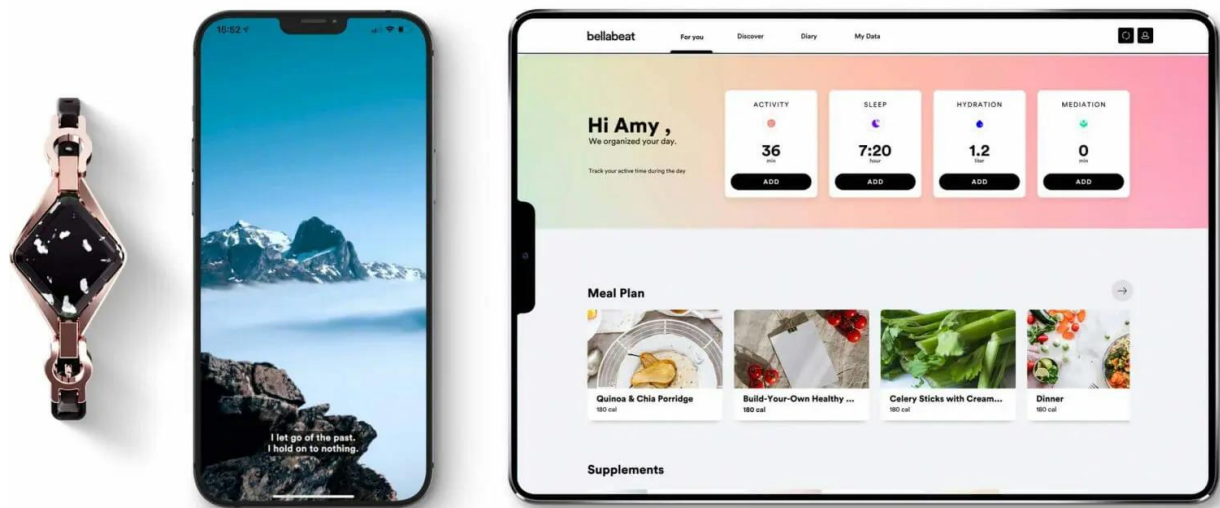# Analysis of Smart Device Usage for Strategic Marketing Insights at Bellabeat

Emmett Demirhan

2023-12-06

Table of Contents

Executive Summary

This analysis presents strategic insights into smart device usage for Bellabeat, a high-tech producer of health-focused products for women. The study used data from Fitbit fitness trackers to understand user behaviours in physical activity, sleep monitoring, and heart rate trends to enhance Bellabeat's marketing strategies.

Key findings indicate that users exhibit significant variability in physical activity, with an average step count of 7,638 steps per day and a considerable standard deviation, suggesting a wide range of individual activity levels. Moreover, sleep data analysis uncovered varying patterns, remarkably less sleep on Tuesdays and Thursdays. Sedentary behaviour was notably predominant, with users averaging over 16 hours of inactivity, stressing the need for features that encourage movement. Additionally, a positive linear relationship was found between total steps and calories burned, underscoring the importance of regular physical activity for energy expenditure and weight management.

The strategic recommendations for Bellabeat based on this analysis include developing personalised sleep insights, sedentary behaviour alerts, and wellness content targeting less active days. Furthermore, expanding data collection to include demographic information would enable more targeted marketing strategies.

These insights could be implemented through product development, marketing campaigns, and customer engagement initiatives that align with the identified user behaviours. Continuous monitoring of strategy effectiveness and data-driven adjustments are essential for sustained impact.

In conclusion, while the dataset provides valuable insights, it is limited by its size, scope, and the fact that it is somewhat dated. To maximise the efficacy of the derived strategies, Bellabeat is advised to invest in more comprehensive and current data collection, including demographic details, to ensure that the insights and recommendations reflect its target market's behaviours and needs.

Business Task Overview

Objective:

The primary objective of this analysis is to provide Bellabeat, a high-tech manufacturer of health-focused products for women, with actionable insights derived from the analysis of smart device usage data. This analysis is aimed at informing and enhancing the company's marketing strategies.

Background:

Bellabeat, established in 2013, has been at the forefront of developing innovative, health-focused smart products designed specifically for women. With a range of products, including the Bellabeat app, Leaf, Time, and Spring, the company has made significant strides in integrating technology with wellness. However, to keep pace with the dynamic smart device market and to capitalise on potential growth opportunities, a deeper understanding of consumer behaviour and device usage is essential.

Problem Statement:

Despite Bellabeat's success in creating aesthetically pleasing and functional wellness products, there is a need to understand the broader trends in smart device usage. This understanding is crucial to effectively tailor Bellabeat's marketing strategies, ensuring they resonate with the target audience's habits and preferences.

Significance:

By analysing existing smart device usage data, specifically from non-Bellabeat products like Fitbit, I aim to uncover trends and patterns that can inform Bellabeat's marketing strategies. This analysis will help Bellabeat align its products with consumer needs, preferences, and lifestyle choices, enhancing customer engagement and market share.

Approach:

My approach involves a detailed analysis of the FitBit Fitness Tracker Data to explore user habits in areas such as physical activity, heart rate monitoring, and sleep patterns. Though external, this data provides a valuable proxy to understand the potential behaviours and preferences of Bellabeat's current and prospective customer base.

Ask

Analysis Goals:

In this phase of the analysis, the goal is to identify and clarify the key questions that will drive the data exploration and analysis. These questions are designed to uncover insights into smart device usage trends that are applicable and beneficial to Bellabeat's marketing strategy and product development.

