2024

Depression Severity Analysis

CS-C4100 - DIGITAL HEALTH AND HUMAN BEHAVIOR

CONTENTS

1. Introduction 2

1. Problem 2

1.2 Outline 2

2. Data Description and Preprocessing 4

2.1 Dataset Overview 4

2.2 Data Cleaning 6

3. Data Analysis 7

3.1 Exploring the Relationship Between Steps and PHQ-9 Scores 7

3.2 Relationship Between Steps and PHQ-9 Severity 7

3.3 Relationship Between Sleep Metrics and PHQ-9 Scores 8

3.4 Sleep Efficiency and PHQ-9 Severity 10

3.5 Analysis of Weekday and Weekend Sleep Patterns by PHQ-9 Severity 12

4 Methodology 15

4.1 Why I used Clustering and K-means 15

4.2 Determining the Optimal Number of Clusters 16

4.3 Cluster Characteristics and Observations 17

4.4 Demographic Analysis by Cluster 18

4.4.2 Life Stress and Alcohol Reduction by Cluster 19

4.4.3 Conclusion 19

4.5 Correlation Between Clusters and PHQ-9 Scores 20

5. Results 23

6. Discussion 24

7. Conclusion 26

8. References: 27

1. Introduction

1. Problem

Depression is one of the most common mental health conditions, affecting millions of people worldwide. It can have a serious impact on both quality of life and productivity. According to the World Health Organization (WHO), around 280 million people globally suffer from depression (WHO, 2023). People experience depression differently based on their personal circumstances, environments, and biological factors. The reasons for depression can vary widely, such as stress, lack of sleep, physical inactivity, or other health conditions.

Research shows that depression is a major cause of disability and health problems worldwide, affecting about 15% of people during their lifetime (Wang et al., 2021). Tracking behavior and health data is important to better understand mental health. Depression is especially common in older adults, where it greatly increases the risk of disability and death (Zenebe et al., 2021). However, almost half of depression cases go undiagnosed, particularly in older people, making it crucial to find better ways to address this issue.

In this project, I focus on analyzing the relationship between health behaviors tracked by wearable devices like steps taken, sleep quality, and moderate-to-vigorous physical activity and the severity of depression. Depression severity is measured using PHQ-9 scores, a commonly used self-report questionnaire.

The study uses clustering techniques and statistical analyses to identify patterns in behavior and how they relate to mental health outcomes. These findings provide useful insights for creating targeted interventions to improve sleep, increase physical activity, and support better mental health outcomes.

This project aims to explore how technology, like wearable devices, can be used to better understand mental health and develop strategies to reduce depression severity. By identifying patterns in health behavior, we can take steps toward more personalized approaches to treating and managing depression.

1.2 Outline

In Section 1, I introduce the topic of depression, highlighting its global prevalence and the motivation behind leveraging wearable data to better understand its behavioral correlates. I also outline the objectives and scope of the project.

In Section 2, I present the PSYCHE-D dataset, detailing its structure, key variables, and the preprocessing steps undertaken to ensure data quality for analysis.

In Section 3, I describe the exploratory data analysis conducted to investigate the relationships between physical activity, sleep patterns, and depression severity, as measured by PHQ-9 scores.

In Section 4, I explain the clustering methodology, focusing on the use of K-Means to segment participants into meaningful behavioral groups and the rationale behind selecting the optimal number of clusters.

In Section 5, I present the results of the clustering analysis, highlighting the characteristics of each cluster and the key insights regarding their behavioral and demographic profiles.

In Section 6, I discuss the implications of our findings for mental health research and interventions, addressing the study's limitations and proposing directions for future research.

In Section 7, I summarize the conclusions drawn from the project, emphasizing the potential of wearable data and clustering techniques to support personalized mental health care.

2. DATA DESCRIPTION AND PREPROCESSING

2.1 DATASET OVERVIEW

The dataset used in this project is the PSYCHE-D Dataset, which is publicly available on Zenodo (Makhmutova et al., 2024). It comprises data collected from 4,036 unique participants as part of the DiSCover Project, a one-year longitudinal study conducted between January 2018 and January 2020. The dataset was curated by researchers including Mariko Makhmutova and Raghu Kainkaryam.

The primary goal of the study was to monitor changes in depression severity over time and investigate the relationship between depression measured using PHQ-9 scores and lifestyle factors such as physical activity, sleep patterns, and medication changes. Data collection was facilitated through Fitbit wearable devices and self-reported surveys on socio-demographics, lifestyle behaviors, and mental health.

Key Dataset Components**:**

1. Wearable Data: Includes step counts and sleep metrics derived from Fitbit devices.
2. PHQ-9 Scores: Measures of depression severity, collected quarterly using the 9-item PHQ-9 questionnaire.
3. Lifestyle Surveys: Monthly updates on lifestyle and medication changes (e.g., increased physical activity or reduced alcohol consumption).
4. Demographic Data: Static participant information, including age, gender, and comorbidities.
5. MVPA (Moderate-to-Vigorous Physical Activity): Represents physical activity requiring moderate to high effort, characterized by activities that raise the heart rate significantly.

The dataset consists of 35,694 rows with 154 features . Notably, approximately 78 features have detailed descriptions provided by the dataset documentation.

The features with the description:

Feature name	Description
sex	Sex
race_white	Race, white
race_black	Race, black
race_hispanic	Race, hispanic
race_asian	Race, asian
race_other	Race, other
birthyear	Birth year
educ	Education level
height	Height, in inches
weight	Weight, in lbs
bmi	BMI
pregnant	Pregnant, at baseline
birth	Given birth, past year
trauma	Experienced trauma resulting in injury (e.g. accident, fall), past year
insurance	Has health insurance
money	Sufficient money, past month
money_assistance	Money assistance from government, past month
household	Number of individuals in household
comorbid_cancer	Diagnosed: cancer
comorbid_diabetes_typ1	Diagnosed: type 1 diabetes
comorbid_diabetes_typ2	Diagnosed: type 2 diabetes
comorbid_gout	Diagnosed: gout
comorbid_migraines	Diagnosed: migraines
comorbid_ms	Diagnosed: multiple sclerosis
comorbid_osteoporosis	Diagnosed: osteoporosis
comorbid_neuropathic	Diagnosed: fibromyalgia, peripheral neuropathic pain, central neuropathic pain (sum)
comorbid_arthritis	Diagnosed: osteoarthritis, rheumatoid arthritis (sum)
num_migraine_days	Number of migraines days, per month, on average
meds_migraine	Takes daily prescription migraine medication, at baseline
med_start	Started a new medication, past month
med_stop	Stopped a medication, past month
med_dose	Changed dosage of at least one medication, past month
nonmed_start	Started new non-medication therapy, past month
nonmed_stop	Stopped non-medication therapy, past month
med_nonmed_dnu	Do not take medication/non-medication therapies
life_meditation	Started meditation/other relaxation techniques, past month
life_stress	Reduced stress-inducing activities, past month
life_activity_eating	Increased physical activity or improved eating habits, past month
life_red_stop_alcoh	Reduced or stopped alcohol consumption, past month
steps_awake_mean	Mean number of steps taken while awake
steps_mvpa_iqr	IQR of number of minutes with MVPA-range steps
steps_lpa_iqr	IQR of number of minutes with LPA-range steps
steps_awake_sum_iqr	IQR of number of steps taken while awake
steps__active_day_count_	Number of days with >10k steps, last 7 days
steps__sedentary_day_count_	Number of days with <5k steps, last 7 days
steps_mvpa_sum_recent	Number of MVPA-range step minutes, last 4 days
steps_lpa_sum_recent	Number of LPA-range step minutes, last 4 days
steps_rolling_6_median_recent	Median number of 6 minute rolling number of steps, last 4 days
steps_rolling_6_max_recent	Max number of 6 minute rolling number of steps, last 4 days
sleep_asleep_weekday_mean	Mean number of minutes asleep per night on weekdays
sleep_asleep_weekend_mean	Mean number of minutes asleep per night on weekends
sleep_in_bed_weekday_mean	Mean number of minutes in bed per night on weekdays
sleep_in_bed_weekend_mean	Mean number of minutes in bed per night on weekends
sleep_ratio_asleep_in_bed_weekday_mean	Mean ratio between number of minutes asleep / in bed on weekdays
sleep_ratio_asleep_in_bed_weekend_mean	Mean ratio between number of minutes asleep / in bed on weekends
sleep_in_bed_iqr	IQR of minutes in bed per night
sleep_asleep_iqr	IQR of minutes asleep per night
sleep_ratio_asleep_in_bed_iqr	IQR of ratio between number of minutes asleep / in bed
sleep_main_start_hour_adj_median	Median start hour of main sleep
sleep_main_start_hour_adj_iqr	IQR of start hour of main sleep
sleep_main_start_hour_adj_range	Max start hour of main sleep - min start hour of main sleep
sleep__hypersomnia_count_	Number of nights with >10 hours asleep, last 7 days
sleep__hyposomnia_count_	Number of nights with <5 hours asleep, last 7 days
sleep_asleep_mean_recent	Mean number of minutes asleep per night, last 4 days
sleep_in_bed_mean_recent	Mean number of minutes in bed per night, last 4 days
sleep_ratio_asleep_in_bed_mean_recent	Mean ratio between number of minutes asleep / in bed, last 4 days
phq9_cat_end	PHQ-9 category at the end of the 3 month period

