Bellabeat Case Study

MERIT AMERICA

GOOGLE DATA ANALYTICS CAPSTONE PROJECT

BY: BIANNETT CHEVALIER 10/7/2023

BIANNETTCHEVALIER@GMAIL.COM

LINKEDIN

PORTFOLIO

Ask: About Bellabeat, it's stakeholders and the business task

Founded in 2013, Bellabeat is a high-tech company that manufactures health-focused smart products.

Bellabeat focuses on collecting data on activity, sleep, stress, and reproductive health through different products, ranging from an app on a smartphone to a water bottle that collects data.

Bellabeat is by women, for women. The company is cofounded by Urška Sršen and Sando Mur, two women wanting to empower other women with knowledge about their own health and habits.

During this case study, we will determine trends in smart device usage and apply insights to Bellebeat products in order to reveal more opportunities for growth.

Questions we'll be answering

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

Prepare: Upload and inspect data

- To start, I inspected and downloaded the Fitbit dataset from Kaggle to use in my analysis.
- Of the 14 .csv files included, I moved the dailyActivity_merged, sleepDay_merged and weightLogInfo_merged into a separate folder named “Files used” which would make it quicker and easier for me to import into SQL later.
- I decided to focus on these 3 tables because the other tables had similar information and I felt these would be most insightful. I decided to work with these in SQL because of how large the data sets were, it would not be efficient to work with in Excel.

Data Limitations

- There were some limitations with this data, the biggest ones being size and when the data was collected.
- There were around 30 participants and not all of them tracked the different types of activity, like daily activity, sleep and weight. This made the insights a bit limited.
- The data was also collected in 2015, making it 8 years old and outdated.
- The data is also missing demographics, meaning we have no way of knowing if this data will give us biased results.

Process

In SQL, I started by making a new database named “Bellabeat”. I began importing the excel sheets as flat file tables. Of the 18 provided tables, I imported the ones named:

- dailyActivity_merged,
- sleepDay_merged
- weightLogInfo_merged

I renamed all tables to:

- daily_activity
- sleep_day
- weight_log

This allowed me to write queries into SQL quicker and more readable.

Lastly, as I imported the flat files I made sure to change the data types of each column so they were the same across each table. This included changing “floats” into integers so I could later do calculations.

Analyze and Share:
Daily Activity

To start, wanted to see how often users wore their trackers and then I categorized them into 3 different categories: Active, Moderate, and Light users.

```
-- How many times users wore their trackers and categorize users by type

SELECT
  Id,
  COUNT(Id) AS Total_Uses,
CASE
  WHEN COUNT (Id) BETWEEN 20 AND 31 THEN 'Active'
  WHEN COUNT (Id) BETWEEN 10 AND 19 THEN 'Moderate'
  WHEN COUNT (Id) BETWEEN 0 AND 9 THEN 'Light'
END as Activity
FROM
  bellabeat.dbo.daily_activity
GROUP BY
  Id
ORDER BY
  Total_Uses DESC;
```