1. **Trend Analysis**: What are the prevailing trends in smart device usage, especially in areas relevant to Bellabeat's products (such as activity tracking, sleep monitoring, and stress management)?

2. **Consumer Behaviour**: How do these trends reflect the behaviours and preferences of potential Bellabeat customers?

3. **Strategic Application**: How can these identified trends be leveraged to enhance Bellabeat's marketing strategies and product offerings?

4. **Market Opportunities**: Are there unexplored areas in the smart device market that Bellabeat can capitalise on, based on the trends and patterns observed in the data?

5. **Competitive Advantage**: How do Bellabeat's products compare with non-Bellabeat smart devices in terms of features and user engagement? What unique value propositions could Bellabeat offer to stand out in the market?

Approach to Addressing the Questions:

To address those questions, a systematic approach will be taken to explore the FitBit Fitness Tracker Data. It will involve:

- **Data Exploration**: Examining the dataset to understand its structure, content, and quality.
- **Data Analysis**: Conducting descriptive statistical analyses to uncover patterns and trends in the data.
- **Comparative Analysis**: Comparing findings from the FitBit data with Bellabeat's product features and market positioning.

Expected Outcomes:

The outcome of this phase will be a clear set of objectives for the data preparation, processing, and analysis stages. The insights gained from addressing these questions will directly inform the strategic marketing decisions and potential innovation areas for Bellabeat.

Stakeholder Engagement:

Throughout this phase, engagement with key stakeholders, including Bellabeat's marketing team and product development unit, would be vital. Their input would ensure that the analysis remains aligned with the company's business goals and market positioning strategies.

Prepare

Data Description and Storage

Source and Licensing:

The data for this analysis comes from the FitBit Fitness Tracker Data, provided under the CC0: Public Domain license by Möbius and available on Kaggle. This dataset comprises personal fitness tracker data from thirty Fitbit users, including minute-level output for physical activity, heart rate, and sleep monitoring.

Storage:

The dataset consists of 18 CSV files stored locally on my machine for easy access and analysis. The files are organised in a dedicated directory, ensuring efficient management and retrieval during analysis.

Data Organization and Structure

The dataset is divided into 18 CSV files, each representing different aspects of the users' fitness and health data. These files include information on daily activity, steps, sleep patterns, heart rate, and more. The data is organised in a long format, which is ideal for time-series analysis and tracking individual user metrics over time.

Data Credibility and Integrity

ROCCC Analysis:

- **Reliability**: With data from only 30 Fitbit users, there's a potential bias as this sample might not represent the broader population. Plans to augment this dataset with additional data sources should be considered to enhance its representativeness.
- **Originality**: The data is sourced from a third party (Möbius) and not directly from the service provider, which might affect its originality.
- **Comprehensiveness**: The dataset lacks demographic information such as age, gender, and specific device types, which limits the comprehensiveness of our analysis.
- **Currency**: The data was collected between December 3rd and 5th, 2016, which may impact its relevance to current market trends.
- **Cited**: The dataset is properly cited and sourced from a reliable crowdsourcing platform.

Addressing Limitations:

To mitigate these limitations, particularly the issues of small sample size and lack of comprehensiveness, one plans to:

- Explore additional datasets that could complement and enhance this dataset.

- Conduct a thorough exploratory data analysis to identify and understand any gaps or inconsistencies in the data.

Data Privacy and Ethics

The data was collected with the consent of the participants and is publicly available, respecting the privacy and ethical considerations of personal health data usage. However, it's essential to maintain this standard of privacy and ethics throughout the analysis process.

Preliminary Data Check

Before diving into detailed analysis, an initial check of the data's integrity and structure will be performed. This involves:

- Loading each CSV file into R to understand its structure and content.
- Checking for missing values, outliers, and inconsistencies.
- Summarising the key features of each dataset to ensure a good understanding of the available data.

Process

Data manipulation steps were excluded from this report. If you would like to examine them, you can find them in the R markdown file.

Analyse and Share Phase

Analysing Daily Activities Dataset
*Summary Statistics of Daily Activity Variables*

Variable	Mean	Median	Min	Max	SD
total_steps	7637.91	7405.50	0	36019.00	5087.15
total_distance	5.49	5.24	0	28.03	3.92
logged_activities_distance	0.11	0.00	0	4.94	0.62
very_active_distance	1.50	0.21	0	21.92	2.66
moderately_active_distance	0.57	0.24	0	6.48	0.88
light_active_distance	3.34	3.36	0	10.71	2.04
sedentary_active_distance	0.00	0.00	0	0.11	0.01
very_active_minutes	21.16	4.00	0	210.00	32.84
fairly_active_minutes	13.56	6.00	0	143.00	19.99
lightly_active_minutes	192.81	199.00	0	518.00	109.17
sedentary_minutes	991.21	1057.50	0	1440.00	301.27
calories	2303.61	2134.00	0	4900.00	718.17

The summary statistics of the 'daily_activity' dataset offer insightful revelations about daily physical activity and health metrics. On average, individuals take around 7,638 steps daily, but there's considerable variation, as indicated by the high standard deviation of about 5,087 steps. This variability is also reflected in the median of 7,405 steps, which is lower than the mean, hinting at skewed data with some individuals achieving significantly higher step counts. The total distance covered daily averages at approximately 5.49 km, but again, the range is broad, spanning from no activity to over 28 km.

