Fitness Tracker Case Study

Tyler Dussuau

2023-02-13

Business Task

The goal is to identify trends and patterns in those who use fitness trackers. The drawn conclusions will drive the marketing strategy of Bellabeat, a company that manufactures smart devices focused on women. Specifically, the recommendations provided will be for the BellaBeat Time, a smart watch.

The dataset contains tables describing: - activity details by day, hour, and minute - weight - heart rate

Within these tables, the fields included are:

- daily:
  - id
  - log id
  - date
  - total steps
  - total calories
  - total distance
  - distance based on activity level (sedentary, lightly active, moderately active, very active)
  - duration based on activity level (sedentary, lightly active, moderately active, very active)
  - weight
  - BMI
  - Fat
- hourly:
  - id
  - log id
  - date
  - steps in hour
  - total intensity
  - average intensity
  - calories
- minute:
  - id
  - log id
  - date
  - steps
  - METS
  - calories
  - sleep

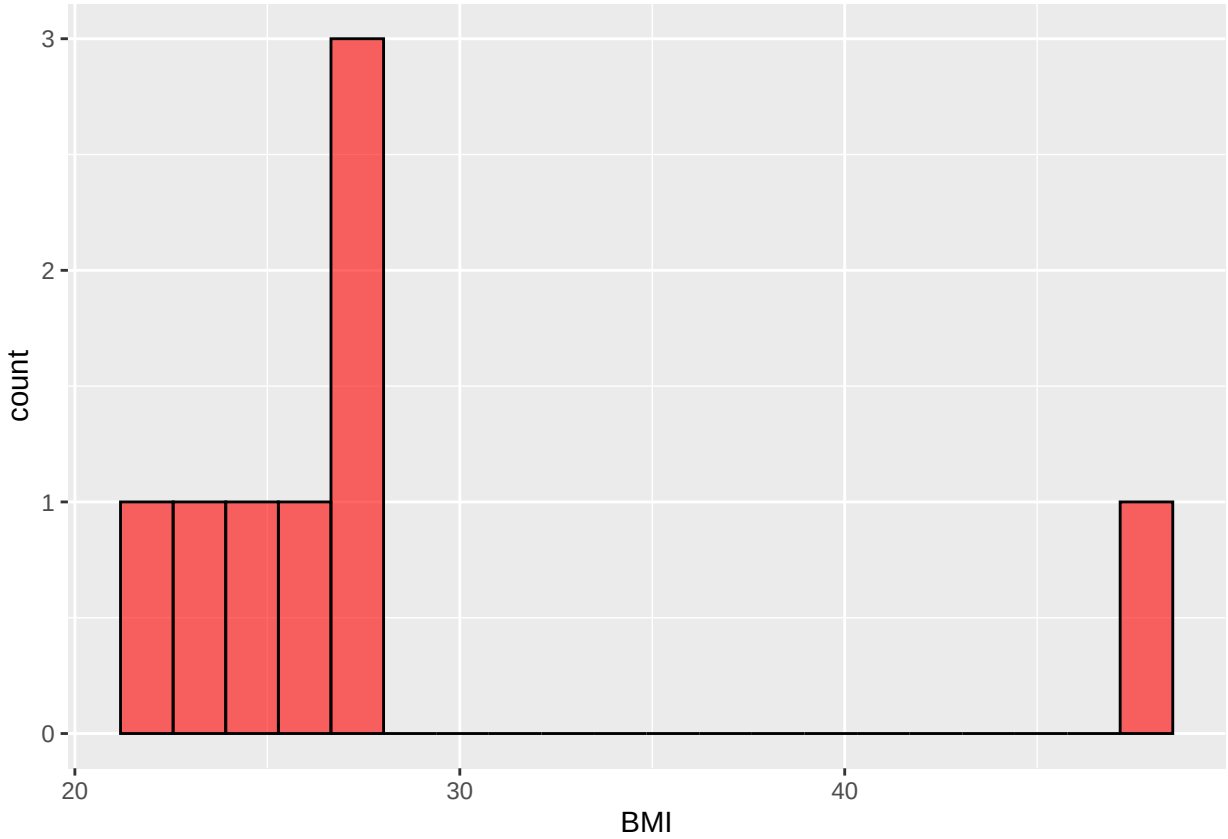
The tables with information based on day include 33 people, except the weight table. The weight table included 8 people. All the tables with information based on hour and minute include 33 people. Heart rate inlcudes data for 14 people. This data is from March 12, 2016 to May 12, 2016. It was collected by a survey on Amazon Mechanical Turk. 30 users consented for their FitBit data to be used.

