

Bellabeat Case Study

Andrew Davis

2023-05-03

Introduction

Hello! My name is Andrew Davis. For my Google Analytics course, I was tasked with completing a case study and recommend a marketing strategy within Bellabeat: a high-tech manufacturer of health-focused products for women

Scenario

You are a junior data analyst working on the marketing analyst team at Bellabeat, a high-tech manufacturer of health-focused products for women. Bellabeat is a successful small company, but they have the potential to become a larger player in the global smart device market. Urška Sršen, cofounder and Chief Creative Officer of Bellabeat, believes that analyzing smart device fitness data could help unlock new growth opportunities for the company. You have been asked to focus on one of Bellabeat's products and analyze smart device data to gain insight into how consumers are using their smart devices. The insights you discover will then help guide marketing strategy for the company. You will present your analysis to the Bellabeat executive team along with your high-level recommendations for Bellabeat's marketing strategy.

Characters and Products

- **Characters**
 - **Urška Sršen:** Bellabeat's cofounder and Chief Creative Officer
 - **Sando Mur:** Mathematician and Bellabeat's cofounder; key member of the Bellabeat executive team
 - **Bellabeat marketing analytics team:** A team of data analysts responsible for collecting, analyzing, and reporting data that helps guide Bellabeat's marketing strategy. You joined this team six months ago and have been busy learning about Bellabeat's mission and business goals — as well as how you, as a junior data analyst, can help Bellabeat achieve them.
- **Products**
 - **Bellabeat app:** The Bellabeat app provides users with health data related to their activity, sleep, stress, menstrual cycle, and mindfulness habits. This data can help users better understand their current habits and make healthy decisions. The Bellabeat app connects to their line of smart wellness products.
 - **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. The Time watch connects to the Bellabeat app to provide you with insights into your daily wellness.
 - **Spring:** This is a water bottle that tracks daily water intake using smart technology to ensure that you are appropriately hydrated throughout the day. The Spring bottle connects to the Bellabeat app to track your hydration levels.
 - **Bellabeat membership:** Bellabeat also offers a subscription-based membership program for users. Membership gives users 24/7 access to fully personalized guidance on nutrition, activity, sleep, health and beauty, and mindfulness based on their lifestyle and goals.

About the company

Urška Sršen and Sando Mur founded Bellabeat, a high-tech company that manufactures health-focused smart products. Sršen used her background as an artist to develop beautifully designed technology that informs and inspires women around the world. Collecting data on activity, sleep, stress, and reproductive health has allowed Bellabeat to empower women with knowledge about their own health and habits. Since it was founded in 2013, Bellabeat has grown rapidly and quickly positioned itself as a tech-driven wellness company for women.

By 2016, Bellabeat had opened offices around the world and launched multiple products. Bellabeat products became available through a growing number of online retailers in addition to their own e-commerce channel on their website. 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.

Sršen knows that an analysis of Bellabeat's available consumer data would reveal more opportunities for growth. She has asked 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, she would like high-level recommendations for how these trends can inform Bellabeat marketing strategy.

Ask

To find new growth opportunities for the company through analyzing smart device fitness data, focusing on one of Bellabeat's products to gain insight into how consumers are using their smart device. The final deliverable is presenting this analysis to the Bellabeat executive team along with high-level recommendations for the marketing strategy.

My hypothesis is that there are a significant percent of individuals who do not have any sleep data, who we would be able to market the Leaf or Time product to.

Prepare and Process

Loading Packages

For this I will be loading the following packages:

```
library(tidyverse)
library(janitor)
library(skimr)
library(lubridate)
```

Load the following data sources:

- Daily Activity Data
- Daily Calories
- Sleep Data
- Weight Log

```
dailyActivity <- read_csv("fitbit_data/dailyActivity_merged.csv")
dailyCalories <- read_csv("fitbit_data/dailyCalories_merged.csv")
sleepDay <- read_csv("fitbit_data/sleepDay_merged.csv")
weightLogInfo <- read_csv("fitbit_data/weightLogInfo_merged.csv")
```

Calculate total minutes for:

- Daily Activity
- Unaccounted Time not logged from the device
- Daily Usage, based on Unaccounted Time

```
dailyActivity <- dailyActivity %>%
  mutate(TotalTime = VeryActiveMinutes + FairlyActiveMinutes + LightlyActiveMinutes + SedentaryMinutes) %>%
  mutate(UnaccountedTime = 1440 - TotalTime) %>%
  mutate(DailyUsage = (1440 - UnaccountedTime) / 1440) %>%
  mutate(DayofWeek = wday(mdy(ActivityDate), label = TRUE)) %>%
  rename(ActivityDay = ActivityDate)
```

