometimes"). This plot suggests that students who miss social events "Always" or "Sometimes" have higher stress levels, while those who "Never" miss social events report lower stress levels.

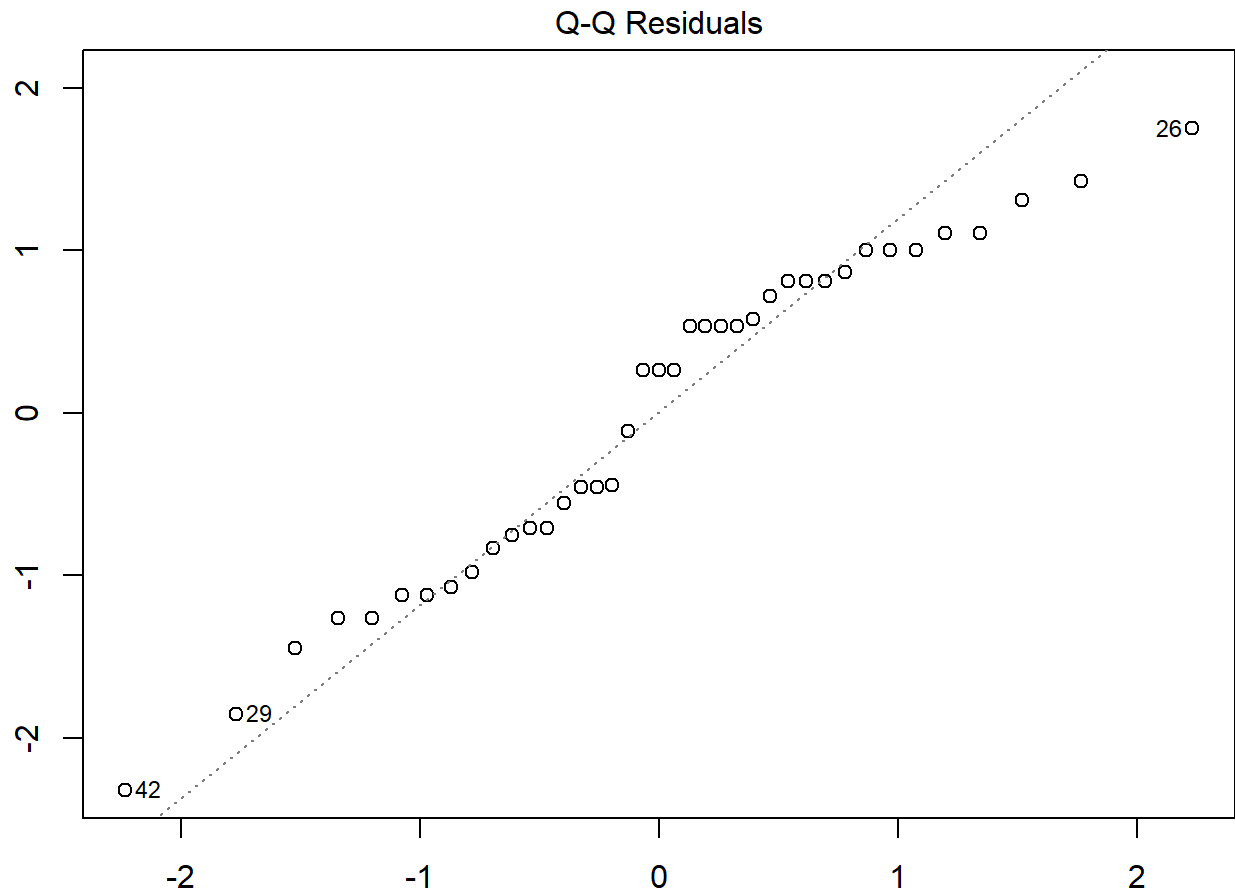


## Necessary Assumptions

To investigate the impact of missed social events and living situations on stress, we use a multiple linear regression model. The response variable is stress (in numeric form), and the predictors are missed social events and living situation. We verify the following assumptions:







