Google Data Analytics Capstone

Case Study 2: How Can a Wellness Technology Company Play It Smart?

Feb 2023

Content

Case study summary

Case study summary

Company:	Bellabeat
Scope:	High-tech manufacturer of health-focused products for women
Products:	**Bellabeat app:** The Bellabeat app provides users with health data related to their activity, sleep, stress, menstrual cycle, and mindfulness habits. **Leaf:** Bellabeat's classic wellness tracker can be worn as a bracelet, necklace, or clip. The Leaf tracker connects to the Bellabeat app to track activity, sleep, and stress. **Time:** This wellness watch combines the timeless look of a classic timepiece with smart technology to track user activity, sleep, and stress. **Spring:** This is a water bottle that tracks daily water intake using smart technology to ensure that you are appropriately hydrated throughout the day.
Task	An analysis of Bellabeat's available consumer data would **reveal more opportunities for growth**. The marketing analytics team to focus on a Bellabeat product and analyze smart device usage data in order **to gain insight into how people are already using their smart devices**. Then, using this information, high-level recommendations to be produced for how these **trends can inform Bellabeat marketing strategy.** Select one Bellabeat product to apply these insights to in presentation.

Summary of business task

Summary of business task
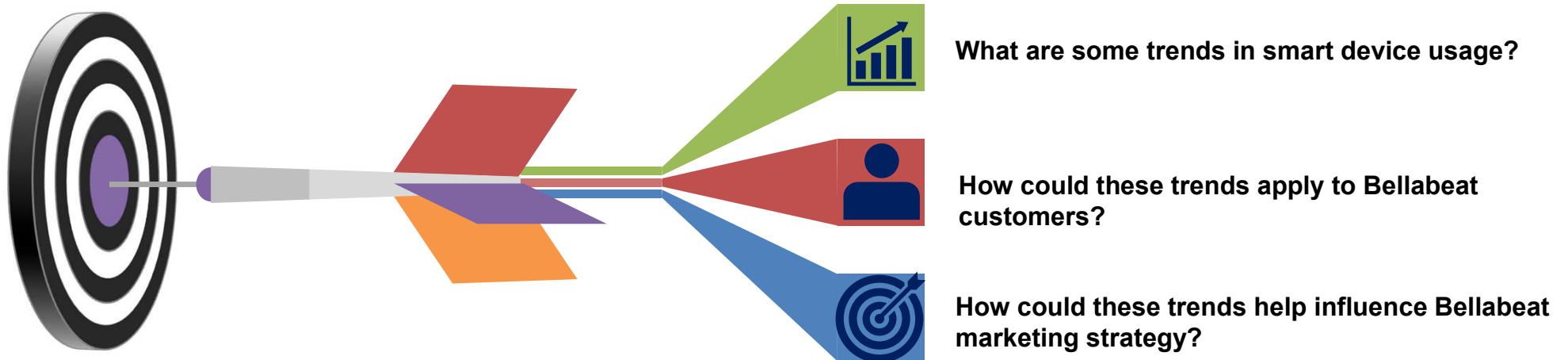
What are some trends in smart device usage?

How could these trends apply to Bellabeat customers?

How could these trends help influence Bellabeat marketing strategy?

Following questions may help find insights:
1.What parts of users' day are occupied by bed time, sedentary or active time
2.What time the majority of users wakeup/go sleep
3.What is the average amount of steps users do per day
4.Is there any day per week with maximum/minimum activity
5.Is there any day per week with maximum/minimum time sleep time
6.Maximum steps, calories per hour
7.Is there any correlation btw calories burnt and sleep time
8.What is the main difference btw people who burn relatively larger calories amount in comparison with those who burn less.

Data sources used

Data sources used description

Data source 01

This dataset generated by respondents to a distributed survey via Amazon Mechanical Turk. Thirty eligible Fitbit users consented to the submission of personal tracker data, including minute-level output for physical activity, heart rate, and sleep monitoring. Individual reports can be parsed by export session ID (column A) or timestamp (column B). Variation between output represents use of different types of Fitbit trackers and individual tracking behaviors / preferences.

https://www.kaggle.com/datasets/arashnic/fitbit

Data source 02

https://www.thinkwithgoogle.com/future-of-marketing/emerging-technology/smart-device-use-statistics/

smart-device-use-statistics

Data source 03

https://www2.deloitte.com/dk/da/pages/technology-media-and-telecommunications/topics/digital-consumer-trends/new-gadgets-early-trends-in-future-smart-life.html

Analysis of health monitoring devices in Scandinavia by Deloitte.

Data source 01: [FitBit Fitness Tracker Data | Kaggle](#)

