

Proposed Marketing Strategy Focus for Bellabeat Application

Charlene A. Robinson

9/1/2021

Proposed Marketing Strategy Focus for Bellabeat Application

What is Bellabeat?

Bellabeat is a health-tech company focused on empowering women's health through collecting data on stress, sleep, activity, and reproductive health and turning that data into useful information for their customers. The Bellabeat app provides insights to the customer's entire body ecosystem, from reproductive health to weight loss to stress management. For more information, see <https://bellabeat.com/science/>.

Business Objective:

Using public dataset from competing organizations using similar smart technology devices to measure health, this proposal aims to uncover insights in trends that apply to the Bellabeat application and how this could shape Bellabeat's marketing strategy for the Bellabeat app.

Data Analysis Process:

Data from 30 Fitbit participant survey was collected by Amazon Mechanical Turk between March 12, 2016 to May 12, 2016 (found here).

The following steps were taken:

1. Download csv files
2. Upload csv files on daily activity, heart rate, and sleep to RStudio workspace
3. Installed and loaded the following packages:
 - tidyverse
 - here
 - skimr
 - janitor
 - dplyr
4. Created dataframes for each csv file and reviewed data (Daily_Activity, Heartrate_Seconds, Daily_Sleep)
5. Opened sleep and heart rate csv files in Excel to remove time stamp on date/time column
6. Re-uploaded sleep and heart rate csv files
7. Removed old sleep and heart rate dataframes and created new dataframes from new csv files (Heart_Rate_By_Date and Daily_Sleep, respectively)
8. Renamed date column in Daily_Sleep and Heart_Rate_By_Date to match column name in Daily_Activity dataframe (ActivityDate)
9. In Heart_Rate_By_Date dataframe, grouped by first Id then ActivityDate to calculate mean = Daily_Average_HR, making new dataframe New_Heart_Rate_By_Date
10. Full joined New_Heart_Rate_By_Date and Daily_Activity to create Daily_No_Sleep dataframe
11. Full joined Daily_No_Sleep and Daily_Sleep to create Daily_All_Data dataframe
12. Created new column for TotalActiveMinutes by combining VeryActiveMinutes + FairlyActiveMinutes + LightlyActiveMinutes

