How Can a Wellness Technology Company Play It Smart?

Scenario

Bellabeat is a high-tech manufacturer of health-focused products for women. It is a small company that has the potential to become a larger player in the global smart device market. Urška Sršen, co-founder and Chief Creative Officer of Bellabeat, believes that analyzing smart device fitness data could help unlock new growth opportunities for the company. The objective is to analyze smart device data to gain insight into how consumers are using their smart devices, which will then help guide marketing strategy for the company.

Overarching Questions

- What are some trends in smart device usage?
- How could these trends apply to Bellabeat customers?
- How could these trends help influence Bellabeat marketing strategy?

Phase 1: Ask

1.1 Business Task

Formulate recommendations for an effective marketing strategy using the insights discovered from analyzing smart device data.

1.2 Stakeholders

- Urška Sršen - Bellabeat's cofounder and Chief Creative Officer
- Sando Mur - Mathematician and Bellabeat's cofounder; key member of the Bellabeat executive team
- Bellabeat marketing analytics team - A team of data analysts responsible for collecting, analyzing, and reporting data that helps guide Bellabeat's marketing strategy
- Bellabeat app development team - A team of software developers focused on consistently improving the Bellabeat app in an attempt to meet user needs and preferences

Phase 2: Prepare

2.1 Data Source

The data source is the Fitbit Fitness Tracker Data shared by Mobius and stored on Kaggle. There are a total of 18 CSV files.

2.2 Data Format

It is considered long data since multiple rows belong to the same subject but at multiple points in time.

2.3 Data Credibility

The dataset exhibits sampling bias as it comprises only 33 participants. It is unclear whether random sampling was employed to mitigate bias since the eligibility criteria for participants were not explicitly disclosed on the Kaggle page. Notably, the dataset's Usability score on Kaggle is a perfect 10.00, reflecting a full score (100%) in the dimensions of completeness, credibility, and compatibility.

In evaluating the data by the ROCCC (Reliable, Original, Comprehensive, Current, & Cited) standard, the following criteria are as observed:

- Reliable - The dataset demonstrates a high level of credibility, as evidenced by its perfect score of 100% in the Credibility criterion on Kaggle. However, its reliability is questionable because the dataset includes data from only 33 participants over a brief two-month period in 2016. This limited sample size and narrow timeframe may not accurately represent the entire population of fitness smart device consumers in the year 2023.
- Original - The data's provenance can be traced back to its origins as crowd-sourced Fitbit datasets, which was initially hosted on the website zenodo.org.
- Comprehensive -The dataset cannot be considered comprehensive due to its limited scope, as it only surveyed 33 participants during a specific time frame from April 12, 2016, to May 12, 2016. The small sample size and narrow data collection period may not provide a representative and comprehensive view of the broader population or a more extended period..
- Current - Dataset was last updated in 2016. Not Current
- Cited - The dataset's provenance is attributed to crowd-sourced Fitbit datasets on zenodo.org, while the survey was conducted through Amazon Mechanical Turk.

2.4 Licensing, Privacy, Security, and Accessibility

- Licensing: The dataset has a CC0: Public Domain License which means that the owner relinquished all copyright and similar rights held and has dedicated those rights to the public domain.
- Privacy: The dataset does not include any personal information. A numerical identifier has been assigned to each individual.
- Security & Accessibility: It is open-source and accessible to the public. Since sensitive information is not included in the dataset, security does not stand out as an obvious issue.

2.5 Data Integrity

Upon employing the ROCCC method to review the data, it becomes evident that data integrity is compromised due to the absence of unbiased sampling. The inclusion of only 33 participants from a two-month period in 2016 does not provide a representative sample of the entire fitness smart device consumer population. Acknowledging this limitation and considering its significance, the analysis will proceed with the available data to extract insights.

