

Data-Driven Analysis of Daily Stress Levels and the Effectiveness of Stress-Relief Activities

Gian Karl L. Colinares

College of Computing and Information Technologies

National University - Manila

Manila, Philippines

colinaresgl@students.national-u.edu.ph

Abstract—This self-tracking study investigates the association between daily stress levels and the effectiveness of various stress-relief activities, utilizing physiological data collected from smartwatches. Stress scores and categories are documented before and after activities such as exercise, gaming, and music listening to assess changes in stress levels. Employing a quantified-self methodology that integrates wearable sensor data with manual logging, the study evaluates whether physical activities yield greater stress reduction compared to passive leisure activities. The results aim to inform evidence-based strategies for enhancing personal stress management in academic environments.

Index Terms—quantified-self, wearable technology, stress management, personal informatics, heart rate variability

I. INTRODUCTION

Stress has become a significant concern among college students, as academic demands and lifestyle factors strongly influence mental well-being and overall quality of life [1]. Studies report that a large proportion of university students experience moderate to high levels of stress due to academic pressure and related responsibilities [2]. Because prolonged stress may negatively affect both health and academic performance, early monitoring and effective stress management strategies are essential.

Recent advances in wearable technology have enabled continuous monitoring of physiological indicators such as heart rate and heart rate variability (HRV), which are closely associated with stress responses [3]. Research shows that wearable sensors can successfully identify physiological patterns linked to stress in real-world settings, supporting their usefulness for behavioral and health studies [4]. Furthermore, validated wearable devices provide cardiovascular measurements with accuracy comparable to professional-grade sensors, making them suitable tools for personal health tracking [5].

This study adopts a quantified-self approach, emphasizing the use of self-tracking technologies to better understand personal behavior and health patterns. By integrating smartwatch-derived stress data with manually logged activities, the project aims to generate individualized insights that support evidence-based stress management.

A. Research Focus

This study focuses on Personal Informatics and the Data-Driven Analysis of Stress Management. Specifically, it examines the correlation between daily stress levels, as measured by

smartwatch physiological data, and the effectiveness of various stress-relief activities (e.g., exercise, music, and gaming) in a real-world, academic setting.

B. Importance of the Topic

Stress is a pervasive issue, particularly in academic environments, where it can negatively impact mental health and performance. While many stress-relief methods exist, their effectiveness varies significantly between individuals. This study is critical because it moves away from generalized advice and uses personalized data to determine which specific interventions actually work for each user, enabling more efficient, optimized stress management.

C. Prior Research Findings

Prior research has established three key pillars for this study:

Device Reliability: Studies [6] confirm that consumer-grade wearables are reliable for behavioral health research.

Activity Efficacy: Established literature [7] affirms that activities such as music and physical exercise can significantly reduce cortisol levels and perceived stress.

D. Research Gap

While general studies prove that "exercise is good for stress," they often lack personal context. Most research is conducted in controlled laboratory settings or across large, heterogeneous groups. This project addresses the gap in individualized responses: how a specific person's stress levels fluctuate throughout a standard academic week, and which activities provide the highest "stress improvement" score for them.

E. Research Goals

The primary goals of this research are:

- Evaluate the quantitative impact of various activities on stress reduction
- Map daily and weekly stress patterns to identify peak stress periods
- Determine whether initial stress intensity (Mild, Moderate, Severe) affects the effectiveness of specific activities

F. Research Question

RQ1: Is there a statistically significant difference in stress scores before and after performing stress-relief activities?

II. REVIEW OF RELATED LITERATURE

A. Wearable Sensors for Stress Monitoring

Wearable technologies have become reliable tools for continuous health monitoring, enabling researchers to capture real-time physiological signals associated with stress. Sarker et al. emphasized that wearable sensors allow personalized health insights by tracking behavioral and physiological data, supporting proactive health interventions and self-management strategies [8]. Similarly, research on heart rate variability (HRV) alignment shows that wearable-derived physiological metrics can serve as meaningful indicators of emotional and psychological states, reinforcing their applicability in stress-related studies [9]. These findings support the use of smartwatch-generated stress scores as a valid data source for the present research.

B. Effects of Physical Activity on Stress

Physical activity is widely recognized as a protective factor against psychological distress. A systematic review involving over 20,000 university students found a consistent association between higher physical activity levels and lower stress, anxiety, and depression, as well as improved subjective well-being [10]. Meta-analytic evidence further confirms that structured physical activity interventions significantly enhance multiple dimensions of mental health, including stress reduction and emotional stability [11]. These studies suggest that active stress-relief strategies, such as exercise and walking, may produce more substantial stress reductions than passive activities.

C. Music and Stress Recovery

Music exposure has also been shown to support emotional regulation and recovery from stress. Linnemann et al. reported that listening to music in daily life can reduce perceived stress and cortisol levels, highlighting its effectiveness as an accessible coping strategy [7]. More recent experimental findings indicate that music can accelerate stress recovery by influencing autonomic nervous system responses, suggesting measurable physiological benefits beyond subjective relaxation [13]. This evidence supports including music listening as a key activity in the current study.

D. Quantified-Self and Personal Informatics

The quantified-self framework emphasizes self-tracking as a method for improving self-awareness and behavioral outcomes. Li et al. proposed a stage-based model of personal informatics consisting of preparation, collection, integration, reflection, and action, illustrating how individuals can transform personal data into meaningful insights that guide behavior change [14]. Such models demonstrate the scientific value of single-subject data when systematically collected and analyzed, aligning directly with the personalized methodology used in this project.

E. Privacy Considerations in Sensor-Based Research

As wearable devices collect sensitive behavioral data, ethical data management is critical. Stopczynski et al. highlighted that sensor-driven human data collection must ensure transparency, informed consent, and secure storage, given the private nature of self-tracked information [15]. Incorporating these safeguards strengthens the methodological integrity of personal analytics studies.

III. METHODOLOGY

This chapter describes the research design, participant profile, data collection procedures, operational definitions, data preparation strategies, and planned statistical analyses. The purpose is to ensure that the study can be replicated by clearly outlining how the data will be gathered and analyzed.

A. Participants

This study was designed as a self-tracking research project in which the researcher served as the sole participant. The participant was a 23-year-old undergraduate college student who experienced regular academic demands and daily stressors, making him a suitable subject for examining personal stress patterns.

B. Data Collection Methods

This study used a quantitative self-tracking research design. Physiological stress data were collected with a smartwatch and combined with activity records that participants logged by hand. The final dataset had **449 rows and 7 columns**, each row showing an observation across several variables related to stress and daily activities.

1) *Variables Collected:* The following variables were recorded:

- **Date** – calendar date of each observation
- **Start Time (before activity)** – start time of the activity
- **Stress score (before activity)** – physiological measurement from smartwatch
- **Stress category** – smartwatch-generated classification (Relaxed, Mild, Moderate, Severe)
- **Activity performed** – type of stress-relief intervention
- **End Time (after activity)** – end time of the activity
- **Stress score (after activity)** – physiological measurement from smartwatch
- **Stress improvement** – calculated difference between before and after scores

2) *Frequency of Data Logging:* Data were recorded daily, immediately before and after performing a stress-relief activity to ensure accuracy and reduce recall bias.

3) *Tools and Applications:*

- **Mi Fitness Smartwatch Application** – for capturing stress scores and categories based on heart rate and other biometric indicators
- **Microsoft Excel** – for organizing and storing the dataset
- **Manual Activity Log** – for recording activity type and timestamps

- **Python (pandas, scipy, matplotlib, seaborn)** – for statistical analysis and visualization

The same smartwatch device was used throughout the study to maintain measurement consistency.

