

1. Executive Summary

Company: Bellabeat

Analyst: Shivali Muthukumar

Tools Used: SQL (DB Browser for SQLite), Python (Pandas), Tableau Public

This project analyses open-source Fitbit user data to uncover **how consumers use their smart fitness devices** and identify behavioural patterns that can inform Bellabeat's future product and marketing strategies.

Through SQL queries, Python analysis, and Tableau visualization, I have explored user activity, calories burned, sleep habits, and overall engagement patterns. The final insights help Bellabeat understand what habits correlate most strongly with active lifestyles and health outcomes.

2. Business Task

2.1 Goal:

To analyse smart device usage data and understand consumer activity trends, sleep habits, and energy expenditure patterns that can help **Bellabeat design and market** smarter wellness devices.

2.2 Key Questions:

1. How active are users on average?
2. Is there a correlation between physical activity (steps) and calories burned?
3. What are the average sleep durations among users?
4. How do different activity metrics relate to each other?

3. Data Sources

- Public Fitbit Fitness Tracker Data (collected via Amazon Mechanical Turk, 2016)
- Covers ~30 users over 2 months (April–May 2016)
- Data Files Used (18 CSVs including hourly/minute datasets for deeper granularity.):
 - dailyActivity_merged.csv
 - dailyCalories_merged.csv
 - dailyIntensities_merged.csv
 - dailySteps_merged.csv
 - sleepDay_merged.csv
 - weightLogInfo_merged.csv

3.1 Integrity Notes:

- No missing or null values found in major datasets.
- Dates standardized to YYYY-MM-DD.
- Minor inconsistencies (e.g., inactive days, missing distance units) were addressed during cleaning.

4. Tools and Methods

Tool	Purpose
SQL (SQLite)	Data exploration, aggregation, user activity summaries
Python (Pandas)	Data cleaning, merging, trend analysis, correlations
Tableau Public	Visual analytics and storytelling dashboard creation

5. SQL Analysis Overview

- **Unique Users:** 33
- **Average Daily Steps per User:** ~11,000
- **Top 10 Active Users:** ranged 9,000–16,000 steps/day
- **Average Steps by Weekday:** Activity peaks on **Tuesdays and Saturdays**
- **Correlation (Steps vs Calories): 0.59** — strong positive relationship
- **Average Sleep (Hours):** ~6.6 hours

→ **Insight:**

Users who are consistently active (more steps/day) also burn more calories. Sleep averages are within the healthy range, but few users maintain consistent patterns.

6. Python (Pandas) Analysis Summary

- **Data Cleaning:** Merged datasets by 'Id' and 'ActivityDate', standardized date formats, removed duplicates and zeros.
- **Correlation Analysis:**
 - Steps vs Calories → 0.592
 - Active Minutes vs Calories → 0.65
 - Sleep vs Steps → Weak negative (−0.08)

→ **Insight:**

Higher movement intensity (very active minutes) drives calorie burn more than total steps alone. Sleep time has minimal correlation with activity — showing users may not balance rest and activity optimally.

7. Tableau Visualizations & Insights

7.1 Visualization 1: Daily Steps Over Time

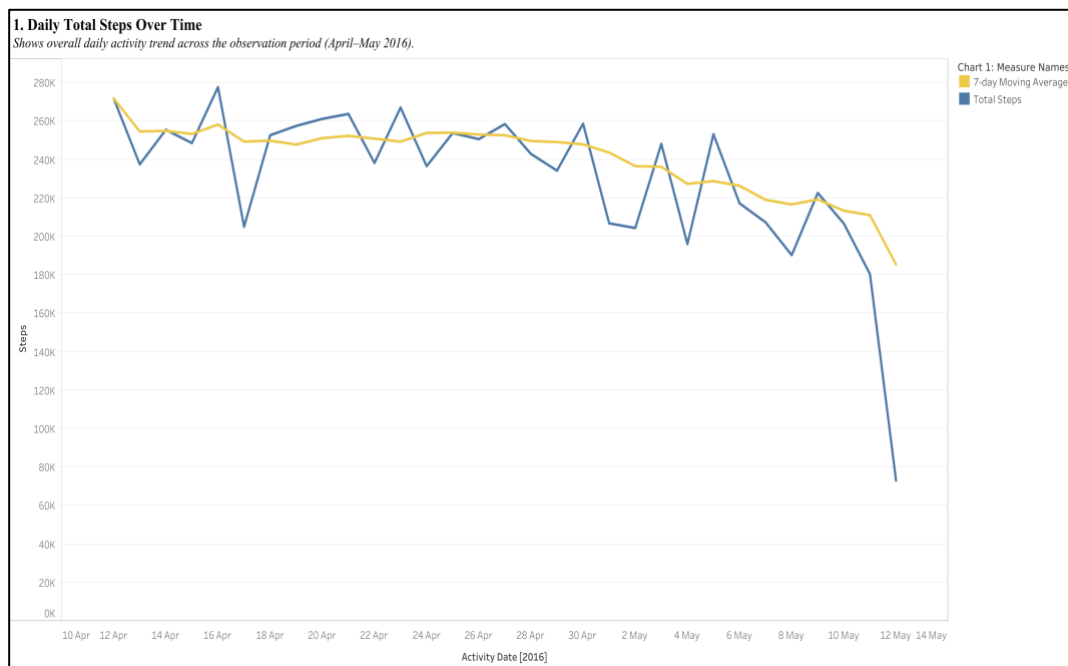


Figure 1.: Shows overall user activity trend from April–May 2016.