Reformat Sleep Data Date to allow for joining and calculate total hours for each measurement.

```
sleepDay <- sleepDay %>%
  distinct(Id, SleepDay, .keep_all = TRUE) %>%
  mutate(ActivityDay = gsub("0(\\d\\D)", "\\1", strftime(as.POSIXct(SleepDay, format = '%m/%d/%Y %H:%M:%S'), format = '%m/%d/%Y')) %>%
  mutate(TotalHoursInBed = round(TotalTimeInBed / 60, 1)) %>%
  mutate(TotalHoursAsleep = round(TotalMinutesAsleep / 60, 1)) %>%
  mutate(TotalHoursAwakeInBed = TotalHoursInBed - TotalHoursAsleep)

sleepDay <- sleepDay %>%
  select(-c(SleepDay, TotalSleepRecords))
```

Remove duplicate dates from Weight Log, reformat date to allow for joining, and filter out unnecessary columns.

```
weightLogInfo <- weightLogInfo %>%
  distinct(Id, Date, .keep_all = TRUE) %>%
  mutate(ActivityDay = gsub("0(\\d\\D)", "\\1", strftime(as.POSIXct(Date, format = '%m/%d/%Y %H:%M:%S'), format = '%m/%d/%Y')) %>%
  weightLogInfo <- weightLogInfo %>%
  select(-c(Date, LogId))
```

Join all 4 tables

```
complete_data <- merge(dailyActivity, dailyCalories, by = c("Id", "ActivityDay"))
complete_data <- merge(complete_data, sleepDay, by = c("Id", "ActivityDay"), all.x = TRUE)
complete_data <- merge(complete_data, weightLogInfo, by = c("Id", "ActivityDay"), all.x = TRUE)
```

Preview the completed data

```
skim(complete_data)
```

Data summary

Name	complete_data
Number of rows	940
Number of columns	30
Column type frequency:	
character	1

factor	1
logical	1
numeric	27

Group variables	None
-----------------	------

Variable type: character

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
ActivityDay	0	1	8	9	0	31	0

Variable type: factor

skim_variable	n_missing	complete_rate	ordered	n_unique	top_counts
DayofWeek	0	1	TRUE	7	Tue: 152, Wed: 150, Thu: 147, Fri: 126

Variable type: logical

skim_variable	n_missing	complete_rate	mean	count
IsManualReport	873	0.07	0.61	TRU: 41, FAL: 26

Variable type: numeric

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	
Id	0	1.00	4.855407e+09	2.424805e+09	1.50396e+09	2.320127e+09	4.445115e+09	6.962181e+09	8.877689
TotalSteps	0	1.00	7.637910e+03	5.087150e+03	0.00000e+00	3.789750e+03	7.405500e+03	1.072700e+04	3.601900
TotalDistance	0	1.00	5.490000e+00	3.920000e+00	0.00000e+00	2.620000e+00	5.240000e+00	7.710000e+00	2.803000
TrackerDistance	0	1.00	5.480000e+00	3.910000e+00	0.00000e+00	2.620000e+00	5.240000e+00	7.710000e+00	2.803000
LoggedActivitiesDistance	0	1.00	1.100000e-01	6.200000e-01	0.00000e+00	0.000000e+00	0.000000e+00	0.000000e+00	4.940000
VeryActiveDistance	0	1.00	1.500000e+00	2.660000e+00	0.00000e+00	0.000000e+00	2.100000e-01	2.050000e+00	2.192000
ModeratelyActiveDistance	0	1.00	5.700000e-01	8.800000e-01	0.00000e+00	0.000000e+00	2.400000e-01	8.000000e-01	6.480000
LightActiveDistance	0	1.00	3.340000e+00	2.040000e+00	0.00000e+00	1.950000e+00	3.360000e+00	4.780000e+00	1.071000
SedentaryActiveDistance	0	1.00	0.000000e+00	1.000000e-02	0.00000e+00	0.000000e+00	0.000000e+00	0.000000e+00	1.100000
VeryActiveMinutes	0	1.00	2.116000e+01	3.284000e+01	0.00000e+00	0.000000e+00	4.000000e+00	3.200000e+01	2.100000
FairlyActiveMinutes	0	1.00	1.356000e+01	1.999000e+01	0.00000e+00	0.000000e+00	6.000000e+00	1.900000e+01	1.430000
LightlyActiveMinutes	0	1.00	1.928100e+02	1.091700e+02	0.00000e+00	1.270000e+02	1.990000e+02	2.640000e+02	5.180000
SedentaryMinutes	0	1.00	9.912100e+02	3.012700e+02	0.00000e+00	7.297500e+02	1.057500e+03	1.229500e+03	1.440000
Calories.x	0	1.00	2.303610e+03	7.181700e+02	0.00000e+00	1.828500e+03	2.134000e+03	2.793250e+03	4.900000
TotalTime	0	1.00	1.218750e+03	2.659300e+02	2.00000e+00	9.897500e+02	1.440000e+03	1.440000e+03	1.440000
UnaccountedTime	0	1.00	2.212500e+02	2.659300e+02	0.00000e+00	0.000000e+00	0.000000e+00	4.502500e+02	1.438000
DailyUsage	0	1.00	8.500000e-01	1.800000e-01	0.00000e+00	6.900000e-01	1.000000e+00	1.000000e+00	1.000000
Calories.y	0	1.00	2.303610e+03	7.181700e+02	0.00000e+00	1.828500e+03	2.134000e+03	2.793250e+03	4.900000
TotalMinutesAsleep	530	0.44	4.191700e+02	1.186400e+02	5.80000e+01	3.610000e+02	4.325000e+02	4.900000e+02	7.960000
TotalTimeInBed	530	0.44	4.584800e+02	1.274600e+02	6.10000e+01	4.037500e+02	4.630000e+02	5.260000e+02	9.610000
TotalHoursInBed	530	0.44	7.640000e+00	2.120000e+00	1.00000e+00	6.730000e+00	7.700000e+00	8.800000e+00	1.600000
TotalHoursAsleep	530	0.44	6.990000e+00	1.980000e+00	1.00000e+00	6.000000e+00	7.200000e+00	8.200000e+00	1.330000
TotalHoursAwakeInBed	530	0.44	6.500000e-01	7.800000e-01	0.00000e+00	3.000000e-01	4.000000e-01	7.000000e-01	6.200000
WeightKg	873	0.07	7.204000e+01	1.392000e+01	5.26000e+01	6.140000e+01	6.250000e+01	8.505000e+01	1.335000
WeightPounds	873	0.07	1.588100e+02	3.070000e+01	1.15960e+02	1.353600e+02	1.377900e+02	1.875000e+02	2.943200
Fat	938	0.00	2.350000e+01	2.120000e+00	2.20000e+01	2.275000e+01	2.350000e+01	2.425000e+01	2.500000
BMI	873	0.07	2.519000e+01	3.070000e+00	2.14500e+01	2.396000e+01	2.439000e+01	2.556000e+01	4.754000

