

HOW CAN A WELLNESS TECHNOLOGY
COMPANY PLAYS IT SMART?

An abstract graphic design featuring two thin, dark gray lines that intersect on a light gray background. One line runs diagonally from the top-left towards the bottom-right, while the other runs from the top-right towards the bottom-left. The intersection point is located to the left of the text.

A FITBIT FITNESS TRACKER DATA

Two thin, dark grey lines intersect on a light grey background. One line is nearly vertical, starting from the top center and extending towards the bottom right. The other line is diagonal, starting from the top left and extending towards the bottom right, crossing the first line.

THIS STUDY WILL PASS THROUGH THE SIX
PHASES OF THE DATA LIFE CYCLE WHICH ARE;

ASK


PREPARE

PROCESS

ANALYZE

SHARE

ACT

A series of thin, black, intersecting lines forming a complex geometric pattern in the top-left corner of the slide.

Bellabeat, a high-tech wellness company founded in 2013, creates smart health products for women, tracking data on activity, sleep, stress, and reproductive health. This data empowers women by providing insights into their health and habits. Urška Sršen, Bellabeat's co-founder and Chief Creative Officer, believes that analyzing fitness data from smart devices can reveal new growth opportunities. In this case study, the focus is on one of Bellabeat's products, with the goal of using consumer data insights to guide the company's marketing strategy and uncover growth potential.

Bellabeat has several products, but for this analysis, we will focus on the application. The Bellabeat app provides users with health data related to their activity, sleep, stress, menstrual cycle, and mindfulness habits. This data can help users understand their current practices and make healthy decisions.

BUSINESS TASK

The task is to analyze smart device usage data to uncover trends and insights into how consumers are using their devices. These insights will then be applied to one of Bellabeat's products to help inform and optimize Bellabeat's marketing strategy. the analysis should focus on understanding consumer behavior with smart devices and identifying actionable trends that can be leveraged to better engage Bellabeat's customers and drive growth in the competitive smart device market.

THE DATASET

These datasets were made available through Mobius in Kaggle and licensed under the CCO Public Domain.

This dataset generated by respondents to a distributed survey via Amazon Mechanical Turk between 03.12.2016-05.12.2016.

I reviewed each dataset to understand what I'd be working with and observed the following:

- 1. The data is organized in a long format.**
- 2. The dataset lacks integrity and reliability. While Zenodo.org reported that 30 Fitbit users consented to share their personal tracker data, further examination showed 33 users over a 31-day span. Additionally, since the data was collected in 2016, it is outdated for the current analysis. Although the timeframe was listed as December 3, 2016, to May 12, 2016, the data actually covered only 31 days.**
- 3. The small sample size introduces bias, and a larger sample would have provided a more accurate representation. Moreover, the absence of demographic details such as gender limits the analysis, particularly since Bellabeat's target audience is primarily women. Including demographic data would have greatly improved the quality of the analysis and recommendations.**

For this analysis, I will be using the following datasets:

- dailyactivity_merged**
- dailycalories_merged**
- dailyIntensities_merged**
- dailysteps_merged**

Data cleaning was performed in Google Sheets and SQL, following these steps for each dataset:

1. Sort and filter data.
2. Check for empty cells.
- 3- Link tables with duplicate columns.
4. Format date and time columns to MM/DD/YYYY.

After completing the initial cleaning, I moved to SQL to proceed with the analysis.

ANALYZE

I used the SQL Import Wizard to upload all the files in preparation for analysis.

1-Check data size and distribution:

```
#total record  
SELECT COUNT(*) AS total_records FROM `first-one-  
434623.Google_project.dailyActivity_merged`
```

2. Analysis of users' average daily steps:

```
SELECT Id,  
       AVG(TotalSteps) AS avg_daily_steps,  
       MIN(TotalSteps) AS min_daily_steps,  
       MAX(TotalSteps) AS max_daily_steps  
FROM `first-one-434623.Google_project.dailyActivity_merged`  
GROUP BY Id  
ORDER BY avg_daily_steps DESC;
```

ANALYZE

3. Daily calorie analysis of users:

```
SELECT Id,  
       AVG(Calories) AS avg_daily_calories,  
       MIN(Calories) AS min_daily_calories,  
       MAX(Calories) AS max_daily_calories  
FROM `first-one-434623.Google_project.dailyActivity_merged`  
GROUP BY Id  
ORDER BY avg_daily_calories DESC;
```