BMI Distribution

```
## Rows: 67 Columns: 7
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (1): Date
## dbl (5): Id, WeightKg, WeightPounds, BMI, LogId
## lgl (1): IsManualReport
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

Using a histogram, we can see that of the eight users, most of them are overweight or near overweight. A healthy BMI is 18.5 and 24.9, according to the CDC.

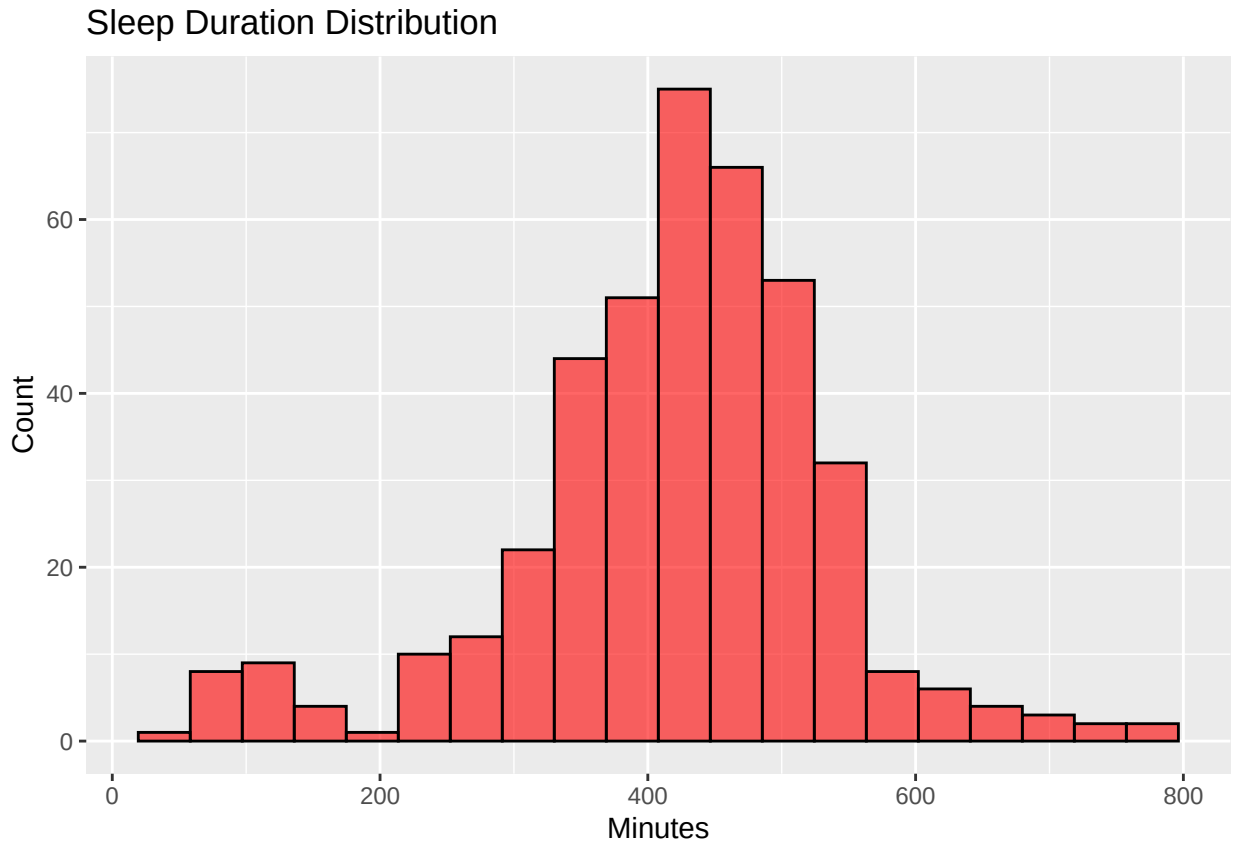
Reference: https://www.cdc.gov/healthyweight/assessing/index.html#:~:text=If%20your%20BMI%20is%20less,falls%20witl



Sleep Duration

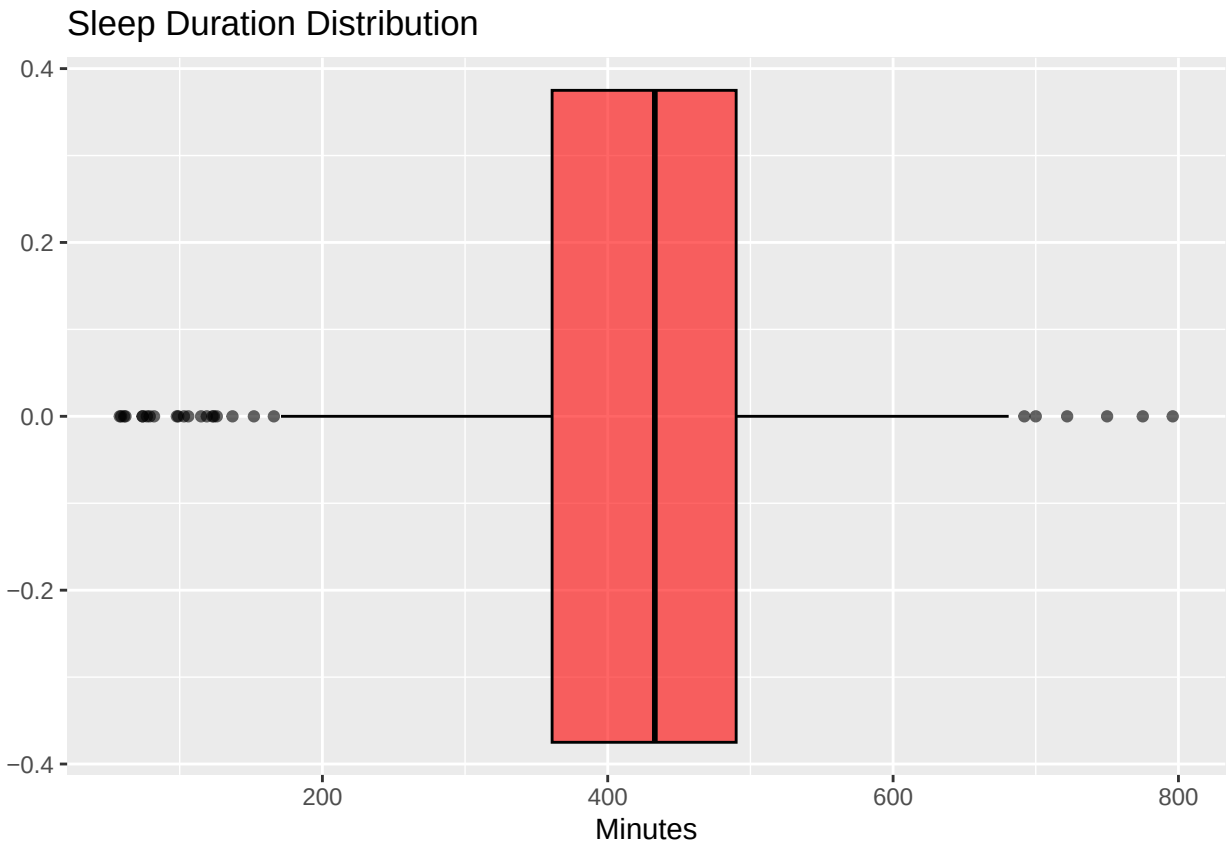
Do users get enough sleep? We will create a histogram to see the distribution of sleep durations.

