**Introduction:**

Bellabeat is a high-tech company, founded in 2013 by Urska Srsen and Sando Mur, that focuses on development and manufacturing of smart devices targeted to women that promote their customers’ health and wellness with features such as tracking of daily activity, sleep, stress and hydration levels as well as reproductive health. Urska Srsen and Sando Mur capitalized on their talents in art and mathematics, respectively, to collaborate on the design and technical elements to create their first-class product line.

The company’s smart products include the Bellabeat apps: Leaf, which is a tracker that can be worn as a bracelet, necklace or clip; Time, which is worn as a classically styled watch; and Spring, a unique water bottle with dual functionality as a hydration tracker. All these devices connect directly to the BellaBeat app to provide customers with data analytics such as their activity, sleep, menstrual cycles, stress management and mindfulness habits. These smart devices allow women to take control of their health and well-being and look fantastic while doing it. BellaBeat also currently offers a subscription-based membership program that provides customers with around-the-clock access to personalized guidance in such areas as sleep, nutrition, activity and mindfulness, as well as health and beauty. This guidance is uniquely tailored to each individual’s personal lifestyle and goals in which they desire to accomplish.

Bellabeat has been actively taking advantage of multiple advertising platforms, both traditional forms and more modern-day advertising tactics, such as maximizing social media presence. Urska recently has approached her marketing analytics team to ask for an analysis of consumer data to reveal more opportunities for growth. She strongly feels that by analyzing user data from other smart device products on the market, the company can glean insights into user trends to assist in identification of growth opportunities for BellaBeat. She has asked the team to analyze a dataset made available through Mobius and uploaded from Kaggle. The dataset is titled, “FitBit Fitness Tracker Data”.

**Key stakeholders to consider in this analysis:**

Urska Srsen: Chief Creative Officer and cofounder of Bellabeat

Sando Mur: Mathematician and co-founder of BellaBeat

**Other individuals essential to this analysis:**

Bellabeat marketing analytics team

**Bellabeat product line:**

Bellabeat app-provides users with data related to their health such as daily activity, sleep, stress management

Leaf-Wellness tracker that can be worn as a bracelet, necklace or clip that connects to the BellaBeat app to track health and wellness

Time-Watch with classical timepiece look in addition to smart functionality to connect to the BellaBeat app and track health and wellness

Spring-Water bottle that tracks daily hydration levels by connecting to the BellaBeat app

**ASK**

**Problem Statement:**

The underlying problem we are trying to solve is how Bellabeat can enhance the health and wellness features on their smart devices to enhance user experience among customers and drive sales growth opportunities for the corporation. This analysis will provide high-level recommendations for BellaBeat to incorporate into their marketing strategies.

**Solution (How my insights drive business decisions):**

The solution will be to implement upgrade features to current line of smart devices or create a new smart device that will maximize features to drive customer satisfaction and sales growth.

**Key questions:**

The key questions I asked to be able to initiate and guide my data analysis were:

* What are some trends in smart device usage?

Data from FitBit users will be analyzed to identify current trends in smart device usage.

* How could these trends appeal to Bellabeat customers?

Data will be analyzed to determine which features appealed most to users in order to determine how those features could be capitalized on in order to drive growth.

* How could these trends help influence Bellabeat marketing strategy?

Data will be analyzed to determine how the identified trends can help influence BellaBeat marketing strategy.

**PREPARE**

**Data Sources:**

This data was retrieved from the Kaggle website. The FitBit Fitness Tracker Data dataset was made available through Mobius. Thirty Fitbit users participated in a survey via Amazon Mechanical Turk during a time period from 03/12/2016-05/12/2016, consenting to their personal health and wellness information being tracked. Upon downloading and cleaning the data, it was noted that there were actually 33 users in total, hence it is believed these additional 3 users also consented to participation in the survey, therefore the information was included in this analysis.

**Data Format & Storage:**

The data was formatted in long format in CSV files. They were transferred to Google Sheets and most initial data cleaning was done in Google Sheets. Files were unable to be uploaded from Google Sheets to BigQuery, so were transferred to Excel in CSV format then uploaded to BigQuery for querying in SQL.

**Files:**

There was initially a total of 18 CSV files, however for the purposes of this analysis, only three were used as the others did not have any impact on the analysis being conducted. The files that were saved were:

* dailyActivity\_merged.csv
* sleepDay\_merged.csv
* weightLogInfo\_merged.csv

The three files were renamed as:

* DailyActivity.csv
* TimeAsleep.csv
* WeightLog.csv

**Data Limitations:**

Limitations to this dataset are as follows:

* Sample size is only 33 users, not representative of a greater population.
* Data is six years old, however the trends remain fairly relevant in 2022.
* Data does not identify gender of survey respondents. This could pose an issue as the BellaBeat products are designed only for women at this time.
* All users did not consistently track sleep-Only 24 users tracked their sleep.
* Most users did not track weight. Only 8 users tracked weight and of those 8, only 2 did so with any relevant degree of consistency. This was therefore later eliminated from the final analysis.
* Data is from a survey through Amazon Mechanical Turk; therefore, reliability is low due to this being third-party data.
* Data was relatively comprehensive, and data was well cited, so reliability here is better than in the above-mentioned areas.
* No additional datasets were able to be located online to enable a more comprehensive analysis.