phq9_cat_start		PHQ-9 category at the start of the 3 month period
phq9_score_end		PHQ-9 score at the end of the 3 month period
phq9_score_start		PHQ-9 score at the start of the 3 month period
phq9_cat_end		PHQ-9 category at the end of the 3 month period
phq9_cat_start		PHQ-9 category at the start of the 3 month period
phq9_score_end		PHQ-9 score at the end of the 3 month period
phq9_score_start		PHQ-9 score at the start of the 3 month period
phq9_cat_end		PHQ-9 category at the end of the 3 month period
phq9_cat_start		PHQ-9 category at the start of the 3 month period
phq9_score_end		PHQ-9 score at the end of the 3 month period
phq9_score_start		PHQ-9 score at the start of the 3 month period

2.2 DATA CLEANING

I focused on addressing missing values in critical columns to ensure data quality for analysis. I identified and handled missing values in key variables, including PHQ-9 scores phq9_score_start, phq9_score_end, wearable , steps_awake_mean, sleep_asleep_weekday_mean , and other lifestyle and survey data. Columns with high missingness were either removed or imputed based on their importance to the analysis.

Steps I Took:

1. PHQ-9 Scores: I filled missing values using forward-filling for sequential consistency or imputed them based on participant-level trends.
2. Wearable Data: I imputed missing values using mean or median values to maintain consistency and accuracy.
3. Cleaning Summary: After the cleaning process, I successfully addressed all missing values in critical columns, ensuring the dataset is complete and ready for further analysis.

By following this process, I ensured the dataset was consistent and prepared for in-depth analysis.

3. Data Analysis

3.1 Exploring the Relationship Between Steps and PHQ-9 Scores

I will explore the the relationship between step counts and PHQ-9 scores, I analyzed correlations and visualized trends. The scatter plot indicates a weak negative correlation, suggesting that higher step counts may be associated with slightly lower PHQ-9 scores. Both PHQ-9 scores at the start and end of the study exhibit similar trends. The figure 1 graph below highlights the relationship between step counts and PHQ-9 scores:

FIGURE 1

3.2 Relationship Between Steps and PHQ-9 Severity

I analyzed the connection between average step counts and PHQ-9 severity levels, which include Minimal, Mild, Moderate, Moderately Severe, and Severe, at both the beginning and end of the study. The aim of the analysis was to understand how physical activity changes with varying levels of depression severity.

To carry out the analysis, I categorized the PHQ-9 scores into severity levels based on standard cutoffs. I calculated the average step counts for each severity level at both the start and end of the study. The results are presented in Figure 2, which compares the step counts across the different severity levels, with separate bars for PHQ-9 scores at the start and at the end of the study.

The results showed that participants with minimal depression severity had the highest average step counts. In contrast, step counts progressively decreased as the severity of depression increased. This trend was consistent at both the beginning and the end of the study. Figure 2 highlights that participants in the severe depression category experienced a slight decline in step counts over time.

The findings suggest a possible relationship between physical activity levels and depression severity. Reduced step counts may indicate higher levels of depression.

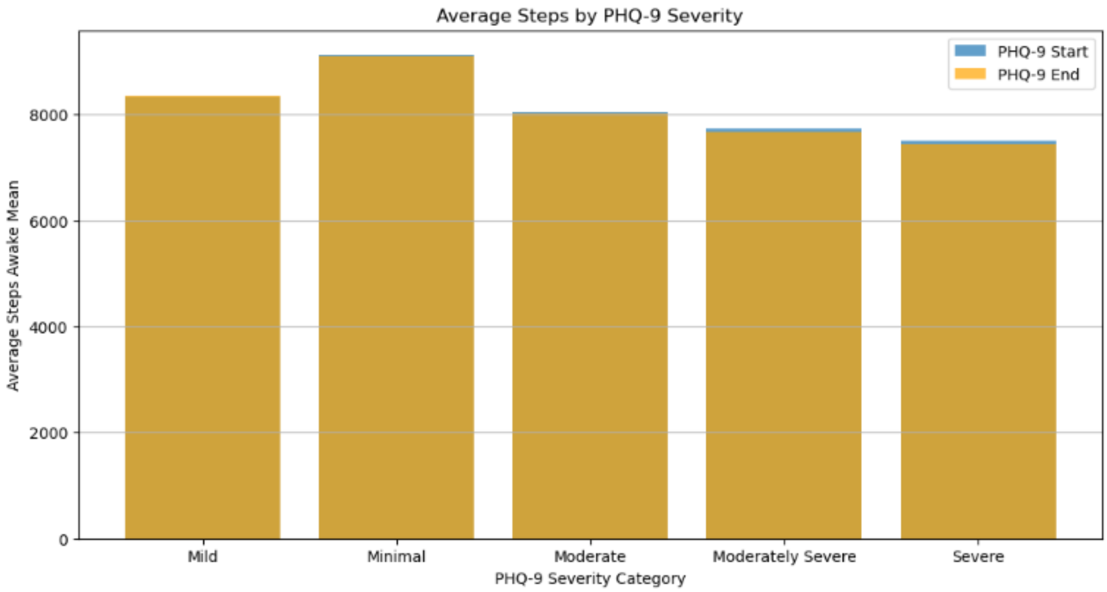


FIGURE 2

This analysis in figure 2 highlights the importance of exploring physical activity as a factor in depression management.

