

Bellabeat Case Study

Rahul Sakharkar

2022-11-02



About the Company

Bellabeat is a high-tech company that manufactures health-focused smart products. Founded by Urška Sršen and Sando Mur, Bellabeat aims to empower women with knowledge about their own health and habits. Website- <http://bellabeat.com>

Ask

Sršen believes an analysis of Bellabeat's available consumer data would reveal more opportunities for growth. Sršen asks you to analyze smart device usage data in order to gain insight into how consumers use non-Bellabeat smart devices. Using this information, she would like high-level recommendations for how these trends can inform Bellabeat marketing strategy.

Business Task

To analyse smart device's usage data - uncover trends and patterns and to find how can bellabeat's marketing strategy be improved and how its customer's can benefit from it.

Cleaning and Processing Data

Cleaning of the data was done using spreadsheets. The following operations were performed -

- Removed Duplicate Rows
- Removed whitespaces using trim function
- Splitting Day_Date combined coloumns into separate Day and Date columns using split function
- Removed inconsistencies

Loading Packages

```
library(tidyverse)
library(lubridate)
library(ggplot2)
library(dplyr)
```

Importing Datasets

```
activity<-read_csv('dailyActivity_merged.csv')
sleep<-read_csv('sleepDay_merged.csv')
weight<-read_csv('weightLogInfo_merged.csv')
```

Viewing datasets

```
View(activity)
View(sleep)
View(weight)
```

Finding distinct observations in each dataset

- n_distinct(activity\$id)
- n_distinct(sleep\$id)
- n_distinct(weight\$id)

In Activity -

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

In sleep -

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

In weight -

```
## [1] 8
```

Summarizing datasets

```
activity_sleep= activity %>% inner_join(sleep,by=c('Id', 'Date'))
col1=c('Calories','TotalSteps','TotalMinutesAsleep')
summary(activity_sleep[col1])
```

Joining activity and sleep datasets by inner join

```
##      Calories      TotalSteps      TotalMinutesAsleep
##  Min.   : 257   Min.   : 17   Min.   : 58.0
##  1st Qu.:1850  1st Qu.: 5206  1st Qu.:361.0
##  Median :2220   Median : 8925  Median :433.0
##  Mean   :2398   Mean   : 8541  Mean   :419.5
##  3rd Qu.:2926  3rd Qu.:11393 3rd Qu.:490.0
##  Max.   :4900   Max.   :22770  Max.   :796.0
```

```
summary(weight['WeightKg'])
```

```
##      WeightKg
##  Min.   : 52.60
##  1st Qu.: 61.40
##  Median : 62.50
##  Mean   : 72.04
##  3rd Qu.: 85.05
##  Max.   :133.50
```