In terms of active and sedentary behaviours, the data shows that, on average, individuals spend about 21 minutes in very active states and 14 minutes in fairly active states daily. However, the significantly higher median for lightly active minutes (199 minutes) compared to very and fairly active minutes suggests that people tend to engage more in light-intensity activities. Sedentary behaviour is predominant, with an average of 991 minutes (over 16 hours) spent sedentary, highlighting a potentially concerning trend in lifestyle habits.

The calories burned per day average around 2,304, with a wide range from zero to 4,900 calories, underscoring the diversity in individual metabolic rates and activity levels. These statistics reveal a pattern of moderate activity interspersed with significant periods of sedentary behaviour, emphasising the need for targeted interventions to promote more active and less sedentary lifestyles.

Plotting the correlation between variables
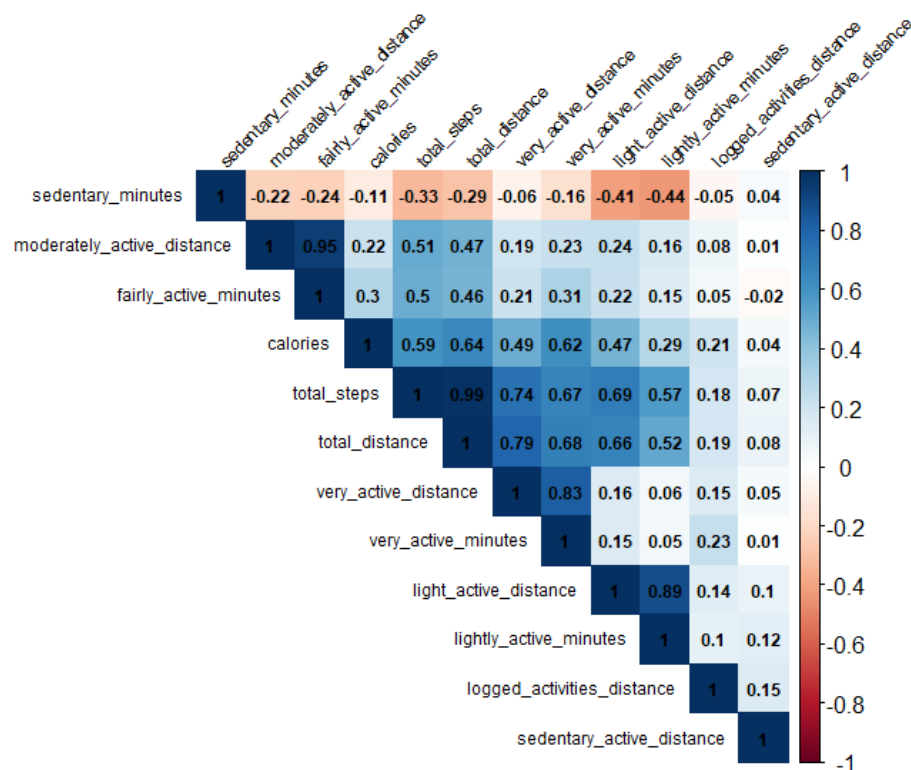
The correlation analysis reveals some notable insights. Particularly striking is the high correlation between total steps and distance, indicating a strong and direct relationship

between these two fundamental activity metrics, as expected. Also interesting is the negative correlation between sedentary minutes and lightly active minutes, highlighting a clear trade-off between active and inactive periods.

The relationship between Total Steps vs. Calories with a regression line

Total Steps vs. Calories
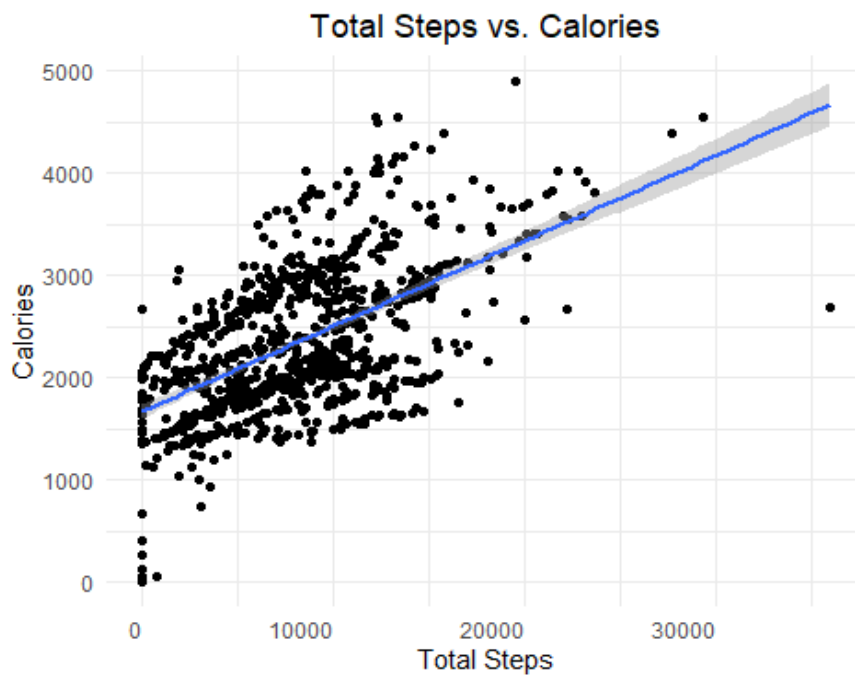
The positive linear relationship can be seen above again with a scatter plot.