4. Daily activity analysis based on activity levels:

```
SELECT Id,  
       AVG(VeryActiveMinutes) AS avg_very_active_minutes,  
       AVG(FairlyActiveMinutes) AS avg_fairly_active_minutes,  
       AVG(LightlyActiveMinutes) AS avg_lightly_active_minutes,  
       AVG(SedentaryMinutes) AS avg_sedentary_minutes  
FROM `first-one-434623.Google_project.dailyIntensities_merged`  
GROUP BY Id  
ORDER BY avg_very_active_minutes DESC;
```

ANALYZE

5. Heart rate analysis throughout the day:

```
SELECT Id,  
       EXTRACT(HOUR FROM Time) AS hour,  
       AVG(Value) AS avg_heart_rate  
FROM `first-one-434623.Google_project.heartrate_seconds_merged`  
GROUP BY Id, hour  
ORDER BY Id, hour
```

6. Analysis of calories burned per hour:

```
SELECT Id,  
       EXTRACT(HOUR FROM ActivityHour) AS hour,  
       AVG(Calories) AS avg_hourly_calories  
FROM `first-one-434623.Google_project.hourlyCalories_merged`  
GROUP BY Id, hour  
ORDER BY Id, hour;
```

ANALYZE

7. Analysis of the relationship between the number of steps and calories:

```
SELECT a.Id,  
       a.ActivityDate,  
       a.TotalSteps,  
       c.Calories  
FROM `first-one-434623.Google_project.dailyActivity_merged` a  
JOIN `first-one-434623.Google_project.dailyCalories_merged` c  
ON a.Id = c.Id AND a.ActivityDate = c.ActivityDay;
```

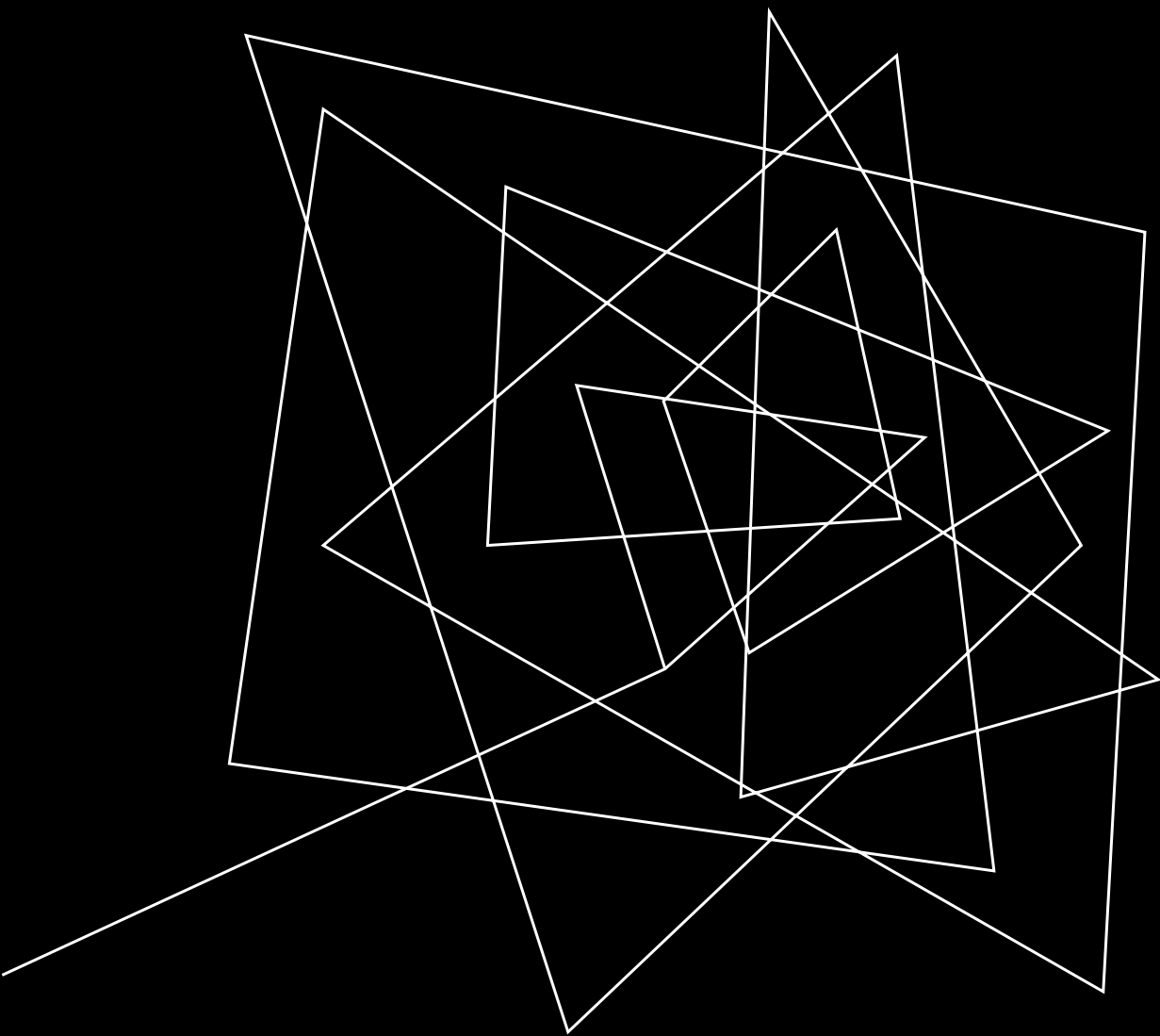
8. Weekly trends:

```
SELECT da.Id,  
       FORMAT_DATE('%A', DATE(da.ActivityDate)) AS day_of_week,  
       AVG(da.TotalSteps) AS avg_steps,  
       AVG(dc.Calories) AS avg_calories  
FROM `first-one-434623.Google_project.dailyActivity_merged` AS da  
JOIN `first-one-434623.Google_project.dailyCalories_merged` AS dc  
ON da.Id = dc.Id  
   AND da.ActivityDate = dc.ActivityDay  
GROUP BY da.Id, day_of_week  
ORDER BY day_of_week;
```



