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.

5. Share

6. 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, 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

Fields

Type	Field Name	Physical Table	Remote Field Name
#	Id	dailyActivity!all	Id
📅	Activity Date	dailyActivity!all	ActivityDate
Abc	ID ActivityDate	dailyActivity!all	ID_ActivityDate
#	Total Steps	dailyActivity!all	TotalSteps
#	Total Distance	dailyActivity!all	TotalDistance
#	Tracker Distance	dailyActivity!all	TrackerDistance
#	Logged Activities Distance	dailyActivity!all	LoggedActivitiesDis...
#	Very Active Distance	dailyActivity!all	VeryActiveDistance
#	Moderately Active Distance	dailyActivity!all	ModeratelyActiveDi...
#	Light Active Distance	dailyActivity!all	LightActiveDistance
#	Sedentary Active Distance	dailyActivity!all	SedentaryActiveDis...
#	Very Active Minutes	dailyActivity!all	VeryActiveMinutes

dailyActivity!all
874 rows 18 fields

dailyActivity!all Calories	dailyActivity!all Daily_sleep	dailyActivity!all DailySteps	dailyActivity!all FairlyActiveMinutes	dailyActivity!all Id	dailyActivity!all LightActiveDistance	dailyActivity!all LightlyActiveMinute
50	5.4500	13,162	0	1,503,960,366	0.1300	
1,985	5.4500	13,162	13	1,503,960,366	6.0600	
1,797	6.4000	10,735	19	1,503,960,366	4.7100	
1,776	*Null*	10,460	11	1,503,960,366	3.9100	
1,745	6.8700	9,762	34	1,503,960,366	2.8300	
1,863	5.6700	12,669	10	1,503,960,366	5.0400	
1,728	11.6700	9,705	20	1,503,960,366	2.5100	
1,921	*Null*	13,019	16	1,503,960,366	4.7100	
2,035	5.0700	15,506	31	1,503,960,366	5.0300	
1,786	6.0000	10,544	12	1,503,960,366	4.2400	
1,775	5.4200	9,819	8	1,503,960,366	4.6500	
1,827	*Null*	12,764	27	1,503,960,366	2.2400	
1,949	6.0200	14,371	21	1,503,960,366	5.3600	
1,788	7.1700	10,039	5	1,503,960,366	3.2800	
2,013	4.6200	15,355	14	1,503,960,366	3.9400	
1,970	4.0800	13,755	23	1,503,960,366	5.5400	
2,159	*Null*	18,134	11	1,503,960,366	5.4100	

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;
```

min_steps	max_steps	avg_steps	min_calories	max_calories	avg_calories	min_sleep	max_sleep	avg_sleep
17	22988	8696.75	52	5517	2380.21	0.97	17.14	7.04

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;
```

ActivityDate	sum_steps	avg_calories
2016-04-12	6916728	2898.78
2016-04-13	6911963	2340.42
2016-04-14	6966879	2427.03
2016-04-15	7179539	2399.72
2016-04-16	6845293	2444.26
2016-04-17	6495962	2329.45
2016-04-18	7110429	2374.58
2016-04-19	6946258	2445.8
2016-04-20	6985549	2446.88
2016-04-21	6836179	2488.81