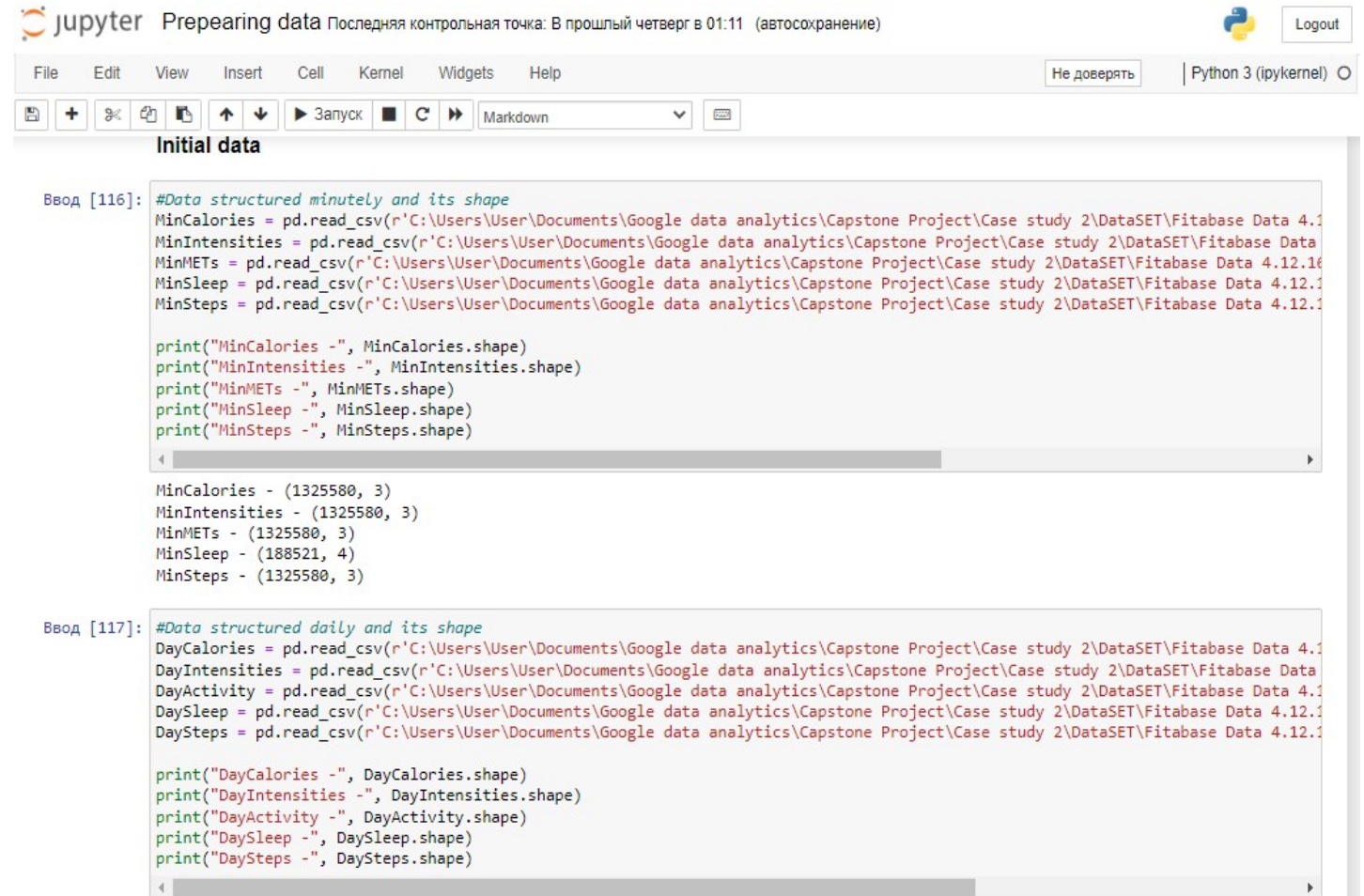**Initially, data is presented by 15 csv files:**

- Structured daily - **"Activity"**, **"Calories"**, **"Intensities"***, **"Sleep time"**, **"Steps"**;
- Structured hourly - **"Calories"**, **"Intensities"**, **"Steps"**;
- Structured minutely - **"Calories"**, **"Intensities"**, **"METs"***, **"Sleep time"**, **"Steps"**;
- Occurring every second - **"Heart Rate"**;
- And also - **"Weight"**.

*Intensity Minutes are earned based on your current heart rate when compared to your average resting heart rate or the number of steps taken per minute. For example, you can earn intensity minutes once a brisk walk or run is detected.

** **MET (metabolic equivalent)** - the amount of oxygen consumed while sitting at rest and is equal to 3.5 ml O2 per kg body weight x min. The MET concept represents a simple, practical, and easily understood procedure for expressing the energy cost of physical activities as a multiple of the resting metabolic rate.

Cleaning and manipulation of data documentation

Data cleaning and manipulation documentation

Data cleaning and manipulation was done using the python pandas library.
Here are screen shots of some lines.
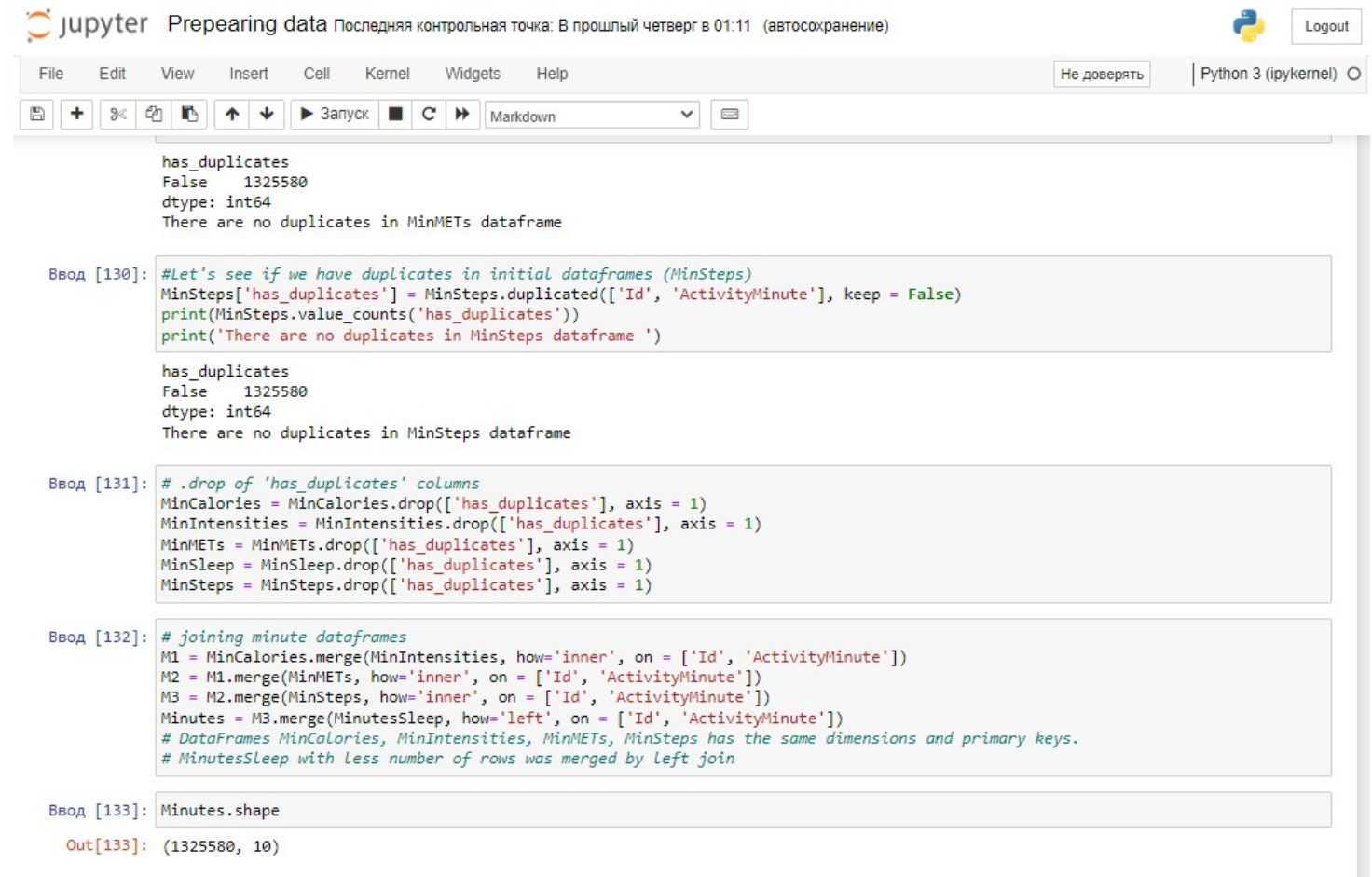The file "Prepearing_data.ipynb" is attached