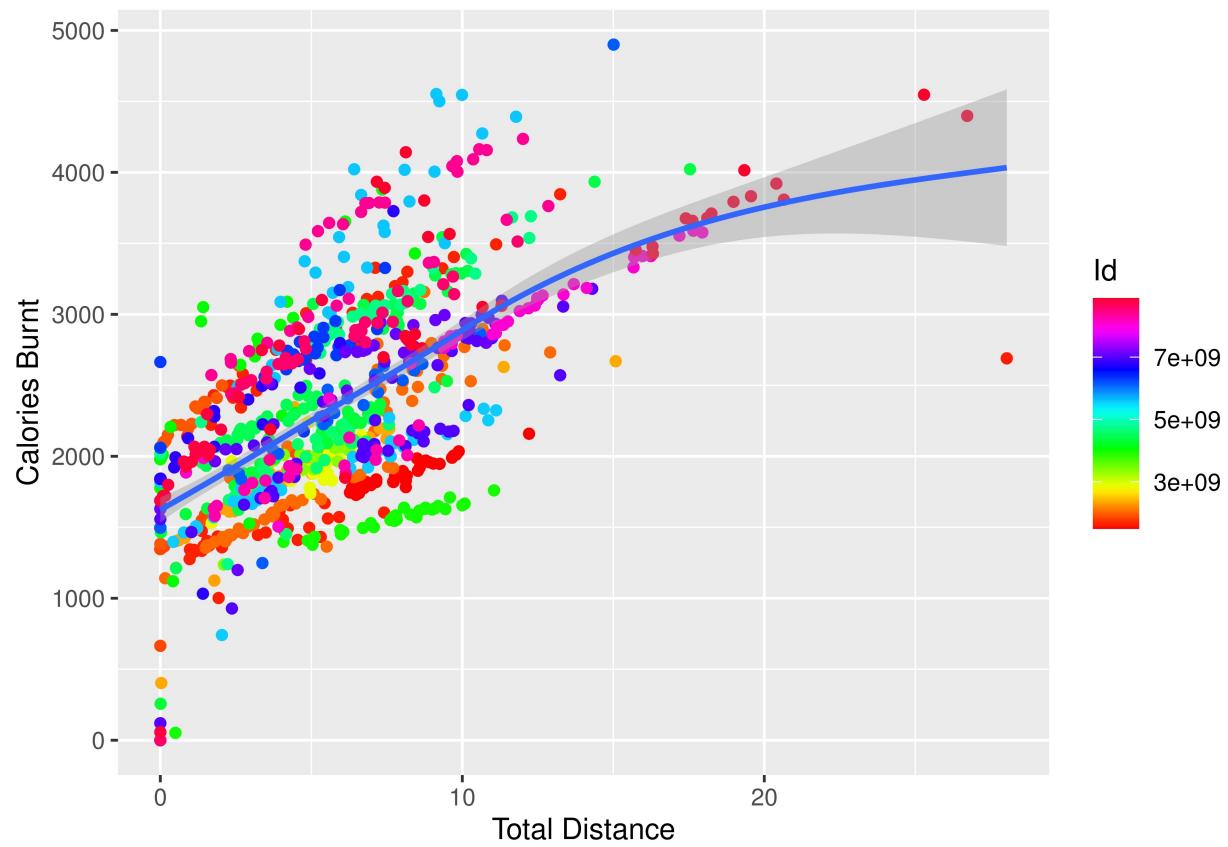
Important insights from the summaries

- Average Calories burnt: 2400
- Average number of steps are 8500 per day. Recommended number are 10,000 steps/day. So this should be improved
- Average number of hours slept by a person is 7
- Average weight of a person is 72 kg. This can be reduced for a much healthier lifestyle.

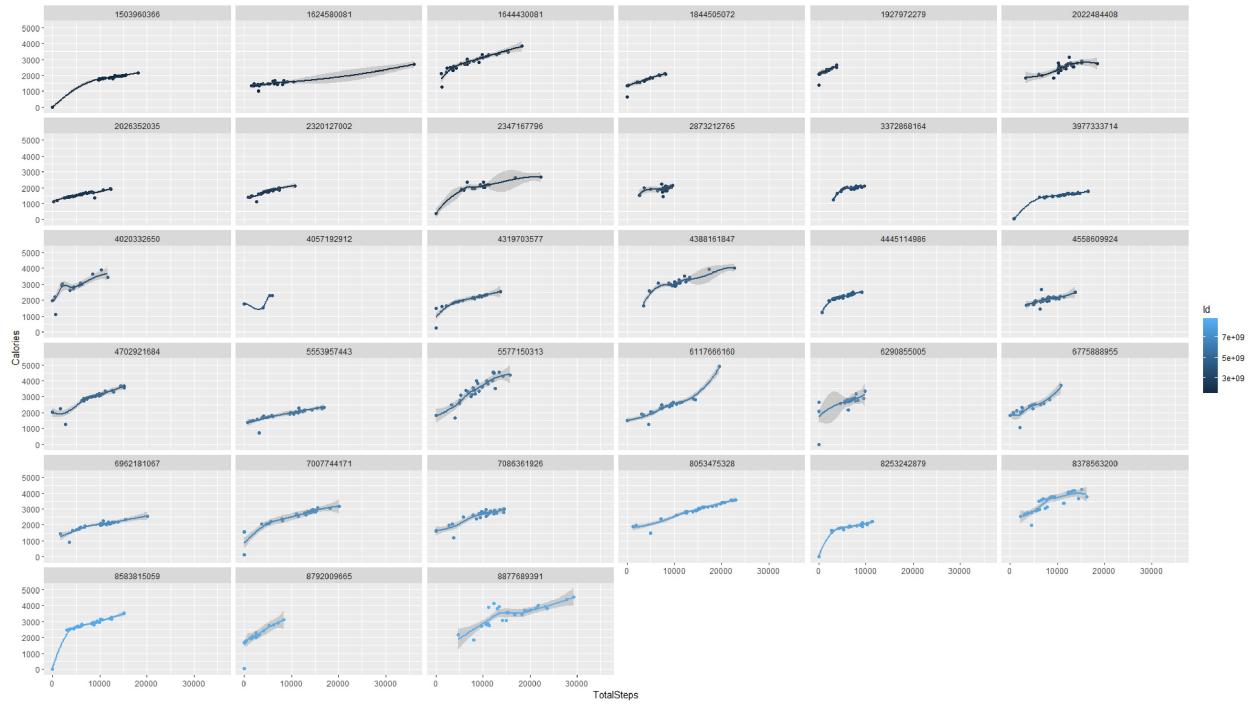
Visualizations

Plotting a graph between distance/steps and calories burnt with color representing different people. Same color points represents the same person walking on different dates. Second graph shows separate plots for steps vs calories burnt for different people over time

```
ggplot(data=activity, aes(x=TotalDistance,y=Calories,color=Id)) + geom_point()+
geom_smooth(method="gam") + xlab('Total Distance') + ylab('Calories Burnt')+
scale_color_gradientn(colours = rainbow(33))
```

