# Bellabeat Wellness Device Case Study:

# Analyzing User Activity and Wellness

**Project Title**: Bellabeat Wellness Device Case Study: Analyzing User Activity and Wellness Patterns

**Introduction**: The Bellabeat Wellness Device Case Study explores user activity patterns, sleep behavior, and calorie expenditure to gain insights into wellness trends. This analysis was conducted to help Bellabeat, a wellness technology company, understand user behavior and enhance their products’ engagement and impact on health. By leveraging data from Bellabeat devices, this study aims to identify actionable insights to inform product design and marketing strategies.

The study follows a structured 6-step framework, encompassing asking key questions, preparing and processing data, performing analyses, sharing findings through visualizations, and making data-driven recommendations. Each step is documented in detail to demonstrate the analytical process and outcomes.

A person holding a phone

Description automatically generated

[1. Ask](#Ask)

[2. Prepare](#Prepare)

[3. Process](#Process)

[4. Analyze](#Analyze)

[5. Share](#Share)

[6. Act](#Act)

**1. Ask**

* **Objective**: Define the business task and set clear objectives for the analysis.
* **Business Task**: Analyze activity and wellness data from Bellabeat devices to understand user behavior and trends. The goal is to identify actionable insights that can guide product improvements and marketing strategies for Bellabeat’s wellness products.
* **Questions**:
  + What are the daily activity patterns of Bellabeat users?
  + How do steps, calories, and sleep patterns vary across the week?
  + Are there any correlations between steps taken and calories burned?

**2. Prepare**

* **Data Sources**: [Bellabeat wellness data,](https://www.kaggle.com/datasets/arashnic/fitbit" \t "_blank) which includes metrics such as daily steps, calories burned, activity levels, and sleep duration.
* **Data Cleaning and Preparation**:
  + Merged and cleaned data to remove duplicates and ensure consistency.
  + Performed checks for missing values and handled any discrepancies.
  + Verified data types to ensure accurate analysis

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

**Count of Records and Unique Users**

#Count of records

SELECT COUNT(\*) as total\_records

FROM `Bellabeat\_wellness.daily\_activity\_all`;

#Unique users

SELECT COUNT(DISTINCT id) as total\_records

FROM `Bellabeat\_wellness.daily\_activity\_all`;

**Total records: 874 Unique users: 33**

**Handling Missing Values of Daily Sleep**

Missing values in the Daily\_sleep column have been updated to ensure data consistency for analysis.

#Handling missing values of daily sleep records

UPDATE `merit-america-data-project-fb.Bellabeat\_wellness.daily\_activity\_all`

SET Daily\_sleep = 0

WHERE Daily\_sleep = '#N/A';

**3. Process**

* **Exploratory Data Analysis (EDA)**:
  + Initial data exploration was conducted using SQL to identify distributions and patterns within each variable.
  + Calculated Basic statistics (Min, Max, Average).

# Basic statistics (Min, Max, Average)

WITH distinct\_steps AS(

  SELECT DISTINCT id,ActivityDate,ID\_ActivityDate,DailySteps

  FROM `Bellabeat\_wellness.daily\_activity\_all`

),

aggr\_calories AS(

  SELECT id,ActivityDate,SUM(Calories) AS total\_calories

  FROM `Bellabeat\_wellness.daily\_activity\_all`

  GROUP BY Id,ActivityDate

  ),

  aggr\_sleep AS(

    SELECT id,ActivityDate,SUM(Cast(Daily\_sleep as float64)) AS total\_sleep

    FROM `Bellabeat\_wellness.daily\_activity\_all`

    WHERE Daily\_sleep IS NOT NULL ANd Daily\_sleep <> ''

    GROUP BY id,ActivityDate

  )

