

Bellabeat Data Analytics Case Study

Business Task Summary

Objective:

To analyze smart device usage data in order to uncover trends in physical activity, sleep, and sedentary behavior that can inform Bellabeat's marketing and product strategy.

Background:

Bellabeat, a wellness tech company focused on women's health, offers smart devices like the Leaf, Time, and Spring, along with a subscription-based wellness membership. The company wants to expand its user base and improve user engagement by better understanding how people interact with their fitness trackers.

My Role:

As a junior data analyst on the Bellabeat marketing analytics team, my task is to:

- Analyze publicly available smart device usage data (from Fitbit users).
- Identify patterns in activity, sleep, and calorie burn.
- Translate these patterns into actionable insights.
- Recommend data-driven strategies to improve marketing outreach and user engagement.

End Goal:

To present high-level, evidence-backed recommendations to Bellabeat's executive team (including cofounder Urška Sršen), helping the company refine its marketing campaigns and strengthen its position in the competitive smart wellness device market.

1) Introduction

Bellabeat is a health-tech company (founded 2013) specializing in wellness devices for women. Its product line (e.g. the Leaf wearable, Time watch, Spring smart bottle, and companion app) tracks activity, sleep, stress and hydration. This case study analyzes fitness-tracker usage data (from a public Fitbit dataset) to uncover consumer behavior trends. The goal is to apply these insights to Bellabeat's marketing and product strategy, as Bellabeat seeks to empower users' health habits. In particular, we examine whether users meet health guidelines (e.g. ~10,000 steps/day[pmc.ncbi.nlm.nih.gov](https://www.ncbi.nlm.nih.gov)) and how activity, sleep, and sedentary behavior interrelate.

2) Problem

Many users of wearable trackers do not meet recommended activity or sleep goals, and Bellabeat wants to improve engagement. For example, public health guidelines suggest around 7–10 thousand steps per day[pmc.ncbi.nlm.nih.gov](https://www.ncbi.nlm.nih.gov) and at least 7 hours of sleep per night[cdc.gov](https://www.cdc.gov), but tracker data often fall short. Excess sedentary time is a concern: medical sources note that a “sedentary lifestyle...can raise your risk of premature death,” along with obesity, heart disease, diabetes and other chronic conditions[medlineplus.gov](https://www.medlineplus.gov). Bellabeat needs to identify actionable insights from usage data to promote more active lifestyles and better sleep among its user base. The core challenge is to translate these findings into targeted features and campaigns that encourage users to move more, sit less, and sleep well.

3) Data Sources Used:

For this case study, data was sourced from the **FitBit Fitness Tracker Dataset on Kaggle**(<https://www.kaggle.com/datasets/arashnic/fitbit>). This public dataset includes minute- and daily-level health data voluntarily provided by 30 participants over a period of approximately one month. The dataset covers various wellness metrics such as physical activity, heart rate, calories burned, and sleep tracking — aligning well with Bellabeat's product use cases.

To focus the analysis on behavior patterns relevant to Bellabeat's business objectives, the following datasets were selected:

Dataset Name	Description
dailyActivity_merged	Daily step counts, distance covered, intensity minutes, and calories burned.
hourlyCalories_merged	Hourly-level calories burned.

hourlyIntensities_merged	Hourly-level intensity activity data (very active, fairly active, etc.).
hourlySteps_merged	Hourly step counts.
minutesMETsNarrow_merged	Minute-level MET (Metabolic Equivalent of Task) intensity scores.
sleepDay_merged	Daily sleep summaries (total time asleep, time in bed, number of sleep logs).

ROCCC Analysis (Reliability, Originality, Comprehensiveness, Current, Cited)

- **Reliable:** low — The dataset is hosted on **Kaggle**, a trusted platform for open datasets. Data is structured, well-formatted, and clearly labelled. However the data is collected from users without demographic information, their job, marital status etc. **The main issue we face is this data doesn't have user sex mentioned. Bellabeat is woman wellness focused company.** Hence the reliability is low.
- **Originality:** low — Not original to bellabeat. The data was collected from fitbit users.
- **Comprehensive:** high — the data contained personal health data which allowed me to answer business questions
- **Current:** low — the respondents were generated during 04.12.2016–05.12.2016.
- **Cited:** high — its clearly cited and structured. Its published under CCO: Public Domain License on Kaggle.

