

Case Study: How to Improve Bellabeat's Marketing Strategy

Abstract

Bellabeat is a high-tech manufacturer of health-focused products for women that wants to expand its market share in the global smart device market. My task was to analyze the FitBit dataset to gain insight into how consumers use non-Bellabeat smart devices and I applied these insights to the Bellabeat app to guide the company's marketing strategy. By using public data provided by FitBit Fitness Tracker Data, I aimed to identify trends, such as the positive correlation between the number of steps taken and the amount of calories burned, the slight difference in step-taking behavior between weekdays and weekends, and the impact of weight on calorie burn using RStudio. The main goal of the case study is to help Bellabeat's business and marketing goals.

Introduction

Bellabeat is a high-tech manufacturer of health-focused products for women that wants to expand its market share in the global smart device market. To achieve this, according to the company's request I analyzed smart device usage data and applied insights to guide Bellabeat's marketing strategy in the future. I used public data provided by FitBit Fitness Tracker Data to identify trends, patterns and suggestions in smart device usage.

The programming language I used to analyze the data was R, and various statistical techniques were applied to draw conclusions according to my hypotheses. The results indicated sometimes strong sometimes weak correlations between certain factors and the participants' choices, providing insights into consumer behavior and decision-making processes.

Ultimately, these insights can help Bellabeat develop targeted marketing campaigns and messaging that address their specific fitness goals and concerns, unlock new growth opportunities by gaining a better understanding of its customers' needs and preferences and by that improving Bellabeat app, based on my recommendations to the Bellabeat executive team, who will use the findings to guide the company's future marketing strategy.

Research Questions

In order to improve their business and marketing strategy of Bellabeat, I decided to conduct my case study considering three main research questions. The first main goal was to investigate the relationship between the number of steps taken and the amount of calories burned among FitBit users. Therefore, the first research question is: "Is there a relationship between the number of steps taken and the amount of calories burned among FitBit users?" The results of this analysis could be used to inform the development of Bellabeat's fitness tracking products, such as providing users with more accurate estimations of calorie burn based on their step count.

The second question focuses on the differences in the number of steps taken and the amount of calories burned between weekdays and weekends among FitBit users: “Is there a difference in the number of steps taken and the amount of calories burned between weekdays and weekends among FitBit users?” This could be valuable information for designing targeted promotions or features in Bellabeat app to encourage users to stay active on weekdays when they may be more likely to be sedentary due to work or other obligations.

The final question aims to investigate how the relationship between the number of steps taken and the amount of calories burned differs between users of different weight groups: “How does the relationship between the number of steps taken and the amount of calories burned differ between users of different weight groups?”. This information could be used to provide personalized recommendations for users on how to optimize their activity level to achieve their fitness goals. Overall, by investigating these research questions, I aim to better understand the behavior and preferences of Bellabeat’s target audience and to use these insights to enhance its products and services.

Hypotheses

Bellabeat's business and marketing strategy could benefit greatly from testing several hypotheses using the FitBit database. The first hypothesis states that there is a positive correlation between the number of steps taken and the amount of calories burned among FitBit users. By analyzing the data, Bellabeat can use this knowledge to develop products or features that motivate users to increase their step count and improve their health.

The second hypothesis states that FitBit users will take more steps on weekdays than on weekends, but the difference is not significant. If this would be with a significant result, this information could be used to tailor marketing campaigns that encourage users to be more active during the week, and to design features that are particularly helpful during the workweek. However, since I expect an insignificant result, this information is to help Bellabeat to avoid focusing on weekdays and weekends as it might be a misleading track.

Finally, the third hypothesis suggests that users with a higher weight will burn more calories when taking the same number of steps compared to users with lower weight. This could be an important finding for Bellabeat, as it could allow them to design products or features that are specifically targeted to each users of higher or lower weight. With these findings, Bellabeat can improve its business, app and better meet the needs of its users.

Data

For my analyses I use FitBit Fitness Tracker Data, which is a public dataset made available through Mobius (“FitBit Fitness Tracker Data” n.d.). Among the many available CSV files I will use 3 for my main questions and findings. Firstly, daily activity (“dailyActivity_merged.csv”) which contains daily activity data for a group of individuals. The data includes various variables such as step count, distance, calories burned, and minutes of activity. This is basically the main dataset for my use, as calory and steps are a main pillar of my research.