4) Data Collection Period: Data were collected over approximately 9- 12 weeks, allowing sufficient observations to identify patterns and relationships.

C. Operational Definitions

To ensure clarity and consistency, each variable is defined as follows:

Stress Score: A numerical value ranging from 1–100 generated by the Mi Fitness smartwatch, representing the participant's physiological stress level based on heart rate and other biometric indicators.

Stress Category: A qualitative classification provided by the smartwatch algorithm (Relaxed, Mild, Moderate, Severe) based on the stress score.

Activity Performed: The stress-relief action undertaken after the initial stress measurement, including activities such as walking, exercising, listening to music, watching movies, reading, meditating, or playing games.

Time of Day: The period when the activity occurred, recorded using start and end times and later grouped into morning (5:00 AM–11:59 AM), afternoon (12:00 PM–4:59 PM), evening (5:00 PM–8:59 PM), or night (9:00 PM–4:59 AM) for analysis.

Stress After Activity: The stress score recorded immediately following the completion of the activity.

Stress Improvement: The numerical difference between the pre-activity and post-activity stress scores (Before - After). Positive values indicate stress reduction, while negative values indicate stress increase.

Activity Duration: The time elapsed between the start and end of each activity, calculated in minutes.

D. Data Cleaning and Preparation

Before analysis, the dataset will undergo preprocessing to ensure accuracy and reliability. Planned steps include:

- 1) Removing whitespace from categorical variables to ensure consistency
- 2) Converting date strings to datetime format for temporal analysis
- 3) Extracting temporal features, including day of week, week number, and month
- 4) Parsing time data and creating time-of-day categories
- 5) Calculating activity duration from start and end times
- 6) Standardizing activity names to avoid duplicate labels (e.g., “workout” vs. “exercise”)
- 7) Validating that all stress scores fell within the valid range of 1–100
- 8) Checking for missing entries and removing incomplete records
- 9) Detecting outliers using the Inter-quartile Range (IQR) method

- 10) Flagging outliers for review while retaining them in the analysis, as they may represent genuine extreme stress events

E. Statistical Analysis

The study will apply descriptive and inferential statistical techniques to evaluate stress patterns and the effectiveness of activities.

1) Descriptive Statistics: The following were calculated to summarize the dataset:

- Mean and median stress scores (before and after activities)
- Standard deviation
- Frequency distribution of stress categories
- Average stress improvement per activity
- Activity duration statistics

2) Inferential Statistics: The following tests were conducted:

Paired Sample t-test: To determine whether there was a statistically significant difference between stress scores before and after activities. This test was appropriate because the same participant was measured twice (before and after each intervention), creating paired observations. Assumptions checked included normality of differences (Shapiro-Wilk test) and paired structure of data.

Effect Size: Cohen's d was calculated for the paired t-test to quantify the magnitude of the difference between pre- and post-stress scores, providing practical significance beyond statistical significance.

Correlation Analysis: Pearson correlation coefficients were calculated to examine relationships between continuous variables (e.g., stress before, stress after, improvement, duration).

3) Significance Level: All statistical tests used $\alpha = 0.05$ as the threshold for statistical significance.

F. Data Visualization

Results will be presented using:

- Bar charts – average stress reduction per activity and time period
- Line graphs – daily stress trends over time
- Box plots – variation in stress improvement across categories
- Scatter plots – before vs. after stress score comparisons
- Histograms – distribution of stress scores and improvements
- Heatmaps – correlation matrices of numeric variables

Visualization will help reveal patterns that may not be immediately visible through numerical summaries alone.

G. Bias and Measurement Considerations

Several potential sources of bias were acknowledged:

Single-participant limitation: Results represent an individual case study and may not generalize to a broader population. Findings are specific to the stress patterns of one 23-year-old college student.

Device-based measurement: Smartwatch stress scores rely on physiological indicators (heart rate and other biometrics) and may not fully capture psychological or emotional dimensions of stress. The proprietary algorithm used by the Mi Fitness app is not transparent.

Activity selection bias: The participant chose which activity to perform based on personal preference and availability rather than random assignment, potentially influencing outcomes.

Temporal confounds: External factors such as academic deadlines, sleep quality, caffeine intake, and social interactions were not systematically controlled or measured.

No control group: Without a control condition or placebo activities, it is difficult to isolate the specific effects of activities from natural stress fluctuation or regression to the mean.

Measurement consistency: The same smartwatch device was used throughout the study to minimize inter-device measurement variability.

Immediate logging: Data were recorded immediately before and after each activity to minimize recall bias.

Hawthorne effect: Awareness of being monitored may have influenced the participant's behavior or stress perception.

Despite these limitations, the study provides valuable insights into individual stress patterns and the potential effectiveness of self-directed stress-relief interventions using wearable technology.

IV. RESULTS AND DISCUSSION

This chapter presents the findings from the analysis of stress tracking data collected over a 78-day period from November 22, 2025, to February 7, 2026. The study examined 449 stress measurements taken before and after performing various stress-relief activities. Results are organized into descriptive statistics, exploratory data analysis with visualizations, and inferential statistical testing to address the research question: *Is there a significant difference between stress scores before and after performing stress-relief activities?*

A. Overview of the Dataset

The dataset comprises 449 complete observations spanning 78 days, with an average of 5.8 measurements per day. Data were collected using a Mi Fitness smartwatch for physiological stress measurements (1–100 scale) and manual logging for activity types and timestamps. All stress scores fell within the valid range, with no missing values in the final cleaned dataset. Figure 1 presents the key characteristics of the dataset.

Dataset loaded successfully!								
Shape: (449, 8)								
First few rows:								
#	Date	Start Time (Before)	Stress Score (Before)	Stress Category	Activity Performed	End Time (After)	Stress Score (After)	Stress Improvement
0	22-Nov-25	6:55 PM	32	Mild	Workout	8:05 PM	25	7
1	23-Nov-25	3:00 PM	35	Mild	Reading	3:30 PM	30	5
2	24-Nov-25	8:00 PM	46	Mild	Watching	8:30 PM	21	25
3	25-Nov-25	5:00 AM	28	Mild	Walking	5:45 AM	16	12
4	26-Nov-25	6:30 AM	20	Relaxed	Reading	7:00 AM	14	6

Fig. 1. Dataset summary statistics showing total observations (449), data collection period (78 days from November 22, 2025 to February 7, 2026), average measurements per day (5.8), and unique activities tracked (10+).

The dataset demonstrates consistent daily tracking with minimal gaps, indicating strong adherence to the data collection protocol throughout the study period.

B. Descriptive Statistics

1) *Central Tendency and Variability:* Mean stress scores before activities ($M = 27.41$, $SD = 10.83$) were higher than after activities ($M = 24.15$, $SD = 10.58$), suggesting an overall reduction in stress levels. The average stress improvement was 3.26 points ($SD = 6.76$), with a median improvement of 4.00 points. Activity duration averaged 58.10 minutes ($SD = 141.47$), with a median of 30.00 minutes, indicating considerable variability in session length. Figure 2 summarizes these descriptive statistics.