Insight: Slight decline in activity toward May. Weekend spikes suggest higher engagement on non-workdays.

7.2 Visualization 2: Calories Burnt vs Total Steps

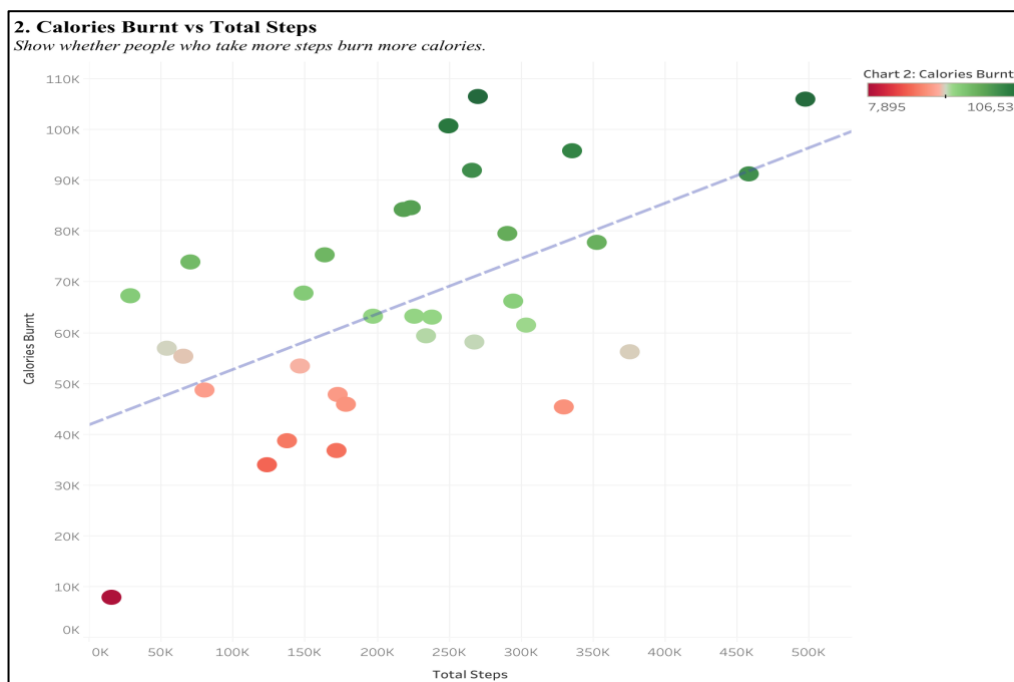


Figure 2.: Scatter plot showing the relationship between steps and calories.

Insight: Positive correlation confirms higher steps = higher calories burned ($R^2 = 0.326$, $p = 0.0005$).

7.3 Visualization 3: Distribution of Sleep Duration (Hours)

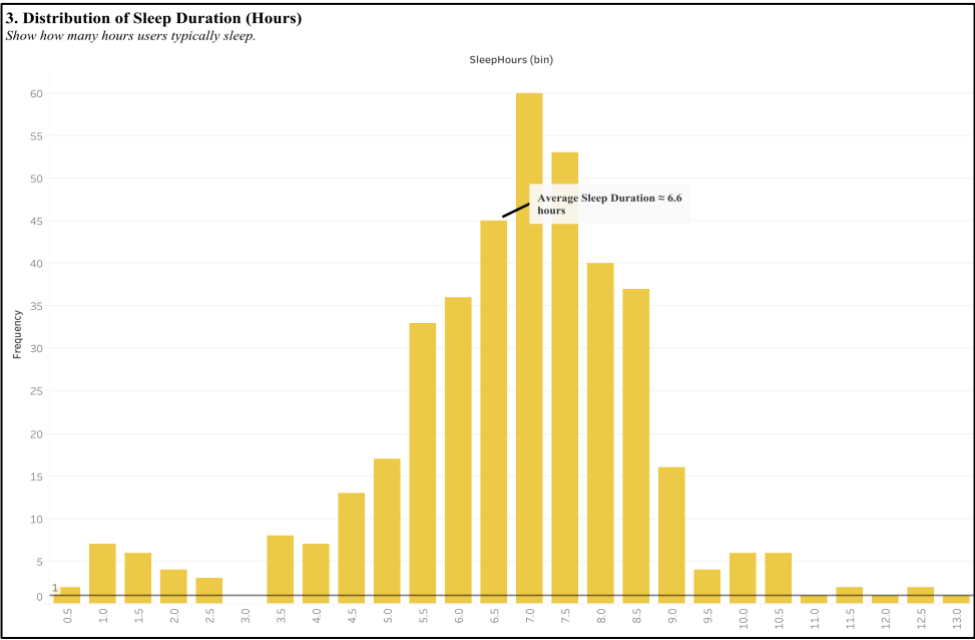


Figure 3.: Histogram of user sleep hours.