Cleaning and manipulation of data documentation

This screenshot shows the process of merging several tables

The full process of data cleaning and manipulation you can find in attached 'ipynb' files:

- "Prepearing_data.ipynb"
- "Data_analysis.ipynb"

Visualizations
and key findings

Visualizations and key findings

Following questions may help find insights: **1. What parts of users' day are occupied by bed time, sedentary or active time**

Sum of sedentary/sleep/active average daily minutes
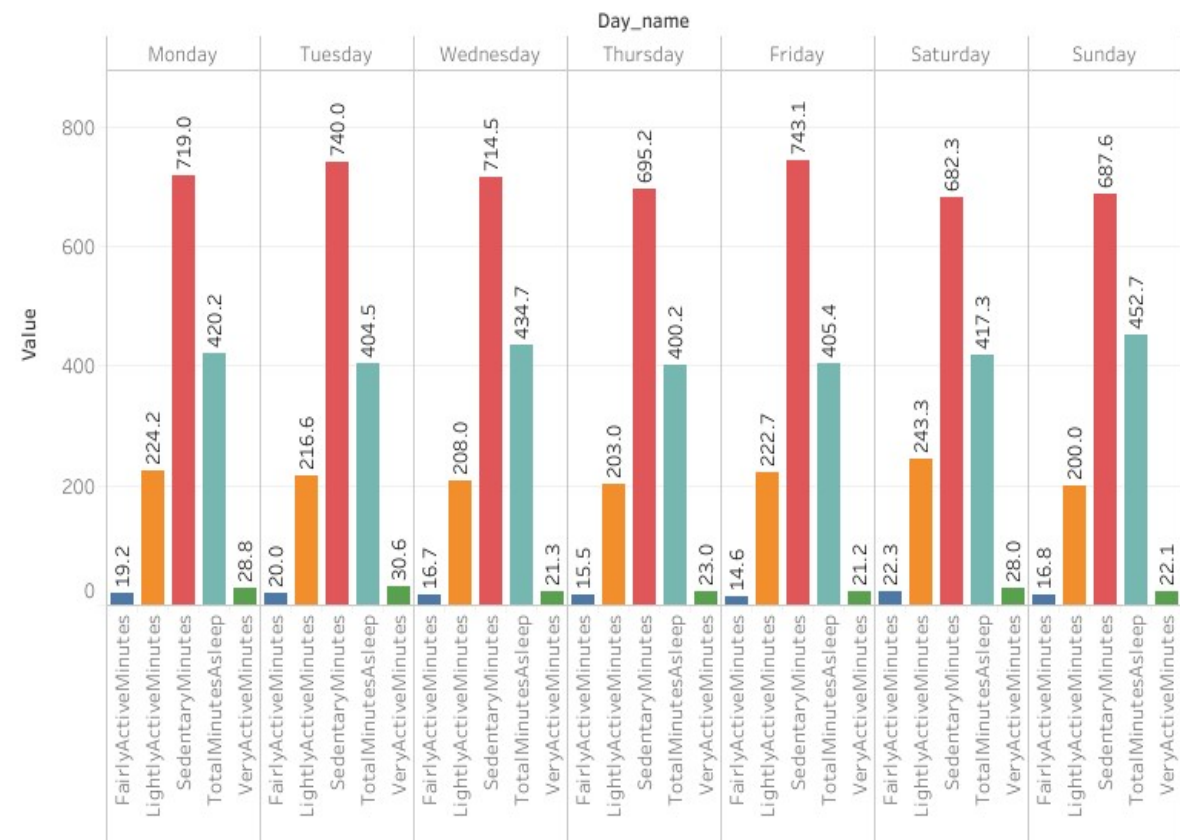
The longest sedentary time can be observed in Friday and Tuesday.
A little shorter sedentary time in Saturday and Sunday.

Minutes asleep has the maximum level in Wednesday, followed by Sunday.

Lightly active minutes has the maximum level in Saturday.

On average, users of smart devices have the following distribution of time during the day:

Sedentary_time	11.86 hours
Sleep_time	6.99 hours
Lightly_Active	3.61 hours
Very_Active	0.42 hours
Fairly_Active	0.3 hours

which is 23.18 hours in total.

The rest 42 minutes is probably average time occupied by time in bed not sleeping.