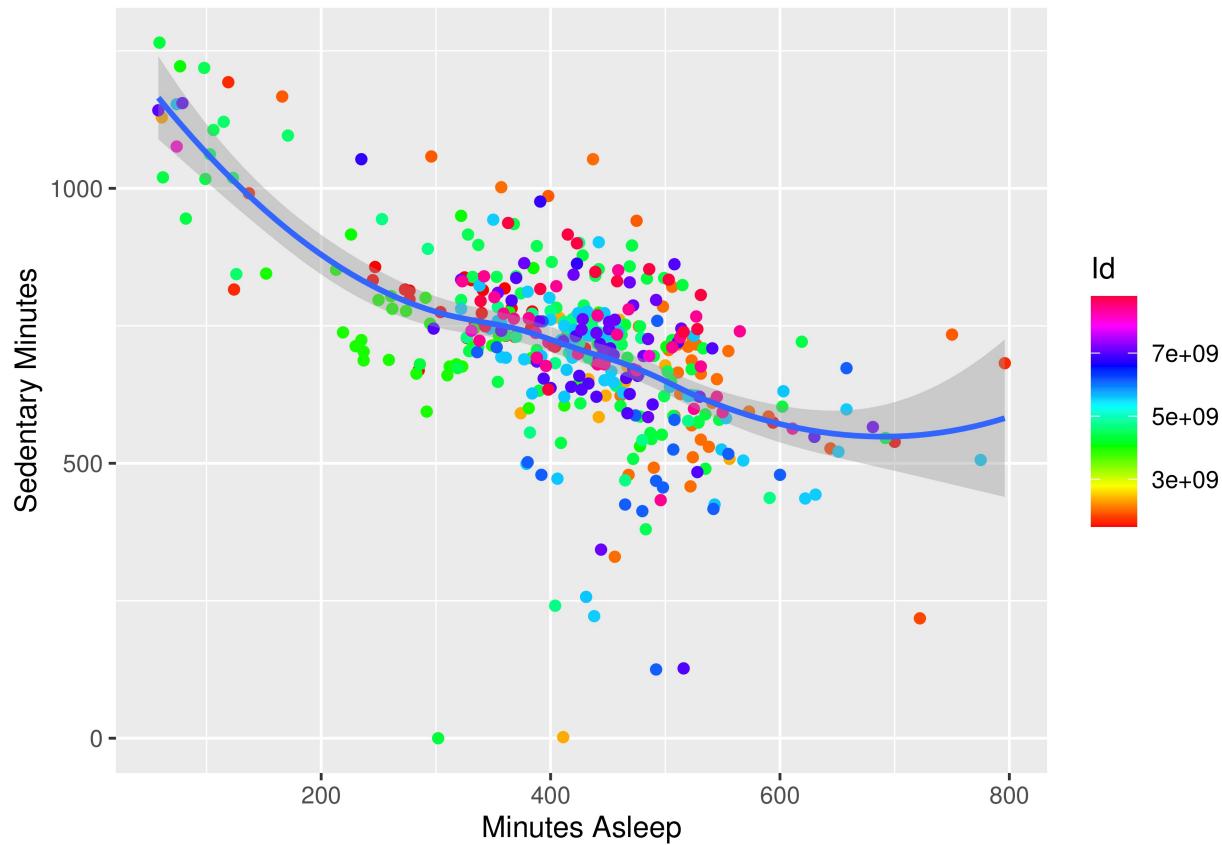


```
ggplot(data=activity, aes(x=TotalSteps,y=Calories,color=Id)) + geom_point() +  
geom_smooth(method="loess") + facet_wrap(~Id)
```



There is clear linear correlation between distance/steps and calories burnt. As the distance or number of steps increase - the more calories are burnt.

