

# How Can a Wellness Technology Company Play it Smart?

## Google Data Analytics Capstone Project

Babou Clion Muhoza

### Introduction:

The Capstone Project is the final step towards completing the Google Data Analytics Professional Certificate. The following includes the steps taken during the analysis.

### Background:

Bellabeat is an upcoming small business with potential for higher success, the company engages in manufacturing high-tech health centered products for women. The purpose of the case study is to gain insights about trends in smart device usage so as to improve Bellabeat marketing strategy. Bellabeat's stakeholders believe that the analysis from Bellabeat devices data will be crucial for opportunity creation and growth.

The data source provided does not fulfill the requirement for a good data source. The data included is considered unreliable with only 30 participants. The source is not original since the data is from a third party provider. In addition, the data is both not current being more than 5 years old and poorly cited. Although significant insights can be extracted from the data, any recommendations would not be as effective considering the integrity of the data.

### Business Task:

Acquiring insights and trends from Bellabeat's customers data with the goal of providing reliable recommendations to Bellabeat's marketing strategy.

- The tool of choice for data cleaning, manipulation, analysis and finally visualization is RStudio.

### Processing Data:

I started with loading the different packages that will be needed to work with the data.

```
```{r}
if (!require('tidyverse'))
{
  install.packages('tidyverse');
  library(ggplot2);
}
```
```

```
```{r}
if (!require('janitor'))
{
  install.packages('janitor');
  library('janitor');
}
```
```

```
```{r}
if (!require('sqldf'))
{
  install.packages('sqldf');
  library("sqldf");
}
```
```

```
```{r}
if (!require('dplyr'))
{
  install.packages('dplyr');
  library(dplyr);
}
```
```

```
```{r}
if (!require('lubridate'))
{
  install.packages('lubridate');
  library(lubridate);
}
```
```

```
```{r}
if (!require('ggplot2'))
{
  install.packages('ggplot2');
  library(ggplot2);
}
```
```

```
```{r}
if (!require('skimr'))
{
  install.packages('skimr');
  library("skimr");
}
```
```

The first step was to install and load the required packages. The packages are to be used for analysis, the functions derived will be essential for data manipulation, visualization, cleaning, and utilizing queries.

Next, I installed the “readr” package for uploading and importing CSV files to Rstudio.

```
```{r}
if (!require('readr'))
{
  install.packages('readr');
  library(readr);
}
```
```

### Importing Data:

For the next step, I will be importing the different datasets to be used in the analysis process.

```
```{r}
day_activity <- read_csv("~/Documents/capstone_project/Fitabase Data 4.12.16-5.12.16/dailyActivity_merged.csv")
```
```{r}
daily_calories <- read_csv("~/Documents/capstone_project/Fitabase Data 4.12.16-5.12.16/dailyCalories_merged.csv")
```
```{r}
daily_intensities <- read_csv("~/Documents/capstone_project/Fitabase Data 4.12.16-5.12.16/dailyIntensities_merged.csv")
```
```{r}
daily_steps <- read_csv("~/Documents/capstone_project/Fitabase Data 4.12.16-5.12.16/dailySteps_merged.csv")
```
```{r}
daily_sleep <- read_csv("~/Documents/capstone_project/Fitabase Data 4.12.16-5.12.16/sleepDay_merged.csv")
```
```

```
```{r}
hourly_calories <- read_csv("~/Documents/capstone_project/Fitabase Data 4.12.16-5.12.16/hourlyCalories_merged.csv")
```
```{r}
hourly_intensities <- read_csv("~/Documents/capstone_project/Fitabase Data 4.12.16-5.12.16/hourlyIntensities_merged.csv")
```
```{r}
hourly_steps <- read_csv("~/Documents/capstone_project/Fitabase Data 4.12.16-5.12.16/hourlySteps_merged.csv")
```
```

```
```{r}
minute_steps <- read_csv("~/Documents/capstone_project/Fitabase Data 4.12.16-5.12.16/minuteStepsNarrow_merged.csv")
```
```{r}
minute_calories <- read_csv("~/Documents/capstone_project/Fitabase Data 4.12.16-5.12.16/minuteCaloriesNarrow_merged.csv")
```
```{r}
minute_intensities <- read_csv("~/Documents/capstone_project/Fitabase Data 4.12.16-5.12.16/minuteIntensitiesNarrow_merged.csv")
```
```{r}
minute_METs <- read_csv("~/Documents/capstone_project/Fitabase Data 4.12.16-5.12.16/minuteMETsNarrow_merged.csv")
```
```

```
```{r}
heartrate_seconds <- read_csv("~/Documents/capstone_project/Fitabase Data 4.12.16-5.12.16/heartrate_seconds_merged.csv")
```
```{r}
weight_log <- read.csv("~/Documents/capstone_project/Fitabase Data 4.12.16-5.12.16/weightLogInfo_merged.csv")
```
```

Among the provided datasets, those above were the ones chosen for evaluation considering the relevance of data and the company objectives.

## Data Manipulation and Cleaning:

In this step, I will be employing various measures and techniques to inspect the data for inconsistencies, nulls, and other kinds of errors in the datasets.

The reason behind using R above other data analysis tools like SQL and spreadsheets is because of the efficiency, visualizations and time management in general. The next phase of the analysis is to involve inspecting and reviewing the datasets with the goal of checking for consistency, duplicates, and other forms of insurances to verify the data integrity.

```
##{r}
colnames(daily_steps)
##

[1] "Id"          "ActivityDay" "StepTotal"
```

```
##{r}
colnames(daily_activity)
##

[1] "Id"          "Activity_Date"
[3] "Total_Steps" "Total_Distance"
[5] "Tracker_Distance" "Logged_Activities_Distance"
[7] "Very_Active_Distance" "Moderately_Active_Distance"
[9] "Light_Active_Distance" "Sedentary_Active_Distance"
[11] "Very_Active_Minutes" "Fairly_Active_Minutes"
[13] "Lightly_Active_Minutes" "Sedentary_Minutes"
[15] "Calories" "Total_Active_Minutes"
[17] "Total_Active_Hours" "Dates"
```

```
##{r}
colnames(daily_calories)
##

[1] "Id"          "ActivityDay" "Calories"
```

```
##{r}
colnames(daily_intensities)
##

[1] "Id"          "ActivityDay" "SedentaryMinutes"
[4] "LightlyActiveMinutes" "FairlyActiveMinutes" "VeryActiveMinutes"
[7] "SedentaryActiveDistance" "LightActiveDistance" "ModeratelyActiveDistance"
[10] "VeryActiveDistance"
```

```
##{r}
colnames(heartrate_seconds)
##

[1] "Id"          "Time" "Value"
```

```
'''{r}
colnames(hourly_calories)
'''
```

```
[1] "Id"          "ActivityHour" "Calories"
```

```
'''{r}
colnames(hourly_intensities)
'''
```

```
[1] "Id"          "ActivityHour"  "TotalIntensity" "AverageIntensity"
```

```
'''{r}
colnames(hourly_steps)
'''
```

```
[1] "Id"          "ActivityHour" "StepTotal"
```

```
'''{r}
colnames(minute_calories)
'''
```

```
[1] "Id"          "ActivityMinute" "Calories"
```

```
'''{r}
colnames(minute_intensities)
'''
```

```
[1] "Id"          "ActivityMinute" "Intensity"
```

```
'''{r}
colnames(minute_METs)
'''
```

```
[1] "Id"          "ActivityMinute" "METs"
```

```
'''{r}
colnames(weight_log)
'''
```

```
[1] "Id"          "Date"          "WeightKg"      "WeightPounds"  "Fat"
[6] "BMI"         "IsManualReport" "LogId"
```

```
'''{r}
colnames(minute_steps)
'''
```

```
[1] "Id"          "ActivityMinute" "Steps"
```

```

```{r}
sapply(list(daily_activity, daily_calories, daily_intensities, daily_steps, daily_sleep, hourly_calories,
hourly_intensities, hourly_steps, heartrate_seconds, weight_log), head)
```

```

| <b>Id</b><br><dbl> | <b>Activity_Date</b><br><chr> | <b>Total_Steps</b><br><dbl> | <b>Total_Distance</b><br><dbl> | <b>Tracker_Distance</b><br><dbl> | <b>Logged_Activities_Distance</b><br><dbl> |
|--------------------|-------------------------------|-----------------------------|--------------------------------|----------------------------------|--|
| 1503960366         | 4/12/2016                     | 13162                       | 8.50                           | 8.50                             | 0  |
| 1503960366         | 4/13/2016                     | 10735                       | 6.97                           | 6.97                             | 0  |
| 1503960366         | 4/14/2016                     | 10460                       | 6.74                           | 6.74                             | 0  |
| 1503960366         | 4/15/2016                     | 9762                        | 6.28                           | 6.28                             | 0  |
| 1503960366         | 4/16/2016                     | 12669                       | 8.16                           | 8.16                             | 0  |
| 1503960366         | 4/17/2016                     | 9705                        | 6.48                           | 6.48                             | 0  |

