REPORT AND ANALYSIS

A Time Series Correlation Analysis Among Patients

Time Series Monitoring data

Prepared By

Brylle Matthew A. Lupac

Naomi Christienne Tiama

1. Introduction

This dataset contains longitudinal health monitoring data collected from three individual patients across a span of 42 months. It focuses on three key wellness indicators: average monthly step count, self-reported stress level, and Body Mass Index (BMI), with separate values recorded for each patient. The **Month_numerical** column serves as a time-tracking variable, starting from month 1 up to month 42, allowing researchers to examine behavioral patterns and health changes over an extended period. The **Patient_avg_steps** columns reflect each patient's physical activity level, measured by the average number of steps taken each month, which is a common indicator of lifestyle and cardiovascular fitness. The **Patient_Stress_Level** variables capture subjective stress ratings on a scale of 1 to 10, where 10 indicates very high stress and 1 indicates minimal stress, providing insights into emotional and psychological health. Meanwhile, the **Patient_BMI** values provide an estimate of body fat based on height and weight, allowing observation of weight trends and potential links to activity and stress levels. This structured time series data can be used for detailed trend analysis, patient-specific comparisons, or multi-variable modeling to understand how physical activity and stress relate to weight stability or changes. It is particularly useful for exploring personalized health trajectories, identifying periods of high or low activity and stress, and analyzing whether consistent exercise or elevated stress levels are associated with changes in BMI over time.

2. Methods

2.1. Data Cleaning and Preparation

- The dataset was first imported from a .csv file using read.csv().
- Column names were cleaned and standardized using the clean_names() function.
- Missing values were checked and handled using the na.omit() function.
- All variables were coerced into numeric form using the mutate(across(...)) function.

2.2. Descriptive Statistical Analysis

- For each numeric variable, the group computed the following:
    - **Mean** – the average value
    - **Median** – the middle value in the ordered dataset
    - **Standard Deviation** – a measure of how spread out the values are from the mean
    - **Variance** – the square of the standard deviation, indicating variability
- Measures were calculated using a function called describe_variable() and applied to all relevant columns using lapply().

2.3. Data Visualization

- **Histogram**: A histogram of Patient 1 Average Steps was created using geom_histogram().
- **Boxplot**: A boxplot comparing the BMI of all three patients was created by reshaping the data into a long format with pivot_longer().
- **Scatter Plot**: A scatter plot was made for Patient 1 Average Steps vs. Stress Level using geom_point() and geom_smooth(method = "lm").
- **Bar Graph**: The average stress levels for each patient were calculated and plotted using a bar graph.

2.4. Advanced Statistical Test

- To see the relationship between average steps per day and stress level, the group performed the Pearson correlation test using the cor.test() function for each of the three patients. This test measures the strength and direction of the linear relationship between two continuous variables.
- The correlation coefficient (r) was interpreted as follows:
    - r ≈ 1: strong positive relationship
    - r ≈ -1: strong negative relationship
    - r ≈ 0: no linear relationship
- The corresponding p-value indicated whether the observed correlation was statistically significant.

3. Results and Discussion

3.1. Summary Statistics

Table 1. Descriptive Statistics of Step Count, Stress Level, and BMI per Patient

Variable	Subject	Mean	Median	Standard Deviation
Step Count	Patient 1	3,648.74	3,530.00	2,214.48
	Patient 2	2,916.10	2,783.50	1,578.09
	Patient 3	4,081.62	3,623.00	2,849.03
Stress Level	Patient 1	6.48	7.00	2.54
	Patient 2	7.17	7.00	1.70
	Patient 3	5.76	6.50	2.87
BMI	Patient 1	24.24	24.00	0.96
	Patient 2	24.42	24.35	0.29

	Patient 3	24.92	25.15	0.96

- **Step Count**
  - Patient 3 had the highest mean step count (4,081.62), indicating greater physical activity compared to the others.
  - Patient 2 had the lowest mean (2,916.10), suggesting the least active lifestyle.
  - Patient 3 also showed the greatest variability (SD = 2,849.03), meaning inconsistent step counts over time
- **Stress Level**
  - Patient 2 had the highest average stress level (7.17), possibly linked to low physical activity.
  - Patient 3 reported the lowest mean stress (5.76), which may relate to higher physical activity.
  - Stress levels for all patients had moderate to high variability, especially for Patient 3 (SD = 2.87).
- **BMI**
  - All patients had similar mean BMI values (~24–25), which fall within the normal weight range.
  - Patient 2 had the most stable BMI (SD = 0.29), indicating minimal fluctuation.
  - Patient 3 had the highest mean BMI (24.92), slightly higher than the others but still within a healthy range.

