Case Study 2: How Can a Wellness Technology Company Play It Smart?

You are a junior data analyst working on the marketing analyst team at Bellabeat, a high-tech manufacturer of health-focused products for women. Bellabeat is a successful small company, but it 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. You have been asked to focus on one of Bellabeat's products and analyze smart device data to gain insight into how consumers are using their smart devices. The insights you discover will then help guide the marketing strategy for the company. You will present your analysis to the Bellabeat executive team along with your high-level recommendations for Bellabeat's marketing strategy.

To accomplish my task, I will follow the steps of the data analysis process: Ask, Prepare, Process, Analyze, Share, and Act.

Tools: Microsoft Excel and Microsoft SQL Server. Tableau is used for visualization.

Introduction

Welcome to the Bellabeat data analysis case study! In this case study, you will perform many real-world tasks of a junior data analyst. You will imagine you are working for Bellabeat, a high-tech manufacturer of health-focused products for women, and meet different characters and team members. To answer the key business questions, you will follow the steps of the data analysis process: ask, prepare, process, analyze, share, and act. Along the way, the Case Study Roadmap tables — including guiding questions and key tasks — will help you stay on the right path. By the end of this lesson, you will have a portfolio-ready case study. Download the packet and reference the details of this case study anytime. Then, when you begin your job hunt, your case study will be a tangible way to demonstrate your knowledge and skills to potential employers.

Step 1: Ask

Examine fitness data from smart devices to extract valuable insights about consumer behavior, pinpointing trends that can shape the marketing strategy for Bellabeat During the ask phase of the data analysis process, one must acknowledge the existing problem, articulate the business task, take relevant stakeholders into account, and determine how insights and findings will benefit these stakeholders.

The key stakeholders for this project include:

- **Urška Sršen**: Bellabeat's co-founder and Chief Creative Officer

- **Sando Mur**: Mathematician and Bellabeat co-founder; a 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

These three questions will guide our analysis:

- What are some trends in smart device usage?

- How could these trends apply to Bellabeat customers?

- How could these trends help influence Bellabeat's marketing strategy?

Step 2: Prepare

Why SQL and Tableau?
We have opted for Microsoft SQL Server Management Studio for our data analysis due to their strong focus on data integrity, which guarantees the reliability of our analysis results. By using this studio, we facilitate complex querying and data manipulation, a crucial aspect of data analysis. Moreover, it provides robust support for various data types, making it adaptable to different data sources and formats. Additionally, this tool efficiently handles large datasets, allowing us to analyze substantial volumes of information effectively.

For visualization we are using Tableau, another powerful tool for data visualization and reporting because it provides powerful, user-friendly tools to create interactive and insightful visualizations, helping us convey our data analysis results effectively to both internal and external stakeholders.

Data Sources

The datasets used in this analysis were sourced from Mobius on Kaggle and are covered under the CCO Public Domain license. These datasets were compiled from responses collected through a distributed survey conducted via Amazon Mechanical Turk during the period of March 12, 2016, to May 12, 2016.

The datasets downloaded were in a zip file from which I extracted 18 CSV files representing 18 datasets but of these datasets, I will be using 10.

Step 3: Process

The following CSV files were used for analysis.

- dailyactivity_merged
- dailycalories_merged
- dailyIntensities_merged
- dailysteps_merged
- hourlycalories_merged
- hourlyintensities_merged
- hourlysteps_merged
- minutesleep_merged
- sleepday_merged
- weightLogInfo_merged
-

Cleaning was done using Excel. The following steps were taken for each dataset. I began by organizing and refining the data, followed by examining for any empty fields. Then, I scrutinized for any repeated values, resulting in the removal of 543 duplicates from the minutesleep_merged dataset, leaving 187978 entries, and 3 duplicates from the sleepday_merged dataset, leaving 410 unique entries. Subsequently, I standardized the date and time columns to MM/DD/YYYY and hh:mm:ss formats respectively and separated the combined date/time column using INT(). Lastly, I eliminated rows containing data not conforming to the required format. Step 4: Analyze and share.

In the fourth phase, data analysis takes center stage. The primary objective here is to uncover trends, patterns, and relationships within the data, all aimed at addressing the core business objective. This involves delving deeper into the data, sorting it, and

formatting it in a way that elucidates the behavior of various groups both independently and in correlation with each other. The aim is to derive meaningful insights from the data, effectively telling the story that it holds.