**PROCESS**

**Data Cleaning and Wrangling:**

Google Sheets was primarily used for most of the data cleaning and wrangling. The following steps were taken to accomplish this:

1. Uploaded files into Google Sheets.
2. Reformatted date/times to proper formatting conventions of YYYY-MM-DD.
3. Applied conditional formatting tool to all three tables to discover any columns with no information. On the table named weightLogInfo\_merged, several cells in the column titled “fat” were missing any value. Since this variable was not necessary to conduct the analysis, this column was removed.
4. Removed column named “WeightKg” from the table named “WeightLogInfo” as it was not necessary to have weight measured in both kilograms and pounds.
5. Removed column named “LogId” from table named “WeightLogInfo” as it was not relevant to the analysis.
6. Removed column named “TotalSleepRecords” from table named “SleepDay” as it did not create any significant impact upon this data analysis project.
7. Changed sheet named “SleepDay” to “TimeAsleep” to create more relevancy pertaining to the data contained within that table.
8. Removed column named “Total Time in Bed” as it did not create any significant impact on the data analysis project.
9. On the sheet/table named “WeightLogInfo”, selected A1:E68 range, then sorted by ID to keep all weight logs for each logged user together for easier interpretation of data.
10. Used COUNTUNIQUE function in Google Sheets to determine the number of distinct users that tracked their weight and BMI with their smart device. =COUNTUNIQUE(A2:A68). Only eight users did so and of those eight, only 2 users tracked with any consistency.
11. Ran AVERAGE function in Google Sheets for average weight (158.8) and average BMI (25.1) of users. However, only eight users logged weight and BMI and of those eight, only two of the users consistently tracked this information. Therefore, there is not enough data in this table to perform any valuable analysis.
12. Used COUNTUNIQUE function in Google Sheets to determine the number of distinct users that tracked their sleep with their smart device. =COUNTUNIQUE(A2:A414). Only 24 users tracked their sleep with their smart device.
13. Used AVERAGE function in Google Sheets to determine average number of minutes asleep. Average number of minutes asleep tracked by users was 419.4673123.

Table

Description automatically generated

1. Divided minutes asleep by minutes in an hour to obtain average hours asleep. Average hours of sleep tracked by users was 6.991121872. This was rounded to an average of 7 hours of sleep per day tracked by users.
2. Used COUNTUNIQUE function in Google Sheets to determine number of distinct users that tracked their total steps with their smart device. =COUNTUNIQUE(A2:A941). There were 33 users that tracked their total steps with their smart device.
3. Used AVERAGE function in Google Sheets to determine average total daily steps.

Table

Description automatically generated

1. Several users missed days of tracking steps and some missed days of tracking calories. In those cases, their average daily steps and daily average calories was documented in place of the missing values.
2. Removed columns D-N (Distance, Logged activities, levels of activities) as these columns were not relevant to this analysis.
3. Removed three duplicate rows in the table named “TimeAsleep”. No duplicates were identified in the remaining tables used.
4. Could not upload the Google Sheets tables to BigQuery without a paid BigQuery account. Worked around this obstacle by saving the Google Sheets as CSV files in Excel then uploading those to the BigQuery platform.

**ANALYZE**

1. Ran SQL query to double check for distinct number of users. Total users are verified again as 33.

Graphical user interface, text, application, email

Description automatically generated

1. Ran SQL query on table named TimeAsleep to double check for distinct number of users. Total users are verified again as 24.

Graphical user interface, text, application, email

Description automatically generated

1. Ran SQL query on table named WeightLog for distinct number of users. Total users verified again as 8.

Graphical user interface, text, application, email

Description automatically generated

1. Double checked that each table had ID numbers with exactly 10 characters.

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

1. Performed an inner join to join the three tables together in one view.

Graphical user interface, text, application

Description automatically generated

**SHARE**

I shared my graphic visualizations on Tableau Public (See link below) and also took some screenshots of the visualizations to share in this summary (See below).

<https://public.tableau.com/app/profile/jennifer8101>

There seemed to be no correlation between total daily calories and total daily steps.

Chart, scatter chart

Description automatically generated

There seemed to be no correlation between average minutes asleep and average total steps. What factors in as something that skews these results is the lack of participation of FitBit users in consistently recording sleep time.

A picture containing table

Description automatically generated

There seemed to be no correlation average calories, average minutes asleep and average steps.

Chart, scatter chart

Description automatically generated

**ACT**

**Recommendations:**

This analysis was very challenging wholly in part due to its limitations of sample size and consistency of contributions by users. Therefore, my first recommendation would be to conduct further analysis for more definitive directions in how to capitalize on marketing strategies for BellaBeat. However, based on this limited analysis, I can make the following conclusions and accompanying recommendations:

*Observation 1:*

Users seemed most interested in the tracking of activity levels. They did not maintain consistency with recording sleep, weight, and daily caloric intake. This is indicative of busy lifestyles and no time to fully engage in all the features of the app.

*Recommendation 1:*

To get users more engaged, capitalize on the exercise components of this product. Consumers want to be active, so get them more engaged by incorporating exercise reminder notifications, goal achievement notifications and awards, daily challenges and competitions with other users. This will get BellaBeat users interacting more with the app and in turn, opening them up to truly visualizing and utilizing all of the features of the app.

*Observation 2:*

Users kept track of daily caloric intake more than they did weight loss.

*Recommendation 2:*

To get users more engaged with tracking daily calories, incorporate recipe ideas in the app. Allow users to engage with the app to share recipes with others and review recipes.

*Observation 3:*

In today’s world everyone is in a hurry and doesn’t necessarily take the time to engage with the app. They do take the time to engage with friends, family and the world through social media.

*Recommendation 3:*

Capitalize on the power of social media to reach your consumer base. Get BellaBeat on Instagram, Facebook, Twitter, TikTok. Promote the app, daily challenges, recipes, and user testimonials. Allow BellaBeat users to engage with one another. Hire a social media marketing manager to create daily Instagram Reels and TikTok videos sharing recipes, showing weight loss journeys of actual BellaBeat consumers, highlighting product and app features.