

Bellabeat Case Study using Microsoft Excel, SQL and Power BI

How Can a Wellness Technology Company Play It Smart?



Introduction

In this case study, I analyze historical smart device usage data from Bellabeat, a high-tech wellness company specializing in health-focused products for women. The goal is to identify key trends in consumer behavior and provide data-driven recommendations to optimize Bellabeat's marketing strategy.

To answer the key business questions, I will follow the Google Data Analytics process: Ask, Prepare, Process, Analyze, Share, and Act.

The main tools I used are Microsoft Excel for merging table at ease, SQL for data processing and Power BI for data visualization.

◆ Quick Links:

- ✓ Data Source: Fitbit smart device dataset
- ✓ SQL Queries
- ✓ Data Visualizations: Power BI

A more detailed breakdown of the case study scenario is provided below, followed by my full report.

Scenario

As a junior data analyst on the Bellabeat marketing analytics team, I have been tasked with analyzing smart device usage data to uncover key trends and insights. Our team is responsible for collecting, analyzing, and reporting on consumer behavior to guide Bellabeat's marketing strategy.

Urška Sršen, Bellabeat's Chief Creative Officer, believes that analyzing consumer behavior can provide valuable insights into user habits and help the company optimize its marketing efforts. The goal is to explore how consumers interact with Bellabeat's smart devices and use these findings to develop a data-driven marketing strategy.

For this case study, I will:

- ✓ Analyze smart device usage data to identify consumer trends.
- ✓ Compare patterns in activity, sleep, stress, and hydration tracking.
- ✓ Discover how Bellabeat's marketing team can leverage these insights.
- ✓ Present high-level recommendations to the Bellabeat executive team.

By the end of this analysis, Bellabeat will have actionable insights that can shape its marketing strategy and support its continued growth in the wellness technology market.

Background

Bellabeat is a high-tech wellness company that designs smart health-focused products for women. Founded in 2013 by Urška Sršen and Sando Mur, the company has quickly positioned itself as a leading brand in the smart device market. Bellabeat offers a range of wellness products, including the Leaf wellness tracker, Time smartwatch, and Spring smart water bottle, all of which connect to the Bellabeat app to provide users with valuable health insights on activity, sleep, stress, hydration, and mindfulness.

Despite its success, Bellabeat aims to expand further and become a global leader in the smart device industry. To achieve this, the company heavily invests in digital marketing through platforms like Google Search, Facebook, Instagram, YouTube, and the Google Display Network. However, Bellabeat's leadership believes that a data-driven marketing approach could unlock new growth opportunities by better understanding how customers use their smart devices.

Phase 1. Ask – Define the Business Problem

Business Task:

Bellabeat wants to use **smart device data** to gain insights into how customers interact with their products and identify **opportunities for growth**. The marketing team needs data-driven recommendations to refine their **marketing strategy** and attract more customers.

Key Questions to Answer

1. What are the current **trends** in smart device usage?
2. How can these trends be **applied to Bellabeat customers**?
3. How can these insights **shape Bellabeat's marketing strategy**?

Stakeholders

- **Urška Sršen** (Co-founder & Chief Creative Officer) – Seeks **data-driven insights** to inform business strategy.
 - **Bellabeat Marketing Team** – Uses insights to **optimize digital marketing campaigns**.
 - **Bellabeat Executive Team** – Requires recommendations for **expanding market presence**.
-

Phase 2. Prepare – Understanding and Preparing the Data

Data Source

The dataset used for this analysis is the [FitBit Fitness Tracker Data](#), publicly available on Kaggle under a **CC0 Public Domain license**. It contains fitness tracking data from **30 Fitbit users**, including **daily activity levels, step counts, heart rate monitoring, and sleep patterns**.

Data Organization & Storage

- The dataset consists of multiple **CSV files**, each representing different fitness tracking metrics.
- It follows a **wide format**, where each row represents a single day's activity for an individual user.
- The data was **imported into SQL** for cleaning and analysis, and **Power BI** was used for visualization.

Data Credibility & Limitations

- The dataset is **limited to 30 users**, which may not fully represent **the broader population**.
- There is **potential bias**, as the data only includes users who consented to share their information.
- The dataset **lacks demographic details**, making it difficult to segment users by age, gender, or location.

