

BellaBeat Case Study Report

Author: Aadit Shetty

Report date: 8/14/2021



Introduction

This case study was made as a final project for the Google Data Analytics Certificate. Here, I have been appointed as a data analyst in the analytics team of BellaBeat and am asked to analyze certain data and help the company grow successfully by guiding them using the insights gleaned therein.

The Mission Statement

[BellaBeat](#) is a high-tech company that manufactures health-focused smart products wants to analyse the usage of one of their products in order to gain insight into how people are already using their smart devices. Then, using this information, she would like high-level recommendations for how these trends can inform Bellabeat marketing strategy.

Purpose

Provide insight on trends in smart device usage, which can influence Bellabeat marketing strategy

Business Task

1. Identify trends in smart device usage
2. Find trends which can be applied to Bellabeat customers
3. Propose ideas, which can help influence Bellabeat marketing strategy

Key Stakeholders

1. Urška Sršen: Bellabeat's cofounder and Chief Creative Officer
2. Sando Mur: Mathematician and Bellabeat's cofounder
3. Bellabeat marketing analytics team

Bellabeat Target Audience

Helping and encouraging women in monitoring and improving physical and mental health.

Data Source and Assessing Data Quality

Source : I will analyse FitBit Fitness Tracker Data to make useful insights for Bellabeat. [Click here](#) to access the data source

1. Reliability (Low)
2. Small sample size Originality (Low)
3. Collected by the third party - Amazon Mechanical Turk Comprehensive (Low) May not have the same target customers as Bellabeat Does not included key parameters crucial for analysis (eg. age, gender)
4. Currency (Medium) Collected in 2016 -> has started to loose its up to date status
5. Cited (Low) The data has been downloaded and sited. However, the data set is mostly being used by junior data analysts, thus, it potentially lowers its credibility.

Documentation of Data Analysis Procedures

Identify the Business

Analyze FitBit fitness tracker data to gain insights into how consumers are using the FitBit app and discover trends for Bellabeat marketing strategy.

Prepare and Process data

I have used Excel to,

1. Observe the data
2. Adjust data types
3. Add data
4. Remove duplicate values
5. Remove empty rows
6. Other data manipulation purposes

I have stated below the steps I followed for Preparing, Processing and Cleaning data in R Studio Cloud.

Packages I have used:

1. tidyverse
2. lubridate
3. tidyr

Setting up the directory and listing the files:

```
setwd("/cloud/project/Data")  
getwd()  
list.files()
```

Reading the contents of the files and analyzing the data structure

```
daily_activity <- read.csv("dailyActivity_merged.csv")
hourly_calories <- read.csv("hourlyCalories_merged.csv")
sleep_day <- read.csv("sleepDay_merged.csv")
str(daily_activity)
str(hourly_calories)
str(sleep_day)
```

Looking for unique users

```
unique(daily_activity$id)
unique(hourly_calories$id)
unique(sleep_day$id)
```

Cleaning data:

- **Separating date and time from ActivityHour column**
hourly_calories <- separate(hourly_calories, ActivityHour, into = c('Date', 'Time'), sep = ' ')
head(hourly_calories)

- **Changing the data type**
daily_activity\$ActivityDate <- mdy(daily_activity\$ActivityDate)
hourly_calories\$Date <- mdy(hourly_calories\$Date)
sleep_day\$SleepDate <- mdy(sleep_day\$SleepDate)
glimpse(daily_activity) glimpse(hourly_calories) glimpse(sleep_day)

- **Computing and arranging the average calories burnt per hour**

Computation:

```
avg_calperhr <- hourly_calories %>% group_by(Time) %>% summarise(avg_calperhr =  
mean(Calories))  
str(avg_calperhr)
```

Arranging the columns in order:

```
avg_calperhr %>% arrange(match(Time, c("0:00", "1:00", "2:00",  
"3:00", "4:00", "5:00", "6:00", "7:00", "8:00", "9:00", "10:00", "11:00", "12:00", "13:00", "14:00",  
"15:00", "16:00", "17:00", "18:00", "19:00", "20:00", "21:00", "22:00", "23:00")), desc(avg_calperhr))
```

