 BELLABEAT

# **INTRODUCTION**

BELLABEAT is a high-tech company that manufactures health-focused smart products for women. Their products include trackers in a form of necklaces and smart watches, clips, apps and water bottles. They also have different memberships for their users. Stakeholders asked for recommendations from insights gained when analyzing smart usage data based on the following questions:

1. Urska Srsen – Bellabeat’s co-founder and Chief Creative Officer.
2. Sando Mur – Mathematician and Bellabeat’s co-founder.
3. Bellabeat’s marketing analytics team – Data Analysts.

# **Ask phase**

**Business Task:**

Analyze smart device usage data to gain insight into how customers use non-Bellabeat smart devices.

**Stakeholders:**

* Urska Srsen – Bellabeat’s co-founder and Chief Creative Officer.
* Sando Mur – Mathematician and Bellabeat’s co-founder.
* Bellabeat’s marketing analytics team – Data Analysts.

# **Prepare phase**

**Datasets:**

* Datasets used is Fitbit Fitness Tracker Data
* Found in Kaggle at CC0: Public Domain, available through Mobius.
* 30 eligible Fitbit users gave consent to the submission of personal tracker data, including minute-level output for physical activity, heart rate and sleep monitoring.
* The data include information about daily activity, steps, and heart rate that can be used to explore user’s habits.

After downloading the datasets, I prepared my data by cleaning it in uploading it in Google Sheets. I changed the format of data to 2 decimals, checked for spelling errors. I removed empty columns and values. I sorted the datasets to display data from the highest calories lost.

**Limitations:**

* Srsen, one of the Stakeholder mentioned that this data might have some limitations and encouraged to consider adding more data to address the limitations.
* Data recorded data is between April – May 2016, meaning it is outdated. User behavior may have changed.
* The data is from 30 users, however, Fitbit has approximately 30 million users. Using a population size of 30 million users and only using 30 sample size gives out a Margin of Error = 17.9% which concludes that the given sample size is little to make accurate predictions.
* I used Google Big Query to count the number of users using SQL and found 33 users instead of 30.

-- Checking the number of number of users

SELECT

DISTINCT Id

FROM `bellabeat-392006.Bellabeat.dailyActivity`

* The data was recorded from 2016-04-12 to 2016-05-12, 30 days of data collection.

-- Checking the time frame of the recording of data

SELECT

DISTINCT Id, ActivityDate

FROM `bellabeat-392006.Bellabeat.dailyActivity`

ORDER BY ActivityDate

# **Process phase**

For data cleaning, Google sheets and SQL were used to ensure the integrity and cleanliness of the data.

* Checked Id user length consistency.
* Removed empty columns.
* Rounded the values to 2 decimal places.
* Checked for duplicates.
* Extra space within cells.
* Checked for date formats.
* Checked for nulls and empty spaces.

The following data cleaning was used.

File: ***dailyActivity\_Merged***

* Checked Id length consistency.
* Removed empty columns and rows.
* Rounded all numeric values to 2 decimal places.
* Removed the VeryActiveDistance, ModeratelyActiveDistance and LightActiveDistance columns.
* Focused on the Id, ActivityDate, TotalSteps, TrackerDistance,SedentaryMinutes and Calories columns.
* Renamed SedentaryMinutes to IdleMinutes.

File: ***weightLogInfo\_Merged***

* Used the LEFT & RIGHT Function to separate the Date and Time.
* Rounded the WeightKg numerical values to 2 decimal places.

File: ***sleepDay\_Merged***

* Separated the Date and Time Function.
* The data was clean.