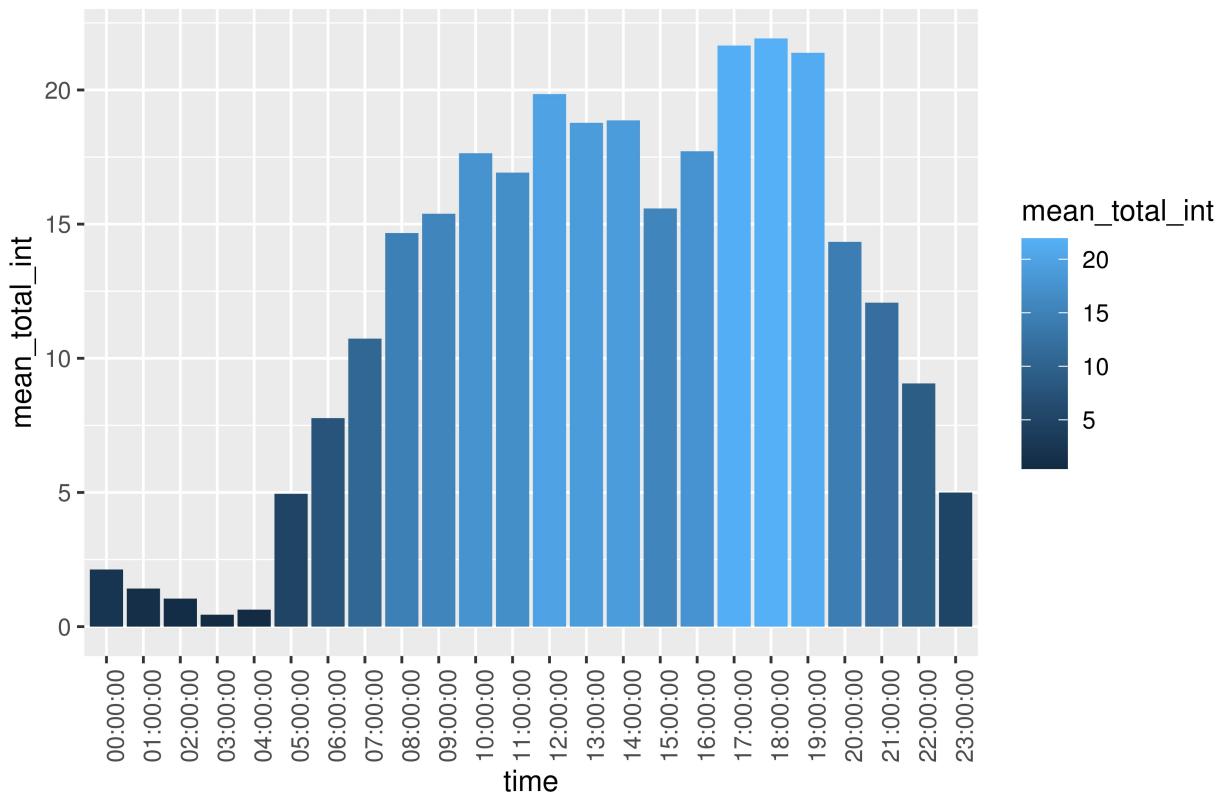
```
ggplot(data=activity_sleep,aes(x=TotalMinutesAsleep,y=SedentaryMinutes,color=Id))+  
  geom_point() +geom_smooth() + xlab('Minutes Asleep') +  
  ylab('Sedentary Minutes')+scale_color_gradientn(colours = rainbow(24))
```



- The above data shows as number of hours a person sleeps reduces, his sedentary time (non-active) time increases. One plausible reason could be that sleep replenishes our energy and reduces stress, thus making us more active.
- The data shows an average of 500 minutes(8 hrs) results in least sedentary time, while sleeping less or oversleeping leads to increased sedentary time.

```
ggplot(data=int_new, aes(x=time, y=mean_total_int, fill=mean_total_int)) +
  geom_bar(stat = "identity") + theme(axis.text.x = element_text(angle = 90)) +
  labs(title="Average Total Intensity vs. Time")
```

Average Total Intensity vs. Time



- The graph shows that people are most active between 7 am - 6 pm

Recommendations

- Bellabeat can give timely notifications to users on their step count and calories burnt. It can also give motivational texts or animations to urge people to walk and complete their daily goal.
- It can give a reminder for sleep as well as track sleep to give more insights on the sleeping patterns of a user, thus making them more aware and conscious of their sleep schedule.
- It can also give a notification if there is an inactivity for an hour or more, reminding people to do light exercise or walk.
- Bellabeat app can also be used to inform people of good habits that lead to a healthier and happier lifestyle

Since Bellabeat is a women focused brand, the marketing team can focus on three major groups-

- For older women, Bellabeat can be used to give timely alerts for their medicines or other daily activity.
- It can highlight the hectic and cumbersome life of working women or moms. Not giving time to their health because of work and children is very common.
- For younger women or girls, mainly students, studying or sitting for long hours without much physical movement can lead to back pain and high cholesterol levels. Another point that can be highlighted is the excess use of electronic devices and its blue rays which negatively impacts the sleep quality.

Thus implementing the above mentioned recommendations can prove very useful to all three classes of women.