Introduction#

This is Bellabeat (Wellness Technology Company) data analysis case study. This analysis process will be followed five steps as prepare, process, analysis, share, act.

##Background of the company##

Bellabeat is a high-tech company that manufactures health focused smart products. Their target market is woman. Product objective is empowered woman with knowledge about their health and habits. It was founded in 2013, grew rapidly & quickly as a tech-driven wellness company for woman. Their product's objective is collecting data on activity, sleep, stress & reproductive health of woman through products of Leaf, Time and Spring. These products has connected to Bellabeat app to understand their own details. By 2016, Bellabeat had opened offices around the world and launched multiple products. These products were available via online retailers, e-commerce channel on their website.

#Ask Phase#

Analyze the fitness tracker data to gain insight about trends and patterns on fitbit's consumer usage of fitness trackers that could be useful insights for Bellabeat marketing strategy.

##Business Objective(Questions for the analysis)

1. What are some trends in smart device usage?
2. How could these trends to apply Bellabeat customers?
3. How could these trends help influence Bellabeat's marketing strategy?

##Business task##

Unlock new growth opportunities by analyzing smart device fitness data and gaining insight into how consumers use their smart devices, which will help guide the marketing strategy for the company to achieve their potential.

##Deliverables##

Three high-Level recommendations for the marketing team.

##Main stakeholders##

Urška Sršen - Bellabeat's cofounder and Chief Creative Officer

Sando Mur -Mathematician and Bellabeat's cofounder; key member of the Bellabeat executive team

Bellabeat marketing analytics team - A team of data analysts responsible for collecting, analyzing, and reporting data that helps guide Bellabeat's marketing strategy.

#Prepare Phase#

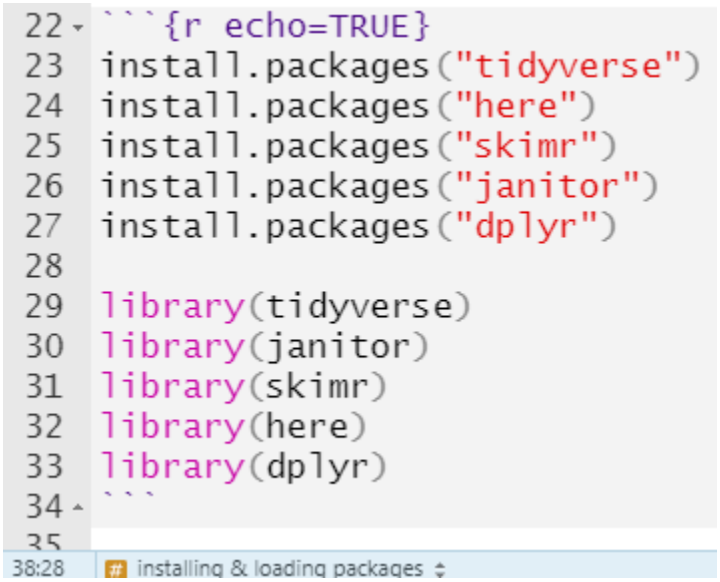
##Data Source

The data for this analysis will come from FitBit Fitness Tracker Data on Kaggle. These 18 datasets were generated by respondents to a distributed survey via Amazon Mechanical Turk between 03.12.2016–05.12.2016.

##Data Limitations ##

Data has no demographic information to confirm whether it is woman or not. It is belong to 2016 , which little outdated

##Installing & loading packages##



##Importing datasets##

The csv files were first opened in Excel. Then formatted dates/time from "general" to date/time as appropriate data format.

#Process Phase#

Viewing the data frames closely

To ensure the data frames were imported correctly, the head() function is used. The colnames() and glimpse() functions were used to explore the data frames.

```r
```{r}
head(daily_activity)
colnames(daily_activity)
glimpse(daily_activity)

head(daily_calory)
colnames(daily_calory)
glimpse(daily_calory)

head(daily_intencities)
colnames(daily_intencities)
glimpse(daily_intencities)

head(daily_steps)
colnames(daily_intencities)
glimpse(daily_intencities)

head(minute_METs)
colnames(minute_METs)
glimpse(minute_METs)

head(sleep_day)
colnames(sleep_day)
glimpse(sleep_day)

head(weight_log)
colnames(weight_log)
glimpse(weight_log)

head(heart_rate)
colnames(heart_rate)
glimpse(heart_rate)
```
```

```
R 4.2.0 · ~/fitbit_analysis/
> head(daily_calory)
          Id ActivityDay Calories
1 1503960366   4/12/2016     1985
2 1503960366   4/13/2016     1797
3 1503960366   4/14/2016     1776
4 1503960366   4/15/2016     1745
5 1503960366   4/16/2016     1863
6 1503960366   4/17/2016     1728
> colnames(daily_calory)
[1] "Id"          "ActivityDay" "Calories"
> glimpse(daily_calory)
Rows: 940
Columns: 3
$ Id          <dbl> 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, 150396036…
$ ActivityDay <chr> "4/12/2016", "4/13/2016", "4/14/2016", "4/15/2016", "4/16/2016", "4/17/2016", "4/18/2016", "4/19/2016", "4/20/2016", "4/21/2016",…
$ Calories    <int> 1985, 1797, 1776, 1745, 1863, 1728, 1921, 2035, 1786, 1775, 1827, 1949, 1788, 2013, 1970, 2159, 1898, 1837, 1947, 1820, 2004, 199…
>
```