TRENDS AND RELATIONSHIPS:

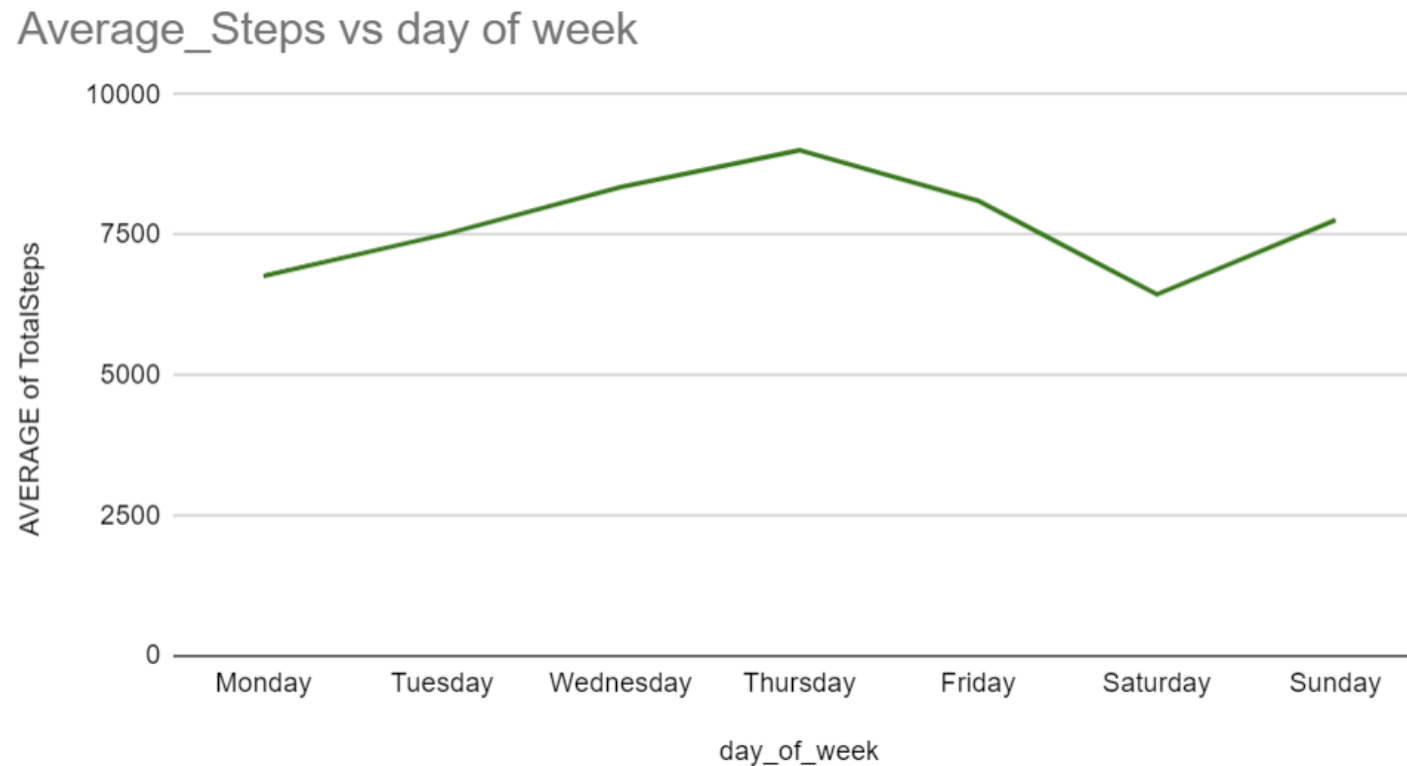
- Users generally show increased step counts, heart rates, and calorie burn during mid-morning and early evening.
- A positive correlation exists between daily steps and calories burned.
- Higher sedentary minutes are associated with lower average daily steps and calorie burn, highlighting potential areas to promote more active lifestyles.
- Activity levels tend to be lower on weekdays compared to weekends, suggesting users are more active on weekends.