Violin Plot of Total Steps by Day of the Week

Summary statistics of total steps by day of the week
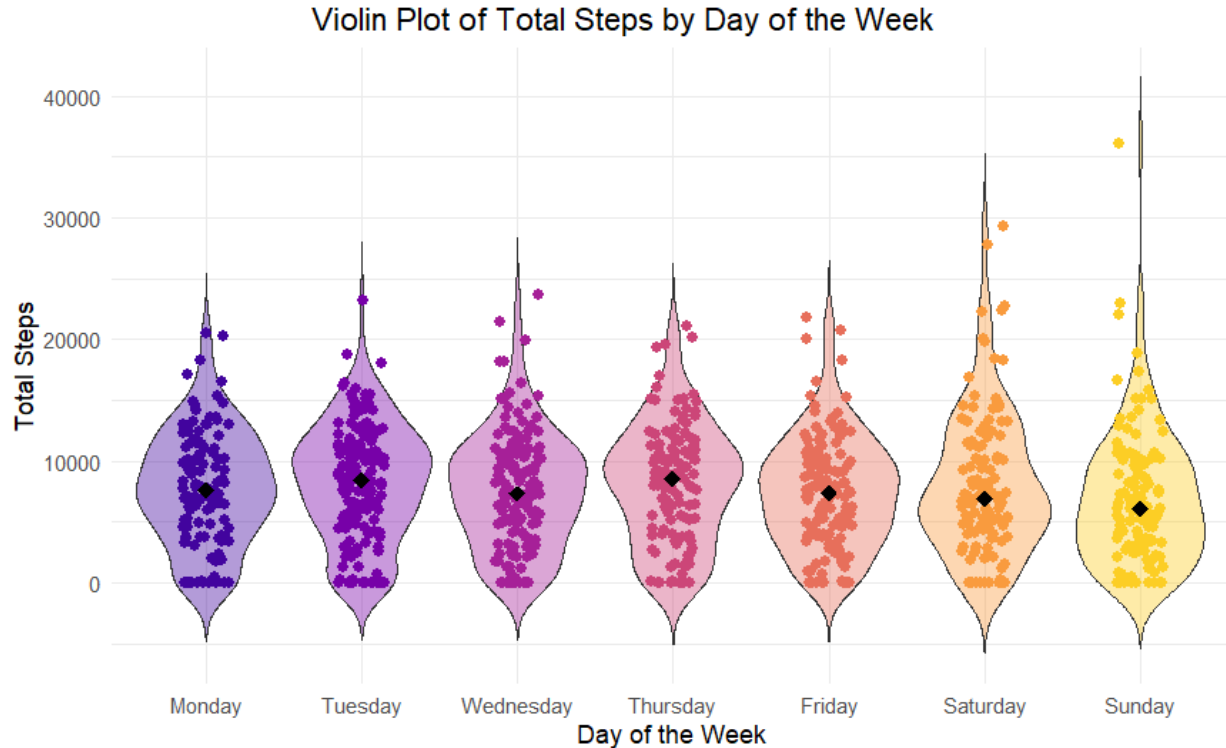
Day of week	Mean	Median	Min	Max	SD
Monday	7780.87	7625.5	0	20500	4717.82
Tuesday	8125.01	8411.0	0	23186	4833.81
Wednesday	7559.37	7317.0	0	23629	4810.10
Thursday	7405.84	7860.0	0	21129	5007.99
Friday	7448.23	7408.0	0	21727	4648.04
Saturday	8152.98	6946.0	0	29326	5944.30
Sunday	6933.23	6083.0	0	36019	5644.24

The summary statistics of total steps taken across different days of the week reveal a nuanced pattern of physical activity. Notably, Tuesdays and Thursdays stand out with higher median steps, suggesting more consistent activity levels on these days. In contrast, while weekends, particularly Sundays, register the lowest median steps, they exhibit the highest maximum steps, indicating sporadic bursts of high activity, possibly linked to leisure or recreational activities.

This data suggests that while weekdays, especially Tuesdays and Thursdays, are characterised by steady, moderate activity levels, weekends are marked by greater variability in physical activity. The high standard deviations on weekends, especially Saturdays, further underscore this inconsistency. These insights highlight an opportunity

for targeted health and wellness interventions. Encouraging consistent activity throughout the week, particularly on weekends, could be beneficial. Additionally, since daily step counts do not consistently reach the recommended 10,000 steps, integrating motivational features or personalised health recommendations into the app could nudge users towards more active lifestyles. Understanding these patterns and incorporating demographic data for a more profound customer segmentation analysis could lead to more effective, personalised user engagement strategies.

Analysing Heart Rate Dataset

Summary statistics for heart rate values

ID	Mean	Median	Min	Max	SD
2022484408	80.24	76	38	203	17.60
2026352035	93.78	95	63	125	12.61
2347167796	76.72	73	49	195	15.52
4020332650	82.30	83	46	191	15.91
4388161847	66.13	62	39	180	15.85
4558609924	81.67	80	44	199	14.13
5553957443	68.63	64	47	165	14.82
5577150313	69.56	62	36	174	20.96
6117666160	83.75	84	52	189	14.12
6775888955	92.03	91	55	177	17.60
6962181067	77.72	73	47	184	17.00
7007744171	91.12	90	54	166	13.99
8792009665	72.49	70	43	158	13.71
8877689391	83.61	72	46	180	30.60