Steps Distribution

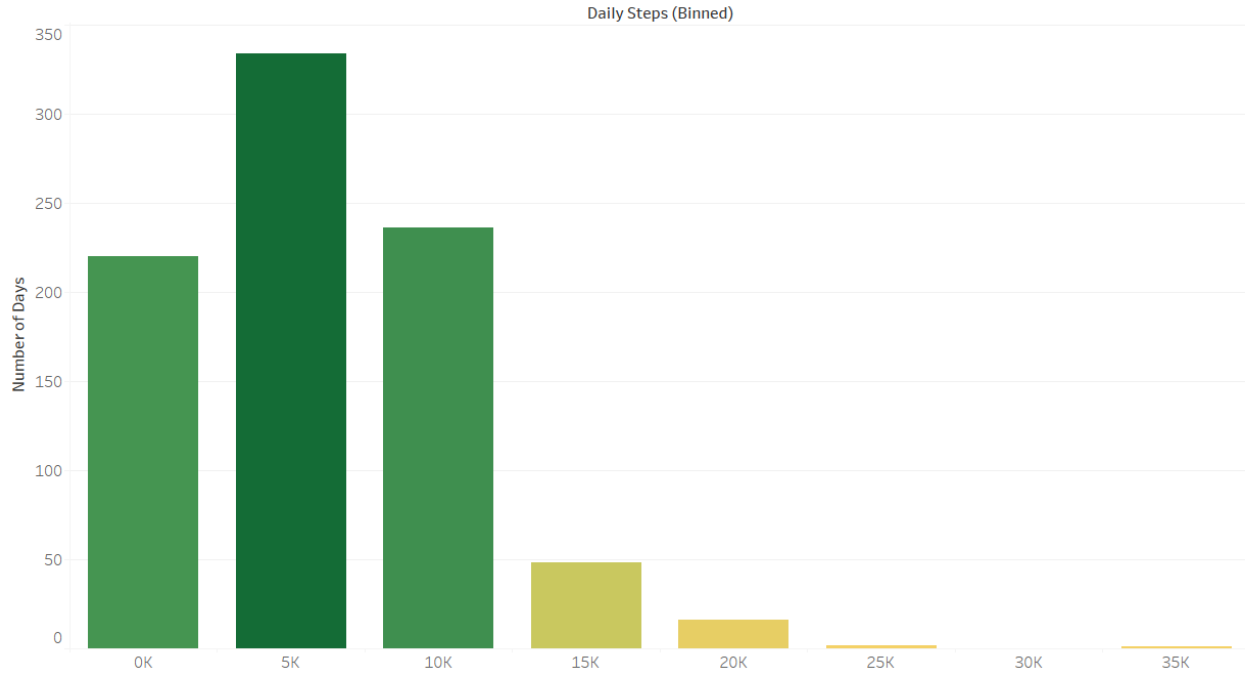
```
WITH distinct_steps AS(
  SELECT DISTINCT id,ActivityDate,ID_ActivityDate,DailySteps
  FROM `Bellabeat_wellness.daily_activity_all`
)


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;
```

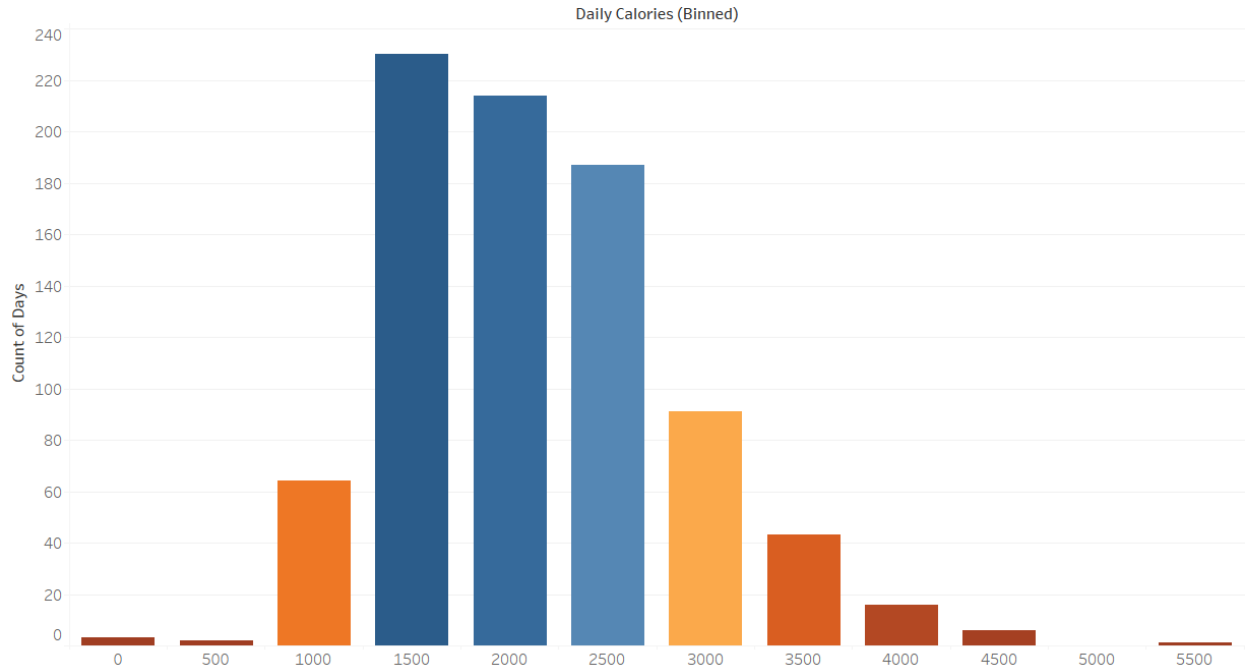
step_range	count_days
(1) <5000	220
(2) 5000 - 10000	334
(3) 10001 - 15000	236
(4) >15000	67

