

Bellabeat_Case Study

Zaryn Ooi

7/29/2021

Phase 1 : Ask

About the Company

Bellabeat is a high tech company that manufactures health-focused smart products. Since it was founded in 2013, Bellabeat has grown rapidly and positioned itself as a tech-driven wellness company for women.

Key Stakeholders

Primary stakeholders:

- Urška Sršen, cofounder and Chief Creative Officer
- Sando Mur, Mathematician and Bellabeat's cofounder

Secondary stakeholders:

- Bellabeat marketing analytics team

Business task

To identify trends in smart devices usage to help Bellabeat make marketing strategy improvement.

Phase 2 : Prepare

Data Source

Fitbit Fitness Tracker Data (CC0: Public Domain, dataset made available through Mobius)

<https://www.kaggle.com/arashnic/fitbit>

This data set contains personal fitness tracker from thirty Fitbit users. Thirty eligible Fitbit users consented to the submission of personal tracker data, including minute-level output for physical activity, heart rate, and sleep monitoring. It includes information about daily activity, steps, and heart rate that can be used to explore users' habits.

Bias/ Credibility (ROCCC Analysis)

- Reliability = YES
- Original = YES
- Comprehensive = YES
- Current = NO
- Cited = YES

Limitations

- Data is not current as it was generated between 03.12.2016 - 05.12.2016.
- Sample size of 30 is too small to represent the population.
- The data does not show the demographic information of the respondent. Hence, the data cannot ensure that the respondent is a representative of female users, given that the target audience for Bellabeat is female.
- The data were collected by different types of Fitbit trackers and individual tracking behaviors / preferences. However, the data does not indicate which devices were used.

Data Selection

For this analysis, i have selected 2 datasets that provided daily activity and the sleep time of the user because there may be a correlation between the sleep behaviours and the overall activity of the users, which are helpful for me to identify trends in smart devices usage.

Files selected:

- dailyActivity__merged.csv
- sleepDay__merged.csv

Phase 3 : Process

Install and Load Packages

I have installed necessary packages and datasets for analysis in R.

```
install.packages('tidyverse')

## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/4.1'
## (as 'lib' is unspecified)
install.packages('dplyr')

## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/4.1'
## (as 'lib' is unspecified)
install.packages('ggplot2')

## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/4.1'
## (as 'lib' is unspecified)
install.packages('lubridate')

## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/4.1'
## (as 'lib' is unspecified)
install.packages('skimr')

## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/4.1'
## (as 'lib' is unspecified)
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --
## v ggplot2 3.3.5      v purrr   0.3.4
## v tibble  3.1.3      v dplyr  1.0.7
```

```
## v tidyr 1.1.3 v stringr 1.4.0
## v readr 2.0.0 v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag() masks stats::lag()

library(dplyr)
library(ggplot2)
library(lubridate)

##
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':
##
## date, intersect, setdiff, union

library(skimr)
```

Importing data

I will create a dataframe named 'daily_activity' for the daily activity data and 'sleep_day' for the sleep data.

```
daily_activity <- read.csv("dailyActivity_merged.csv")
```

Let's use the head() function to display the first 6 rows present in the 'daily_activity' data frame.

```
head(daily_activity)
```

```
##           Id ActivityDate TotalSteps TotalDistance TrackerDistance
## 1 1503960366 4/12/2016      13162           8.50           8.50
## 2 1503960366 4/13/2016      10735           6.97           6.97
## 3 1503960366 4/14/2016      10460           6.74           6.74
## 4 1503960366 4/15/2016       9762           6.28           6.28
## 5 1503960366 4/16/2016      12669           8.16           8.16
## 6 1503960366 4/17/2016       9705           6.48           6.48
##   LoggedActivitiesDistance VeryActiveDistance ModeratelyActiveDistance
## 1                        0                1.88                    0.55
## 2                        0                1.57                    0.69
## 3                        0                2.44                    0.40
## 4                        0                2.14                    1.26
## 5                        0                2.71                    0.41
## 6                        0                3.19                    0.78
##   LightActiveDistance SedentaryActiveDistance VeryActiveMinutes
## 1                6.06                    0                25
## 2                4.71                    0                21
## 3                3.91                    0                30
## 4                2.83                    0                29
## 5                5.04                    0                36
## 6                2.51                    0                38
##   FairlyActiveMinutes LightlyActiveMinutes SedentaryMinutes Calories
## 1                13                328                728      1985
## 2                19                217                776      1797
## 3                11                181                1218      1776
## 4                34                209                726      1745
## 5                10                221                773      1863
## 6                20                164                539      1728
```

```
sleep_day <- read.csv("sleepDay_merged.csv")
```

```
head(sleep_day)
```

```
##           Id           SleepDay TotalSleepRecords TotalMinutesAsleep
## 1 1503960366 4/12/2016 12:00:00 AM                1                327
## 2 1503960366 4/13/2016 12:00:00 AM                2                384
## 3 1503960366 4/15/2016 12:00:00 AM                1                412
## 4 1503960366 4/16/2016 12:00:00 AM                2                340
## 5 1503960366 4/17/2016 12:00:00 AM                1                700
## 6 1503960366 4/19/2016 12:00:00 AM                1                304
## TotalTimeInBed
## 1                346
## 2                407
## 3                442
## 4                367
## 5                712
## 6                320
```