13. Filtered out rows where Calories = 0
14. Pivoted Very, Fairly, and LightlyActiveMinutes to create an ActivityType column with ActivityMinutes
15. Summarize data within Daily_All_Data to observe correlations, trends, and further insights (see below)
16. Created plots to understand correlations and trends better (see Unsurprising Trends section)

```
summary(Daily_All_Data)
```

```
##           Id           ActivityDate      TotalSleepRecords TotalMinutesAsleep
##  Min.      :1.504e+09  Length:2817          Min.      :1.000      Min.      : 58.0
##  1st Qu.:2.320e+09    Class :character  1st Qu.:1.000      1st Qu.:361.0
##  Median :4.445e+09    Mode  :character  Median :1.000      Median :433.0
##  Mean   :4.853e+09                    Mean   :1.119      Mean   :419.5
##  3rd Qu.:6.962e+09                    3rd Qu.:1.000      3rd Qu.:490.0
##  Max.   :8.878e+09                    Max.   :3.000      Max.   :796.0
##                                     NA's   :1578      NA's   :1578
##  TotalTimeInBed  Daily_Average_HR  TotalSteps    TotalDistance
##  Min.      : 61.0  Min.      : 59.01  Min.      : 0    Min.      : 0.000
##  1st Qu.:403.0  1st Qu.: 71.25  1st Qu.: 3821  1st Qu.: 2.650
##  Median :463.0  Median : 77.58  Median : 7451  Median : 5.280
##  Mean   :458.6  Mean   : 77.40  Mean   : 7685  Mean   : 5.526
##  3rd Qu.:526.0  3rd Qu.: 83.28  3rd Qu.:10742  3rd Qu.: 7.730
##  Max.   :961.0  Max.   :104.87  Max.   :36019  Max.   :28.030
##  NA's    :1578  NA's    :2418
##  TrackerDistance  LoggedActivitiesDistance  VeryActiveDistance
##  Min.      : 0.000  Min.      :0.00000  Min.      : 0.000
##  1st Qu.: 2.650  1st Qu.:0.00000  1st Qu.: 0.000
##  Median : 5.280  Median :0.00000  Median : 0.220
##  Mean   : 5.512  Mean   :0.1105  Mean   : 1.511
##  3rd Qu.: 7.720  3rd Qu.:0.00000  3rd Qu.: 2.090
##  Max.   :28.030  Max.   :4.9421  Max.   :21.920
##
##  ModeratelyActiveDistance  LightActiveDistance  SedentaryActiveDistance
##  Min.      :0.00000  Min.      : 0.000  Min.      :0.000000
##  1st Qu.:0.00000  1st Qu.: 1.960  1st Qu.:0.000000
##  Median :0.2500  Median : 3.390  Median :0.000000
##  Mean   :0.5733  Mean   : 3.364  Mean   :0.001608
##  3rd Qu.:0.8100  3rd Qu.: 4.800  3rd Qu.:0.000000
##  Max.   :6.4800  Max.   :10.710  Max.   :0.110000
##
##  SedentaryMinutes  Calories  TotalActiveMinutes  TotalMinutesNotAsleep
##  Min.      : 0.0  Min.      : 52  Min.      : 0.0  Min.      : 0.00
##  1st Qu.: 728.0  1st Qu.:1835  1st Qu.:148.0  1st Qu.: 17.00
##  Median :1055.0  Median :2149  Median :249.0  Median : 25.00
##  Mean   : 988.4  Mean   :2317  Mean   :228.9  Mean   : 39.17
##  3rd Qu.:1226.0  3rd Qu.:2798  3rd Qu.:318.0  3rd Qu.: 40.00
##  Max.   :1440.0  Max.   :4900  Max.   :552.0  Max.   :371.00
##                                     NA's    :1578
##  ActivityType  ActivityMinutes
##  Length:2817  Min.      : 0.00
##  Class :character  1st Qu.: 0.00
##  Mode  :character  Median : 18.00
##                                     Mean   : 76.29
##                                     3rd Qu.:131.00
##                                     Max.   :518.00
##
```

Limitations: To clarify, this data does have limitations. As seen in the sleep and heart rate category summary, there are N/A values, meaning those participants did not have data logged for sleep and heart rate for each day. Also, this data does not specify only women participated. Therefore, in looking for trends to apply to marketing strategies for Bellabeat, it is best to keep in mind that the amount of women in this dataset is unknown.