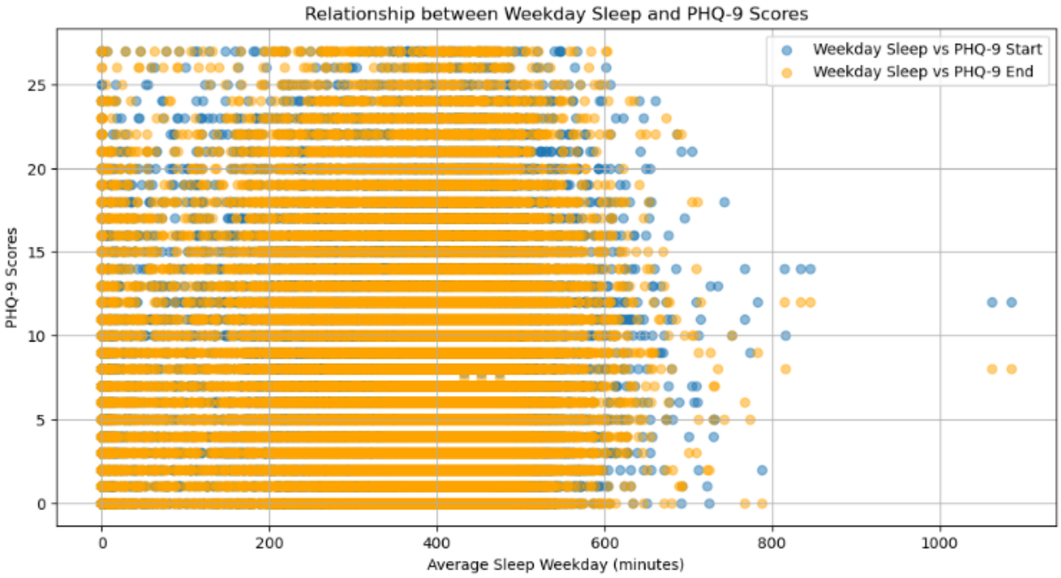
3.3 Relationship Between Sleep Metrics and PHQ-9 Scores

I analyzed the relationship between sleep metrics, including weekday sleep, weekend sleep, and sleep efficiency, and PHQ-9 scores to explore how sleep patterns are associated with depression severity.

To conduct the analysis, I calculated correlations between sleep metrics and PHQ-9 scores at both the beginning and end of the study. The key metrics examined included weekday sleep duration (sleep_asleep_weekday_mean), weekend sleep duration (sleep_asleep_weekend_mean), and sleep efficiency ratios (sleep_ratio_asleep_in_bed_weekday_mean) and sleep_ratio_asleep_in_bed_weekend_mean). Minimal correlations were observed between sleep duration and PHQ-9 scores. I also created scatter plots to visualize the relationship between weekday sleep and PHQ-9 scores at both time points.

The analysis revealed that weekday sleep duration had a minimal negative correlation with PHQ-9 scores, with values of -0.01 at both the start and end of the study. Weekend sleep duration showed a slightly stronger negative correlation, with values of -0.02 at both time points. Sleep efficiency ratios displayed negative correlations ranging from -0.11 to -0.10, indicating that higher sleep efficiency is associated with lower depression severity. Figure 3 highlights these relationships, including the minimal direct correlation between weekday sleep duration and PHQ-9 scores.

The findings suggest that sleep efficiency may provide a more meaningful understanding of depression severity compared to simple measures of sleep duration or time spent in bed. These insights are visually represented in Figure 3.




```
29]: (sleep_asleep_weekday_mean                -0.005868
      sleep_asleep_weekend_mean                -0.011383
      sleep_in_bed_weekday_mean                 0.043114
      sleep_in_bed_weekend_mean                 0.027805
      sleep_ratio_asleep_in_bed_weekday_mean   -0.108033
      sleep_ratio_asleep_in_bed_weekend_mean   -0.097750
      Name: phq9_score_start, dtype: float64,
      sleep_asleep_weekday_mean                -0.012886
      sleep_asleep_weekend_mean                -0.019284
      sleep_in_bed_weekday_mean                 0.037809
      sleep_in_bed_weekend_mean                 0.020749
      sleep_ratio_asleep_in_bed_weekday_mean   -0.112952
      sleep_ratio_asleep_in_bed_weekend_mean   -0.102004
      Name: phq9_score_end, dtype: float64)
```

FIGURE 3

3.4 Sleep Efficiency and PHQ-9 Severity

I explored group-level sleep efficiency patterns across different PHQ-9 severity categories to understand how sleep efficiency relates to depression severity.

To perform the analysis, I grouped participants by PHQ-9 severity categories, including Minimal, Mild, Moderate, Moderately Severe, and Severe. For each group, I calculated the average sleep efficiency ratio, which represents the proportion of sleep time to in-bed time on weekdays, at both the start and end of the study. The results were visualized using a bar chart that compared average sleep efficiency across severity categories for PHQ-9 Start and PHQ-9 End scores.

Participants in higher PHQ-9 severity categories, such as Moderately Severe and Severe, showed slightly lower sleep efficiency. In contrast, participants in lower severity categories, such as Minimal and Mild, exhibited higher sleep efficiency. The trend of reduced sleep efficiency with increasing PHQ-9 severity was consistent for both the start and end PHQ-9 scores. These findings are presented in Figure 4, highlighting the inverse relationship between sleep efficiency and depression severity.

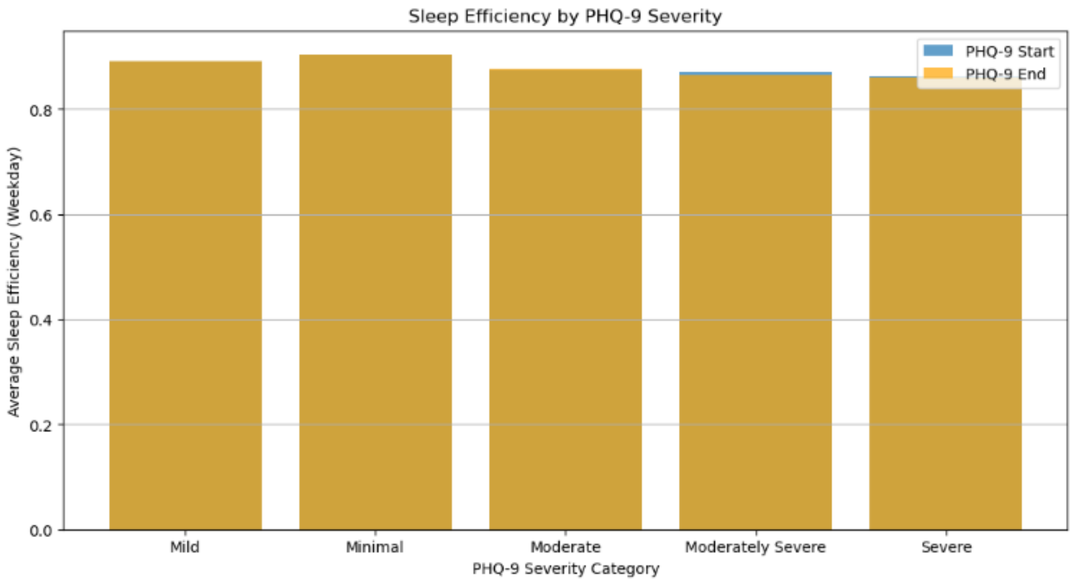


FIGURE 4

3.5 Analysis of Weekday and Weekend Sleep Patterns by PHQ-9 Severity

To investigate the impact of PHQ-9 severity on sleep patterns, I analyzed key metrics such as asleep time, in-bed time, and sleep efficiency for both weekdays and weekends across different severity categories. The results are summarized in three figures.