Visualizations and key findings

Following questions may help find insights: **What time the majority of users wakeup/go sleep**

Group by hour
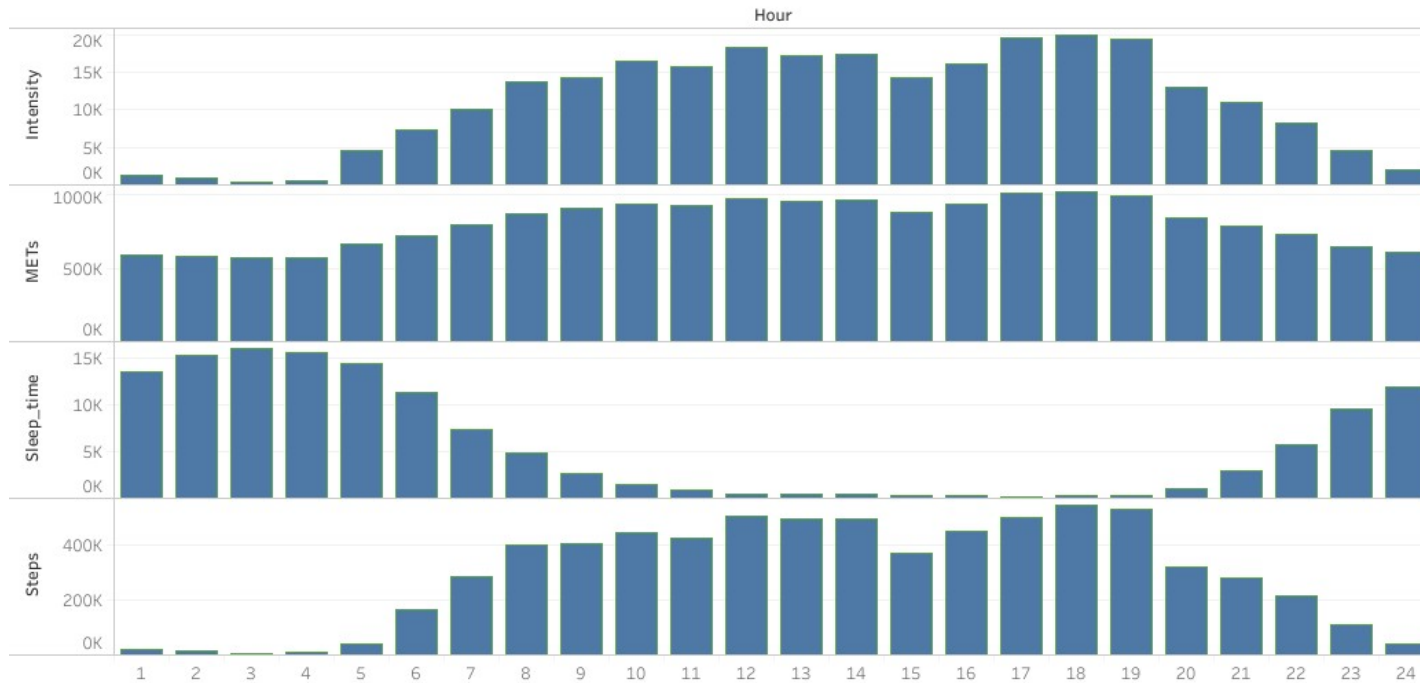
Intensity minutes are earned based on current heart rate when compared to average resting heart rate or the number of steps taken per minute, that is why it repeats steps dynamics.

Most smart device users sleep from 9 pm to 8 am.

Visualizations and key findings

Following questions may help find insights: **What is the average amount of steps users do per day**

Total steps Boxplot
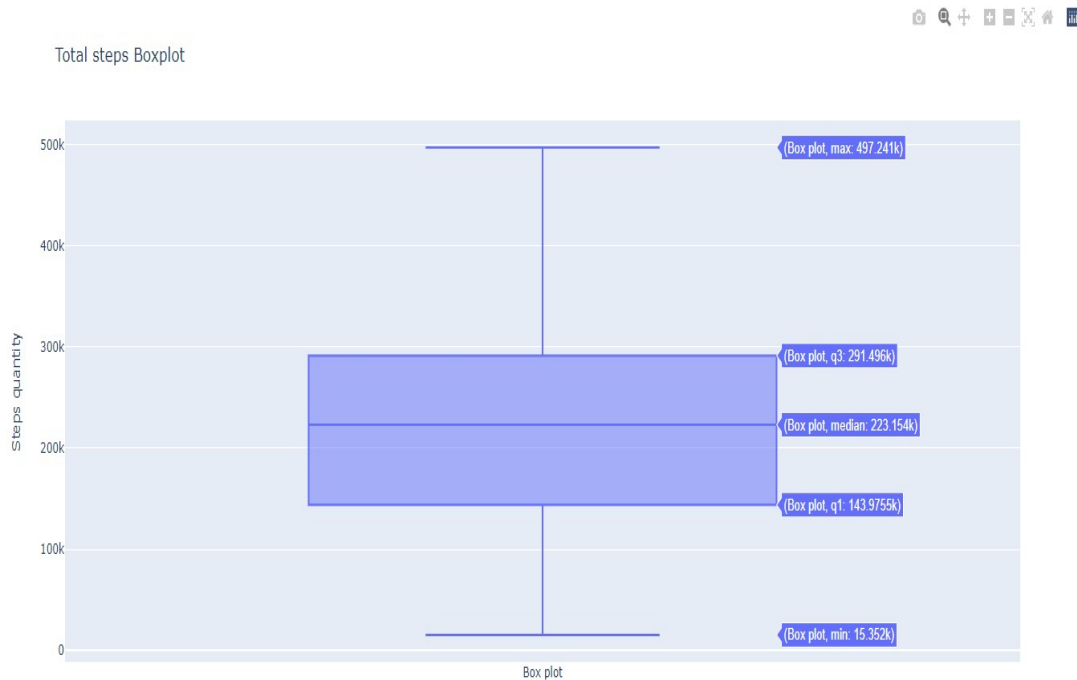
Average number of steps per day is 7018. However, it is likely that some people are more active and take more steps, and some less than the average. It can be useful to know how smart device users can be segmented by the number of steps.

