

Bellabeat Data Analysis Case Study

Jada Towner

Feb 21, 2025

Introduction

Bellabeat, founded in 2013, is a high-tech company specializing in health-focused smart products. The company is known for developing wellness gadgets that promote healthy lifestyles, with a focus on women's health. In this report, I analyze the usage data of non-Bellabeat smart devices, with the aim to understand how users engage with their fitness trackers and related applications. Insights from this analysis can be used to improve Bellabeat's offerings and provide users with better, more personalized health recommendations.

Business Task

The goal of this analysis is to evaluate how consumers use non-Bellabeat smart devices, specifically fitness trackers like Fitbit. By understanding the usage patterns, Bellabeat can enhance its own product offerings and design features that better meet user needs.

Problem

How can Bellabeat leverage insights from the usage of non-Bellabeat smart devices to enhance its products and align them with user preferences?

Data Sources

The dataset used in this analysis was obtained from the Fitbit Fitness Tracker Data available on Kaggle. This dataset includes data collected from 30 Fitbit users between March 12, 2016, and May 12, 2016, and contains information on daily activity, sleep, heart rate, and step counts. The data was made available by Mobius, a data scientist at Healthcare Melbourne, Victoria, Australia. The dataset is deemed credible as it is part of the public domain.

Data Breakdown

The dataset consists of several CSV files with different types of data:

- **dailyActivity_merged.csv:** Includes daily activity data such as steps, calories burned, and intensity levels.
- **sleepDay_merged.csv:** Contains data on total minutes slept, total time in bed, and sleep quality.

- **hourlySteps_merged.csv:** Provides hourly step data to track daily activity levels at a more granular level.

These data were analyzed to understand patterns in activity, sleep, and user behavior across various metrics.

Analysis Methodology

To conduct this analysis, I used R and several key libraries for data cleaning and manipulation, including:

- Tidyverse
- dplyr
- lubridate
- ggplot2
- tidyr

Data was imported, cleaned, and merged to ensure consistency across the datasets, especially with respect to date and user ID alignment. The following steps were performed during the analysis:

1. **Data Cleaning and Preparation:** The dataset columns were cleaned using the `clean_names()` function, and date formats were standardized to ensure uniformity.
2. **Data Exploration:** We explored the dataset using summary statistics to determine key insights such as average daily steps, calories burned, and sleep patterns.
3. **Data Merging:** The datasets were merged by user ID and date to create a central table that allowed for more insightful comparisons across different activity measures.

Key Insights from the Analysis

1. **Activity Data:**
 - The average steps taken per day was 7,638 steps, which is below the recommended 10,000 steps per day.
 - The average calories burned per day was 2,304, which aligns with the recommended daily intake for active individuals.
 - Users spent an average of 991.2 minutes per day sedentary, which is significantly higher than the recommended 4 hours.
2. **Sleep Data:**
 - The average time spent asleep was 419.5 minutes (about 7 hours), which is lower than the recommended 7-9 hours of sleep per night.

- Users spent an average of 458.6 minutes in bed, showing that they generally spend more time in bed than they actually sleep.
3. **Step Data:**
 - The total steps recorded across all users demonstrated trends in activity but highlighted that many users are not meeting the recommended daily step count.
 4. **Visualization:**

The data visualizations revealed patterns such as:

 - A positive relationship between daily steps and calories burned.
 - A negative relationship between sedentary minutes and sleep quality, indicating that increasing activity levels may improve sleep.

Recommendations for Bellabeat

Based on the analysis of the Fitbit data, the following recommendations can be made for Bellabeat to improve its products and better serve its users:

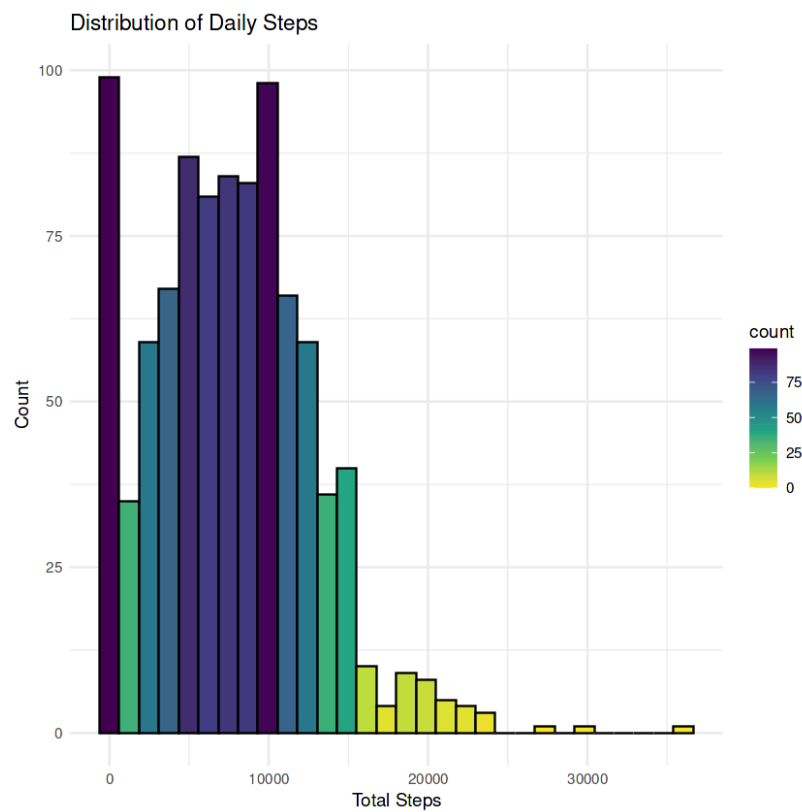
1. **Personalized Activity Goals:** Bellabeat can introduce personalized activity goals based on users' current activity levels. The app should set a daily step goal and encourage users to reach it with motivational messages or gamified features, like weekly step challenges.
2. **Sleep Improvement Features:** Bellabeat should focus on improving sleep tracking by providing more detailed insights into sleep quality. Suggestions on optimal bedtimes and routines could help users improve their sleep patterns. For example, offering bedtime reminders based on activity levels during the day could encourage better sleep habits.
3. **Sedentary Time Reduction:** The app can encourage users to reduce their sedentary time by sending reminders to stand or walk after periods of inactivity. Bellabeat could also integrate gentle reminders to take breaks throughout the day, improving user engagement.
4. **Integration of Activity Data with Menstrual Tracking:** Bellabeat could leverage the existing menstrual cycle data in conjunction with fitness and sleep data to provide more holistic health insights, such as how activity levels may impact menstrual cycles or how sleep patterns change during different phases.
5. **Enhance Social and Community Features:** To increase user engagement, Bellabeat could enhance the community aspect of the app, allowing users to share their achievements and compete in fitness challenges with friends or other users globally.

Conclusion

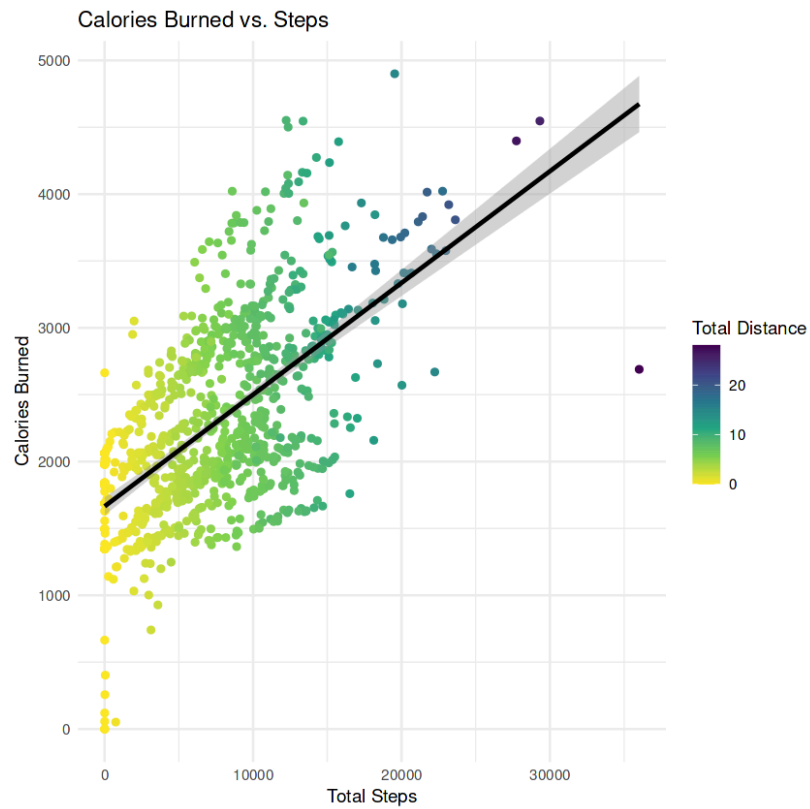
By analyzing non-Bellabeat smart device usage data, we gained valuable insights into how users interact with their fitness trackers. Bellabeat can use these insights to improve their own product, offer more personalized health recommendations, and increase user engagement. Implementing the

suggested changes will help Bellabeat stay competitive in the growing wearables market and foster a more health-conscious community.