Console | Terminal | Jobs

R 4.2.0 · ~/fitbit_analysis/

```
> head(daily_activity)
          Id ActivityDate TotalSteps TotalDistance TrackerDistance LoggedActivitiesDistance
1 1503960366    4/12/2016      13162          8.50            8.50                         0
2 1503960366    4/13/2016      10735          6.97            6.97                         0
3 1503960366    4/14/2016      10460          6.74            6.74                         0
4 1503960366    4/15/2016       9762          6.28            6.28                         0
5 1503960366    4/16/2016      12669          8.16            8.16                         0
6 1503960366    4/17/2016       9705          6.48            6.48                         0
  VeryActiveDistance ModeratelyActiveDistance LightActiveDistance SedentaryActiveDistance
1               1.88                     0.55                6.06                        0
2               1.57                     0.69                4.71                        0
3               2.44                     0.40                3.91                        0
4               2.14                     1.26                2.83                        0
5               2.71                     0.41                5.04                        0
6               3.19                     0.78                2.51                        0
  VeryActiveMinutes FairlyActiveMinutes LightlyActiveMinutes SedentaryMinutes Calories
1                25                  13                  328              728     1985
2                21                  19                  217              776     1797
3                30                  11                  181             1218     1776
4                29                  34                  209              726     1745
5                36                  10                  221              773     1863
6                38                  20                  164              539     1728
> colnames(daily_activity)
 [1] "Id"                       "ActivityDate"             "TotalSteps"
 [4] "TotalDistance"            "TrackerDistance"          "LoggedActivitiesDistance"
 [7] "VeryActiveDistance"       "ModeratelyActiveDistance" "LightActiveDistance"
[10] "SedentaryActiveDistance"  "VeryActiveMinutes"        "FairlyActiveMinutes"
[13] "LightlyActiveMinutes"     "SedentaryMinutes"         "Calories"
> glimpse(daily_activity)
Rows: 940
Columns: 15
$ Id                       <dbl> 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, 1503960366…
$ ActivityDate             <chr> "4/12/2016", "4/13/2016", "4/14/2016", "4/15/2016", "4/16/2016", "4/17…
$ TotalSteps               <int> 13162, 10735, 10460, 9762, 12669, 9705, 13019, 15506, 10544, 9819, 127…
$ TotalDistance            <dbl> 8.50, 6.97, 6.74, 6.28, 8.16, 6.48, 8.59, 9.88, 6.68, 6.34, 8.13, 9.04…
$ TrackerDistance          <dbl> 8.50, 6.97, 6.74, 6.28, 8.16, 6.48, 8.59, 9.88, 6.68, 6.34, 8.13, 9.04…
$ LoggedActivitiesDistance <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0…
$ VeryActiveDistance       <dbl> 1.88, 1.57, 2.44, 2.14, 2.71, 3.19, 3.25, 3.53, 1.96, 1.34, 4.76, 2.81…
$ ModeratelyActiveDistance <dbl> 0.55, 0.69, 0.40, 1.26, 0.41, 0.78, 0.64, 1.32, 0.48, 0.35, 1.12, 0.87…
$ LightActiveDistance      <dbl> 6.06, 4.71, 3.91, 2.83, 5.04, 2.51, 4.71, 5.03, 4.24, 4.65, 2.24, 5.36…
$ SedentaryActiveDistance  <dbl> 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00…
$ VeryActiveMinutes        <int> 25, 21, 30, 29, 36, 38, 42, 50, 28, 19, 66, 41, 39, 73, 31, 78, 48, 16…
$ FairlyActiveMinutes      <int> 13, 19, 11, 34, 10, 20, 16, 31, 12, 8, 27, 21, 5, 14, 23, 11, 28, 12, …
$ LightlyActiveMinutes     <int> 328, 217, 181, 209, 221, 164, 233, 264, 205, 211, 130, 262, 238, 216, …
$ SedentaryMinutes         <int> 728, 776, 1218, 726, 773, 539, 1149, 775, 818, 838, 1217, 732, 709, 81…
$ Calories                 <int> 1985, 1797, 1776, 1745, 1863, 1728, 1921, 2035, 1786, 1775, 1827, 1949…
>
```

```
> head(minute_METs)