3.1. Data Visualizations

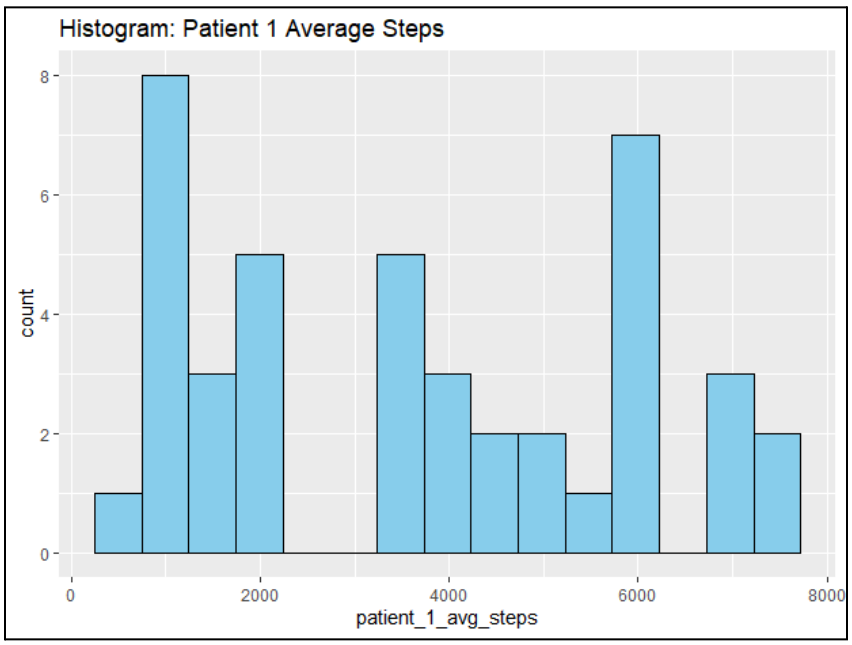

Figure 1. Histogram of Average Monthly Step Counts

- The most common average step count was around 1,500 steps, with 8 occurrences.
- 6,000 steps followed closely with 7 occurrences, indicating a smaller group with higher activity.
- Moderate activity levels like 2,000 and 3,500 steps each occurred 5 times, showing variability across patients.

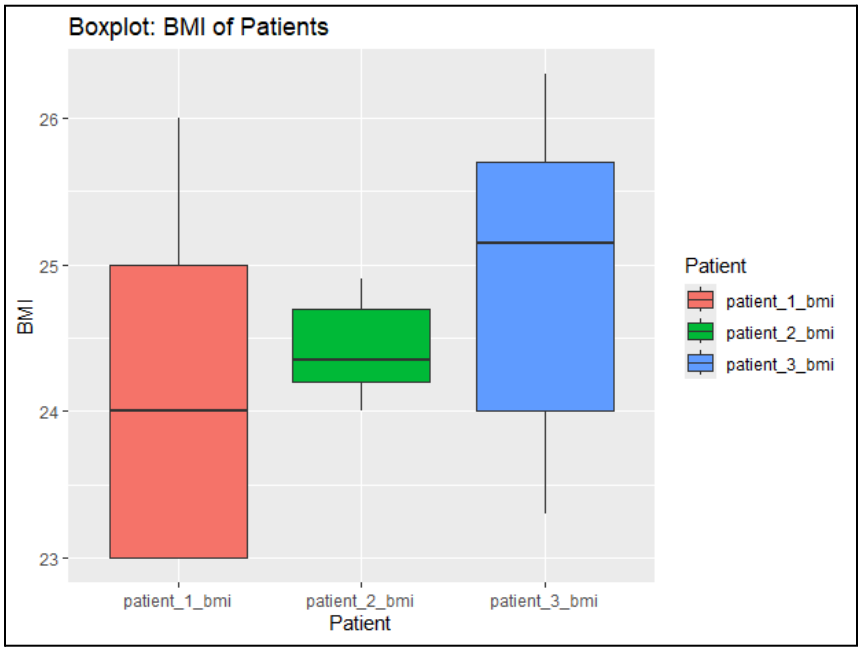

Figure 2. Box Plots of Monthly Step Counts per Patient

- Patient 3 showed the highest median step count and the widest spread, indicating both high and variable physical activity.
- Patient 1 had a moderate median with a more compact range, suggesting fairly consistent activity.
- Patient 2 had the lowest median and the least variability, reflecting consistently low physical activity levels.
- No extreme outliers were observed, but differences in distributions highlight variation in each patient's activity patterns.