Heart Rate Distribution by ID
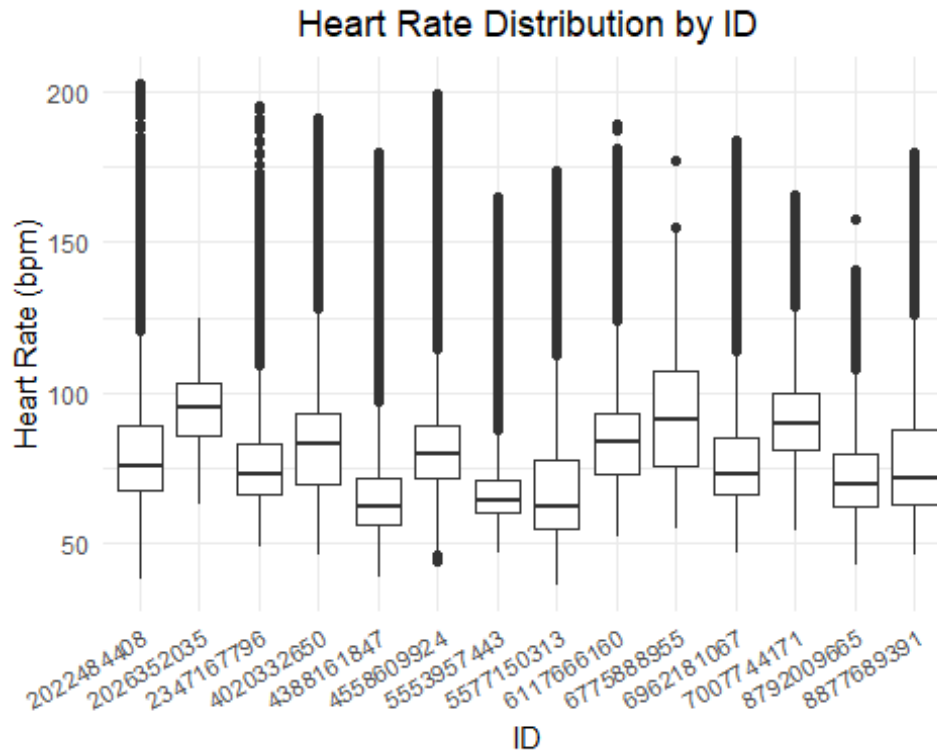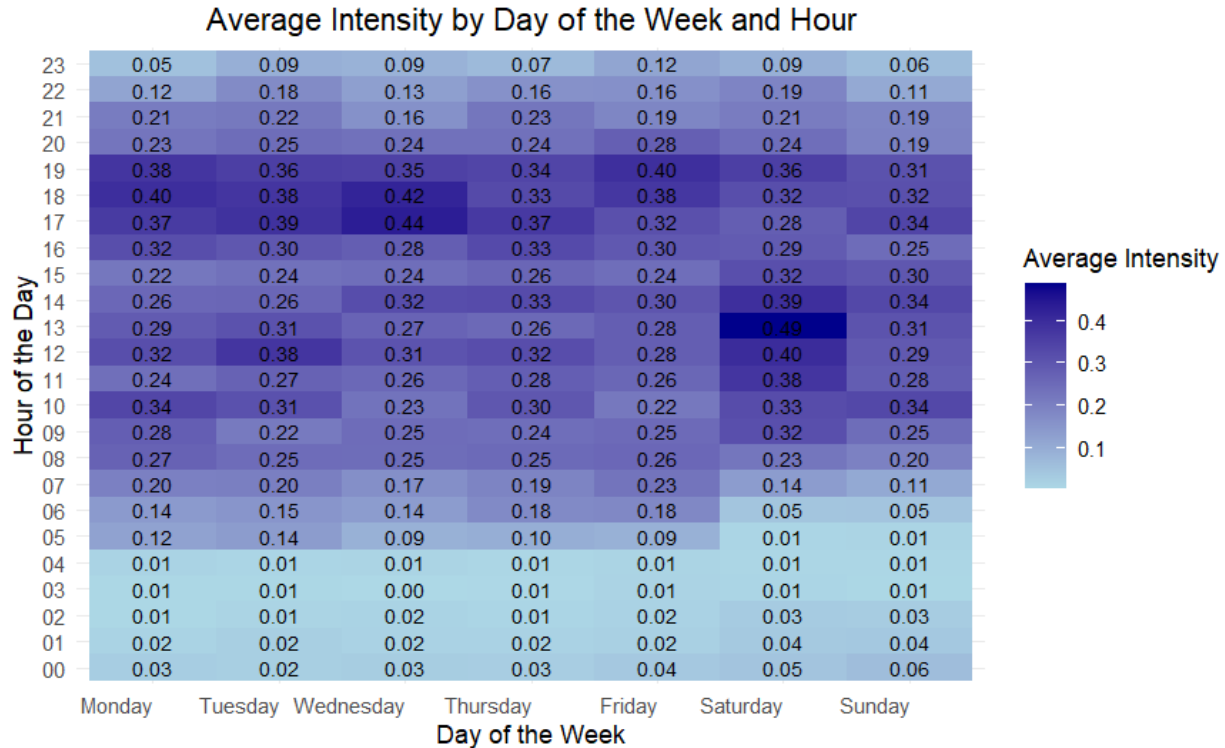

The summary statistics of heart rate by ID illustrate diverse physiological patterns among individuals. Mean heart rates range from approximately 66 to 93 bpm, indicating varied resting or active states. The closely aligned means and medians suggest symmetrical distributions for most individuals. The extensive range in minimum and maximum heart rates, from as low as 36 bpm to as high as 203 bpm, reflects the breadth of physical states, from rest to high activity. The standard deviations, varying significantly across IDs, highlight the distinct fluctuations in heart rate each person experiences.