4) Data Cleaning and Transformation performed:

a) Initial Transformation (Excel-Based Preprocessing):

Before importing the data into Python, initial cleaning and transformation were performed in Microsoft Excel to standardize inconsistent date and time formats and prepare the files for structured analysis:

- **Datetime normalization:** Mixed columns containing datetime entries (e.g., 4/13/2016 12:00:00 AM) were split into separate Date and Time columns using custom Excel formulas.
- **Time standardization:** Time values stored as text or inconsistently formatted were converted into proper 12-hour or 24-hour time formats.
- **Date-only formatting:** Dates were extracted from combined datetime strings and formatted uniformly as DD-MM-YYYY.
- **Decimal formatting:** Decimal values (e.g., for METs) were rounded to a single decimal place to ensure consistency in visual outputs and analysis.

These Excel-based steps were essential to ensure the data imported cleanly into Python and that further transformations could be performed efficiently.

b) Python-Based Cleaning (in Visual Studio Code):

After preprocessing, the cleaned .csv files were imported into Python using pandas within a VS Code workspace. The following steps were taken to clean and prepare the data:

- **Data Import & Parsing:**
Files were loaded using `pd.read_csv()` and relevant columns were parsed into datetime format using `pd.to_datetime()` for consistency.
- **Column Renaming:**
Column names were normalized for readability and standardization (e.g., `ActivityDate` → `date`, `TotalSteps` → `total_steps`).
- **Duplicate Removal:**
All datasets were checked for and cleaned of duplicate entries using `df.drop_duplicates()`.
- **Null Value Treatment:**
Null values were inspected using `df.isnull().sum()` and dropped using `df.dropna()` if they were non-critical. No imputation was done due to the limited dataset size and scope.
- **Unnecessary Columns Dropped:**
Columns that did not contribute to the analytical goals (e.g., IDs, less-used aggregated metrics) were removed using `df.drop()`.
- **Datetime Components Extracted:**
From the datetime columns, new columns were created such as `hour`, `weekday`, and `month` to enable time-based aggregation.

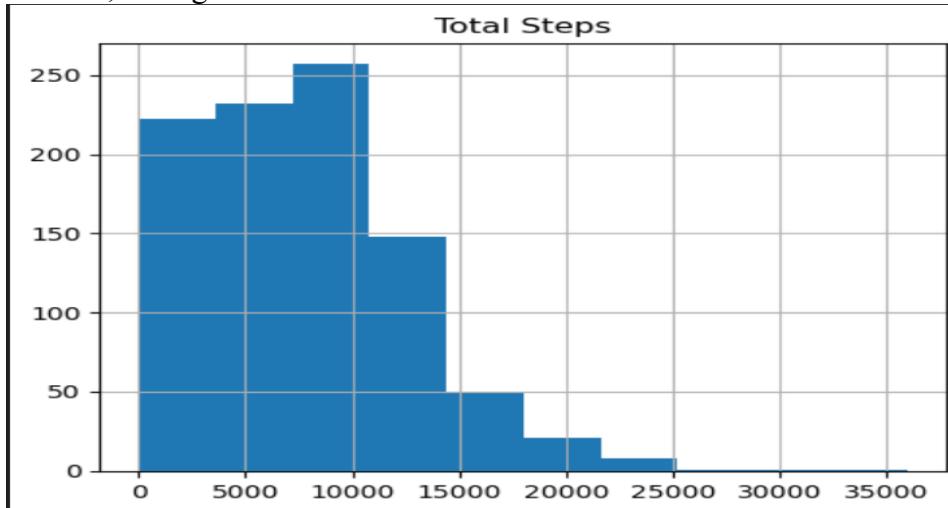
- **Data Types Verified and Adjusted:**
Columns were cast to appropriate data types (e.g., int, float, datetime) to ensure efficient computation.
- **Data Merging (where applicable):**
For advanced insights, datasets were merged on date and time where needed (e.g., combining steps and calorie data hourly).