Figure 5 compares weekday and weekend sleep efficiency across PHQ-9 severity categories. The analysis shows that sleep efficiency decreases as PHQ-9 severity increases. Weekend sleep efficiency is consistently slightly higher than weekday efficiency, suggesting that weekends may offer recovery opportunities. This pattern is more pronounced in participants with moderate to severe depression, where sleep efficiency is significantly lower.

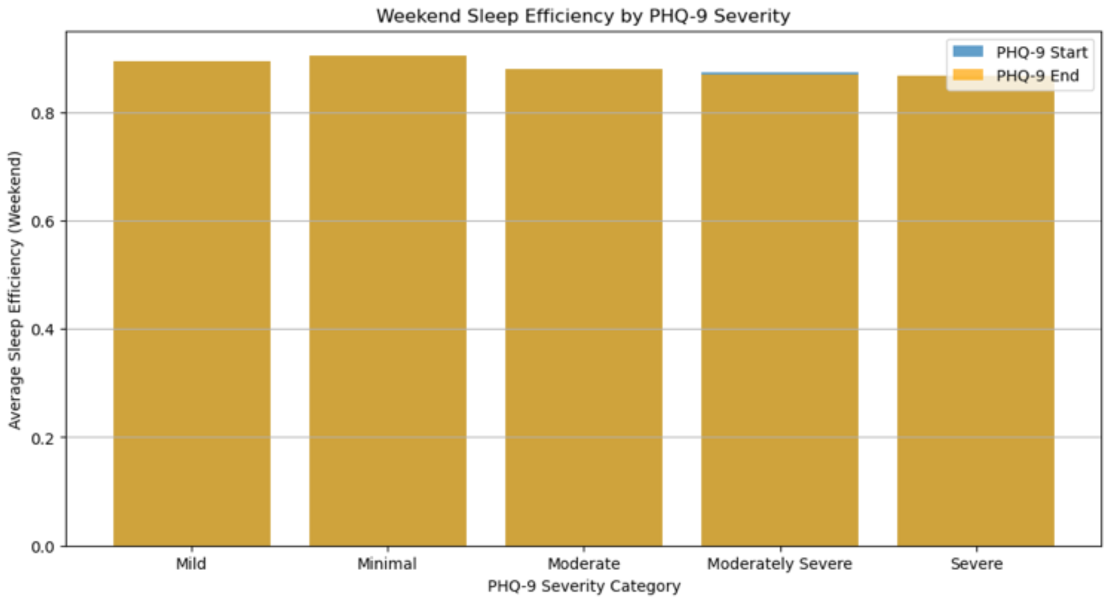


FIGURE 5

Figure 6 examines weekday and weekend sleep patterns by visualizing asleep time and in-bed time across severity categories. Participants with lower PHQ-9 scores, such as Minimal and Mild, have longer asleep times and shorter in-bed durations, indicating higher sleep efficiency. Weekend asleep and in-bed times are longer across all severity categories, with participants with higher PHQ-9 scores showing a slight improvement in sleep patterns over the weekend.

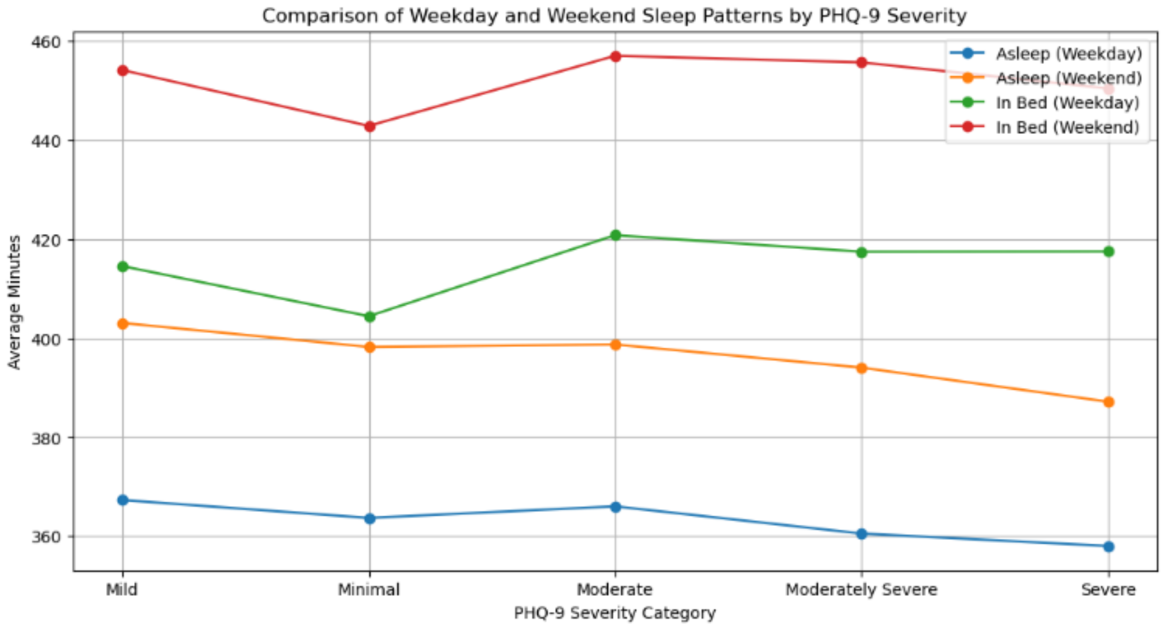


FIGURE 6

Figure 7 focuses on weekend sleep efficiency and compares PHQ-9 scores at the start and end of the study. Participants with minimal or mild depression severity consistently exhibit higher sleep efficiency on weekends. Those with severe depression show the lowest sleep efficiency, reinforcing the negative relationship between sleep efficiency and depression severity.

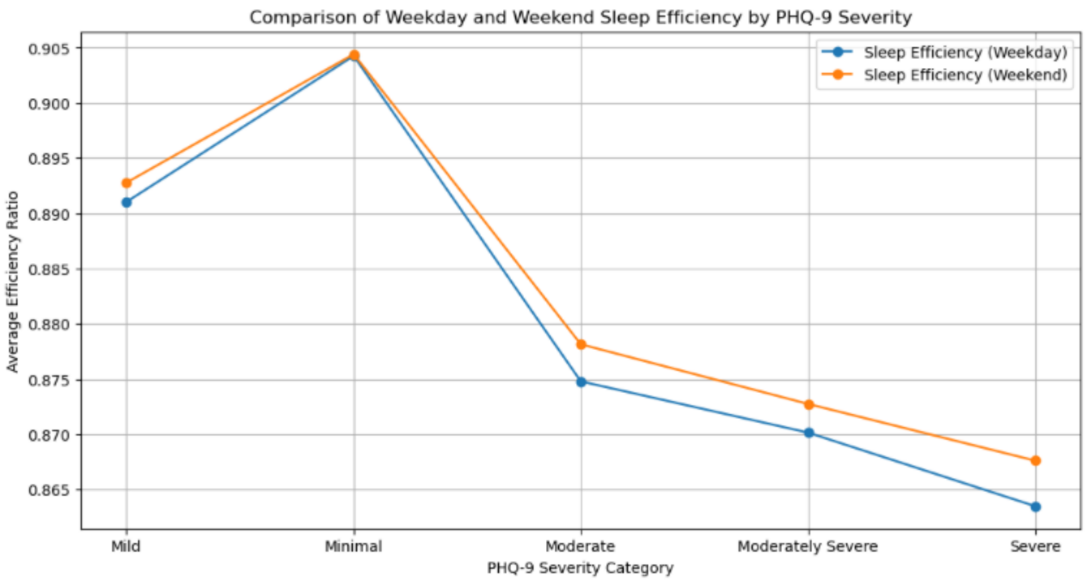


FIGURE 7

These findings are presented in Figures 5, 6, and 7, highlighting the relationship between PHQ-9 severity and various aspects of sleep efficiency and patterns. Participants with higher depression severity experience challenges with both sleep efficiency and duration, with weekends providing some benefits.