calorie_range	count_days
(1) <1000	5
(2) 1000-2000	294
(3) >2000	558

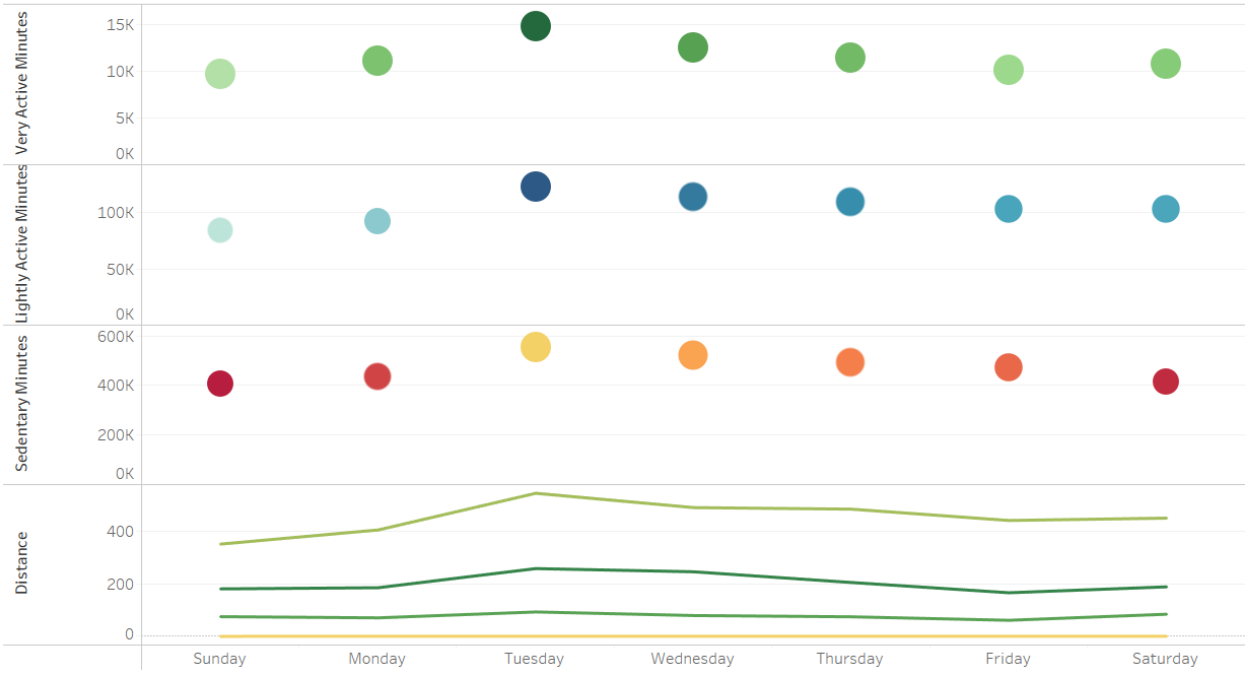
Distribution of Daily Steps Across All Days