Addressing Data Integrity & Limitations

- Verified data quality by checking for **missing values, duplicates, and inconsistencies**.
- Filtered out incomplete records to ensure accuracy in analysis.
- Considered the need for **additional datasets** to supplement missing demographic insights.

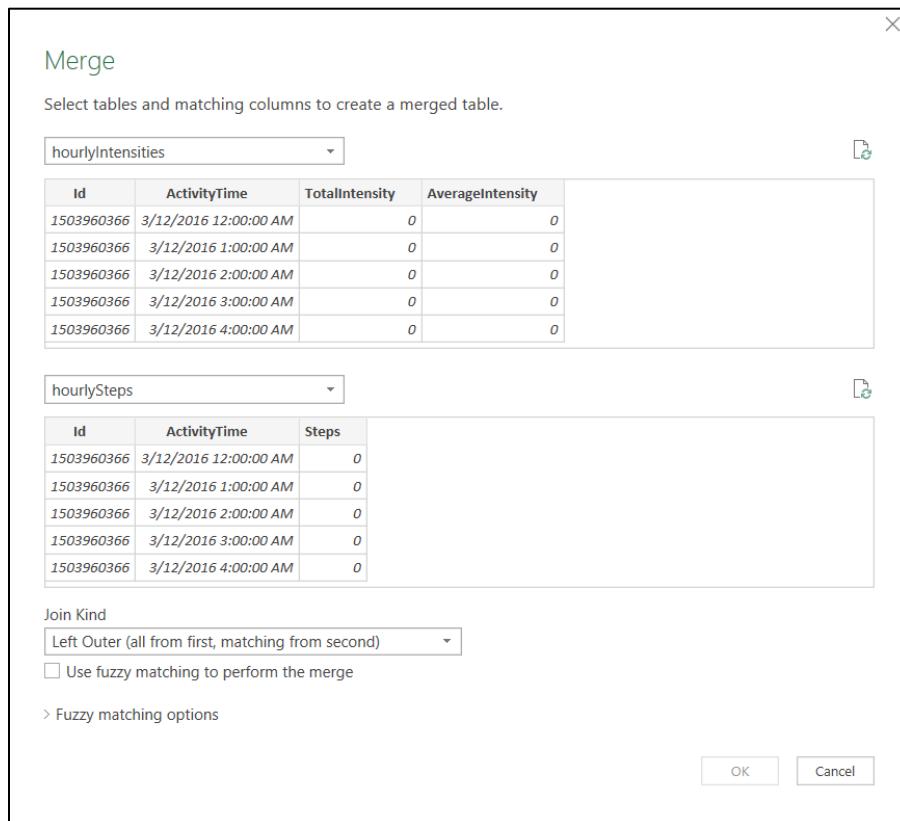
This dataset provides **valuable insights into daily activity, sleep, and heart rate trends**, helping to answer key business questions related to **user behavior and engagement with smart wellness devices**.

Phase 3. Process – Cleaning and Preparing Data for Analysis

Tools Used

To merge easily I first used **Microsoft Excel**. Then, for data cleaning and transformation, **MySQL** was chosen due to its ability to handle large datasets efficiently and perform complex data manipulations using SQL queries. The cleaned data will later be used for analysis and visualization in **Power BI**.

Merging Tables in Microsoft Excel



	A	B	C	D	E	F
1	Id	ActivityTime	TotalIntensity	AverageIntensity	Steps	Calories
2	1503960366	3/12/2016 0:00	0	0.00	0	48
3	1503960366	3/12/2016 1:00	0	0.00	0	48
4	1503960366	3/12/2016 2:00	0	0.00	0	48
5	1503960366	3/12/2016 3:00	0	0.00	0	48
6	1503960366	3/12/2016 4:00	0	0.00	0	48
7	1503960366	3/12/2016 5:00	0	0.00	0	48
8	1503960366	3/12/2016 6:00	0	0.00	0	48
9	1503960366	3/12/2016 7:00	0	0.00	0	48
10	1503960366	3/12/2016 8:00	0	0.00	0	48
11	1503960366	3/12/2016 9:00	1	0.02	8	49
12	1503960366	3/12/2016 10:00	26	0.43	551	89

Data Cleaning & Manipulation Steps in MySQL

The following steps were taken to **clean and prepare** the dataset for analysis:

1. Check for Missing Values and Handle Them

Some records may have missing values, which can impact the analysis.