4 Methodology

4.1 Why I used Clustering and K-means

I used clustering, specifically K-Means, for this project because it is an effective method for segmenting participants into distinct groups based on shared characteristics. This approach was particularly suited for my analysis due to the following reasons:

Why I Chose Clustering:

1. Handling Unlabeled Data: The dataset lacked predefined labels for participant behaviors or outcomes. Clustering allowed me to group participants based on patterns in their data without requiring prior knowledge of their categories.

2. Identifying Patterns: Clustering is a powerful tool for identifying underlying patterns in multidimensional data. It helped me segment participants based on variables such as sleep efficiency, physical activity, and depression severity.

3. Behavioral Profiling: By clustering the data, I could create meaningful behavioral profiles e.g., participants with high sleep efficiency and low depression vs. those with poor sleep efficiency and severe depression. These profiles are crucial for understanding the relationship between behaviors and mental health.

Why I chose K-Means Clustering:

1 efficiency and Simplicity: K-Means is computationally efficient and straightforward to implement, making it suitable for handling the numerical data in this project.

2 Optimal Number of Clusters: Using the Elbow Method, I determined that three clusters provided the best balance between compactness (low variance within clusters) and distinctiveness (high separation between clusters).

3 Interpretability: K-Means provided clear and interpretable results. Each participant was assigned to a specific cluster, enabling me to analyze group characteristics and derive actionable insights.

4.2 Determining the Optimal Number of Clusters

To determine the optimal number of clusters for K-Means clustering, I applied the Elbow Method, which evaluates within-cluster variance across different numbers of clusters. Figure 8 presents the results of this process.

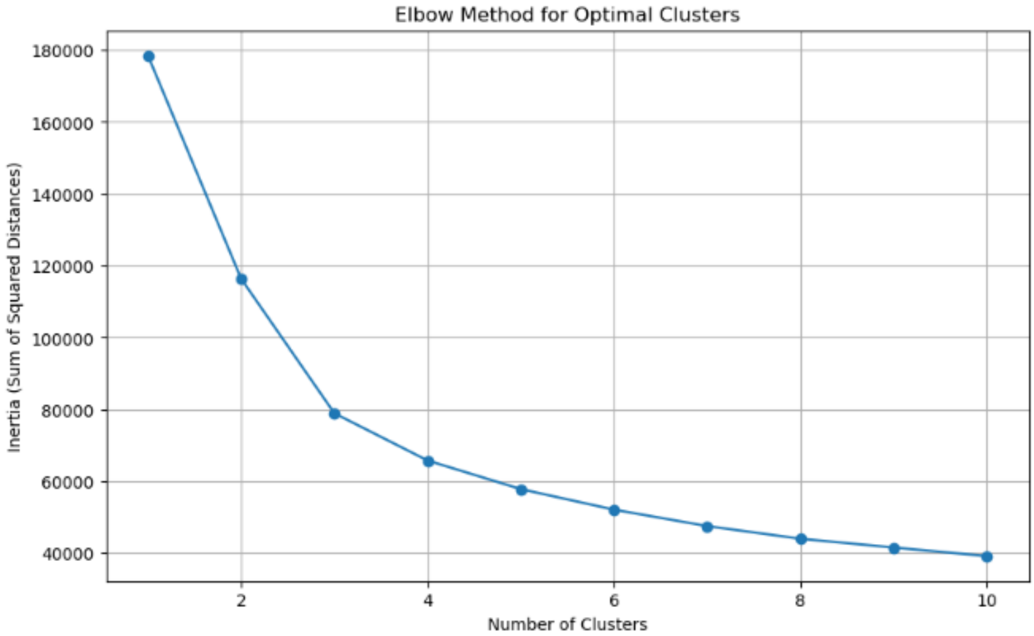


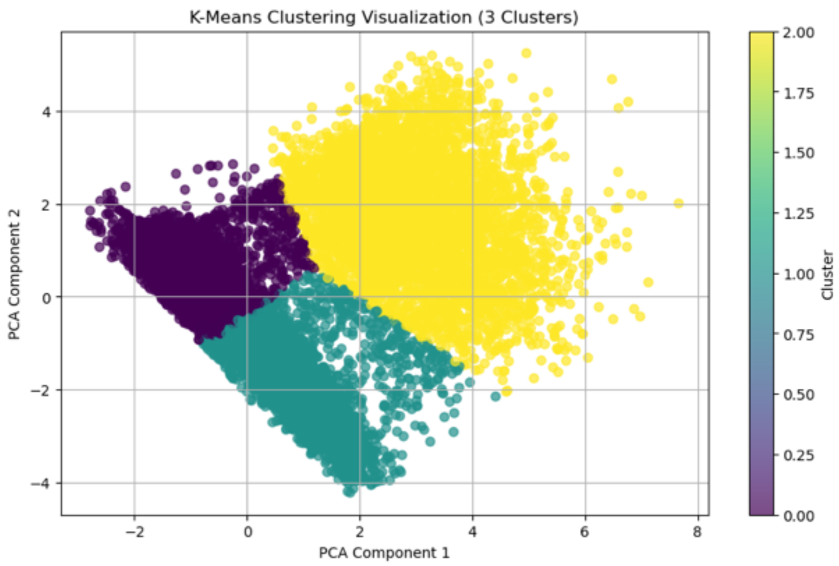
FIGURE 8

The plot in Figure 8 displays the sum of squared distances as the number of clusters increases. The results indicate that inertia decreases steeply from k=1 to k=3, showing that each additional cluster significantly improves grouping quality. Beyond k=3, the rate of decrease slows down, forming an elbow in the graph. This suggests that three clusters provide the best balance between compactness and separation.

Three clusters were selected based on several factors. First, the elbow point at k=3 demonstrates that increasing the number of clusters beyond this point yields diminishing returns. Second, the three clusters offer meaningful interpretations aligned with the study's goals. These clusters represent a group with high physical activity and low depression severity, a group with moderate activity and severe depression severity, and a group with poor sleep efficiency and moderate depression severity. Lastly, choosing three clusters strikes a balance between simplicity and detail, allowing for actionable segmentation without overcomplicating the analysis.

4.3 Cluster Characteristics and Observations

The clustering analysis identified distinct groups of participants based on sleep efficiency, physical activity, and PHQ-9 scores. Insights from the PCA-based clustering visualization, cluster metrics summary, and demographic differences were combined to better understand the behavioral and demographic characteristics of each cluster.



