# Bellabeat Report

## Understand customers

Users of the membership app seem to be mostly female, not important. We are trying to understand user of a competitor, Fitbit. Both are popular with those trying to maintain or get into good shape it seems

## Ask

**Business task(s):**

* Understand users of Fitbit to understand overall trends in use of smart fitness devices
* Apply these insights to our products

**Stakeholders:**

* **Urška Sršen – CCO of Bellabeat and the one**
* **Sando Mur- mathematician who may give insight into new models to use**
* **Analyst Team- will (if this was real) work with us to gather and validate data**

## Prepare

* Where is your data stored?
  + Downloads- Fitabase Data 4.12.16-5.12.16
* How is the data organized? Is it in long or wide format?
  + Long format
* Are there issues with bias or credibility in this data? Does your data ROCCC?
  + It is reliable and I can’t find bias
* How are you addressing licensing, privacy, security, and accessibility?
  + I will acknowledge where I got the
* How did you verify the data’s integrity?
  + Yes, there is some missing data in weightLogInfo and zeroed data, otherwise the data is in tact
* How does it help you answer your question?
  + We can use data from sleepDay, weightLogInfo, and dailyActivity (.csv) to analyze the average daily cycle of the users and how and when they are using the device.
* Are there any problems with the data?
  + Missing fat data in weightLogInfo and some zeroed data

## Process

1. I remove any data without activity from dailyActivity

|  |  |
| --- | --- |
| 8253242879 | 4/30/16 |
| 6290855005 | 5/10/16 |
| 1503960366 | 5/12/16 |
| 8583815059 | 5/12/16 |

1. Upload and read the 3 data sets into R
2. Install packages here, skimr, and janitor
3. Ensured dplyr was loaded
4. Called skim\_without\_charts on all
   1. Fat in weightLogInfo had a 2% completion, so I dropped it using subset
5. Renamed any column referring to date to “Date”
6. Remove time components of date
7. Merge into one table

## Analyze

Mono-variable

A blue square with a line and a black text

Description automatically generated with medium confidence

The average app user takes about 7,500 steps a day but about a quart don’t log any steps and a few will log about 20-30,000

A graph of a number of objects

Description automatically generated with medium confidence

Average of about 5km but goes up to 15-25+ km.

A graph with numbers and dots

Description automatically generated

Even at the third quartile, we see no activity tracking involving step and even at the highest amount, we only see 5 km

Summary of previous graphs

While most users are moving a fair amount, a good number log no movement and very few use the device to actively track their distance during activities. This may suggest the device is used for more weight based activities or that the setting to track is difficult to find or otherwise rarely used.

Let’s look at light, moderate, and heavy activity distance and times

A graph of a graph with a number of rows

Description automatically generated with medium confidence

Most activities are “light” with “very active” being second. Moderate is rare

A graph with a number of different times

Description automatically generated with medium confidence

About the same as above

Summary from previous two graphs

Most users spend most of their active time in light activity, this likely indicated they spend most of their time just walking and doing other similar activities. However, heavy exercise seems more prevalent than normal. This would suggest most users take their exercise seriously when active

Now let us look at calories and BMI, I will use BMI instead of weight because we can’t know heightsA graph with a green square and black lines

Description automatically generated

Average calory intake is pretty normal, though we have some people eating much more and much less, this may be part of their fitness regiments as some athlete eat 4,000+ calories and some weight loss regiments have limits of 300-600 calories

A diagram of a number of users

Description automatically generated

Average BMI is about 23 and range from about 21 to 24, this is normal and really good to see. There are some above this, it seems only 2 are overweight or obese though

Summary of previous 2 graphs

We can see most users are already in a good weight range, this may suggest weight loss isn’t a big deal. Calory intake was relatively diverse but for the most part in the healthy range for an average adult. Some people need more or less calories so the 4,000 and even 300 may be normal for those people given their activity levels and weight goals

A graph with a red square and a white line

Description automatically generated

This is a very diverse data set, we see the average is about 430 minutes just over 7 hours (with about 50% ranging from 360-490 minutes), but some people get less than 3 hours and some more than 10.

A graph of a number of minutes in bed

Description automatically generated

Average time in bed is about 460 minutes with most falling between 400 and 520.

Summary of previous two graphs

Most people are getting a relativity good amount of sleep, but some are getting little to no sleep. The averages for time in bed are roughly 30 minutes higher than sleep, suggesting the average person takes about 30 minutes to fall asleep and/or waking up in the morning.

Bi-variate

A graph of a number of dots

Description automatically generated with medium confidence

We can see aa very slight positive correlation between distance and calories, meaning those that are more active eat more, which makes sense. At distance bellow 10 km, this trend is much weaker.

A graph of a number of black dots

Description automatically generated

After about 10 km we can see a good positive correlation, This is logical as these people likely doing some cardio activity.

A graph showing a number of dots

Description automatically generated

We can see that there is a very slight negative correlation between distance moved and time spent not doing activities.

A graph showing a number of dots

Description automatically generated

Amongst those that do high activity, there is a positive correlation between time spent and calories.

Conclusions from previous section

We can see those that move more are more active and eat more on average. This may suggest that these people are dedicated to intense workouts and thus may need more calories.

A graph with numbers and a line

Description automatically generated

Interesting , no one who has BMI also has calories.

The same is true for other categories I looked at, meaning we can’t really get a lot of info form BMI as well as asleep time

Conclusion, people don’t use multiple functionalities (sleep, exercise, diet) too often on the same day)

A graph with black dots

Description automatically generated

No relationship between sleep and bmi

Insights

We can apply to our Bellabeat app by targeting people already into fitness and who tend to walk a lot. We should advertise in running store, sports ware sections, and on podcasts that focus on that. We may not want to focus on weight and diet as those don’t seem popular and focus instead on metrics relevant to cardio, perhaps heart rate? With this data we can stand out by creating recommendations for excercises.