Phase 3: Process

Tools:
Microsoft SQL Server Management Studio for data processing
Tableau for visualization

3.1 Data Cleaning

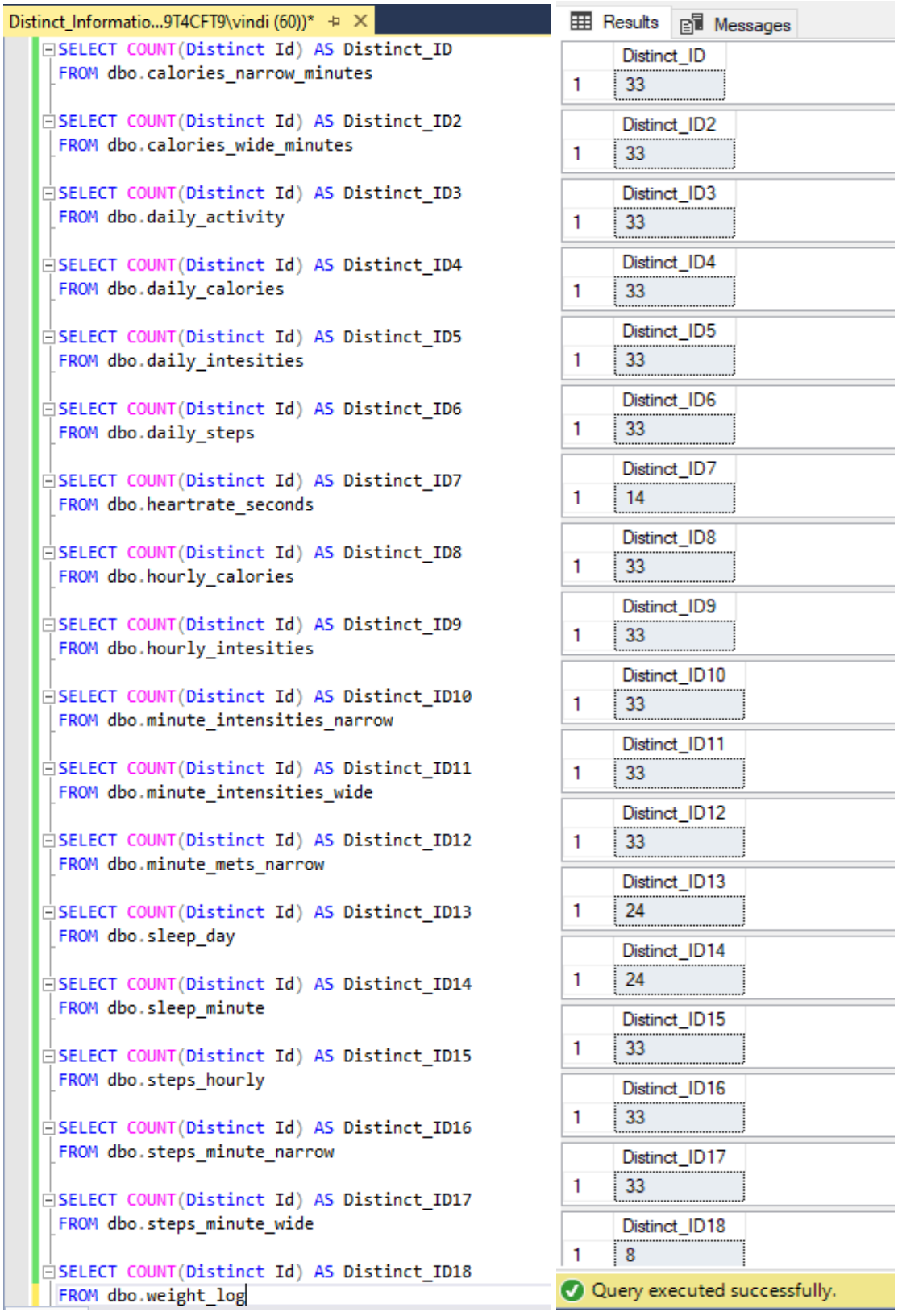
Uploaded 18 CSV Files into Microsoft SQL Server Management Studio locally.
Changed names of files for easier coding in SQL
Original File Names:

Name	Type	Compressed size	Password ...	Size	Ratio	Date modified
dailyActivity_merged	Comma Separated Values ...	26 KB	No	109 KB	77%	12/16/2020 2:20 PM
dailyCalories_merged	Comma Separated Values ...	5 KB	No	25 KB	81%	12/16/2020 2:20 PM
dailyIntensities_merged	Comma Separated Values ...	16 KB	No	69 KB	78%	12/16/2020 2:20 PM
dailySteps_merged	Comma Separated Values ...	6 KB	No	25 KB	79%	12/16/2020 2:20 PM
heartrate_seconds_merged	Comma Separated Values ...	8,381 KB	No	87,489 KB	91%	12/16/2020 2:20 PM
hourlyCalories_merged	Comma Separated Values ...	82 KB	No	783 KB	90%	12/16/2020 2:20 PM
hourlyIntensities_merged	Comma Separated Values ...	87 KB	No	878 KB	91%	12/16/2020 2:20 PM
hourlySteps_merged	Comma Separated Values ...	86 KB	No	778 KB	89%	12/16/2020 2:20 PM
minuteCaloriesNarrow_merged	Comma Separated Values ...	3,868 KB	No	64,887 KB	95%	12/16/2020 2:20 PM
minuteCaloriesWide_merged	Comma Separated Values ...	857 KB	No	22,455 KB	97%	12/16/2020 2:20 PM
minuteIntensitiesNarrow_merged	Comma Separated Values ...	3,132 KB	No	45,273 KB	94%	12/16/2020 2:20 PM
minuteIntensitiesWide_merged	Comma Separated Values ...	201 KB	No	3,235 KB	94%	12/16/2020 2:20 PM
minuteMETsNarrow_merged	Comma Separated Values ...	3,618 KB	No	46,570 KB	93%	12/16/2020 2:20 PM
minuteSleep_merged	Comma Separated Values ...	425 KB	No	8,641 KB	96%	12/16/2020 2:20 PM
minuteStepsNarrow_merged	Comma Separated Values ...	3,484 KB	No	45,442 KB	93%	12/16/2020 2:20 PM
minuteStepsWide_merged	Comma Separated Values ...	411 KB	No	3,400 KB	88%	12/16/2020 2:20 PM
sleepDay_merged	Comma Separated Values ...	4 KB	No	18 KB	83%	12/16/2020 2:20 PM
weightLogInfo_merged	Comma Separated Values ...	2 KB	No	7 KB	78%	12/16/2020 2:20 PM

New File Names:

Name	Status	Date modified	Type	Size
calories_narrow_minutes	✓	7/31/2023 9:59 PM	Comma Separate...	64,887 KB
calories_wide_minutes	✓	7/31/2023 9:59 PM	Comma Separate...	22,455 KB
daily_activity	✓	7/31/2023 9:59 PM	Comma Separate...	109 KB
daily_calories	✓	7/31/2023 9:59 PM	Comma Separate...	25 KB
daily_intesities	✓	7/31/2023 9:59 PM	Comma Separate...	69 KB
daily_steps	✓	7/31/2023 9:59 PM	Comma Separate...	25 KB
heartrate_seconds	✓	7/31/2023 9:59 PM	Comma Separate...	87,489 KB
hourly_calories	✓	7/31/2023 9:59 PM	Comma Separate...	783 KB
hourly_intesities	✓	7/31/2023 9:59 PM	Comma Separate...	878 KB
minute_intensities_narrow	✓	7/31/2023 9:59 PM	Comma Separate...	45,273 KB
minute_intensities_wide	✓	7/31/2023 9:59 PM	Comma Separate...	3,235 KB
minute_mets_narrow	✓	7/31/2023 9:59 PM	Comma Separate...	46,570 KB
sleep_minute	✓	7/31/2023 9:59 PM	Comma Separate...	8,641 KB
sleepDay_merged	✓	7/31/2023 9:59 PM	Comma Separate...	18 KB
steps_hourly	✓	7/31/2023 9:59 PM	Comma Separate...	778 KB
steps_minute_narrow	✓	7/31/2023 9:59 PM	Comma Separate...	45,442 KB
steps_minute_wide	✓	7/31/2023 9:59 PM	Comma Separate...	3,400 KB
weight_log	✓	7/31/2023 9:59 PM	Comma Separate...	7 KB