A. SUMMARY STATISTICS

We understood the type of data and how they are related to the findings from the process step. Now, let's get some high-level summary statistics that can show us how these tables are related to one another, as well as what trends might exist within each.

Unique participants:

How many unique participants are there at each table?

There are 33 users (one user per unique ID) in the daily activity tab.

33 respondents submitted their tracking device data from 12/04/2016 to 12/05/2016 which is 31 days in total. On average the device was used for 29 days with only one responding that had used the device for only 4 days. All respondents used their tracking devices to record:

Of the 33 respondents surveyed, 24 individuals, constituting 73% of the total, utilized devices capable of tracking their sleeping patterns. On average, these users monitored their sleep for 17 days. Additionally, 14 respondents, representing 42%, recorded their heart rate, while only 8 respondents, accounting for 24%, provided their weight measurements. The graphical representation of this data is illustrated in

Other Findings

User Types by Total Steps

According to a Healthline.com article by Sara Lindberg in 2019, referencing a study by

Tudor-Locke et al. from 2011 titled "How many steps/day are enough? for adults," a target of 10,000 steps per day is considered reasonable for maintaining health in adults. Lindberg (2019) categorizes activity levels based on steps into three groups:

    Inactive: Less than 5,000 steps per day

    Average (somewhat active): Ranges from 7,500 to 9,999 steps per day

    Very Active: More than 12,500 steps per day

This breakdown provides a framework for understanding activity levels and their impact on health.

I plan to categorize users within the Daily Activity dataset based on their activity levels using the step categories mentioned in the Healthline article. However, I noticed that these categories have some gaps. Therefore, I've added two additional categories:

Low Active User: 5,000 to 7,499 steps

Active User: 10,000 to 12,499 steps

By incorporating these additional categories, I aim to provide a more comprehensive breakdown of user activity levels within the dataset.
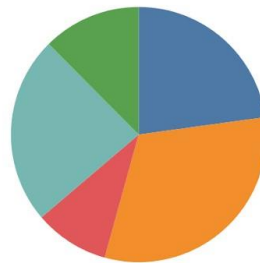
AvgTotalSteps
248,118

UserType
■ Active User
■ Average Active User
■ Inactive
■ Low Active User
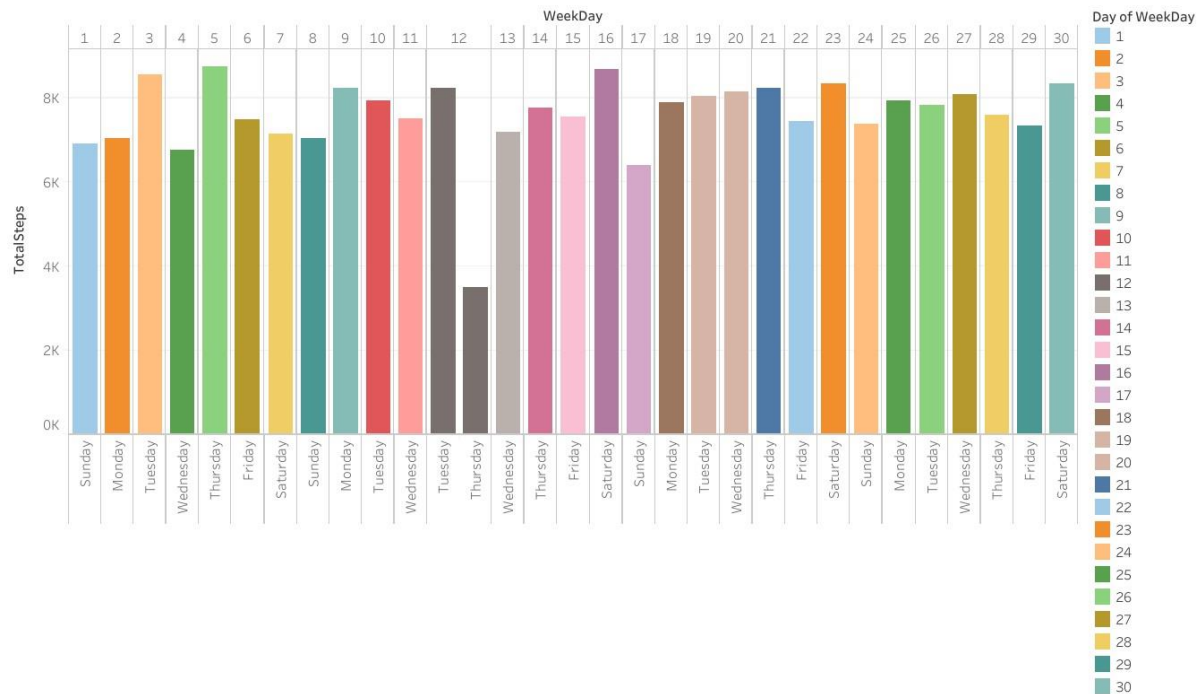■ Very Active User