Overall, the summary table and the boxplot suggest that each individual's heart rate pattern is unique and potentially influenced by various factors, including physical activity, overall health, and lifestyle habits. Such insights can be invaluable for personalised health monitoring and targeted interventions.

Analysing Intensities Data

Average Intensity by Day of the Week and Hour

Hour	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
23	0.05	0.09	0.09	0.07	0.12	0.09	0.06
22	0.12	0.18	0.13	0.16	0.16	0.19	0.11
21	0.21	0.22	0.16	0.23	0.19	0.21	0.19
20	0.23	0.25	0.24	0.24	0.28	0.24	0.19
19	0.38	0.36	0.35	0.34	0.40	0.36	0.31
18	0.40	0.38	0.42	0.33	0.38	0.32	0.32
17	0.37	0.39	0.44	0.37	0.32	0.28	0.34
16	0.32	0.30	0.28	0.33	0.30	0.29	0.25
15	0.22	0.24	0.24	0.26	0.24	0.32	0.30
14	0.26	0.26	0.32	0.33	0.30	0.39	0.34
13	0.29	0.31	0.27	0.26	0.28	0.49	0.31
12	0.32	0.38	0.31	0.32	0.28	0.40	0.29
11	0.24	0.27	0.26	0.28	0.26	0.38	0.28
10	0.34	0.31	0.23	0.30	0.22	0.33	0.34
09	0.28	0.22	0.25	0.24	0.25	0.32	0.25
08	0.27	0.25	0.25	0.25	0.26	0.23	0.20
07	0.20	0.20	0.17	0.19	0.23	0.14	0.11
06	0.14	0.15	0.14	0.18	0.18	0.05	0.05
05	0.12	0.14	0.09	0.10	0.09	0.01	0.01
04	0.01	0.01	0.01	0.01	0.01	0.01	0.01
03	0.01	0.01	0.00	0.01	0.01	0.01	0.01
02	0.01	0.01	0.02	0.01	0.02	0.03	0.03
01	0.02	0.02	0.02	0.02	0.02	0.04	0.04
00	0.03	0.02	0.03	0.03	0.04	0.05	0.06

Average Intensity: 0.4, 0.3, 0.2, 0.1

The analysis of hourly intensities across different days of the week reveals distinct peaks in user activity, particularly in the evening, especially between 5 PM and 7 PM on Wednesdays, indicating a midweek surge in physical activities, possibly due to post-work exercise routines. Additionally, activity levels on Saturday around noon are notably higher than on other days, suggesting that users might engage in outdoor activities or workouts during weekend leisure time. This information benefits Bellabeat, as it can tailor its engagement strategies by sending targeted reminders or motivational content during these peak hours. Encouraging users to maintain or increase physical activities during these times can optimise app engagement and promote a healthier lifestyle, especially for those with a tendency towards sedentary weekday behaviour.

Analysing Daily Sleep Dataset

Box plot for total minus asleep per capita
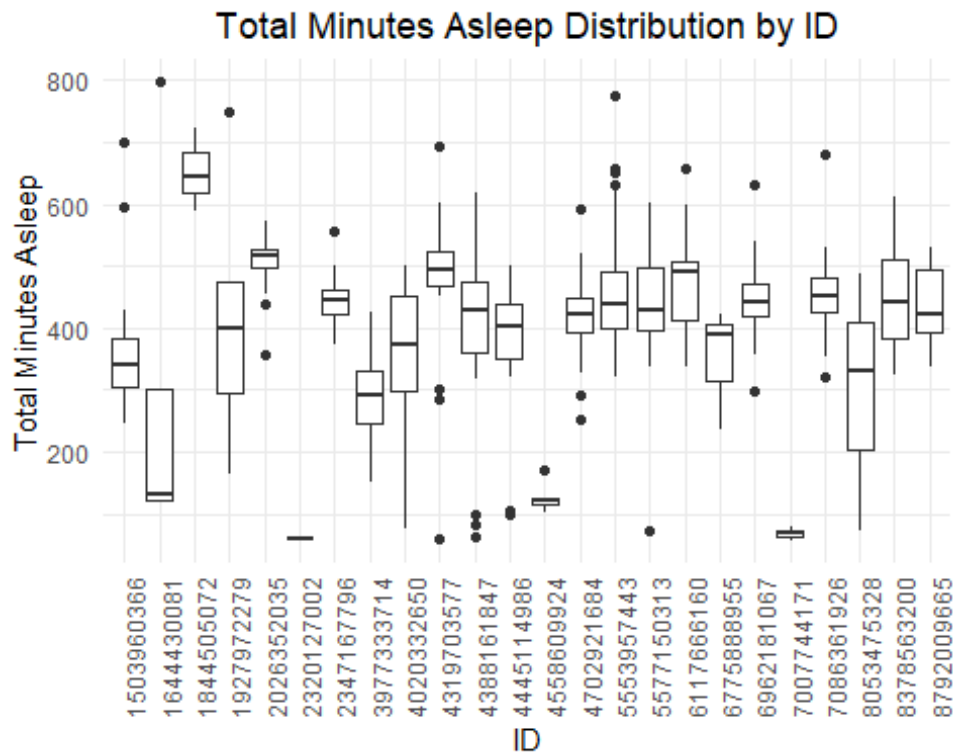
The summary statistics of total minutes asleep reflect significant variability in sleep patterns across the dataset. The data shows a broad range in average sleep duration, with some extremely low or high averages suggesting potential anomalies or data entry errors. Median values generally align with the means, indicating a primarily symmetrical distribution of sleep durations. However, the widespread minimum and maximum sleep times highlight the diversity in sleep habits, ranging from very short to extended durations. The variability in sleep patterns, as evidenced by the standard deviations, suggests differing levels of consistency among individuals, with some showing stable sleep routines and others experiencing more fluctuation. Overall, these statistics underscore the individuality of sleep habits, possibly influenced by a mix of personal, health, and environmental factors.

