Case Study 2: Bellabeat – TimeWatch Analysis

Table of Contents

---

1. Summary

This case study analyzes data collected from FitBit users to identify behavioral trends in physical activity and calorie expenditure. Using SQL and Looker Studio, we uncover insights that support marketing and product development strategies for Bellabeat's **TimeWatch**—a

smart wellness tracker. The analysis informs how Bellabeat can target different user groups based on their habits and promote consistent device usage.

---

2. Ask Phase

2.1 Business Task

Bellabeat wants to promote its TimeWatch by understanding how users interact with similar smart fitness devices. The business task is to:

- Segment users by their activity levels.
- Identify patterns in physical activity, calories burned, and device usage.
- Provide actionable insights for marketing and product decisions.

---

3. Prepare Phase

3.1 Dataset Used

We used publicly available FitBit Fitness Tracker data from Kaggle. Key datasets:

- `dailyActivity_merged.csv`
- `dailyIntensities_merged.csv`
- `dailyCalories.csv`

3.2 Accessibility and Privacy of Data

The dataset is open-source and provided under a CC0 license. All data is anonymized and ethically safe for use.

3.3 Information About Our Dataset

- 33 unique users
- Collected over a 31-day period

- Includes metrics like: Total Steps, Calories, Very/Moderately/Lightly Active Minutes, Sedentary Minutes, Activity Date

3.4 Data Organization and Verification

- Checked for missing values, duplicate rows, and null values
- Standardized column names and converted date formats
- Merged based on `Id` and `ActivityDate`

3.5 Data Credibility and Integrity

- Real-world data from fitness devices
- Limited number of users, but adequate for generating hypotheses and insights
- Merged and validated through SQL for consistency

---

4. Process Phase

4.1 Tools Used

- **Google BigQuery** – Data cleaning, transformation, and querying
- **Looker Studio** – Creating interactive and engaging visualizations

4.2 Importing Datasets

CSV files were uploaded and converted into BigQuery tables:

- `dailyActivity`
- `dailyIntensities`
- `dailyCalories`

4.3 Preview of Datasets

```
SELECT * FROM `wellness_dataset.dailyActivity` LIMIT 1000;
```

4.4 Cleaning and Formatting

- Removed rows with null values in key columns:

```
SELECT * FROM `wellness_dataset.dailyActivity`
WHERE Id IS NULL OR TotalSteps IS NULL OR Calories IS NULL;
```

- Removed duplicate rows using:

```
SELECT Id, ActivityDate, COUNT(*)
FROM `wellness_dataset.dailyActivity`
GROUP BY Id, ActivityDate
HAVING COUNT(*) > 1;
```

4.5 Merging Datasets

```
SELECT
  a.Id, a.ActivityDate, a.TotalSteps, a.Calories,
  i.VeryActiveMinutes, i.FairlyActiveMinutes,
  i.LightlyActiveMinutes, i.SedentaryMinutes
FROM
  `wellness_dataset.dailyActivity` a
JOIN
  `wellness_dataset.dailyIntensities` i
ON
  a.Id = i.Id AND a.ActivityDate = i.ActivityDay;
```

---

5. Analyze and Share Phase

5.1 Type of Users per Activity Level

Users were categorized based on `TotalSteps`:

```
SELECT
  CASE
    WHEN TotalSteps < 5000 THEN 'Low'
```

```sql
    WHEN TotalSteps BETWEEN 5000 AND 9999 THEN 'Moderate'
      ELSE 'High'
    END AS step_category,
    AVG(Calories) AS avg_calories
FROM `wellness_dataset.dailyActivity`
GROUP BY step_category;
```

This segmentation helped identify potential customer groups for TimeWatch.