RECOMMENDATIONS:

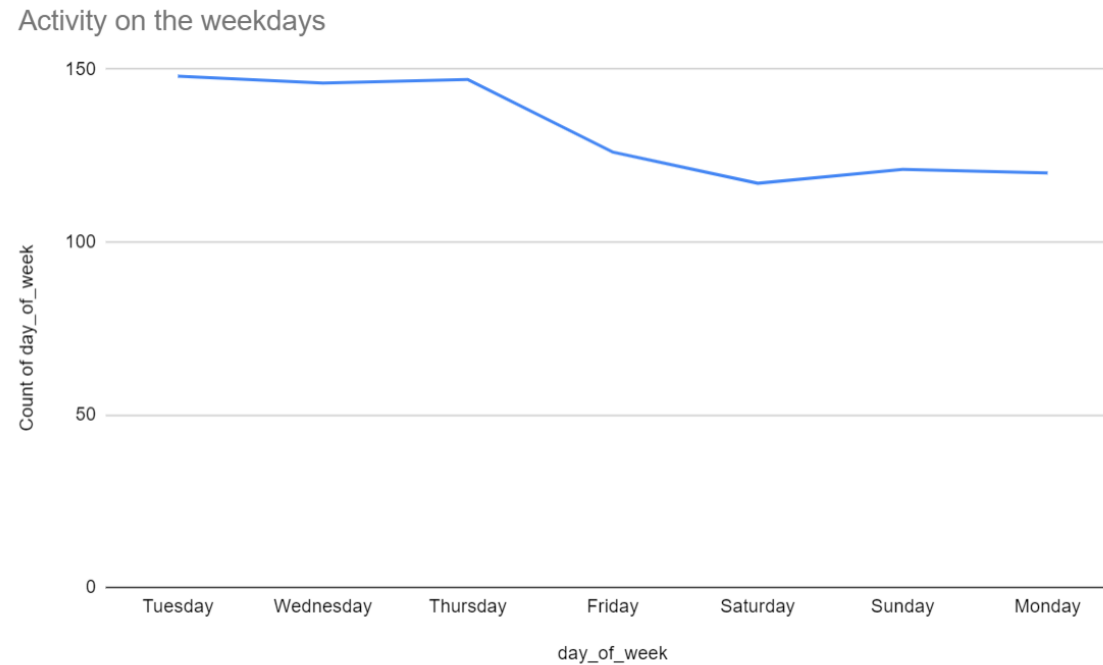
1- Optimize Marketing for Peak Activity Times

- Leverage insights on peak activity hours to craft targeted marketing campaigns, promoting the company's products during users' most active times.