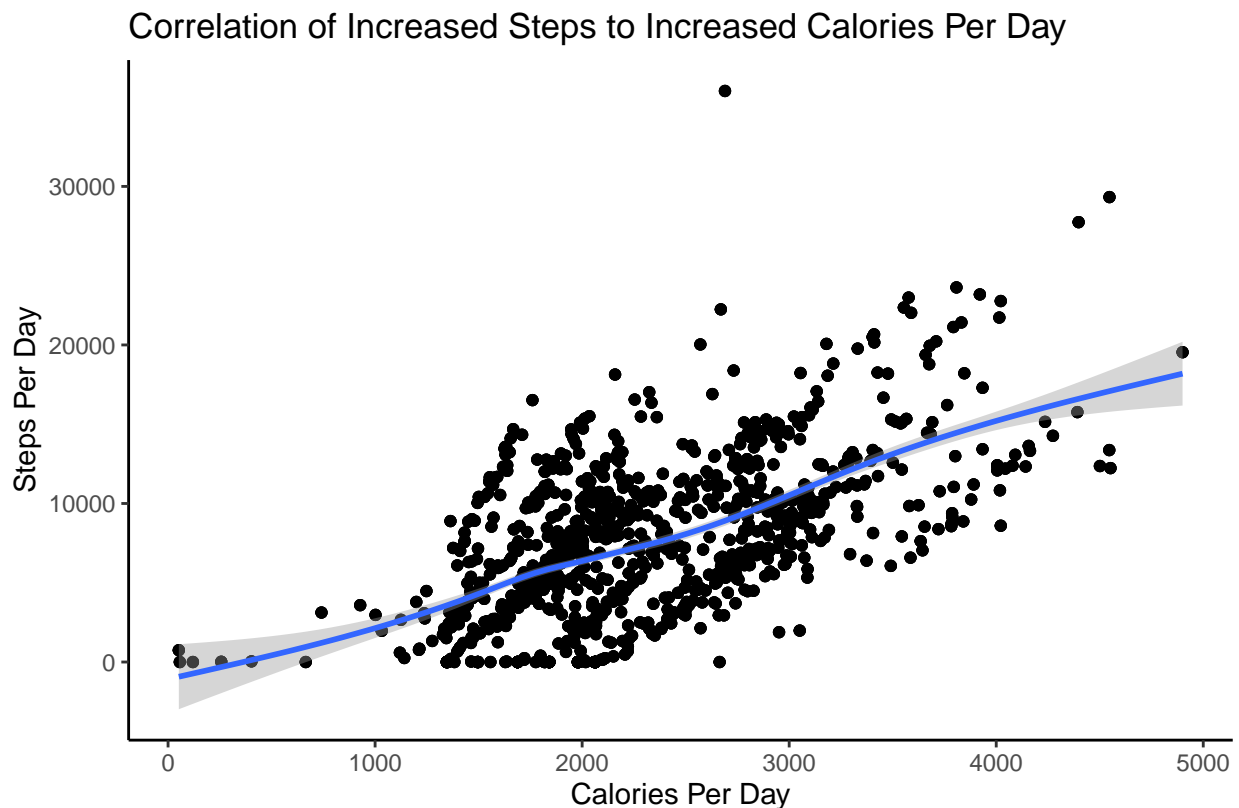
Unsurprising Trends:

It will come as no surprise that an increase in total steps and total active minutes show a trend in increase of calories. Here are some graphs that demonstrate the trend within this data.

```
ggplot(data=Daily_All_Data) +
  geom_point(aes(x=Calories, y=TotalSteps)) +
  geom_smooth(aes(x=Calories, y=TotalSteps)) +
  labs(title="Correlation of Increased Steps to Increased Calories Per Day", caption="CC0: Public Domain") +
  xlab('Calories Per Day') + ylab('Steps Per Day') +
  theme_classic()
```