Cluster	steps_awake_mean	sleep_ratio_asleep_in_bed_weekday_mean	sleep_ratio_asleep_in_bed_weekend_mean	phq9_score_start	phq9_score_end
0	9140.159607	0.935285	0.935702	4.073035	3.865706
1	7597.540659	0.930278	0.933321	13.900310	13.565761
2	7568.992947	0.578072	0.582055	9.351071	9.154453

FIGURE 9

Figure 9 illustrates three distinct clusters, each representing unique participant profiles. Cluster 0 includes participants with high physical activity, high sleep efficiency, and low depression severity. Cluster 1 consists of individuals with moderate activity, moderate sleep efficiency, and severe depression severity. Cluster 2 comprises participants with poor sleep efficiency, moderate physical activity, and moderate depression severity. The table in Figure 9 summarizes key metrics for each cluster, highlighting differences in steps, sleep efficiency, and PHQ-9 scores.

To further interpret the clusters, demographic variables such as life stress, and efforts to reduce alcohol use were analyzed, as shown in Figure 10.

Cluster	life_stress	life_red_stop_alcoh
0	0.034097	0.039616
1	0.045695	0.044144
2	0.047125	0.036979

FIGURE 10

Cluster 0 is characterized by high activity levels and sleep efficiency, with the lowest PHQ-9 scores. Demographically, participants in this cluster report moderate life stress and moderate efforts to reduce alcohol use. This cluster represents a healthy group with optimal behaviors and minimal mental health challenges.

Cluster 1 shows moderate activity and sleep efficiency but the highest PHQ-9 scores. Participants in this cluster have report significant efforts to reduce alcohol use, suggesting potential lifestyle or behavioral challenges. This group may benefit from targeted interventions to improve activity levels and mental health outcomes.

Cluster 2 has poor sleep efficiency but moderate physical activity, paired with moderate depression severity. Participants in this cluster report the highest levels of life stress and the lowest efforts to reduce alcohol use. This cluster highlights the importance of addressing sleep quality and stress management in mental health interventions.

By integrating behavioral and demographic data, the analysis reveals protective factors, such as high activity and good sleep, associated with better mental health in Cluster 0. Cluster 1 identifies participants who may require focused interventions, while Cluster 2 emphasizes the need to improve sleep quality and manage stress to address mental health concerns. These findings provide a comprehensive understanding of the behavioral and demographic distinctions within each cluster.

4.4 Demographic Analysis by Cluster

To gain a deeper understanding of participant profiles, I analyzed demographic characteristics across the three clusters, focusing life stress, and efforts to reduce alcohol use. The bar charts in Figures 11 summarize these differences, revealing key behavioral and demographic trends.

4.4.2 Life Stress and Alcohol Reduction by Cluster

Figure 11 illustrates life stress levels and efforts to reduce alcohol use across the clusters. Cluster 0 participants report the lowest levels of life stress 0.034 and moderate efforts to reduce alcohol use 0.039, aligning with their overall positive behavioral profile. Cluster 1 participants show the second-highest stress levels 0.046 and the highest alcohol reduction efforts 0.044, indicating active attempts to manage their severe depression. Cluster 2 participants report the highest stress levels 0.047 and the lowest alcohol reduction efforts 0.037, highlighting the critical role of stress management in addressing their moderate depression severity.

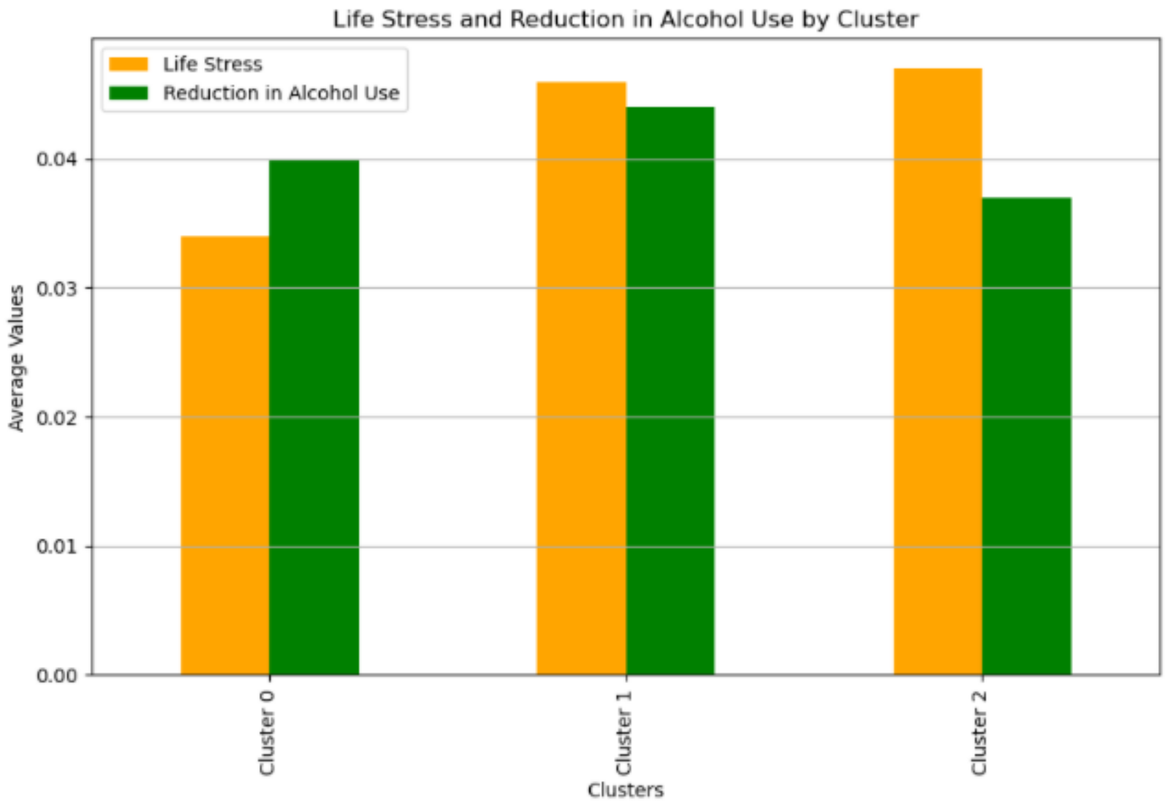


Figure 11

4.4.3 Conclusion

Cluster 0 emerges as the healthiest group, characterized by low stress, moderate alcohol reduction efforts. Cluster 1 faces the greatest challenges, moderate stress, and severe depression severity, despite their efforts to reduce alcohol use. Cluster 2 underscores the significant impact of high stress and poor sleep efficiency on moderate depression severity.

These findings emphasize the importance of integrating demographic and behavioral analyses to better understand the unique needs of each cluster. Such insights can inform targeted interventions to improve mental health outcomes tailored to the specific challenges faced by each group.

4.5 Correlation Between Clusters and PHQ-9 Scores

Figure 13 provides a comparative analysis of PHQ-9 scores at the start and end of the study across the three clusters. Cluster 0 participants have the lowest PHQ-9 scores, indicating minimal depression severity and better mental health outcomes. Their scores decreased slightly from 4.07 at the start to 3.87 at the end of the study. Cluster 1 participants exhibit the highest PHQ-9 scores, reflecting severe depression severity. Their scores showed a minor decline from 13.90 at the start to 13.56 at the end. Cluster 2 participants have moderate PHQ-9 scores, with values of 9.35 at the start and 9.15 at the end, consistent with their classification as a moderately affected group.