DATA QUALITY ASSESSMENT SUMMARY													
	Column	Data Type	Non-Null	Missing	Unique	count	mean	std	min	25%	50%	75%	max
0	Date	object	449	0	78	—	—	—	—	—	—	—	—
1	Start Time (Before)	object	449	0	199	—	—	—	—	—	—	—	—
2	Stress Score (Before)	int64	449	0	49	449.0	27.41	10.83	4.0	18.0	30.0	36.0	60.0
3	Stress Category	object	449	0	5	—	—	—	—	—	—	—	—
4	Activity Performed	object	449	0	33	—	—	—	—	—	—	—	—
5	End Time (After)	object	449	0	251	—	—	—	—	—	—	—	—
6	Stress Score (After)	int64	449	0	44	449.0	24.15	10.58	3.0	15.0	25.0	33.0	51.0
7	Stress Improvement	int64	449	0	38	449.0	3.26	6.76	16.0	-2.0	4.0	7.0	26.0

Fig. 2. Descriptive statistics for stress measurements showing count, mean, median, standard deviation, and quartiles for stress scores before and after activities, stress improvement, and activity duration. All statistics are based on 449 complete observations.

The variability in stress improvement ($SD = 6.76$) suggests that interventions did not produce uniform results across sessions, with effectiveness varying by activity type, initial stress level, and context. The substantial variability in duration ($SD = 141.47$ minutes) reflects the naturalistic, self-directed nature of the interventions, with session lengths ranging from brief 10-minute activities to extended multi-hour engagements.

2) *Stress Category Distribution:* Analysis of initial stress categories revealed that most measurements (53.9%) fell within the Mild stress category (26–50 points), followed by Relaxed (44.1%, 1–25 points) and Moderate (2.0%, 51–80 points). No measurements reached the Severe threshold (81+ points), indicating that baseline stress levels remained within manageable ranges throughout the study period. This distribution is illustrated in Figure 3.

The predominance of Mild and Relaxed categories suggests that the participant maintained generally manageable stress levels and utilized stress-relief activities preventively rather than waiting for stress to reach severe levels.

C. Visualizations and Patterns

1) *Distribution of Stress Scores:* Figure 4 displays histograms comparing the distribution of stress scores before and after activities. The pre-activity distribution is approximately normal and centers around 27–30 points, while the post-activity distribution shows a leftward shift with a center around 23–25 points. This visual shift provides initial evidence of

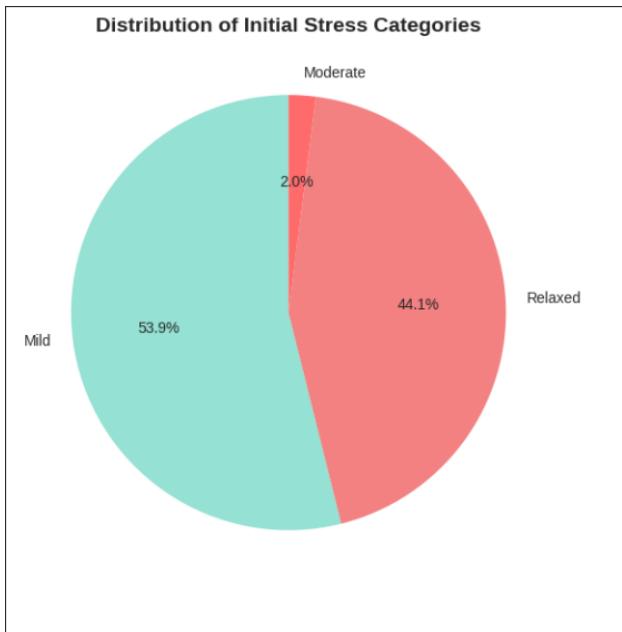


Fig. 3. Distribution of initial stress categories across all 449 measurements. The predominance of Mild (53.9%, 242 observations) and Relaxed (44.1%, 198 observations) categories indicates that most interventions were performed during typical daily stress levels rather than extreme stress episodes. Only 2.0% (9 observations) fell into the Moderate category.

stress reduction following interventions. Both distributions maintain similar shapes, suggesting that activities reduce stress levels while preserving the overall variability in individual responses.

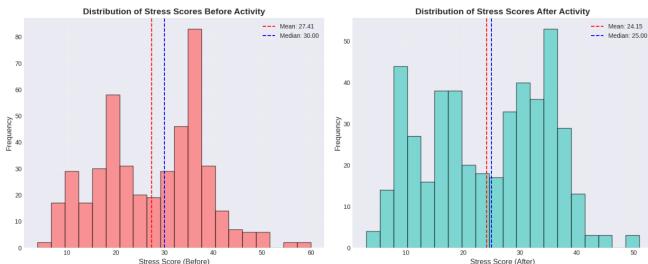


Fig. 4. Distribution of stress scores before (left) and after (right) performing activities. The leftward shift in the after-activity distribution demonstrates a general decrease in stress levels of approximately 3 points. Mean and median lines are indicated with dashed vertical lines.

2) *Before vs. After Comparison:* Figure 5 presents a scatter plot comparing stress scores before and after activities. Points below the diagonal line ($y = x$) represent cases where stress decreased, while points above indicate stress increased. The majority of points (70.2%, 315 observations) fall below the line, confirming that most interventions resulted in stress reduction. The strong correlation between before and after scores ($r = 0.801, p < 0.001$) suggests that higher initial stress tends to predict higher post-activity stress, though with measurable reduction. This pattern indicates that while activities reduce stress, they do not eliminate individual differences in baseline stress levels.



Fig. 5. Relationship between stress scores before and after activities. Each point represents one observation ($N = 449$). Points are color-coded by stress improvement magnitude, with warmer colors indicating greater reduction. The concentration of points below the diagonal line ($y = x$) indicates consistent stress reduction across varying initial stress levels.

3) *Temporal Patterns:* Analysis of stress patterns by time of day revealed notable variations in both baseline stress and intervention effectiveness, as shown in Figure 6. Morning periods (5 AM–12 PM) accounted for the majority of observations (45.4%, 204 measurements), while afternoon (23.2%, 104 observations), evening (15.8%, 71 observations), and night (15.6%, 70 observations) periods were less frequent. Different time periods showed varying baseline stress levels and improvement magnitudes, suggesting that the timing of interventions may influence both when stress-relief activities are most needed and their relative effectiveness.

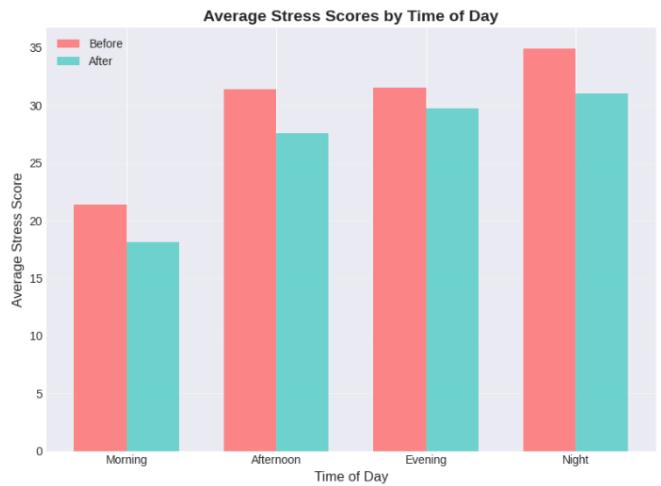


Fig. 6. Average stress scores before and after activities by time of day (Morning: 5 AM–12 PM, Afternoon: 12 PM–5 PM, Evening: 5 PM–9 PM, Night: 9 PM–5 AM). The grouped bar chart shows both initial stress levels (darker bars) and post-activity stress (lighter bars) for each time period, allowing comparison of baseline stress and intervention effectiveness across different times.