Explore the imported data

```
head(daily_activity)
```

```
##           Id ActivityDate TotalSteps TotalDistance TrackerDistance
## 1 1503960366 4/12/2016      13162          8.50          8.50
## 2 1503960366 4/13/2016      10735          6.97          6.97
## 3 1503960366 4/14/2016      10460          6.74          6.74
## 4 1503960366 4/15/2016       9762          6.28          6.28
## 5 1503960366 4/16/2016      12669          8.16          8.16
## 6 1503960366 4/17/2016       9705          6.48          6.48
## LoggedActivitiesDistance VeryActiveDistance ModeratelyActiveDistance
## 1                0                1.88                0.55
## 2                0                1.57                0.69
## 3                0                2.44                0.40
## 4                0                2.14                1.26
## 5                0                2.71                0.41
## 6                0                3.19                0.78
## LightActiveDistance SedentaryActiveDistance VeryActiveMinutes
## 1                6.06                0                25
## 2                4.71                0                21
## 3                3.91                0                30
## 4                2.83                0                29
## 5                5.04                0                36
## 6                2.51                0                38
## FairlyActiveMinutes LightlyActiveMinutes SedentaryMinutes Calories
## 1                13                328                728      1985
## 2                19                217                776      1797
## 3                11                181               1218      1776
## 4                34                209                726      1745
## 5                10                221                773      1863
## 6                20                164                539      1728
```

```
head(sleep_day)
```

```
##           Id           SleepDay TotalSleepRecords TotalMinutesAsleep
```

```
## 1 1503960366 4/12/2016 12:00:00 AM 1 327
## 2 1503960366 4/13/2016 12:00:00 AM 2 384
## 3 1503960366 4/15/2016 12:00:00 AM 1 412
## 4 1503960366 4/16/2016 12:00:00 AM 2 340
## 5 1503960366 4/17/2016 12:00:00 AM 1 700
## 6 1503960366 4/19/2016 12:00:00 AM 1 304
## TotalTimeInBed
## 1 346
## 2 407
## 3 442
## 4 367
## 5 712
## 6 320
```

```
skim_without_charts(daily_activity)
```

Table 1: Data summary

| | |
|------------------------|----------------|
| Name | daily_activity |
| Number of rows | 940 |
| Number of columns | 15 |
| Column type frequency: | |
| character | 1 |
| numeric | 14 |
| Group variables | None |

Variable type: character

| skim_variable | n_missing | complete_rate | min | max | empty | n_unique | whitespace |
|---------------|-----------|---------------|-----|-----|-------|----------|------------|
| ActivityDate | 0 | 1 | 8 | 9 | 0 | 31 | 0 |

Variable type: numeric

| skim_variable | n_missing | complete_rate | mean | sd | p0 | p25 | p50 | p75 | p100 |
|--------------------------|-----------|---------------|---------------|------------|---------|---------------|------------|-------------|------------|
| Id | 0 | 1 | 4.855407e-200 | 24805e-150 | 3960366 | 320127e-409 | 45115e-609 | 62181e-808 | 77689e+09 |
| TotalSteps | 0 | 1 | 7.637910e-503 | 87150e+03 | 0 | 3.789750e-704 | 5500e-103 | 72700e-306 | 1900e+04 |
| TotalDistance | 0 | 1 | 5.490000e-302 | 20000e+00 | 0 | 2.620000e-502 | 40000e-707 | 10000e-208 | 3000e+01 |
| TrackerDistance | 0 | 1 | 5.480000e-301 | 10000e+00 | 0 | 2.620000e-502 | 40000e-707 | 10000e-208 | 3000e+01 |
| LoggedActivitiesDistance | 0 | 1 | 1.100000e-6 | 20000e- | 0 | 0.000000e-900 | 00000e-900 | 00000e-400 | 40000e+00 |
| VeryActiveDistance | 0 | 1 | 1.500000e-206 | 60000e+00 | 0 | 0.000000e-200 | 00000e-2 | 050000e-201 | 92000e+01 |
| ModeratelyActiveDistance | 0 | 1 | 5.700000e-204 | 80000e- | 0 | 0.000000e-200 | 00000e-8 | 00000e-6 | 480000e+00 |
| LightActiveDistance | 0 | 1 | 3.340000e-200 | 40000e+00 | 0 | 1.950000e-306 | 60000e-407 | 80000e-100 | 71000e+01 |
| SedentaryActiveDistance | 0 | 1 | 0.000000e-400 | 00000e- | 0 | 0.000000e-900 | 00000e-900 | 00000e-400 | 00000e- |
| VeryActiveMinutes | 0 | 1 | 2.116000e-302 | 84000e+01 | 0 | 0.000000e-400 | 00000e-302 | 00000e-201 | 00000e+02 |

| skim_variable | n_missing | complete_rate | mean | sd | p0 | p25 | p50 | p75 | p100 |
|----------------------|-----------|---------------|--------------|--------------|----|--------------|--------------|--------------|--------------|
| FairlyActiveMinutes | 0 | 1 | 1.356000e+01 | 1.990000e+01 | 0 | 0.000000e+00 | 0.000000e+00 | 0.000000e+00 | 0.130000e+02 |
| LightlyActiveMinutes | 0 | 1 | 1.928100e+02 | 1.091700e+02 | 0 | 1.270000e+02 | 1.990000e+02 | 2.240000e+02 | 5.028000e+02 |
| SedentaryMinutes | 0 | 1 | 9.912100e+02 | 3.021270e+02 | 0 | 7.297500e+02 | 1.025750e+03 | 1.022950e+03 | 1.044000e+03 |
| Calories | 0 | 1 | 2.303610e+03 | 7.081700e+02 | 0 | 1.828500e+03 | 2.034000e+03 | 2.079325e+03 | 4.090000e+03 |

```
skim_without_charts(sleep_day)
```