- Independence of errors: The residual plot indicates no clear patterns, suggesting that the errors are independent.
- Normality of errors: Since the sample size is large, the Central Limit Theorem implies that the residuals should approximate a normal distribution. Also, a QQ-plot confirms approximate normality.
- Homogeneity of variances: The residuals plot exhibits a roughly horizontal band, indicating constant variance across predicted values.
- No multi-collinearity: We calculated the VIF of each independent variable, and found that  $VIF(\text{missed\_social\_events}) = 1.009045$ , and  $VIF(\text{living\_situation}) = 1.003006$ . Since both VIF values are close to 1, there's no indication of multicollinearity in this model.

With all assumptions adequately met, we proceed with the regression analysis.

## Computation and Statistical Test Output

The regression analysis produced the following results:

Statistic	Estimate	SE	t-value	p-value
Intercept ( $b_0$ )	1.93006	0.27980	6.898	$6.02 \times 10^{-8}$
missed_social_events ( $b_1$ )	0.15560	0.07721	2.015	0.0518

Statistic	Estimate	SE	t-value	p-value
living_on_campus ( $b_2$ )	0.81061	0.39928	2.030	0.0502
living_with_family ( $b_3$ )	0.21206	0.28310	0.749	0.4590
living_with_roommates ( $b_4$ )	0.30493	0.25406	1.200	0.2383

The residual standard error is 0.5745 with 34 degrees of freedom. The  $R^2$  value is 0.1806, the adjusted  $R^2$  value is 0.08416, and the F-statistic is  $F(4, 34) = 1.873$ , with a  $p$ -value of 0.1378.

The equation of the regression line is given by:

Stress

=

1.93006

+

0.15560 · Missed Social Events

+

0.81061 · Living On Campus

+

0.21206 · Living With Family

+

0.30493 · Living With Roommates

### Interpretation

The results indicate that missed social events are a marginally significant predictor of stress levels ( $p = 0.0518$ ). Living situation, specifically living on campus, is also a marginally significant predictor ( $p = 0.0502$ ). However, living with family ( $p = 0.4590$ ) and living with roommates ( $p = 0.2383$ ) do not significantly predict stress.

## 5. Discussion and Results

For Research Question 1, we wished to investigate whether there is a relationship between a student’s academic workload and the prevalence of mental health issues in students. Preliminarily, we plotted a boxplot to look for any potential relationship between the two variables, and we found that on average, students with higher stress levels tend to have a higher academic workload. We later confirmed this with the simple linear regression test and found that academic workload is a significant predictor of stress.

For Research Question 2, we aimed to investigate whether there is a relationship between students’ academic workload and the amount of sleep they get each night. An earlier test indicated no significant relationship between hours of sleep and academic workload. Therefore, we decided to examine the effect of another factor: the number of social events a student misses on their hours of sleep. Initially, we plotted a boxplot of hours of sleep against the number of missed social events. The graph suggested that, on average, the more social events a student misses, the fewer hours of sleep they tend to get. Our analysis confirmed the earlier finding that academic workload does not have a significant relationship with hours of sleep but also revealed a strong relationship between the number of missed social events and students’ daily hours of sleep.

For Research Question 3, we studied whether mental health is affected by students’ living situations and their ability to maintain a work-life balance. To explore this relationship, we first plotted a boxplot of stress against the number of missed social events. The plot suggested that students who sometimes or always miss social events tend to experience more stress, while those who never miss social events experience less

stress. Then, we ran a test that concluded missing social events and living on campus may slightly affect stress levels, although the evidence is not very strong. In contrast, living with family or roommates does not appear to have a significant impact on stress.

Kruskal-Wallis test between stress and academic workload has the calculated *p-value* of 0.09332, which is greater than 0.05. Therefore, we fail to reject the null hypothesis. This means that stress levels and academic workload do not have a significant relationship.