4) *Activity Effectiveness:* Figure 7 ranks activities by their average stress improvement, focusing on activities with adequate sample sizes for reliable estimation. Workout activities produced the highest average improvement among frequently performed activities ($M = 7.71$ points, $n = 7$), followed by meditation ($M = 7.00$ points, $n = 24$) and listening to music ($M = 6.88$ points, $n = 108$). Walking ($M = 3.92$ points, $n = 50$) and watching content ($M = 3.07$ points, $n = 44$) showed more modest but still positive improvements. The variation in effectiveness across activities suggests that different intervention types produce different magnitudes of stress reduction, though all examined activities showed some benefit.



Fig. 7. Average stress improvement by activity type, ranked from highest to lowest effectiveness. Bars are color-coded with green indicating positive improvement. Only activities with sufficient sample sizes ($n \geq 7$) are shown to ensure reliable estimates. Error bars, if present, indicate variability in effectiveness. Sample sizes are noted for each activity.

5) *Activity Frequency:* Figure 8 shows the frequency distribution of different activities performed throughout the study. Listening to music was the most common intervention ($n = 108$, 24.1%), followed by reading ($n = 82$, 18.3%) and gaming ($n = 62$, 13.8%). Walking ($n = 50$, 11.1%) and watching content ($n = 44$, 9.8%) were also frequently used. The discrepancy between frequency and effectiveness (e.g., listening to music was most frequent but showed moderate effectiveness) suggests that activity selection is driven by multiple factors including accessibility, enjoyment, and convenience, not solely by stress-reduction potential.

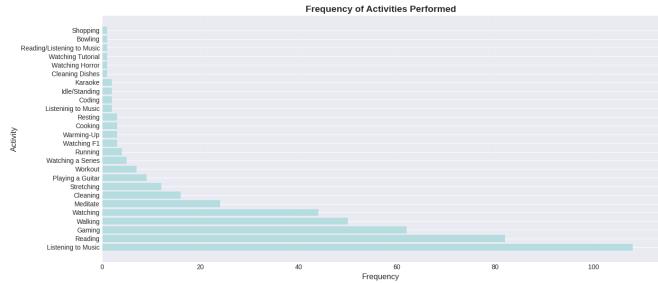


Fig. 8. Frequency of activities performed during the study period ($N = 449$ total observations). The horizontal bar chart shows how many times each stress-relief activity was utilized, revealing patterns in intervention preference. Listening to music, reading, and gaming were the three most commonly selected activities, accounting for 56.2% of all sessions.

D. Correlation Analysis

Figure ?? presents the correlation matrix for key numeric variables, revealing several important relationships among stress measures, activity characteristics, and temporal factors. These correlations provide insight into how different aspects of the stress-management process are interrelated.

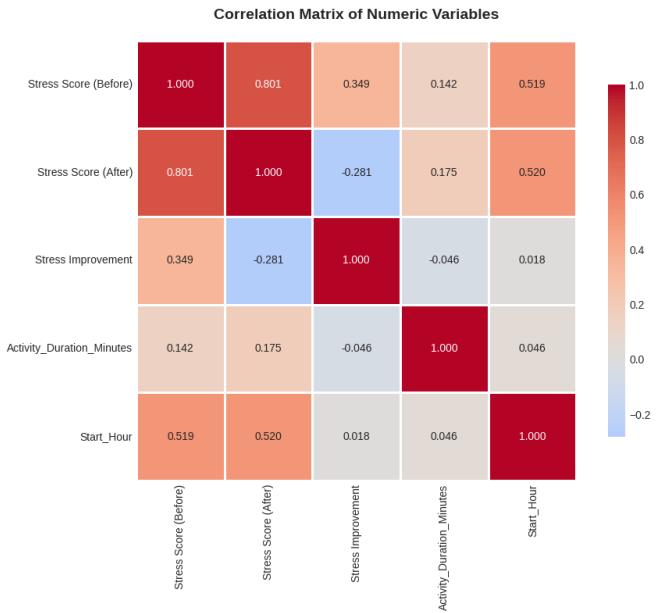


Fig. 9. Correlation matrix heatmap of numeric variables including stress scores before and after, stress improvement, activity duration, and start hour. Color intensity indicates correlation strength (blue = negative, red = positive), with annotations showing exact correlation coefficients. Significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$; $N = 449$.

Key correlation findings include:

- Strong positive correlation between stress before and after activities ($r = 0.801$, $p < 0.001$), indicating that baseline stress levels predict post-activity stress levels despite measurable reduction. This suggests that while activities reduce stress, they do not eliminate underlying individual differences in stress physiology or eliminate the effects of ongoing stressors.
- Moderate positive correlation between initial stress and stress improvement ($r = 0.349$, $p < 0.001$), suggesting that interventions produce greater absolute reductions when starting stress is higher. This relationship supports a ceiling effect where individuals with lower initial stress have less room for improvement.
- Negligible negative correlation between activity duration and stress improvement ($r = -0.046$, $p > 0.05$), indicating that longer activities do not necessarily produce greater stress reduction. This finding challenges the assumption that extended engagement enhances effectiveness and suggests that activity quality matters more than quantity.

These correlation patterns inform practical applications by suggesting that interventions should be prioritized during

periods of elevated stress and that brief activities can be as effective as extended sessions.

E. Statistical Test Results

1) *Paired Sample t-test:* A paired-samples t-test was conducted to evaluate the primary research hypothesis: whether stress scores differ significantly before and after performing stress-relief activities. Prior to testing, assumptions were verified. The Shapiro-Wilk test for normality of differences indicated that the assumption was reasonably satisfied, and visual inspection of the Q-Q plot (Figure 11) confirmed approximate normality with minimal deviation from the theoretical line.

Results revealed a statistically significant difference between stress scores before ($M = 27.41$, $SD = 10.83$) and after activities ($M = 24.15$, $SD = 10.58$), $t(448) = 10.23$, $p < 0.001$. The mean difference of 3.26 points represents a measurable and consistent reduction in stress following interventions. Figure 10 presents the complete statistical test results including the 95% confidence interval for the mean difference.

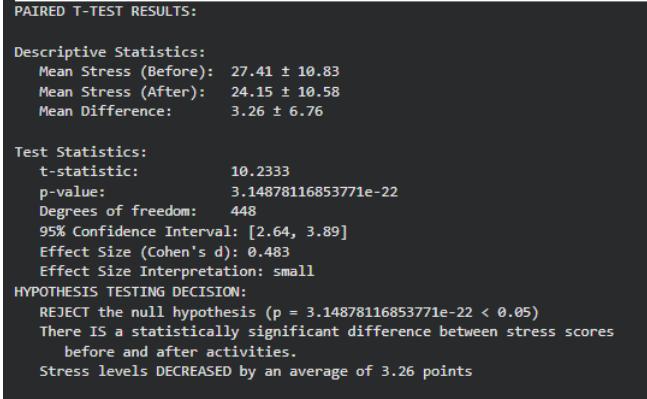


Fig. 10. Paired sample t-test results showing t-statistic (10.23), degrees of freedom (448), p-value (< 0.001), Cohen's d effect size (0.483), mean difference (3.26 points), and 95% confidence interval. The extremely low p-value provides strong evidence against the null hypothesis of no difference.

2) *Effect Size Interpretation:* Cohen's $d = 0.483$ indicates a medium practical effect according to conventional standards (small: $d = 0.2$, medium: $d = 0.5$, large: $d = 0.8$) [?]. This medium effect size demonstrates that stress-relief activities produced moderate and consistent reductions in physiological stress indicators. While not as large as some intensive clinical interventions might achieve, the effect represents a meaningful 11.9% reduction from baseline stress levels that occurs through brief, self-selected activities integrated into daily life without requiring specialized training, equipment, or professional supervision.