Analyze

Going back to my hypothesis: *are there a significant percent of individuals who do not have any sleep data, who we would be able to market the Leaf or Time product to?*

Just looking at the skim data, there is only a 43.6% completion rate for sleep data. I will need to further analyze to see if certain individuals do not have any sleep data.

```
grouped_sleepdata <- complete_data %>%
  group_by(Id) %>%
  summarise(AverageTotalHoursAsleep = mean(TotalHoursAsleep, na.rm = TRUE)) %>%
  mutate(Sleep_data = case_when(AverageTotalHoursAsleep == 'NaN' ~ 'No', .default = 'Yes'))

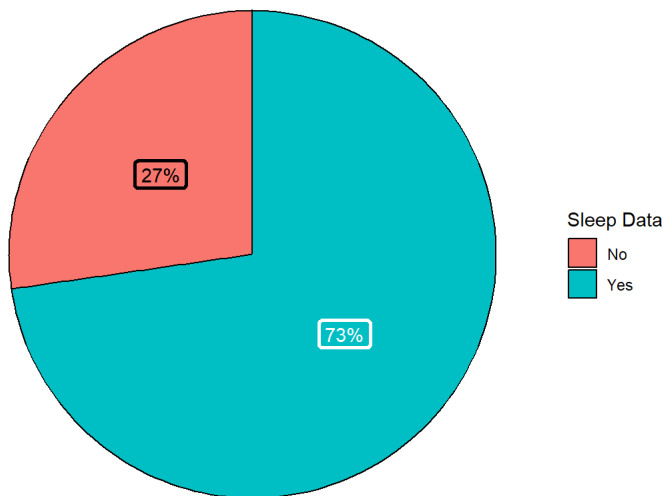
grouped_sleepdata <- grouped_sleepdata %>%
  group_by(Sleep_data) %>%
  tally() %>%
  ungroup() %>%
  mutate(perc = `n` / sum(`n`)) %>%
  arrange(perc) %>%
  mutate(labels = scales::percent(perc))
```

Convert into a sleep chart

```
sleepchart <- grouped_sleepdata %>%
  ggplot(mapping = aes(x = "", y = perc, fill = Sleep_data)) +
  geom_col(color = "black") +
  geom_label(aes(label = labels), color = c(1, "white"),
    position = position_stack(vjust = 0.5),
    label.size = 1,
    show.legend = FALSE) +
  guides(fill = guide_legend(title = "Sleep Data")) +
  coord_polar(theta = "y") +
  theme_void() +
  labs(title = 'Users with Sleep Data')
```

Share

Users with Sleep Data



Act

27% of users do not have any sleep data, which would allow up to market the Leaf and Time products, but would also allow to provide better insight into their current habits and make healthy decisions through the Bellabeat app.