**Summary:**

Bellabeat is a high-tech manufacturer of health-focused products for women. They offer smart devices and apps that collect data on activity, sleep, fitness, and diet to help women track their personal health and make health-conscious decisions.

This case study will focus on analyzing smart devices’ health data to determine new avenues of growth and improved customer satisfaction for Bellabeat. The primary product of this case study will be the Bellabeat app, which provides users with health data related to their activity, sleep, stress, menstrual cycle, and mindfulness habits.

**Business Task:**

To analyze smart device usage data to identify trends and apply them to a selected Bellabeat product in order to inform the company's marketing strategy team and provide a summary of the analysis and top recommendations based on the findings.

The Stakeholders are

* Urška Sršen – Bellabeat co-founder and Chief Creative Officer
* Sando Mur - Bellabeat cofounder and a key member of the Bellabeat executive team
* Bellabeat Marketing Analytics team

**Data Organization**

The data used in this case study was sourced from FitBit Fitness Tracker Data. The datasets were stored in Kaggle and made available through Mobius.

The dataset has tracked information on thirty Fitbit users who have consented to submit personal tracker data, including minute-level outputs for physical activity, heart rate, and sleep monitoring.

The dataset contains 18 .csv files that are either organized in both wide and long formats. The contents of each dataset are as follows:

|  |  |
| --- | --- |
| Data Table | Description |
| Daily Activity | Tracks Activity of 33 users over 31 days. Variables include Total Steps, Calories Burned, Total Distance, Activity Levels |
| Daily Calories | Tracks Calories Burned by 33 users over 31 days. |
| Daily Intensities | Tracks intensity of exercise in minutes and distance of 33 users over 31 days. Categories: Sedentary, Lightly Active, Fairly Active, Very Active |
| Daily Steps | Tracks steps taken in a day for 33 users over 31 days |
| Heartrate Seconds | Tracks heart rate of each minute for 7 individuals |
| Hourly Calories | Hourly Calories burned over 31 days for 33 users |
| Hourly Intensities | Hourly total and average intensity over 31 days of 33 users. Range from 0-180 and 0-3 respectively. |
| Hourly Steps | Hourly Steps over 31 days of 33 users |
| Minute Calories | Calories burned every minute over 31 days for the 33 users |
| Minute Intensities | The intensity of Activity by Minute over 31 days for 33 users. Max is 3, thus it seems intensity is measured ordinally, with 0 being sedentary and 3 being very active. |
| Minutes METs | The Metabolic Rate compared to the resting rate by minute over 31 days for 33 users. In this study, the resting rate seems to be valued at 10, so all ratios should be divided by 10 for use in the analysis. |
| Minutes Sleep | Tracks all times an individual is asleep for 24 individuals. Range of 1-3. Assumed 1 is light sleep with 3 being deep sleep. 0 should represent an individual who is awake, and it is not included in the dataset. |
| Minutes Steps | Tracks Steps taken per minute for 33 users over 31 days. |
| Day Sleep | Tracks Total Sleep, Number of Sleep Sessions, and Total Time in bed per day for 33 users over 31 days |
| Weight Log | Tracked the Weight and calculated the BMI of 8 individuals. Only 3 individuals recorded more than 2 logs, and from them, only two individuals manually recorded their logs. |

**Official Nomenclature**

The above table’s descriptions were defined by what was available in the data. However, Fitbit also fully explains the nomenclature of their captured data in the Fitabase Data Dictionary.

The dictionary also mentions that data for calorie intake is also collected, thus that can be a consideration even though we currently do not have the information in our datasets.

It is also important to note that the vast majority of data in our dataset appears to have been passively collected rather than actively logged, which could influence downstream recommendations.

**Data Integrity and Preliminary Biases**

There are concerns about the potential bias for the generalizability of the dataset, as the sample size is both low and volunteer-based, which could result in confounding based on the qualities of women likely to volunteer. There also is very little demographic data which could result in increased variability in results in addition to sampling bias.

**Licensing**

The data is licensed under CC0 or Public Domain, thus the owner has waived all rights to their work worldwide under copyright law. Thus, we can copy, modify, distribute, and perform the work, even for commercial purposes, without permission needed.

**Processing**

Since the dataset is relatively small, using an integrated development environment such as RStudio would make tracking changes for cleaning, analysis, and visualization much clearer for presentation. I also believe that pipes are an incredibly powerful tool to make code look neat and legible

**Data Cleaning and Analysis Sharing**

**-**continue in R Markdown

**Recommendations:**

Bella beat’s smart devices provide both a stylish and simple way of monitoring data for women’s health.

In order to expand on the possibilities of growth and improved customer service, my first recommendation would be to request access to the in-house personal databases for further analysis. The current datasets were small samples that ranged between 8 to 33 individuals and may be subject to bias. There were also many instances of reduced usage, or a limited usage of all the features that the Bellabeat app could provide such as sleep logs, weight logs, and calorie burn.

Thus, based on the trends found in the analysis, we recommend the following to target a wider audience and improve overall app usage.

|  |  |
| --- | --- |
| Recommendation | Rationale |
| Sleep with Bellabeat | Only a few people sleep with Bellabeat on. Many who do sleep, only do so for short naps. Reasons could be that the device is uncomfortable, or they take it off prior to washing and just don’t put it back on. An information campaign on Bellabeat’s battery and water-proof qualities could mitigate that and increase usage. |
| Weight Tracking Promotions | Only a few of the population tracked their weight. So, either this means people don’t want to, or they don’t have the opportunity to. Bellabeat could potentially increase usage, data, and sales by partnering with a company that sells device adaptable scales in addition to devices. Scales that sync with the app, would also improve convenience, and reduce usage, especially for women with weight loss goals. |
| Leverage Influencers to target the not-so-healthy | Even as a health product, a significant majority of Bellabeat’s users have either sedentary or lightly active lifestyles. An emphasis on the style of Bellabeat’s products, and the sexy UI should help engage with younger women, who are already Bellabeat’s target demographic. |
| Food Journaling | One of the biggest issues for Bellabeat is that it does not have a caloric intake section to match the caloric burn. The next biggest competitor FitBit has provided this as a service. While the app does provide meal recommendations, these are typically time intensive and easily skippable. This addition would increase use cases for the devices and should improve overall usage. |
| Belle-Alarm | A significant portion of users do not wear their Bellabeat’s every day. And an even greater amount do not sleep with them. This can be mitigated with additional functionality or an alarm that users could set. Not only would it wake them up, but it would also bring attention to the Bellabeat so that they are nudged into putting it on before starting their day. |

Overall, most of these recommendations are functionality based to improve utility and use cases. The recommendations intentionally omitted additional notifications or inconveniences that could reduce brand sentiment and customer satisfaction. The biggest recommendation, however, is to receive access to a larger pool of data with additional parameters to better analyze the population to best meet their needs along with the companies.