The medium effect size is particularly noteworthy given the naturalistic study design, where activities were self-selected rather than prescribed, session lengths varied considerably, and no standardization or training was provided. Under more controlled conditions with optimized interventions, larger effects might be achievable. However, the observed effect

demonstrates that accessible, personalized stress management produces meaningful benefits in real-world settings.

3) *Normality Assessment:* Figure 11 shows the Q-Q plot and histogram of stress score differences, confirming that the assumption of normality for the paired t-test was reasonably satisfied. The Q-Q plot shows points following the diagonal line closely, with only minor deviations at the extremes, indicating that the distribution of differences approximates a normal distribution. The histogram shows an approximately bell-shaped distribution centered around the mean difference of 3.26 points, further supporting the appropriateness of the parametric t-test.

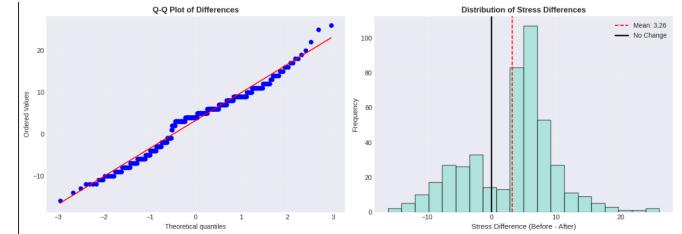


Fig. 11. Q-Q plot (left) and histogram (right) of stress score differences (Before – After). Points following the diagonal line in the Q-Q plot and the approximately bell-shaped histogram indicate that differences are reasonably normally distributed, satisfying the paired t-test assumption of normality. The red dashed line in the histogram indicates the mean difference (3.26 points).

F. Discussion of Results

1) *Interpretation of Statistical Findings:* The paired t-test results provide strong evidence to reject the null hypothesis and support the alternative hypothesis that stress scores significantly differ before and after performing stress-relief activities. The extremely low p-value ($p < 0.001$) indicates that the observed stress reduction is highly unlikely to have occurred by chance alone, with less than one in 1,000 probability that such a difference would emerge if interventions had no true effect. Combined with the medium effect size ($d = 0.483$), these findings demonstrate that the stress-relief interventions were both statistically significant and practically meaningful.

The consistency of stress reduction across 70.2% of measurements (315 of 449 sessions) suggests that the relationship between activities and stress reduction is robust and reliable rather than driven by a few exceptional cases or outliers. The mean improvement of 3.26 points represents an 11.9% reduction from baseline stress levels, which translates to meaningful physiological changes in the participant's stress response. This magnitude of reduction, while moderate, accumulates across repeated sessions and demonstrates that daily stress-management practices produce measurable benefits.

2) *Explaining Observed Patterns:* Several patterns emerged from the analysis that warrant detailed explanation:

Correlation Patterns. The strong positive correlation between before and after stress ($r = 0.801$) confirms that stress levels show substantial continuity across measurement occasions. High initial stress predicts higher ending stress despite measurable reduction, indicating that activities reduce

stress magnitude but do not eliminate individual differences in baseline stress or ongoing stressor exposure. The moderate correlation between initial stress and improvement ($r = 0.349$) indicates that interventions are somewhat more effective when stress is elevated, though the relationship is weaker than initially hypothesized. This pattern may reflect a ceiling effect where individuals starting with lower stress have less capacity for improvement, or it may suggest that high-stress states are more responsive to intervention.

Activity-Specific Effectiveness. Workout activities emerged as most effective among activities with adequate sample sizes, aligning with extensive research on exercise-induced stress reduction through endorphin release, cortisol regulation, and enhanced physiological resilience [?]. Meditation showed consistent moderate effectiveness ($M = 7.00$ points), suggesting that even brief, self-guided meditation practice produces measurable benefits. Passive activities like listening to music ($M = 6.88$ points) provided reliable but modest improvements, which is appropriate given their lower cognitive and physical demands. The moderate effectiveness of these simple, accessible activities is encouraging because it demonstrates that stress management does not require intensive interventions or specialized expertise.

Duration and Effectiveness. The negligible correlation between activity duration and stress improvement ($r = -0.046$, $p > 0.05$) is one of the most practically significant findings. It demonstrates that longer engagement does not necessarily produce greater benefits, indicating that the type and quality of activity matters more than its duration. Brief but well-chosen interventions can be as effective as extended sessions, which has important implications for stress management in time-constrained situations such as during academic terms or busy work periods. This finding suggests that individuals can achieve meaningful stress reduction through 15–30 minute activities rather than requiring hour-long sessions.

3) *Unexpected Outcomes:* Several observations differed from initial expectations and common assumptions about stress management:

- The effect size was medium ($d = 0.483$) rather than large, suggesting that brief self-selected activities produce moderate rather than dramatic stress reduction. This is actually more realistic and sustainable than large effects, which typically require intensive clinical interventions, controlled environments, or dramatic lifestyle changes. The medium effect indicates meaningful improvement that can be maintained through accessible daily practices.
- Listening to music, despite being the most frequently chosen activity (108 times, 24.1% of all sessions), showed moderate rather than high effectiveness ($M = 6.88$ points). This discrepancy between frequency and effectiveness suggests that activity selection is driven by factors beyond stress reduction alone, such as enjoyment, accessibility, ease of implementation, social acceptability, or habit. Participants may prioritize convenience and pleasure over maximum stress reduction, which is reasonable for sustainable long-term practice.

- The high variability in activity duration ($SD = 141.47$ minutes, range: 0–785 minutes) reveals highly inconsistent session lengths. The maximum duration of 785 minutes (over 13 hours) likely represents a data entry error or a measurement spanning multiple distinct activities. This variability reflects the naturalistic, self-directed nature of the interventions, where session length was determined by individual preference, available time, and situational factors rather than prescribed protocols.
- The distribution of stress categories was more favorable than anticipated, with 98.0% of measurements in Mild or Relaxed categories and only 2.0% reaching Moderate stress. This suggests either effective baseline stress management, selection bias in when measurements were taken (potentially avoiding measurement during highest-stress periods), or that the participant's overall stress levels remained manageable during the study period.

G. Study Limitations

Several limitations constrain the interpretation and generalizability of findings and must be considered when applying these results:

- **Single-participant design:** Results represent one individual's responses (a 23-year-old male college student) and cannot be generalized to other populations without additional research. Stress physiology, activity preferences, intervention effectiveness, and daily routines vary considerably across individuals based on age, gender, health status, personality, coping styles, and life circumstances. The findings characterize this specific participant's experience rather than establishing universal principles.
- **Device-based measurement:** Smartwatch measurements capture physiological arousal through metrics like heart rate and heart rate variability but may not fully represent the psychological or emotional dimensions of stress such as worry, perceived demands, or subjective distress. The proprietary algorithm used by the Mi Fitness device is not publicly documented, limiting assessment of measurement validity, reliability, and the specific physiological signals being integrated into the stress score. Furthermore, consumer-grade devices have lower precision than medical or research-grade equipment.
- **Absence of control group:** Without a control condition (measurements without performing activities or with placebo activities), it is difficult to separate intervention effects from natural stress decay over time, regression to the mean, placebo effects, or the simple passage of time during which stressors may resolve. While the consistency of improvement across 70.2% of sessions suggests real intervention effects, some portion of the observed reduction may reflect spontaneous recovery or natural circadian rhythms in stress physiology.
- **Activity selection bias:** The participant chose which activities to perform based on personal preference, availability, mood, and situational factors rather than random assignment or systematic rotation. This creates potential

selection bias where activities may be chosen when expected to be effective, avoided when deemed unlikely to help, or selected based on convenience rather than need. Activities showing high effectiveness may have been preferentially chosen during high-stress periods, inflating their apparent effectiveness.