Here are the Results:

- Inactive User: 8 users

- Low Active User: 9 users

- Average Active User: 9 users

- Active User: 5 users

- Very Active User: 2 users

Total Steps by Day

Look at average steps by day to see if users were more active on certain days of the week.

Average Steps by Weekdays

After running the query, there wasn't a whole lot of difference between each day in terms of average steps. With that said, Saturday had the highest average steps as well as the beginning of each week (Monday and Tuesday). We could potentially infer from this that the users wanted to be more active right after the weekend of rest (Sunday with the lowest total steps & Friday not too far behind) & that Saturday allowed for more time for activity & movement.

Average Hourly Intensity

The average total steps per day is 7628, falling below the recommended 10,000 steps per day.

The average sedentary minutes per day is 991.2, representing the highest average among all activity minutes, indicating that more than 50% of the time is spent sedentary.

The average lightly active minutes per day is 192.8, the highest among all active minutes per day.

On average, individuals spend 91.5% of their time sleeping in bed, calculated by dividing the average total minutes asleep (419.5) by the average total time in bed (458.6).

The average number of calories burnt per hour is 97.39.

The figure below shows the time when the respondents were mostly active by displaying average hourly intensity. Intensity measures the physical activity used by Fitabase (Fitabase, 2018). It is classified as follows:

0 = Sedentary,1 = Light,2 = Moderate,3 = Very Active To get the total intensity these values are added up for all respondents in a particular hour and then divided by the number of respondents to get the average. The higher the Average intensity, the more physically active the respondents were. The average hourly intensity is above 21 from 17:00 to after 19:00 but then drop from 20:00 onwards. This means most respondents were involved in vigorous physical activity from 5 PM to just before 8 PM. The second highest average intensity is from 12:00 to just after 14:00. Another pattern to take note of is that activity starts picking up slowly from 5 AM and reaches the peak at 12 PM and falls slightly.