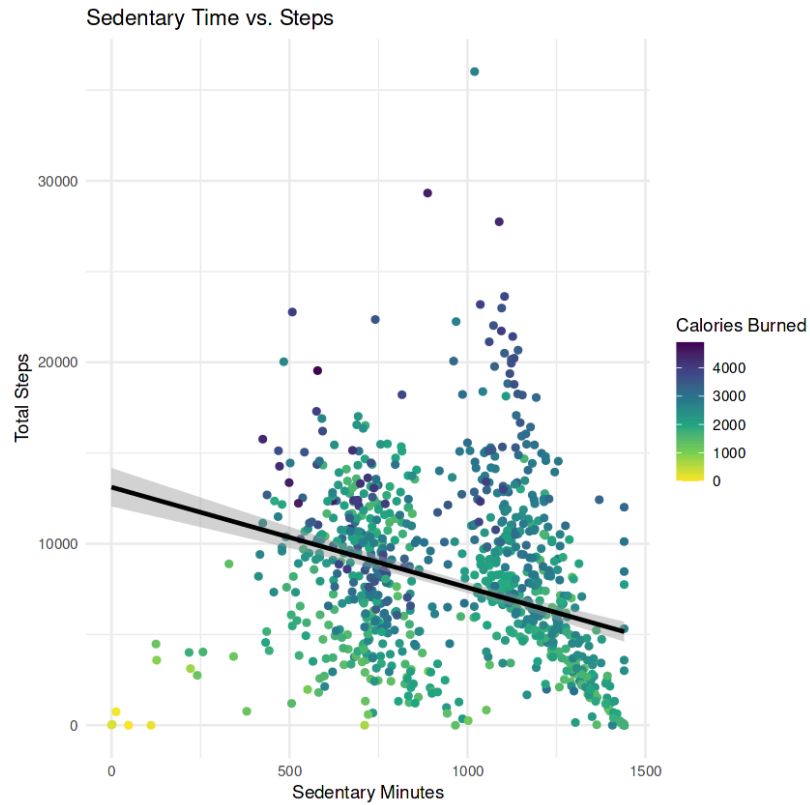
Appendix



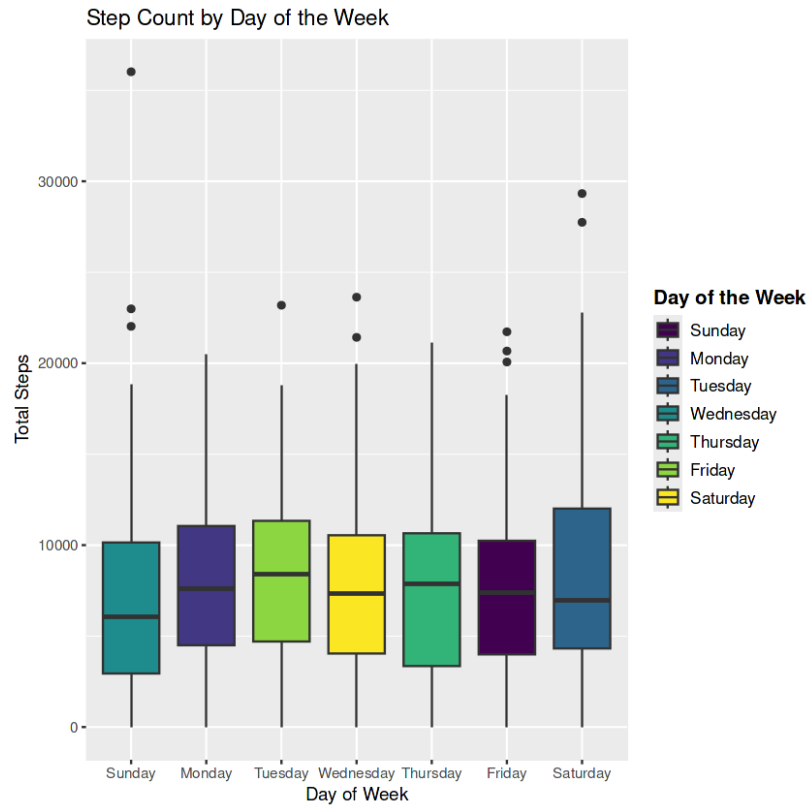
This histogram visualizes the distribution of total daily steps across the dataset. The distribution provides insight into the frequency of different step counts, helping to identify patterns in daily activity levels.