Kruskal-Wallis test between stress and hours of sleep has the calculated *p-value* of 0.02971, which is larger than 0.05. Therefore, we fail to reject the null hypothesis. There is not enough evidence to determine a connection between hours of sleep and levels of stress among students. This suggests that variations in hours of sleep do not significantly relate to variations in stress levels.

Kruskal-Wallis test between stress and missed social events has the calculated *p-value* of 0.61, which is greater than 0.05. Therefore, we fail to reject the null hypothesis. This means that stress levels and missed social events do not have a significant relationship.

For the chi square tests of independence between academic workload and stress, anxiety, the respective *p* values are 0.0002288 and 0.0284 which are both less than the  $\alpha = 0.05$ , so we reject the null hypothesis. For the chi square tests of independence between academic workload and concentration, the *p* value is 0.3696, which is greater than the  $\alpha = 0.05$ , therefore we fail to reject the null hypothesis. This suggests that a student's academic workload has an impact on their stress or anxiety but not their concentration.

Chi square tests between students' living situation and their stress, anxiety and concentration yield *p* values of 0.5446, 0.359, and 0.8241 respectively. All of these *p* values are greater than the  $\alpha = 0.05$ , therefore we fail to reject the null hypothesis for all three of these chi square tests. This suggests that the living situation of a student doesn't have an association with their stress, anxiety and concentration.

Chi square tests between students' time management skills and their stress, anxiety and concentration yield *p* values of 0.2078, 0.5851, and 0.3151 respectively. All of these *p* values are greater than the  $\alpha = 0.05$ , therefore we fail to reject the null hypothesis for all three of these chi square tests. This suggests that a student's ability to manage their time doesn't have an association with their stress, anxiety and concentration.

Chi square tests between students' financials and their stress, anxiety and concentration yield *p* values of 0.7223, 0.7236, and 0.6308 respectively. All of these *p* values are greater than the  $\alpha = 0.05$ , therefore we fail to reject the null hypothesis for all three of these chi square tests. This suggests that a student's financial situation doesn't have an association with their stress, anxiety and concentration.

The results of the multiple and linear regression tests indicate that missed social events are a marginally significant predictor of stress levels ( $p = 0.0518$ ). Living situation, specifically living on campus, is also a marginally significant predictor ( $p = 0.0502$ ). However, living with family ( $p = 0.4590$ ) and living with roommates ( $p = 0.2383$ ) do not significantly predict stress.

Overall, the results of this study show that stress and anxiety among university students is primarily related to academic factors, with lifestyle factors such as sleep and socializing having less of an impact than expected.

## 6. Limitations

Four major limitations that our study faced were failing to reject the null hypotheses, survey limitations, the indirect nature of our research, and our failure to meet the expected counts for each group.

### 1. Failing to Reject Null Hypotheses

Many additional tests not included in this report failed to produce statistically significant results. For example, one linear regression test concluded that there is no relationship between academic workload and some mental health issues. This sort of result seemed counterintuitive and difficult to accept, which suggested there might be other factors that influence mental health that we did not consider.

### 2. Survey Limitations

Students may have rushed through the survey in order to receive course credit, without taking the time to reflect on their responses. Also, students may have not felt comfortable answering questions about their mental health truthfully, which could have led to response bias.

### 3. Indirect Nature of Research

Mental health is a complex topic, and it is difficult to be certain about the reliability and validity of our results when our study relied on indirect measures such as stress and anxiety to represent the broad spectrum of mental health.

### 4. Expected Counts

In regards to the chi square test, we were not able to satisfy the assumption that the expected counts for all groups is greater than or equal to five. In particular, only a single person rated their academic workload less than 3, and very few people responded 'never' to the questions about their stress, anxiety and concentration. The groups could not be combined in a way that ensured the expected counts would be equal to or greater than five, while also still providing meaningful data.