- **Calculating the average calories burnt per day in a week**
avg_calperday <- daily_activity %>% group_by(Day) %>% summarise(avg_calperday =
mean(Calories))
str(avg_calperday)
- **Calculating the average daily steps by users in a week**
avg_stepsperweek <- daily_activity %>% group_by(Day) %>% summarise(avg_stepsperweek =
mean(TotalSteps))
str(avg_stepsperweek)
- **Calculating the average number of sleep minutes per day of the week and also the minimum average sleep**

Avg sleep per day

```
avg_sleepperday <- sleep_day %>% group_by(Day) %>% summarise(avg_sleepperday =  
mean(MinutesAsleep))  
str(avg_sleepperday)
```

Minimum average sleep

```
min_avg_sleepperday <- avg_sleepperday %>% summarise(min_avg_sleepperday  
= min(avg_sleepperday))  
str(min_avg_sleepperday)
```

- **Creating a Data Frame and Exporting it**

Creating the data frame

```
df1 <- data_frame(avg_calperday)  
df2 <- data_frame(avg_calperhr)  
df3 <- data_frame(avg_sleepperday)  
df4 <- data_frame(avg_stepsperweek)
```

Converting the data frame into a .csv file and thereby exporting it:

```
write.csv(df1, "avg_calperday.csv")  
write.csv(df2, "avg_calperhr.csv")  
write.csv(df3, "avg_sleepperday.csv")  
write.csv(df4, "avg_stepsperweek.csv")
```

Analysis and Sharing the Analysis

I have used Tableau for creating visualizations of my findings.

To view the same, click below:

1. [Viz. 1](#)

2. [Viz. 2](#)

1. Users tend to burn more calories mid-week
2. Users calories decrease from mid-week to around the start of the weekend
3. Users calories once again increase during the beginning of the weekend and take another dip when it is almost the start of the new week
4. Calories are at its lowest during sleep which is between 0:00 (12:00 AM) and 4:00 (AM)
5. It starts to increase at 4:00 (AM) and peaks around noon time
6. After lunch, it decreases for a bit and at around 15:00 (3 PM) or 16:00 (4 PM), the calories begin to increase until 19:00 (7 PM)
7. After 19:00 (7 PM) calories continue to decline into the night. Calories are low in the sleeping hours
8. Calories increase when between the morning and the afternoon. Inference: This is because many of the users are productive in the morning (work)
9. After peaking at lunch time, calories decrease until the evening. Inference: People feel sleepy and tired after lunch. They slow down in activities near closing time. 10 Calories increase at the evening and then begin declining into the night. Inference: People have time to do a little bit of exercise. Calories drop when people prepare for bed.

Recommendations

General

1. As people having less active lifestyle is a larger group, Bellabeat may focus on more casual device properties: aesthetics, battery life.

2. Watch can include a custom Calorie Intake tracker and help users track their diet easily and consistently.
3. Exercise alarm/planner: Plan ahead of the week and make sure that users fill up the “required” amount of calories that they need to burn according to their health.
4. Develop more designs to attract customers. Example custom watch bands, watch colours, etc.

Marketing

1. Time product is an elegant watch that would attract more buyers as it offers dual utilities of
 - a. a watch, which can be considered as a ‘necessity goods’ and,
 - b. a health tracking device. Thus a slight rise in the advertising of this product can be beneficial for the company.
2. If feasible, sign contracts with celebrities/athletes to act as ambassadors of the products.
3. Market the elegant look that Time offers thereby emphasizing that the watch can be worn on multiple occasions where a watch with a sporty look doesn’t fit well.

My Experience Completing the Case Study

The entire case study was completed within a period of a week. I, being a non-computer science student, found it a very challenging at times writing codes and using softwares. However, the strong desire to learn Data Analytics kept me going. I referred sites of R, Tableau, Excel, Kaggle, Medium and StackOverflow to find solutions whenever I got stuck. The whole process of this data analytics case study was fun and enlightening. I hope to work on more such projects. Thank you!