          Id      Date Minute METs
1 1503960366 4/12/2016   0:00   10
2 1503960366 4/12/2016   0:01   10
3 1503960366 4/12/2016   0:02   10
4 1503960366 4/12/2016   0:03   10
5 1503960366 4/12/2016   0:04   10
6 1503960366 4/12/2016   0:05   12
> colnames(minute_METs)
[1] "Id"     "Date"   "Minute" "METs"
> glimpse(minute_METs)
Rows: 1,048,575
Columns: 4
$ Id     <dbl> 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, …
$ Date   <chr> "4/12/2016", "4/12/2016", "4/12/2016", "4/12/2016", "4/12/2016", "4/12/2016", "4/12/2016", "4/12/2016", "4/12/2016", "4/…
$ Minute <chr> "0:00", "0:01", "0:02", "0:03", "0:04", "0:05", "0:06", "0:07", "0:08", "0:09", "0:10", "0:11", "0:12", "0:13", "0:14", …
$ METs   <int> 10, 10, 10, 10, 10, 12, 12, 12, 12, 12, 12, 12, 10, 10, 12, 10, 12, 10, 10, 10, 12, 12, 12, 12, 12, 26, 12, 12, 12, 32, …
>
```

```
> head(daily_intencities)
         Id ActivityDay SedentaryMinutes LightlyActiveMinutes FairlyActiveMinutes VeryActiveMinutes SedentaryActiveDistance
1 1503960366   4/12/2016              728                  328                  13                25                       0
2 1503960366   4/13/2016              776                  217                  19                21                       0
3 1503960366   4/14/2016             1218                  181                  11                30                       0
4 1503960366   4/15/2016              726                  209                  34                29                       0
5 1503960366   4/16/2016              773                  221                  10                36                       0
6 1503960366   4/17/2016              539                  164                  20                38                       0
  LightActiveDistance ModeratelyActiveDistance VeryActiveDistance
1                6.06                      0.55               1.88
2                4.71                      0.69               1.57
3                3.91                      0.40               2.44
4                2.83                      1.26               2.14
5                5.04                      0.41               2.71
6                2.51                      0.78               3.19
> colnames(daily_intencities)
 [1] "Id"                       "ActivityDay"              "SedentaryMinutes"         "LightlyActiveMinutes"
 [5] "FairlyActiveMinutes"      "VeryActiveMinutes"        "SedentaryActiveDistance"  "LightActiveDistance"
 [9] "ModeratelyActiveDistance" "VeryActiveDistance"
> glimpse(daily_intencities)
Rows: 940
Columns: 10
$ Id                       <dbl> 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, 150396…
$ ActivityDay              <chr> "4/12/2016", "4/13/2016", "4/14/2016", "4/15/2016", "4/16/2016", "4/17/2016", "4/18/2016", "4/19/2016"…
$ SedentaryMinutes         <int> 728, 776, 1218, 726, 773, 539, 1149, 775, 818, 838, 1217, 732, 709, 814, 833, 1108, 782, 815, 712, 730…
$ LightlyActiveMinutes     <int> 328, 217, 181, 209, 221, 164, 233, 264, 205, 211, 130, 262, 238, 216, 279, 243, 189, 243, 217, 246, 27…
$ FairlyActiveMinutes      <int> 13, 19, 11, 34, 10, 20, 16, 31, 12, 8, 27, 21, 5, 14, 23, 11, 28, 12, 34, 35, 15, 24, 22, 24, 6, 46, 8…
$ VeryActiveMinutes        <int> 25, 21, 30, 29, 36, 38, 42, 50, 28, 19, 66, 41, 39, 73, 31, 78, 48, 16, 52, 33, 41, 50, 36, 45, 24, 37…
$ SedentaryActiveDistance  <dbl> 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, …
$ LightActiveDistance      <dbl> 6.06, 4.71, 3.91, 2.83, 5.04, 2.51, 4.71, 5.03, 4.24, 4.65, 2.24, 5.36, 3.28, 3.94, 5.54, 5.41, 3.79, …
$ ModeratelyActiveDistance <dbl> 0.55, 0.69, 0.40, 1.26, 0.41, 0.78, 0.64, 1.32, 0.48, 0.35, 1.12, 0.87, 0.21, 0.57, 0.92, 0.41, 1.16, …
$ VeryActiveDistance       <dbl> 1.88, 1.57, 2.44, 2.14, 2.71, 3.19, 3.25, 3.53, 1.96, 1.34, 4.76, 2.81, 2.92, 5.29, 2.33, 6.40, 3.54, …
>
```

```
> head(weight_log)
          Id      Date     Time WeightKg WeightPounds Fat   BMI IsManualReport    LogId