Used SQL to find and count the distinct Id in each file. By conducting this analysis, I was able to identify the tables that lack complete information.



```
SELECT COUNT(Distinct Id) AS Distinct_ID
FROM dbo.calories_narrow_minutes

SELECT COUNT(Distinct Id) AS Distinct_ID2
FROM dbo.calories_wide_minutes

SELECT COUNT(Distinct Id) AS Distinct_ID3
FROM dbo.daily_activity

SELECT COUNT(Distinct Id) AS Distinct_ID4
FROM dbo.daily_calories

SELECT COUNT(Distinct Id) AS Distinct_ID5
FROM dbo.daily_intesities

SELECT COUNT(Distinct Id) AS Distinct_ID6
FROM dbo.daily_steps

SELECT COUNT(Distinct Id) AS Distinct_ID7
FROM dbo.heartrate_seconds

SELECT COUNT(Distinct Id) AS Distinct_ID8
FROM dbo.hourly_calories

SELECT COUNT(Distinct Id) AS Distinct_ID9
FROM dbo.hourly_intesities

SELECT COUNT(Distinct Id) AS Distinct_ID10
FROM dbo.minute_intensities_narrow

SELECT COUNT(Distinct Id) AS Distinct_ID11
FROM dbo.minute_intensities_wide

SELECT COUNT(Distinct Id) AS Distinct_ID12
FROM dbo.minute_mets_narrow

SELECT COUNT(Distinct Id) AS Distinct_ID13
FROM dbo.sleep_day

SELECT COUNT(Distinct Id) AS Distinct_ID14
FROM dbo.sleep_minute

SELECT COUNT(Distinct Id) AS Distinct_ID15
FROM dbo.steps_hourly

SELECT COUNT(Distinct Id) AS Distinct_ID16
FROM dbo.steps_minute_narrow

SELECT COUNT(Distinct Id) AS Distinct_ID17
FROM dbo.steps_minute_wide

SELECT COUNT(Distinct Id) AS Distinct_ID18
FROM dbo.weight_log
```

Distinct_Informatio...9T4CFT9\vindi (60))*

Results | Messages

	Distinct_ID
1	33

	Distinct_ID2
1	33

	Distinct_ID3
1	33

	Distinct_ID4
1	33

	Distinct_ID5
1	33

	Distinct_ID6
1	33

	Distinct_ID7
1	14

	Distinct_ID8
1	33

	Distinct_ID9
1	33

	Distinct_ID10
1	33

	Distinct_ID11
1	33

	Distinct_ID12
1	33

	Distinct_ID13
1	24

	Distinct_ID14
1	24

	Distinct_ID15
1	33

	Distinct_ID16
1	33

	Distinct_ID17
1	33

	Distinct_ID18
1	8

Query executed successfully.

3.2 Data Merging/Joining

Used SQL to JOIN select files to look for statistically relevant information.

JOIN Calories and S...9T4CFT9\vindi (72))

```sql
SELECT dc.Id,
       SUM(dc.Calories) AS TotalCalories,
       SUM(ds.StepTotal) AS TotalSteps
FROM daily_calories dc
JOIN daily_steps ds ON dc.Id = ds.Id AND dc.ActivityDay = ds.ActivityDay
WHERE ds.StepTotal > 0
GROUP BY dc.Id
ORDER BY TotalCalories DESC;
```

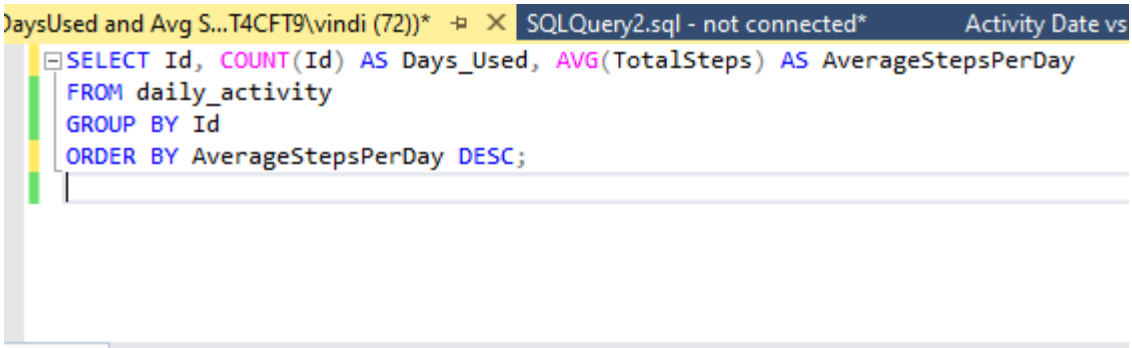
3.3 Data Formatting

Checked while uploading the format of the cells.
In regards to Id cells in each file they were either originally in INT format or BIGINT format. While Uploading to Microsoft SQL Server Management Studio each file had the Id cell formatted to BIGINT

Phase 4: Analyze

In the Analyze stage I continued using SQL to create csv files to use in Tableau to visualize. I began searching for answers to the first business question,"What are some trends in smart device usage?"