6 rows | 1-6 of 18 columns

| <b>Id</b><br><dbl> | <b>ActivityDay</b><br><chr> | <b>Calories</b><br><dbl> |
|--------------------|-----------------------------|--------------------------|
| 1503960366         | 4/12/2016                   | 1985                     |
| 1503960366         | 4/13/2016                   | 1797                     |
| 1503960366         | 4/14/2016                   | 1776                     |
| 1503960366         | 4/15/2016                   | 1745                     |
| 1503960366         | 4/16/2016                   | 1863                     |
| 1503960366         | 4/17/2016                   | 1728                     |

6 rows

| <b>Id</b><br><dbl> | <b>ActivityDay</b><br><chr> | <b>SedentaryMinutes</b><br><dbl> | <b>LightlyActiveMinutes</b><br><dbl> | <b>FairlyActiveMinutes</b><br><dbl> |
|--------------------|-----------------------------|----------------------------------|--------------------------------------|-------------------------------------|
| 1503960366         | 4/12/2016                   | 728                              | 328                                  | 13                                  |
| 1503960366         | 4/13/2016                   | 776                              | 217                                  | 19                                  |
| 1503960366         | 4/14/2016                   | 1218                             | 181                                  | 11                                  |
| 1503960366         | 4/15/2016                   | 726                              | 209                                  | 34                                  |
| 1503960366         | 4/16/2016                   | 773                              | 221                                  | 10                                  |
| 1503960366         | 4/17/2016                   | 539                              | 164                                  | 20                                  |

6 rows | 1-5 of 10 columns

| <b>Id</b><br><dbl> | <b>ActivityDay</b><br><chr> | <b>StepTotal</b><br><dbl> |
|--------------------|-----------------------------|---------------------------|
| 1503960366         | 4/12/2016                   | 13162                     |
| 1503960366         | 4/13/2016                   | 10735                     |
| 1503960366         | 4/14/2016                   | 10460                     |
| 1503960366         | 4/15/2016                   | 9762                      |
| 1503960366         | 4/16/2016                   | 12669                     |
| 1503960366         | 4/17/2016                   | 9705                      |

6 rows

|  | <b>Id</b><br><dbl> | <b>SleepDay</b><br><chr> | <b>TotalSleepRecords</b><br><dbl> | <b>TotalMinutesAsleep</b><br><dbl> | <b>TotalTimeInBed</b><br><dbl> |
|--|--------------------|--------------------------|-----------------------------------|------------------------------------|--------------------------------|
|  | 1503960366         | 4/12/2016 12:00:00 AM    | 1                                 | 327                                | 346                            |
|  | 1503960366         | 4/13/2016 12:00:00 AM    | 2                                 | 384                                | 407                            |
|  | 1503960366         | 4/15/2016 12:00:00 AM    | 1                                 | 412                                | 442                            |
|  | 1503960366         | 4/16/2016 12:00:00 AM    | 2                                 | 340                                | 367                            |
|  | 1503960366         | 4/17/2016 12:00:00 AM    | 1                                 | 700                                | 712                            |
|  | 1503960366         | 4/19/2016 12:00:00 AM    | 1                                 | 304                                | 320                            |

6 rows

|  | <b>Id</b><br><dbl> | <b>ActivityHour</b><br><chr> | <b>Calories</b><br><dbl> |
|--|--------------------|------------------------------|--------------------------|
|  | 1503960366         | 4/12/2016 12:00:00 AM        | 81                       |
|  | 1503960366         | 4/12/2016 1:00:00 AM         | 61                       |
|  | 1503960366         | 4/12/2016 2:00:00 AM         | 59                       |
|  | 1503960366         | 4/12/2016 3:00:00 AM         | 47                       |
|  | 1503960366         | 4/12/2016 4:00:00 AM         | 48                       |
|  | 1503960366         | 4/12/2016 5:00:00 AM         | 48                       |

6 rows

|  | <b>Id</b><br><dbl> | <b>ActivityHour</b><br><chr> | <b>TotalIntensity</b><br><dbl> | <b>AverageIntensity</b><br><dbl> |
|--|--------------------|------------------------------|--------------------------------|----------------------------------|
|  | 1503960366         | 4/12/2016 12:00:00 AM        | 20                             | 0.333333                         |
|  | 1503960366         | 4/12/2016 1:00:00 AM         | 8                              | 0.133333                         |
|  | 1503960366         | 4/12/2016 2:00:00 AM         | 7                              | 0.116667                         |
|  | 1503960366         | 4/12/2016 3:00:00 AM         | 0                              | 0.000000                         |
|  | 1503960366         | 4/12/2016 4:00:00 AM         | 0                              | 0.000000                         |
|  | 1503960366         | 4/12/2016 5:00:00 AM         | 0                              | 0.000000                         |

6 rows

|  | <b>Id</b><br><dbl> | <b>ActivityHour</b><br><chr> | <b>StepTotal</b><br><dbl> |
|--|--------------------|------------------------------|---------------------------|
|  | 1503960366         | 4/12/2016 12:00:00 AM        | 373                       |
|  | 1503960366         | 4/12/2016 1:00:00 AM         | 160                       |
|  | 1503960366         | 4/12/2016 2:00:00 AM         | 151                       |
|  | 1503960366         | 4/12/2016 3:00:00 AM         | 0                         |
|  | 1503960366         | 4/12/2016 4:00:00 AM         | 0                         |
|  | 1503960366         | 4/12/2016 5:00:00 AM         | 0                         |

6 rows

|  | <b>Id</b><br><dbl> | <b>Time</b><br><chr> | <b>Value</b><br><dbl> |
|--|--------------------|----------------------|-----------------------|
|  | 2022484408         | 4/12/2016 7:21:00 AM | 97                    |
|  | 2022484408         | 4/12/2016 7:21:05 AM | 102                   |
|  | 2022484408         | 4/12/2016 7:21:10 AM | 105                   |
|  | 2022484408         | 4/12/2016 7:21:20 AM | 103                   |
|  | 2022484408         | 4/12/2016 7:21:25 AM | 101                   |
|  | 2022484408         | 4/12/2016 7:22:05 AM | 95                    |

6 rows

|   | <b>Id</b><br><dbl> | <b>Date</b><br><chr>  | <b>WeightKg</b><br><dbl> | <b>WeightPounds</b><br><dbl> | <b>Fat</b><br><int> | <b>BMI</b><br><dbl> | <b>IsManualReport</b><br><chr> | <b>LogId</b><br><dbl> |
|---|--------------------|-----------------------|--------------------------|------------------------------|---------------------|---------------------|--------------------------------|-----------------------|
| 1 | 1503960366         | 5/2/2016 11:59:59 PM  | 52.6                     | 115.9631                     | 22                  | 22.65               | True                           | 1.462234e+12          |
| 2 | 1503960366         | 5/3/2016 11:59:59 PM  | 52.6                     | 115.9631                     | NA                  | 22.65               | True                           | 1.462320e+12          |
| 3 | 1927972279         | 4/13/2016 1:08:52 AM  | 133.5                    | 294.3171                     | NA                  | 47.54               | False                          | 1.460510e+12          |
| 4 | 2873212765         | 4/21/2016 11:59:59 PM | 56.7                     | 125.0021                     | NA                  | 21.45               | True                           | 1.461283e+12          |
| 5 | 2873212765         | 5/12/2016 11:59:59 PM | 57.3                     | 126.3249                     | NA                  | 21.69               | True                           | 1.463098e+12          |
| 6 | 4319703577         | 4/17/2016 11:59:59 PM | 72.4                     | 159.6147                     | 25                  | 27.45               | True                           | 1.460938e+12          |

```

```{r}
sapply(list(daily_activity, daily_calories, daily_intensities, daily_steps, daily_sleep, hourly_calories,
hourly_intensities, hourly_steps, heartrate_seconds, weight_log), glimpse)
```