```
## Rows: 413 Columns: 5
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (1): SleepDay
## dbl (4): Id, TotalSleepRecords, TotalMinutesAsleep, TotalTimeInBed
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```



Most users tend to get a healthy amount of sleep. If we take another look, we can see that 50% of sleep durations fall between 350 and 490 minutes. This means that there are also users who are sleep for shorter periods and longer periods than what is recommended.

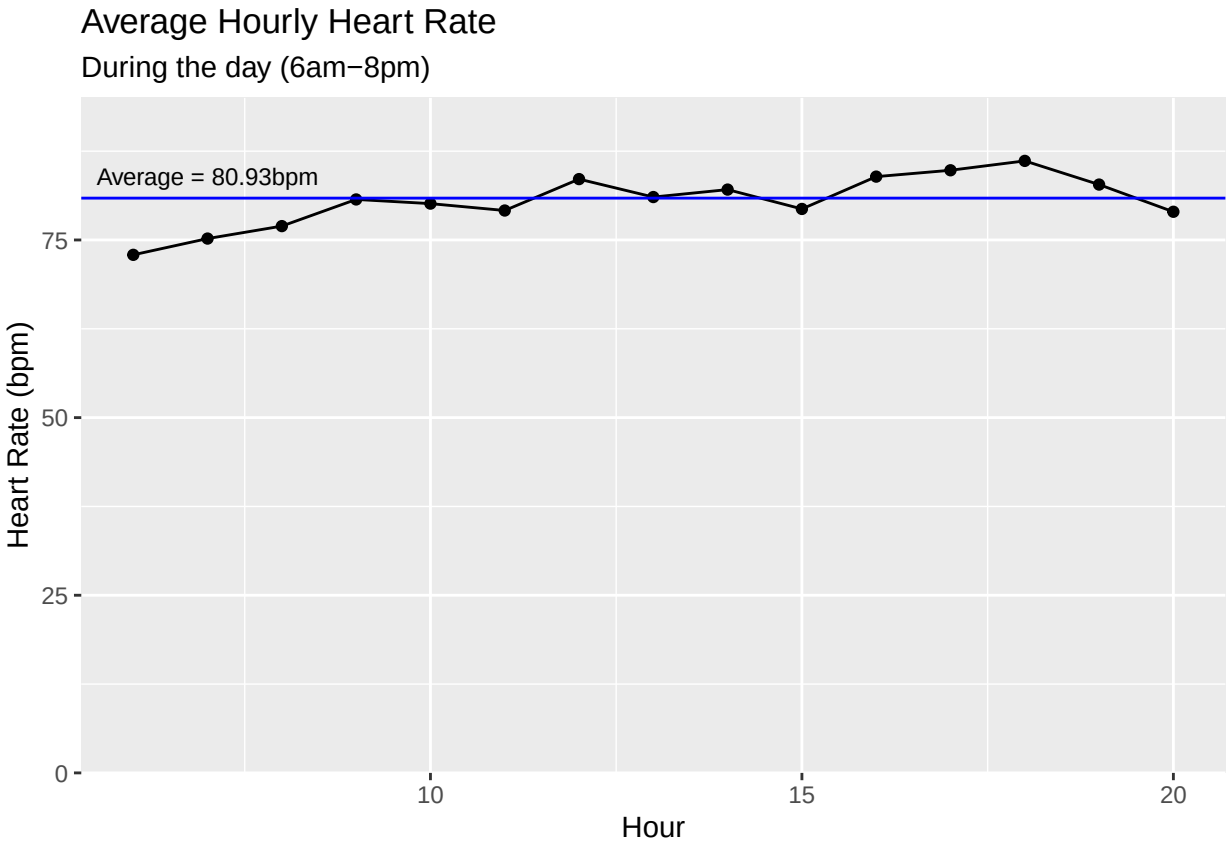
```
## Warning in geom_boxplot(color = "black", fill = "red", alpha = 0.6, bins = 20):
## Ignoring unknown parameters: `bins`
```



