Stress Impacts On Lifestyle Behavior

Gabriel Gladstone
dept. of Computer Engineering
Charlottesville, VA
zwy7ce@virginia.edu

***Abstract*—Stress is a contributor to adverse health effects such as depression and obesity. The explores the relationship between work-related stress and daily lifestyle. The study analyzed the online data of 1 participant over months to years based on the data source. The study concludes that work-related stress is a small but statistically significant factor in reducing a person's healthy lifestyle habits such as sleep, fitness, and gaming.**

I. Introduction

The study explores the effects of stress on behavior to answer the following questions.

- **RQ1**: What does an individual spend time doing under no stress?
- **RQ2**: How does stress influence an individual's lifestyle?

"There is a vast body of literature linking stress and social behavior" [3]. Stress has been explored in rodents and other animals. The physiological responses have been analyzed [2]. Stress is a contributor to many adverse effects such as depression [7] and obesity [10]. This study further explores the effects of stress on lifestyle changes. In this context lifestyle means the things an individual usually does. The study specifically explores the effects of long-term multi-tasking stress on lifestyle behavior. There have been studies using short-term multitasking activities to measure stress, but my online behavior data has allowed me to conduct this study to analyze longer-term correlations.

II. Methodology

A. Study Design

The study focuses on students and uses the online behavior data of 1 willing participant, me. The study compares a proxy measurement of stress with a wide range of activities an individual normally does including fitness, gaming, sleep, and work. However, more measures could easily be incorporated in the future. The timeline of the study varies based on the measure and collection mechanism of a given data source. Data ranges from all 4 years to a couple months. Sample sizes are clearly defined in the discussion. Statistical measurements for time-series correlation and causation are then explored in the discussion. The goal of the study is to draw conclusions about how stress in school can effect the behavior of students and how schools can design their curriculum to reduce stress.

B. Data Exploration

An ideal set of data would fully categorize the lifestyle behavior of a student's normal day. This could include sleep, attending classes, homework, clubs, eating, gaming, relaxing, and more. In this study I explore the possibility of using the following datasets.

- **School Work**
  - ☑ Canvas
  - ☑ Gradescope
  - ☐ Piazza: Difficult to develop measure
- **Fitness**
  - ☑ Iphone Health
- **Social Media**
  - ☐ Iphone analytics: Only past 7 days.
  - ☐ Iphone Screentime: Mac and custom parsing required.
- **Gaming**
  - ☑ Riot Games Data
  - ☐ Steam: Not used often.
- **Sleep**
  - ☑ Browser Search.
  - ☑ Google Location Timeline.

I capture a subset of these data sets for most of the categories in a normal day. With all online behavior studies, there are many data sources, but only data sources that are accessible and routinely used by the participant should be incorporated into an analysis. A future study would capture my social media Screentime data as that is a very fine grained view of daily social media habits.

C. Data Extraction and Processing

1) School Work Data:

The majority of stress in a student's life comes from school. Thus, I use school work, amount of work, difficulty of work, overlapping deadlines, and upcoming deadlines, as a proxy for stress in a student's life. I pull school data from Canvas and Gradescope.

I use davekats' canvas-student-data-export tool [5] to pull all canvas change logs. This includes assignments, discussions, announcements, and modules metadata. My goal was to build a timeline of work meaning I wanted to collect items of work with a start date, end date, and accompanying description for difficulty classification. This led me to use the assignment and discussion objects.

I collected Canvas data from April 2023 to April 2025. Cleaning was necessary for the Canvas Data. Some records were missing a start date, some were missing an end date, and some had a start date that was way too early for a student to begin thinking about the assignment. In some cases this led to over 100 assignments in progress for a student because the assignments were all assigned at the beginning of the year. To deal with missing dates, I assigned default thresholds and max thresholds based on school work categories as shown in Fig 1. I performed a simple string matching on the *title* field of the record to assign it a category. Records with missing assigned or due dates would get the associated dates assigned by the default threshold. Records with too long of a date range would get their assigned date moved to the max threshold with respect to the due dates. The due dates were usually pretty accurate.

```
default_thresholds = {"default": 3, "test": 4, "quiz": 2, "project": 14}
max_thresholds = {"default": 7, "test": 7, "quiz": 5, "project": 21}
```

Fig. 1: Date Thresholds (day) by Work Category

The Canvas preprocessing resulted in the assignment intervals shown in Fig 2.