SQL Query to Identify & Handling Missing Values:

- If values are missing **due to data entry errors**, they can be replaced with averages or median values.
- If records have **too many missing fields**, they can be removed to maintain data integrity.

```
-- REMOVING MISSING VALUES --

-- FROM DAILYACTIVITY TABLE
SELECT *
FROM dailyActivity_merged
WHERE Id IS NULL OR ActivityDate IS NULL OR TotalSteps IS NULL OR Calories IS NULL;

DELETE FROM dailyActivity_merged
WHERE Id IS NULL OR ActivityDate IS NULL OR TotalSteps IS NULL OR Calories IS NULL;

-- FROM HEARTRATE TABLE
SELECT *
FROM heartrate_seconds_merged
WHERE Id IS NULL OR Time IS NULL OR Value IS NULL;

DELETE FROM heartrate_seconds_merged
WHERE Id IS NULL OR Time IS NULL OR Value IS NULL;
```

2. Remove Duplicates

Duplicates can skew the analysis by over-representing certain users.

SQL Query to Identify Duplicates:

```
--Identifying Duplicates
SELECT Id, ActivityDate, COUNT(*)
FROM dailyActivity_merged
GROUP BY Id, ActivityDate
HAVING COUNT(*) > 1;

SELECT Id, COUNT(*)
FROM heartrate_seconds_merged
GROUP BY Id
HAVING COUNT(*) > 1;
```

SQL Query to Remove Duplicates:

```
--Removing Duplicates

DELETE FROM dailyActivity_merged
WHERE Id NOT IN (
    SELECT MIN(Id)
    FROM dailyActivity_merged
    GROUP BY Id, ActivityDate);

DELETE FROM heartrate_seconds_merged
WHERE Id NOT IN (
    SELECT MIN(Id)
    FROM heartrate_seconds_merged
    GROUP BY Id);
```

Summary of Data Cleaning Process

- ✓ Checked for missing values and handled them appropriately
 - ✓ Removed duplicate records
-

Phase 4&5. Analyze Trends and Share Visualization

I used **PowerBi** to analyze trends and to make visuals.

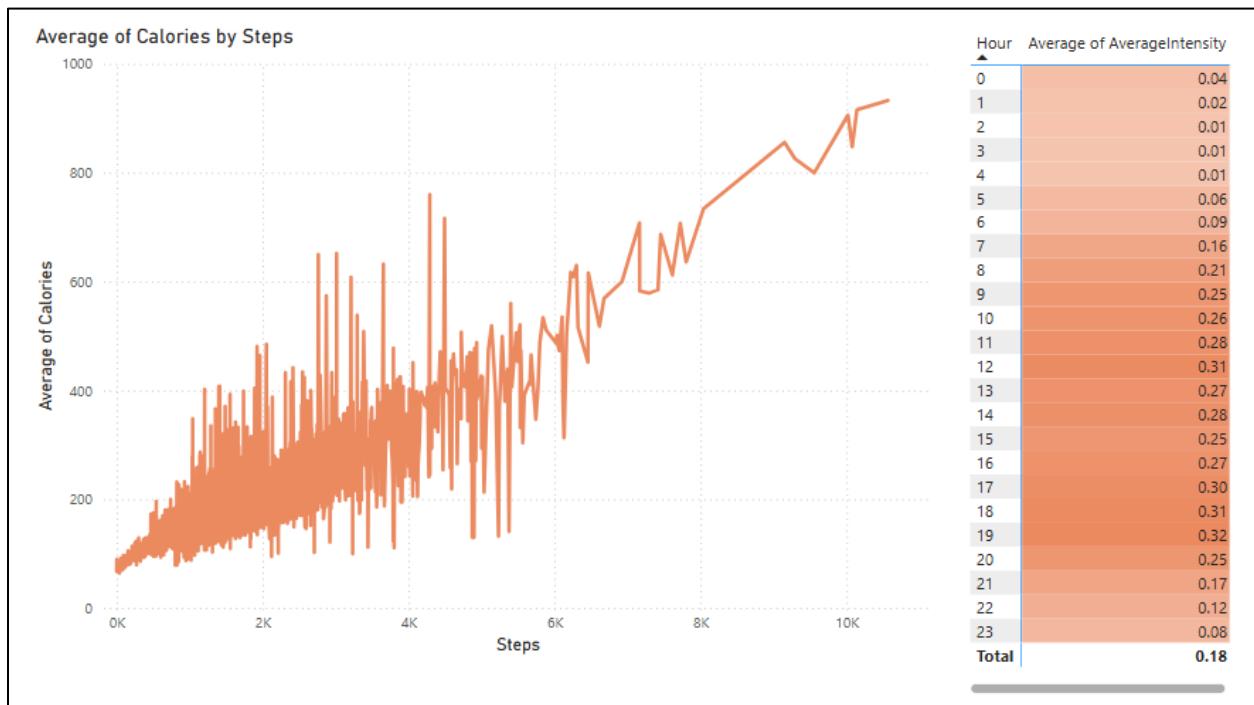
Some trends in smart device usage:

📍 Trend 1: Users Are Most Active in the Morning and Evening

- **Hourly step count and calorie burn peak during 7-9 AM and 6-9 PM**, indicating that users are most active **before and after work hours**.
- **Afternoon activity levels drop significantly**, suggesting that users are less engaged during working hours.

📊 Power BI Visuals:

- Line chart showing **hourly steps & calories burned**.



- Heatmap of **activity intensity throughout the day**.

☒ Key Insights:

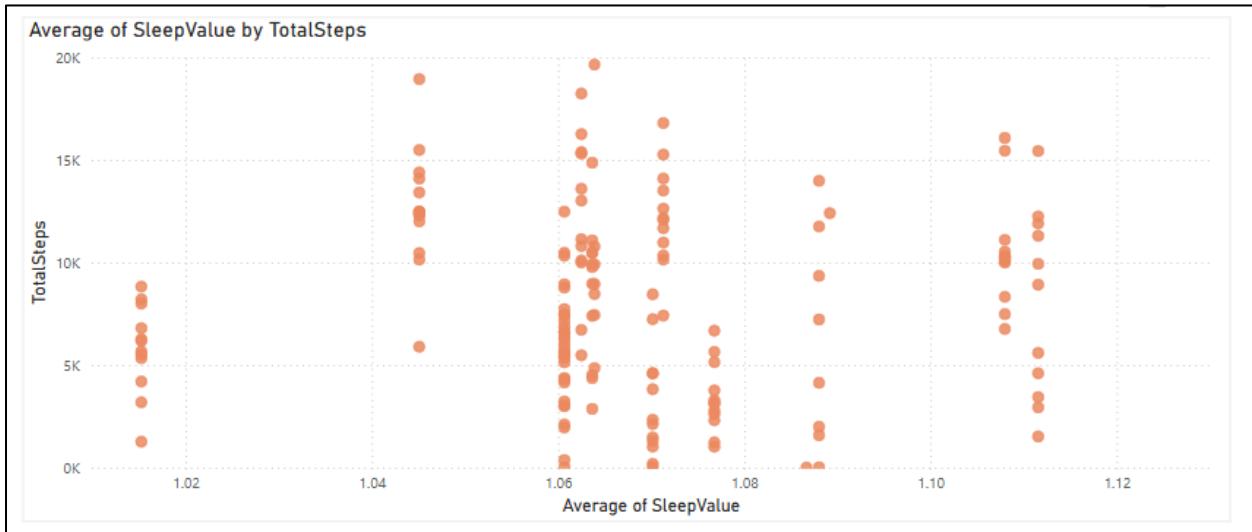
- If **weekend activity is lower**, Bellabeat could **send app reminders** to encourage weekend workouts.
- If **weekday morning workouts are popular**, Bellabeat can **target morning fitness campaigns**.