By determining quartiles, all users can be segmented according to level of their activity
➢ **the first quartile** is when a peson do up to 144K steps in 31 days, (~4.6K per day), can be called passive segment
➢ **the second quartile** - from 144K to 223K (~ 4.6K - 7.2K per day), can be called upper-passive segment
➢ **the third quartile** - from 223K to 291K (~ 7.2K - 9.4K per day), can be called semi-active segment
➢ **the fourth quartile** - from 291K to 497K (~ 9.4K - 16K per day), can be called active segment

Visualizations and key findings

Following questions may help find insights:
4.Is there any day per week with maximum/minimum activity
5.Is there any day per week with maximum/minimum time sleep time

4. Is there any day per week with maximum/minimum activity

```
вод [24]:  # group by Day_name
           Day_name_steps_df = Minutes.groupby(['Day_name'])['Steps'].agg(['sum']).sort_values(by='sum', ascending=False)
           Day_name_steps_df.reset_index(inplace = True)
           Day_name_steps_df.head(10)
```

Out[24]:

	Day_name	sum
0	Tuesday	1204432
1	Wednesday	1117874
2	Thursday	1055176
3	Saturday	1003914
4	Friday	934681
5	Monday	922682
6	Sunday	834790

The most step based active day according to this dataset is Tuesday, the least is Sunday

The most step based active day according to this dataset is Tuesday, the last is Sunday

5. Maximum/minimum sleep time per week

```
вод [25]:  # group by Day_name
           Day_name_steps_df = Minutes.groupby(['Day_name'])['Sleep_time'].agg(['sum']).sort_values(by='sum', ascending=False)
           Day_name_steps_df.reset_index(inplace = True)
           Day_name_steps_df.head(10)
```

Out[25]:

	Day_name	sum
0	Sunday	23026.0
1	Wednesday	22797.0
2	Thursday	21481.0
3	Saturday	20008.0
4	Tuesday	19215.0
5	Monday	15746.0
6	Friday	15064.0

Sunday shows maximum sleep time, Friday - minimum

Visualizations and key findings

Following questions may help find insights: **Maximum steps, calories per hour**

6. Maximum steps, calories per hour

```python
#Groupby Minutes dataframe by Hour column with aggregation of 'Steps'
Hour_steps = Minutes.groupby(['Hour'])['Steps'].agg(['sum']).sort_values(by='sum', ascending=False)
Hour_steps.reset_index(inplace = True)
Hour_steps.head()
```

Ввод [28]:

Out[28]:

	Hour	sum
0	18	542848
1	19	528552
2	12	505848
3	17	498511
4	14	497158

Ввод [29]:

```python
#Groupby Minutes dataframe by Hour column with aggregation of 'Calories'
Hour__calories = Minutes.groupby(['Hour'])['Calories'].agg(['sum']).sort_values(by='sum', ascending=False)
Hour__calories.reset_index(inplace = True)
Hour__calories.head(5)
```

Out[29]:

	Hour	sum
0	18	111874.801325
1	17	111207.904204
2	19	110051.666754
3	12	108051.356172
4	14	106278.275241

Ввод [30]:

```python
# 6, 7 p.m - maximum steps hours
# 6, 5 p.m - maximum calories hours
```

6, 7 p.m - maximum steps hours
6, 5 p.m - maximum calories hours

Visualizations and key findings

Following questions may help find insights: **7. Is there any correlation btw calories burnt and sleep time**
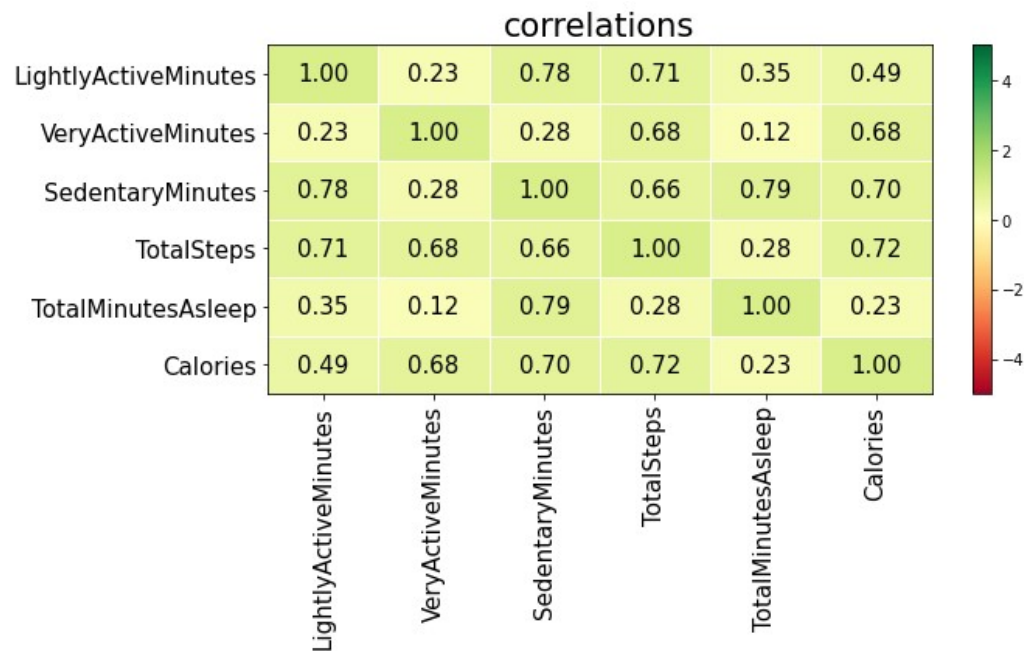
The table above reveal correlations between variables. The closer the value is to 1, the greater correlation.

We can see that total minutes asleep do not influence calories much.
Total steps and very active minutes have much more influence on burning calories process.
Sedentary time has 0,7 value with calories, but it is more likely because of long time of sedentary per day.

Visualizations and key findings

Following questions may help find insights: **8. What is the main difference btw people who burn relatively larger calories amount in comparison with those who burn less.**