2- Encourage Health Coaching and Wellness Challenges

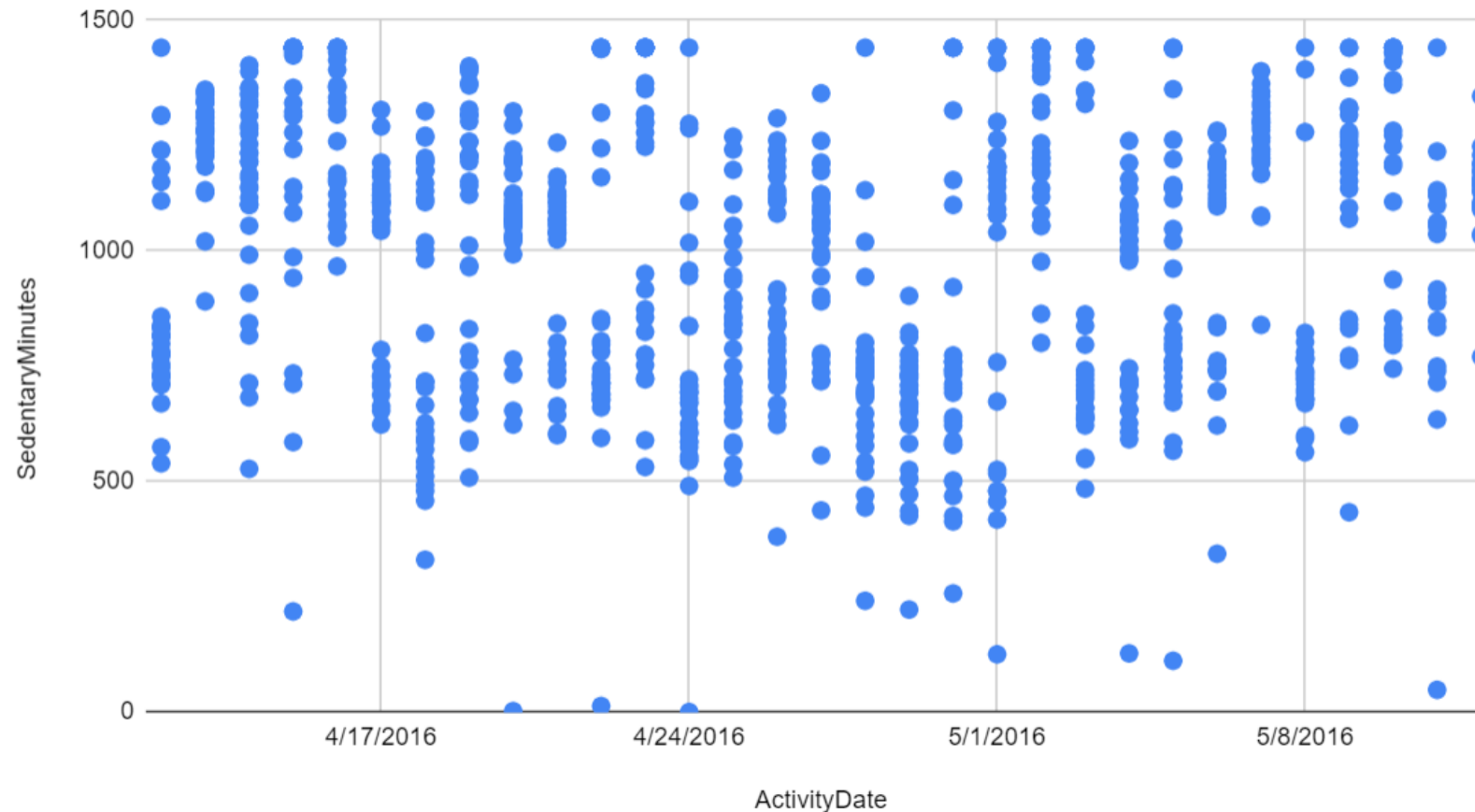
- Develop personalized wellness challenges that resonate with users' activity and sedentary patterns, such as weekend challenges to boost engagement.



3. Leverage Sedentary Data for Preventive Health Messaging

- The chart reveals prolonged sedentary periods, especially during work hours. By offering preventive health tips and micro-coaching, the company can position itself as a proactive wellness brand, appealing to health-conscious consumers.

SedentaryMinutes vs. ActivityDate



A series of white, thin, overlapping geometric lines on a black background, creating a complex, abstract pattern on the left side of the slide.

THANK YOU

Dhay Alsawat

dhayalsawat@gmail.com