Ania Shaheed, Riya Parikh, Adhya Hoskote
DS 110
Professor Kevin Gold
12 December, 2022

Undergraduate Stress and Sleep at Boston University

I. Introduction

Stress, whether acute or chronic, affects a vast majority, if not all, humans. One way stress can physically manifest is through sleep, which is closely linked to stress. Stress can negatively impact the quality and length of sleep, but insufficient sleep can increase stress levels[1]. Stress is a completely normal reaction to everyday pressures, but in a college environment, these pressures are somewhat different. Students deal with increased responsibilities and academic demands. According to the American College Health Association, 45.1% of college students in America reported a stress level of 4 on a 5-point scale, and 12.7% reported a stress level of 5[2]. These students likely experience some sort of disturbance in sleep, as they have high levels of stress. Individual stress levels are caused by a myriad of factors, but which ones are the most deterministic of how stressed a college student will be?

In society, there are widespread stereotypes and assumptions about different academic paths and their relationships to stress. The majority believe that STEM majors are more difficult in course material and that they have a larger workload compared to non-STEM majors. The minority argue that this is not an accurate generalization, because regardless of major, people at the top of their field have comparable amounts of work and stress. Anecdotally, it appears that students in certain academic majors are less stressed than students in other disciplines, having more free time and getting more sleep. If this is indeed true, there should be a pattern showing the relationship between academic discipline and stress or sleep.

The purpose of this study is to determine if these widespread ideas hold any truth. The study will primarily focus on two questions: “Does a student’s perceived stress level depend on their student’s major?” and “Does a student’s average amount of sleep depend on whether or not they are in a STEM field?” As such, the null hypothesis (H_0) for the first question is: There is no statistically significant difference between perceived stress levels among different majors. The alternative hypothesis (H_A) is: There is a statistically significant difference between perceived stress levels among different majors. H_0 for the second question is: There is no statistically significant difference between the amount of sleep students in a STEM major get as compared to students in a non-STEM major. H_A is: There is a statistically significant difference between the amount of sleep students in a STEM major get as compared to students in a non-STEM major.

II. Methodology

To answer the research questions, a qualitative study was conducted. We designed a Google Form asking respondents to detail their personal experiences with stress, academics,

[1] Stress and Sleep
[2] ACHA/NCHA II Fall 2018 Reference Group Data Report

sleep, and related variables[3]. In this study, the source population was undergraduate, full-time students at Boston University. We have the greatest access to these people as their peers, and these constraints were also intended to reduce variability through specific demographics. When designing this study, we determined that a minimum of 100 respondents would be adequate to run statistical tests and assume that the sample size is representative of the source population.

After creating the survey, we obtained 105 responses via text and encouraged participation from students in public. Once there was enough data, we exported the responses to a Google Sheet. We kept a copy of the raw data[4] and used Sheets to clean the data[5]. In this stage, we removed whitespace, screened for outliers, ensured there were no duplicates of answers in different cases or wording, and overall standardized the responses. After this, the table was exported as a CSV file for upload into Google Colab[6].

In Colab, we prepared the CSV file for exploratory data analysis using Pandas. Using Seaborn and Matplotlib, we created count-plots for all variables (they were all categorical). Here, we found that our sample wasn't perfectly representative. There were schools (and obviously majors) that we hadn't sampled. While we tried our best to include multiple grades and majors, due to nonresponse bias and other factors, the demographics were skewed towards freshmen and sophomores in Data Science and Business majors. We also found that many of our variables most pertinent to the study were ordinal, so we treated them as discrete numerical variables. Subsequently, we created histograms for each of these ordinal variables, broken down by different demographic variables. We overlaid one kernel distribution estimation (KDE) plot for each category of a demographic variable onto the histograms to see if their distributions differed. Interestingly, we noted quite a few seemingly prominent differences in the KDE trends underneath the graphs. Later, we found that these observations were not statistically significant (see below).