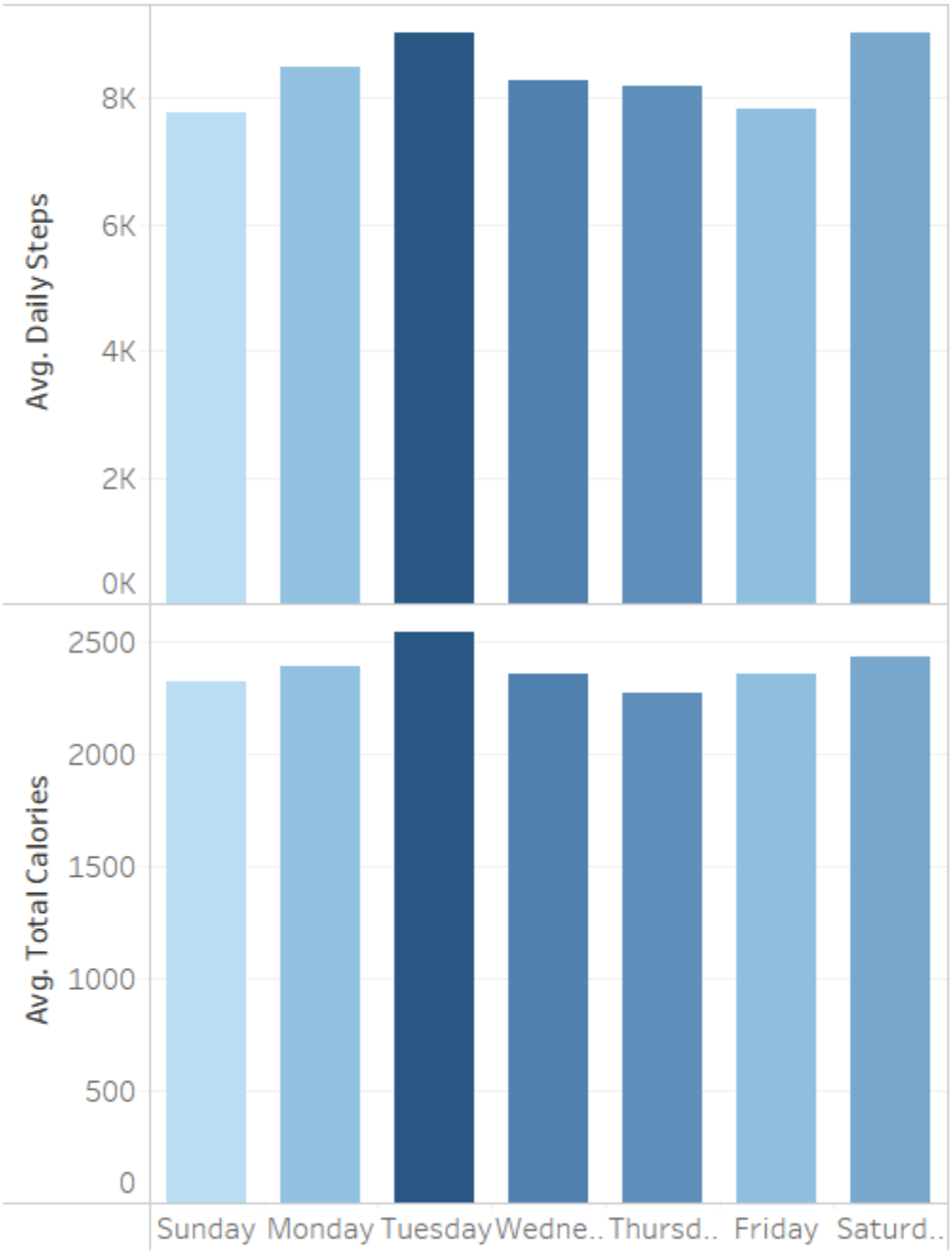
Distribution of Daily Calories Burned



Distribution of Activity Levels Across the Week

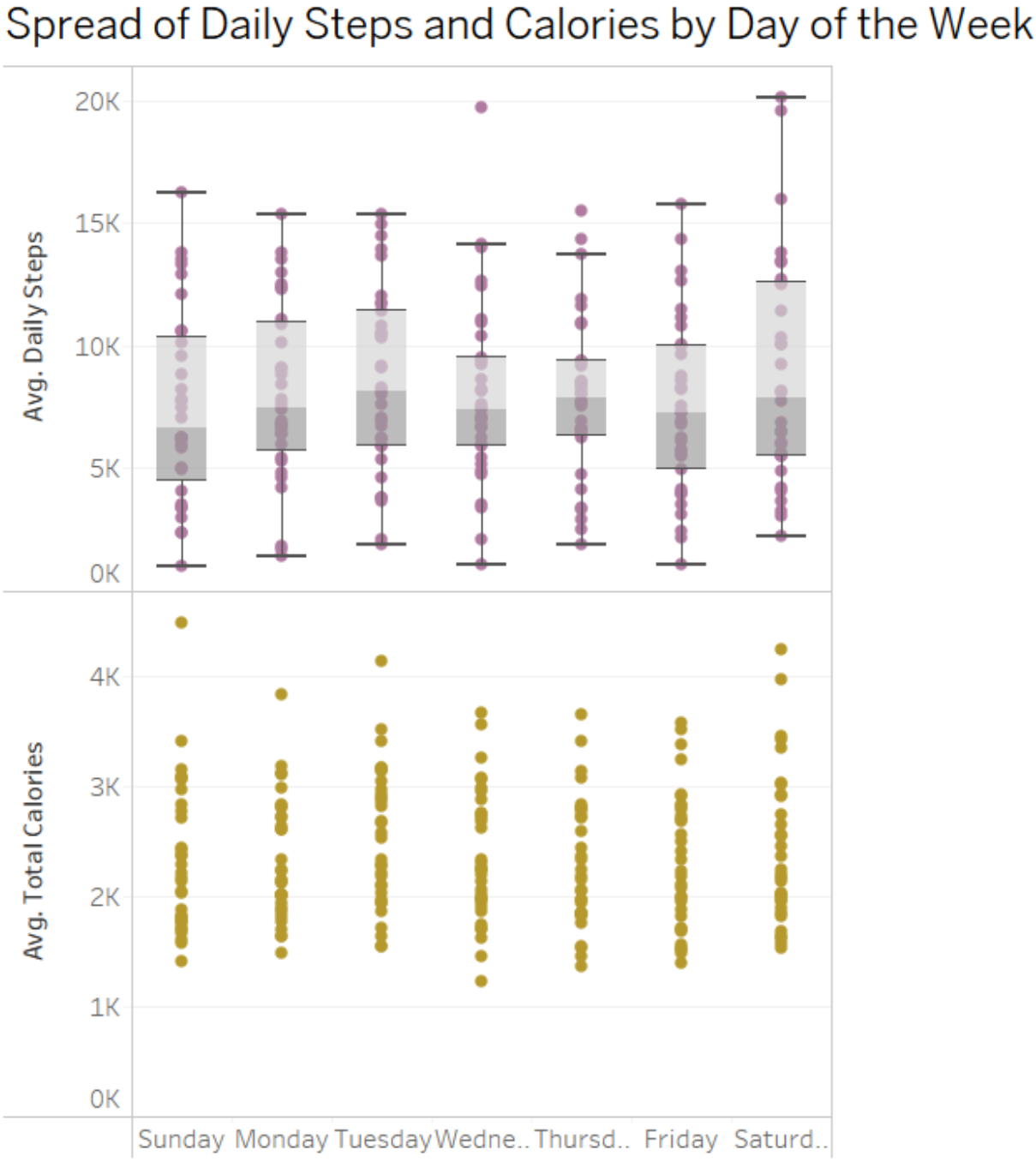


Average Steps/Calories by Day of the Week



```
#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
```

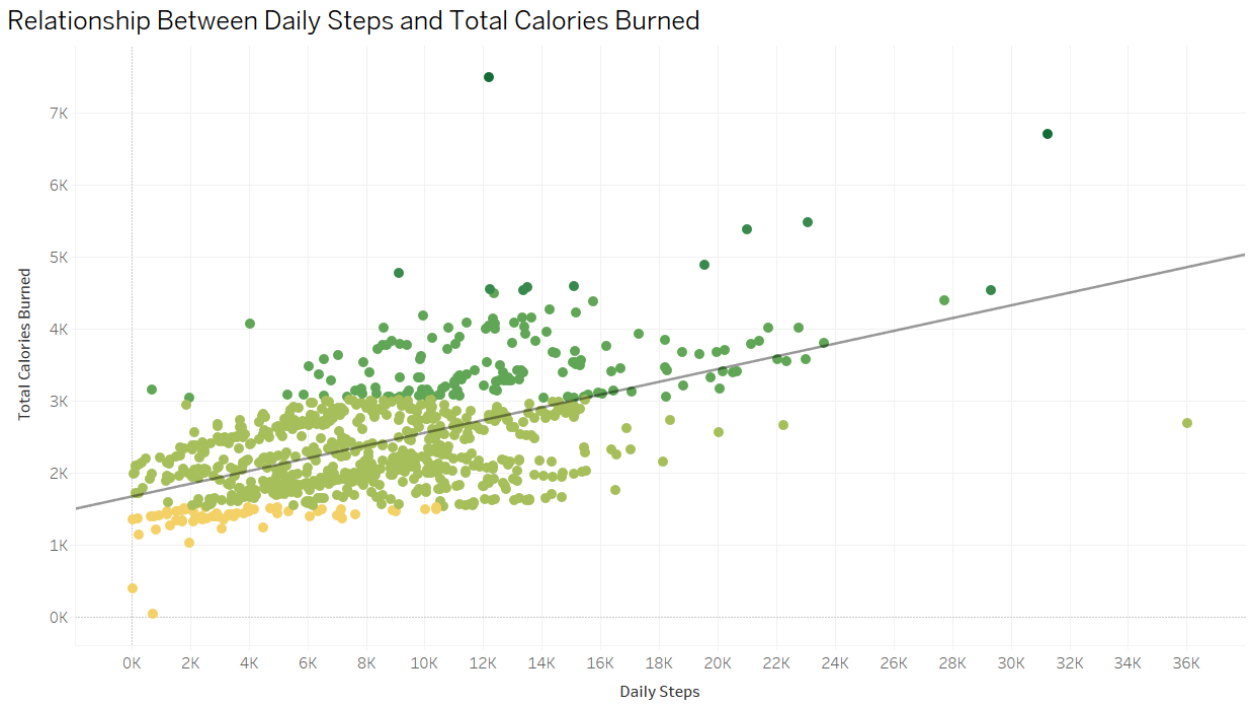
id	ActivityDate	DailySteps	avg_daily_steps	standard_deviation
4319703577	2016-05-12	17	8377.53	4710.19
2347167796	2016-04-29	42	8377.53	4710.19
1844505072	2016-05-06	44	8377.53	4710.19
4020332650	2016-04-18	62	8377.53	4710.19
4020332650	2016-04-14	108	8377.53	4710.19
8792009665	2016-04-21	144	8377.53	4710.19
1927972279	2016-04-22	149	8377.53	4710.19
1927972279	2016-04-25	152	8377.53	4710.19
1844505072	2016-04-19	197	8377.53	4710.19
1927972279	2016-04-18	244	8377.53	4710.19