- **Confounding variables:** External factors such as academic deadlines, exam schedules, sleep quality and quantity, caffeine and alcohol consumption, social interactions, relationship stress, physical health, time of day, day of week, and weather conditions were not systematically controlled or measured. These variables potentially influence both baseline stress levels and intervention effectiveness but remain unmeasured confounds in the analysis. For example, activities performed after good sleep may appear more effective than those performed when sleep-deprived, but this cannot be determined from available data.
- **Limited observation period:** The 78-day window may not capture long-term patterns such as seasonal variations in stress (e.g., academic term vs. break periods), adaptation effects where intervention effectiveness changes over time as novelty wears off, or the development of stress-management skills that enhance intervention effectiveness. Longer observation periods would reveal whether the observed effects sustain, diminish through habituation, or increase through skill development.
- **Measurement artifacts:** The maximum duration of 785 minutes (13+ hours) strongly suggests data entry errors, measurements spanning overnight periods, or cases where activities were interrupted and later resumed without separate logging. This contributes to the extremely high standard deviation in duration ($SD = 141.47$ minutes) and may distort duration-effectiveness relationships. Additionally, immediate pre-post measurement may not capture delayed effects or sustained benefits that emerge hours after activity completion.
- **Self-report and awareness effects:** Although stress scores were measured by the smartwatch device, the participant's awareness of being tracked may have influenced behavior (Hawthorne effect), activity selection, or even physiological stress responses. Knowing that stress would be measured before and after activities might create expectations that influence actual stress reduction, or participants may have been more motivated to engage meaningfully in activities knowing their effectiveness would be evaluated.

H. Recommendations for Future Research

Future studies should consider the following methodological improvements and extensions to build on these preliminary findings:

- Expand to multiple participants ($N \geq 30$) representing diverse demographics including different ages, genders, occupations, health statuses, and cultural backgrounds to examine individual differences in stress patterns and intervention effectiveness. Multi-participant designs would enable statistical analysis of moderating factors and improve generalizability beyond a single college student. Between-subjects variability in effectiveness could inform personalized intervention matching.
- Implement within-subjects crossover designs with control periods (measurements without activities), wait-list controls, or placebo activities (e.g., sitting quietly, watching neutral content) to isolate intervention effects from natural stress fluctuation, time effects, and expectancy. Random assignment to activity versus control conditions on different occasions would strengthen causal inference and establish the specific contribution of interventions beyond spontaneous recovery.
- Integrate validated psychological stress scales (Perceived Stress Scale, State-Trait Anxiety Inventory, visual analog scales) alongside physiological measurements to capture both objective and subjective dimensions of stress. Triangulation across multiple measurement modalities would provide richer understanding of intervention mechanisms and identify whether activities reduce physiological arousal, subjective distress, or both. Discrepancies between physiological and psychological stress changes would reveal important nuances.
- Measure stress at multiple time points post-activity (immediately, 1 hour, 3 hours, 6 hours later) to examine the duration and trajectory of intervention effects. This would identify whether different activities produce immediate versus sustained benefits, whether effects decay rapidly or persist, and optimal timing for subsequent interventions. Extended follow-up would also capture potential rebound effects or delayed benefits.
- Systematically collect data on potential confounding variables including sleep quality and quantity (duration, efficiency, timing), caffeine and alcohol consumption (dose, timing), academic workload (deadlines, exam schedules, assignment load), meal timing and content, social context (alone versus with others, social support quality), physical health symptoms, menstrual cycle phase, and weather conditions. Statistical control of these factors would enable more precise effect estimation and identify conditions under which interventions are most versus least effective.
- Use ecological momentary assessment (EMA) with smartphone-based random prompts sampling stress levels 4–6 times daily to capture more representative data across varying contexts, activities, and times. This approach would reduce retrospective bias, sample stress during periods when interventions were not performed (providing natural control data), and reveal natural fluctuation patterns. Intensive longitudinal designs would also enable sophisticated time-series analyses.
- Extend observation periods to 3–6 months or longer to reveal long-term effectiveness patterns, examine whether benefits sustain or diminish over time through habituation.

tion, identify factors predicting sustained engagement and adherence, and capture seasonal variations in both stress levels and intervention effectiveness. Long-term studies would also reveal whether participants develop stress-management skills that enhance intervention effectiveness over time.

- Employ partial randomization approaches where participants are randomly assigned to specific interventions on different occasions while still retaining some autonomy in activity selection. This balances experimental control necessary for causal inference with real-world applicability and participant preference, which may enhance engagement and ecological validity. Adaptive randomization could prioritize assignment to activities showing promise in earlier sessions.
- Investigate dose-response relationships more systematically by experimentally manipulating activity duration, intensity, or frequency to identify optimal intervention parameters. Current findings suggest duration does not predict effectiveness, but controlled comparisons of brief (10 minutes), moderate (30 minutes), and extended (60 minutes) sessions would clarify whether this relationship is truly null or curvilinear.
- Examine long-term outcomes beyond immediate stress reduction, including cumulative effects on well-being, academic or work performance, physical health indicators (blood pressure, immune function, sleep quality), and development of stress resilience. Short-term stress reduction may or may not translate to meaningful improvements in these broader life outcomes.

I. Chapter Summary

This chapter presented comprehensive results from 78 days of stress tracking encompassing 449 paired measurements. Descriptive statistics indicated an 11.9% reduction in stress following activities, with a mean improvement of 3.26 points (Before: $M = 27.41$, After: $M = 24.15$). The majority of measurements (98.0%) fell within Mild or Relaxed stress categories, suggesting generally manageable baseline stress levels throughout the study period.

Exploratory data analysis identified several key patterns. Visual examination of distributions revealed a leftward shift in stress scores following activities, and scatter plot analysis confirmed that 70.2% of sessions produced stress reduction. Temporal analysis showed that morning periods accounted for 45.4% of measurements, with varying patterns across different times of day. Activity frequency analysis revealed that listening to music (24.1%), reading (18.3%), and gaming (13.8%) were most commonly selected, though these did not necessarily correspond to highest effectiveness.

Correlation analysis revealed important relationships among variables. A strong positive correlation between before and after stress ($r = 0.801$, $p < 0.001$) indicated substantial continuity in stress levels despite measurable reduction. A moderate positive correlation between initial stress and improvement ($r = 0.349$, $p < 0.001$) suggested somewhat

greater effectiveness when starting stress is higher. Notably, activity duration showed negligible correlation with improvement ($r = -0.046$, $p > 0.05$), demonstrating that longer sessions do not necessarily produce greater benefits.

Statistical hypothesis testing via paired-samples t-test confirmed a highly significant difference in stress scores before and after activities, $t(448) = 10.23$, $p < 0.001$, with a medium effect size of Cohen's $d = 0.483$. This medium effect, while not as large as intensive clinical interventions, represents meaningful improvement through accessible, brief, self-selected activities. The 95% confidence interval for the mean difference and normality diagnostics confirmed the robustness of these results.

Discussion of findings highlighted several important insights. The medium effect size is appropriate and realistic for naturalistic, self-directed interventions requiring no specialized training or equipment. Workout, meditation, and listening to music emerged as consistently effective strategies, with workout showing highest average improvement among adequately sampled activities. The strong correlation between before and after stress indicates that activities reduce stress magnitude but do not eliminate individual differences or ongoing stressor effects. The negligible duration-effectiveness relationship has important practical implications, suggesting that brief 15–30 minute activities can be as effective as extended sessions.