Table 4: Data summary

| | |
|------------------------|-----------|
| Name | sleep_day |
| Number of rows | 413 |
| Number of columns | 5 |
| Column type frequency: | |
| character | 1 |
| numeric | 4 |
| Group variables | None |

Variable type: character

| skim_variable | n_missing | complete_rate | min | max | empty | n_unique | whitespace |
|---------------|-----------|---------------|-----|-----|-------|----------|------------|
| SleepDay | 0 | 1 | 20 | 21 | 0 | 31 | 0 |

Variable type: numeric

| skim_variable | n_missing | complete_rate | mean | sd | p0 | p25 | p50 | p75 | p100 |
|--------------------|-----------|---------------|--------------|--------------|---|-----|-----|-----|------|
| Id | 0 | 1 | 5.000979e+02 | 0.06036e+02 | 0.05039603666977333714702921684962181068792009665 | | | | |
| TotalSleepRecords | 0 | 1 | 1.120000e+01 | 0.50000e-01 | 1 | 1 | 1 | 1 | 3 |
| TotalMinutesAsleep | 0 | 1 | 4.194700e+02 | 1.218340e+02 | 58 | 361 | 433 | 490 | 796 |
| TotalTimeInBed | 0 | 1 | 4.586400e+02 | 1.227100e+02 | 61 | 403 | 463 | 526 | 961 |

Fixing Formatting

After exploring the datasets, I have noticed that there are some problems with the timestamp data. Before analysis, I will use the lubridate library's `mdy()` function to convert date strings as the date elements in `daily_activity` data are ordered as month, day and year.

```
daily_activity <-
  daily_activity %>%
  mutate(ActivityDate = mdy(ActivityDate))
```

Take a look and recheck the data.

```
head(daily_activity)
```

```
##           Id ActivityDate TotalSteps TotalDistance TrackerDistance
```

```
## 1 1503960366 2016-04-12 13162 8.50 8.50
## 2 1503960366 2016-04-13 10735 6.97 6.97
## 3 1503960366 2016-04-14 10460 6.74 6.74
## 4 1503960366 2016-04-15 9762 6.28 6.28
## 5 1503960366 2016-04-16 12669 8.16 8.16
## 6 1503960366 2016-04-17 9705 6.48 6.48
## LoggedActivitiesDistance VeryActiveDistance ModeratelyActiveDistance
## 1 0 1.88 0.55
## 2 0 1.57 0.69
## 3 0 2.44 0.40
## 4 0 2.14 1.26
## 5 0 2.71 0.41
## 6 0 3.19 0.78
## LightActiveDistance SedentaryActiveDistance VeryActiveMinutes
## 1 6.06 0 25
## 2 4.71 0 21
## 3 3.91 0 30
## 4 2.83 0 29
## 5 5.04 0 36
## 6 2.51 0 38
## FairlyActiveMinutes LightlyActiveMinutes SedentaryMinutes Calories
## 1 13 328 728 1985
## 2 19 217 776 1797
## 3 11 181 1218 1776
## 4 34 209 726 1745
## 5 10 221 773 1863
## 6 20 164 539 1728
```

I will use the `mdy_hms()` function to convert date strings as the date elements in `sleep_day` data are ordered as month, day, year, hour, minute, and second.

```
sleep_day <-
  sleep_day %>%
  mutate(SleepDay = mdy_hms(SleepDay))
```

let's take a look and recheck the data.

```
head(sleep_day)
```

```
##      Id  SleepDay TotalSleepRecords TotalMinutesAsleep TotalTimeInBed
## 1 1503960366 2016-04-12             1             327           346
## 2 1503960366 2016-04-13             2             384           407
## 3 1503960366 2016-04-15             1             412           442
## 4 1503960366 2016-04-16             2             340           367
## 5 1503960366 2016-04-17             1             700           712
## 6 1503960366 2016-04-19             1             304           320
```

Next, I will rename both 'ActivityDate' and 'SleepDay' column to 'Date'.

```
daily_activity <-
  daily_activity %>%
  rename(Date = ActivityDate)

sleep_day <-
  sleep_day %>%
  rename(Date = SleepDay)
```

Now that i have finsihed formatting, I can start exploring the data.

Summarize the data

I will use `n_distinct()` function to find out the number of respondents in each dataset.

```
n_distinct(daily_activity$Id)
```

```
## [1] 33
```

```
n_distinct(sleep_day$Id)
```

```
## [1] 24
```

There are 33 participants in the `daily_activity` dataset, and only 24 in the `sleep-day` dataset. This shows that not all users have sleep data.

The, I will use the `summary()` function to view the summary statistics of each dataset.