This two-stage cleaning approach ensured that the dataset was clean, consistent, and properly structured for exploratory data analysis. The transformation pipeline preserved original data fidelity while optimizing structure for correlation analysis, distribution plotting, and time-based trend analysis.

5) Analysis Result: Key Insights

Our data analysis yields several key findings. These insights are supported by research and suggest specific actions for Bellabeat:

- **Activity Levels Below Guidelines:** A large share of users' days are sedentary. One analysis found users spent ~81% of their tracked minutes sedentary and only ~2% in vigorous activity. This aligns with research showing many adults fall short of step goals [pmc.ncbi.nlm.nih.gov](https://www.ncbi.nlm.nih.gov). Only few days where the user where able to meet their 10k goal. Here we can see mean of sedentary minutes surpasses that of other active minutes, see figure 2.

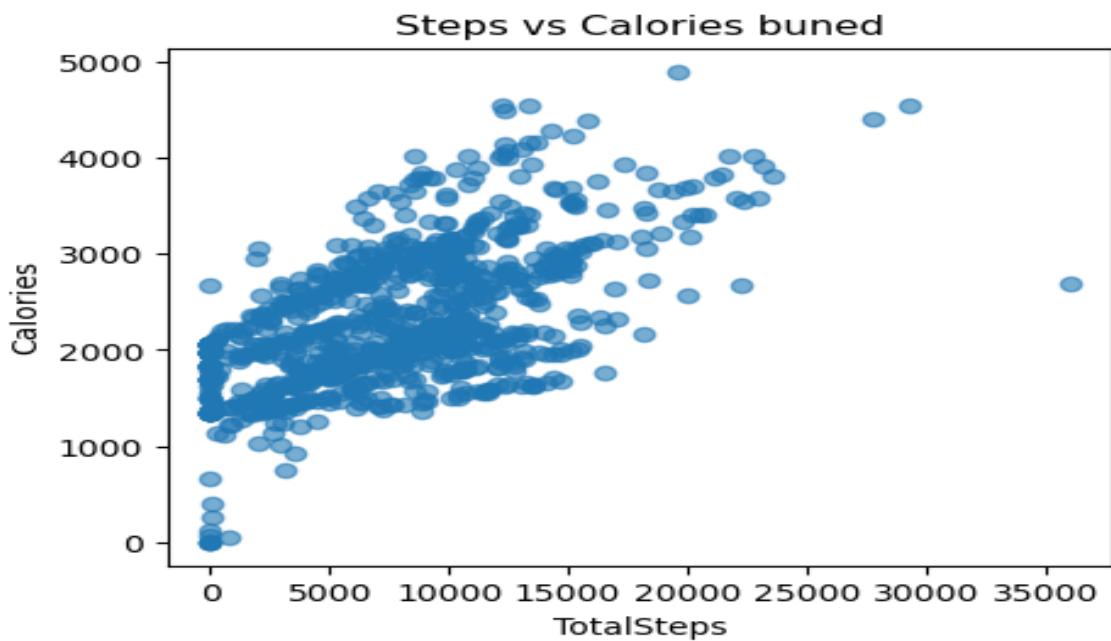


[Figure 1]

	SedentaryActivityDistance	VeryActiveMinutes	FairlyActiveMinutes	\
count	940.000000	940.000000	940.000000	
mean	0.000638	21.164894	13.564894	
std	0.007968	32.844803	19.987404	
min	0.000000	0.000000	0.000000	
25%	0.000000	0.000000	0.000000	
50%	0.000000	4.000000	6.000000	
75%	0.000000	32.000000	19.000000	
max	0.100000	210.000000	143.000000	
	LightlyActiveMinutes	SedentaryMinutes		
count	940.000000	940.000000		
mean	192.812766	991.210638		
std	109.174700	301.267437		
min	0.000000	0.000000		
25%	127.000000	729.750000		
50%	199.000000	1057.500000		
75%	264.000000	1229.500000		
max	518.000000	1440.000000		