Usage and Average Steps:



	Id	Days_Used	AverageStepsPerDay
1	8877689391	31	16040
2	8053475328	31	14763
3	1503960366	31	12116
4	2022484408	31	11370
5	7007744171	26	11323
6	3977333714	30	10984
7	4388161847	31	10813
8	6962181067	31	9794
9	2347167796	18	9519
10	7086361926	31	9371
11	8378563200	31	8717
12	5553957443	31	8612
13	4702921684	31	8572
14	5577150313	30	8304
15	4558609924	31	7685
16	2873212765	31	7555
17	1644430081	30	7282
18	4319703577	31	7268
19	8583815059	31	7198
20	6117666160	28	7046
21	3372868164	20	6861
22	8253242879	19	6482
23	1624580081	31	5743
24	6290855005	29	5649
25	2026352035	31	5566
26	4445114986	31	4796
27	2320127002	31	4716
28	4057192912	4	3838
29	1844505072	31	2580
30	6775888955	26	2519
31	4020332650	31	2267
32	8792009665	29	1853
33	1927972279	31	916

I noticed from our distinct search for ids earlier that some tables had 9 users not use the device during bedtime. And analyzed the data to see if there was any trend there.

SQLQuery5.sql - DE...9T4CFT9\vindi (69))* | TotalSleepsbyuser.s...9T4CFT9\vindi (55))

```sql
SELECT s.Id, s.DaysWornSleeping, d.Days_Used
FROM (
    SELECT Id, COUNT(Id) AS DaysWornSleeping
    FROM dbo.sleep_day
    GROUP BY Id
) s  -- The alias "s" for the subquery calculating DaysWornSleeping
JOIN (
    SELECT Id, COUNT(Id) AS Days_Used
    FROM dbo.daily_activity
    GROUP BY Id
) d  -- The alias "d" for the subquery calculating Days_Used
ON s.Id = d.Id;
```