1 1503960366  5/2/2016 11:59 PM     52.6     115.9631  22 22.65           TRUE 1.46e+12
2 1503960366  5/3/2016 11:59 PM     52.6     115.9631  NA 22.65           TRUE 1.46e+12
3 1927972279 4/13/2016  1:08 AM    133.5     294.3171  NA 47.54          FALSE 1.46e+12
4 2873212765 4/21/2016 11:59 PM     56.7     125.0021  NA 21.45           TRUE 1.46e+12
5 2873212765 5/12/2016 11:59 PM     57.3     126.3249  NA 21.69           TRUE 1.46e+12
6 4319703577 4/17/2016 11:59 PM     72.4     159.6147  25 27.45           TRUE 1.46e+12
> colnames(weight_log)
[1] "Id"             "Date"           "Time"           "WeightKg"       "WeightPounds"   "Fat"
[7] "BMI"            "IsManualReport" "LogId"
> glimpse(weight_log)
Rows: 67
Columns: 9
$ Id             <dbl> 1503960366, 1503960366, 1927972279, 2873212765, 2873212765, 4319703577, 4319703577, 4558…
$ Date           <chr> "5/2/2016", "5/3/2016", "4/13/2016", "4/21/2016", "5/12/2016", "4/17/2016", "5/4/2016", …
$ Time           <chr> "11:59 PM", "11:59 PM", "1:08 AM", "11:59 PM", "11:59 PM", "11:59 PM", "11:59 PM", "11:5…
$ WeightKg       <dbl> 52.6, 52.6, 133.5, 56.7, 57.3, 72.4, 72.3, 69.7, 70.3, 69.9, 69.2, 69.1, 90.7, 62.5, 62.…
$ WeightPounds   <dbl> 115.9631, 115.9631, 294.3171, 125.0021, 126.3249, 159.6147, 159.3942, 153.6622, 154.9850…
$ Fat            <int> 22, NA, NA, NA, NA, 25, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, …
$ BMI            <dbl> 22.65, 22.65, 47.54, 21.45, 21.69, 27.45, 27.38, 27.25, 27.46, 27.32, 27.04, 27.00, 28.0…
$ IsManualReport <lgl> TRUE, TRUE, FALSE, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, FALSE, TRUE, TR…
$ LogId          <dbl> 1.46e+12, 1.46e+12, 1.46e+12, 1.46e+12, 1.46e+12, 1.46e+12, 1.46e+12, 1.46e+12, 1.46e+12…
```

```
> head(sleep_day)
          Id      Date     Time TotalSleepRecords TotalMinutesAsleep TotalTimeInBed
1 1503960366 4/12/2016 12:00 AM                 1                327            346
2 1503960366 4/13/2016 12:00 AM                 2                384            407
3 1503960366 4/15/2016 12:00 AM                 1                412            442
4 1503960366 4/16/2016 12:00 AM                 2                340            367
5 1503960366 4/17/2016 12:00 AM                 1                700            712
6 1503960366 4/19/2016 12:00 AM                 1                304            320
> colnames(sleep_day)
[1] "Id"                 "Date"               "Time"               "TotalSleepRecords"  "TotalMinutesAsleep"
[6] "TotalTimeInBed"
> glimpse(sleep_day)
Rows: 413
Columns: 6
$ Id                 <dbl> 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, 1503960366, …
$ Date               <chr> "4/12/2016", "4/13/2016", "4/15/2016", "4/16/2016", "4/17/2016", "4/19/2016", "4/20/…
$ Time               <chr> "12:00 AM", "12:00 AM", "12:00 AM", "12:00 AM", "12:00 AM", "12:00 AM", "12:00 AM", …
$ TotalSleepRecords  <int> 1, 2, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, …
$ TotalMinutesAsleep <int> 327, 384, 412, 340, 700, 304, 360, 325, 361, 430, 277, 245, 366, 341, 404, 369, 277,…
$ TotalTimeInBed     <int> 346, 407, 442, 367, 712, 320, 377, 364, 384, 449, 323, 274, 393, 354, 425, 396, 309,…
>
```

```
> head(heart_rate)
          Id      Date    Time Value
1 2022484408 4/12/2016 7:21 AM    97
2 2022484408 4/12/2016 7:21 AM   102
3 2022484408 4/12/2016 7:21 AM   105
4 2022484408 4/12/2016 7:21 AM   103
5 2022484408 4/12/2016 7:21 AM   101
6 2022484408 4/12/2016 7:22 AM    95
> colnames(heart_rate)
[1] "Id"    "Date"  "Time"  "Value"
> glimpse(heart_rate)
Rows: 1,048,575
Columns: 4
$ Id    <dbl> 2022484408, 2022484408, 2022484408, 2022484408, 2022484408, 2022484408, 2022484408, 2022484408, 202…
$ Date  <chr> "4/12/2016", "4/12/2016", "4/12/2016", "4/12/2016", "4/12/2016", "4/12/2016", "4/12/2016", "4/12/20…
$ Time  <chr> "7:21 AM", "7:21 AM", "7:21 AM", "7:21 AM", "7:21 AM", "7:22 AM", "7:22 AM", "7:22 AM", "7:22 AM", …
$ Value <int> 97, 102, 105, 103, 101, 95, 91, 93, 94, 93, 92, 89, 83, 61, 60, 61, 61, 57, 54, 55, 58, 60, 59, 57,…
> ```
```