Trend 2: Higher Activity Leads to Better Sleep

- Users who **walk more than 5,000 steps per day** or engage in **higher intensity workouts** have better sleep patterns (**higher sleep value 1 occurrences**).
- Users with **low activity levels (<5,000 steps per day)** tend to have more **inconsistent sleep**.

Power BI Visuals:

- Scatter plot showing **total steps vs. sleep value**.



Key Insights:

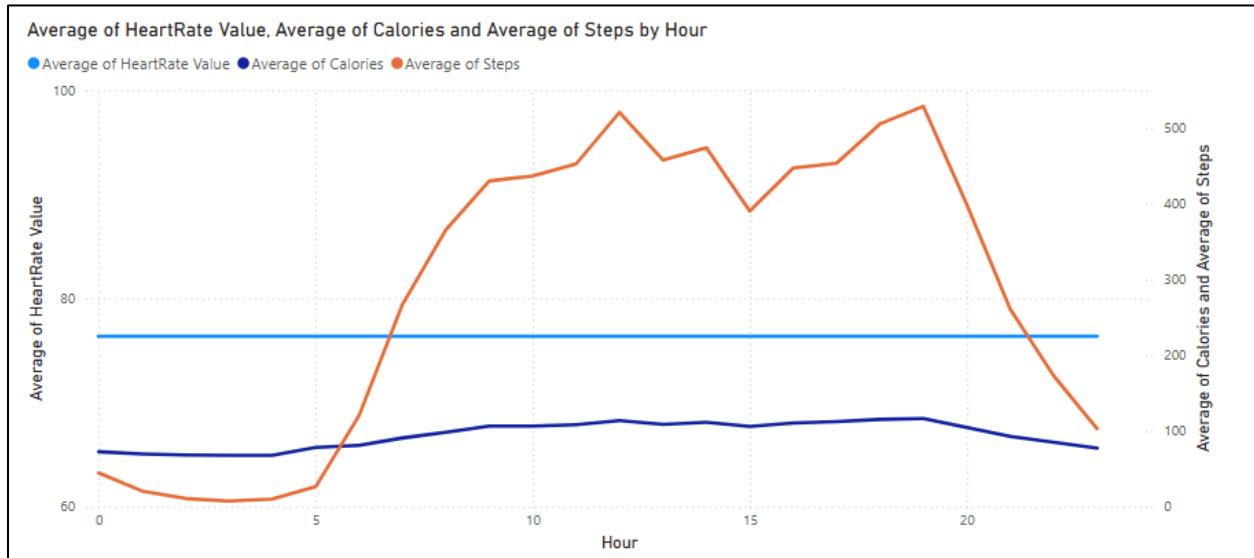
- If higher activity leads to better sleep, Bellabeat can market its smart devices as sleep improvement tools.
- App notifications can encourage evening walks for better sleep quality.

Trend 3: High Heart Rate Correlates with Intense Activity Periods

- Heart rate spikes align with peak activity times (morning and evening), confirming strong engagement during workouts.
- Low heart rate patterns appear during late-night hours, showing when users are at rest.

Power BI Visuals:

- Line graph comparing hourly heart rate vs. activity levels.



Key Insights:

- If heart rate data shows consistent morning workouts, Bellabeat can promote early morning fitness challenges.
- Heart rate tracking reminders can be added to fitness plans for optimal engagement.

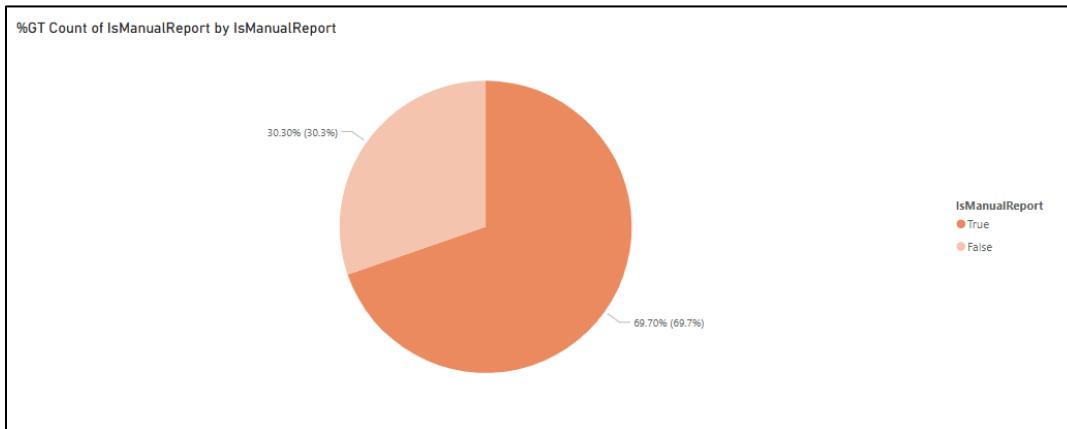
⌚ Trend 4: Weight Tracking Behavior Varies Among Users