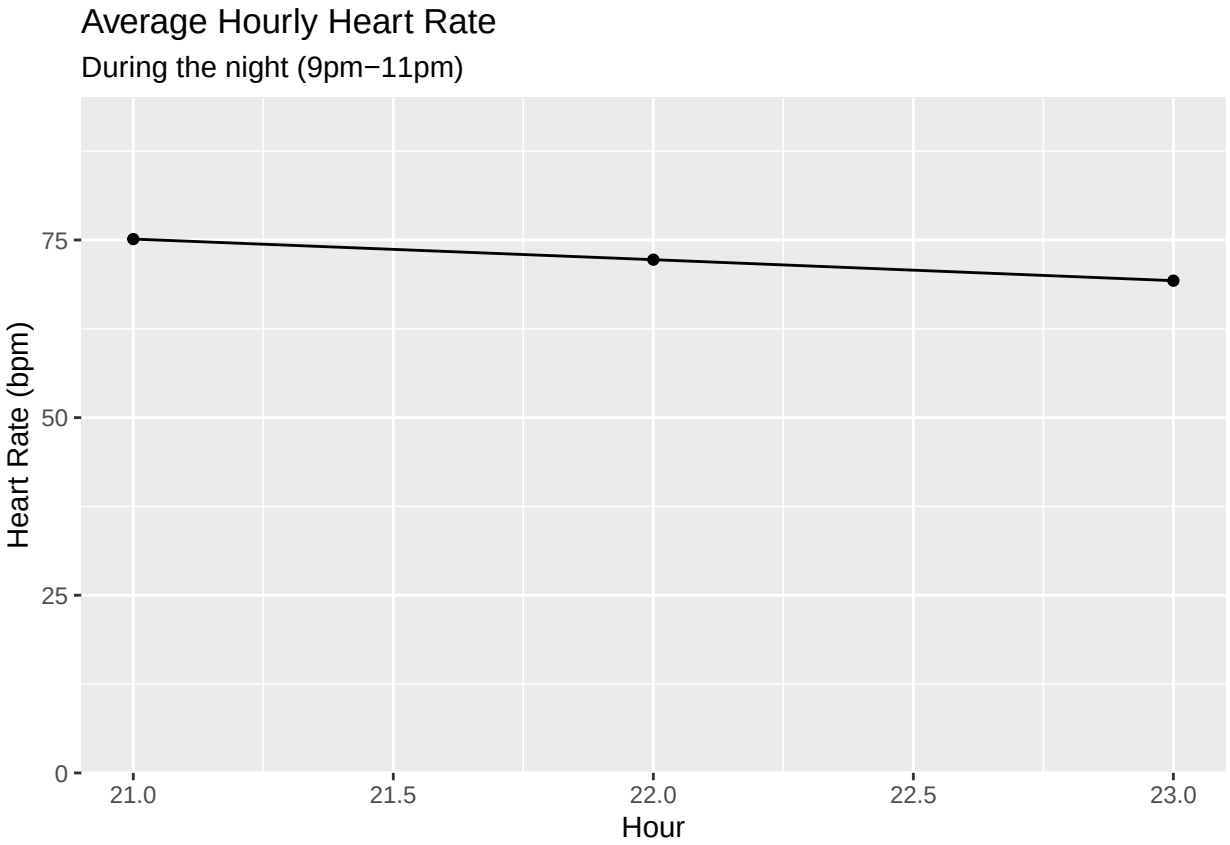
Average Heart Rate

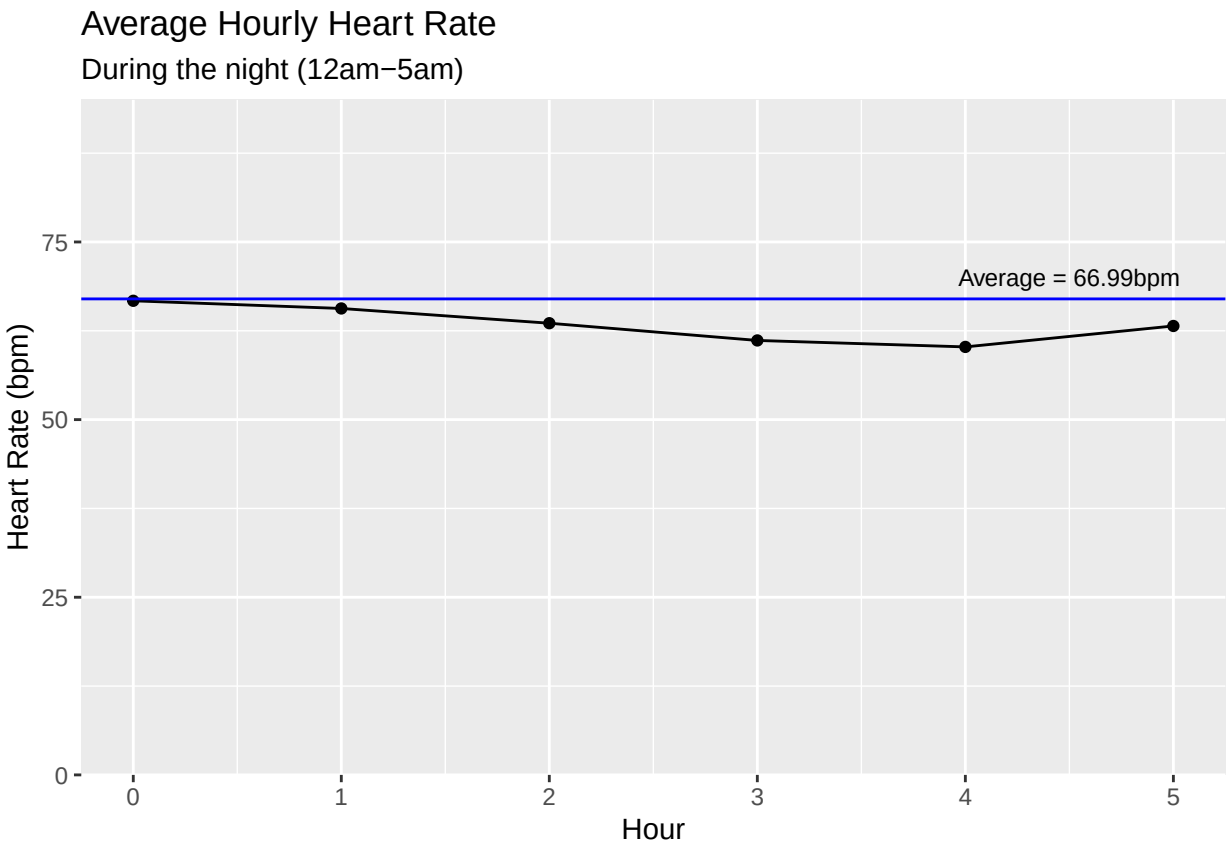
Let's look at how users' heart rates change during the day compared to sleep. According to Harvard Health, a healthy heart rates during sleep should be around 20% - 30% less than the resting heart rate. We will take the average heart rate during the day (6am - 9pm)

```
## Rows: 2483658 Columns: 3
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (1): Time
## dbl (2): Id, Value
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```