SELECT

  MIN(DailySteps) as min\_steps,

  MAX(DailySteps) as max\_steps,

  Round(AVG(DailySteps),2) as avg\_steps,

  MIN(ac.total\_calories) as min\_calories,

  MAX(ac.total\_calories) as max\_calories,

  Round(AVG(ac.total\_calories),2) as avg\_calories,

  MIN(asl.total\_sleep) as min\_sleep,

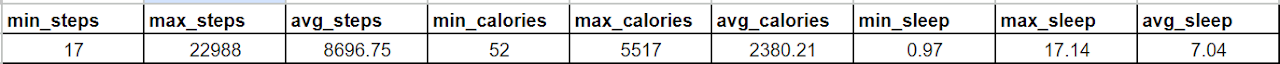
  MAX(asl.total\_sleep) as max\_sleep,

  Round(AVG(asl.total\_sleep),2) as avg\_sleep

FROM distinct\_steps ds

INNER JOIN aggr\_calories ac ON ds.id = ac.id

Inner join aggr\_sleep asl ON asl.id = ds.id;



**4. Analyze**

# Daily steps and average calories over time

WITH distinct\_steps AS(

  SELECT DISTINCT id,ActivityDate,ID\_ActivityDate,DailySteps

  FROM `Bellabeat\_wellness.daily\_activity\_all`

),

aggr\_calories AS(

  SELECT id,ActivityDate,SUM(Calories) AS total\_calories

  FROM `Bellabeat\_wellness.daily\_activity\_all`

  GROUP BY Id,ActivityDate

  )

SELECT ac.ActivityDate,SUM(DailySteps) as sum\_steps,

  ROUND(AVG(ac.total\_calories),2) as avg\_calories

FROM distinct\_steps ds

INNER JOIN aggr\_calories ac ON ac.id = ds.id

GROUP BY ActivityDate

ORDER BY ActivityDate;

A screenshot of a calculator

Description automatically generated

# Steps Distribution

WITH distinct\_steps AS(

  SELECT DISTINCT id,ActivityDate,ID\_ActivityDate,DailySteps

  FROM `Bellabeat\_wellness.daily\_activity\_all`

)

SELECT

  CASE

    WHEN DailySteps < 5000 THEN '(1) <5000'

    WHEN DailySteps BETWEEN 5000 AND 10000 THEN '(2) 5000 - 10000'

    WHEN DailySteps BETWEEN 10001 AND 15000 THEN '(3) 10001 - 15000'

    ELSE '(4) >15000'

  END AS step\_range,

  Count(id) as count\_days

  FROM distinct\_steps

  GROUP BY step\_range

  ORDER BY step\_range;

# Calories Distribution

WITH aggr\_calories AS(

  SELECT id,ActivityDate,SUM(Calories) AS total\_calories

  FROM `Bellabeat\_wellness.daily\_activity\_all`

  GROUP BY Id,ActivityDate

  )

  SELECT

    CASE

      WHEN aggr\_calories.total\_calories < 1000 THEN '(1) <1000'

      WHEN aggr\_calories.total\_calories BETWEEN 1000 AND 2000 THEN '(2) 1000-2000'

      ELSE '(3) >2000'

    END AS calorie\_range,

    Count(id) as count\_days

  FROM aggr\_calories

  GROUP BY calorie\_range

  ORDER BY calorie\_range;

A screenshot of a computer

Description automatically generated

A screenshot of a google calendar

Description automatically generated

A graph of a bar

Description automatically generated with medium confidence

A graph of calories burned

Description automatically generated

A graph of a graph with colored dots

Description automatically generated with medium confidence

A graph of blue bars

Description automatically generated with medium confidence

#Outlier Detection

#Identify days with unusually high/low steps:

WITH distinct\_steps AS(

  SELECT DISTINCT id,ActivityDate,ID\_ActivityDate,DailySteps

  FROM `Bellabeat\_wellness.daily\_activity\_all`

)

(SELECT id,ActivityDate,DailySteps,(SELECT Round(AVG(DailySteps),2) FROM distinct\_steps) as avg\_daily\_steps,(Select Round(STDDEV(DailySteps),2) FROM distinct\_steps) as standard\_deviation

FROM distinct\_steps ds

WHERE DailySteps > (SELECT AVG(DailySteps) + 2 \* STDDEV(DailySteps) FROM distinct\_steps)

)

UNION ALL

(

SELECT id,ActivityDate,DailySteps,(SELECT Round(AVG(DailySteps),2) FROM distinct\_steps) as avg\_daily\_steps,(Select Round(STDDEV(DailySteps),2) FROM distinct\_steps) as standard\_deviation

FROM distinct\_steps

WHERE DailySteps < 1000

)