```

| <b>Id</b><br><dbl> | <b>Activity_Date</b><br><chr> | <b>Total_Steps</b><br><dbl> | <b>Total_Distance</b><br><dbl> | <b>Tracker_Distance</b><br><dbl> | <b>Logged_Activities_Distance</b><br><dbl> |
|--------------------|-------------------------------|-----------------------------|--------------------------------|----------------------------------|--|
| 1503960366         | 4/12/2016                     | 13162                       | 8.50                           | 8.50                             | 0.000000                                   |
| 1503960366         | 4/13/2016                     | 10735                       | 6.97                           | 6.97                             | 0.000000                                   |
| 1503960366         | 4/14/2016                     | 10460                       | 6.74                           | 6.74                             | 0.000000                                   |
| 1503960366         | 4/15/2016                     | 9762                        | 6.28                           | 6.28                             | 0.000000                                   |
| 1503960366         | 4/16/2016                     | 12669                       | 8.16                           | 8.16                             | 0.000000                                   |
| 1503960366         | 4/17/2016                     | 9705                        | 6.48                           | 6.48                             | 0.000000                                   |
| 1503960366         | 4/18/2016                     | 13019                       | 8.59                           | 8.59                             | 0.000000                                   |
| 1503960366         | 4/19/2016                     | 15506                       | 9.88                           | 9.88                             | 0.000000                                   |
| 1503960366         | 4/20/2016                     | 10544                       | 6.68                           | 6.68                             | 0.000000                                   |
| 1503960366         | 4/21/2016                     | 9819                        | 6.34                           | 6.34                             | 0.000000                                   |

1-10 of 940 rows | 1-6 of 18 columns

Previous **1** 2 3 4 5 6 ... 94 Next

| <b>Id</b><br><dbl> | <b>ActivityDay</b><br><chr> | <b>Calories</b><br><dbl> |
|--------------------|-----------------------------|--------------------------|
| 1503960366         | 4/12/2016                   | 1985                     |
| 1503960366         | 4/13/2016                   | 1797                     |
| 1503960366         | 4/14/2016                   | 1776                     |
| 1503960366         | 4/15/2016                   | 1745                     |
| 1503960366         | 4/16/2016                   | 1863                     |
| 1503960366         | 4/17/2016                   | 1728                     |
| 1503960366         | 4/18/2016                   | 1921                     |
| 1503960366         | 4/19/2016                   | 2035                     |
| 1503960366         | 4/20/2016                   | 1786                     |
| 1503960366         | 4/21/2016                   | 1775                     |

1-10 of 940 rows

Previous **1** 2 3 4 5 6 ... 94 Next

| <b>Id</b><br><dbl> | <b>ActivityDay</b><br><chr> | <b>SedentaryMinutes</b><br><dbl> | <b>LightlyActiveMinutes</b><br><dbl> | <b>FairlyActiveMinutes</b><br><dbl> |
|--------------------|-----------------------------|----------------------------------|--------------------------------------|-------------------------------------|
| 1503960366         | 4/12/2016                   | 728                              | 328                                  | 13                                  |
| 1503960366         | 4/13/2016                   | 776                              | 217                                  | 19                                  |
| 1503960366         | 4/14/2016                   | 1218                             | 181                                  | 11                                  |
| 1503960366         | 4/15/2016                   | 726                              | 209                                  | 34                                  |
| 1503960366         | 4/16/2016                   | 773                              | 221                                  | 10                                  |
| 1503960366         | 4/17/2016                   | 539                              | 164                                  | 20                                  |
| 1503960366         | 4/18/2016                   | 1149                             | 233                                  | 16                                  |
| 1503960366         | 4/19/2016                   | 775                              | 264                                  | 31                                  |
| 1503960366         | 4/20/2016                   | 818                              | 205                                  | 12                                  |
| 1503960366         | 4/21/2016                   | 838                              | 211                                  | 8                                   |

1-10 of 940 rows | 1-5 of 10 columns

Previous **1** 2 3 4 5 6 ... 94 Next



| <b>Id</b><br><dbl> | <b>ActivityDay</b><br><chr> | <b>StepTotal</b><br><dbl> |
|--------------------|-----------------------------|---------------------------|
| 1503960366         | 4/12/2016                   | 13162                     |
| 1503960366         | 4/13/2016                   | 10735                     |
| 1503960366         | 4/14/2016                   | 10460                     |
| 1503960366         | 4/15/2016                   | 9762                      |
| 1503960366         | 4/16/2016                   | 12669                     |
| 1503960366         | 4/17/2016                   | 9705                      |
| 1503960366         | 4/18/2016                   | 13019                     |
| 1503960366         | 4/19/2016                   | 15506                     |
| 1503960366         | 4/20/2016                   | 10544                     |
| 1503960366         | 4/21/2016                   | 9819                      |

1-10 of 940 rows

Previous [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) ... [94](#) Next

| <b>Id</b><br><dbl> | <b>SleepDay</b><br><chr> | <b>TotalSleepRecords</b><br><dbl> | <b>TotalMinutesAsleep</b><br><dbl> | <b>TotalTimeInBed</b><br><dbl> |
|--------------------|--------------------------|-----------------------------------|------------------------------------|--------------------------------|
| 1503960366         | 4/12/2016 12:00:00 AM    | 1                                 | 327                                | 346                            |
| 1503960366         | 4/13/2016 12:00:00 AM    | 2                                 | 384                                | 407                            |
| 1503960366         | 4/15/2016 12:00:00 AM    | 1                                 | 412                                | 442                            |
| 1503960366         | 4/16/2016 12:00:00 AM    | 2                                 | 340                                | 367                            |
| 1503960366         | 4/17/2016 12:00:00 AM    | 1                                 | 700                                | 712                            |
| 1503960366         | 4/19/2016 12:00:00 AM    | 1                                 | 304                                | 320                            |
| 1503960366         | 4/20/2016 12:00:00 AM    | 1                                 | 360                                | 377                            |
| 1503960366         | 4/21/2016 12:00:00 AM    | 1                                 | 325                                | 364                            |
| 1503960366         | 4/23/2016 12:00:00 AM    | 1                                 | 361                                | 384                            |
| 1503960366         | 4/24/2016 12:00:00 AM    | 1                                 | 430                                | 449                            |

1-10 of 413 rows

Previous [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) ... [42](#) Next

| <b>Id</b><br><dbl> | <b>ActivityHour</b><br><chr> | <b>Calories</b><br><dbl> |
|--------------------|------------------------------|--------------------------|
| 1503960366         | 4/12/2016 12:00:00 AM        | 81                       |
| 1503960366         | 4/12/2016 1:00:00 AM         | 61                       |
| 1503960366         | 4/12/2016 2:00:00 AM         | 59                       |
| 1503960366         | 4/12/2016 3:00:00 AM         | 47                       |
| 1503960366         | 4/12/2016 4:00:00 AM         | 48                       |
| 1503960366         | 4/12/2016 5:00:00 AM         | 48                       |
| 1503960366         | 4/12/2016 6:00:00 AM         | 48                       |
| 1503960366         | 4/12/2016 7:00:00 AM         | 47                       |
| 1503960366         | 4/12/2016 8:00:00 AM         | 68                       |
| 1503960366         | 4/12/2016 9:00:00 AM         | 141                      |

1-10 of 22,099 rows

Previous [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) ... [100](#) Next

| <b>Id</b><br><dbl> | <b>ActivityHour</b><br><chr> | <b>TotalIntensity</b><br><dbl> | <b>AverageIntensity</b><br><dbl> |
|--------------------|------------------------------|--------------------------------|----------------------------------|
| 1503960366         | 4/12/2016 12:00:00 AM        | 20                             | 0.333333                         |
| 1503960366         | 4/12/2016 1:00:00 AM         | 8                              | 0.133333                         |
| 1503960366         | 4/12/2016 2:00:00 AM         | 7                              | 0.116667                         |
| 1503960366         | 4/12/2016 3:00:00 AM         | 0                              | 0.000000                         |
| 1503960366         | 4/12/2016 4:00:00 AM         | 0                              | 0.000000                         |
| 1503960366         | 4/12/2016 5:00:00 AM         | 0                              | 0.000000                         |
| 1503960366         | 4/12/2016 6:00:00 AM         | 0                              | 0.000000                         |
| 1503960366         | 4/12/2016 7:00:00 AM         | 0                              | 0.000000                         |
| 1503960366         | 4/12/2016 8:00:00 AM         | 13                             | 0.216667                         |
| 1503960366         | 4/12/2016 9:00:00 AM         | 30                             | 0.500000                         |