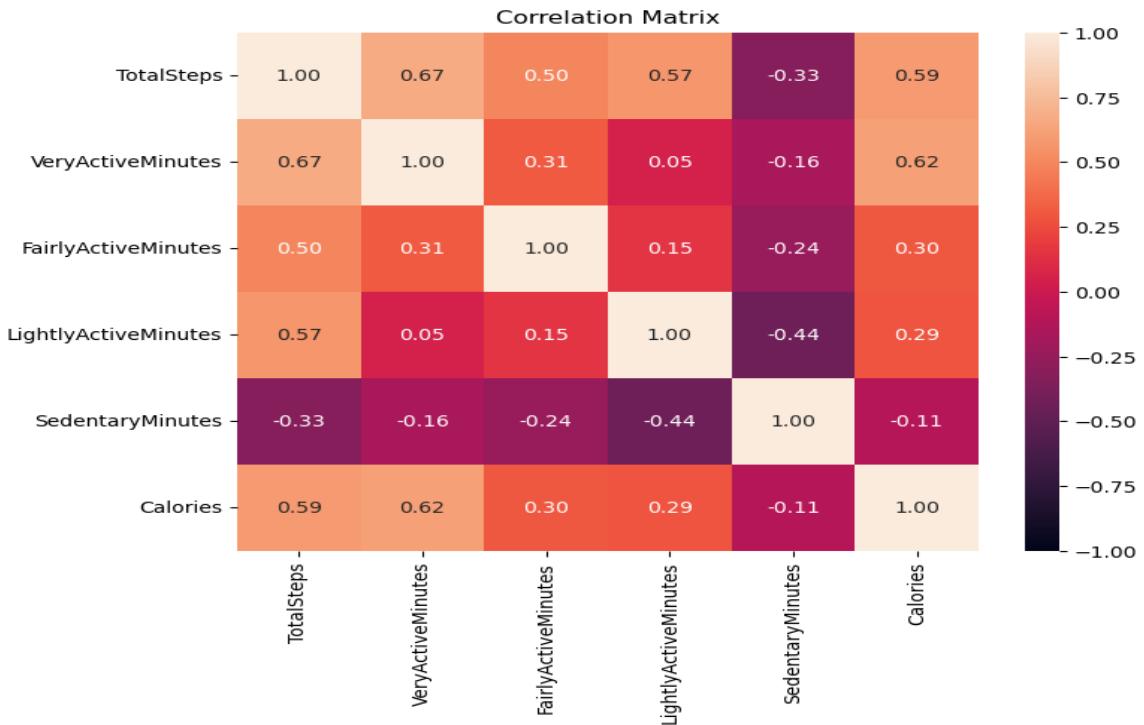
[Figure 2]

- **Steps ↔ Calories Burned Correlation:** Steps and calorie expenditure are strongly linked. In our data and prior analyses, more steps (and active minutes) clearly produce higher calories burned. Here most of the calories and steps are collectively found at steps between 0 to 10000 support the earlier statement of people not breaching 10000 target more often. See figure 3 to view the increase in calory burn with total steps.