The average heart rate is 80.93bpm. So, the average sleeping heart rate should be between 56.51bpm and 64.74bpm. The graph will be split into two. The first part of the graph is between 9pm-11pm. The second graph is between 12am-5am.


Average Hourly Heart Rate
During the night (9pm−11pm)
Heart Rate (bpm)
75
50
25
0
21.0
21.5
22.0
22.5
23.0
Hour

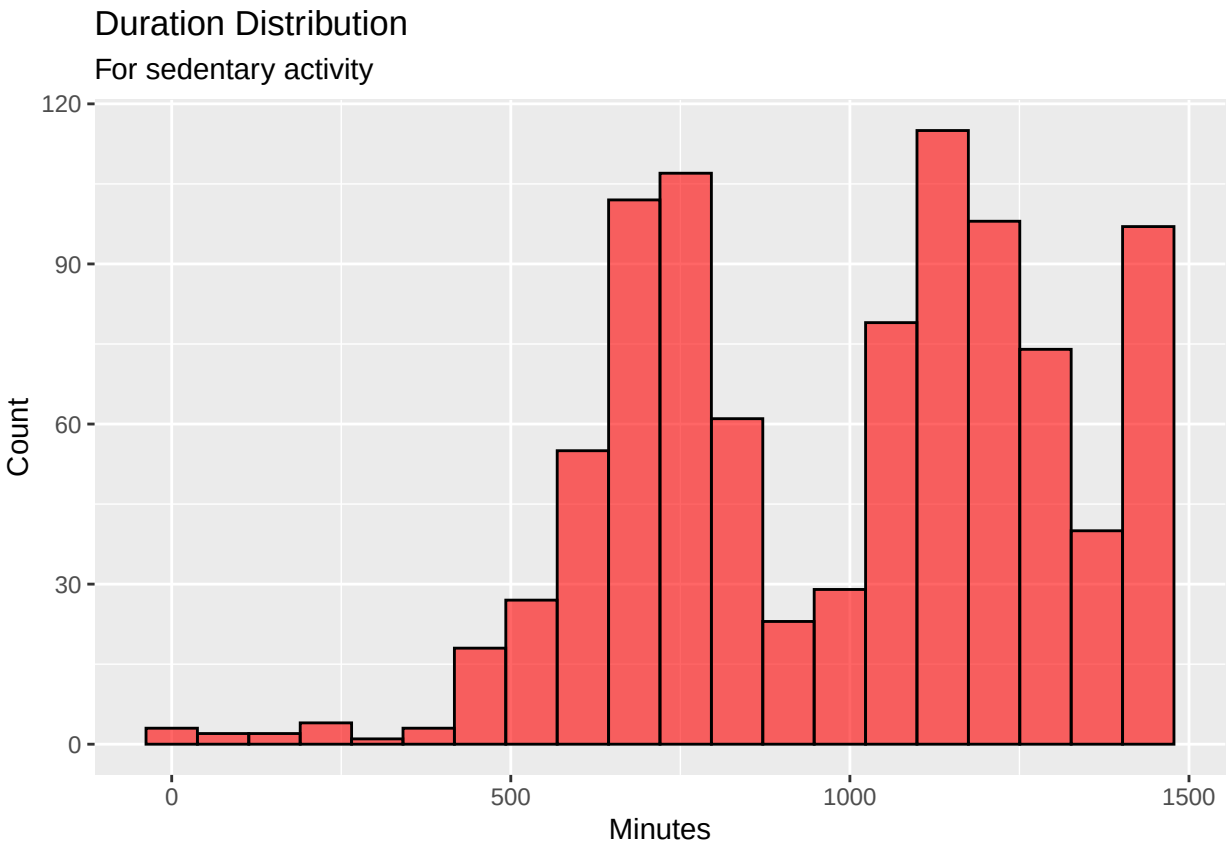


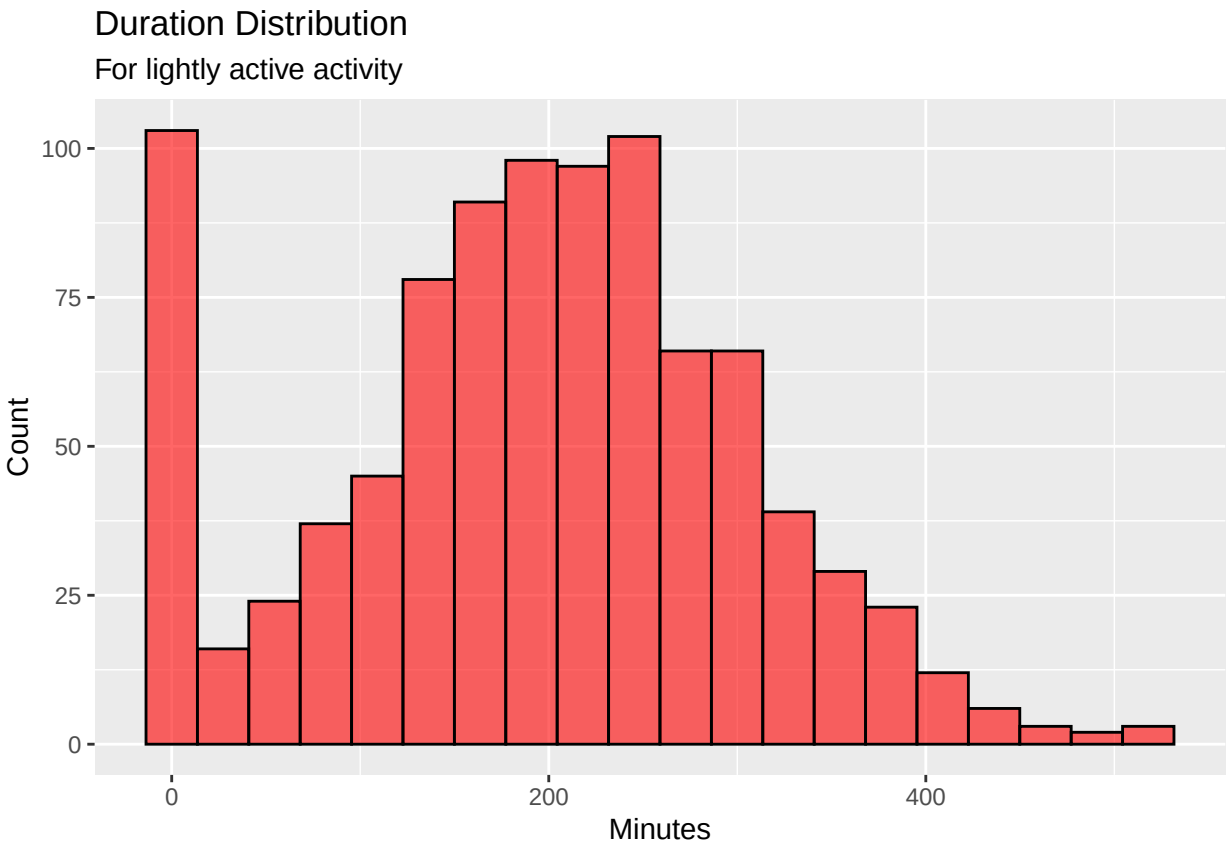
The average heart rate throughout the day (6am-9pm) is 80.61 bpm. One thing this fails to consider is not all of these heart rates are druing resting heart rates. Of