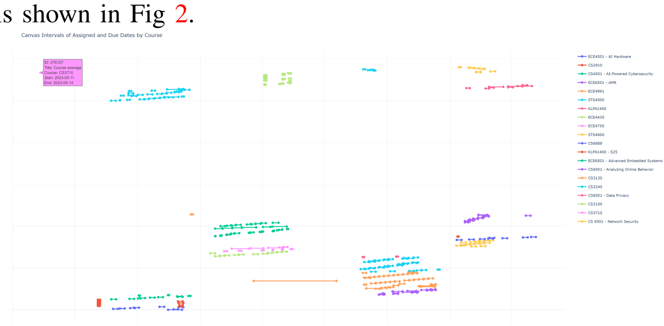

Fig. 2: Canvas Work Dates by Course

Unfortunately, UVA's migration to Canvas started in Spring 2023 so I needed another way to collect work data as I enrolled in Fall 2021. For this data, I pulled Gradescope, the primary assignment posting platform used before Canvas. Gradescope did not have a built-in or custom tool for exporting data so I used a custom HTML parse to pull the records from each course page. From initial observations, Gradescope dates were accurate so no preprocessing was used.

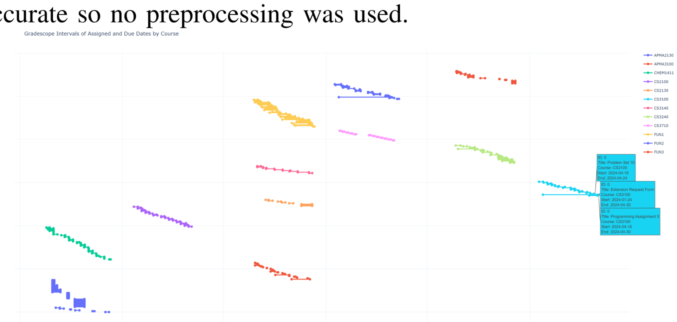

Fig. 3: Gradescope Work Dates by Course

Fig 3 shows Gradescope data from Jul 2021 to Jan 2024. This means that some records will overlap. This specifically happened for CS3100. Because the Gradescope dates were more accurate, Gradescope records were prioritized when the datasets were combined. To combine the datasets, a manual mapping of course name to course id was required as Gradescope and Canvas track ids differently.

Fig. 4: All Work Dates by Course

The combined work interval data is shown in Fig 4. This is is a pretty fine-grained data set of schoolwork the participant did over the course of their 4 year college education.

Now that we have school work data, I develop a measure to estimate stress.

$$S(t) = \sum_{i=1}^{N} \frac{w_i}{D_i + 1} + \alpha * O(t)$$

- t = day
- N = Total # of tasks.
- w_i = Weight of task.
- D_i = Days remaining until due date.
- α = Scaling factor
- $O(t)$ = Number of overlapping tasks on a day t.

The stress formula considers the importance of the task by category, assumes that a task becomes more stressful as the due date approaches, and assumes that a person experiences more stress as the number of overlapping tasks increases. I assign categories an arbitrary value of stress where tests are of course more stressful than regular work. I believe this measure to be a good representation of stress for a student based off of work. Unfortunately, I could not track when a student finished an assignment, which would mean that that assignment's stress should not be considered. All assignments are assumed to cause stress up until the due date. The arbitrary measure of stress will not effect the results of the correlation analysis, but may impact the magnitude of a causation analysis as this is an arbitrary score. Stress level and active school work are shown in Fig 5 only stress level is used in the analysis, but it can be seen they are related as active School Work is a factor in the stress formula.