##Merge data frames##

All data frames associated with Id column. daily-activity column contain data for daily_-calory, daily_-intensities, daily-steps. Then it is possible merge them if needed. Sqldf package use to determine if the values of daily calories, daily steps, the number of observations must be the same and the observations must match for each ID number.

However, the number of columns must be the same between the data frames, so a temporary data frame with the important columns is created first.

```
99
100 ```{r}
101 daliy_activity2 <- daily_activity %>%
102     select(Id,ActivityDate,Calories)
103       head(daily_activity2)
104
105 sql_check <-  sqldf('SELECT * FROM daily_activity2 INTERSECT SELECT * FROM daily_calory')
106 head(sql_check)
107
108 daily_activity3 <- daily_activity %>%
109   select(Id,ActivityDate,SedentaryMinutes,LightlyActiveMinutes,FairlyActiveMinutes,VeryActiveMinutes,SedentaryA
    ctiveDistance,LightActiveDistance,ModeratelyActiveDistance,VeryActiveDistance)
110 head(daily_activity3)
111
112 sql_check_2 <- sqldf('SELECT * FROM daily_activity3 INTERSECT SELECT * FROM daily_intencities')
113 head(sql_check_2)
114
115 daily_activity4 <- daily_activity %>%
116   select(Id,ActivityDate,TotalSteps)
117 head(daily_activity4)
118
119
120 sql_check3 <- sqldf('SELECT * FROM daily_activity4 INTERSECT SELECT * FROM daily_steps')
121 head(sql_check3)
122
123
```

Console | Terminal | Jobs

R 4.2.0 · ~/fitbit_analysis/

```
> n_distinct(daily_activity$Id)
[1] 33
>
> n_distinct(heart_rate$Id)
[1] 7
> n_distinct(minute_METs$Id)
[1] 27
> n_distinct(sleep_day$Id)
[1] 24
> n_distinct(weight_log$Id)
[1] 8
> nrow(daily_activity)
[1] 940
> nrow(heart_rate)
[1] 1048575
> nrow(minute_METs)
[1] 1048575
> nrow(sleep_day)
[1] 413
> nrow(weight_log)
[1] 67
>
```

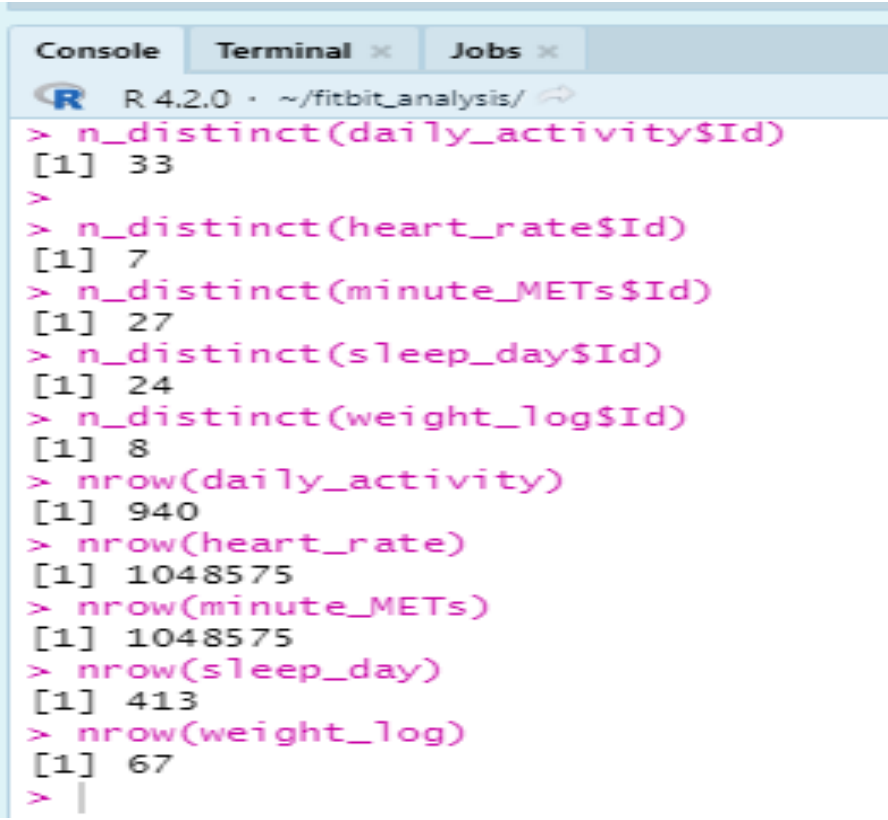
```
> head(daily_activity3)
         Id ActivityDate SedentaryMinutes LightlyActiveMinutes FairlyActiveMinutes
1 1503960366    4/12/2016              728                  328                  13
2 1503960366    4/13/2016              776                  217                  19
3 1503960366    4/14/2016             1218                  181                  11
4 1503960366    4/15/2016              726                  209                  34
5 1503960366    4/16/2016              773                  221                  10
6 1503960366    4/17/2016              539                  164                  20
  VeryActiveMinutes SedentaryActiveDistance LightActiveDistance ModeratelyActiveDistance
1                25                       0                6.06                      0.55
2                21                       0                4.71                      0.69
3                30                       0                3.91                      0.40
4                29                       0                2.83                      1.26
5                36                       0                5.04                      0.41
6                38                       0                2.51                      0.78
  VeryActiveDistance
1               1.88
2               1.57
3               2.44
4               2.14
5               2.71
6               3.19
> head(sql_check_2)
         Id ActivityDate SedentaryMinutes LightlyActiveMinutes FairlyActiveMinutes
1 1503960366    4/12/2016              728                  328                  13
2 1503960366    4/13/2016              776                  217                  19
3 1503960366    4/14/2016             1218                  181                  11
4 1503960366    4/15/2016              726                  209                  34
5 1503960366    4/16/2016              773                  221                  10
6 1503960366    4/17/2016              539                  164                  20
  VeryActiveMinutes SedentaryActiveDistance LightActiveDistance ModeratelyActiveDistance
1                25                       0                6.06                      0.55
2                21                       0                4.71                      0.69
3                30                       0                3.91                      0.40
4                29                       0                2.83                      1.26
5                36                       0                5.04                      0.41
6                38                       0                2.51                      0.78
  VeryActiveDistance
1               1.88
2               1.57
3               2.44
4               2.14
5               2.71
6               3.19
> nrow(sql_check_2)
[1] 940
```

```
> head(daily_activity4)
         Id ActivityDate TotalSteps
1 1503960366    4/12/2016      13162
2 1503960366    4/13/2016      10735
3 1503960366    4/14/2016      10460
4 1503960366    4/15/2016       9762
5 1503960366    4/16/2016      12669
6 1503960366    4/17/2016       9705
> head(sql_check3)
         Id ActivityDate TotalSteps
1 1503960366    4/12/2016      13162
2 1503960366    4/13/2016      10735
3 1503960366    4/14/2016      10460
4 1503960366    4/15/2016       9762
5 1503960366    4/16/2016      12669
6 1503960366    4/17/2016       9705
> nrow(sql_check3)
[1] 940
```