Summary Statistics of `daily_activity` Data:

```
daily_activity %>%
```

```
  select(TotalSteps,  
         TotalDistance,  
         SedentaryMinutes,  
         Calories) %>%
```

```
  summary()
```

```
##   TotalSteps   TotalDistance   SedentaryMinutes   Calories  
##   Min.      :    0   Min.      : 0.000   Min.      :    0.0   Min.      :    0  
##   1st Qu.: 3790   1st Qu.: 2.620   1st Qu.: 729.8   1st Qu.:1828  
##   Median : 7406   Median : 5.245   Median :1057.5   Median :2134  
##   Mean    : 7638   Mean    : 5.490   Mean     : 991.2   Mean     :2304  
##   3rd Qu.:10727   3rd Qu.: 7.713   3rd Qu.:1229.5   3rd Qu.:2793  
##   Max.    :36019   Max.    :28.030   Max.     :1440.0   Max.     :4900
```

Based on the summary, I have discovered that the average total steps of the respondents is 7638 per day, the average total distance is 5.490, and the average sedentary time is 991 minutes (16 hours). The goal of 10,000 steps is the recommended daily step target for healthy adults to achieve health benefits.

Summary Statistics of `sleep_day` Data:

```
sleep_day %>%
```

```
  select(TotalSleepRecords,  
         TotalMinutesAsleep,  
         TotalTimeInBed) %>%
```

```
  summary()
```

```
##   TotalSleepRecords   TotalMinutesAsleep   TotalTimeInBed  
##   Min.      :1.000     Min.      : 58.0     Min.      : 61.0  
##   1st Qu.:1.000     1st Qu.:361.0     1st Qu.:403.0  
##   Median :1.000     Median :433.0     Median :463.0  
##   Mean    :1.119     Mean    :419.5     Mean     :458.6  
##   3rd Qu.:1.000     3rd Qu.:490.0     3rd Qu.:526.0  
##   Max.    :3.000     Max.    :796.0     Max.     :961.0
```

From this summary, the average time asleep of the respondents is 419.5 minutes (7 hours). National Sleep Foundation guidelines advise that healthy adults need between 7 and 9 hours of sleep per night.

Merged the data

To identify the relationship between both datasets, I need to merge(inner join) both datasets using `inner_join()` function by “Id” and “Date”.

```
combined_data <- inner_join(daily_activity, sleep_day, by = c("Id", "Date"))
```

```
n_distinct(combined_data$Id)
```

```
## [1] 24
```

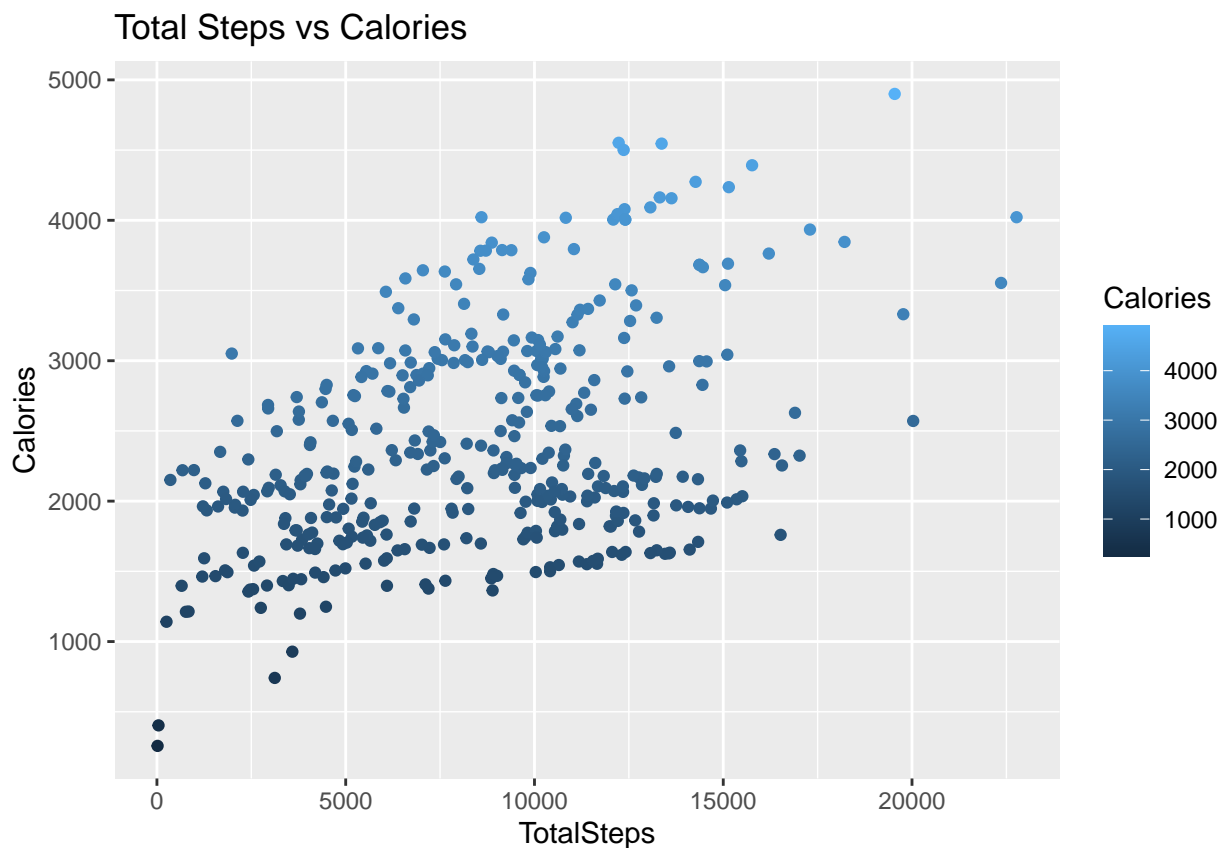
Now that I have merged the two datasets, there will only be 24 respondents.