Violin Plot of Total Hours Asleep by Day of the Week
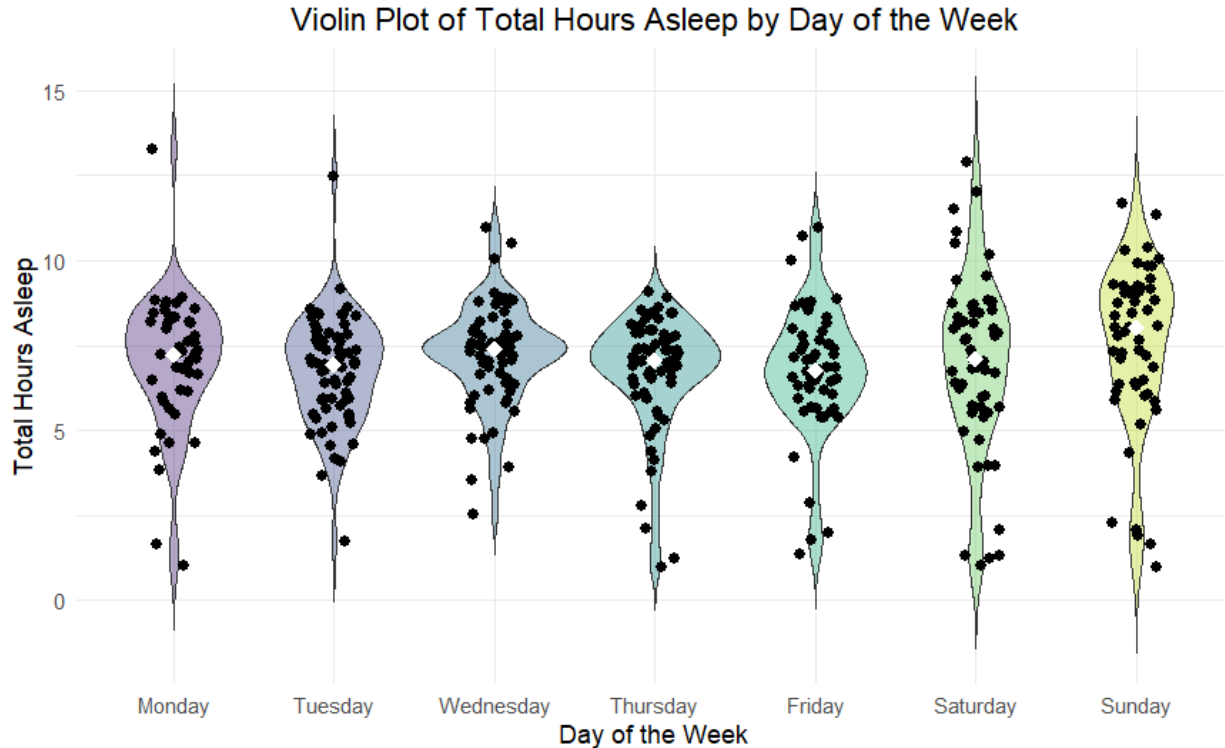
Summary statistics of total hours asleep by day of the week

Day of week	Mean	Median	Min	Max	SD
Monday	6.99	7.23	1.03	13.27	2.00
Tuesday	6.74	6.95	1.72	12.50	1.60
Wednesday	7.24	7.41	2.53	10.97	1.50
Thursday	6.69	7.06	0.98	9.08	1.71
Friday	6.76	6.75	1.37	10.97	1.88
Saturday	6.98	7.10	1.02	12.92	2.59
Sunday	7.55	8.02	0.97	11.67	2.42

The sleep data reveals a distinct pattern throughout the week, with the median sleep duration indicating that individuals tend to sleep less on Tuesdays and Thursdays, averaging around 6.95 and 7.06 hours, respectively. These days stand out as the ones with the shortest median sleep times, suggesting midweek nights might be the busiest or most sleep-challenged for many individuals. In contrast, Sundays show a marked increase in sleep duration, with the highest median sleep time of 8.02 hours, indicating that people tend to catch up on sleep during the weekend.