ORDER BY DailySteps

A screenshot of a data

Description automatically generated

A screenshot of a graph

Description automatically generated

#Correlation between Steps and Calories

WITH distinct\_steps AS(

  SELECT DISTINCT id,ActivityDate,ID\_ActivityDate,DailySteps

  FROM `Bellabeat\_wellness.daily\_activity\_all`

),

aggr\_calories AS(

  SELECT id,ActivityDate,SUM(Calories) AS total\_calories

  FROM `Bellabeat\_wellness.daily\_activity\_all`

  GROUP BY Id,ActivityDate

  )

  SELECT CORR(DailySteps,total\_calories) AS correlation\_steps\_calories

  FROM distinct\_steps ds

  INNER JOIN aggr\_calories ac ON ac.id = ds.id;

**correlation\_steps\_calories: 0.24**

A graph showing a number of calories burned

Description automatically generated

A graph with a bar

Description automatically generated

A screen shot of a graph

Description automatically generated

**Key Insights**:

* **Daily Calories Burned**: Most users burned between 1,500 and 2,500 calories daily, with fewer high-calorie days.
* **Daily Steps**: Users generally took between 5,000 and 10,000 steps per day, with a noticeable decrease on weekends.
* **Activity Levels Across the Week**: Sedentary minutes were significantly high, especially during weekends.
* **Correlation Between Steps and Calories**: A positive but weak correlation (0.24), suggesting other factors impact calorie burn beyond steps taken.
* **Sleep Duration**: Most users averaged around 6-8 hours of sleep per night.

**5. Share**

**Dashboard 1: Overview of Daily Activity Patterns and Wellness Metrics**

A screenshot of a graph

Description automatically generated

This dashboard provides an overview of key daily wellness metrics tracked by Bellabeat users. It includes the distribution of daily calories burned, daily steps taken, sleep duration, and activity levels across the week. Additionally, it shows the correlation between steps and calories burned.

**Dashboard 2: Weekly Trends in Activity, Calories, and Sleep**

A screenshot of a data presentation

Description automatically generated

This dashboard explores weekly trends in steps, calories, and sleep duration. It visualizes the spread of daily steps and calories by each day of the week, average sleep patterns, and the distribution of activity distances. By analyzing these metrics over the week, we can observe fluctuations in user engagement and identify days with higher or lower activity levels.

**6. Act**

**Recommendations for Bellabeat**

Based on the insights gathered from the analysis, the following recommendations can help Bellabeat enhance user engagement and promote healthier lifestyles:

**1. Introduce Activity Challenges and Reminders on Weekends:**

Analysis showed that users tend to be more sedentary during the weekends. Bellabeat can introduce weekend challenges or send reminders to encourage users to increase their activity levels.

**2.** **Promote Notifications for Low Activity or Calorie Burn Days**

Bellabeat could implement notifications for users who have low daily steps or calorie burn, encouraging them to stay active. Personalized messages could remind users of their health goals and suggest achievable activities to help them stay on track. This would create a more supportive experience and help users develop consistent wellness habits.

**Next Steps**

To further refine Bellabeat’s offerings, consider these next steps:

* **User Segmentation Analysis**: Identify distinct user segments based on activity, calorie burn, and sleep patterns. This will enable Bellabeat to offer more personalized recommendations and tailor marketing strategies to different user needs.
* **Enhanced Sleep Insights**: Explore the relationship between sleep patterns and activity levels to offer advice on sleep quality improvements, aligning with holistic wellness goals.

**Tools and Skills Demonstrated**

* **SQL**: Used for data cleaning, aggregation, and calculating summary statistics.
* **Spreadsheet**: Assisted with data preparation, handling missing values, and preprocessing.
* **Tableau**: Enabled visualizations and dashboard creation to communicate insights effectively.

* **Data Cleaning**: Ensured consistency and accuracy by handling missing and incorrect values.
* **Data Analysis**: Conducted exploratory data analysis (EDA) to uncover patterns and insights.
* **Data Visualization**: Created engaging visualizations to illustrate key findings and support recommendations.

**Conclusion**

This analysis of Bellabeat’s wellness data highlights the potential for enhancing user engagement through data-driven insights. By identifying trends in user activity, sleep, and calorie expenditure, Bellabeat can introduce personalized features that align with users’ wellness goals.