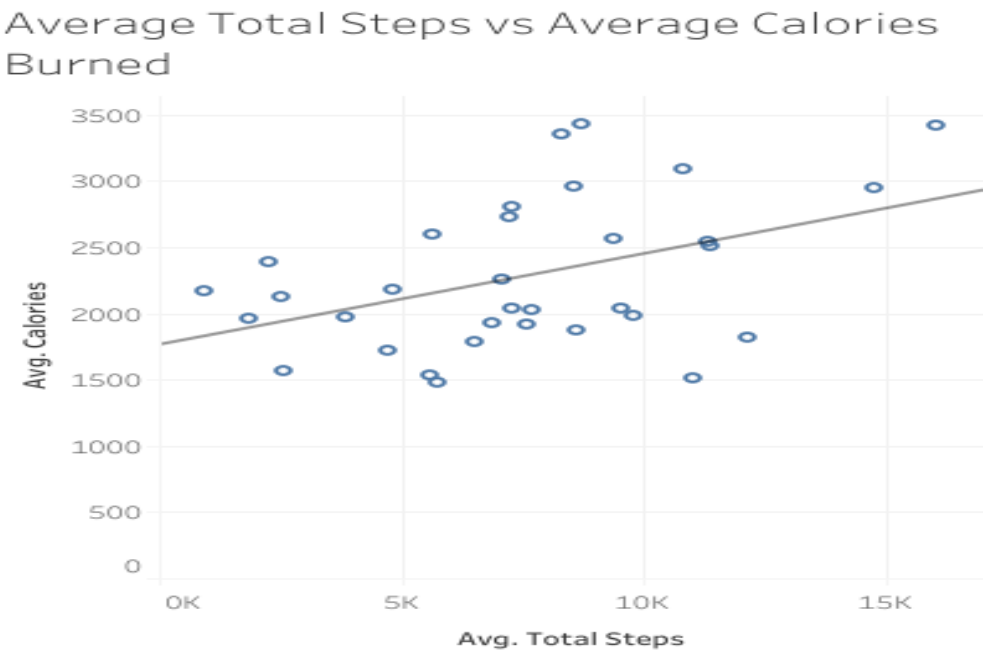
Results:

1. There were 33 users tracking their activity
2. The data was collected for 31 days
3. Of those 33 users, 30 were considered “active” users, tracking their activity 20+ days
4. 2 users were considered “moderate” users, tracking their activity 18-19 out of 31 days
5. Only 1 user was considered a “light” user, only tracking 4 out of 31 days

Daily Activity cont.

Next I wanted to see the average total steps each user took daily and I wanted to compare that to the average calories burned daily.
After running this query, I noticed a general trend that higher step count and higher distance usually meant higher calories burned daily which was confirmed with the trend line in the scatter point graph.

```
-- Avg total steps, distance and calories burned by user grouped by id
SELECT
    Id,
    AVG(TotalSteps) AS avg_total_steps,
    AVG(TotalDistance) AS avg_total_distance,
    AVG(Calories) AS avg_calories_burned
FROM
    bellabeat.dbo.daily_activity
GROUP BY
    Id
ORDER BY
    avg_total_steps desc;
```