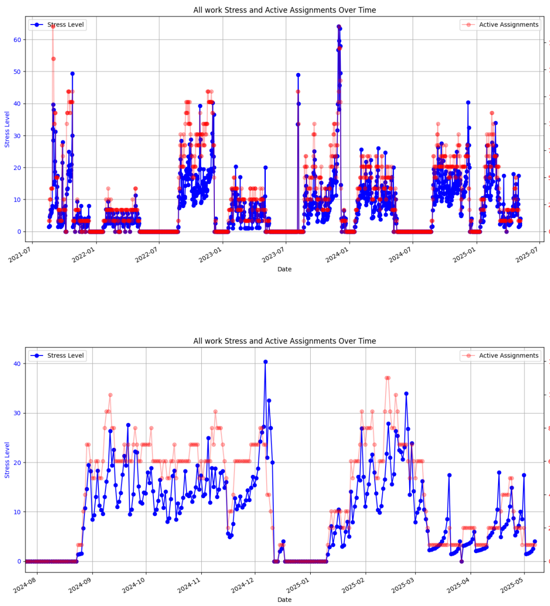

Fig. 5: Stress Level of Active School Work

2) *Sleep Data:*

The sleep data methodology from my previous study was used for this study. Search history data across all major browsers (Edge, Chrome, Firefox) as well as Google Timeline timestamp data was used. Refer to my study on a bespoke data collection framework [6]. The previous study proposed a method for estimating a user's sleep intervals and validated its accuracy. The start and end times of a user's sleep history can be seen in Fig 6. As shown, the sleep history is pretty expansive with a lot of data points.

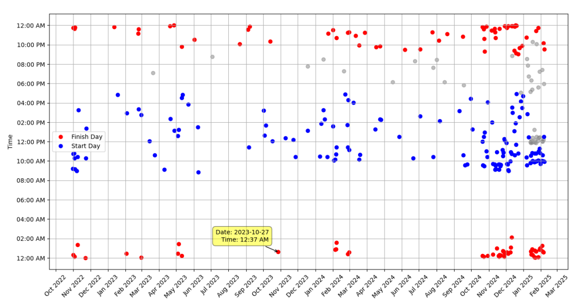

Fig. 6: Sleep Data

This study uses the tool to determine how many hours of sleep a participant receives per day shown in Fig 7. It can be seen that hours slept on average is higher than expected. This is because the study doesn't consider lag times in the method. My previous study determined a 1-2 hour lag time under very constricted circumstances. It can also be seen that the samples dramatically decreased in the method. This is because a sleep interval requires a start time with an accompanying finish time and much of the data is spread out from cleaning. I do not further process the data under any assumptions as my sleep behavior has changed over such a large time span.

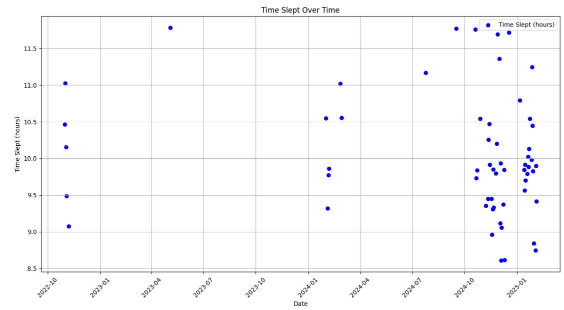

Fig. 7: Hours Slept per Day (N = 58)

3) *Fitness Data:*

Fitness data was exported from the iPhone Health App. I track steps per day, active energy burned, and basal energy burned. This data should be somewhat accurate as I always carry my phone with me unless doing strenuous activity (eg. running, climbing, soccer). Steps per day 8 seems reasonable as an adult should target about 10,000 steps per day [1] and while I don't hit this every day, I am within a reasonable range. The combined results of calories burned from active and basal burn is also reasonable as an adult can burn anywhere from 1,300 - 2000 calories on a normal day [4]. No further processing was performed after the initial sanity checks of the data. As shown, all fitness data has a decent sample size for analysis.

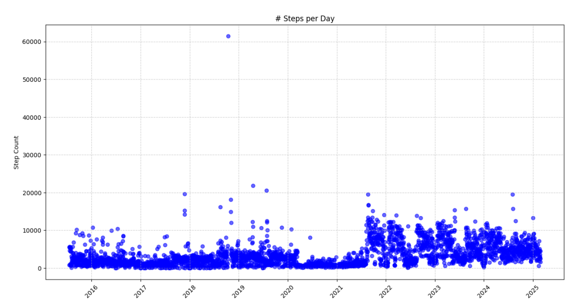

Fig. 8: # Steps per Day (N = 3473)

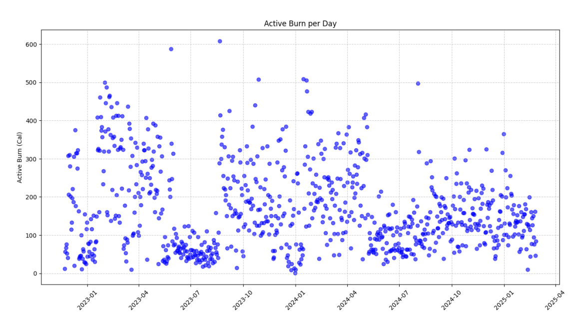

Fig. 9: Active Burn (Cal) per Day (N = 828)