In the future, to address these limitations, we plan to incorporate attention-check questions to identify rushed or careless responses. We will also explore additional variables that may influence mental health to ensure a comprehensive understanding of the complexities of mental health. Finally, some open-ended questions could help students elaborate on their responses.

## 7. Conclusion

The linear regression analysis conducted for the three research questions of this study provides valuable insights into the factors affecting student stress and mental health. For Research Question 1, we found a significant positive relationship between academic workload and stress levels, suggesting that students with heavier workloads tend to experience higher levels of stress.

In Research Question 2, while academic workload did not significantly impact the amount of sleep, missed social events were found to significantly reduce hours of sleep, which may contribute to poor mental health.

For Research Question 3, missed social events and living situations were marginally significant predictors of stress, indicating that students who miss social events or live on campus tend to report higher stress levels.

These findings highlight the importance of considering multiple factors in understanding a student's overall well-being.

However, the limitations of this analysis, including potential issues with sample randomization and measurement quality, should be addressed in future research to enhance the robustness and accuracy of the results. Future studies should also explore other possible variables, such as social support and coping mechanisms, which may further explain the mental health challenges faced by students.

We also conducted Kruskal-Wallis testing for the three research questions for this study.

Research question 1 (RQ1), questions if academic workload contributes to a student's stress levels. The results of the Kruskal-Wallis test between the variables, stress and academic\_workload also suggested that there is not a significant relationship between a student's academic workload and levels and stress.

Next, the ANOVA test for the second research question, if a student's mental health is affected by their quality of sleep, suggested that there is not enough evidence to imply that there is a relationship between the variables hours\_sleep and stress.

Finally, the ANOVA test for RQ3 gave us insight on whether a student's mental health is affected by their ability to keep a work-life balance. The results of this analysis did not show evidence pointing towards a significant relationship between the variables, missed\_social\_events and stress.

We also performed the Chi-Squared Test of Independence for the first and third research questions. For research question 1, the results of the chi square test suggest there is a relationship between academic\_workload and the variables stress and anxiety. For research question 3, the chi square test did not indicate a significant relationship between the variables living\_situation, time\_management, financial\_problems and the stress, anxiety or concentration levels of a student.

Future and further research would greatly benefit from having a larger population size, since we could not go with the assumption that each group has a normal population distribution in each categorical group, since the CLT could not be satisfied due to the fact that each group did not have more than 30 observations.

## Documentation of AI Tool Usage

We utilized generative AI tools, namely ChatGPT, to help refine our ideas and improve readability. These tools were used primarily for brainstorming, grammar corrections, and the explanations of certain sections. However, all the analytical work, including statistical methods and interpretation of results, was done by the authors.

## 8. Citations:

1. <https://chatgpt.com/share/674a3a7d-83e4-8013-a381-4192bb2028da>
2. Beiter R, Nash R, McCrady M, Rhoades D, Linscomb M, Clarahan M, Sammut S. The prevalence and correlates of depression, anxiety, and stress in a sample of college students. J Affect Disord. 2015 Mar

1;173:90-6. doi: 10.1016/j.jad.2014.10.054.

(<https://www.sciencedirect.com/science/article/abs/pii/S0165032714006867?via%3Dihub>). Epub 2014 Nov 8. PMID: 25462401.

3. Hysing M, Harvey AG, Linton SJ, Askeland KG, Sivertsen B. Sleep and academic performance in later adolescence: results from a large population-based study. *J Sleep Res.* 2016 Jun;25(3):318-24. [doi:10.1111/jsr.12373](https://doi.org/10.1111/jsr.12373). Epub 2016 Jan 30. PMID: 26825591.
4. Hershner SD, Chervin RD. Causes and consequences of sleepiness among college students. *Nat Sci Sleep.* 2014 Jun 23;6:73-84. [doi:10.2147/NSS.S62907](https://doi.org/10.2147/NSS.S62907). PMID: 25018659; PMCID: PMC4075951.