After exploring the data, it was time to run statistical tests. Since the data was categorical and we were examining the relationship between two unpaired groups, we determined that chi-square analysis was suitable for this study. One prerequisite for chi-square is that all expected and observed values must be greater than 5. To test this, we created a function, *five()* that took an independent and a dependent variable as input, creating a crosstab out of them using Pandas (essentially an array of observed values) and running them through *chi2_contingency* from SciPy.Stats (which would return an array of expected values). Our function iterated over the expected and observed arrays. If 80% of either array constituted of values less than or equal to 5, we encoded *five()* to return a boolean value of False. We also created another function, *chi()*, which would take an independent and dependent variable. These variables would be used to create a crosstab, which was then inputted into *chi2_contingency*. Based on the p-value, *chi()* would print whether or not there was a statistical relationship between the two variables at an alpha level of 0.05 for 95% confidence.

---

[3] Undergraduate Stress Survey
[4] Raw Data
[5] Cleaned Data
[6] Code

Using *five()*, we determined that three of our variable pairs were unsuitable for chi-square. An alternative test for small samples is Fisher's exact test, but the only functions in Python that run the test use 2 x 2 contingency tables. Therefore, we decided to introduce a threshold on the dependent variables for pairs that were unsuitable. To do this, we made the dependent variable binary, the threshold being the determiner. In the end, all of the variable pairs that explicitly answered our research questions were not statistically significant, since their p-values were too high. We also created a function to pair the remaining variables in every combination. We iterated over these pairs using *five()* and *chi()*, but there weren't any noteworthy findings.

III. Results

We failed to reject the null hypothesis for the research question about stress and major. We determined that the best way to display this result was through a histogram and KDE plot hybrid, where the categories for "Major" are separated by hue.

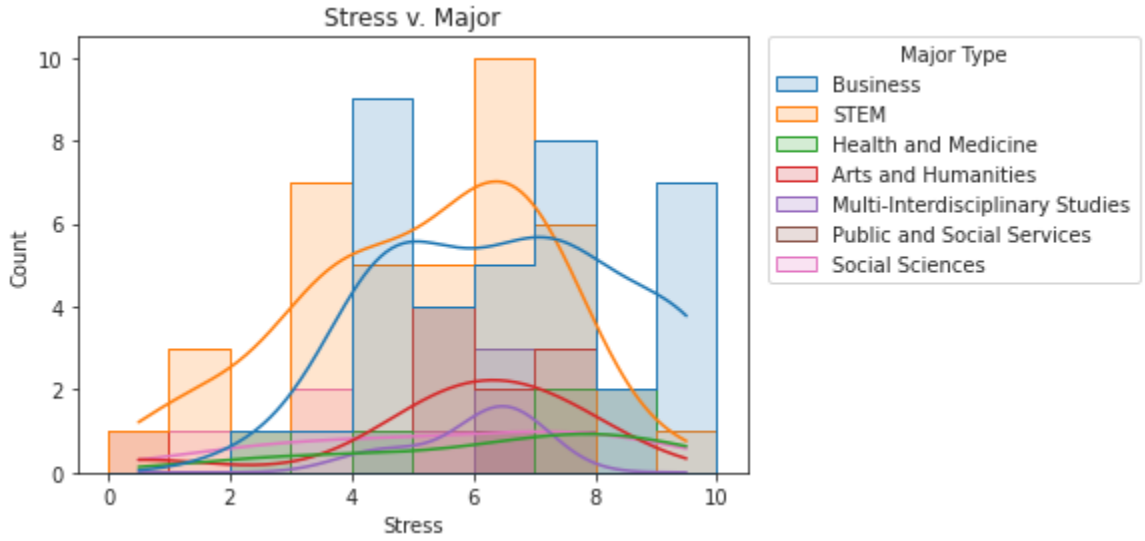

Figure 1: Bar graph and trend lines associated with Stress vs. Major

We also failed to reject the null hypothesis for the research question about sleep and STEM/Non-STEM. This result is best displayed in a stacked bar plot, which shows the uniformity in average hours of sleep on a weekday between STEM students and non-STEM students.