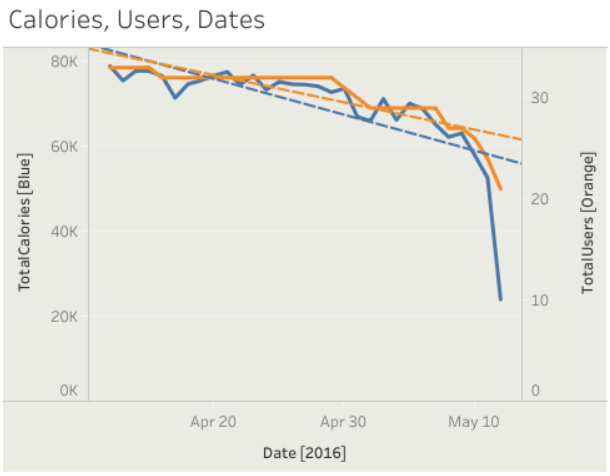
100 %

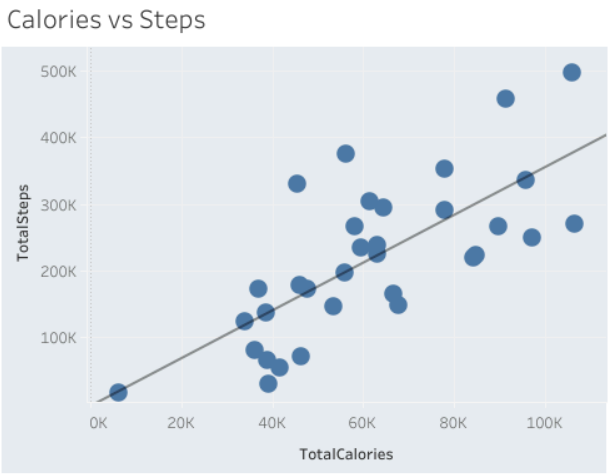
Results | Messages

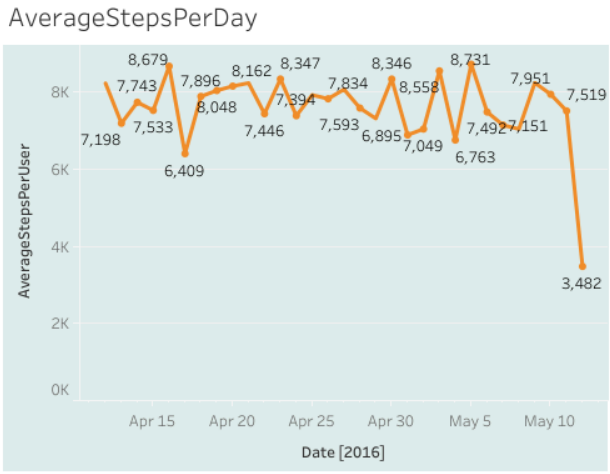
	Id	DaysWornSleeping	Days_Used
1	1503960366	25	31
2	1644430081	4	30
3	1844505072	3	31
4	1927972279	5	31
5	2026352035	28	31
6	2320127002	1	31
7	2347167796	15	18
8	3977333714	28	30
9	4020332650	8	31
10	4319703577	26	31
11	4388161847	24	31
12	4445114986	28	31
13	4558609924	5	31
14	4702921684	28	31
15	5553957443	31	31
16	5577150313	26	30
17	6117666160	18	28
18	6775888955	3	26
19	6962181067	31	31
20	7007744171	2	26
21	7086361926	24	31
22	8053475328	3	31
23	8378563200	32	31
24	8792009665	15	29

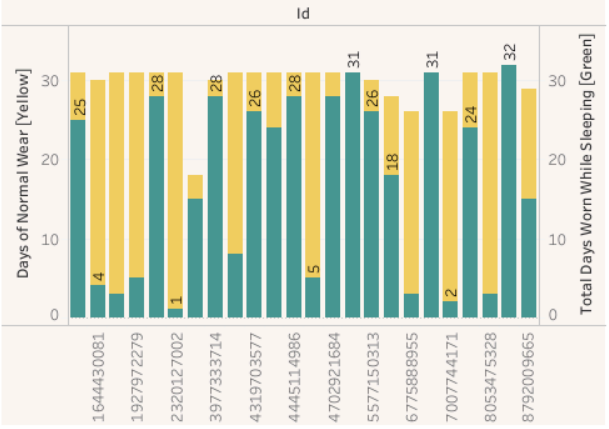
Phase 5: Share

Utilizing Tableau, an integrated dashboard was meticulously designed to present a comprehensive collection of visualizations. These visualizations encompass a dual-axis line chart, a scatterplot graph, a standard line graph, and a dual-axis bar chart, all seamlessly combined into a single, cohesive interface.









Phase 6: Act

Answering the Questions:

6.1 What are some trends in smart device usage?

1. A notable observation within the user base is the variance in the utilization of the fitness tracker's sleep monitoring feature. While some users opt to wear the device continuously, others refrain from using it during sleep. Remarkably, there were nine users who did not record any monitored sleep data. Among the users who did engage

with this functionality, the frequency ranged from occasional usage (once) to consistent daily monitoring.
2. Total Steps and Total Calories exhibit a positive correlation, highlighting the direct relationship between physical activity and caloric expenditure. This finding has important implications for health and fitness strategies.
3. During the course of the tracking period, a discernible trend emerged, indicating a decrease in Consistent Daily Usage over the last week. Users' engagement with the fitness tracker exhibited a notable decline during this temporal interval, prompting further investigation into the factors contributing to this observed pattern.

6.2 How could these trends apply to Bellabeat customers?

End of the Month Drop in Daily Usage:

Bellabeat Ivy fitness tracker might observe a similar trend of decreased usage towards the end of the month due to factors such as users' busy schedules, declining motivation, or lapses in tracking habits. Understanding this trend can help Bellabeat Ivy to implement strategies to keep users engaged and motivated throughout the month.