1-10 of 22,099 rows

Previous [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) ... [100](#) Next

| <b>Id</b><br><dbl> | <b>ActivityHour</b><br><chr> | <b>StepTotal</b><br><dbl> |
|--------------------|------------------------------|---------------------------|
| 1503960366         | 4/12/2016 12:00:00 AM        | 373                       |
| 1503960366         | 4/12/2016 1:00:00 AM         | 160                       |
| 1503960366         | 4/12/2016 2:00:00 AM         | 151                       |
| 1503960366         | 4/12/2016 3:00:00 AM         | 0                         |
| 1503960366         | 4/12/2016 4:00:00 AM         | 0                         |
| 1503960366         | 4/12/2016 5:00:00 AM         | 0                         |
| 1503960366         | 4/12/2016 6:00:00 AM         | 0                         |
| 1503960366         | 4/12/2016 7:00:00 AM         | 0                         |
| 1503960366         | 4/12/2016 8:00:00 AM         | 250                       |
| 1503960366         | 4/12/2016 9:00:00 AM         | 1864                      |

1-10 of 22,099 rows

Previous  2 3 4 5 6 ... 100 Next

| <b>Id</b><br><dbl> | <b>Time</b><br><chr> | <b>Value</b><br><dbl> |
|--------------------|----------------------|-----------------------|
| 2022484408         | 4/12/2016 7:21:00 AM | 97                    |
| 2022484408         | 4/12/2016 7:21:05 AM | 102                   |
| 2022484408         | 4/12/2016 7:21:10 AM | 105                   |
| 2022484408         | 4/12/2016 7:21:20 AM | 103                   |
| 2022484408         | 4/12/2016 7:21:25 AM | 101                   |
| 2022484408         | 4/12/2016 7:22:05 AM | 95                    |
| 2022484408         | 4/12/2016 7:22:10 AM | 91                    |
| 2022484408         | 4/12/2016 7:22:15 AM | 93                    |
| 2022484408         | 4/12/2016 7:22:20 AM | 94                    |
| 2022484408         | 4/12/2016 7:22:25 AM | 93                    |

1-10 of 2,483,658 rows

Previous  2 3 4 5 6 ... 100 Next

| <b>Id</b><br><dbl> | <b>Date</b><br><chr>  | <b>WeightKg</b><br><dbl> | <b>WeightPounds</b><br><dbl> | <b>Fat</b><br><int> | <b>BMI</b><br><dbl> | <b>IsManualReport</b><br><chr> | <b>LogId</b><br><dbl> |
|--------------------|-----------------------|--------------------------|------------------------------|---------------------|---------------------|--------------------------------|-----------------------|
| 1503960366         | 5/2/2016 11:59:59 PM  | 52.6                     | 115.9631                     | 22                  | 22.65               | True                           | 1.462234e+12          |
| 1503960366         | 5/3/2016 11:59:59 PM  | 52.6                     | 115.9631                     | NA                  | 22.65               | True                           | 1.462320e+12          |
| 1927972279         | 4/13/2016 1:08:52 AM  | 133.5                    | 294.3171                     | NA                  | 47.54               | False                          | 1.460510e+12          |
| 2873212765         | 4/21/2016 11:59:59 PM | 56.7                     | 125.0021                     | NA                  | 21.45               | True                           | 1.461283e+12          |
| 2873212765         | 5/12/2016 11:59:59 PM | 57.3                     | 126.3249                     | NA                  | 21.69               | True                           | 1.463098e+12          |
| 4319703577         | 4/17/2016 11:59:59 PM | 72.4                     | 159.6147                     | 25                  | 27.45               | True                           | 1.460938e+12          |
| 4319703577         | 5/4/2016 11:59:59 PM  | 72.3                     | 159.3942                     | NA                  | 27.38               | True                           | 1.462406e+12          |
| 4558609924         | 4/18/2016 11:59:59 PM | 69.7                     | 153.6622                     | NA                  | 27.25               | True                           | 1.461024e+12          |
| 4558609924         | 4/25/2016 11:59:59 PM | 70.3                     | 154.9850                     | NA                  | 27.46               | True                           | 1.461629e+12          |
| 4558609924         | 5/1/2016 11:59:59 PM  | 69.9                     | 154.1031                     | NA                  | 27.32               | True                           | 1.462147e+12          |

1-10 of 67 rows

Previous  2 3 4 5 6 7 Next

```

```{r}
sapply(list(daily_activity, daily_calories, daily_intensities, daily_steps, daily_sleep, hourly_calories,
hourly_intensities, hourly_steps, heartrate_seconds, weight_log), str)
```

```

```

spc_tbl_ [940 × 18] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
 $ Id                : num [1:940] 1.5e+09 1.5e+09 1.5e+09 1.5e+09 1.5e+09 ...
 $ Activity_Date      : chr [1:940] "4/12/2016" "4/13/2016" "4/14/2016" "4/15/2016" ...
 $ Total_Steps        : num [1:940] 13162 10735 10460 9762 12669 ...
 $ Total_Distance     : num [1:940] 8.5 6.97 6.74 6.28 8.16 ...
 $ Tracker_Distance   : num [1:940] 8.5 6.97 6.74 6.28 8.16 ...
 $ Logged_Activities_Distance: num [1:940] 0 0 0 0 0 0 0 0 0 ...
 $ Very_Active_Distance : num [1:940] 1.88 1.57 2.44 2.14 2.71 ...
 $ Moderately_Active_Distance: num [1:940] 0.55 0.69 0.4 1.26 0.41 ...
 $ Light_Active_Distance : num [1:940] 6.06 4.71 3.91 2.83 5.04 ...
 $ Sedentary_Active_Distance : num [1:940] 0 0 0 0 0 0 0 0 0 ...
 $ Very_Active_Minutes : num [1:940] 25 21 30 29 36 38 42 50 28 19 ...
 $ Fairly_Active_Minutes : num [1:940] 13 19 11 34 10 20 16 31 12 8 ...
 $ Lightly_Active_Minutes : num [1:940] 328 217 181 209 221 164 233 264 205 211 ...
 $ Sedentary_Minutes    : num [1:940] 728 776 1218 726 773 ...
 $ Calories             : num [1:940] 1985 1797 1776 1745 1863 ...
 $ Total_Active_Minutes : num [1:940] 1094 1033 1440 998 1040 ...
 $ Total_Active_Hours   : num [1:940] 18 17 24 17 17 13 24 19 18 18 ...
 $ Dates               : Date[1:940], format: "2016-04-12" "2016-04-13" ...
- attr(*, "spec")=
.. cols(
..   Id = col_double(),
..   ActivityDate = col_character(),
..   TotalSteps = col_double(),
..   TotalDistance = col_double(),
..   TrackerDistance = col_double(),
..   LoggedActivitiesDistance = col_double(),
..   VeryActiveDistance = col_double(),
..   ModeratelyActiveDistance = col_double(),
..   LightActiveDistance = col_double(),
..   SedentaryActiveDistance = col_double(),
..   VeryActiveMinutes = col_double(),
..   FairlyActiveMinutes = col_double(),
..   LightlyActiveMinutes = col_double(),
..   SedentaryMinutes = col_double(),
..   Calories = col_double()
.. )
- attr(*, "problems")=<externalptr>

```

```

spc_tbl_ [940 × 3] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
 $ Id                : num [1:940] 1.5e+09 1.5e+09 1.5e+09 1.5e+09 1.5e+09 ...
 $ ActivityDay       : chr [1:940] "4/12/2016" "4/13/2016" "4/14/2016" "4/15/2016" ...
 $ Calories          : num [1:940] 1985 1797 1776 1745 1863 ...
- attr(*, "spec")=
.. cols(
..   Id = col_double(),
..   ActivityDay = col_character(),
..   Calories = col_double()
.. )
- attr(*, "problems")=<externalptr>

```

```

spec_tbl_ [940 × 10] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
 $ Id          : num [1:940] 1.5e+09 1.5e+09 1.5e+09 1.5e+09 1.5e+09 ...
 $ ActivityDay : chr [1:940] "4/12/2016" "4/13/2016" "4/14/2016" "4/15/2016" ...
 $ SedentaryMinutes : num [1:940] 728 776 1218 726 773 ...
 $ LightlyActiveMinutes : num [1:940] 328 217 181 209 221 164 233 264 205 211 ...
 $ FairlyActiveMinutes : num [1:940] 13 19 11 34 10 20 16 31 12 8 ...
 $ VeryActiveMinutes : num [1:940] 25 21 30 29 36 38 42 50 28 19 ...
 $ SedentaryActiveDistance : num [1:940] 0 0 0 0 0 0 0 0 0 ...
 $ LightActiveDistance : num [1:940] 6.06 4.71 3.91 2.83 5.04 ...
 $ ModeratelyActiveDistance: num [1:940] 0.55 0.69 0.4 1.26 0.41 ...
 $ VeryActiveDistance : num [1:940] 1.88 1.57 2.44 2.14 2.71 ...
- attr(*, "spec")=
.. cols(
..   Id = col_double(),
..   ActivityDay = col_character(),
..   SedentaryMinutes = col_double(),
..   LightlyActiveMinutes = col_double(),
..   FairlyActiveMinutes = col_double(),
..   VeryActiveMinutes = col_double(),
..   SedentaryActiveDistance = col_double(),
..   LightActiveDistance = col_double(),
..   ModeratelyActiveDistance = col_double(),
..   VeryActiveDistance = col_double()
.. )
- attr(*, "problems")=<externalptr>