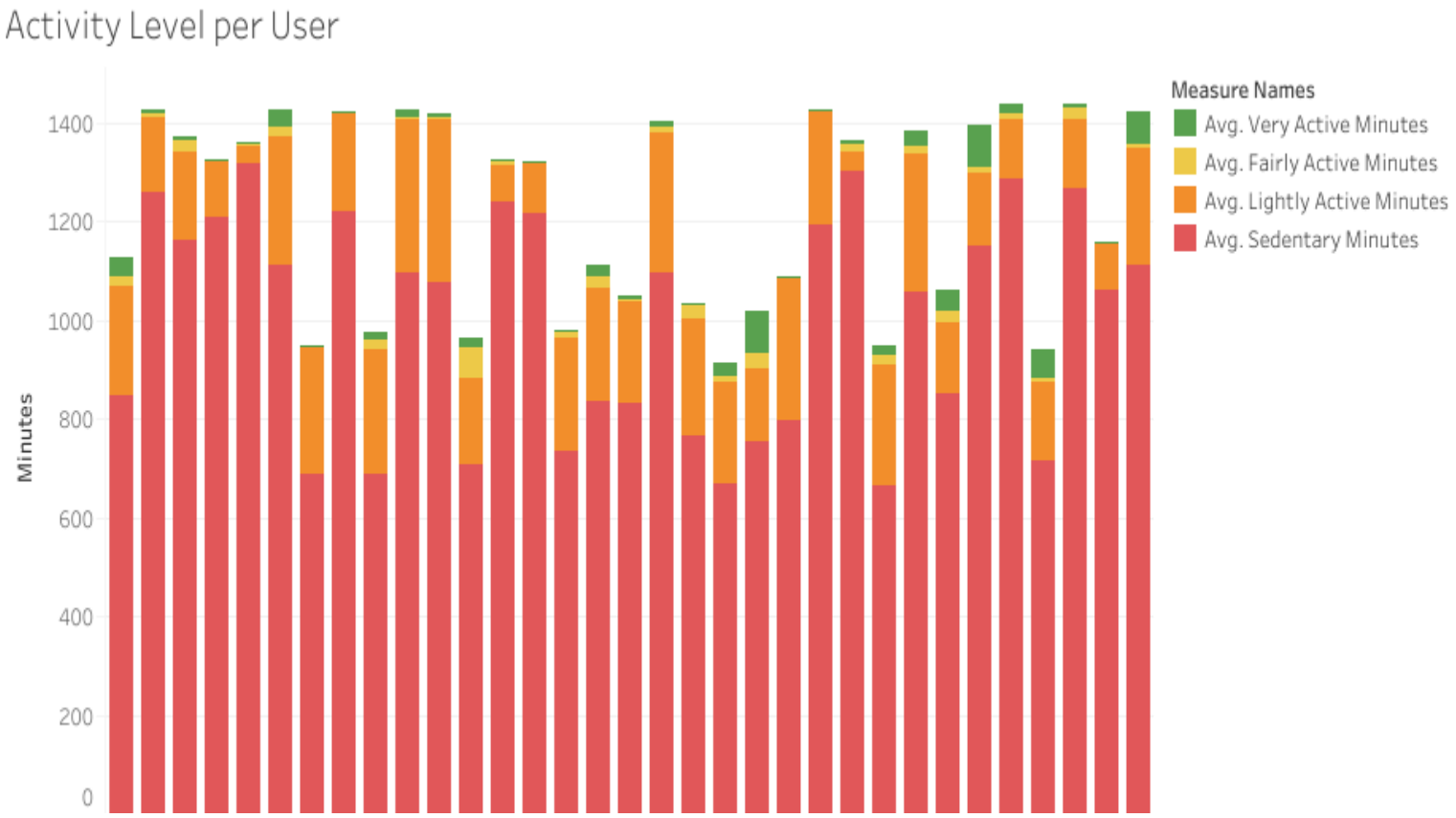
Daily Activity cont.

I wanted to see how long each user was in each activity type on average, so I ran a query which allowed me to notice that every single user is in Sedentary status significantly longer than any other activity type.

```
-- Avg time spent in each activity level
SELECT
     Id,
     AVG(VeryActiveMinutes) AS very_active_minutes,
     AVG(FairlyActiveMinutes) AS fairly_active_minutes,
     AVG(LightlyActiveMinutes) AS lightly_active_minutes,
     AVG(SedentaryMinutes) AS sedentary_minutes
FROM
     bellabeat.dbo.daily_activity
GROUP BY
     Id
ORDER BY
     Id
```