Another CSV file is the “sleepDay_merged.csv” containing sleep data for the same group of individuals as the daily activity data. The data includes observations such as time in bed, minutes asleep, or sleep efficiency. The main purpose of this dataset is some preliminary calculations with the dataset to understand the sleeping habits of the FitBit users.

The thirdly user data is “weightLogInfo_merged.csv” which is a CSV file containing weight data for the same group of individuals as in the dailyActivity and sleepDay data. The data includes weight, BMI, and body fat percentage of the users, but weight is the most reliable variable in this dataset as the other two is lacking many observations.

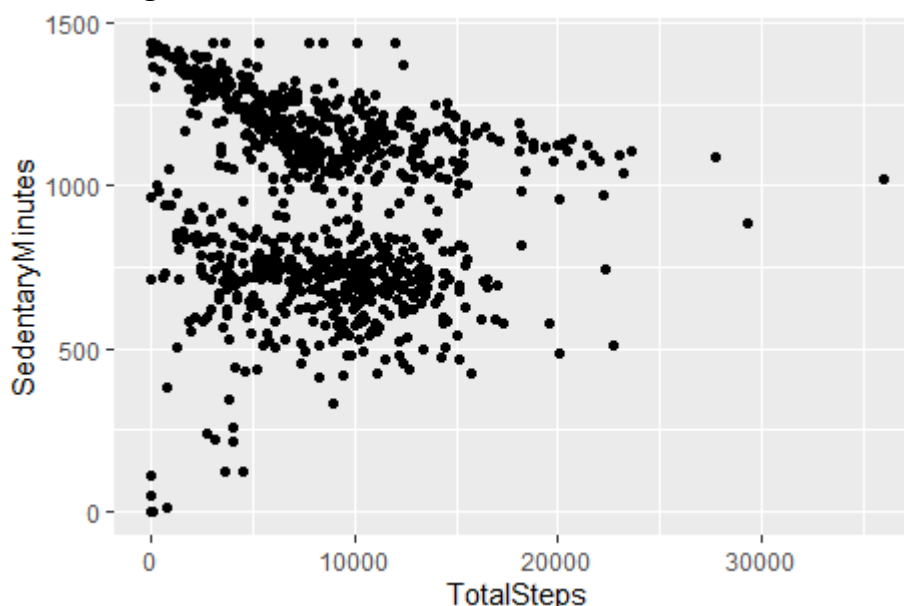
Cleansing and Manipulation of Data

The code conducted in R can be divided in two main parts. In the first part, the necessary packages are installed, the data is read, and a quick exploration of the data is conducted. The data is then cleaned and merged, and some basic plots are created to visualize the data. The code begins by installing and loading the required libraries for data manipulation and visualization: tidyverse, lubridate, ggplot2, readxl, and dplyr. The data is then read into the R environment using read.csv. The three datasets used in the analysis are dailyActivity, sleepDay, and weightLogInfo, the equivalents of the abovementioned CSV files. The number of distinct participants is checked using n_distinct, and a summary of the relevant columns in each dataset is printed using summary. Data cleansing and merging is performed where I combine the first main dataset with the sleepDay and weightLogInfo datasets. For analytic purposes I’m interested in observations where the participants are present in both datasets.

Visualization and Evaluation of the Data

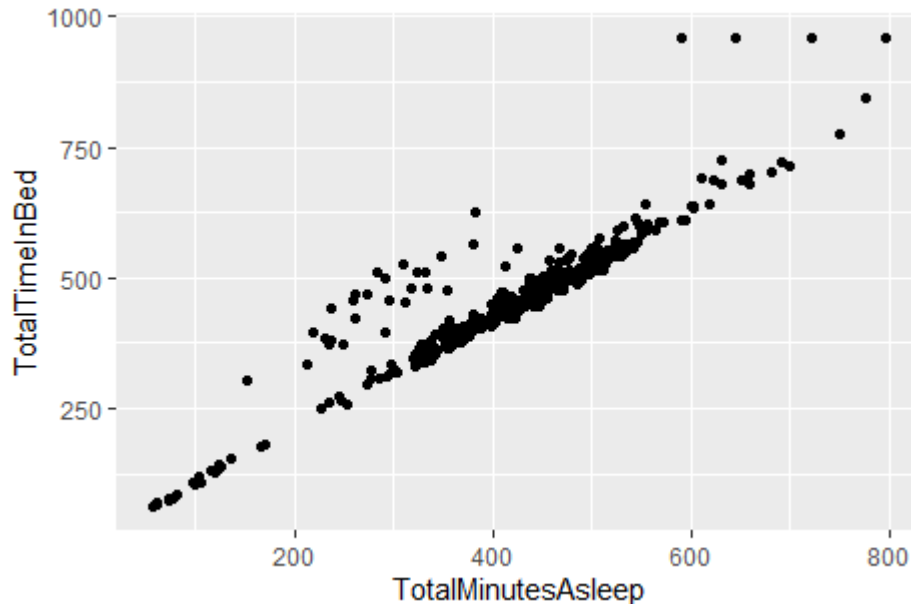
In the end of the first part of the research two plots are created: one comparing total steps and sedentary minutes (figure 1), and another comparing minutes asleep and time in bed (figure 2).