The outputs of the head() function of the temporary data frames created, match the outputs of the head() function for the original data frames. The outputs of the head() function of the SQL data frames match the outputs of the head() function for the temporary data frames. The number of observations for each SQL data frame are equal to 940. It can be concluded that the data for the daily-_calories, daily_-intensities, and daily_-steps data frames are contained in daily_-activity. These three data frames will be removed from the analysis for simplicity.

##Summarizing the data##

The n-distinct() and nrow() functions are used to determine the number of unique values and the number of rows in a data frame, respectively.



The heart rate and weight log data frames contain a very low number of participants based on above unique values explorations. That is not a sufficient to make good recommendations. The summary() function is used to pull key statistics about the data frames.

```r
daily_activity %>%
  select(TotalSteps,TotalDistance,SedentaryMinutes,LightlyActiveMinutes,FairlyActiveMinutes,VeryActiveMinutes,Calories)
%>%
summary()
```

```
   TotalSteps    TotalDistance    SedentaryMinutes LightlyActiveMinutes FairlyActiveMinutes
 Min.   :    0   Min.   : 0.000   Min.   :   0.0   Min.   :  0.0        Min.   :  0.00
 1st Qu.: 3790   1st Qu.: 2.620   1st Qu.: 729.8   1st Qu.:127.0        1st Qu.:  0.00
 Median : 7406   Median : 5.245   Median :1057.5   Median :199.0        Median :  6.00
 Mean   : 7638   Mean   : 5.490   Mean   : 991.2   Mean   :192.8        Mean   : 13.56
 3rd Qu.:10727   3rd Qu.: 7.713   3rd Qu.:1229.5   3rd Qu.:264.0        3rd Qu.: 19.00
 Max.   :36019   Max.   :28.030   Max.   :1440.0   Max.   :518.0        Max.   :143.00
 VeryActiveMinutes    Calories
 Min.   :  0.00    Min.   :   0
 1st Qu.:  0.00    1st Qu.:1828
 Median :  4.00    Median :2134
 Mean   : 21.16    Mean   :2304
 3rd Qu.: 32.00    3rd Qu.:2793
 Max.   :210.00    Max.   :4900
```

This summary shows the average user is taking 7638 steps a day, recommended 10,000 steps for health by the CDC. On average, users are getting 21.16 minutes of very active for a activity a day, this equates to 148.12 minutes a week. The CDC recommends 75 minutes of energetic activity a week, so the typical Fitbit user is doing well in this area and achieving additional health benefits. In contrast, participants are averaging 991.2 minutes, or 16.52 hours of sedentary time a day!

This is a significant amount of time and can lead to other health issues because the body functions best straight. Scientists have determined that 40 minutes of moderate to vigorous activity a day will balance out the effects of sitting up to 10 hours a day.

Furthermore, this summary shows the average user is burning 2304 calories a day. Studies show the average person in the population burns 1800 calories a day, but burning 3500 is needed to lose a pound of weight. The Fitbit users in this case are burning more than the norm, and are on track to lose a few pounds a week if they so choose.

Despite the low number of users in the heart rate data frame, the average heartrate of 77 beats per minute (bpm) fits within the “normal” range. Normal is considered to be between 50 to 80 bpm for men, and 53 to 82 bpm for women. However, research suggests that it is more important for individuals to determine what is a normal and healthy heartrate for them, and not compare to population levels. This is because resting heart rates between different people can vary by as much as 70 bpm. Changes in resting heartrate over days can be a sign of infection, menstrual cycle effects, or other acute triggers. Thus, making heartrate a vital health characteristic to monitor.