Insight:

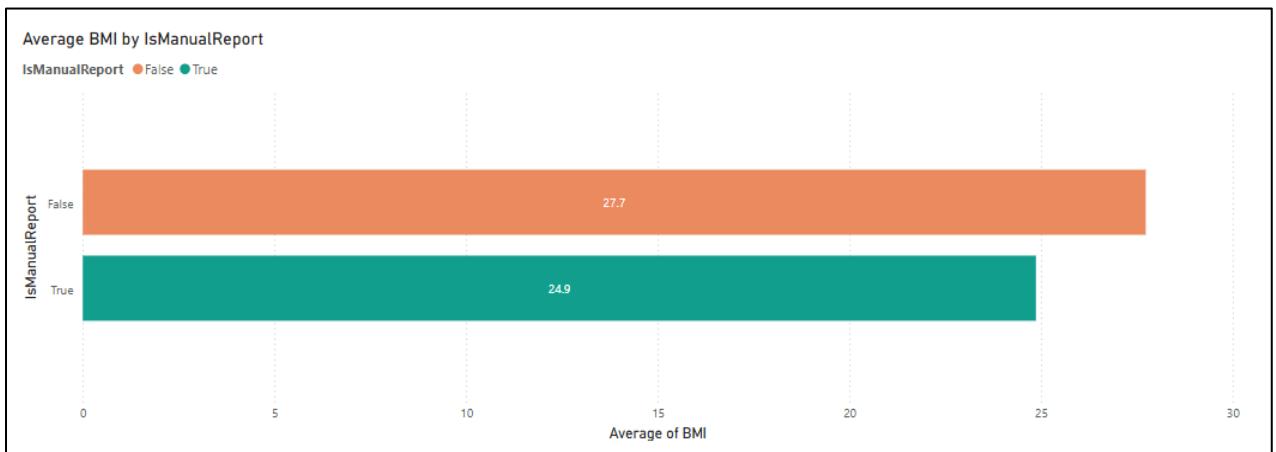
- Some users track weight and BMI consistently, while others rarely log weight data.
- Users who manually track weight have lower BMI and higher engagement with wellness tracking.

◆ Power BI Visuals

📊 Pie Chart – Percentage of Users Who Manually Track Weight



📊 Bar Chart – Manual Weight Tracking vs. BMI Trends



✓ Key Insights:

- If manual weight tracking users have lower BMI, Bellabeat can encourage others to use this feature.
- App push notifications can remind users to log weight weekly.

Phase 6. Act- Recommendations

1. Promote Bellabeat as a Sleep & Wellness Solution

- Market the connection between **activity levels and better sleep**.
- Use **push notifications & emails** to educate users on improving sleep through movement.

2. Introduce Smart Workout & Sleep Reminders

- Send **morning workout notifications (6-9 AM)** and **evening wind-down reminders (6-9 PM)** when users are most active.
- Launch **fitness challenges** based on peak engagement hours.

3. Encourage Weight & BMI Tracking

- Promote **manual weight logging** since engaged users show better wellness tracking behavior.
- Use **reminders & gamification (badges, streaks)** to increase tracking consistency.

4. Develop Personalized Health Insights

- Implement **AI-driven personalized fitness & sleep recommendations** in the Bellabeat app.
 - Offer **weekly reports & adaptive challenges** to motivate different user segments.
-