Calories and Steps

```
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```



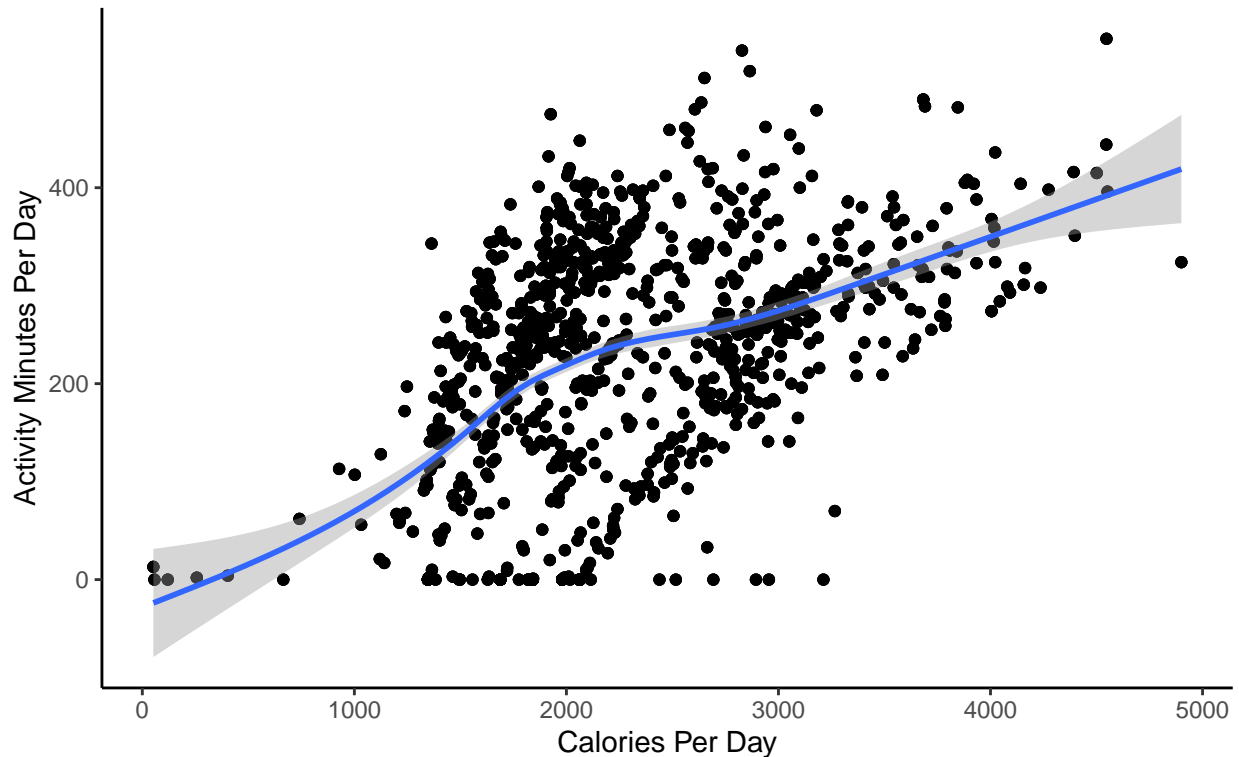
```
ggplot(data=Daily_All_Data) +
  geom_point(aes(x=Calories, y=TotalActiveMinutes)) +
  geom_smooth(aes(x=Calories, y=TotalActiveMinutes)) +
  labs(title="Correlation of Increased Activity to Increased Calories Per Day", caption="CC0: Public Domain") +
  xlab('Calories Per Day') + ylab('Total Active Minutes')
```

```
xlab('Calories Per Day') + ylab('Activity Minutes Per Day') +  
theme_classic()
```

Calories and Activity