course, exercise will increase the value of heart rate. It is most likely lower. At night, it is 65.56bpm. While it is not over the threshold by much, this is also means many users have high heart rates during sleep, indicating bad sleep quality.

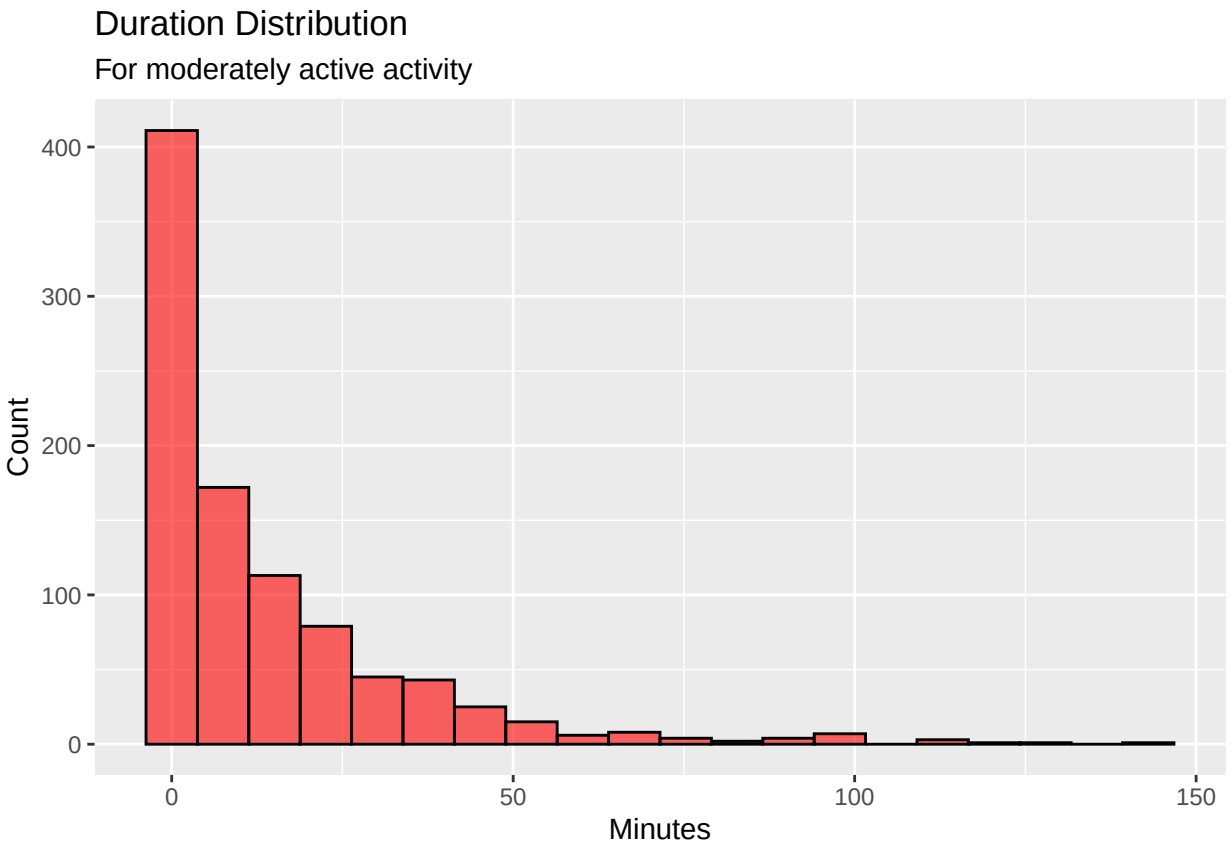
Activity Levels

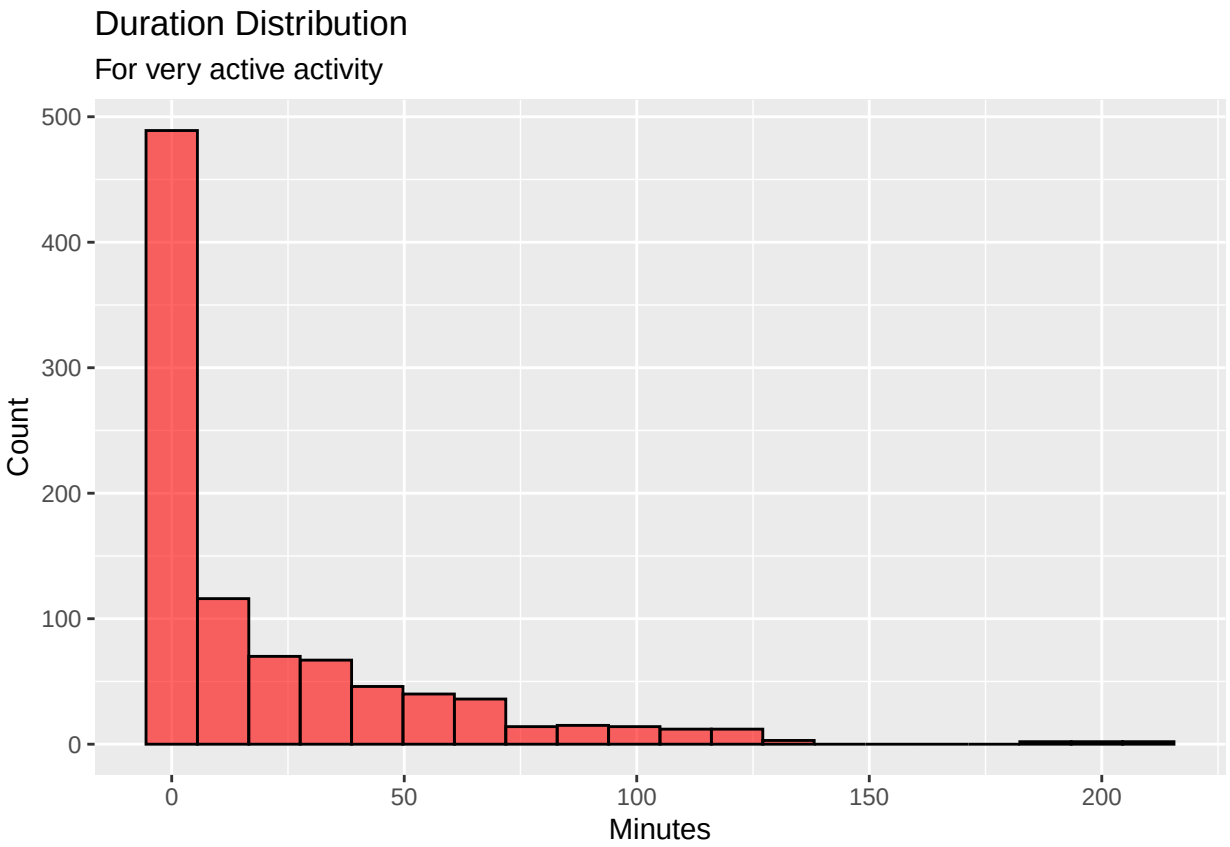
```
## Rows: 940 Columns: 15
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (1): ActivityDate
## dbl (14): Id, TotalSteps, TotalDistance, TrackerDistance, LoggedActivitiesDi...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