Steps vs. Calories Burnt:

The positive correlation between steps and calories burnt suggests that users of Bellabeat Ivy who take more steps tend to burn more calories. This finding could help Bellabeat Ivy emphasize the significance of step count in achieving fitness and health goals, encouraging users to set step-based targets for better calorie management.

Different Usage Patterns (Day and Night):

The data indicating that not all Bellabeat Ivy users track their activities both day and night can provide insights into diverse user preferences and lifestyles. Bellabeat Ivy could use this information to offer different tracking modes or features, catering to users who prefer to track specific activities during certain times of the day.

Targeted Marketing and Features:

Bellabeat Ivy can use insights from user behavior to design targeted marketing campaigns and develop features that resonate with users. For instance, if users tend to be more active during specific times, Bellabeat Ivy could promote relevant activities or challenges during those times to boost engagement.

User Retention and Engagement:

Addressing the drop in daily usage towards the end of the month can be a priority for Bellabeat Ivy to enhance user retention. Implementing engagement strategies like personalized

recommendations, challenges, rewards, or reminders can help maintain user motivation and commitment.

Bellabeat Ivy Product Differentiation:

Analyzing user data can help identify unique selling points for the Bellabeat Ivy fitness tracker. If the data shows a strong correlation between specific activities and calorie burning, Bellabeat Ivy could position itself as specialized for those activities, distinguishing it from other fitness trackers in the market.

User Insights and Feedback:

Continuously monitoring user trends and feedback can lead to product improvements and innovation for Bellabeat Ivy. Analyzing the data can help identify user pain points and work towards enhancing the overall user experience.
By leveraging data insights and addressing user preferences, Bellabeat Ivy can offer a more personalized, engaging, and effective fitness tracking experience, ultimately improving user satisfaction and retention.

6.3 How could these trends help influence Bellabeat marketing strategy?

End of the Month Drop in Daily Usage:

Marketing Campaigns: Bellabeat can design targeted marketing campaigns aimed at increasing user engagement and motivation during the end of the month. These campaigns could include special challenges, rewards, or promotions to keep users active and committed to their fitness goals.

Reminders and Notifications: Implementing automated reminders and notifications towards the end of the month can encourage users to stay consistent with their tracking and fitness routines.

Steps vs. Calories Burnt:

Emphasize Step Count: Based on the positive correlation between steps and calories burnt, Bellabeat can place a strong emphasis on step count in its marketing messaging. Highlighting the impact of steps on overall calorie management and fitness achievements can motivate users to set and achieve step-based targets.

Step-Based Challenges: Bellabeat can create step-based challenges or competitions within its app to foster a sense of achievement and encourage users to strive for higher step counts.

Different Usage Patterns (Day and Night):

Customizable Tracking Modes: Bellabeat can promote the flexibility of its fitness tracker by offering customizable tracking modes that cater to users' preferred activity tracking times, whether it's day or night.

Highlight Unique Features: Marketing efforts can showcase features that allow users to seamlessly switch between day and night tracking modes, addressing diverse user needs.

Targeted Marketing and Features:

User Segmentation: Bellabeat can leverage the insights to segment its user base based on activity patterns and preferences. Tailored marketing messages and features can be created for different user segments to improve engagement.

Personalized Recommendations: By analyzing user behavior, Bellabeat can offer personalized activity recommendations and tips to users based on their unique tracking data.

User Retention and Engagement:

Loyalty Programs: Bellabeat can introduce loyalty programs, offering rewards and incentives to long-term users who consistently engage with the fitness tracker.

Social Community: Building a supportive and active user community can enhance engagement. Bellabeat can encourage users to share their achievements and progress with others, fostering a sense of community and motivation.

Bellabeat Ivy Product Differentiation:

Marketing Messaging: Based on unique insights, Bellabeat can differentiate its product through marketing messaging that emphasizes its specialized features and benefits related to specific activities.

Influencer Partnerships: Collaborating with fitness influencers who align with the tracker's specialized offerings can help amplify the product's unique selling points.

User Insights and Feedback:

Continuous Improvement: Bellabeat can use user insights and feedback to drive product improvements and feature updates. Marketing efforts can highlight the brand's commitment to listening to users and enhancing their experience.
By incorporating these trends into its marketing strategy, Bellabeat can strengthen user engagement, improve user retention, and position itself as a fitness tracker that understands and caters to users' individual preferences and goals.