```

```

spec_tbl_ [940 × 3] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
 $ Id          : num [1:940] 1.5e+09 1.5e+09 1.5e+09 1.5e+09 1.5e+09 ...
 $ ActivityDay: chr [1:940] "4/12/2016" "4/13/2016" "4/14/2016" "4/15/2016" ...
 $ StepTotal  : num [1:940] 13162 10735 10460 9762 12669 ...
- attr(*, "spec")=
.. cols(
..   Id = col_double(),
..   ActivityDay = col_character(),
..   StepTotal = col_double()
.. )
- attr(*, "problems")=<externalptr>

```

```

spec_tbl_ [413 × 5] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
 $ Id          : num [1:413] 1.5e+09 1.5e+09 1.5e+09 1.5e+09 1.5e+09 ...
 $ SleepDay    : chr [1:413] "4/12/2016 12:00:00 AM" "4/13/2016 12:00:00 AM" "4/15/2016 12:00:00 AM"
"4/16/2016 12:00:00 AM" ...
 $ TotalSleepRecords : num [1:413] 1 2 1 2 1 1 1 1 1 ...
 $ TotalMinutesAsleep: num [1:413] 327 384 412 340 700 304 360 325 361 430 ...
 $ TotalTimeInBed : num [1:413] 346 407 442 367 712 320 377 364 384 449 ...
- attr(*, "spec")=
.. cols(
..   Id = col_double(),
..   SleepDay = col_character(),
..   TotalSleepRecords = col_double(),
..   TotalMinutesAsleep = col_double(),
..   TotalTimeInBed = col_double()
.. )
- attr(*, "problems")=<externalptr>

```

```

spc_tbl_ [22,099 x 3] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
 $ Id      : num [1:22099] 1.5e+09 1.5e+09 1.5e+09 1.5e+09 1.5e+09 ...
 $ ActivityHour: chr [1:22099] "4/12/2016 12:00:00 AM" "4/12/2016 1:00:00 AM" "4/12/2016 2:00:00 AM" "4/12/2016
3:00:00 AM" ...
 $ Calories   : num [1:22099] 81 61 59 47 48 48 48 47 68 141 ...
- attr(*, "spec")=
.. cols(
..   Id = col_double(),
..   ActivityHour = col_character(),
..   Calories = col_double()
.. )
- attr(*, "problems")=<externalptr>

```

```

spc_tbl_ [22,099 x 4] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
 $ Id      : num [1:22099] 1.5e+09 1.5e+09 1.5e+09 1.5e+09 1.5e+09 ...
 $ ActivityHour : chr [1:22099] "4/12/2016 12:00:00 AM" "4/12/2016 1:00:00 AM" "4/12/2016 2:00:00 AM" "4/12/2016
3:00:00 AM" ...
 $ TotalIntensity : num [1:22099] 20 8 7 0 0 0 0 0 13 30 ...
 $ AverageIntensity: num [1:22099] 0.333 0.133 0.117 0 0 ...
- attr(*, "spec")=
.. cols(
..   Id = col_double(),
..   ActivityHour = col_character(),
..   TotalIntensity = col_double(),
..   AverageIntensity = col_double()
.. )
- attr(*, "problems")=<externalptr>


































































```

```

spc_tbl_ [22,099 x 3] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
 $ Id      : num [1:22099] 1.5e+09 1.5e+09 1.5e+09 1.5e+09 1.5e+09 ...
 $ ActivityHour: chr [1:22099] "4/12/2016 12:00:00 AM" "4/12/2016 1:00:00 AM" "4/12/2016 2:00:00 AM" "4/12/2016
3:00:00 AM" ...
 $ StepTotal   : num [1:22099] 373 160 151 0 0 ...
- attr(*, "spec")=
.. cols(
..   Id = col_double(),
..   ActivityHour = col_character(),
..   StepTotal = col_double()
.. )
- attr(*, "problems")=<externalptr>

```

## Checking for Nulls and N/A(s).

|  |   |
|--|---|
| <pre>```{r} sum(is.na(daily_activity)) ```</pre>     |          |
| [1] 0  |          |
| <pre>```{r} sum(is.na(daily_calories)) ```</pre>     |          |
| [1] 0  |          |
| <pre>```{r} sum(is.na(daily_steps)) ```</pre>        |          |
| [1] 0  |          |
| <pre>```{r} sum(is.na(daily_intensities)) ```</pre>  |          |
| [1] 0  |          |
| <pre>```{r} sum(is.na(hourly_calories)) ```</pre>    |          |
| [1] 0  |          |
| <pre>```{r} sum(is.na(hourly_intensities)) ```</pre> |    |
| [1] 0  |    |
| <pre>```{r} sum(is.na(hourly_steps)) ```</pre>       |    |
| [1] 0  |    |
| <pre>```{r} sum(is.na(daily_sleep)) ```</pre>        |    |
| [1] 0  |    |
| <pre>```{r} sum(is.na(minute_steps)) ```</pre>       |    |
| [1] 0  |    |
| <pre>```{r} sum(is.na(minute_calories)) ```</pre>    |    |
| [1] 0  |    |
| <pre>```{r} sum(is.na(minute_intensities)) ```</pre> |    |
| [1] 0  |    |

```

{r}
sum(is.na(minute_METs))

```

```
[1] 0
```

```

{r}
sum(is.na(heartrate_seconds))

```

```
[1] 0
```

```

{r}
sum(is.na(weight_log))

```

```
[1] 65
```

The results show that there are 65 missing values in weight\_log dataframe. After close examination, I found out that the missing values are from the "fat" observation. This lead me to delete the column as it provided no insights thus irrelevant to the analysis.

## Data Analysis:

```

{r}
summary(daily_activity)

```

| Id   | Activity_Date    | Total_Steps        | Total_Distance | Tracker_Distance |
|--|------------------|--------------------|----------------|------------------|
| Min. :1.504e+09  | Length:940       | Min. : 0           | Min. : 0.000   | Min. : 0.000     |
| 1st Qu.:2.320e+09  | Class :character | 1st Qu.: 3790      | 1st Qu.: 2.620 | 1st Qu.: 2.620   |
| Median :4.445e+09  | Mode :character  | Median : 7406      | Median : 5.245 | Median : 5.245   |
| Mean :4.855e+09  |                  | Mean : 7638        | Mean : 5.490   | Mean : 5.475     |
| 3rd Qu.:6.962e+09  |                  | 3rd Qu.:10727      | 3rd Qu.: 7.713 | 3rd Qu.: 7.710   |
| Max. :8.878e+09  |                  | Max. :36019        | Max. :28.030   | Max. :28.030     |
| Logged_Activities_Distance Very_Active_Distance Moderately_Active_Distance |                  |                    |                |                  |
| Min. :0.0000   | Min. : 0.000     | Min. :0.0000       |                |                  |
| 1st Qu.:0.0000   | 1st Qu.: 0.000   | 1st Qu.:0.0000     |                |                  |
| Median :0.0000   | Median : 0.210   | Median :0.2400     |                |                  |
| Mean :0.1082   | Mean : 1.503     | Mean :0.5675       |                |                  |
| 3rd Qu.:0.0000   | 3rd Qu.: 2.053   | 3rd Qu.:0.8000     |                |                  |
| Max. :4.9421   | Max. :21.920     | Max. :6.4800       |                |                  |
| Light_Active_Distance Sedentary_Active_Distance Very_Active_Minutes        |                  |                    |                |                  |
| Min. : 0.000   | Min. :0.000000   | Min. : 0.00        |                |                  |
| 1st Qu.: 1.945   | 1st Qu.:0.000000 | 1st Qu.: 0.00      |                |                  |
| Median : 3.365   | Median :0.000000 | Median : 4.00      |                |                  |
| Mean : 3.341   | Mean :0.001606   | Mean : 21.16       |                |                  |
| 3rd Qu.: 4.782   | 3rd Qu.:0.000000 | 3rd Qu.: 32.00     |                |                  |
| Max. :10.710   | Max. :0.110000   | Max. :210.00       |                |                  |
| Fairly_Active_Minutes Lightly_Active_Minutes Sedentary_Minutes Calories    |                  |                    |                |                  |
| Min. : 0.00  | Min. : 0.0       | Min. : 0.0         | Min. : 0       |                  |
| 1st Qu.: 0.00  | 1st Qu.:127.0    | 1st Qu.: 729.8     | 1st Qu.:1828   |                  |
| Median : 6.00  | Median :199.0    | Median :1057.5     | Median :2134   |                  |
| Mean : 13.56   | Mean :192.8      | Mean : 991.2       | Mean :2304     |                  |
| 3rd Qu.: 19.00   | 3rd Qu.:264.0    | 3rd Qu.:1229.5     | 3rd Qu.:2793   |                  |
| Max. :143.00   | Max. :518.0      | Max. :1440.0       | Max. :4900     |                  |
| Total_Active_Minutes Total_Active_Hours Dates                              |                  |                    |                |                  |
| Min. : 2.0   | Min. : 0.00      | Min. :2016-04-12   |                |                  |
| 1st Qu.: 989.8   | 1st Qu.:16.00    | 1st Qu.:2016-04-19 |                |                  |
| Median :1440.0   | Median :24.00    | Median :2016-04-26 |                |                  |
| Mean :1218.8   | Mean :20.31      | Mean :2016-04-26   |                |                  |
| 3rd Qu.:1440.0   | 3rd Qu.:24.00    | 3rd Qu.:2016-05-04 |                |                  |
| Max. :1440.0   | Max. :24.00      | Max. :2016-05-12   |                |                  |

```
summary(daily_calories)
```