5.2 Visual Insights

Visualizations created in Looker Studio:

- Pie Chart: Proportion of Active vs Least Active users
- Bar Chart: Average Calories by Step Category

- Trend Line: Calories over time

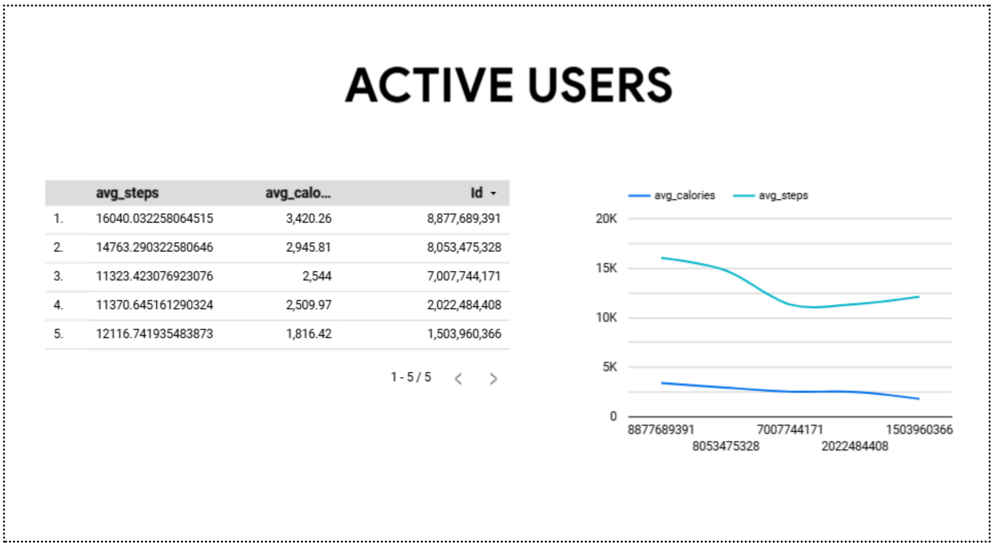
- Donut Chart: User segments by activity

- Combined Chart: Steps vs Active Minutes

These visualizations showed:

Strong correlation between higher steps and calories burned

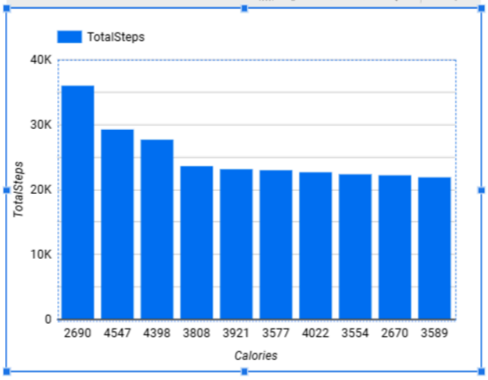
Moderate and Low active users are a significant base to target



CALORIES VS STEPS

	Calories	TotalSteps ▾
1.	2690	36,019
2.	4547	29,326
3.	4398	27,745
4.	3808	23,629
5.	3921	23,186
6.	3577	22,988
7.	4022	22,770
8.	3554	22,359
9.	2670	22,244
10.	3589	22,026

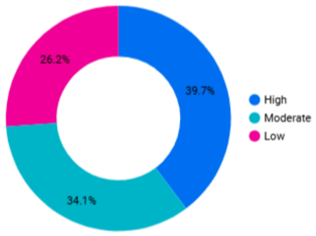
1 - 10 / 10



CUSTOMERS CATEGORISED

	step_category ▾	avg_calories
1.	Moderate	2,355.16
2.	Low	1,806.81
3.	High	2,743.58