```
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```

Correlation of Increased Activity to Increased Calories Per Day



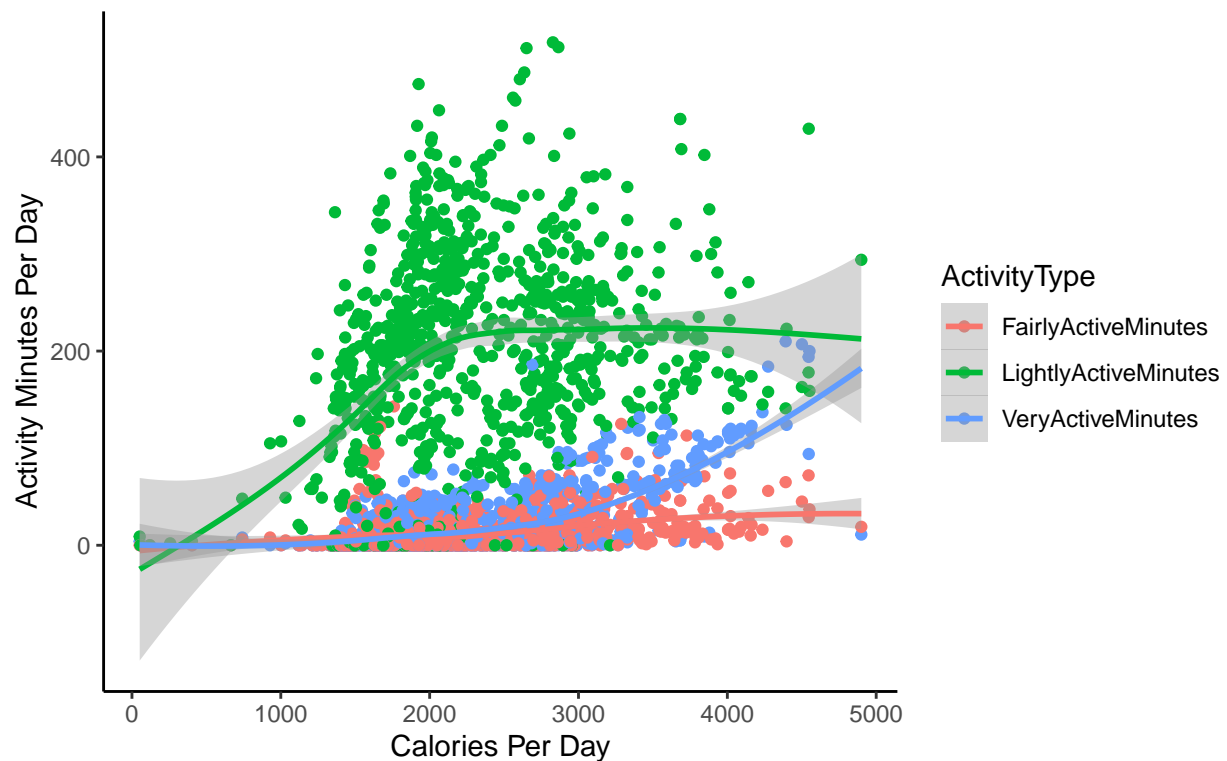
CC0: Public Domain, survey through Amazon Mechanical Turk between 03.12.2016–05.12.2016