Phase 4 & Phase5 : Analyze & Share

Visualization 1 - Total Steps vs Calories

First, I would like to identify the relationship between total steps taken in a day and the calories burn.

```
ggplot(data = combined_data) +  
  geom_point(mapping = aes(x = TotalSteps, y = Calories, color = Calories)) + labs(title = "Total Steps vs Calories")
```

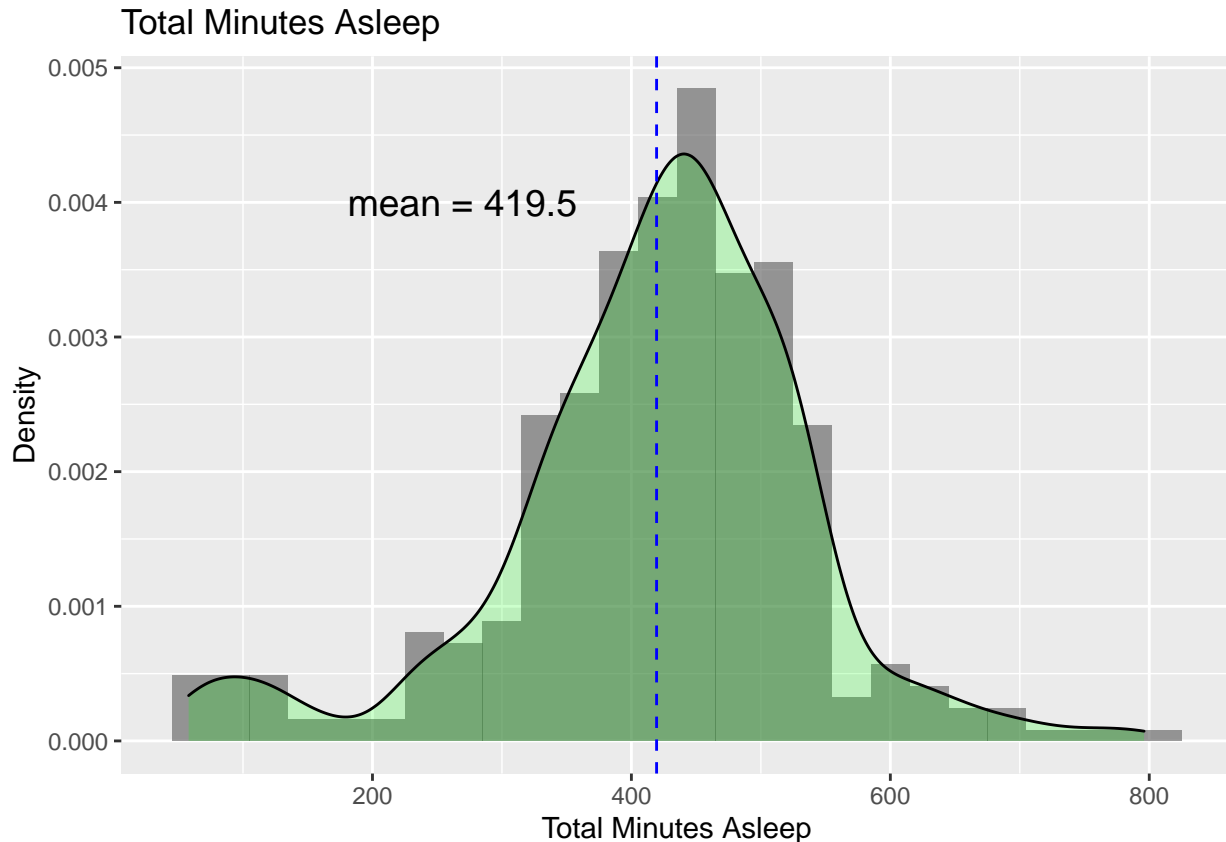


The visualization shows that there is a positive correlation between total steps taken in a day and the calories burn. The more steps we taken, the more calories we burn. The marketing analytics team can position this as a way to motivate users who plan to lose weight to walk more.

Visualization 2 - Total Minutes Asleep

Next, I would like to explore the sleep behaviors of the respondents.

```
ggplot(combined_data, aes(x=TotalMinutesAsleep)) +
  geom_histogram(aes(y=..density..), binwidth=30, alpha=0.6) +
  geom_density(alpha=0.2, fill="green") +
  geom_vline(aes(xintercept=mean(TotalMinutesAsleep)), color="blue", linetype="dashed") +
  annotate(geom="text", x= 419.5 - 150, y = 0.004, size = 5, label = "mean = 419.5") +
  labs(title = "Total Minutes Asleep", x= "Total Minutes Asleep", y="Density")
```



The histogram shows that most respondents sleep around 419.5 minutes, which is 7 hours a day. This means that most respondents get enough hours of sleep and do not have bad sleeping habits.

Visualization 3 - Time Awake in Bed

I will identify the time awake in bed to recheck if the respondents have bad sleeping habits. Before plotting the visualization, I will add a new column called “TotalAwakeInBed” with the `mutate()` function to the data frame to identify if there are respondents who have bad sleeping habits like insomnia or sleep in.