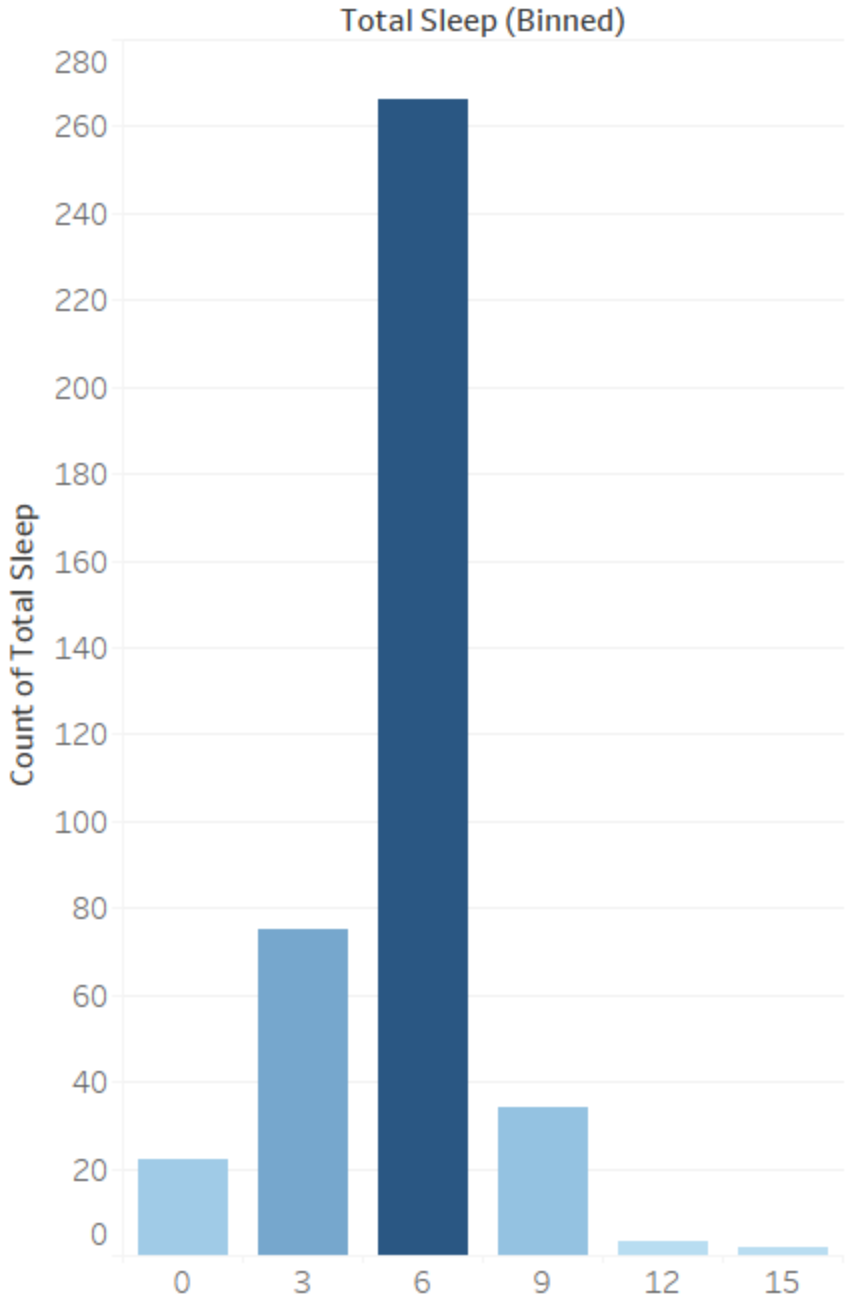

```
#Correlation between Steps and Calories
WITH distinct_steps AS(
 SELECT DISTINCT id,ActivityDate,ID_ActivityDate,DailySteps
 FROM `Bellabeat_wellness.daily_activity_all`
),
```