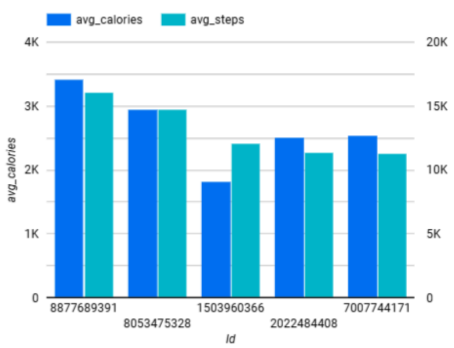
1 - 3 / 3

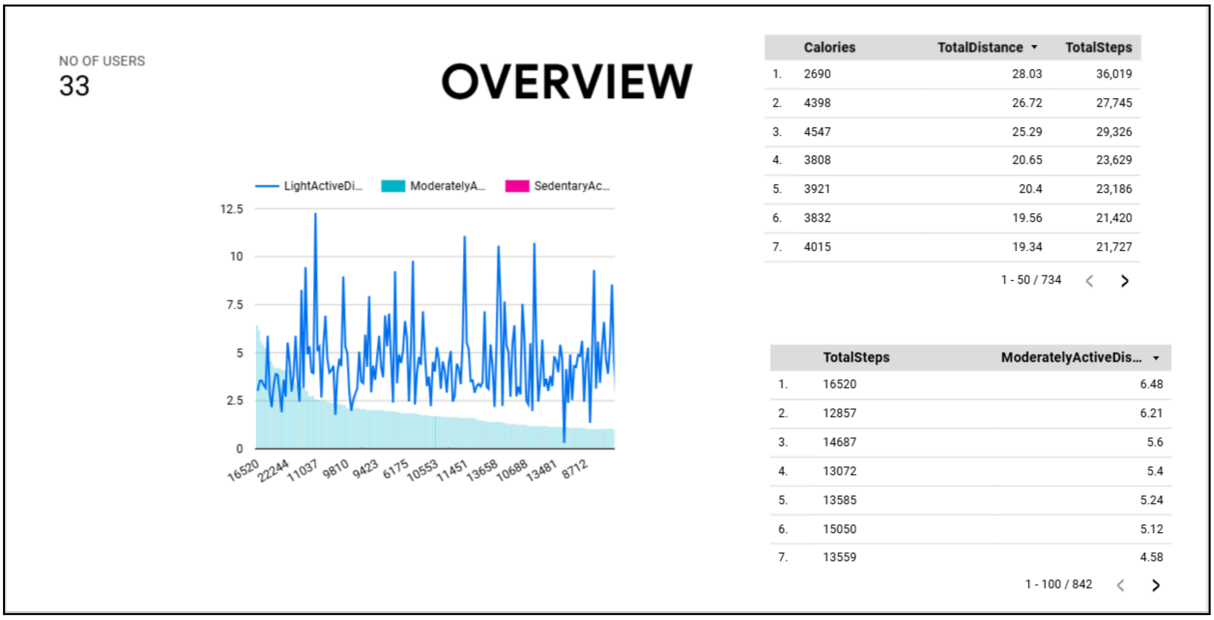


AVG CALORIES VS AVG STEPS

	avg_steps	Id ▾
1.	16040.032258064515	8,877,689,391
2.	14763.290322580646	8,053,475,328
3.	11323.423076923076	7,007,744,171
4.	11370.645161290324	2,022,484,408
5.	12116.741935483873	1,503,960,366

1 - 5 / 5





5.3 Weekly/Hourly Trends (Future Scope)

Hourly and sleep-based insights were not available in the current dataset, but future data sources can explore:

- Peak activity hours
- Device engagement by time of day
- Weekly consistency

5.4 Smart Device Usage

Analysis showed:

- Highly active users log significantly more calories
- Consistent usage leads to healthier patterns
- Moderate users are an ideal target for marketing nudges

5.5 SQL Code Summary

```
-- Top performing users
SELECT
  Id,
  AVG(TotalSteps) AS avg_steps,
  AVG(Calories) AS avg_calories
```

```
FROM `wellness_dataset.dailyActivity`
GROUP BY Id
ORDER BY avg_steps DESC
LIMIT 5;

-- Average activity metrics
SELECT
  AVG(i.VeryActiveMinutes) AS avg_very_active,
  AVG(i.LightlyActiveMinutes) AS avg_light_active,
  AVG(a.Calories) AS avg_calories
FROM
   `wellness_dataset.dailyActivity` a
JOIN
  `wellness_dataset.dailyIntensities` i
ON
  a.Id = i.Id AND a.ActivityDate = i.ActivityDay;
```

---

6. Conclusion (Act Phase)

The analysis shows that:

- Activity level correlates strongly with calories burned
- Moderate and low users should be targeted with personalized nudges and features
- TimeWatch can offer reminders, smart goals, and social motivation features to boost engagement

Recommendation:
Bellabeat should promote TimeWatch as a daily health companion, emphasizing motivation, personalization, and ease of use. The focus should be on nudging moderate users toward consistent activity with in-app rewards, streaks, and fitness challenges.