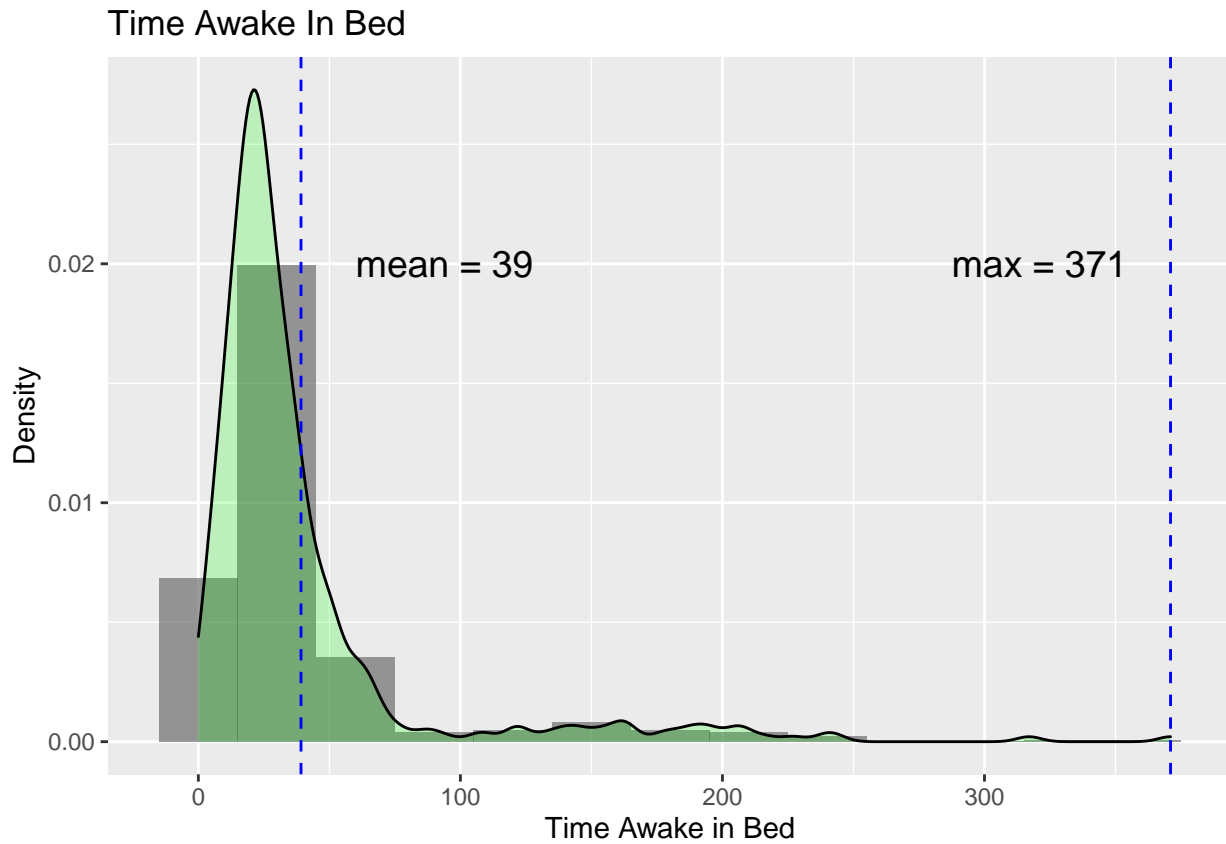
```
combined_data <-
  combined_data %>%
    mutate(TimeAwakeInBed = TotalTimeInBed - TotalMinutesAsleep)
```

Now, I can use the “TimeAwakeInBed” column to plot the visualization to examine if the respondents are spending too much time awake in bed. I will use the `mean()` function to find the average time spent awake in bed.

```
mean(combined_data$TimeAwakeInBed)
```

```
## [1] 39.17191
```

```
ggplot(combined_data, aes(x=TimeAwakeInBed)) +
  geom_histogram(aes(y=..density..), binwidth=30, alpha=0.6)+
  geom_density(alpha=0.2, fill="green") +
  geom_vline(aes(xintercept=mean(TimeAwakeInBed)), color="blue", linetype="dashed") +
  geom_vline(aes(xintercept=max(TimeAwakeInBed)), color="blue", linetype="dashed") +
  annotate(geom="text", x=39 + 55, y = 0.02, size = 5, label = "mean = 39") +
  annotate(geom="text", x=371 - 50, y = 0.02, size = 5, label = "max = 371") +
  labs(title = "Time Awake In Bed", x = "Time Awake in Bed", y="Density")
```