Calories and Activity Type A little more interesting here: the data shows that lightly active time has an initial spike on calories per day before leveling off, but very active has a spike past 3,000 calories per day.

```
ggplot(data=Daily_All_Data) +  
  geom_point(aes(x=Calories, y=ActivityMinutes, color=ActivityType)) +  
  geom_smooth(aes(x=Calories, y=ActivityMinutes, color=ActivityType))+  
  labs(title="Correlation of Activity Types to Calories Per Day", caption="CC0: Public Domain, survey t  
  xlab('Calories Per Day') + ylab('Activity Minutes Per Day') +  
  theme_classic()
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

Correlation of Activity Types to Calories Per Day



: Domain, survey through Amazon Mechanical Turk between 03.12.2016–05.12.2016

Interesting Trends to Focus:

To focus in on marketing strategies regarding the Bellabeat app, some more interesting trends were focused around the average heart rate. A lower heart rate can also be seen as a reduction in stress (see here). The data showed that an increase in total active minutes, increased steps, and more sleep help to lower the participant's heart rate and, likely, also stress level.

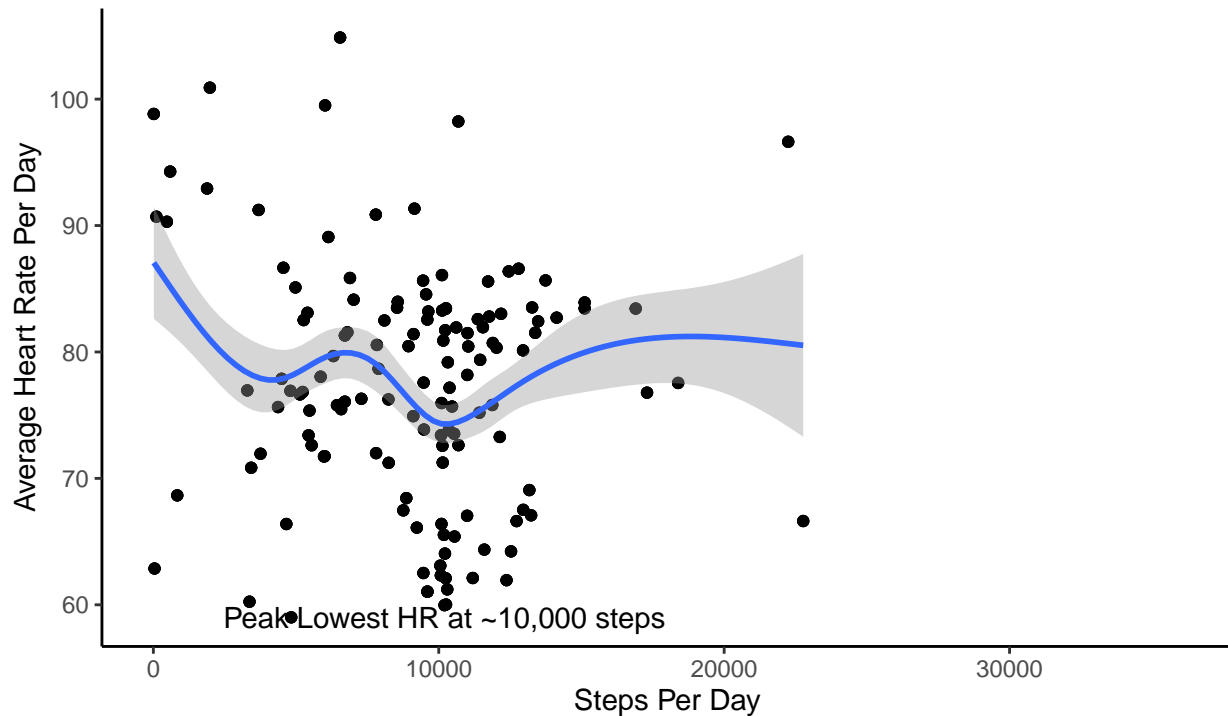
```
ggplot(data=Daily_All_Data) +
  geom_point(aes(x=TotalSteps, y=Daily_Average_HR)) +
  geom_smooth(aes(x=TotalSteps, y=Daily_Average_HR)) +
  labs(title="Correlation of Increased Steps to Average Heart Rate Per Day", subtitle="Not all participants") +
  annotate("text", x=10201, y=59, label="Peak Lowest HR at ~10,000 steps") +
  xlab('Steps Per Day') + ylab('Average Heart Rate Per Day') +
  theme_classic()
```

More Steps, Lower Heart Rate