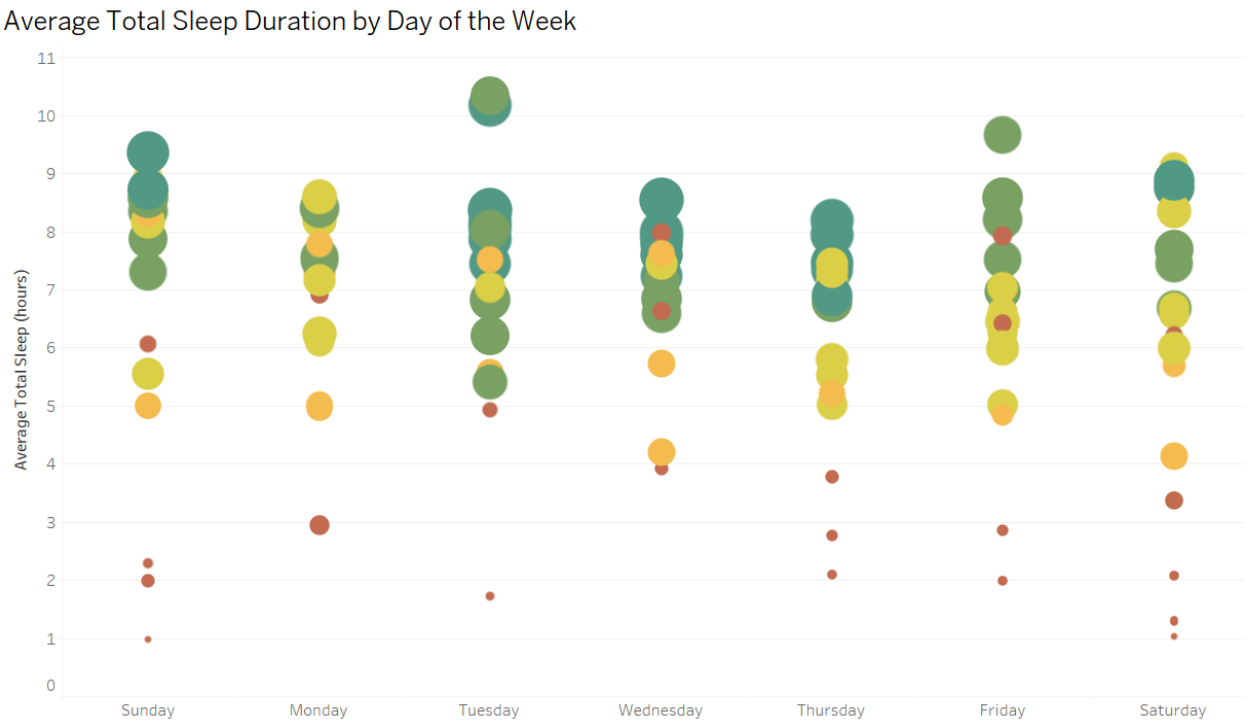
```
aggr_calories AS(
 SELECT id,ActivityDate,SUM(Calories) AS total_calories
 FROM `Bellabeat_wellness.daily_activity_all`
 GROUP BY Id,ActivityDate
 )


 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



Distribution of Daily Sleep



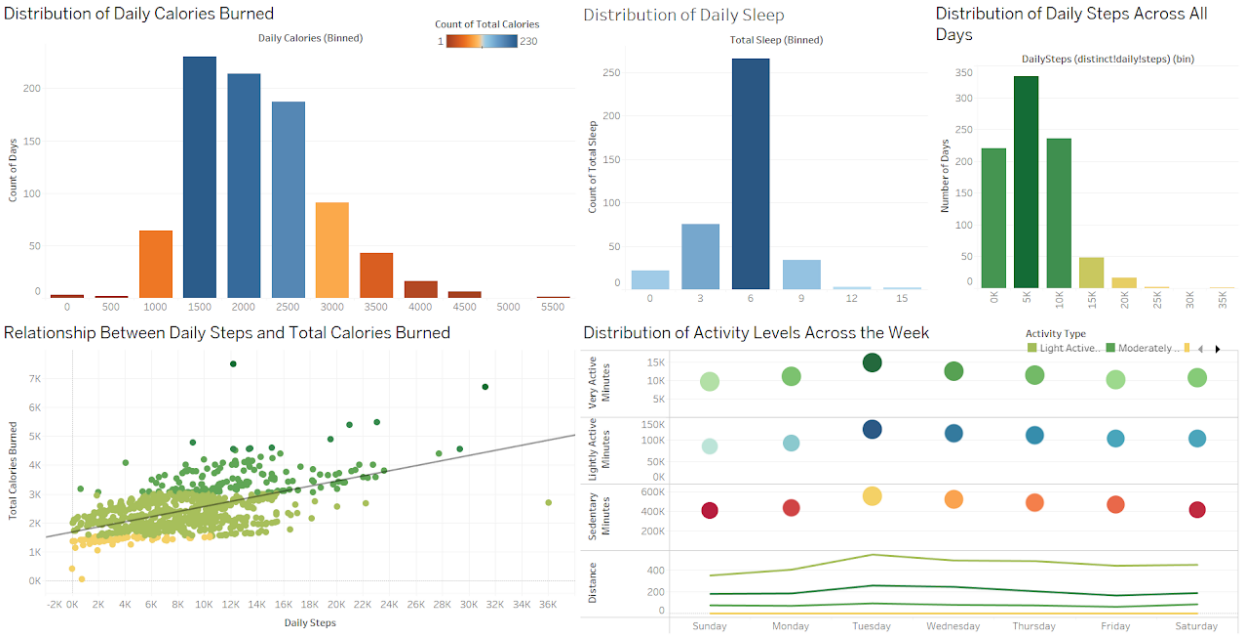