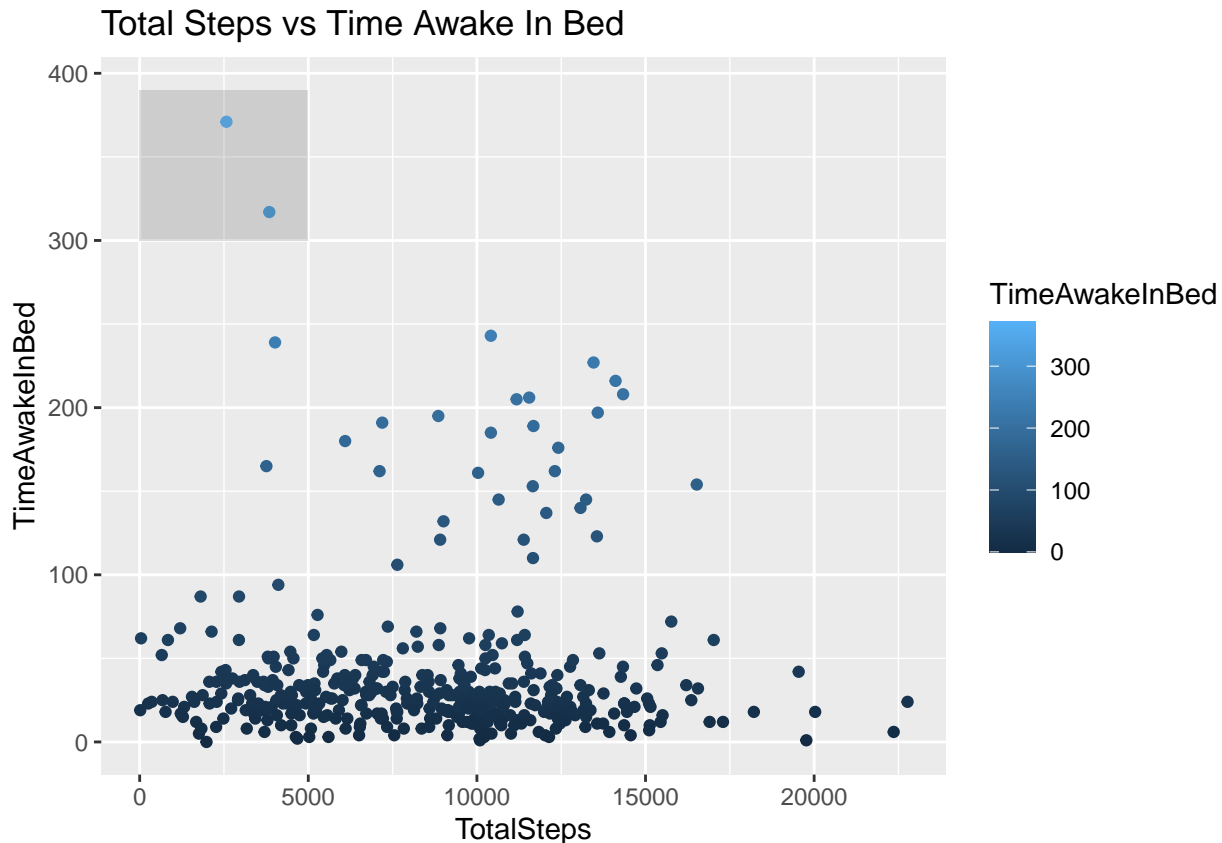


From the histogram, I discovered that the average time of respondents awake in bed is 39 minutes, which shows that most of the respondents do not have bad sleeping habits. Interesting to note, there is a small amount of respondents who spent 371 minutes time awake in bed, which signals that some of the respondents might suffered from insomnia.

Visualization 4 - Total Steps vs Time Awake In Bed

From this analysis, I will construct a visualization to find out if there is a relationship between activity and the sleep quality. I will use “TotalSteps” data to represent the activity and “TimeAwakeInBed” data to represent the sleep quality.

```
ggplot(data = combined_data) +
  geom_point(mapping = aes(x = TotalSteps, y = TimeAwakeInBed, color = TimeAwakeInBed)) + annotate("r
```



By looking at this scatterplot, we're able to identify that users who suffered from insomnia are those who have taken less than 5000 steps a day. We're also able to see that users who took more than 15000 steps do not spend more than 100 minutes awake in bed, which is less than an hour. This further suggests that these users do not develop bad sleeping habits.

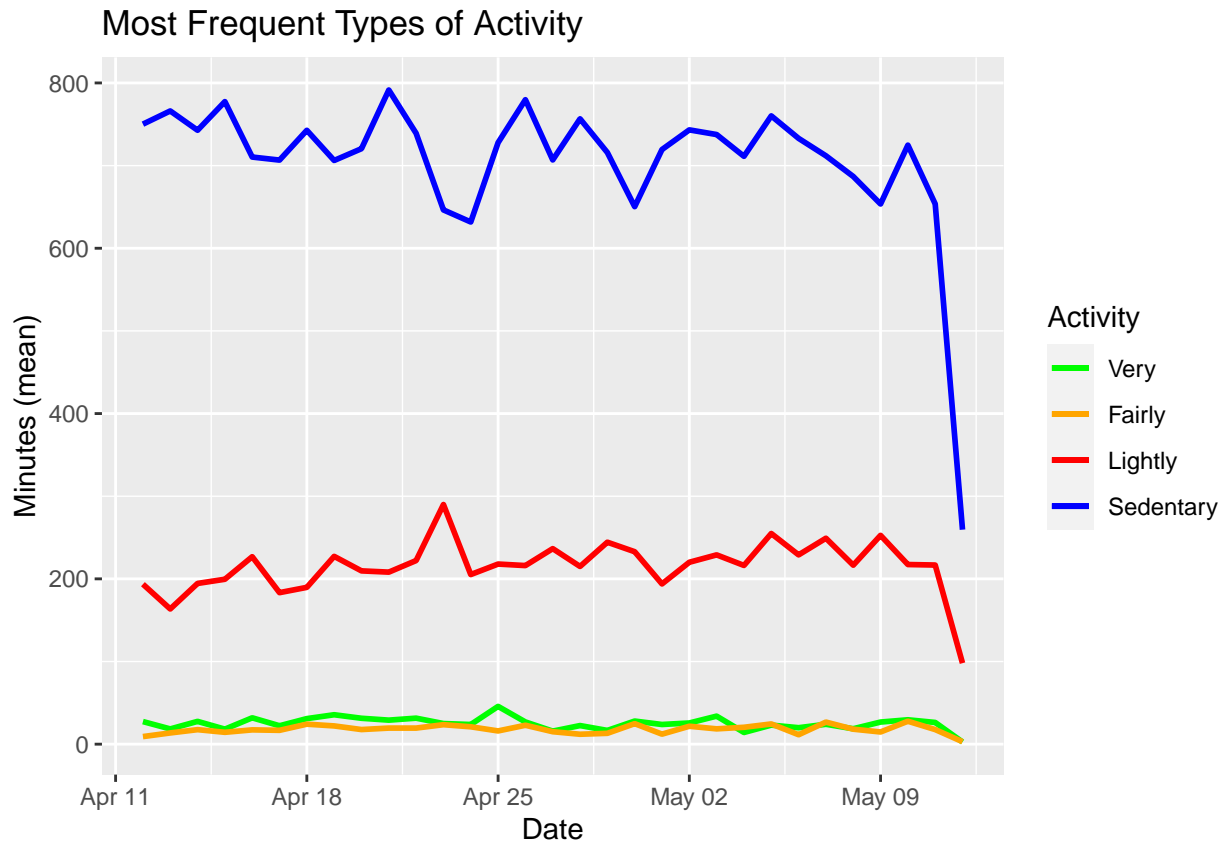
In short, daily step count influences sleep quality, and that higher daily step count leads to better sleep quality.