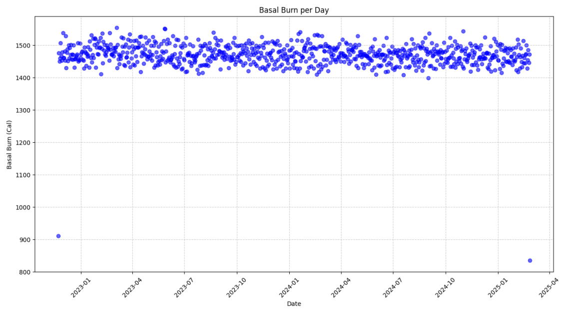

Fig. 10: Basal Burn (Cal) per Day (N = 828)

4) Gaming Data:

I am not much of a gamer, but when I game, I usually play on Riot Games. While there are other platforms, like Steam, I could have tracked. I chose only to pull my Riot Games data because of my gaming habits. I used Tainass' lol-data-fetcher [9] to pull my Riot Games match history shown in 11. This plot shows my gaming history over the past year with a reasonable sample size. No further processing or cleaning was done on the data as matches are tracked exactly when they begin and end.

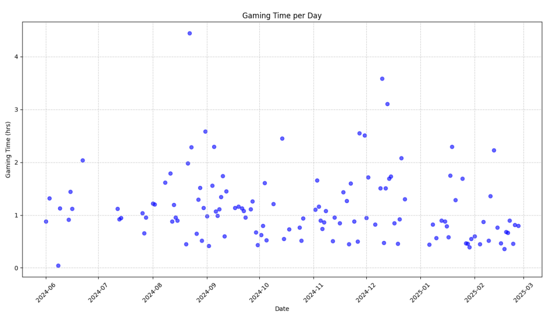

Fig. 11: Gaming Time per Day (N = 126)

D. Statistical Analysis Decisions

To better understand the data. I will perform a correlation and causation analysis.

To track correlation between school work (as a proxy for Stress) and other aspects of a student's daily life I use the Pearson correlation coefficient (PCC) with an associated P value. PCC measures the linear correlation between two variables from -1 to 1 [8]. I track correlation to better understand Stress' relationship with other factors. PCC does not comment on the cause of lifestyle changes, but a clear correlation may be used if future lifestyle prediction models.

To better understand causation linearly, I use ordinary least squares (OLS). This is a type of linear regression model. OLS can be applied to these time-series data sets under the following assumptions. I consider the relationship between lifestyle factors linear, even though stress doesn't cause an equal change in behavior at all levels. I also consider days to be independent even though a very stressful day will probably carry over to the next day. However, this is a reasonable analysis for an initial understanding of causation, where there are many more sample observations than predictors.

III. Results

A. Correlations

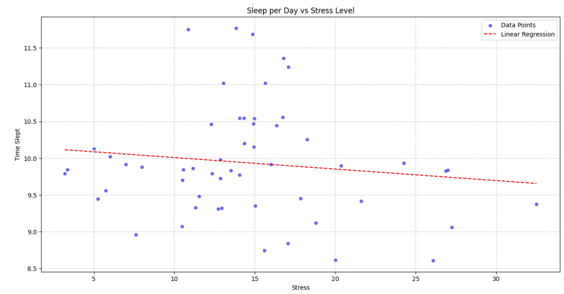

Fig. 12: Sleep vs Stress, $r = -0.13$, $p = 0.356 < 0.5$, $N = 54$

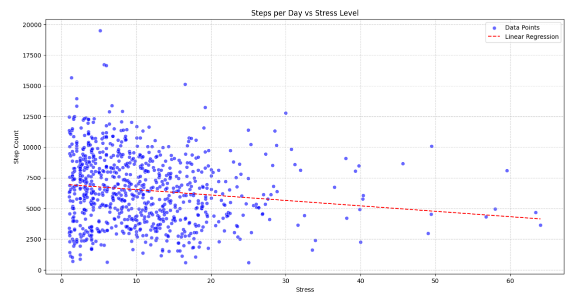

Fig. 13: Steps vs Stress, $r = -0.14$, $p = 0.356 \times 10^{-5} < 0.001$, $N = 824$