Figure 14 illustrates the correlation between clusters and PHQ-9 start and end scores through a heatmap. The analysis reveals a strong positive correlation between Cluster 1 and PHQ-9 scores, confirming that participants in this group experience severe depression. Cluster 0 shows a weak positive correlation with PHQ-9 scores, which aligns with their low depression severity and better mental health outcomes. Cluster 2 demonstrates a moderate correlation with PHQ-9 scores, aligning with its classification as participants with moderate depression severity.

Cluster 0 participants, with their weak positive correlation to PHQ-9 scores, represent individuals with better mental health outcomes, likely due to higher activity levels and better sleep efficiency. In contrast, Cluster 1, showing a strong positive correlation with PHQ-9 scores, highlights participants who experience severe depression and may require significant mental health interventions. Cluster 2, with a moderate correlation to PHQ-9 scores, underscores the importance of targeted interventions to address stress and improve sleep quality in this group.

This analysis of PHQ-9 scores validates the clustering approach, clearly differentiating mental health profiles for each group. The findings emphasize the need for tailored interventions, with Cluster 0 benefiting from maintaining positive behaviors, Cluster 1 requiring intensive mental health support, and Cluster 2 focusing on stress management and sleep quality improvement.

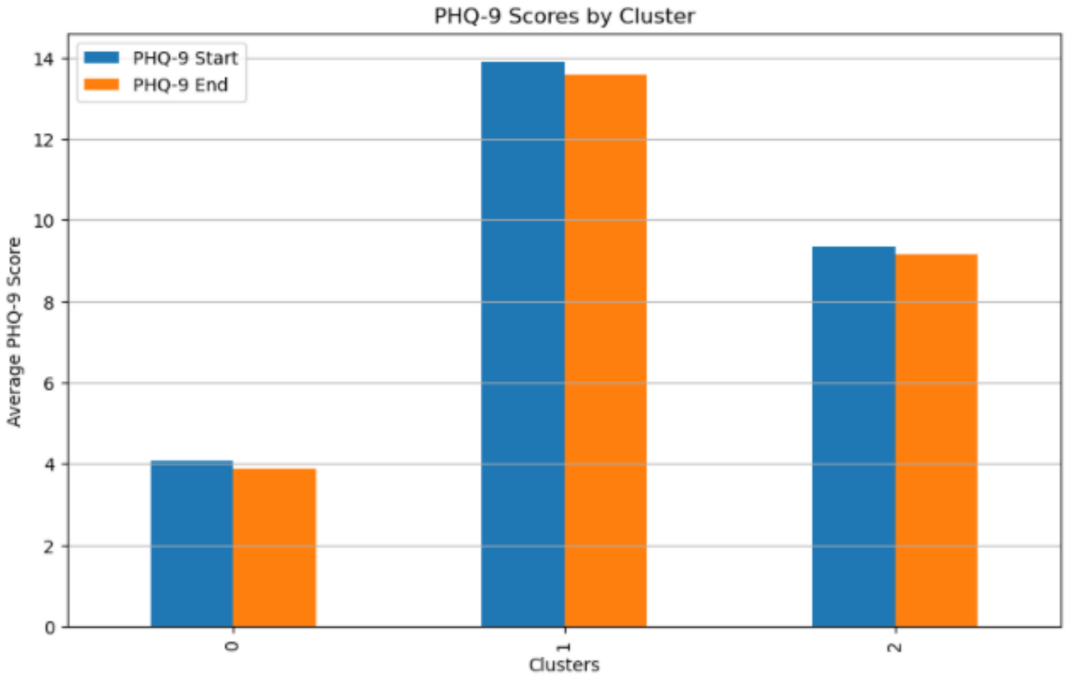

PHQ-9 Scores by Cluster
PHQ-9 Start
PHQ-9 End
Average PHQ-9 Score
0
2
4
6
8
10
12
14
Clusters
0
1
2

FIGURE 12

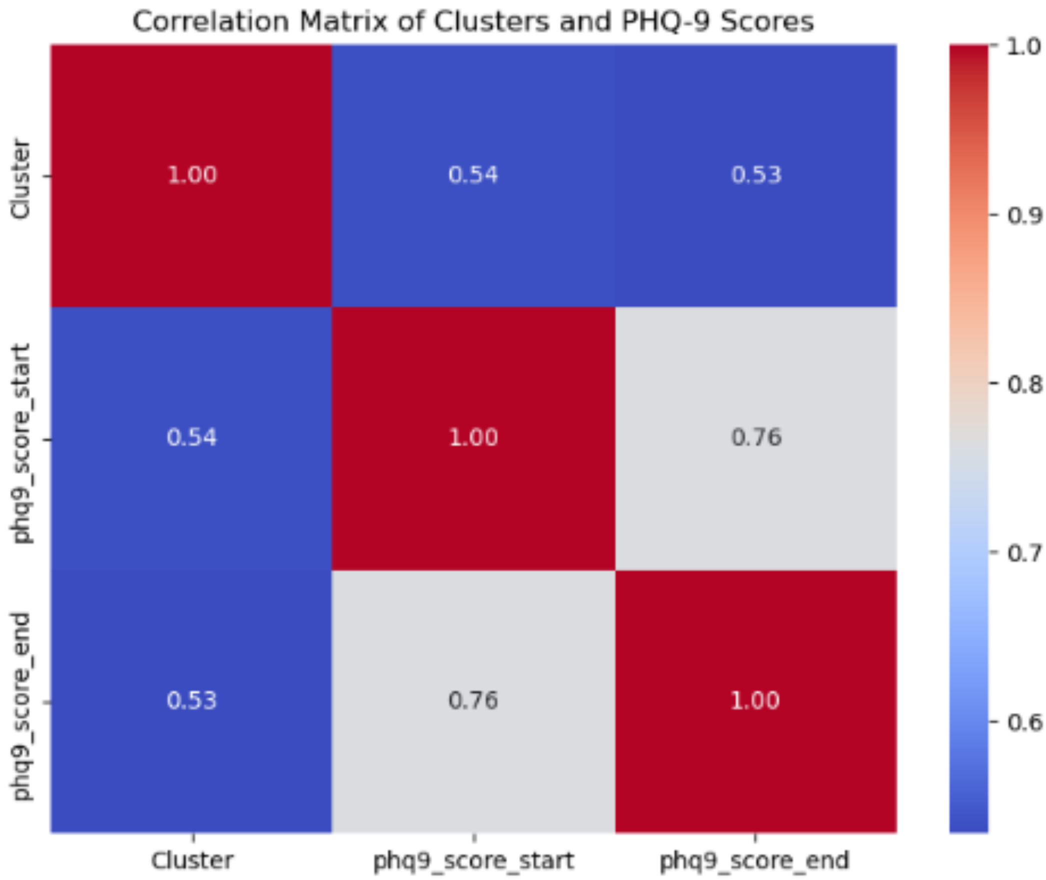

Correlation Matrix of Clusters and PHQ-9 Scores
Cluster
phq9_score_start
phq9_score_end
1.00
0.54
0.53
0.54
1.00
0.76
0.53
0.76
1.00
Cluster
phq9_score_start
phq9_score_end
1.0
0.9
0.8
0.7
0.6

FIGURE 13

5. RESULTS

The results of the analysis reveal meaningful insights into the relationship between wearable metrics and depression severity, as measured by PHQ-9 scores. The clustering approach successfully segmented participants into three distinct groups based on shared behavioral characteristics, including physical activity, sleep efficiency, and mental health indicators.