# **Analyze Phase**

Checked the number of Id Users in

dailyActvity\_Merged data

-- Cheking the total number of users

SELECT COUNT(DISTINCT Id) AS Total\_Ids

FROM `bellabeat-392006.Bellabeat.dailyActivity`

sleepDay\_Merged data

SELECT

DISTINCT Id

FROM `bellabeat-392006.Bellabeat.sleepDay`

weightInfo\_Merged data

SELECT

DISTINCT Id

FROM `bellabeat-392006.Bellabeat.weightInfo`

heartrate\_Merged data

SELECT

DISTINCT Id

FROM `bellabeat-392006.Bellabeat.heartrate`

dailyStep\_Merged data

SELECT

DISTINCT Id

FROM `bellabeat-392006.Bellabeat.dailySteps`

From the SQL query, the number of users were the following:

dailyActivity = 33

sleepDay = 24

weightInfo = 8

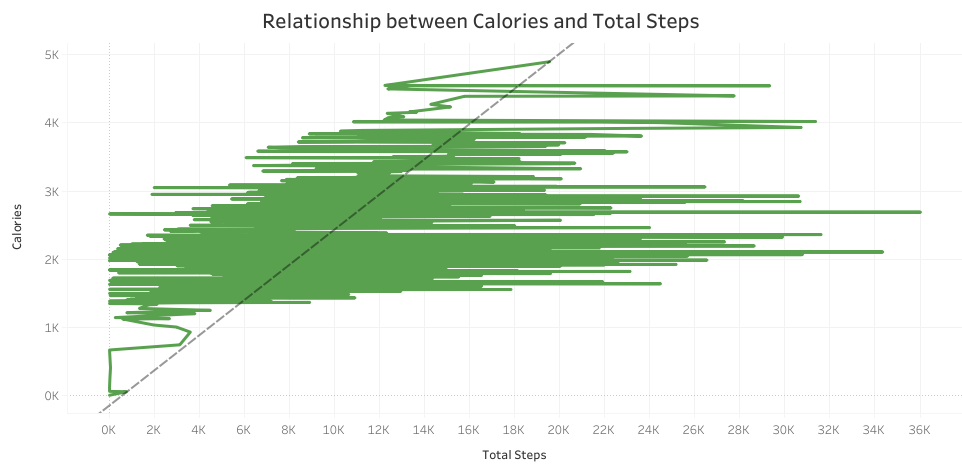
heart rate = 14

dailySteps = 33

* Most users used the dailyActivity and dailySteps features. Therefore, they will be used to further analyze the data.
* The weightInfo and heartrate feature were used less therefore the data won’t give out accurate predictions.

**Analyze the relationship between Total steps and Calories.**

I used Tableau to analyze the relationship between total steps and calories.



*Figure 1: Analyze the relationship between Total steps and Calories.*

Results show that there is a relation between Calories and Total Steps. The more steps the user takes, the more calories the user burns.

* According to this data, users track their steps to burn more calories.

**Examine the days the Total Steps and Calories features are most active/used.**

With the main objective to determine which products Bellabeat beat uses the most and why it is so, I examined the most used product features, which are the daily activity and daily steps product feature. According to the days users burn more calories, I used Google sheets to rename the ActivityDate column into weekday and then proceeded to SQL to find the average number of steps and Tracker distance.