Top 5 and last 5 calories burners and their activity level
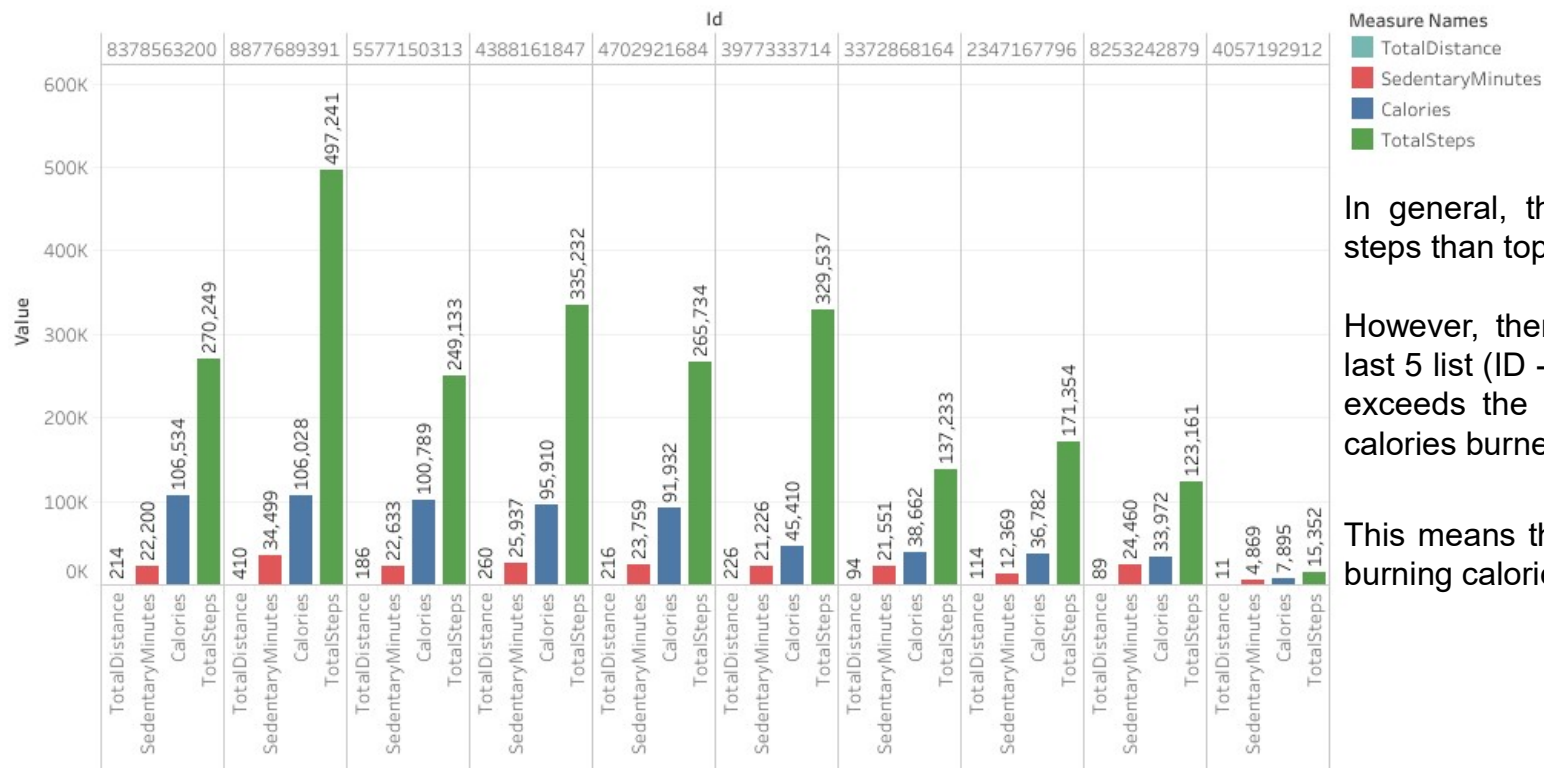
In general, the last 5 calorie burners have fewer steps than top 5.

However, there is one smart device user from the last 5 list (ID - 3977333714), whose number of steps exceeds the number of leader's steps in terms of calories burned (ID - 8378563200).

This means that there is another important factor in burning calories

Visualizations and key findings

Following questions may help find insights: **8. What is the main difference btw people who burn relatively larger calories amount in comparison with those who burn less.**

Top 5 and last 5 calories burners and their activity level 2
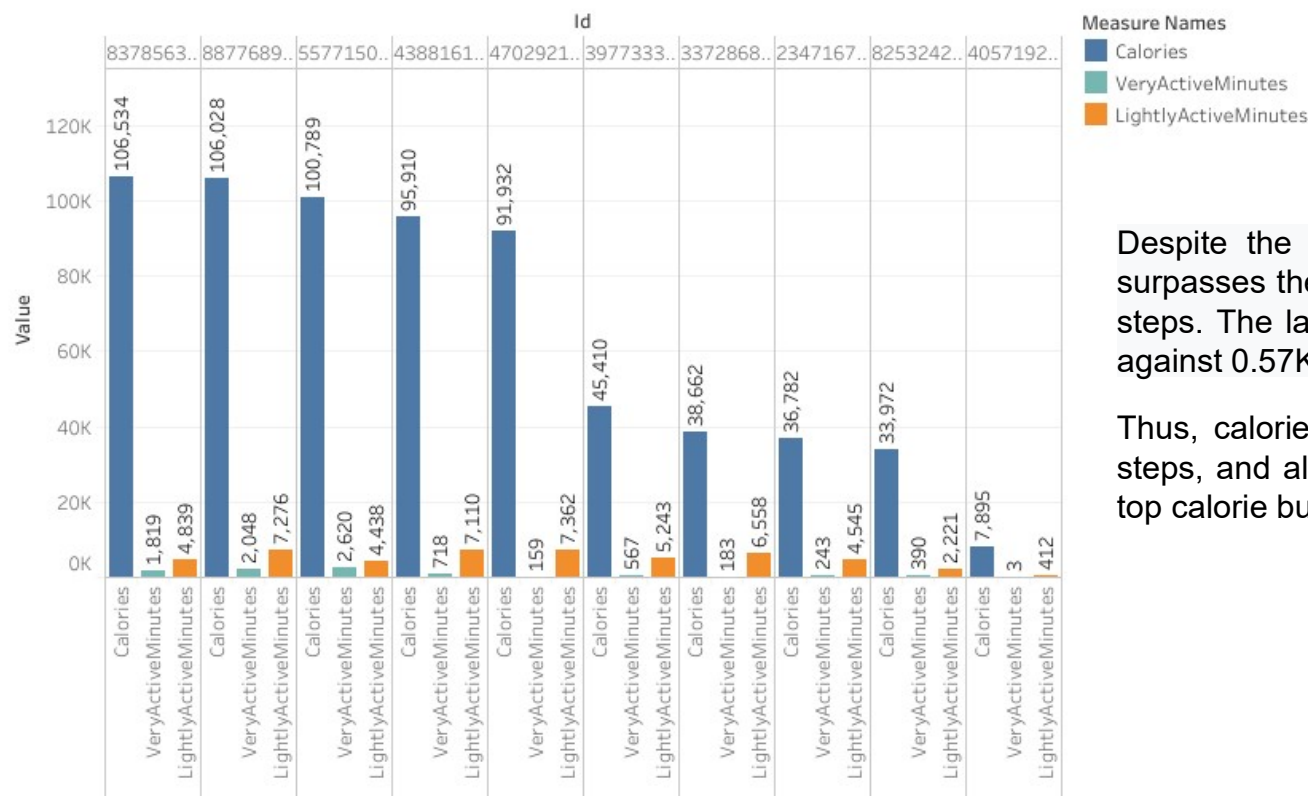
Despite the fact that user from the last5 list (ID - 3977333714) surpasses the top calorie burner (ID - 8378563200) in the number of steps. The latter outperforms by VeryActiveMinutes, which are 4.8K against 0.57K.