Average Hourly Intensity
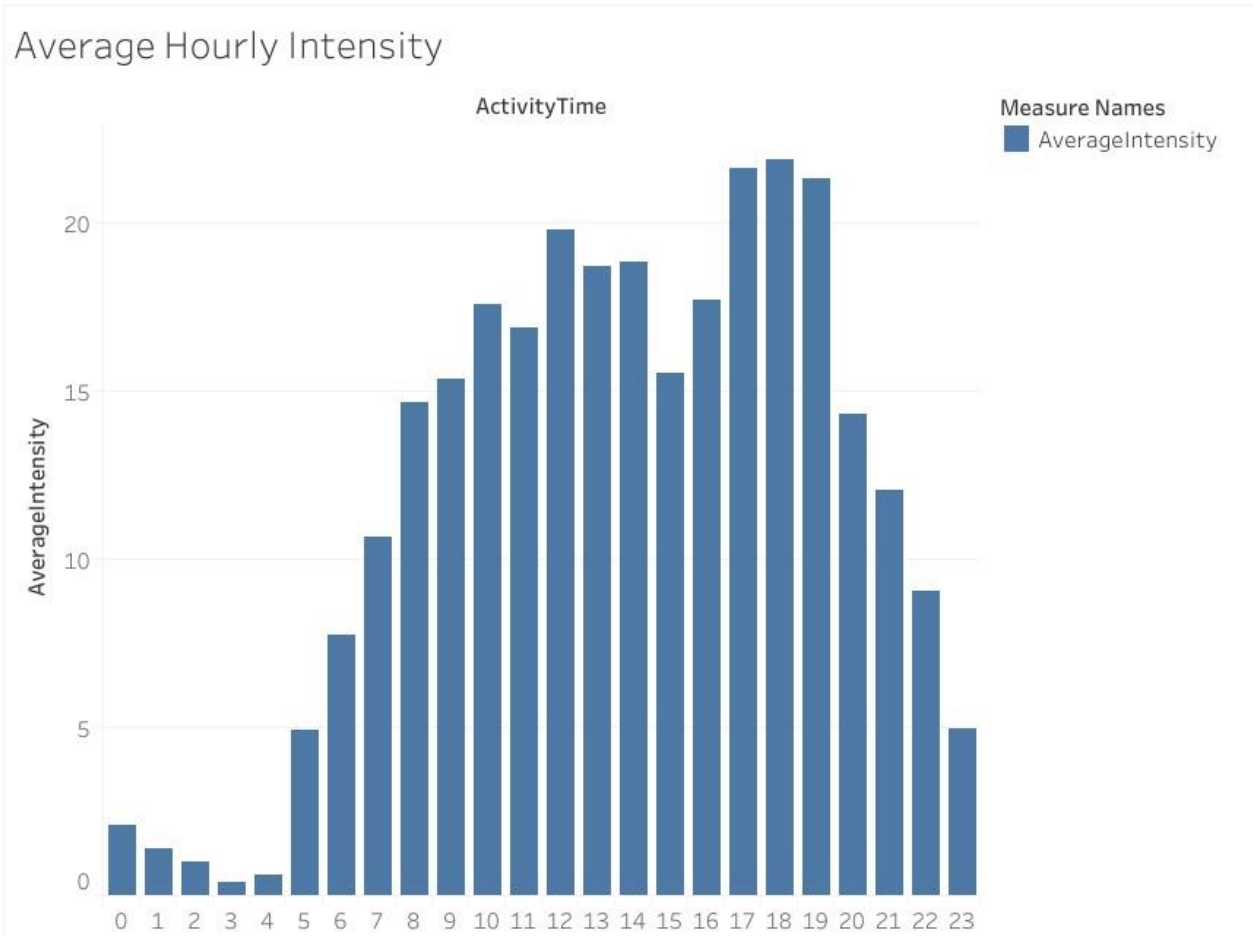
Daily Average Steps

 Average intensities are paired with METs which stands for metabolic equivalents. MET is defined as the amount of oxygen consumed while at rest, that is 1 MET (Healthline, 2019). For example, 2 MET means a person is consuming two times the amount of oxygen consumed while at rest. Classifications of METs are as follows: • <3 METs are light activities • 3 to 6 METS are Moderate activities • 6 METs or more are vigorous activities Highest MET recorded is 15.7 however on average the maximum MET is less than 2 METs because only a few minutes were spent doing physical activity

Respondents clocked an average of 7600 steps a day which is about 5.5 km a day and burnt an average of 2300 kilocalories a day (estimated energy expenditure). Only 7 respondents traveled more than 10 000 steps a day (8 km) on average (refer to Figure 6 below) which is the recommended number of steps per day (Jennifer Huizen, 2021). The response for dataset 2 clocked 8284 steps per day on average and the response for dataset 3 clocked 6849 steps per day on average.