-- The weekdays the users are most active

SELECT

DISTINCT Weekday,

AVG(TotalSteps) AS AVG\_TotalSteps,

AVG(TrackerDistance) AS AVG\_TrackerDistance,

AVG(Calories) AS AVG\_Calories

FROM `bellabeat-392006.Bellabeat.dailyActivity`

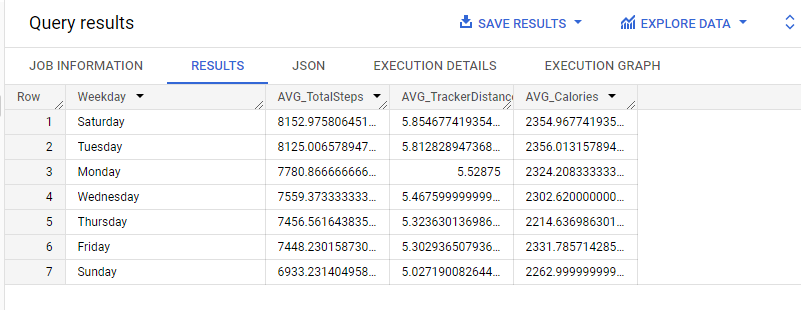
GROUP BY Weekday

ORDER BY AVG\_TotalSteps DESC

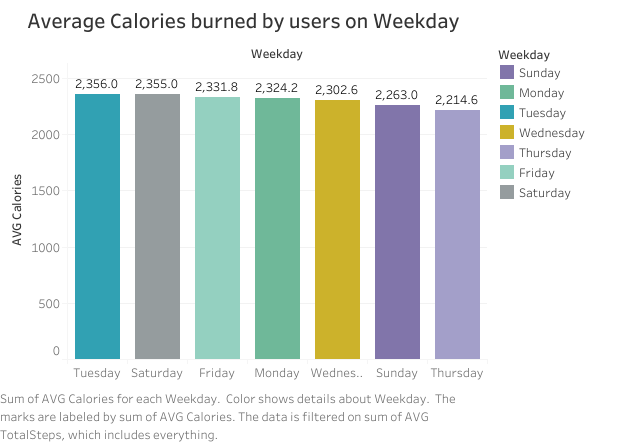
The results show that users are more active on Saturday and less active on Sunday. The relationship between Total steps and Calories is also displayed in this data, although Tuesday is the second highest day in which users are active, it is not the day in which users most active.

* This tells us that the number of calories lost is not merely determined by the Total Steps.

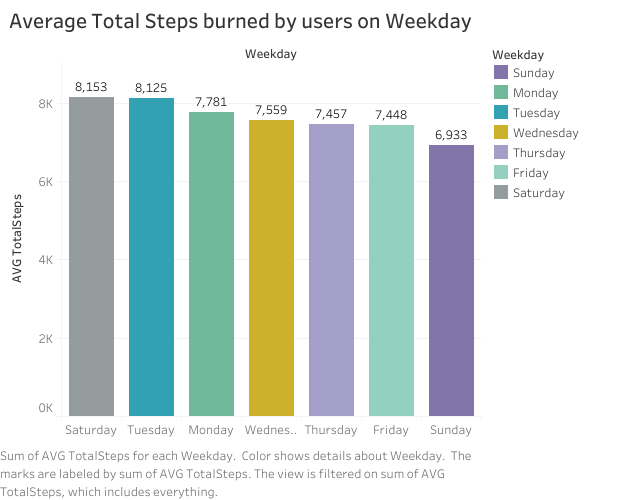
This data shows that Saturday is the day that Bellabeat users were most active.



I used Tableau to visualize the data:



*Figure 2: Average Calories burned by users on Weekdays.*



*Figure 3: Average Total Steps burned by users on Weekdays.*

* Users lost most calories on Tuesday, but users were most active on Saturday. Such inconsistency may be affected by the small sample size.

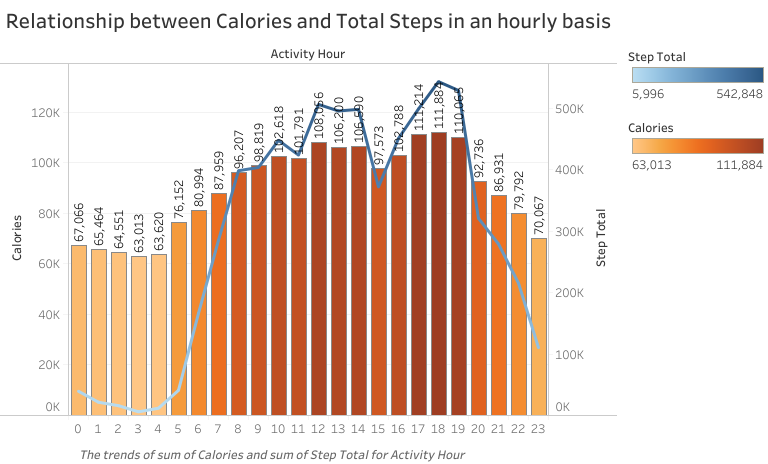
I separated the date and time in the hourlyCalories\_merged file using Google Sheets and then combined it with the hourlySteps\_merged file. I changed the name of the file to “effectiveRelation\_merged” file. This will help determine the time users are more active, and the time they lose more calories. To determine the relationship between Time, Calories burnt and Total Steps, I uploaded the effectiveRelation\_merged file to Big Query ad used SQL to further analyze data.