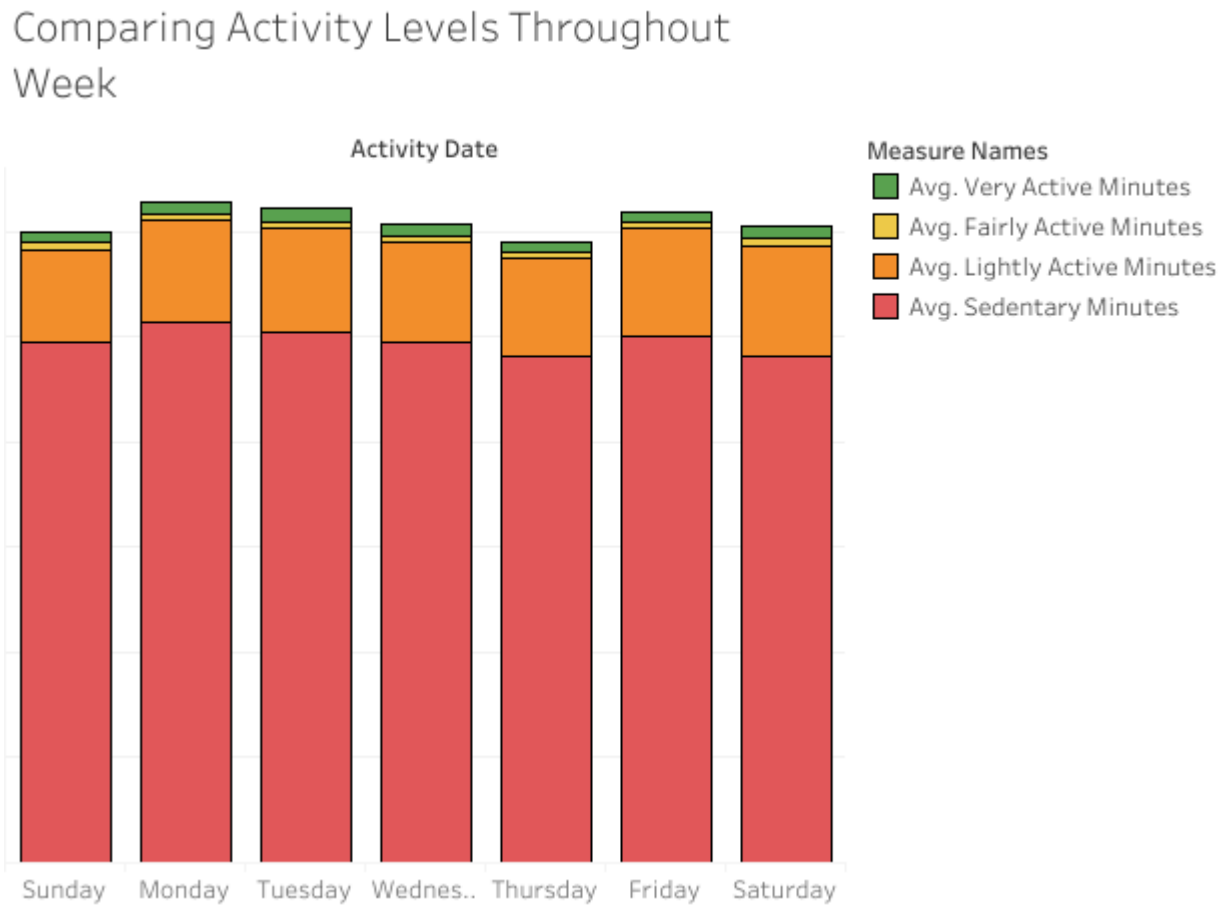
Activity Level Bar Graph

The following graph shows how long each user was in each activity type on average. This confirms my insight that people were in sedentary status significantly longer than any other type.



Activity levels throughout the week

After the results of the last query, I wanted to see if there was any relationship between the users activity types on different days of the week. This graph shows that there is no direct correlation between days of the week and how active users are.



Sleep and Weight

Next, I wanted to see how many users tracked their sleep and weight and what insights I could get from those.

```sql
-- Users tracking sleep
SELECT
COUNT(DISTINCT Id) as num_of_sleep_users
FROM
    bellabeat.dbo.sleep_day

-- Avg sleep hours
SELECT
    Id,
    AVG(TotalMinutesAsleep)/60.00 AS average_hours_asleep
FROM
    bellabeat.dbo.sleep_day
GROUP BY
    Id
ORDER BY
    average_hours_asleep DESC;
```

From these queries, I found that 24 users tracked their sleep and of those 24, about half of them got the recommended 7-8 hours of sleep a night.