Thus, calorie expenditure is strongly influenced by the number of steps, and also by the number of very active minutes per day. The top calorie burner has about 7 hours a week of very active time.

Visualizations and key findings

61% of people own a smart device.

Think with Google

OMD, Global (Australia, Belgium, China, Greece, Ireland, Italy, KSA, Netherlands, Spain, Sweden, U.K., U.S.), Retail Revolution, Online Respondents n=11582, Sept.–Nov. 2019.

74% Tech shoppers spend 74% of their time researching online versus 26% offline.

Think with Google

Google/LRWGreenberg, U.S., Consumer Electronics Path to Purchase Study, Base: consumers who have purchased (in the past three months) or intend to purchase (within six months) consumer tech products, n of 2,243 survey respondents and n of 1,499 clickstream panelists, Aug. 2019.

Tech devices have ranked as one of the top 5 most-purchased gift categories each of the past two years.

Think with Google

Google/Ipsos, U.S., Holiday Study, n=5,543 online Americans 18+ who shopped in the past two days, Nov. 2018–Jan. 2019.

Smart speakers, smartwatches, wireless headphones, and wireless speakers are among the top consumer electronic devices purchased on a mobile phone. Gaming consoles, smartphones, and smart TVs are still typically purchased in stores.

Think with Google

Google/LRWGreenberg, U.S., Consumer Electronics Path to Purchase Study, Base: consumers who have purchased (in the past three months) or intend to purchase (within six months) consumer tech products, n of 2,243 survey respondents and n of 1,499 clickstream panelists, Aug. 2019.

The company has invested in traditional advertising media, such as radio, out-of-home billboards, print, and television, but focuses on digital marketing extensively. Bellabeat invests year-round in Google Search, maintaining active Facebook and Instagram pages, and consistently engages consumers on Twitter. Additionally, Bellabeat runs video ads on Youtube and display ads on the Google Display Network to support campaigns around key marketing dates.
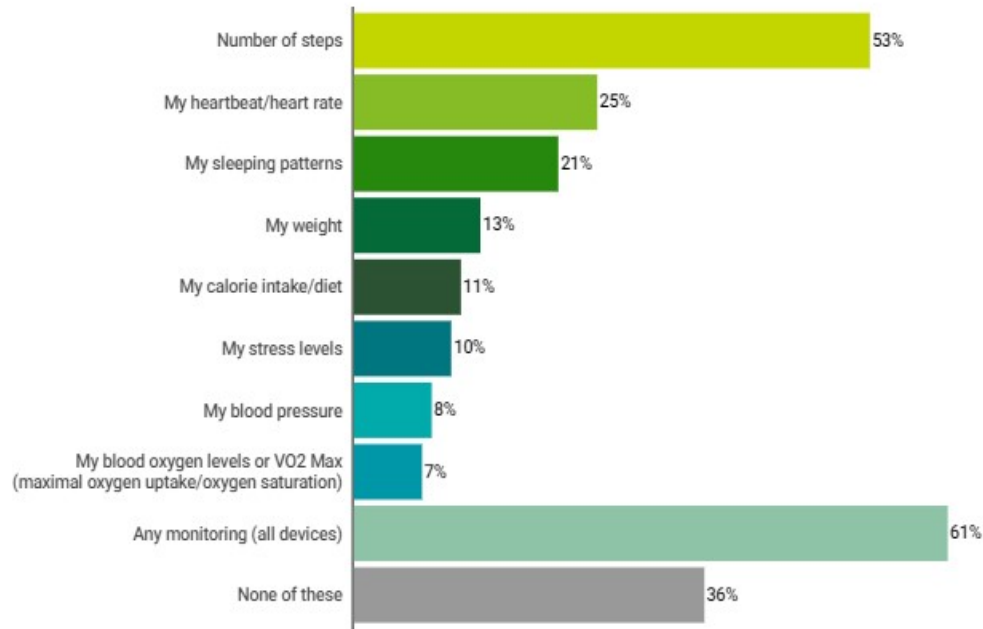
- The company should consider decreasing investment in traditional advertising media, such as radio, out-of-home billboards, print, and television.

- Special gift packages or birthday discounts might be implemented

Visualizations and key findings

Analysis of health monitoring devices in Scandinavia by Deloitte.
Source: https://www2.deloitte.com/dk/da/pages/technology-media-and-telecommunications/topics/digital-consumer-trends/new-gadgets-early-trends-in-future-smart-life.html

Most popular metrics in Scandinavia

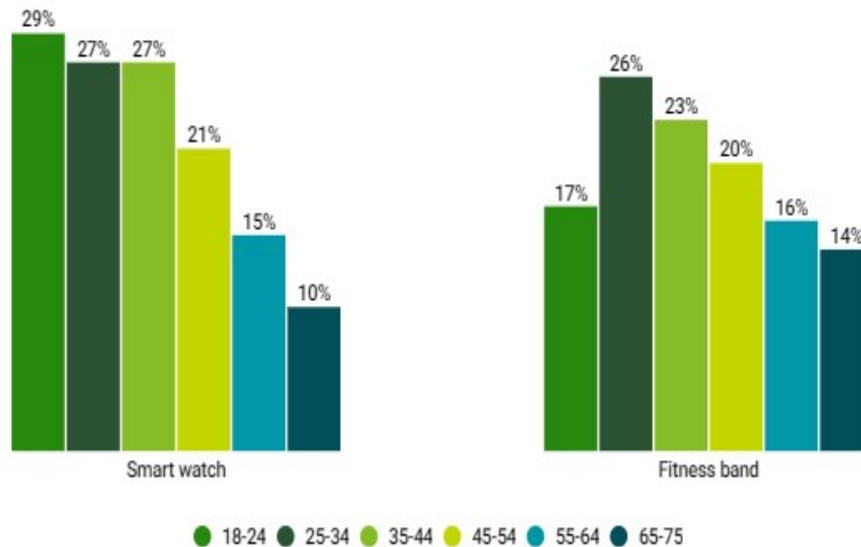| Metric	Percentage
Number of steps	53%
My heartbeat/heart rate	25%
My sleeping patterns	21%
My weight	13%
My calorie intake/diet	11%
My stress levels	10%
My blood pressure	8%
My blood oxygen levels or VO2 Max (maximal oxygen uptake/oxygen saturation)	7%
Any monitoring (all devices)	61%
None of these	36%