```
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## Warning: Removed 2418 rows containing non-finite values (stat_smooth).
## Warning: Removed 2418 rows containing missing values (geom_point).
```

Correlation of Increased Steps to Average Heart Rate Per Day

Not all participants had heart rate data



CC0: Public Domain, survey through Amazon Mechanical Turk between 03.12.2016–05.12.2016

Peak lowest heart rate was around 10,000 steps.

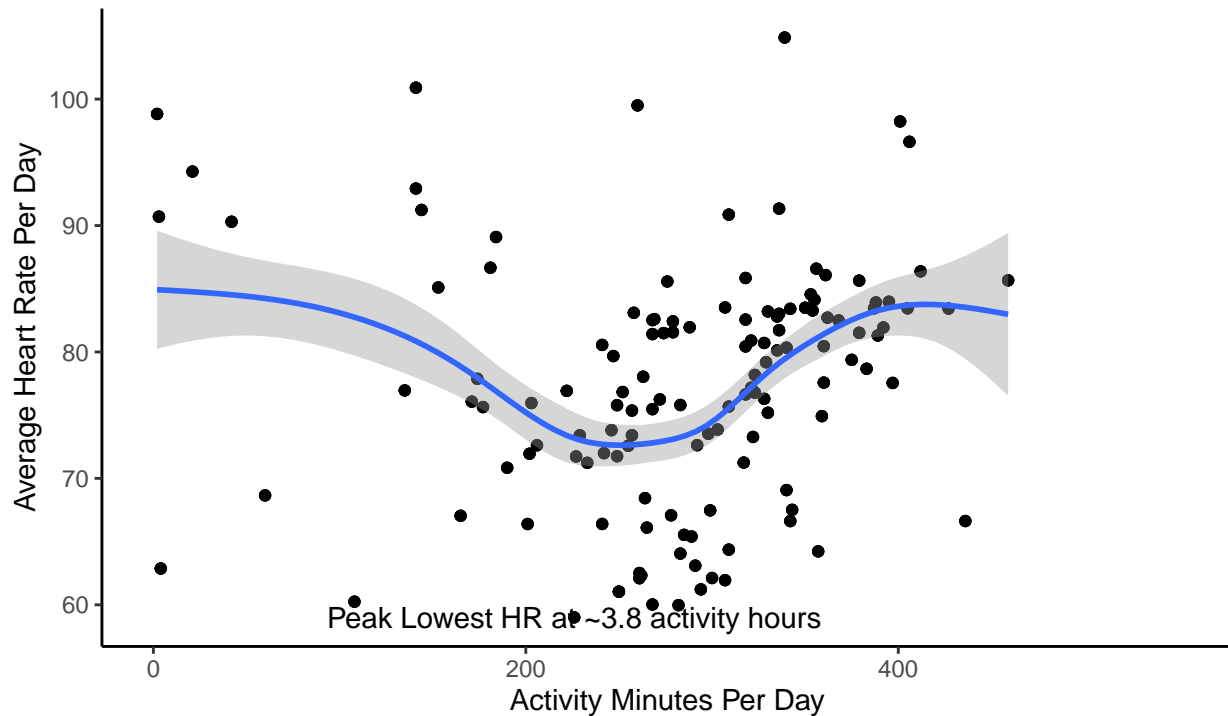
```
ggplot(data=Daily_All_Data) +
  geom_point(aes(x=TotalActiveMinutes, y=Daily_Average_HR)) +
  geom_smooth(aes(x=TotalActiveMinutes, y=Daily_Average_HR)) +
  labs(title="Correlation of Increased Activity to Average Heart Rate Per Day", subtitle="Not all participants had heart rate data") +
  annotate("text", x=226, y=59, label="Peak Lowest HR at ~3.8 activity hours") +
  xlab('Activity Minutes Per Day') + ylab('Average Heart Rate Per Day') +
  theme_classic()
```

More Activity, Lower Heart Rate