Relationship between Minutes Asleep and Sedentary Minutes
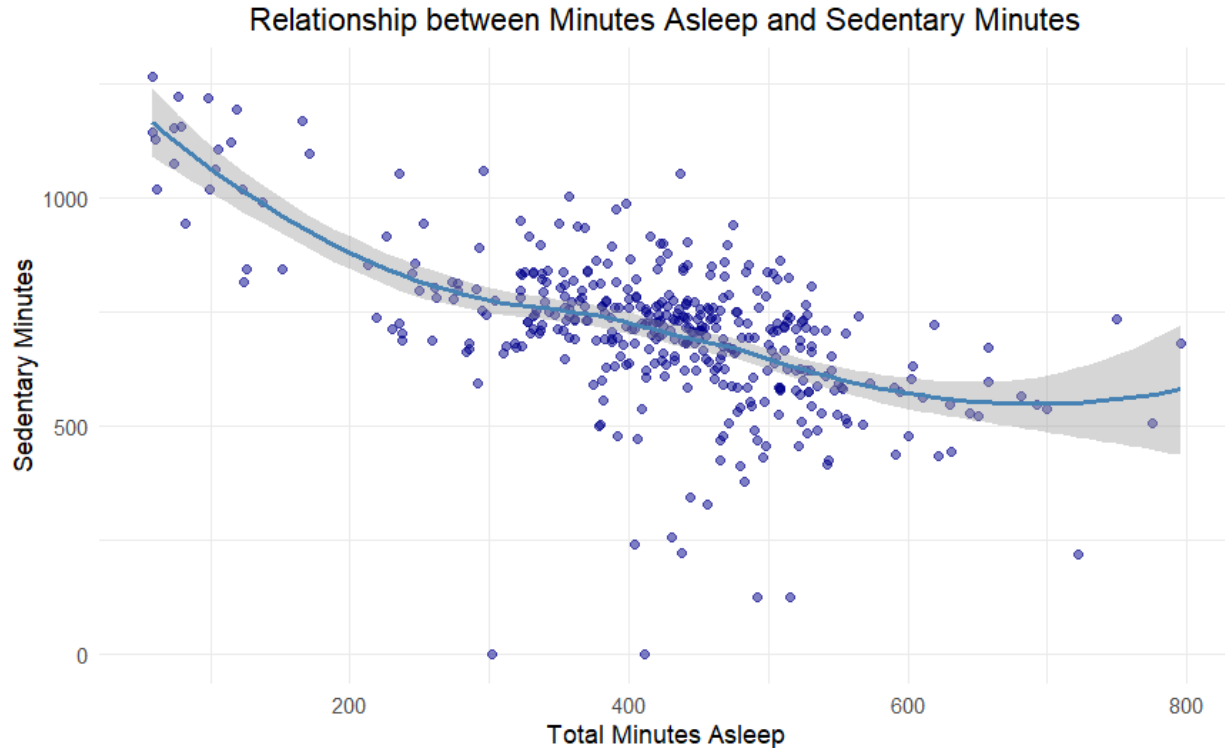
From the plot, I can observe that individuals with fewer minutes of sleep tend to have higher sedentary minutes, which could indicate less restful sleep or shorter sleep duration correlating with longer periods of inactivity during waking hours. On the other hand, those with more minutes asleep tend to have fewer sedentary minutes, suggesting that better-rested individuals might be more active.

Act

Strategic Recommendations

Based on the analysis of the Fitbit Fitness Tracker Data, several actionable strategies for Bellabeat can be proposed:

1. **Personalised Sleep Insights**: Offer personalised sleep analysis and improvement recommendations, as our data indicates a broad range of sleep behaviours and a potential need for better sleep management.

2. **Sedentary Behaviour Alerts**: Implement alerts to encourage movement after periods of inactivity, especially during high sedentary times.

3. **Optimize Engagement for Peak Activity Hours:** Create engagement strategies and motivational content targeting peak activity periods, like weekday evenings and Saturday Noons.

4. **Targeted Wellness Content**: Develop interventions for consistent physical activity throughout the week, leveraging insights from daily activity patterns.

5. **Enhanced Data Collection**: Expand data collection efforts to include demographic information such as age, gender, and lifestyle factors. This enriched dataset will allow for a more segmented and targeted analysis, enabling Bellabeat to tailor its products and marketing strategies to specific consumer groups and individual preferences.

Possible Implementation Plan

1. **Product Development**: Integrate the insights into the product roadmap, focusing on feature updates aligning with the identified activity and sleep patterns.

2. **Marketing Campaigns**: Develop campaigns highlighting the benefits of consistent activity and good sleep, showcasing how Bellabeat products support these goals.

3. **Customer Engagement**: Use the data to inform customer engagement strategies, such as personalised notifications and rewards for reaching activity and sleep targets.

4. **Continuous Monitoring**: Establish a framework for ongoing data analysis to monitor the effectiveness of the implemented strategies and make adjustments as needed.

Conclusion

The analysis of the Fitbit Fitness Tracker Data has yielded several actionable insights that can inform Bellabeat's strategic decisions regarding product development and marketing. I've identified key patterns in physical activity, sleep, sedentary behaviour and peak activity hours that can guide the creation of personalized user experiences and health-driven community engagement.

However, it is vital to acknowledge the limitations of this analysis. The dataset used is relatively small and dated, as it represents a sample from 2016 and may not reflect current trends or the broader demographic Bellabeat targets. The absence of demographic information within the dataset limits the depth of our customer segmentation and the personalisation of insights. Additionally, the data is sourced from a third party and may not fully represent the user engagement that Bellabeat's products would elicit.

Despite these limitations, the strategic recommendations offer a foundation for Bellabeat to enhance its user engagement and product features. Bellabeat must continue investing in data collection and analysis, mainly gathering more current and comprehensive data that includes demographic details. This will enable Bellabeat to refine its strategies, ensuring they are based on accurate and representative insights. Bellabeat can maintain a competitive edge in the health and wellness technology market by continually adapting to its users' evolving needs and preferences.