**Business Task**

My task was to analyze data provided by 30 eligible FitBit users. I would try to find trends and learn more about the different kind of behaviors of the users. I would then identify ways in which these trends and behaviors could help drive the market strategies of Bellabeat. Although, I would have to apply these strategies to only one of the products offered by Bellabeat.

**Data Sources**

The dataset was made up of multiple csv files. Most of these files were appended into one single csv file named “dailyActivity\_merged.csv”, so I decided to use this file for my analysis. I also used the “sleepDay\_merged.csv” file since the data found in this file was not found in the “dailyActivity.csv” file.

**Data Cleaning and Manipulation Documentation**

I removed the text “AM” from the entire “Date” column in sleepDay\_merged.csv to be able to use the database in BigQuery. Also, I removed duplicates using the “Remove duplicates” tool in Google Sheets. I also used the tool to remove any whitespaces but there was not any.

I decided that I would only use sleepDay\_merged.csv and dailyActivity\_merged.csv. Since both had the “Id” column with a unique ID for each user, I used BigQuery to inner join them. I used an inner join using the columns “Id” and “Date” as the fields.

I used the newly joined data and uploaded to Rstudio. Here, I used the count() function to see how many instances each user had. If the user had less than 5 instances I wrote the ID number on a notepad.

After documenting each ID number with less than 5 instances, I uploaded the data to Google sheets. Here, I used conditional formatting to find each instance for every ID number and manually deleted them. After I was done I was left with a total of 18 different users.

I then uploaded the modified data to Rstudio. I decided to use a random sample of five instances from each user. Some users had a lot more instances than the other users so I decided that five would be an ideal number to make sure I did not have any bias in my data. I took samples using the sample() function in R. Every instance in the sample was unique.

**Analysis Summary**

After performing my analysis, I identified a few trends and behaviors. First, I saw a correlation between the time an individual spent in bed with the minutes of sleep an individual obtained. As one increased, the other increased as well. From this, I could conclude that most of the individuals did not have any trouble going to sleep or obtaining sleep. Also, I was able to find that most individuals obtained from 6 hours of sleep to 10 hours of sleep daily. There was also a strong correlation between light active distance and total number of steps. In other words, as the light active distance increased the total number of steps also increased. What does this tell us? This tells us that users mostly performed light activities throughout the day. We also saw this kind of relationship with light active distance and sedentary minutes. Sedentary minutes are minutes spent sitting down or without moving from a single place. Although, in this relationship as sedentary minutes decreased, light active distance increased. Thus, most of the time people left their seat was to perform light activities. With these findings I was able to conclude that these users are normal individuals who might spend most of their day at work or might have constraints that keep them very occupied throughout the day. Nonetheless, these individuals care about being healthy.

**High-level Insights**

Based on the data, the average working-class individual should be our target audience. The best fit for this kind of individual would be the Leaf Bellabeat product. Leaf has great potential since it can be a bracelet, necklace, and even a clip. To get this product to appeal to individuals the product would have to constantly stay up to date with the latest fashion trends. To do this, the product should be able to have interchangeable parts. Doing this would allow individuals to customize their device in many ways. They would be able to customize it to be appropriate for any kind of setting. I believe that having a product that does not get in the way of everyday life would motivate people to want to know the insights of their health. Also, the product also must be practical to be able to be taken into any setting or scenario. Lastly, these users might have constraints that keep them from performing more intense activity. I believe more research is needed to identify if there are any constraints and if there is, identify what these constraints are. Identifying the types of constraints could heavily influence market strategies.