Figure 1



There is no strict correlation between total steps and sedentary minutes, however, until 10000 steps there is a group of people who sits more if he/she walks less. Although the visualized data doesn't let us presuppose any correlation, we can understand that most of the people sit more and walk less. Fitbit products could help to turn this bad habit around by making people aware of it through their products.

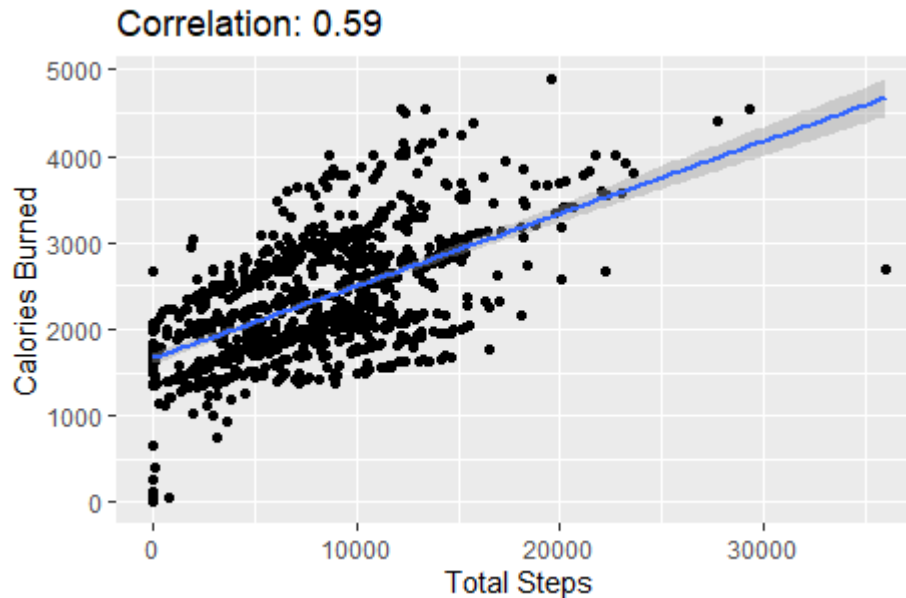
Figure 2



There is a strong correlation between total minutes asleep and total time in bed. There is only a few unexpected trends for those between 200 and 400 and between around 600 and more minutes sleeping, who tends to spend more time in bed. Bellabeat products could target people to improve people's bad sleeping habits and spent time in bed and could also use this already available public data, extend it with their own records and improve it to make their products more effective and appealing.