| Id                | ActivityDay      | Calories     |
|-------------------|------------------|--------------|
| Min. :1.504e+09   | Length:940       | Min. : 0     |
| 1st Qu.:2.320e+09 | Class :character | 1st Qu.:1828 |
| Median :4.445e+09 | Mode :character  | Median :2134 |
| Mean :4.855e+09   |                  | Mean :2304   |
| 3rd Qu.:6.962e+09 |                  | 3rd Qu.:2793 |
| Max. :8.878e+09   |                  | Max. :4900   |

```
summary(daily_intensities)
```

| Id                       | ActivityDay        | SedentaryMinutes        | LightlyActiveMinutes |
|--------------------------|--------------------|-------------------------|----------------------|
| Min. :1.504e+09          | Length:940         | Min. : 0.0              | Min. : 0.0           |
| 1st Qu.:2.320e+09        | Class :character   | 1st Qu.: 729.8          | 1st Qu.:127.0        |
| Median :4.445e+09        | Mode :character    | Median :1057.5          | Median :199.0        |
| Mean :4.855e+09          |                    | Mean : 991.2            | Mean :192.8          |
| 3rd Qu.:6.962e+09        |                    | 3rd Qu.:1229.5          | 3rd Qu.:264.0        |
| Max. :8.878e+09          |                    | Max. :1440.0            | Max. :518.0          |
| FairlyActiveMinutes      | VeryActiveMinutes  | SedentaryActiveDistance | LightActiveDistance  |
| Min. : 0.00              | Min. : 0.00        | Min. :0.000000          | Min. : 0.000         |
| 1st Qu.: 0.00            | 1st Qu.: 0.00      | 1st Qu.:0.000000        | 1st Qu.: 1.945       |
| Median : 6.00            | Median : 4.00      | Median :0.000000        | Median : 3.365       |
| Mean : 13.56             | Mean : 21.16       | Mean :0.001606          | Mean : 3.341         |
| 3rd Qu.: 19.00           | 3rd Qu.: 32.00     | 3rd Qu.:0.000000        | 3rd Qu.: 4.782       |
| Max. :143.00             | Max. :210.00       | Max. :0.110000          | Max. :10.710         |
| ModeratelyActiveDistance | VeryActiveDistance |                         |                      |
| Min. :0.0000             | Min. : 0.000       |                         |                      |
| 1st Qu.:0.0000           | 1st Qu.: 0.000     |                         |                      |
| Median :0.2400           | Median : 0.210     |                         |                      |
| Mean :0.5675             | Mean : 1.503       |                         |                      |
| 3rd Qu.:0.8000           | 3rd Qu.: 2.053     |                         |                      |
| Max. :6.4800             | Max. :21.920       |                         |                      |

```
summary(daily_steps)
```

| Id                | ActivityDay      | StepTotal     |
|-------------------|------------------|---------------|
| Min. :1.504e+09   | Length:940       | Min. : 0      |
| 1st Qu.:2.320e+09 | Class :character | 1st Qu.: 3790 |
| Median :4.445e+09 | Mode :character  | Median : 7406 |
| Mean :4.855e+09   |                  | Mean : 7638   |
| 3rd Qu.:6.962e+09 |                  | 3rd Qu.:10727 |
| Max. :8.878e+09   |                  | Max. :36019   |

```
summary(heartrate_seconds)
```

| Id                | Time             | Value          |
|-------------------|------------------|----------------|
| Min. :2.022e+09   | Length:2483658   | Min. : 36.00   |
| 1st Qu.:4.388e+09 | Class :character | 1st Qu.: 63.00 |
| Median :5.554e+09 | Mode :character  | Median : 73.00 |
| Mean :5.514e+09   |                  | Mean : 77.33   |
| 3rd Qu.:6.962e+09 |                  | 3rd Qu.: 88.00 |
| Max. :8.878e+09   |                  | Max. :203.00   |



```
summary(hourly_calories)
```

| Id                | ActivityHour     | Calories       |
|-------------------|------------------|----------------|
| Min. :1.504e+09   | Length:22099     | Min. : 42.00   |
| 1st Qu.:2.320e+09 | Class :character | 1st Qu.: 63.00 |
| Median :4.445e+09 | Mode :character  | Median : 83.00 |
| Mean :4.848e+09   |                  | Mean : 97.39   |
| 3rd Qu.:6.962e+09 |                  | 3rd Qu.:108.00 |
| Max. :8.878e+09   |                  | Max. :948.00   |

```
summary(hourly_intensities)
```

| Id                | ActivityHour     | TotalIntensity | AverageIntensity |
|-------------------|------------------|----------------|------------------|
| Min. :1.504e+09   | Length:22099     | Min. : 0.00    | Min. :0.0000     |
| 1st Qu.:2.320e+09 | Class :character | 1st Qu.: 0.00  | 1st Qu.:0.0000   |
| Median :4.445e+09 | Mode :character  | Median : 3.00  | Median :0.0500   |
| Mean :4.848e+09   |                  | Mean : 12.04   | Mean :0.2006     |
| 3rd Qu.:6.962e+09 |                  | 3rd Qu.: 16.00 | 3rd Qu.:0.2667   |
| Max. :8.878e+09   |                  | Max. :180.00   | Max. :3.0000     |

```
summary(hourly_steps)
```

| Id                | ActivityHour     | StepTotal      |
|-------------------|------------------|----------------|
| Min. :1.504e+09   | Length:22099     | Min. : 0.0     |
| 1st Qu.:2.320e+09 | Class :character | 1st Qu.: 0.0   |
| Median :4.445e+09 | Mode :character  | Median : 40.0  |
| Mean :4.848e+09   |                  | Mean : 320.2   |
| 3rd Qu.:6.962e+09 |                  | 3rd Qu.: 357.0 |
| Max. :8.878e+09   |                  | Max. :10554.0  |

```
summary(minute_calories)
```

| Id                | ActivityMinute   | Calories        |
|-------------------|------------------|-----------------|
| Min. :1.504e+09   | Length:1325580   | Min. : 0.0000   |
| 1st Qu.:2.320e+09 | Class :character | 1st Qu.: 0.9357 |
| Median :4.445e+09 | Mode :character  | Median : 1.2176 |
| Mean :4.848e+09   |                  | Mean : 1.6231   |
| 3rd Qu.:6.962e+09 |                  | 3rd Qu.: 1.4327 |
| Max. :8.878e+09   |                  | Max. :19.7499   |

```
summary(minute_intensities)
```

| Id                | ActivityMinute   | Intensity      |
|-------------------|------------------|----------------|
| Min. :1.504e+09   | Length:1325580   | Min. :0.0000   |
| 1st Qu.:2.320e+09 | Class :character | 1st Qu.:0.0000 |
| Median :4.445e+09 | Mode :character  | Median :0.0000 |
| Mean :4.848e+09   |                  | Mean :0.2006   |
| 3rd Qu.:6.962e+09 |                  | 3rd Qu.:0.0000 |
| Max. :8.878e+09   |                  | Max. :3.0000   |

```
summary(minute_steps)
```