##Heart-rate##

```
R 4.2.0 · ~/fitbit_analysis/
> heart_rate %>%
+ select(value) %>%
+ summary()
     Value
 Min.   : 38.00
 1st Qu.: 64.00
 Median : 75.00
 Mean   : 77.02
 3rd Qu.: 87.00
 Max.   :203.00
```

##Minute METs##

```
> minute_METs %>%
+ select (METs) %>%
+ summary()
      METs
 Min.   :  0.00
 1st Qu.: 10.00
 Median : 10.00
 Mean   : 14.47
 3rd Qu.: 11.00
```

##Weight log##

```
 Max.   :3.000     Max.   :796.0      Max.   :
> weight_log %>%
+ select(WeightPounds,BMI) %>%
+ summary()
  WeightPounds        BMI
 Min.   :116.0   Min.   :21.45
 1st Qu.:135.4   1st Qu.:23.96
 Median :137.8   Median :24.39
 Mean   :158.8   Mean   :25.19
 3rd Qu.:187.5   3rd Qu.:25.56
 Max.   :294.3   Max.   :47.54
>
```

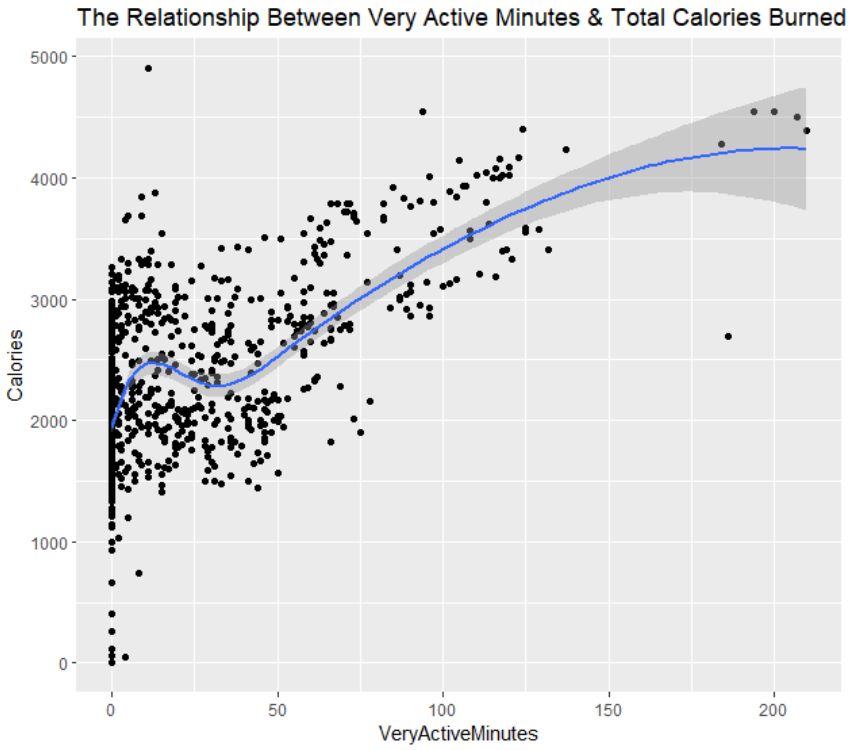
##Sleep day##

```
> sleep_day %>%
+ select(TotalSleepRecords,TotalMinutesAsleep,TotalTimeInBed) %>%
+ summary()
 TotalSleepRecords TotalMinutesAsleep TotalTimeInBed
 Min.   :1.000     Min.   : 58.0      Min.   : 61.0
 1st Qu.:1.000     1st Qu.:361.0      1st Qu.:403.0
 Median :1.000     Median :433.0      Median :463.0
 Mean   :1.119     Mean   :419.5      Mean   :458.6
 3rd Qu.:1.000     3rd Qu.:490.0      3rd Qu.:526.0
 Max.   :3.000     Max.   :796.0      Max.   :961.0
```

#Share Phase#

The ggplot() function of R Studio was used to create data visualizations that depict patterns and trends found in the data frames.

```
199
200 ```{r}
201 ggplot(data=daily_activity, aes(x=VeryActiveMinutes, y=Calories)) +
202   geom_point() + geom_smooth() +
203   labs(title="The Relationship Between Very Active Minutes & Total Calories Burned in
    a day")
204 ```
```