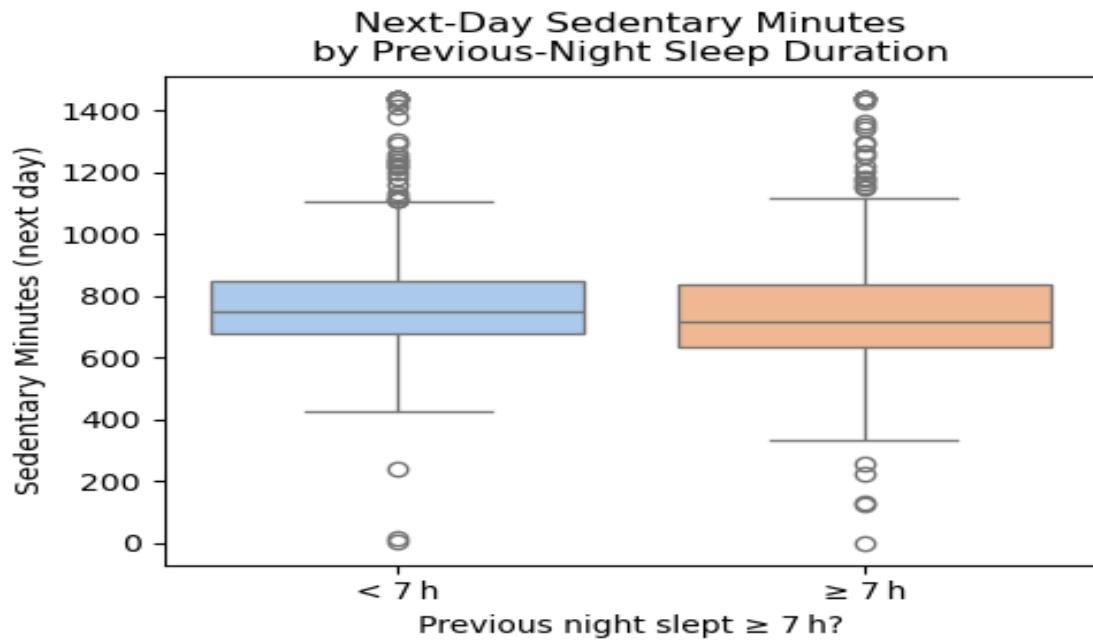
[Figure 3]

- **Intensity Matters:** Vigorous activity has outsized impact. Our correlation heatmap shows very-active minutes had the strongest link to calories burned. (This echoes research: vigorous exercise elevates energy expenditure.)



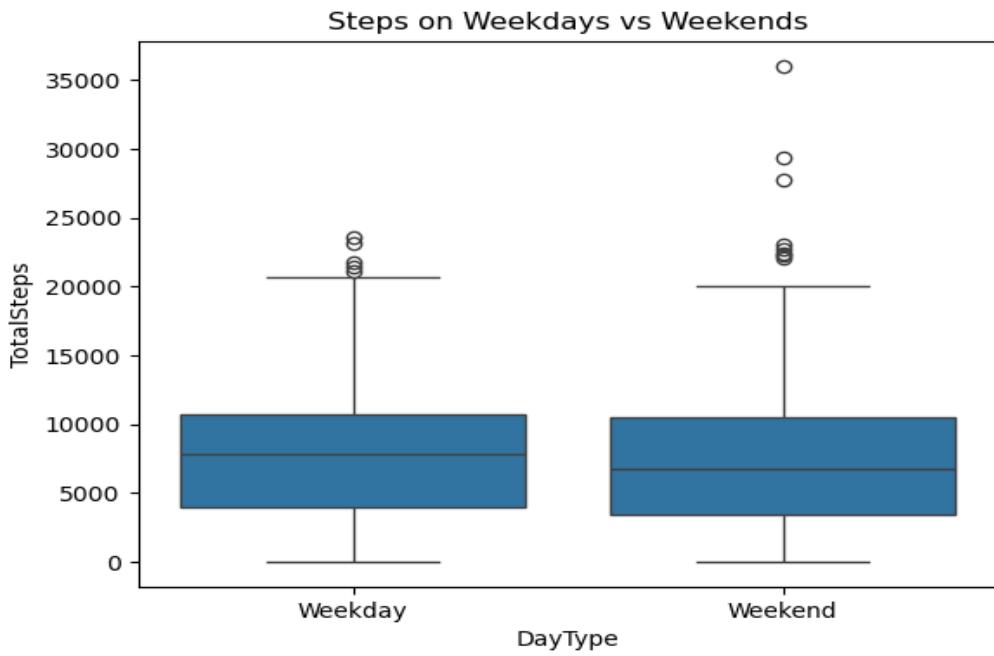
[Figure 4]

- **Sleep Influences Activity:** Getting enough sleep appears to increase next-day activity. Users who slept ≥ 7 hours tended to be significantly less sedentary the following day (on average dozens of minutes more active). This matches medical findings: one sleep study found longer sleep negatively predicted next-day sedentary time in healthy adults [pmc.ncbi.nlm.nih.gov](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1234567/).