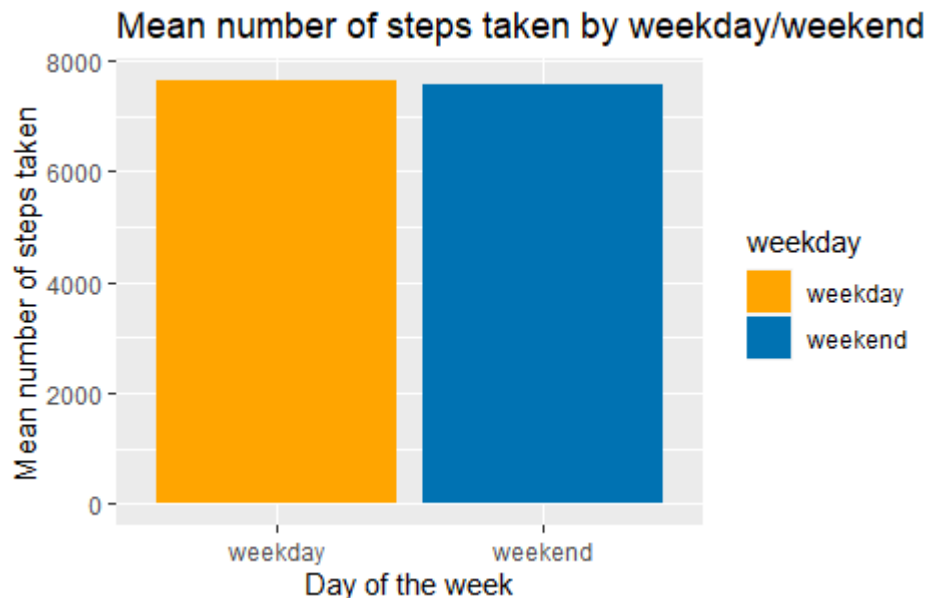
In the second part, three research questions are posed, and the corresponding hypotheses are stated. The first question is whether there is a relationship between the number of steps taken and the amount of calories burned among FitBit users. The hypothesis is that there is a positive correlation between the two variables. A correlation coefficient is calculated and a scatter plot with a linear regression line is created (Figure 3). The results show a positive correlation between the number of steps taken and the amount of calories burned, supporting the first hypothesis.

Figure 3



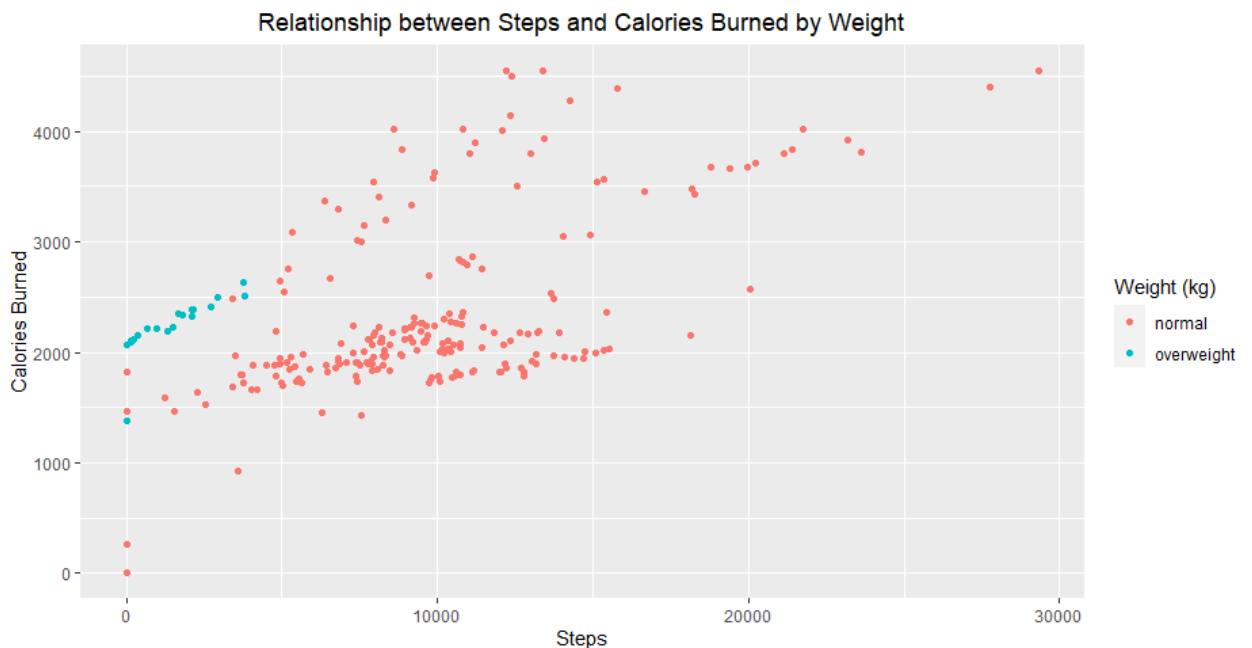
The second research question is whether there is a difference in the number of steps taken and the amount of calories burned between weekdays and weekends among FitBit users. The hypothesis is that users will take more steps on weekdays than on weekends, but the difference is not significant. Figure 4 shows this very clearly. The weekdays are just slightly higher than weekends, which positively answers the research question. Although, the result is not significant, due to the slight difference, it can still help Bellabeat in a way that daily activity measured in steps does not seem to be relevant when it comes to their marketing or business purposes. To double check my result, I ran a t-test to see whether our result is significant or not, where the p-value (0.847) proves that the result is insignificant as it does not meet the 5% p-value criterion.