If we compare histograms for each activity level, as activity level increases, the graphs turn from negatively skewed to postiviely skewed, indicating users are more often lightly active and sedentary active. There is no insight here since we cannot expect users to be very active and moderately active for long.

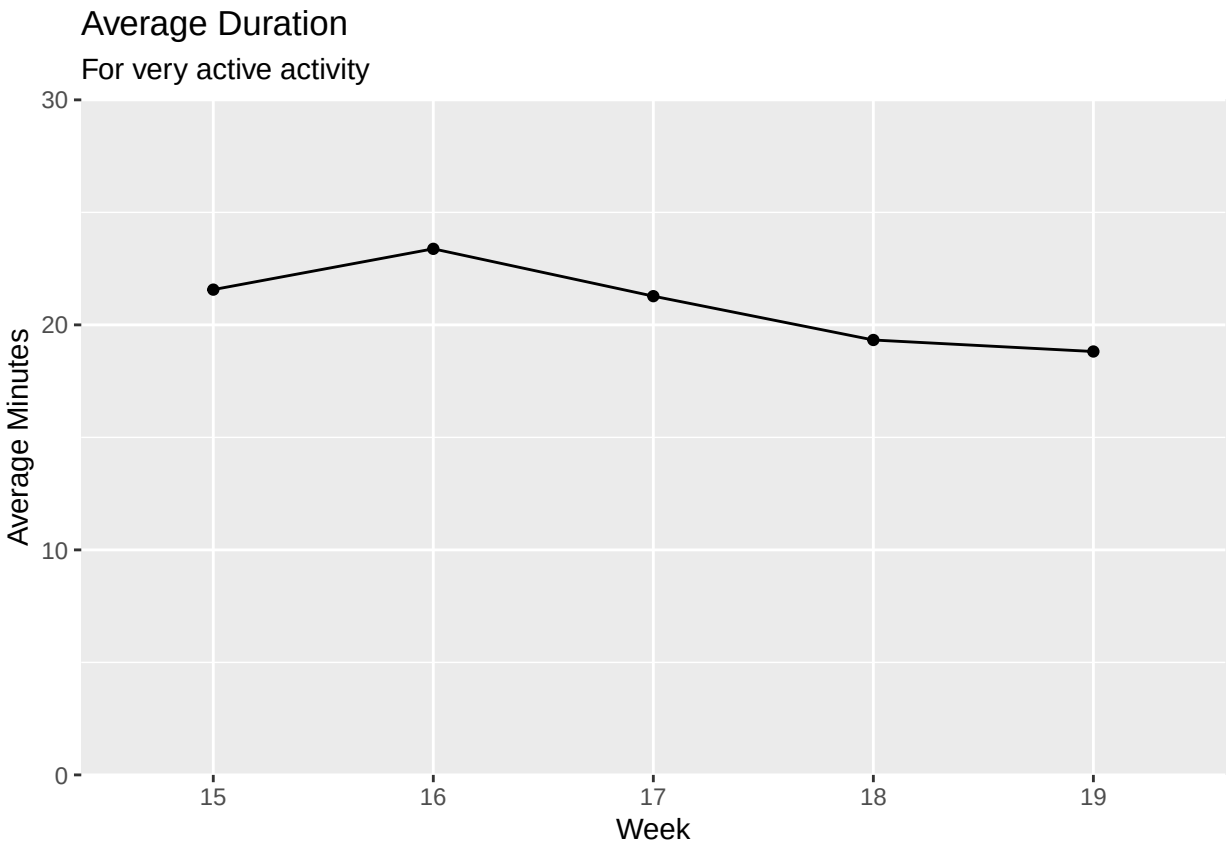

Duration Distribution
For sedentary activity
120
90
60
30
0
Count
0
500
1000
1500
Minutes