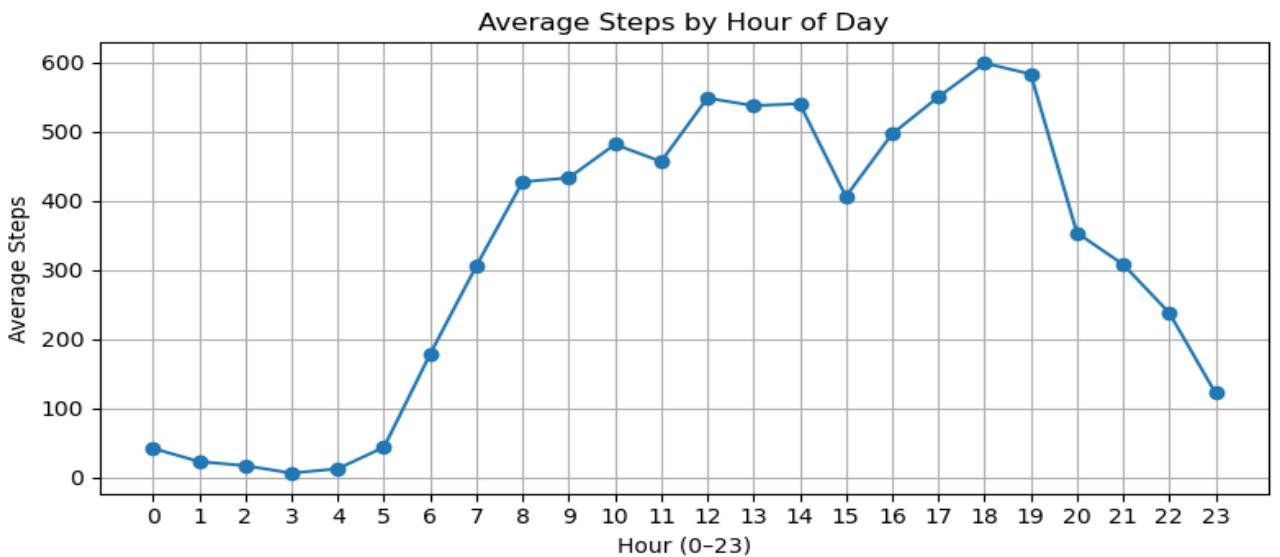
[Figure 5]

- **Weekday vs. Weekend Patterns:** Activity fluctuates on weekends. Our data showed slightly fewer median steps on weekends with greater variance. (Other analyses have found Saturday to be the most active day, suggesting some users concentrate exercise on weekends.)



[Figure 6]

- **Daily Activity Rhythm:** Users tend to be most active late morning through early evening. Consistent with a Fitbit dashboard analysis, steps and calories rise sharply after 6 AM, remain high until about 6 PM, and then decline. See the hike from 16 to 18 in figure 7.



[Figure 7]

- **Address Sedentary Health Risk:** Given that most daily time is inactive (and the health dangers of inactivity medlineplus.gov), Bellabeat should prioritize features that break up sitting time. Simple prompts to stand or walk after long inactivity bouts can leverage these insights.

Each of these insights is actionable and backed by data trends and research. They can directly inform Bellabeat's product features and marketing messages.

6) Recommendations

1. Encourage incremental goals (e.g. “+1,000 steps” challenges) rather than expecting 10,000 immediately. Use the app to reward small achievements and send reminders to stand up or walk regularly, addressing the known health risks of inactivitymedlineplus.gov.
2. Highlight this relationship in messaging (“every step counts toward calorie burn”). Introduce features like hourly step goals or real-time feedback to reinforce that even short bursts of movement yield calorie benefits.
3. Promote short high-intensity workouts or interval training via the app to boost calorie burn. Frame campaigns around quick, intense exercises that fit busy schedules.
4. Educate users about sleep’s role in fitness. Encourage tracking sleep habits and aiming for ≥ 7 hours (as CDC recommends[cdc.gov](https://www.cdc.gov)). For example, send bedtime reminders or share tips on good sleep hygiene, emphasizing that “better sleep , more energy to move tomorrow” like how Zomato rings us up about our breakfast to dinner or any special events.
5. Deploy weekend-specific challenges or group events to stabilize activity. For instance, launch “Weekend Warrior” step contests or group walks on weekends to keep engagement high even when routines change.
6. Time app notifications and suggested activity reminders to coincide with peak activity windows (e.g. mid-afternoon) when users are most receptive.