Figure 4



The third research question is how the relationship between the number of steps taken and the amount of calories burned differs between users of different weight groups. This is the most relevant and I believe the most important hypothesis among all. The hypothesis is that users with a higher weight will burn more calories when taking the same number of steps compared to users with lower weight. Firstly, I created a merged file using the weightLogInfo and dailyActivity databases to make the analyses possible. Then I used three approaches to find the answer my hypothesis the most accurately. In the first approach I created 4 weight groups "underweight" (0-50 kg), "normal" (51-100 kg), "overweight" (101-150 kg) and "obese" (151-200 kg). Then I created a new database grouped by weight category, included the mean values of daily steps and calories burned. Plotting this result does not help answering the research question, but we can see that there are only two weight groups available in the database, normal and overweight. After checking the correlation coefficient between steps and calories and running a correlation test with significant p-value, I plotted the steps and calories burned to see if my hypothesis is valid (Figure 5).

Figure 5



The visualized data shows that the users with a higher weight will burn more calories when taking the same number of steps compared to users with lower weight. However, this seems to be true only until 5000 steps as there is a deviant group above it. As there are only 2 weight groups available of 4, the low sample size or wrong weight groups could both be problems.

Hence, I moved on to the second approach to answer the hypothesis. I decided to make a database grouped by the users and mean weight, steps and calories so that one user will take only one place in the visualized graph. The result (in Figure 6) seems to be too narrow, as we get rid of many valuable information and trends.

Figure 6



It does not seem wise to limit ourselves to individuals only, as people's steps, weight, burned calories and all the other data (except the user ids) might change day to day. Therefore, it seems to be wiser to take every input for steps and burned calories as unique ones, no matter if many dots are connected to the same user.

This moves us further to the third approach where I decided to reorganize the weight groups user in the first approach. This seems to be necessary, since the lightest user is 53 kg and the heaviest is 134 kg heavy. I used the MMA weight categories and narrowed it down to four elements which included "Flyweight-Lightweight" (52-70 kg), "Welterweight-Middleweight" (71-84 kg), "Light Heavyweight-Heavyweight" (85-120 kg), "Super Heavyweight" (121-150 kg) groups. After this I checked the correlation coefficient and ran the correlation test, just as in the first attempt. After this I checked the relationship between steps and calories burned by weight (Figure 7). The purple and blue dots are the heavyweight and super heavyweight FitBit users, while the green and pink dots are the Flyweight-Middleweight users. The data shows that the pink group (lightweight) is burning the least calories with a given amount of steps, while the green group tends to burn more calories and the purple and blue (heavyweight) users burn the most amount of calories. As an example, a lightweight user might walk around 10000 steps, and they would burn around 2000 calories in a day. For heavyweight users this number is 4000 calories, twice as much as for lightweight users.

Figure 7



There is another working way for this approach which further approves our findings by giving a slightly different visualization of the data where we can just ignore the weight categories and use the available data of weight in kilograms instead (Figure 8). In this case the trend is even more visible.

Figure 8



In conclusion, we can say that there is a correlation between the number of steps taken and the amount of calories burned among FitBit users of different weight groups.

Key Findings

Based on the analysis of the Bellabeat Fitness Tracker Data, there were several key findings and supporting visualizations that could be of interest to the Bellabeat business task. Firstly, there is a positive correlation between the number of steps taken and the amount of calories burned among FitBit users. This means that as the number of steps taken by FitBit users increases, their calorie burn also increases.

Secondly, FitBit users will take more steps on weekdays than on weekends, but the difference is not significant. This suggests that there is a slight pattern in FitBit users' step-taking behavior based on weekdays versus weekends, but it is not strong enough to be statistically significant.