```
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## Warning: Removed 2418 rows containing non-finite values (stat_smooth).
## Warning: Removed 2418 rows containing missing values (geom_point).
```

Correlation of Increased Activity to Average Heart Rate Per Day

Not all participants had heart rate data



CC0: Public Domain, survey through Amazon Mechanical Turk between 03.12.2016–05.12.2016

Peak lowest heart rate was around 3.8 hours of activity.

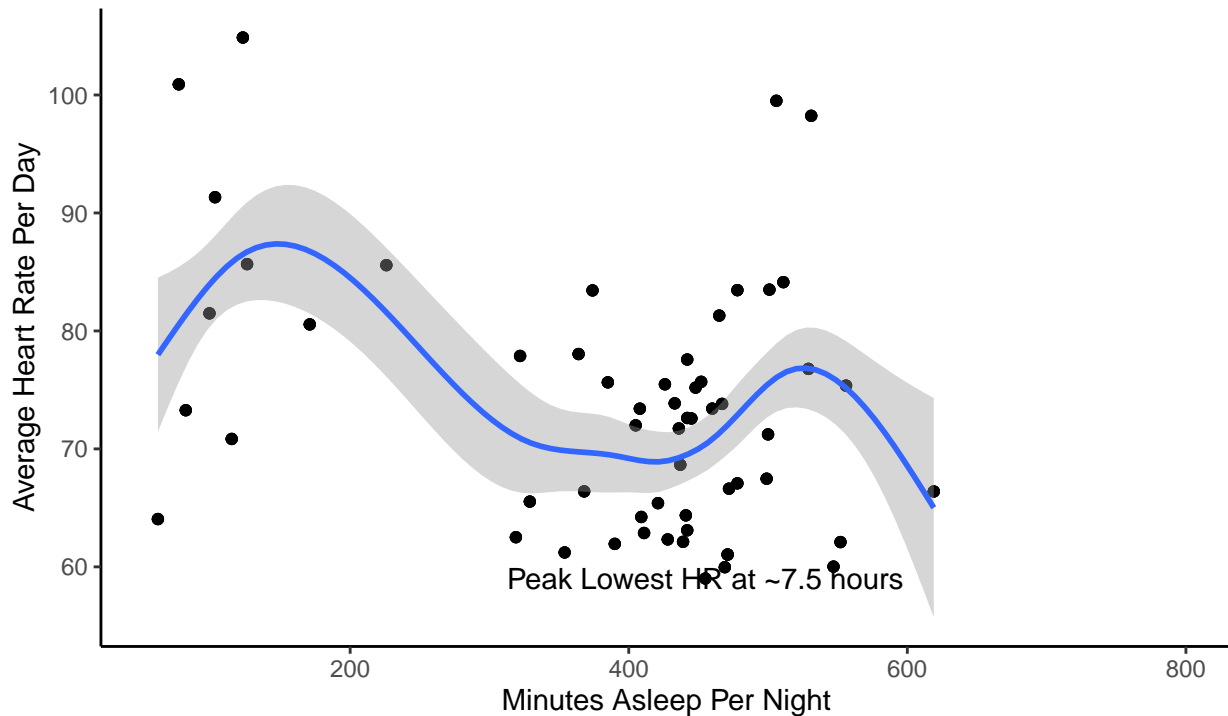
```
ggplot(data=Daily_All_Data) +
  geom_point(aes(x=TotalMinutesAsleep, y=Daily_Average_HR)) +
  geom_smooth(aes(x=TotalMinutesAsleep, y=Daily_Average_HR)) +
  labs(title="Correlation of Increased Sleep to Average Heart Rate Per Day", subtitle="Not all participants had heart rate data") +
  annotate("text", x=455, y=59, label="Peak Lowest HR at ~7.5 hours") +
  xlab('Minutes Asleep Per Night') + ylab('Average Heart Rate Per Day') +
  theme_classic()
```

More Sleep, Lower Heart Rate

```
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
## Warning: Removed 2643 rows containing non-finite values (stat_smooth).
## Warning: Removed 2643 rows containing missing values (geom_point).
```

Correlation of Increased Sleep to Average Heart Rate Per Day

Not all participants had heart rate data



CC0: Public Domain, survey through Amazon Mechanical Turk between 03.12.2016–05.12.2016

Peak lowest heart rate was around 7.5 hours of sleep.

Recommended Marketing Strategy Focus

Let's face it, people need a better way to decrease and cope with stress ([link](#)) and women are more likely to experience stress than men ([link](#)). According to Cleveland Clinic ([link here](#)), ways to reduce stress include exercising regularly and getting adequate sleep. The data from this Fitbit survey indicate that increased steps, activity, and sleep all lower heart rates and potentially lower stress. Bellabeat could run an analysis on its own data. Since Bellabeat does not currently track heart rate, analysis could be done between the correlation of activity, sleep, and stress. Even with this data, Bellabeat could focus one of their marketing strategies towards women with stress. By using the Bellabeat app, women who experience stress can watch as the amount of stress trends down as they focus on increasing activity level and sleep.