Daily Average Steps
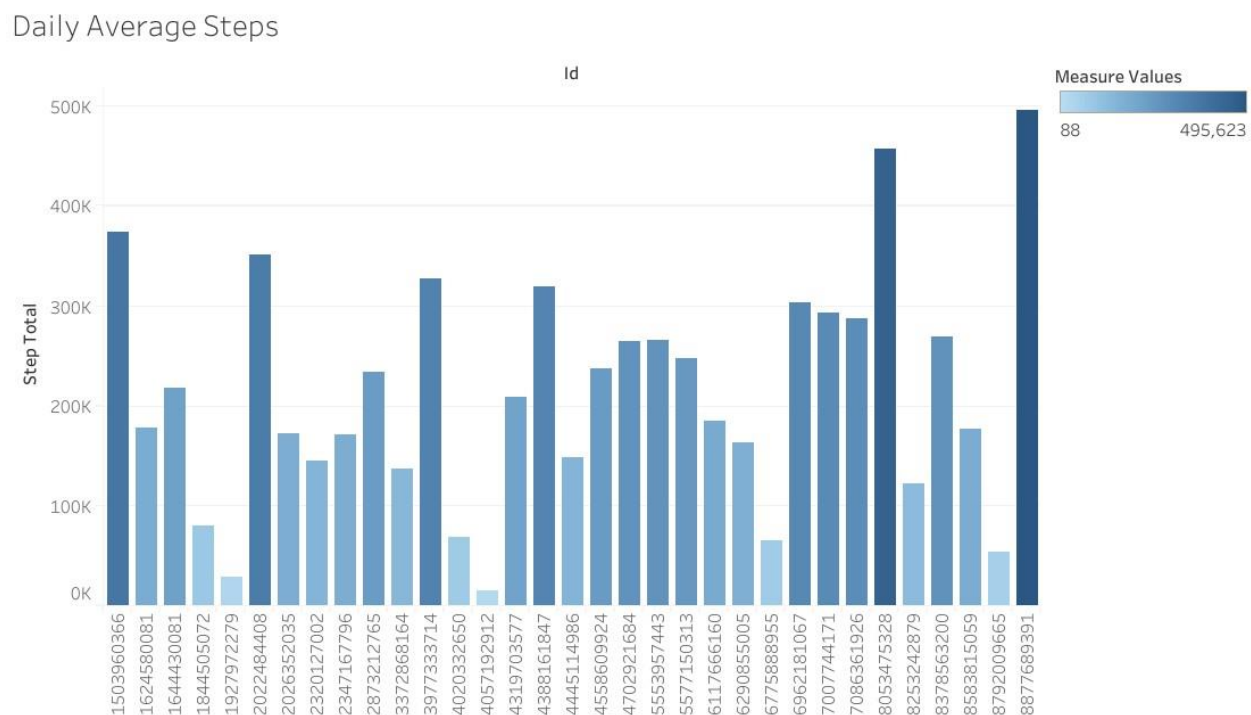
Visualization of average total steps against average total minutes slept to see any type of correlation

Conclusion

The prevalence of female users in tracking devices presents Bellabeat with a strategic advantage, especially considering the projected growth of the market. While the sample size

analyzed in this study is relatively small compared to the overall user population, leveraging additional data from literature studies can enhance Bellabeat's marketing strategy. Below are the key conclusions drawn from the analysis.

The majority of respondents demonstrated consistent usage of health tracking devices, monitoring various metrics such as calories, steps, physical activity, and duration. Interestingly, while 72% tracked their sleep patterns, 42% monitored their heart rate, and only 24% measured their weight. Notably, respondents tended to be more inactive overall. Active periods were observed from 5:00 p.m. to just before 20:00 and midday from noon to just before 15:00. On average, respondents logged approximately 7,600 steps per day, falling below the recommended threshold. However, those who tracked their sleeping patterns generally met the recommended sleep duration. Additionally, among respondents tracking heart health, 42% exhibited a normal heart rate.

Share

The main recommendations for Bellabeat are as follows:

- Encourage users to integrate health-tracking devices into their daily routines beyond just gym sessions. Bellabeat's marketing campaigns should emphasize not only health benefits but also the fashion-forward and stylish nature of their products.

- Prioritize the protection of users' privacy as these devices become integral parts of their lives.

- Leverage social media platforms for advertising, recognizing that most users of health devices are young adults.

- Optimize advertising time slots, targeting early mornings when people prepare for work and afternoons when they head to the gym.

- Provide value-added services such as:

  - Health campaigns promoting holistic lifestyles and offering comprehensive health.
    insights.
  - Highlight the importance of heart rate monitoring and provide alerts for critical
    thresholds, with the option to link devices to medical services.
  - Send timely training reminders during peak activity periods.
  - Ensure devices have long-lasting batteries for extended use.
  - Offer clients comprehensive reports and lifestyle improvement tips.
  - Implement a rewards system to incentivize clients using devices to enhance their health.

These recommendations aim to enhance Bellabeat's market presence and better serve its
clientele.