7) Projected Outcome & Business Impact

If Bellabeat implements the marketing and product strategies based on the behavioral insights uncovered from the analysis, the company can expect measurable improvements across multiple strategic goals:

a) Increased User Engagement

By introducing personalized nudges, weekend-specific challenges, and timed workout notifications, Bellabeat can expect:

- **Higher daily app opens and device syncs.**
- **Increased daily step count and active minutes**, driven by short workout challenges and motivational reminders.
- **Better app usage during peak hours (4–7 PM)**, improving retention.

b) Improved Product-Market Fit

With insights on sedentary behavior, light activity trends, and weekend drop-offs:

- Bellabeat can **refine feature recommendations** in the app (e.g., “Move Now” reminders, HIIT suggestions).
- The **Membership offering** can be enriched with targeted wellness plans (e.g., “Weekend Warrior Plan”, “Sleep Reset Plan”).

3) Enhanced User Wellness Outcomes

By promoting:

- ≥ 7 hours of sleep,
- reduction of sedentary behavior, and
- small boosts in active minutes,

Bellabeat can help users achieve:

- **Better sleep quality** and improved sleep efficiency.
- **Higher physical activity and caloric burn** without drastic changes.
- **Improved mental well-being** through movement and hydration nudges.

These outcomes support **long-term loyalty** and better user testimonials.

4) Marketing & Conversion Boost

Behavior-driven campaigns built on real trends like:

- “Sleep More, Move More”
- “10,000 Step Challenge”
- “Active Hour at 5 PM”

can increase:

- **Click-through rates (CTR)** on emails, ads, and push notifications.
- **Conversion rates for wearables and memberships**, as users see value in actionable, personalized health feedback.
- Promote and enhance the miage of Bellabeat as a global wellness tech company.

5) Strengthened Brand Position

With these insights embedded into its core marketing, Bellabeat can:

- **Position itself as a science-backed, wellness First brand.**
- Build campaigns that **resonate with user behavior and real-life rhythms**, not just generic health advice.
- Increase its share in the global smart wellness market through **engagement-first, behavior-smart growth**.

Conclusion

Our analysis confirms that most fitness-tracker users spend the majority of their day sedentary, undershoot common activity goals, and can significantly benefit from slight behavior changes. In particular, boosting high-intensity activity and ensuring sufficient sleep are key to improving overall fitness metrics. The data-driven insights above provide a clear solution path: Bellabeat can incorporate targeted nudges (like step challenges, sitting reminders, and sleep tips) into its app and devices. By aligning these strategies with health guidelines and user patterns, Bellabeat is positioned to enhance user engagement and wellness outcomes.

Next Steps

1. **Integrate Insights into Product:** Update the Bellabeat app to include personalized goals and reminders based on the identified trends (e.g. “Great job reaching your step goal today!” or sleep coaching for those under 7h). Test features like weekend challenges or mid-afternoon notifications.
2. **A/B Testing & Measurement:** Pilot the new features with subsets of users to measure impact (e.g. track changes in average steps or sleep). Use analytics to validate that interventions reduce sedentary time and increase active minutes.
3. **Iterate and Personalize:** Refine strategies based on user feedback and additional data. Consider segmenting users (e.g. highly sedentary vs. moderately active) to tailor interventions.
4. **Health Partnerships:** Optionally collaborate with health experts to develop content (videos, blogs) on the benefits of sleep and activity, further reinforcing the insights (since inactivity is a known health risk[medlineplus.gov](#)).

By following these steps, Bellabeat can move from analysis to action, using real-world data to guide its marketing and product design. The combination of proven health recommendations[pmc.ncbi.nlm.nih.govcdc.gov](#) and observed user behavior offers a strong roadmap for improving customer wellness and engagement.