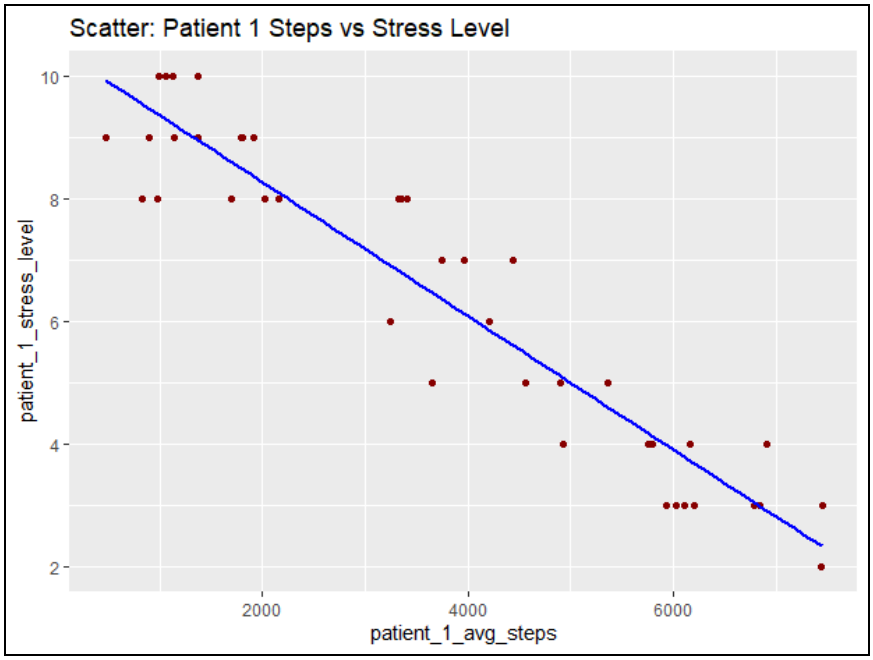

Figure 3. Scatter Plot of Patient 1 Steps vs Stress Level

- The scatter plot shows a strong negative relationship between step count and stress level for Patient 1.
- As step count increases, the stress level tends to decrease, suggesting that higher physical activity may help reduce stress.
- The pattern is tightly clustered along a downward trend, supporting a high negative correlation.

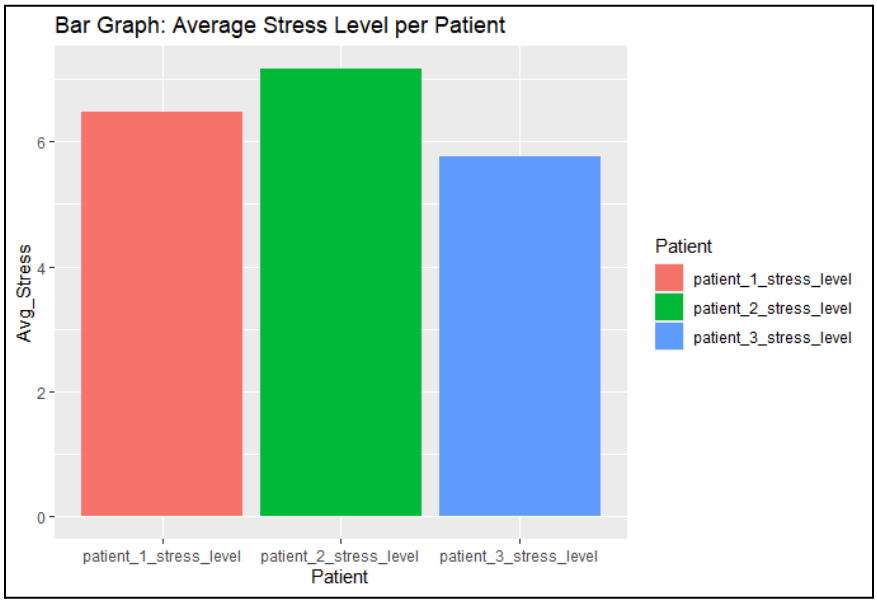

Figure 4. Bar Graph of Monthly Average BMI per Patient

- Patient 2 had the highest average stress level among the three, indicating greater overall stress.

- Patient 1 showed a moderate average stress level, lower than Patient 2 but higher than Patient 3.
- Patient 3 had the lowest average stress level, suggesting better stress regulation or lower stress exposure.

2.3 Advanced Statistical Insight

Table 2. Pearson Correlation: Average daily steps and stress levels for each patient

Patient	Correlation Coefficient (r)	p-value	Interpretation
1	-0.949	1.0887×10^{-21}	Strong negative correlation; statistically significant
2	-0.856	4.7399×10^{-13}	Strong negative correlation; statistically significant
3	-0.963	1.7278×10^{-24}	Very strong negative correlation; highly statistically significant

- All three patients showed a strong negative correlation between average daily steps and stress levels.
- As step count increased, stress levels decreased.
- Patient 3 had the strongest inverse relationship, followed by Patient 1 and then Patient 2.

4. Conclusion

The analysis of longitudinal health data from three patients over 42 months revealed meaningful patterns in physical activity, stress levels, and BMI. Across all patients, higher monthly average step counts were associated with lower stress levels, as confirmed by strong negative Pearson correlation coefficients. Patient 3 showed the most consistent trend with the highest physical activity, lowest stress, and a slightly elevated but stable BMI. Patient 2, in contrast, exhibited the lowest physical activity, highest stress, and the most stable BMI. Overall, the findings suggest that regular physical activity may help reduce stress without causing major fluctuations in BMI. These results highlight the potential of personalized, long-term monitoring in promoting health and managing stress through lifestyle habits.