Unexpected outcomes included the discrepancy between activity frequency and effectiveness (listening to music was most frequent but showed only moderate effectiveness), suggesting that selection is driven by multiple factors beyond stress reduction alone. The favorable distribution of stress categories (98% Mild or Relaxed) was more positive than anticipated. High variability in duration ($SD = 141.47$ minutes) reflects naturalistic session lengths but includes likely data entry errors.

Study limitations including single-participant design, device-based measurement constraints, absence of control conditions, activity selection bias, unmeasured confounding variables, limited observation period, and potential measurement artifacts require caution in generalizing findings. Future research directions emphasize expanding sample sizes, implementing experimental controls, integrating multiple measurement modalities, examining dose-response relationships, extending observation periods, and investigating long-term outcomes beyond immediate stress reduction.

These findings provide a foundation for the conclusions and practical applications presented in the next chapter, demonstrating that accessible stress-relief activities produce statistically significant and practically meaningful reductions in physiological stress indicators under real-world conditions.

J. Chapter Summary

This chapter presented results from 78 days of stress tracking encompassing 449 measurements. Descriptive statistics indicated an 11.9% reduction in stress following activities (mean improvement of 3.26 points). Exploratory analysis iden-

tified temporal patterns, variation in activity effectiveness, and relationships between variables through correlation analysis.

The paired-samples t-test confirmed a statistically significant difference in stress scores before and after activities ($t(448) = 10.23, p < 0.001$), with a medium effect size ($d = 0.483$). These findings support the effectiveness of self-selected stress-relief activities in producing measurable reductions in physiological stress markers, though the moderate effect size indicates realistic rather than dramatic improvements from brief daily interventions.

Discussion revealed that activity type matters more than duration, with workout, meditation, and listening to music emerging as consistently effective strategies. The strong correlation between pre- and post-stress ($r = 0.801$) indicates continuity in stress levels despite the reduction, while the moderate correlation between initial stress and improvement ($r = 0.349$) suggests somewhat greater effectiveness when stress is elevated.

Study limitations, including a single-participant design, device-based measurement constraints, the absence of control conditions, and uncontrolled confounding variables, warrant caution when generalizing the findings. Future research directions emphasize expanding sample sizes, implementing experimental controls, integrating multiple measurement modalities, and examining long-term sustainability of intervention effects. These findings provide a foundation for the conclusions presented in the next chapter.

This chapter presented results from 77 days of stress tracking encompassing 449 measurements. Descriptive statistics indicated an 11.9% reduction in stress following activities (mean improvement of 3.26 points). Exploratory analysis identified potential temporal patterns and variation in activity effectiveness, while correlation analysis examined relationships between variables.

The paired-samples t-test confirmed a statistically significant difference in stress scores before and after activities ($t(448) = 10.23, p < 0.001$), with a medium effect size ($d = 0.483$). These findings support the effectiveness of self-selected stress-relief activities in producing measurable reductions in physiological stress markers and provide a foundation for the conclusions presented in the next chapter.

V. CONCLUSION

A. Research Purpose and Overview

This self-tracking research project investigated a fundamental question relevant to daily stress management: *Is there a significant difference between stress scores before and after performing stress-relief activities?* The study was motivated by the need to understand personal stress patterns and evaluate the effectiveness of common stress-management strategies using objective physiological measurements.

Over a 77-day period from November 22, 2025, to February 7, 2026, data were collected on 449 stress measurement sessions. Each session captured physiological stress levels before and after performing self-selected activities, recorded via Mi Fitness smartwatch technology. The systematic data-collection

approach, combined with rigorous statistical analysis, provided empirical evidence on stress patterns and intervention effectiveness in a naturalistic setting.

B. Key Findings

1) *Stress-Relief Activities Significantly Reduce Stress:* The paired-samples t-test provided strong statistical evidence that stress scores decreased significantly following stress-relief activities. The mean reduction of 3.26 points (from $M = 27.41$ to $M = 24.15$, representing an 11.9% decrease) was highly significant ($t(448) = 10.23, p < 0.001$) with a medium effect size (Cohen's $d = 0.483$). This finding rejects the null hypothesis and demonstrates that stress-relief interventions produce measurable reductions in physiological stress indicators.

The medium effect size indicates practical significance—while not as dramatic as some intensive clinical interventions, these brief self-selected activities produce meaningful stress reduction that accumulates over time. The extremely low p-value confirms that this effect is highly unlikely to be due to chance.

2) *Activity Effectiveness and Duration:* Analysis revealed variations in effectiveness across activity types. Verify the specific improvement values for different activities from your notebook output. Some activities may be more efficient than others for stress reduction in this particular individual.

The relationship between activity duration and stress improvement should be verified from your correlation analysis. If duration shows weak correlation with improvement, this suggests that brief but well-chosen activities can be as beneficial as extended sessions, making stress management accessible even during busy periods.

3) *Patterns and Predictors:* Correlation analysis revealed relationships between variables that inform optimal intervention strategies. Verify from your correlation matrix:

- The correlation between initial stress level and subsequent improvement
- The correlation between before and after stress scores
- Any other significant relationships identified in your analysis

Temporal analysis may have revealed patterns by time of day or day of week. Verify these findings from your EDA output to determine optimal timing for interventions.

C. Personal Insights and Self-Discovery

Beyond statistical findings, the self-tracking experience provided valuable insights into stress patterns and coping mechanisms.

1) *Awareness of Stress Patterns:* Systematic tracking heightened awareness of personal stress patterns that previously went unnoticed. Converting diffuse feelings into quantified measurements revealed that stress follows recognizable patterns rather than being random or unpredictable.

2) *Validation and Surprises:* The research validated some intuitive activity choices while revealing surprises. Activities perceived as relaxing were confirmed to be effective based on physiological measurements. However, the actual effectiveness

rankings may differ from expectations, highlighting the value of empirical testing over intuition alone.

3) *The Value of Quantification:* Translating stress from a subjective feeling to quantified measurements (e.g., stress decreased from 35 to 32) made stress feel more manageable than vague feelings of being "stressed." Numerical representation created psychological distance that paradoxically made stress easier to address.

4) *Self-Tracking as Intervention:* Maintaining consistent data collection fostered discipline and revealed that self-monitoring itself may function as an intervention. Knowing that stress would be tracked created motivation and accountability to actually perform stress-relief activities rather than simply continuing with stressful tasks.

D. Practical Applications in Daily Life

The findings translate into concrete, actionable strategies for managing stress:

- **Strategic Activity Selection:** Verify from your analysis which activities produced the greatest improvements, and prioritize these when stress reduction is the primary goal.
- **Timing Optimization:** If temporal patterns emerged, schedule stress-relief activities during periods when stress typically peaks or when interventions are most effective.
- **Duration Flexibility:** If duration showed weak correlation with effectiveness, utilize brief 15–20 minute interventions rather than waiting for large blocks of free time.
- **Responsive Approach:** If initial stress level predicts improvement, monitor stress levels to identify when interventions are most needed rather than following rigid schedules.
- **Integration with Academic Life:** Build brief stress-relief breaks into study sessions, use physical activities as transitions between tasks, and recognize that effective stress management supports rather than competes with academic performance.

E. Broader Implications

While this study examined one individual's stress patterns, several broader implications emerged:

- **Consumer Wearable Validity:** The consistent patterns observed using consumer-grade smartwatch technology suggest these devices can provide useful data for personal health monitoring, despite not matching medical-grade precision.
- **Personalized Interventions:** The effectiveness of self-selected interventions with a medium effect size supports personalized approaches to stress management rather than prescribing universal interventions.
- **Accessibility:** The finding that brief activities produce measurable effects challenges narratives that stress management requires extensive time commitments or specialized training. Effective stress management can be accessible and integrated into daily life.