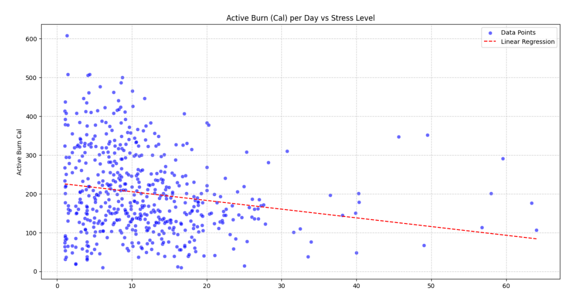

Fig. 14: Active Burn vs Stress, $r = -0.19$, $p = 1.035 \times 10^{-5} < 0.001$, $N = 533$

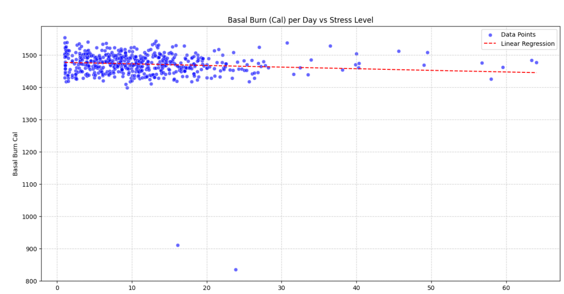

Fig. 15: Basal Burn vs Stress, $r = -0.10$, $p = 0.022 < 0.05$, $N = 533$

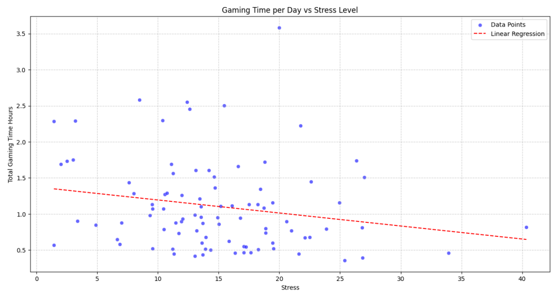

Fig. 16: Gaming vs Stress, $r = -0.20$, $p = 0.056 \approx 0.05$, $N = 91$

Test	Pearson Coefficient	P value	Samples
Sleep vs Stress	-0.13	0.5	54
Steps vs Stress	-0.14	0.001	824
Active Burn vs Stress	-0.19	0.001	533
Basal Burn vs Stress	-0.10	0.05	533
Gaming vs Stress	-0.20	0.05	91

TABLE I: Correlation Summary

There are some meaningful linear relationships between stress in daily lifestyle behavior as shown in the correlation plots and summary Tab I. For all relationships, analysis at 0 stress was removed because it was the data was too random. This is because no stress in this measure means no class assignments where students are likely to be on vacation and have completely different daily lifestyle behaviors as a result.

Sleep has a weak negative correlation with stress. This makes sense because stress may cause lack of sleep, in this case possibly from working too much. Unfortunately, with the low sample size and high p value, I cannot confidently draw conclusions for this.

The rest of the tests also have a weak negative correlation with stress. This makes sense for the same reasons, stress from work will lead a student to spend more time on completing that work, which may take away from other lifestyle behaviors. With much lower p values and much higher number of samples, I am more confident in the conclusion that stress has aversive effects on most other activities of a student's life.

B. Causation

I guess at the reasoning behind these correlations, but the Pearson coefficient doesn't comment on causation. To better understand causation, I perform a OLS linear regression.

Target	β	P Value	R^2	N
Sleep	-0.0157	0.356	0.016	54
Steps	-44	$<$ 0.0001	0.019	824
Active Burn	-2.2	$<$ 0.0001	0.036	533
Basal Burn	-0.5	$<$ 0.022	0.01	533
Gaming	-0.018	0.05	0.04	91

TABLE II: OLS Model Summary

The results of stress on the target activities are shown in Tab II. Stress' effect on the number of steps a person takes in a day and the calories they burn in a day (active and basal) is statistically significant, but negligible, accounting for $\leq 4\%$ or variance in the time-series data sets. The weak negative relationship between hours a person games in a day is less statistically significant and the hours a person sleeps in a day is not statistically significant.

Overall, stress is a weak but statistical cause of decreases in some lifetime activities. This is understandable as there are many factors that affect how much time a person spends on an activity and some of them like school work or other obligations are simply out of the person's control.

IV. Discussion

1) Future Work:

There is a lot of work that could be done to improve this study.

First, other sources of stress could have been analyzed. Most of the assignment record metadata included grades. Bad grades of course can be stressful. This factor could have been considered in the creation of a stress measure. Known anti-stressors could have also been considered in the formula including holidays, short breaks, and birthdays, although this would require more detailed record-keeping for the participant and could be hard to quantify.

