Health Monitoring and Analysis Report

Introduction

Health Monitoring and Analysis involves systematically using health data to track and evaluate individuals' or populations' health status over time. This report presents an analysis of a dataset that includes various health metrics for 500 patients, aiming to monitor their health, identify patterns, and group them based on health standards.

Dataset Overview

The dataset contains the following columns:

- PatientID: Numerical identifier for the patient.
- Age: Age of the patient in years.
- Gender: Gender of the patient.
- HeartRate: Heart rate in beats per minute.
- BloodPressure: Blood pressure readings, formatted inconsistently.
- RespiratoryRate: Respiratory rate in breaths per minute.
- BodyTemperature: Body temperature in Fahrenheit.
- ActivityLevel: Activity level at the time of the measurement.
- OxygenSaturation: Oxygen saturation percentage.
- SleepQuality: Quality of sleep reported by the patient.
- StressLevel: Reported level of stress.
- Timestamp: Date and time of the measurement.

Data Cleaning

The initial inspection of the dataset revealed missing values in the BodyTemperature and OxygenSaturation columns. The missing values were filled using the median values of the respective columns.

Summary Statistics and Distributions

- **Age**: The ages range from 18 to 84 years, with a mean of 51.1 years.
- **Heart Rate**: The heart rate ranges from 60.2 to 99.9 bpm, with a mean of 80.1 bpm.
- **Respiratory Rate**: The respiratory rate ranges from 12 to 23 breaths per minute, with a mean of 17.5 breaths per minute.

- **Body Temperature**: The body temperature ranges from 97.1 to 99.5°F, with a mean of 98.6°F.
- **Oxygen Saturation**: The oxygen saturation ranges from 94% to 99%, with a mean of 96.3%.

Gender Distribution and Correlations

The gender distribution is nearly even, with males comprising 51.2% of the dataset. The correlation matrix showed no strong correlations between the variables, indicating that changes in one metric are not strongly associated with changes in the others.

Health Metrics Analysis

1. **Heart Rate by Activity Level**:
    - The median heart rate increases from resting to walking but does not significantly increase further during running. This is unusual as more strenuous activity typically results in a higher heart rate.
    - There is considerable overlap in heart rate ranges for walking and running activities.
2. **Blood Pressure Distribution**:
    - The systolic blood pressure shows a more spread-out distribution with common readings around 120 mmHg and 140 mmHg.
    - The diastolic blood pressure has a narrower distribution with a significant peak around 80 mmHg.
3. **Health Metrics by Gender**:
    - Heart rate and oxygen saturation show no significant differences between males and females. The median values and interquartile ranges are nearly identical for both genders.
4. **Heart Rate and Oxygen Saturation by Sleep Quality and Stress Levels**:
    - Heart rate remains relatively consistent across different levels of sleep quality and stress.
    - Oxygen saturation shows a minimal decrease in median values from excellent to poor sleep quality but remains largely unchanged across stress levels.
5. **Respiratory Rate and Body Temperature by Activity Levels**:
    - Respiratory rate increases with higher activity levels, as expected.
    - Body temperature remains relatively stable across different activity levels.

Conclusion

The analysis of the health monitoring dataset provides insights into the health status and patterns among the patients. The data shows expected trends such as increased heart and respiratory rates with higher activity levels and relatively stable body temperature across activities. However, the lack of significant differences in heart rate and oxygen saturation across genders, sleep quality, and stress levels indicates that these factors may not strongly influence these metrics within this dataset.

Future analyses could benefit from more detailed data on patient history, lifestyle choices, and genetic information to understand better the factors affecting these health metrics. Additionally, exploring machine learning models could help predict health risks and recommend personalized interventions based on the data.