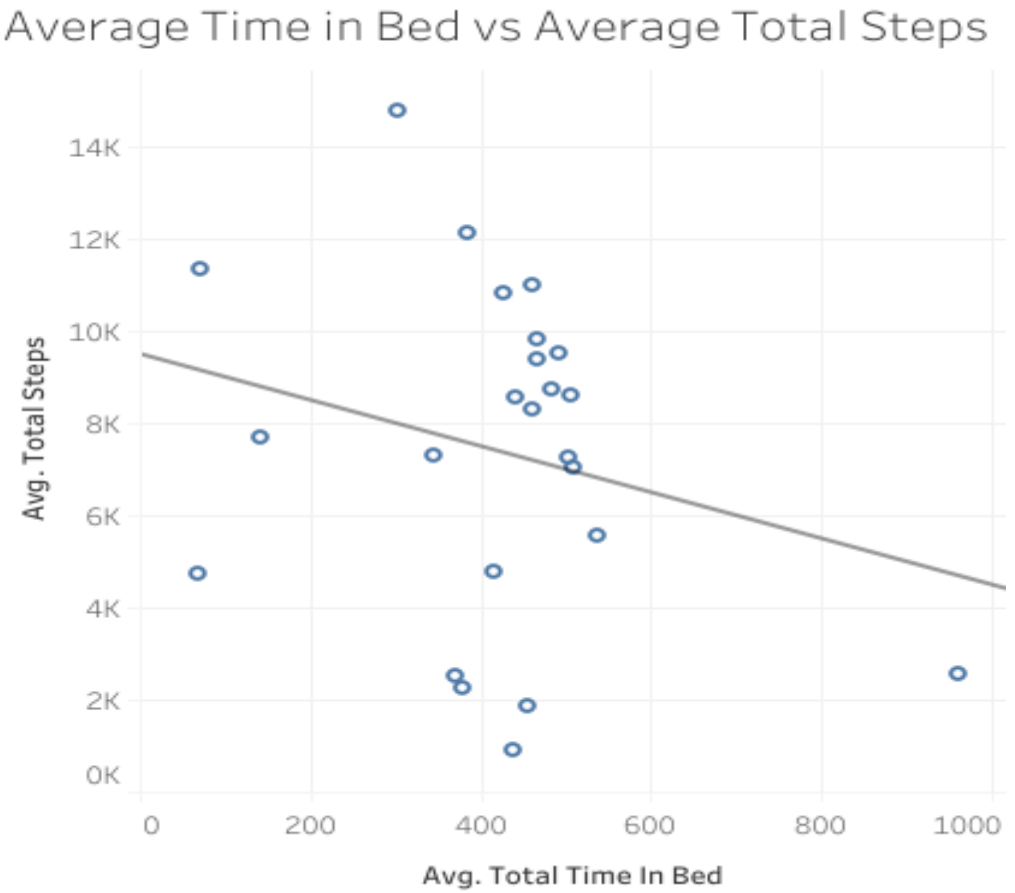
```sql
-- Users tracking weight
SELECT
 COUNT(DISTINCT Id) AS num_of_weight_users
FROM
    bellabeat.dbo.weight_log

-- Avg weight and BMI grouped by ID
SELECT
    Id,
    AVG(WeightPounds) AS avg_weight_in_pounds,
    AVG(BMI) AS avg_bmi
FROM
    bellabeat.dbo.weight_log
GROUP BY
    Id
```

From these, I found out that only 8 users tracked their weight and of the 8, only 3 were considered to have "standard" BMI, while the other 5 were considered to have a high BMI.

Sleep and weight cont.

Lastly, I wanted to see if there was any correlation between how long each user spent in bed and how many steps they took daily on average.

The downward trend on this shows that the longer people spend in bed, the less steps they take on average which means they burn less calories on average as well.



Act:

Trends and insights:

- *Most users choose to track their activity and sleep, and skip tracking their weight.*
- *Users who track their activity spend more time sedentary than they do active.*
- *Users who have a higher step count burn more calories.*
- *Half the users who track their sleep get the recommended 7-8 hours of sleep a night. The other half does not get the recommended amount of sleep.*
- *The more time users spent in bed, the less active they were.*
- *Less than 25% of users tracked their weight, and 62.5% of those users are considered to have a high BMI.*

Recommendations

Bellabeat can use these trends to market their under-utilized features more. For example, only a quarter of the users are using the weight log feature. If Bellabeat can find a way to incentivize logging your weight and hitting goals, users will be more likely to use this feature.

They can set themselves apart from their competitors by promoting a feature that sends you reminders to be more active when they've been sedentary for a certain amount of time, or even develop a feature that helps users get the perfect amount of sleep nightly.

If there's a way Bellabeat can get even more in depth in the sleep logs, users can improve their sleeping habits and quality of sleep, and in turn they will be more active and get better use of the products and features.