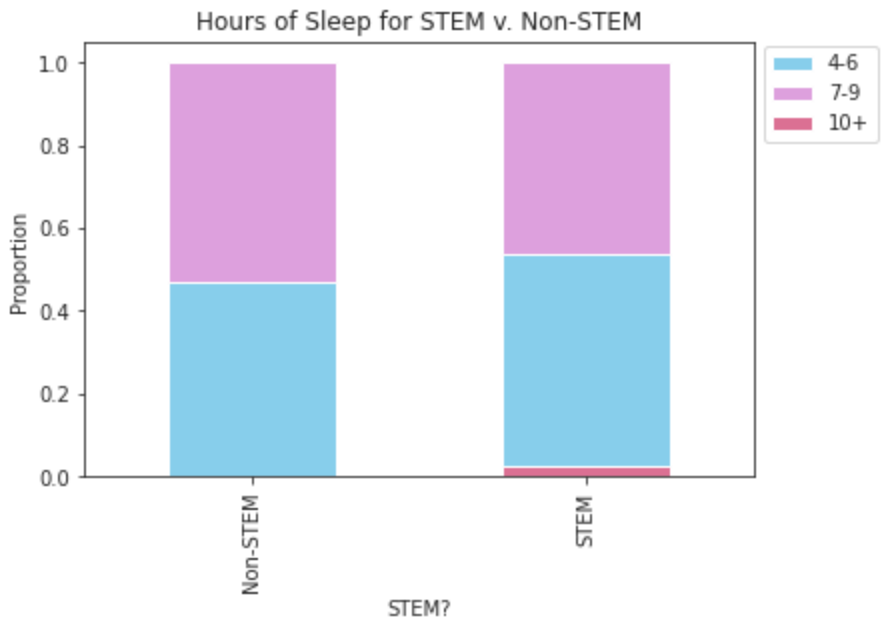

Figure 2: Stacked bar graph showing hours of sleep for STEM vs. Non-STEM students

Stress vs. Academic Path		Sleep vs. Academic Path			
Major vs. Avg. Stress (1-10)	p = 0.999 for a threshold of 7	Major vs. Hours Sleep Weekday	p = 0.999 for a threshold of 7 hours	Major vs. Hours Sleep Weekend	p = 0.996 for a threshold of 7 hours
Major Type vs. Avg. Stress (1-10)	p = 0.29	Major Type vs. Hours Sleep Weekday	p = 0.999	Major Type vs. Hours Sleep Weekend	p = 0.246
School vs. Avg. Stress (1-10)	p = 0.41	School vs. Hours Sleep Weekday	p = 0.999	School vs. Hours Sleep Weekend	p = 0.988
STEM vs. Avg. Stress (1-10)	p = 0.82	STEM vs. Hours Sleep Weekday	p = 0.918	STEM vs. Hours Sleep Weekend	p = 0.434

Table 1: p-values for relevant chi-square tests

The rest of the results, which are not very relevant, can be found on the Google Colab.

IV. Conclusions - What did you conclude from your results? Is there anything else your team learned from the experience?

From our results, we concluded that it is difficult to determine whether or not there are associations between stress and major as well as sleep and whether a student is in the STEM field among undergraduates at Boston University. This is partially due to the complexity of stress and potentially confounding variables that we did not analyze. From this experience, we've learned about the importance of data collection and sample size and how it can impact the results of research. We strongly believe that if we had a larger sample of data to draw conclusions from, we could have been more successful in analyzing our hypotheses, possibly rejecting them. Having a larger sample size would have given us more overlap/higher counts for our chi-squared tests and this would have increased the power of our tests. If we were given the opportunity to conduct a study like this in the future, we would definitely work towards getting more responses from a random sample of Boston University students.