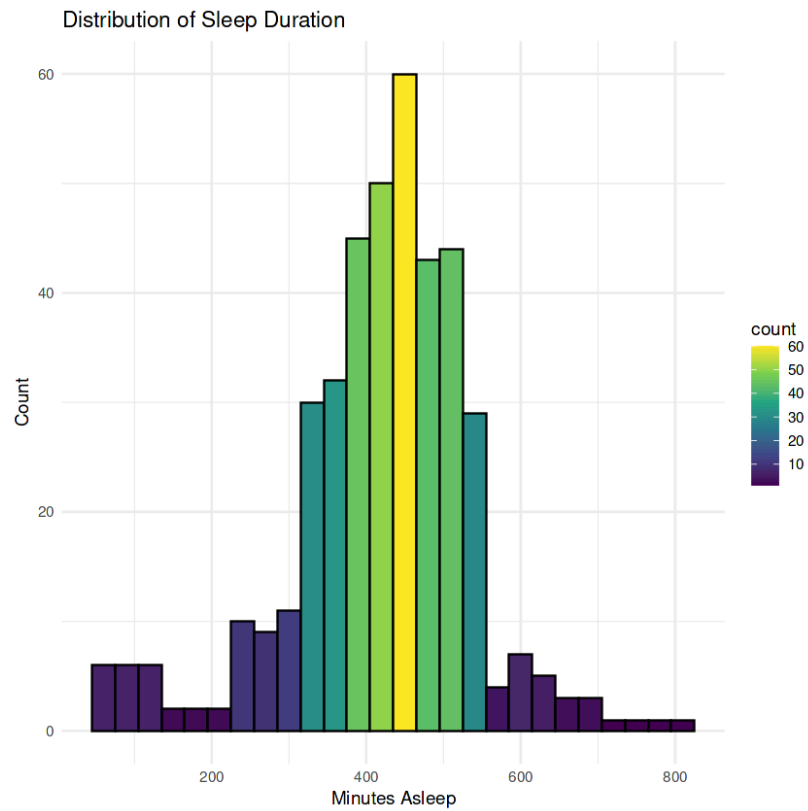
This scatter plot illustrates the relationship between the total steps taken and the calories burned. The trend suggests a positive correlation, indicating that as the number of steps increases, the number of calories burned tends to increase as well.



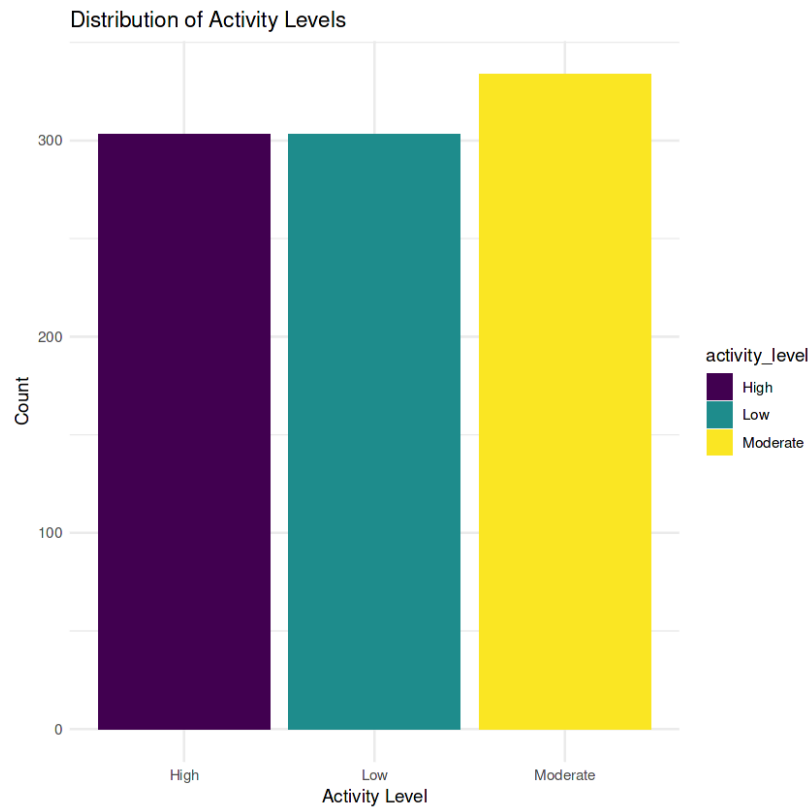
This scatter plot compares sedentary time with total steps taken. A negative correlation is observed, where higher sedentary time corresponds to fewer steps, suggesting that more sedentary behavior is linked with lower physical activity.



This box plot displays the step count distribution across the days of the week. It shows the variation in step counts, highlighting the days with the highest and lowest physical activity levels. The plot suggests that certain days may exhibit more consistent or higher step counts.



This histogram visualizes the distribution of sleep duration across the dataset. It provides an overview of the typical sleep patterns of individuals, revealing common sleep durations and any outliers in the data.



This bar chart depicts the frequency of various activity levels, helping to understand how often individuals engage in different levels of physical activity. It highlights the most common activity levels and provides insight into overall activity trends within the dataset.