Insight: Most users sleep 6–7 hours per night; a few record irregular patterns or missing data. Average Sleep Duration ≈ 6.6 hours.

7.4 Visualization 4: Correlation Between Activity Metrics

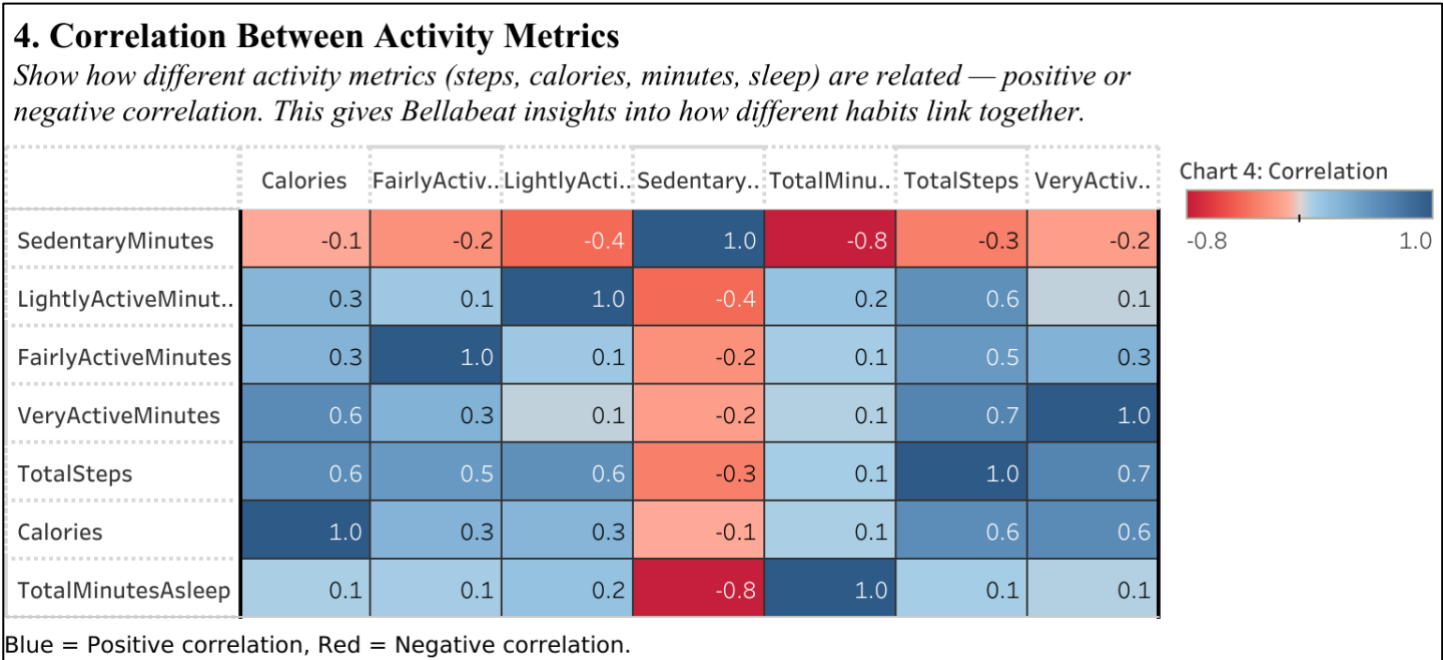


Figure 4.: Heatmap comparing steps, calories, sleep, sedentary, and active minutes.

Insight:

- Steps - Calories: strong positive (0.6)
- Sedentary - Sleep: weak negative
- Very Active Minutes - Total Steps: strongest relationship (0.7)

8. Key Insights Summary

Area	Finding	Business Interpretation
Activity Trends	Avg 11,000 steps/day	Users are moderately active but drop off on weekdays
Calories	Strong link with steps	Encourage consistent movement goals
Sleep	Avg 6.6 hrs/day	Users may not get sufficient rest
Correlations	Steps ↔ Calories ↔ Active minutes	Reinforces movement-based motivation features

9. Recommendations for Bellabeat

- 1. **Personalized Activity Targets:**
Integrate step-based calorie goals (e.g., “10,000 steps = 2000 calories/day”).
- 2. **Rest Balance Reminders:**
Include prompts for users showing high activity but low sleep.
- 3. **Weekend Challenges:**
Leverage weekend activity spikes with group events or badges.
- 4. **Smart Sleep Tracking:**
Provide tips when sleep duration drops below 6 hours.

10. Conclusion

Bellabeat can leverage these findings to enhance product engagement and promote healthier behaviour. The Fitbit dataset reveals that users are moderately active but inconsistent. Activity levels strongly influence calorie expenditure, while sleep patterns remain irregular.

This analysis demonstrates how smart device data can help Bellabeat’s design and marketing teams tailor features that encourage daily movement, recovery, and wellness tracking.

11. Final Tableau Dashboard