To better understand the relationship between stress and daily activities non-linear models like Mutual Information (MI) regression could have been explored. MI regression quantifies the information obtained about 1 random variable by observing another random variable. Although, in this case, I don't believe a non-linear analysis would have brought more insights as both the causation and correlation was weak.

2) Ethical Considerations:

The singular participant in this study consented to their data collection and analysis. While the data may be used to infer sensitive characteristics about the individual such as if they are experiencing depressive disorders, no mental characterization of the individual was carried out in this study. All sensitive data was also anonymized before being used or published.

3) Limitations:

To make the study as scalable as possible, I created a measure of stress using work and multi-tasking as a proxy. This of course has limits as there are many factors that influence stress. There were also some limitations in the datasets, especially sleep data which, after processing, had a small sample size. There were even limitations in the work interval data from imprecise assigned dates and course-grained/subjective categorizations. A future study would implement a continuous and qualitative survey component throughout the time frame of the study.

V. Conclusion

In this paper I show that work-related stress is a weak, but statistically significant cause for a decrease in lifestyle activities. I analyzed a wide range of activities including measures on sleep, fitness, and gaming to characterize a person's day. The research concludes that daily stress is only one small factor in a person's lifestyle change which is promising as stress should not be a controlling factor in a person's life. I understand further work and activity analysis needs to take place, but I hope this work is used as a starting point.

References

[1] *Average Steps Per Day by Age, Gender, Occupation, and Country*. Healthline. Nov. 27, 2018. URL: https://www.healthline.com/health/average-steps-per-day (visited on 03/24/2025).

[2] Linda Becker et al. "Physiological stress responses to digital single- and multitasking demands in younger and older adults". In: *Psychoneuroendocrinology* 174 (Apr. 1, 2025), p. 107376. ISSN: 0306-4530. DOI: 10.1016/j.psyneuen.2025.107376. URL: https://www.sciencedirect.com/science/article/pii/S030645302500099X (visited on 03/24/2025).

[3] Annaliese K. Beery and Daniela Kaufer. "Stress, social behavior, and resilience: Insights from rodents". In: *Neurobiology of Stress* 1 (2015), pp. 116–127. ISSN: 2352-2895. DOI: https://doi.org/10.1016/j.ynstr.2014.10.004. URL: https://www.sciencedirect.com/science/article/pii/S2352289514000125.

[4] *Burn Calories By Reading This Article!* Cleveland Clinic. URL: https://health.clevelandclinic.org/calories-burned-in-a-day (visited on 03/24/2025).

[5] David. *davekats/canvas-student-data-export*. original-date: 2019-08-16T05:36:38Z. Mar. 20, 2025. URL: https://github.com/davekats/canvas-student-data-export (visited on 03/24/2025).

[6] Gabriel Gladstone. *gladstone-9/cs6501-bespoke-framework*. original-date: 2025-02-11T23:56:20Z. Feb. 12, 2025. URL: https://github.com/gladstone-9/cs6501-bespoke-framework (visited on 03/24/2025).

[7] Constance Hammen. "Stress and Depression". In: *Annual Review of Clinical Psychology* 1 (Volume 1, 2005 Apr. 27, 2005). Publisher: Annual Reviews, pp. 293–319. ISSN: 1548-5943, 1548-5951. DOI: 10.1146/annurev.clinpsy.1.102803.143938. URL: https://www.annualreviews.org/content/journals/10.1146/annurev.clinpsy.1.102803.143938 (visited on 03/24/2025).

[8] *Pearson correlation coefficient*. In: *Wikipedia*. Page Version ID: 1280362748. Mar. 14, 2025. URL: https://en.wikipedia.org/w/index.php?title=Pearson_correlation_coefficient&oldid=1280362748 (visited on 03/24/2025).

[9] Rui Pereira. *Tyainss/lol-data-fetcher*. original-date: 2024-05-18T23:47:59Z. Oct. 5, 2024. URL: https://github.com/Tyainss/lol-data-fetcher (visited on 03/24/2025).

[10] Susan J. Torres and Caryl A. Nowson. "Relationship between stress, eating behavior, and obesity". In: *Nutrition* 23.11 (2007), pp. 887–894. ISSN: 0899-9007. DOI: https://doi.org/10.1016/j.nut.2007.08.008. URL: https://www.sciencedirect.com/science/article/pii/S0899900707002493.

VI. Appendix