|         | Id         | ActivityMinute   | Steps          |
|---------|------------|------------------|----------------|
| Min.    | :1.504e+09 | Length:1325580   | Min. : 0.000   |
| 1st Qu. | :2.320e+09 | Class :character | 1st Qu.: 0.000 |
| Median  | :4.445e+09 | Mode :character  | Median : 0.000 |
| Mean    | :4.848e+09 |                  | Mean : 5.336   |
| 3rd Qu. | :6.962e+09 |                  | 3rd Qu.: 0.000 |
| Max.    | :8.878e+09 |                  | Max. :220.000  |

```
summary(minute_METs)
```

|         | Id         | ActivityMinute   | METs           |
|---------|------------|------------------|----------------|
| Min.    | :1.504e+09 | Length:1325580   | Min. : 0.00    |
| 1st Qu. | :2.320e+09 | Class :character | 1st Qu.: 10.00 |
| Median  | :4.445e+09 | Mode :character  | Median : 10.00 |
| Mean    | :4.848e+09 |                  | Mean : 14.69   |
| 3rd Qu. | :6.962e+09 |                  | 3rd Qu.: 11.00 |
| Max.    | :8.878e+09 |                  | Max. :157.00   |

```
summary(weight_log)
```

|         | Id         | Date             | WeightKg       | WeightPounds  | Fat           |
|---------|------------|------------------|----------------|---------------|---------------|
| Min.    | :1.504e+09 | Length:67        | Min. : 52.60   | Min. :116.0   | Min. :22.00   |
| 1st Qu. | :6.962e+09 | Class :character | 1st Qu.: 61.40 | 1st Qu.:135.4 | 1st Qu.:22.75 |
| Median  | :6.962e+09 | Mode :character  | Median : 62.50 | Median :137.8 | Median :23.50 |
| Mean    | :7.009e+09 |                  | Mean : 72.04   | Mean :158.8   | Mean :23.50   |
| 3rd Qu. | :8.878e+09 |                  | 3rd Qu.: 85.05 | 3rd Qu.:187.5 | 3rd Qu.:24.25 |
| Max.    | :8.878e+09 |                  | Max. :133.50   | Max. :294.3   | Max. :25.00   |

|         | BMI    | IsManualReport   | LogId             |
|---------|--------|------------------|-------------------|
| Min.    | :21.45 | Length:67        | Min. :1.460e+12   |
| 1st Qu. | :23.96 | Class :character | 1st Qu.:1.461e+12 |
| Median  | :24.39 | Mode :character  | Median :1.462e+12 |
| Mean    | :25.19 |                  | Mean :1.462e+12   |
| 3rd Qu. | :25.56 |                  | 3rd Qu.:1.462e+12 |
| Max.    | :47.54 |                  | Max. :1.463e+12   |

```
summary(daily_sleep)
```

|         | Id         | SleepDay         | TotalSleepRecords | TotalMinutesAsleep |
|---------|------------|------------------|-------------------|--------------------|
| Min.    | :1.504e+09 | Length:413       | Min. :1.000       | Min. : 58.0        |
| 1st Qu. | :3.977e+09 | Class :character | 1st Qu.:1.000     | 1st Qu.:361.0      |
| Median  | :4.703e+09 | Mode :character  | Median :1.000     | Median :433.0      |
| Mean    | :5.001e+09 |                  | Mean :1.119       | Mean :419.5        |
| 3rd Qu. | :6.962e+09 |                  | 3rd Qu.:1.000     | 3rd Qu.:490.0      |
| Max.    | :8.792e+09 |                  | Max. :3.000       | Max. :796.0        |

|         | TotalTimeInBed |
|---------|----------------|
| Min.    | : 61.0         |
| 1st Qu. | :403.0         |
| Median  | :463.0         |
| Mean    | :458.6         |
| 3rd Qu. | :526.0         |
| Max.    | :961.0         |

## Average User Findings From the Datasets:

### **1. Daily Activity:**

- Total average steps= 7638.
- Total average distance= 5.490 km.
- Total average calories= 2304 kcal.
- Total average time spend sedentary= 16.5 hours.

### **2. Calories burned:**

- 2304 kcal/day

### **3. Heart-rate/min:**

- 97.39 kcal/hour

### **4. Hourly Steps:**

- 320 steps/day

### **5. Hourly Intensities:**

- 12.04/hour, the result is achieved by adding together the intensity values within the hour (sedentary, lightly active, fairly active, and very active).

### **6. Calories Consumed:**

- 97.39 kcal/hour

### **7. Hourly Steps:**

- 320 steps/day

### **8. METs (Metabolic Equivalent of Task):**

- 1.447 METs/minute

### **9. Sleep time:**

- 7-8 hours/day

### **10. Weight:**

- Average weight= 72kg.
- Average BMI= 25.19
- Average body fat percentage= 23.5%

## Data Manipulations:

The first step will be converting daily\_activity column (ActivityDate) into date format so this way I can be able to utilize the data for visualizations.

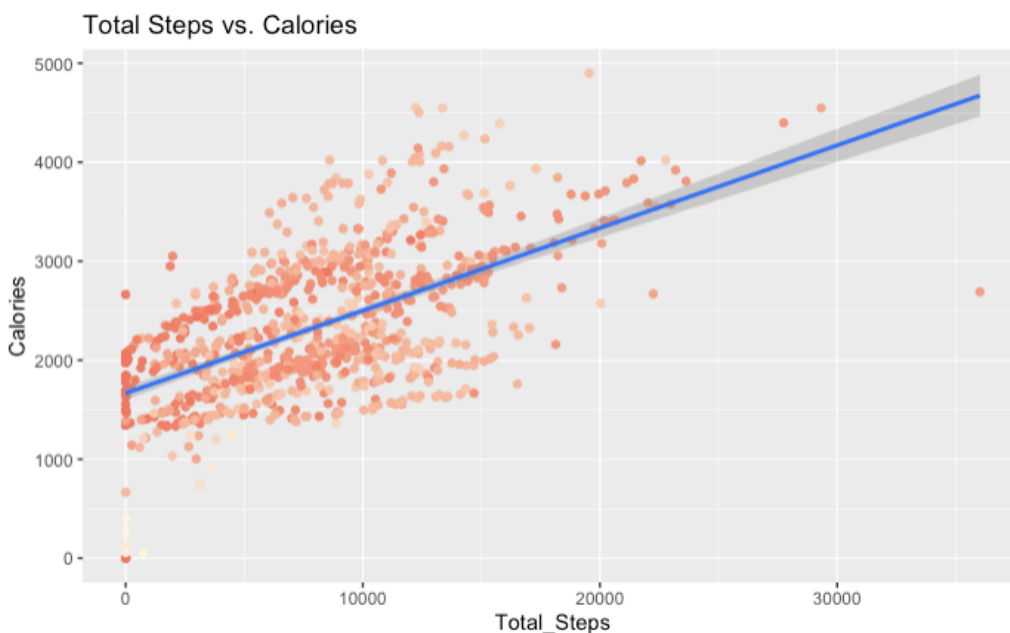
```
library(tidyverse)
daily_activity <- daily_activity %>% mutate(Weekday = weekdays(as.Date(ActivityDate, "%m/%d/%Y")))
```

## Data Visualization:

```
library(tidyverse)
ggplot(data=daily_activity, aes(x=Total_Steps, y = Calories, color=Sedentary_Minutes))+
  geom_point() +
  stat_smooth(method=lm) +
  scale_color_gradient(low="cornsilk", high="coral1") +
  labs(title="Total Steps vs. Calories") +
  theme(legend.position="none")
```



R Console

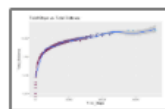


Although the correlation between the total steps and calories burned seems obvious, there are intriguing findings from the visualization. There is a number of users who managed 28,000-36,000 steps yet burned the same amount of calories as others who got 13,000-15,000 steps in. Additionally, there are users who did not register any steps but burned 1,000-2,000 calories.