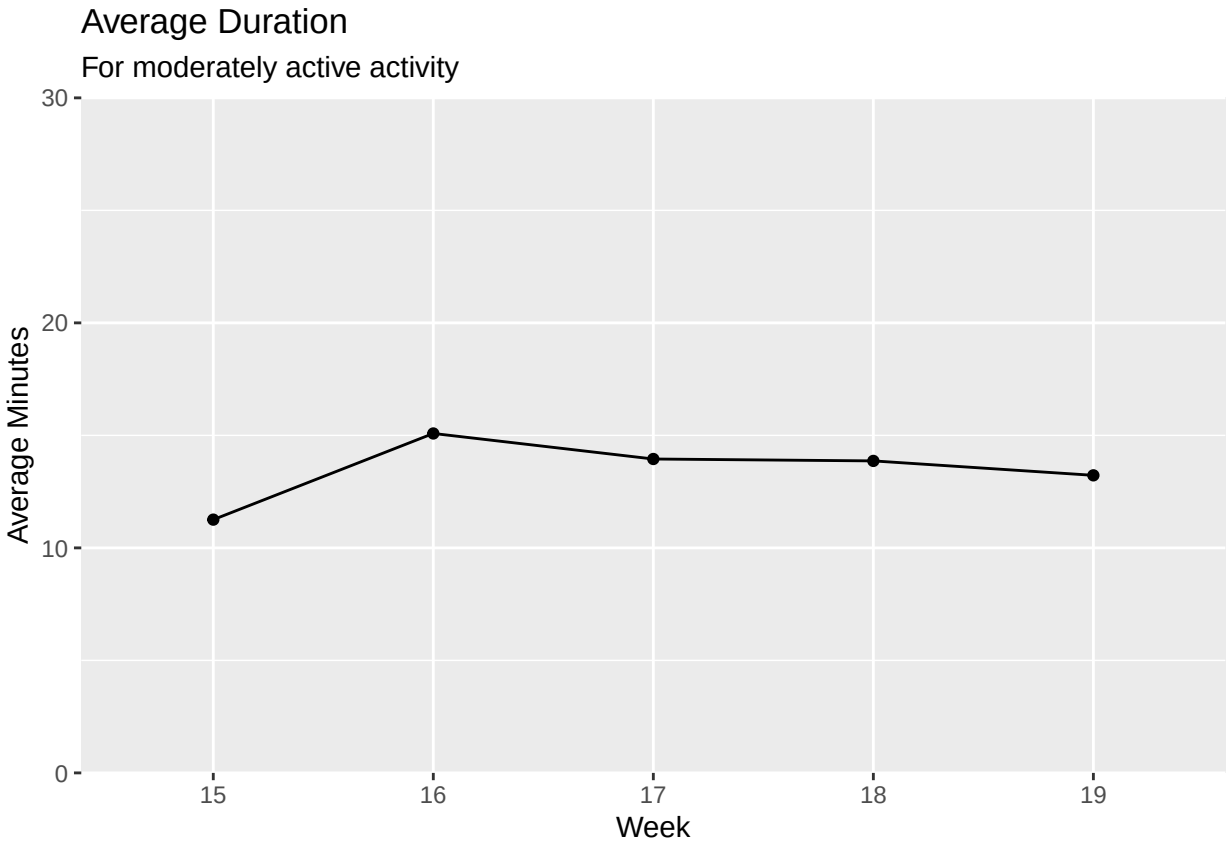

Duration Distribution
For lightly active activity
Count
100
75
50
25
0
0
200
400
Minutes


Duration Distribution
For moderately active activity
Count
400
300
200
100
0
0
50
100
150
Minutes



Let's look another aspect. It is recommended for users to be moderately active for at least 150 minutes or very active for 75 minutes a week. Are users achieving this goal?


Average Duration
For very active activity
30
20
10
0
Average Minutes
15
16
17
18
19
Week



Most users are not getting the recommended amount of exercise during the week.

Recommendations

Note: There are limitations to these conclusions. This data is from seven years ago. Also, the data only reflects thirty users. For some tables, there are as a little as eight people. If possible, it would be best to recollect data to get a more accurate view of users. In a new study, it should include more users.

Based on the available data, I have three recommendations:

1) We should market towards people trying to lose weight or those who are helping other lose weight. The BMI distribution chart shows that most people who have FitBits are overweight. So, one demographic to cater is dietitans. Their clients would need a way to monitor their weight loss. The Bellabeat Leaf could monitor activity to see how users are reaching their daily amount of exercise.

2) Sleep durations vary from 5.5-9 hours. However, if we look at users' heart rates during sleep times, they are higher than they should be. We should market how the Bellabeat Time can help users monitor their sleep patterns. We know Bellabeat caters to women. So, it would help especially during the menstrual cycle, when users mmight have difficultly sleeping.

Reference: https://www.health.harvard.edu/blog/how-does-sleep-affect-your-heart-rate-2021012921846

3) Users are sedentary for a majority of the day. They rarely engange in moderate or very active exercise. The Bellabeat Time could be great for mothers who may only have time or energy for light activity. Advertisements on YouTube and Google for content related to motherhood is a possibility.

*Next steps: - In a new survey, record more weights and BMIs to see if users are losing weight with their FitBit. - Track users over a 3-month period to see how different activitly levels affect heart rate. - Look at if family size affects activity levels - Track what hours users get up and go to sleep