Plot-1

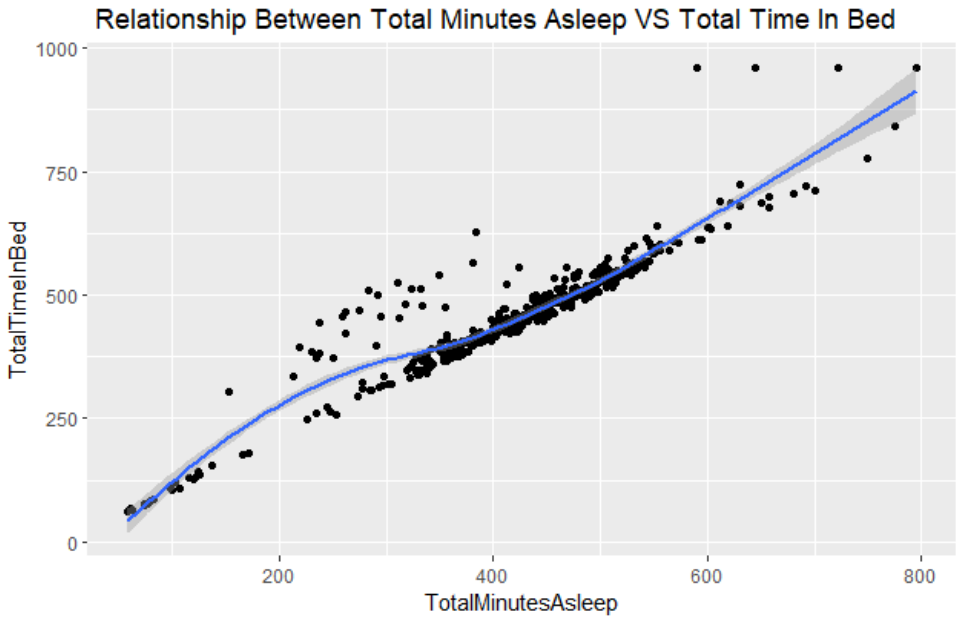
depicts a positive relationship between very active minutes and total daily calories burned. This means that the more vigorous physical activity the participant did, the more calories they burned.

```r
ggplot(data=daily_activity, aes(x=VeryActiveMinutes, y=Calories)) +
  geom_point() + geom_smooth() +
  labs(title="The Relationship Between Very Active Minutes & Total Calories Burned in a day")
```

Plot-2

Plot2 shows a positive trend between total distance and total daily calories burned. As the participants moved a greater distance, the number of calories they burned also increased.


```r
ggplot(data=daily_activity, aes(x=VeryActiveMinutes, y=Calories)) +
  geom_point() + geom_smooth() +
  labs(title="The Relationship Between Very Active Minutes & Total Calories Burned in a day")
```

Plot-3

Plot-4 depicts a positive relationship between total minutes asleep and total time in bed. For the most part, the time participants spent asleep and the time they spent in bed was very similar.

#Act Phase#

Bellabeat has been successful since its foundation by empowering women though providing data on their activity, sleep, stress, hydration levels, and reproductive health. Based on analyzing how Fitbit consumers use and respond to features, recommendations can be made to promote further growth for Bellabeat.

The Bellabeat app should be completely transformed. Rather than simply providing data on user's health, the app should further encourage users to meet fitness goals and become a social media platform.

The CDC recommends working out with a friend in order to feel more motivated, be more adventurous in trying workouts, and to become consistent. The CDC even recommends the use of a social media workout app to connect with friends and reach your goals. The Bellabeat app could become that social media workout app that women turn to, by creating a sisterhood of supportive women ready to prioritize their health.

##Recommendations for (Bellabeat App)##

1. Enable social networking so users can post their favorite workouts, wellness tips, healthy meals, etc.
2. Enable users to add friends and view each other's activity.
3. Create weekly fitness and wellness challenges to encourage use.
4. Have health and fitness companies pay for advertising.
5. Recommend users to get 10,000 steps a day and enable alert notifications to encourage users to meet goal.
6. Recommend users to get at least 7 hours of sleep a night and enable alert notifications to encourage users to meet this.

##Recommendations (Bellabeat membership)##

1. Offer 30-day free trial subscription.
2. Offer reduced subscription fee when a member refers a friend.
3. Offer discounts for Bellabeat smart device products with membership.
4. Partner with health & fitness companies and offer discounts for members.

##Recommendations for (Bellabeat products)

1. Heavily market Spring as Fitbit does not track hydration levels.

2. Offer a bundle deal for the Spring and Leaf together.

Sources

- "3 Reasons to Work out with a Friend." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 23 Apr. 2021, www.cdc.gov/diabetes/library/spotlights/workout-buddy.html
- "About Adult Bmi." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 17 Sept. 2020, www.cdc.gov/healthyweight/assessing/bmi/adult_bmi/index.html
- "CDC — How Much Sleep Do I Need? — Sleep and Sleep Disorders." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 2 Mar. 2017, www.cdc.gov/sleep/about_sleep/how_much_sleep.html
- "The Dangers of Sitting: Why Sitting Is the New Smoking." *The Dangers of Sitting: Why Sitting Is the New Smoking — Better Health Channel*, 22 Aug. 2020, www.betterhealth.vic.gov.au/health/healthyliving/the-dangers-of-sitting
- Gornall, Lucy. "How to Lose Weight: How Many Calories Should i Eat to Lose Weight?" *GoodtoKnow*, 12 Aug. 2020, www.goodto.com/wellbeing/diets-exercise/what-is-calorie-how-many-lose-weigt-425557
- Grey, Heather. "Heart Rates Can Vary by 70 Bpm: What That Means for Your Health." *Healthline*, Healthline Media, 9 Feb. 2020, www.healthline.com/health-news/what-your-heart-rate-says-about-your-health
- "How Much Physical Activity Do Adults Need?" *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 7 Oct. 2020, www.cdc.gov/physicalactivity/basics/adults/index.html