Participants in Cluster 0 displayed high levels of physical activity, optimal sleep efficiency, and the lowest depression severity, reflecting a positive mental health profile. Cluster 1 participants exhibited moderate activity levels and sleep efficiency but the highest depression severity, highlighting a need for targeted mental health interventions. Cluster 2, characterized by poor sleep efficiency and moderate depression severity, demonstrated the importance of improving sleep quality and managing stress for better mental health outcomes.

Correlation analyses showed weak to moderate relationships between physical activity, sleep patterns, and depression severity. While moderate-to-vigorous physical activity had minimal direct correlation with PHQ-9 scores, integrating multiple factors like sleep and activity levels provided deeper insights. Notably, sleep efficiency emerged as a significant metric, inversely associated with depression severity.

Demographic analyses further enriched the clustering results, revealing differences in stress levels, and efforts to reduce alcohol use across clusters. For instance, Cluster 1 participants, with severe depression and reported the most efforts to reduce alcohol consumption. Cluster 2 showed the highest life stress levels, further emphasizing the need for stress management strategies.

These results demonstrate that combining wearable data with clustering techniques can provide actionable insights into mental health management, enabling the identification of at-risk groups and the development of targeted interventions.

6. DISCUSSION

This project aimed to analyze the relationship between health behaviors, including physical activity and sleep patterns, and depression severity using wearable data and PHQ-9 scores. The findings provide insights into how specific behavioral patterns influence mental health outcomes and highlight the potential of wearable technology to support personalized interventions for depression.

The analysis revealed several important findings. Participants with higher physical activity levels, indicated by step counts and better sleep efficiency reported lower PHQ-9 scores, reflecting minimal depression severity. Conversely, participants with poor sleep efficiency and reduced physical activity were more likely to report higher PHQ-9 scores, indicative of severe depression symptoms. These results emphasize the value of combining activity and sleep data for a more comprehensive understanding of depression severity.

The clustering approach identified three distinct participant groups. Cluster 0 included participants with high physical activity, good sleep efficiency, and minimal depression severity. Cluster 1 consisted of individuals with moderate activity, poor sleep efficiency, and severe depression severity. Cluster 2 comprised participants with moderate depression severity, poor sleep efficiency, but moderate physical activity levels. These clusters provided meaningful subgroups that revealed patterns in the interplay between sleep quality, activity levels, and depression, which would not have been evident through single-variable analysis.

Sleep efficiency emerged as a key metric in understanding depression severity. Unlike raw sleep duration, sleep efficiency was more strongly associated with PHQ-9 scores, with participants reporting lower efficiency consistently showing higher depression severity. This finding underscores the importance of focusing on sleep quality in mental health research and interventions.

The findings have several implications for mental health research and interventions. Wearable devices offer unique opportunities for continuous monitoring of physical activity and sleep patterns, which can be combined with clinical assessments for a more comprehensive understanding of mental health. Data from wearables can support the early detection of behavioral changes linked to depression and provide real-time feedback for patients and clinicians. The clustering results revealed subgroups with distinct behavioral profiles, paving the way for personalized intervention strategies. For example, participants in Cluster 1 could benefit from targeted interventions addressing both sleep quality and physical activity, while those in Cluster 2 may require stress management programs and strategies to improve sleep efficiency. These tailored approaches ensure that interventions address individual needs, potentially enhancing their effectiveness.

The study also highlights the protective role of consistent physical activity and efficient sleep in maintaining mental health. Encouraging healthy sleep and regular physical activity could serve as preventative measures against severe depression. Public health campaigns and mental health programs can use these findings to design strategies for depression prevention and management.

Despite its strengths, the project has limitations. The study identified associations between health behaviors and depression severity but did not establish causation. It is unclear whether improving physical activity or sleep directly reduces depression severity or if other factors mediate this relationship. Longitudinal studies or intervention-based research would be necessary to confirm causality. The PSYCHE-D dataset, while rich, had limitations such as missing data that required imputation, a lack of detailed contextual information about participants' lives, and potential sample bias, which may limit the generalizability of findings. Additionally, the K-Means clustering approach assumes that clusters are spherical and equally distributed, which might oversimplify the relationships between variables. Future research could explore more advanced clustering techniques to capture non-linear relationships.

Building on these findings, several future directions can be pursued. Conducting longitudinal studies could establish causal links between health behaviors and depression severity, providing stronger evidence for intervention design. Incorporating additional metrics, such as heart rate variability, stress levels, or dietary patterns, could offer a more comprehensive understanding of how lifestyle factors interact with mental health. Future research could also test and evaluate personalized interventions targeting specific clusters, such as combined sleep and activity programs for Cluster 1 participants and stress management and sleep improvement for Cluster 2 participants. Integrating wearable data with mobile apps or telemedicine platforms could enable real-time monitoring and feedback, enhancing the effectiveness of mental health interventions.

7. Conclusion

This study explored the relationship between wearable metrics and depression severity using clustering and statistical analyses. The findings reveal that combining behavioral and demographic data provides a comprehensive understanding of mental health profiles and highlights key factors influencing depression severity.

The clustering approach segmented participants into three meaningful groups, each representing unique combinations of activity levels, sleep efficiency, and depression severity. Cluster 0 participants, with high activity and sleep efficiency, had the best mental health outcomes. In contrast, Clusters 1 and 2 highlighted challenges such as severe depression, poor sleep efficiency, and high stress, suggesting a need for tailored mental health interventions.

While sleep efficiency emerged as a critical factor, highlighting its relevance in managing mental health. Demographic insights, such as life stress and alcohol reduction efforts, further supported the behavioral profiles of each cluster.

This project demonstrates the potential of wearable technology in monitoring and managing mental health. By leveraging data-driven approaches like clustering, we can identify at-risk groups and design personalized interventions to improve sleep quality, increase physical activity, and reduce stress. These findings pave the way for future research and innovation in mental health care, emphasizing the importance of integrating behavioral insights with technological tools for a holistic approach to mental health management.

8. References:

1. Makhmutova, M., Kainkaryam, R., Ferreira, M., Min, J., Jaggi, M. and Clay, I., 2024. PSYCHE-D: predicting change in depression severity using person-generated health data (DATASET). [Dataset] Available at: https://zenodo.org/records/5085146 [Accessed 6 Dec. 2024].
2. World Health Organization (WHO), 2024. *Depression*. [Online] Available at: https://www.who.int/news-room/fact-sheets/detail/depression [Accessed 6 Dec. 2024].
3. Sivertsen, H., Bjørkløf, G., & Engedal, K., 2021. *Depression in older adults: Prevalence and determinants*. Annals of General Psychiatry. Available at: https://annals-general-psychiatry.biomedcentral.com/articles/10.1186/s12991-021-00375-x [Accessed 6 Dec. 2024].
4. Wang, Z., Huang, Q., & Lin, D. (2021). Evolution and emerging trends in depression research from 2004 to 2019. *Frontiers in Psychiatry*, [online] 12, p.705749. Available at: https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsyt.2021.705749/full [Accessed 6 Dec. 2024].