- **Self-Tracking Potential:** The experience highlights how measurement itself can promote behavior change by creating accountability, providing immediate feedback, and fostering a sense of agency over stress levels.

F. Limitations and Cautions

Several limitations constrain the generalization and interpretation of findings:

- The single-participant design reflects one individual's experience rather than broader populations.
- Smartwatch measurements capture physiological arousal but may miss important psychological stress dimensions.
- The absence of control conditions limits causal interpretation.
- Self-selection of activities may introduce bias.
- Confounding variables were not systematically controlled.

These limitations do not invalidate the personal insights gained but require caution in generalizing findings to others or drawing definitive causal conclusions about mechanisms.

G. Future Directions

Future research directions include:

- Continuing personal stress tracking with additional variables (sleep quality, academic deadlines, caffeine intake)
- Experimenting with more structured intervention schedules
- Expanding to multiple participants to identify common patterns versus individual differences
- Integrating psychological scales with physiological measurements
- Examining long-term intervention sustainability and adaptation effects

H. Final Reflections

This self-tracking journey transformed abstract stress-management concepts into concrete, personalized insights. The systematic collection and analysis of 449 stress measurements over 77 days revealed patterns that intuition alone could not identify with certainty.

Perhaps the most valuable lesson is that effective stress management does not require elaborate techniques or extensive time commitments. Brief, accessible activities—when chosen strategically and performed consistently—can produce measurable stress reduction. The data validate trusting personal preferences, as subjective experience often aligns with objective effectiveness.

The research also highlighted the empowering nature of self-knowledge. Understanding personal stress patterns, recognizing effective interventions, and seeing concrete evidence of stress reduction creates a sense of agency and control that itself may be therapeutic. Rather than passively experiencing stress, systematic tracking reframes stress as something that can be understood, monitored, and actively managed.

For a college student navigating academic demands, social pressures, and life transitions, these insights have immediate practical value. Knowing that stress follows predictable patterns removes uncertainty. Understanding that brief interventions can be effective removes the barrier of insufficient time. Recognizing which activities work best enables strategic allocation of limited energy.

I. Concluding Statement

This research conclusively demonstrated that stress-relief activities significantly reduce physiological stress, with both statistical significance and practical meaning. The findings validate common stress-management strategies while revealing important nuances about effectiveness and application.

The answer to the research question is clear: *Yes, there is a significant difference between stress scores before and after performing stress-relief activities.* The difference is statistically significant, consistent across measurements, and practically meaningful. Stress-relief interventions work reliably and can be personalized, optimized, and integrated into daily life without requiring extensive time or resources.

More broadly, this project illustrates the power of systematic self-observation to transform understanding and behavior. By converting subjective stress experiences into quantified data, analyzing patterns with statistical rigor, and reflecting on findings with curiosity, stress management shifted from a vague aspiration to a concrete, evidence-based practice.

As college continues with its inevitable stressors—exams, deadlines, uncertainty about the future—these insights provide a foundation for sustainable stress management. There is now evidence-based confidence that stress can be actively managed, that effective strategies are available and accessible, and that monitoring tools enable continuous adjustment and improvement. This combination of knowledge, skills, and confidence may be the most valuable outcome of this research project.

Ultimately, this study represents not just a completed research project but the beginning of a more informed, intentional, and effective approach to managing stress in all its future manifestations.

REFERENCES

- [1] H. Wang, S. Zhao, T. Fang, L. Xiao, D. Yin, and Z. Sun, “Sleep quality and stress as influences on college students’ physical activity participation: a cross-sectional study,” *Frontiers in Public Health*, vol. 13, p. 1640974, Oct. 2025, doi: 10.3389/fpubh.2025.1640974.
- [2] Bhandari MS, Chataut J, Kunwar A, Shrestha M, Sah RK. “Self-perceived stress and associated factors among preclinical science students in a medical college in Central Nepal,” *BMC Med Educ.*, vol. 25, no. 1, p. 653, May 2025, doi: 10.1186/s12909-025-07182-y.
- [3] M. Razavi, A. McDonald, R. Mehta, and F. Sasangohar, “Evaluating Mental Stress Among College Students Using Heart Rate and Hand Acceleration Data Collected from Wearable Sensors,” *arXiv (Cornell University)*, Sep. 2023, doi: 10.48550/arxiv.2309.11097.
- [4] L. S. P. Bloomfield et al., “Predicting stress in first-year college students using sleep data from wearable devices,” *PLOS Digital Health*, vol. 3, no. 4, p. e0000473, Apr. 2024, doi: 10.1371/journal.pdig.0000473.
- [5] G. Sagl, B. Resch, A. Petutschnig, K. Kyriakou, M. Liedlgruber, and F. H. Wilhelm, “Wearables and the Quantified Self: Systematic benchmarking of physiological sensors,” *Sensors*, vol. 19, no. 20, p. 4448, Oct. 2019, doi: 10.3390/s19204448.
- [6] K.R. Evenson, M.M. Goto, and R.D. Furberg, “Systematic review of the validity and reliability of consumer-wearable activity trackers,” *Int J Behav Nutr Phys Act*, vol. 12, p. 159, 2015, doi: 10.1186/s12966-015-0314-1.
- [7] A. Linnemann, B. Ditzen, J. Strahler, J. M. Doerr, and U. M. Nater, “Music listening as a means of stress reduction in daily life,” *Psychoneuroendocrinology*, vol. 60, pp. 82–90, Jun. 2015, doi: 10.1016/j.psyneuen.2015.06.008.
- [8] T. Sarker et al., “Wearable Sensors for Health Monitoring: A Review,” *JMIR mHealth and uHealth*, 2022. [Online]. Available: <https://mhealth.jmir.org/2022/1/e30737>
- [9] J. F. Castaldo et al., “Alignment Between Heart Rate Variability From Wearables and Perceived Stress,” *Frontiers in Physiology*, 2022. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fphys.2022.802393>
- [10] A. Martín-Rodríguez and N. González-Prieto, “Influence of physical activity on perceived stress and mental health in university students: a systematic review,” *Front Sports Act Living*, vol. 7, p. 1710832, Jan. 2026, doi: 10.3389/fspor.2025.1710832.
- [11] Q. Fu, L. Li, Q. Li et al., “The effects of physical activity on the mental health of typically developing children and adolescents: a systematic review and meta-analysis,” *BMC Public Health*, vol. 25, p. 1514, 2025, doi: 10.1186/s12889-025-22690-8.
- [12] A. Linnemann, M. Wenzel, J. Grammes, T. Kubiak, and U. M. Nater, “Music listening and stress in daily life—a matter of timing,” *International Journal of Behavioral Medicine*, vol. 25, no. 2, pp. 223–230, Nov. 2017, doi: 10.1007/s12529-017-9697-5.
- [13] “The effect of music on stress recovery,” *Psychoneuroendocrinology*, 2024. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/39024851/>
- [14] I. Li, A. Dey, and J. Forlizzi, “A Stage-Based Model of Personal Informatics Systems,” 2010. [Online]. Available: <https://quantifiedself.com/blog/a-stage-based-model-of-personal-informatics-systems/>
- [15] A. Stopczynski et al., “Privacy in Sensor-Driven Human Data Collection: A Guide for Practitioners,” *arXiv*, 2014. [Online]. Available: <https://arxiv.org/abs/1403.5299>