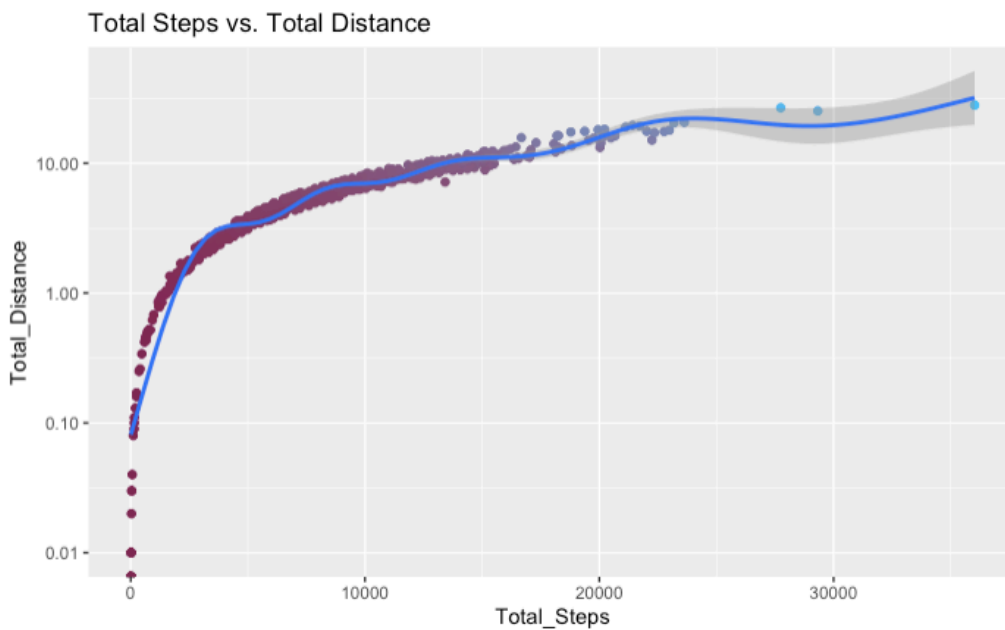
```

{r}
ggplot(data=daily_activity, aes(x=Total_Steps, y = Total_Distance, color=Total_Distance))+
  geom_point() +
  geom_jitter() +
  scale_y_log10() +
  geom_smooth(method="gam", formula = y ~s(x)) +
  scale_color_gradient(low="deeppink4", high="deepskyblue") +
  labs(title="Total Steps vs. Total Distance") +
  theme(legend.position="none")

```



R Console

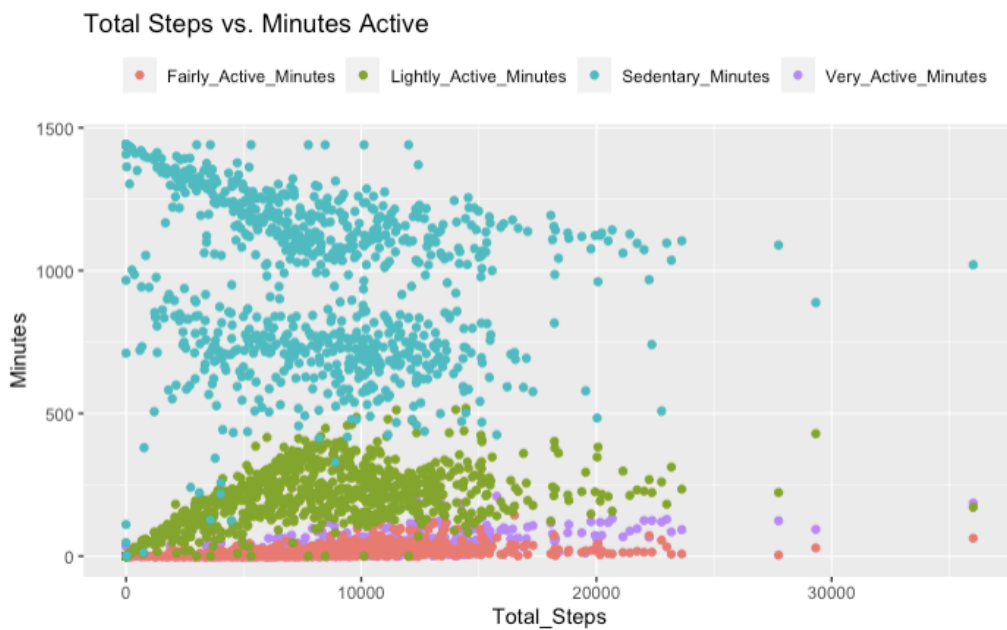


From the visualization, we can see the link between the total steps and the total distance traveled. Evidently, the relationship is linear and illustrating that most of the users take between 1,000-16,000 steps and the total distance of 1-10 km.

For the next visualization, I want to illustrate the link between total steps and active minutes.

```
```{r}
ActiveMinutes <- gather(daily_activity, key = "ActiveMinutes", value = "Minutes", 11:14)
```
```

```
```{r}
ggplot(data=ActiveMinutes, aes(x=Total_Steps, y = Minutes, color=ActiveMinutes))+
  geom_point() +
  labs(title="Total Steps vs. Minutes Active") +
  theme(legend.position="top") +
  theme(legend.title = element_blank())
```
```

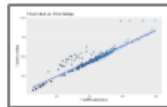


A significant number of users take between 5,000-15,000 steps data points indicate. In addition, the data shows that there are 2 sedentary time clusters, some users spend from 11-16 hours active while others spend 40minutes to 7 hours.

```

{r}
ggplot(data=daily_sleep, aes(x=TotalMinutesAsleep, y=TotalTimeInBed)) +
  geom_point(aes(color=TotalMinutesAsleep)) +
  stat_smooth(method = lm) +
  labs(title="Time in bed vs. Time Asleep") +
  theme(legend.position="none")

```

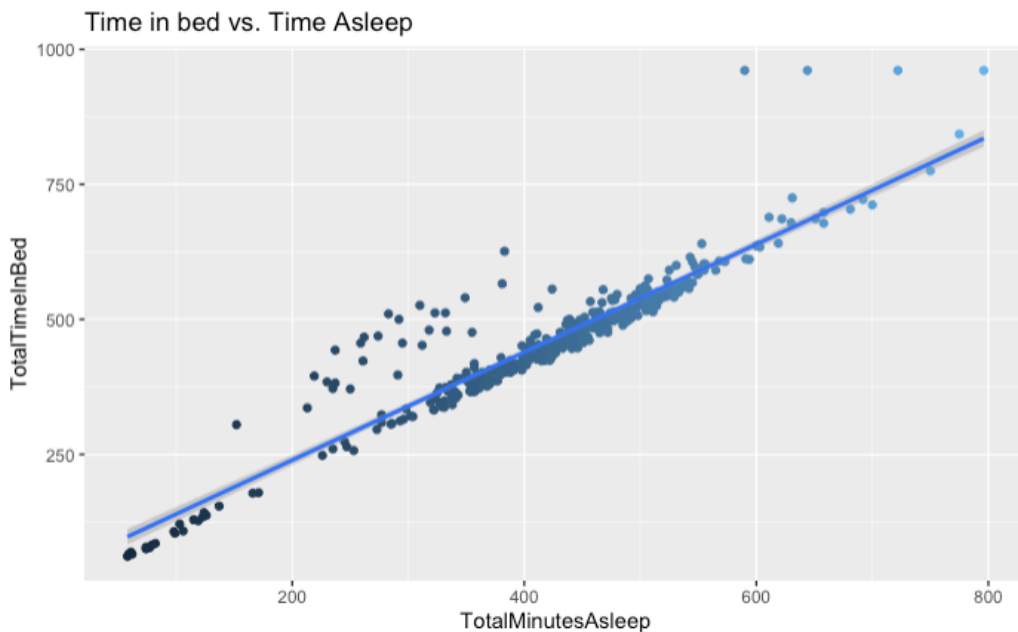


```

geom_smooth() using formula 'y ~ x'

```

R Console



The scatter plot above shows the association between time spent in bed and time asleep. The connection is linear which is not a surprise however data points vary in such a way that the users are categorized in two. The first group spends a total of 200-400 minutes asleep and 370-550 minutes in bed total which means there is a 3 hours that is spent awake in bed. The second group spends around 600-800 minutes asleep and 900 minutes in bed total. This indicates that the second group members spend anywhere around 2-5 hours awake in bed. The second group data points show that the members sleep more hours than the recommended 8, which might be a result of sleep difficulties, using electronics, or some form of discomfort.

## Recommendations and Insights:

### Business Goals:

#### 1. What are some trends in smart device usage?

- The average user's total steps is 7638, which is less than the 10,000 recommended.
- The average user spends approximately 16 hours per day sedentary.
- A significant number of users spend approximately 3 hours awake in bed thus leading to sleeping less than the recommended.

#### 2. How could these trends apply to Bellabeat customers?

Bellabeat customers would enjoy benefits that come in form of the user-friendly and easy to understand data from their sleep, activity, and overall health. Otherwise, more data is required to be in a position to provide an efficient and valuable advice from the analysis process.

#### 3. How could these trends help influence Bellabeat marketing strategy?

- Bellabeat could include a feature that encourages the user to get involved in physical activity after detecting longer sedentary periods.
- Bellabeat devices could encourage the user to get enough sleep (through pop-up notification, sleep-time reminder, etc. ) considering that almost half of the users do not get the recommended 7-8 hours.
- Bellabeat could put in place a simple reward system to encourage the user to meet their health goals and favor consistency.