SELECT \*

FROM `bellabeat-392006.Bellabeat.effectiveRelation`

ORDER BY StepTotal DESC

I used Tableau to visualize the relationship between Activity Day and time, Calories, and total steps.



The most Calories burned are highlighted by the dark orange color. The graph above shows that users were most active at 17, 18 and 19 hours of the day, which the also the exact time where most users lost more Calories.

# **Share Phase**

**CONCLUSION**

It can be clearly seen that active users get more positive results than idle users. This analysis is shown in the relationship between calories and total steps in an hourly basis graph. Users burned twice the calories in the hour they were most active compared to the hours they were least active. On a weekday, users were most active on Saturday and Tuesday and lost most calories on Tuesday and Saturday, respectively.

There is a clear trend between burning calories and taking more steps daily.

# **Act Phase**

**Recommendations to Business**

1. More users prefer to use daily activity features than night features like sleep and heart rate features.
2. Only 8 out of 33 users used the weight info feature, but used the calories feature and steps tracking features.

* Which can mean that users prefer to focus on the process than recording their outcome.

1. Bellabeat can advertise more features that focus on the process, their users want to be encouraged and reminded daily about being active.
2. Bellabeat can insert a function in the App and in the Leaf/Time to notify users to have a certain number of steps in an hourly basis that will estimate the number of Calories they can use in a daily basis.

* This can also have users set their daily and weekly goals. This function can remind then hourly to move around regularly.

1. Bellabeat can also target working females and have Marketing campaigns on how Bellabeat products are a fit to use for after hours.

* Bellabeat is an excellent product to be used on weekdays and during weekends.

1. Lastly, Bellabeat has proven to work and has grown over the years. The service the products offer can also be utilizes by men and therefore, it will be an exciting time to target males. Bellabeat products can be advertised as unisexual because good health does not only specify to 1 gender but to all human beings.
2. Total steps confirmed that movement result in burning more calories, so the Bellabeat app, Leaf tracker, the watch and Spring bottle can be used in the gym. And from the analysis most users are active between 17 – 19 hours of the day, a time where people are going home, and others are going to the gym.

* Bellabeat membership can also include a partnership with a gym membership.

1. Bellabeat membership can be sold with a gym membership including all 4 Bellabeat products (Bellabeat app, Leaf tracker, the watch and Spring bottle). Therefore, all products are being sold and according to the insights, users will benefit greatly because most of their users are more active on Saturday and Tuesday, between 17 – 19 hours a day.
2. And users use Bellabeat products that focus more on the process than on the results, this can be proved because only 8 users used the weight info and sleep day feature was used by 14.

# **Marketing Recommendations**

Bellabeat users are more focused on putting in the work, tracking their activity which is great to introduce gym memberships including all 4 Bellabeat products (Bellabeat app, Leaf tracker, the watch and Spring bottle) as a selling package. The second-highest number of active users was between 12 and 14 hours, which is normally lunch time on a weekday. Targeting 9-5 professional working women will also be a great business move and using a marketing strategy to highlight the success of 17-19 hours users. Bellabeat can also target Males since keeping healthy does not only apply to the gender of women only.

--------------------------------------------------------------------------------------------------------------------------