Key Insights:

- **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.

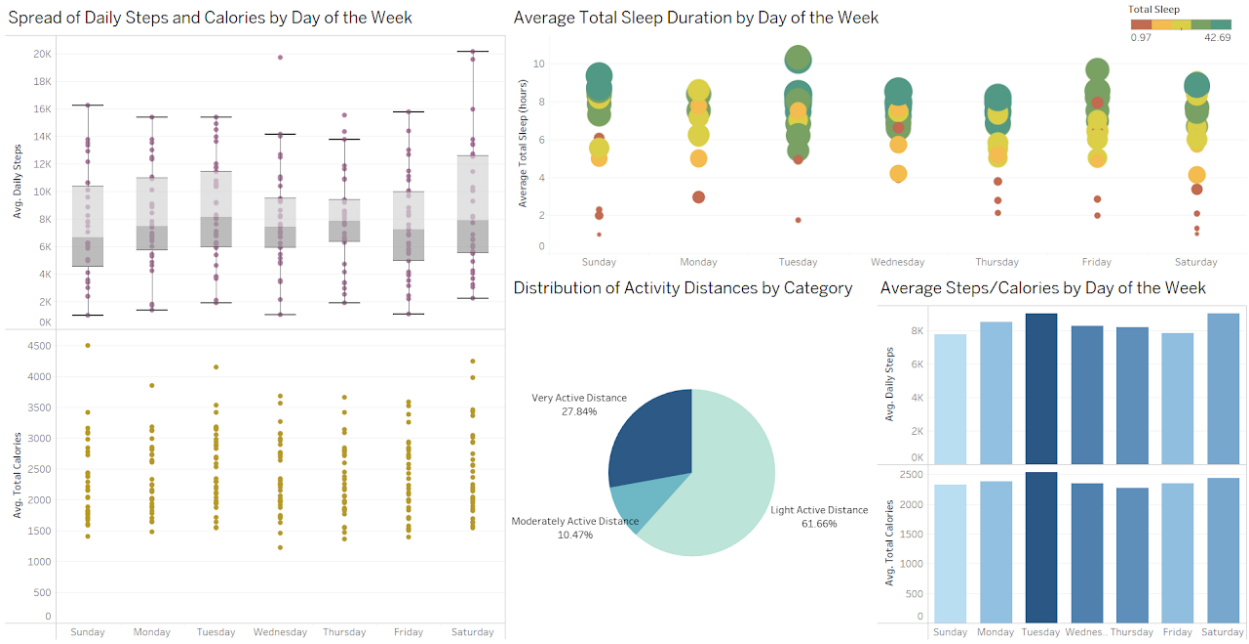
5. Share

Dashboard 1: Overview of Daily Activity Patterns and Wellness Metrics



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



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.