Visualization 5 - Most Frequent Types of Activity

I will construct a data frame and use the `mean()` function to calculate the mean for each types of Active Minutes to find out the most frequent types of activity among the respondents.

```
TypesofActivityByDate <-
  combined_data %>%
    group_by(Date) %>%
    summarize(MeanVeryActiveMinutes = mean(VeryActiveMinutes),
              MeanFairlyActiveMinutes = mean(FairlyActiveMinutes),
              MeanLightlyActiveMinutes = mean(LightlyActiveMinutes),
              MeanSedentaryMinutes = mean(SedentaryMinutes))

colors <- c("Very" = "green", "Fairly" = "orange", "Lightly" = "red", "Sedentary" = "blue")
ggplot(data = TypesofActivityByDate, aes(x = Date)) +
  geom_line(aes(y = MeanVeryActiveMinutes, color = "Very"), size = 1) +
  geom_line(aes(y = MeanFairlyActiveMinutes, color="Fairly"), size = 1) + geom_line(aes(y = MeanLightlyActiveMinutes, color="Lightly"), size = 1) +
  geom_line(aes(y = MeanSedentaryMinutes, color="Sedentary"), size = 1) +
  labs(title = "Most Frequent Types of Activity", x = "Date", y = "Minutes (mean)", color = "Activity") +
  scale_color_manual(values = colors)
```



By looking at this graph, we can see that the blue line appears to be top among the other lines. This means that sedentary activity is the most frequent type of activity carried out by the users, followed by 'Lightly' activity, 'Very' activity, while the least activity type is 'Fairly' activity.

Given that "Sedentary Activity" is the most frequent types of activity carried out by the users, this means that most of the users are working adults who spent long hours sitting in front of the computer/ meeting (Sedentary). And they also spent some time doing some house chores after work (Lightly).

Phase 6 : Act

I have gathered 3 trends from my analysis:

3 Trends in Smart Device Usage

1) Target Audience By analysing the most frequent types of activity as well as the sleep behavior data, I am able to identify how does the target audience looks like. This can help Bellabeat identify the characteristic of a typical consumer.

Target Audience of Bellabeat:

- Female working adults who spent most of the time doing sedentary activities, and rarely exercises.
- Have a relatively good amount of sleep but may spend some time awake in bed.

2) Sleep Behavior Analysis Given the fact that out of 33 respondents, only 24 of them record sleep data, this means that not every user will wear Bellabeat devices while sleeping.

Reason: • Do not know that Bellabeat app can track sleep data • Do not like wearing Bellabeat devices while sleeping • Do not know the importance of tracking sleep data

3) Daily steps count improves overall well being. By analysing the “TotalSteps” data and the “Time Spent Awake in Bed” data, I found out that higher daily step count leads to better sleep quality.

By analysing the daily step data and the calorie data, I’ve also discovered that higher daily step count leads to more calories burn.

Recommendations for Bellabeat App

After understanding the trends that I have discovered, Bellabeat should start:

1) Target potential customer Bellabeat can consider to target the potential customer. This can be done by crafting marketing messages in marketing campaigns/ Ads campaign to target people who have a similar activity pattern.

2) Utilize the identified trend to improve user experience The next strategy that Bellabeat could bring into effective action, is to utilise the trends that I’ve discovered to improve the user experience of their Bellabeat App.

Bellebeat can do this by first tracking the personal metrics of the users. For example, based on the personal metrics, Bellebeat identified that user A, Jamie is someone who suffer from insomnia, Bellebeat can send notifications to Jamie to remind her to go for a run, go to gym or to increase her frequency of higher intensive activities in order to improve her sleep quality. Additionally, these app notifications should combine with health knowledge such higher daily step counts leads to better sleep quality in order to motivate Jamie.

For users who wants to lose weight, the app can send them notifications to remind these users to do more exercise to burn their calories.

For users who do not record sleep data, Bellabeat App should send notifications to users to remind them to wear Bellabeat devices while sleeping in order to track their sleep data, combined with knowledge of the importance of tracking sleeping data. In addition, Bellabeat can allow users who used Bellabeat app to track their sleep data to earn and accrue reward points, which are then redeemable from wide range of gifts.