Finally, users with a higher weight will burn more calories when taking the same number of steps compared to users with lower weight. This implies that users with a higher weight will burn more calories even if they take the same number of steps as users with a lower weight. Overall, these findings and supporting visualizations could be valuable to the Bellabeat business task in terms of understanding user behavior and identifying potential areas for improvement in their product offerings.

Application to Bellabeat Products

The key findings can apply to Bellabeat customers in several ways. Bellabeat customers who use the Bellabeat app, can expect to see a positive correlation between the number of steps they take and the amount of calories burned. This can help them track their physical activity levels and make adjustments to their fitness routine as needed. Their daily steps and burned calories or even sedentary minutes or other available data the tracker or other devices may provide can help the users to make more informed decisions about when to schedule their workouts, how to plan their daily activities or when to rest.

The third hypothesis helps us understand that weight is a significant factor for setting fitness goals. If the users have a higher weight, they will more likely burn more calories with lower physical activity. Therefore, for them a potential primary goal could be not to improve activity, but to apply a healthier diet with the higher calorie burn and lower calorie intake if their main goal is to lose weight. This knowledge and the application of this knowledge to Bellabeat app and potentially to other Bellabeat products can help the users to set up more realistic and achievable personalized fitness goals according to their habits, weight, main purpose of their activity, etc. Therefore, briefly, Bellabeat could provide personalized guidance to customers based on their weight, activity levels, and fitness goals.

Marketing Strategy Recommendations

These trends can help influence Bellabeat's marketing strategy by emphasizing the positive correlation between the number of steps taken and the amount of calories burned can be a key selling point for fitness-oriented smart devices. This messaging can highlight how the devices can help users track their physical activity levels and monitor their calorie burn more accurately. By highlighting the difference in step-taking behavior between weekdays and weekends can be useful

for companies looking to market their devices to individuals who are looking to improve their fitness routine or establish a more structured workout schedule (even if there is no significant difference between weekday and weekend daily steps in general). Also, companies could leverage the knowledge that users with higher weight burn more calories when taking the same number of steps as users with lower weight. This can be especially useful for marketing to individuals who are looking to lose weight or manage their weight more effectively. Companies can highlight how their devices can help users track their calorie burn and adjust their physical activity levels to achieve their weight loss goals more efficiently.

Limitations

The absence of gender separation in the dataset is a problem for Bellabeat as it is a company that primarily focuses on female users. It is important to note that although the study did not separate male and female users, the findings may still provide meaningful insights for Bellabeat's business and marketing strategy. However, as Bellabeat primarily focuses on female users, the absence of gender separation is a bias and limitation that should be addressed in future studies to ensure a more comprehensive understanding of the user base.

Furthermore, the small number of participants in the database may introduce bias and limit the generalizability of the conclusions drawn from the analysis. It is important to acknowledge this limitation and consider the need for further research with a larger sample size to provide more robust and reliable insights.

Conclusion

In my case study, I prepared and analyzed the data provided by FitBit Fitness Tracker. I followed several steps to prepare the data for analysis, such as the loading of data, giving them the most reasonable names, manipulated the data using mean numbers of the data where it was necessary, for example to measure the mean number of steps taken by weekday and weekend. I merged the daily activity, sleep, and weight datasets in a way to make it straightforward for analyses. This means that there were two main datasets combined using a common identifier (the "Id" variable) for each individual. Instead of two I could have made one dataset out of the three, but that would not have made any difference. I subset the data as needed to focus on specific research questions. For example, I subsetted the data to focus on weekdays or weekends.

I conducted the analyses based on my 3 hypotheses. I found that there is a positive correlation between the number of steps taken and the amount of calories burned among FitBit users. Although, FitBit users take more steps on weekdays than on weekends, but the difference is not significant. The most relevant and detailed finding was that users with a higher weight burn more calories when taking the same number of steps compared to users with lower weight. These findings could help Bellabeat in his future market strategy, on what or what not to focus on.

In conclusion, the analysis of smart device usage data has provided valuable insights that can help Bellabeat improve its app and other product offerings and marketing strategy. By leveraging these trends, Bellabeat can develop targeted marketing campaigns and messaging that resonates with its target audience and addresses their specific fitness goals and concerns.

Citations

“FitBit Fitness Tracker Data.” n.d. Accessed March 6, 2023.
<https://www.kaggle.com/datasets/arashnic/fitbit>.