This global trend is also visible in Scandinavia, where 5% of the survey respondents say that they use devices to monitor their health more than they did before the pandemic. 62% of Scandinavians with access to a smart device, including a smartphone, now monitor one or more health statistics, with the most popular metrics being the number of steps taken, heart rate and sleeping patterns.

Visualizations and key findings

Analysis of health monitoring devices in Scandinavia by Deloitte.
Source: https://www2.deloitte.com/dk/da/pages/technology-media-and-telecommunications/topics/digital-consumer-trends/new-gadgets-early-trends-in-future-smart-life.html

Age distribution among health monitoring gadgets in Scandinavia

18-24 ● 25-34 ● 35-44 ● 45-54 ● 55-64 ● 65-75

The type of health gadget used varies across age, gender and region
Smartwatches are the most popular device for monitoring health among the 18- to 44-year olds, whereas 55- to 75-year olds prefer fitness bands. Smartwatches are slightly more popular among men, while fitness bands are more popular among women. Furthermore, 35- to 75-year olds are the most frequent users of smartwatches or fitness bands to monitor health.

High-level recommendations

High-level recommendations

The analyzed data may not be statistically significant due to the small number of participants and short time period. It is recommended to involve the company's own clients for a longer-term collection of information (preferably on an ongoing basis).

The Bellabeat app provides users with health data related to their activity, sleep, stress, menstrual cycle, and mindfulness habits. It is also possible to provide users with some insights, gained from collected data, such as average number of steps, level of activity of burning calories leaders and so on, so that users can compare their activity level, which can be motivational factor.

According to FitBit Fitness Tracker Data, users can be segmented by four groups depending on daily number of steps. This will allow to target each group individually.
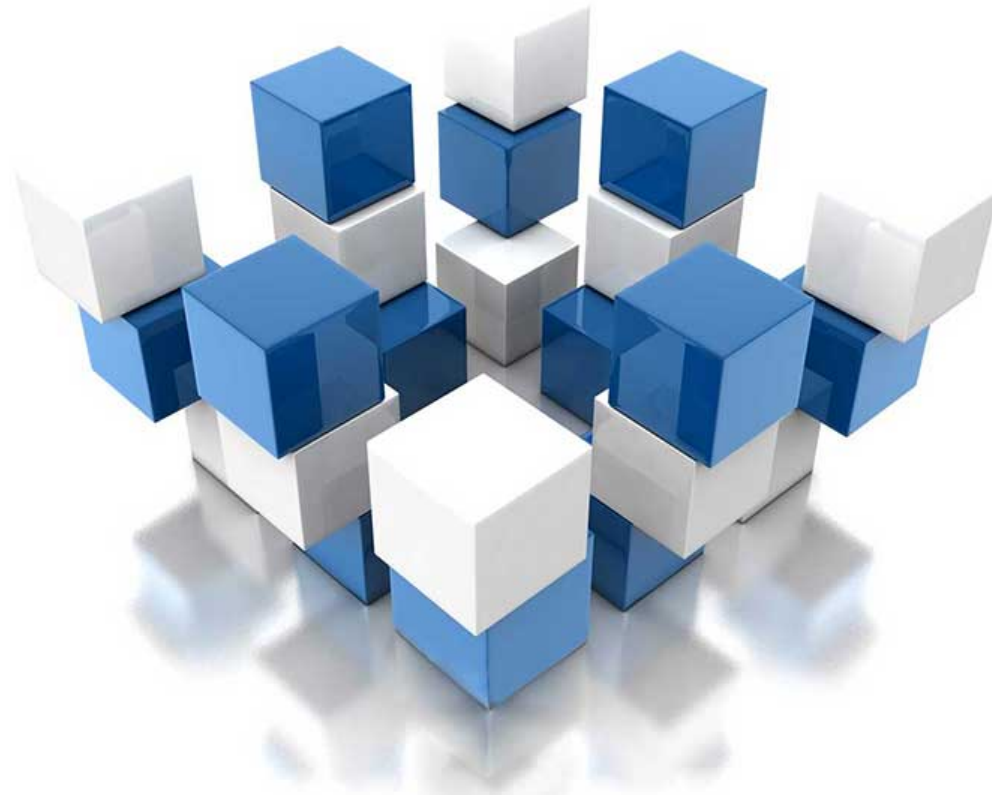The company may consider an extension of product range to offer. For example, rings with a special design for women that will also track activity, sleep, and stress. Or sneakers with the same functionality with targeted ads for women from semi active or active segments.

According to Analysis of health monitoring devices in Scandinavia by Deloitte, smartwatches are the most popular device for monitoring health among the 18- to 44-year olds, whereas 55- to 75-year olds prefer fitness bands. Smartwatches are slightly more popular among men, while fitness bands are more popular among women. Furthermore, 35- to 75-year olds are the most frequent users of smartwatches or fitness bands to monitor health.
First of all it will be useful to organize such kind of survey in all countries of Bellabeat presence.
Although I believe that the company should consider decreasing investment in traditional advertising media, such as radio, out-of-home billboards, print, and television, it must be taken into account that many 55- to 75-year olds people may be active users of resources listed above. So we can assume that, targeted to 55- to 75-year old women, marketing activities of "leaf" promotion, using traditional advertising media might bring positive results and decrease overall marketing budget.

According to 'think with google' source, tech devices have ranked as one of the top 5 most-purchased gift categories. The company might consider special gift packages or birthday discounts implementation.

THANK YOU