

Bellabeat_Project

Cristina Rouson

2024-09-23

##Bellabeat is a high-tech company that manufactures health-focus smart products.

Stakeholders are: -Urska Srzen: Co-founder and Chief Creative Officer -Sando Mur: Co-founder and key member of the Bellabeat executive team

I have been asked to analyze a non-Bellabeat smart devices to gain insight on how consumers are using them.

I need to answer the following questions: 1. What are some trends in smart device usage? 2. How could these trends apply to Bellabeat customers? 3. How could these trends help influence Bellabeat marketing strategy?

I am using FitBit Fitness Tracker Data. This dataset made available through Mobius. The dataset contains personal fitness tracker for 30 eligible Fitbit users. The data is pertaining to woman but does not give any details on the woman's social status, age or physical abilities.

```
#Start off by installing all the needed packages for R programming to clean, analyze and
#Installing packages
install.packages("tidyverse")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.4'
## (as 'lib' is unspecified)
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## vforcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr    1.3.1
## v purrr    1.0.2
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()   masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
install.packages("dplyr")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.4'
## (as 'lib' is unspecified)
library(dplyr)
install.packages("lubridate")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.4'
## (as 'lib' is unspecified)
library(lubridate)
install.packages("ggplot2")
```

```

## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.4'
## (as 'lib' is unspecified)
library(ggplot2)
install.packages("tidyverse")

## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.4'
## (as 'lib' is unspecified)
library(tidyverse)
install.packages("readr")

## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.4'
## (as 'lib' is unspecified)
library(readr)
install.packages("tibble")

## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.4'
## (as 'lib' is unspecified)
library(tibble)
getwd()

## [1] "/cloud/project"
install.packages("here")

## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.4'
## (as 'lib' is unspecified)
library(here)

## here() starts at /cloud/project
install.packages("skimr")

## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.4'
## (as 'lib' is unspecified)
library(skimr)
install.packages("janitor")

## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.4'
## (as 'lib' is unspecified)
library(janitor)

##
## Attaching package: 'janitor'
##
## The following objects are masked from 'package:stats':
##
##     chisq.test, fisher.test
install.packages("rmarkdown")

## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.4'
## (as 'lib' is unspecified)
library(rmarkdown)
install.packages("scales")

```

```

## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.4'
## (as 'lib' is unspecified)
library(scales)
```

```

##
## Attaching package: 'scales'
##
## The following object is masked from 'package:purrr':
##
##      discard
##
## The following object is masked from 'package:readr':
##
##      col_factor

#Importing csv data
calories<-read.csv("dailyCalories_merged.csv")
activity<-read.csv("dailyActivity_merged.csv")
intensities<-read.csv("dailyIntensities_merged.csv")
steps<-read.csv("dailySteps_merged.csv")
sleep<-read.csv("sleepDay_merged.csv")
```

##Data_Organization -activity 940 obs. of 15 variables -calories 940 obs of 3 variables -intensities 940 obs of 15 variables -sleep 413 obs. 5 variables -steps 940 obs. of 3 variables

```
#Preview data set using head() to see a small sampling of the each dataset
head(activity)
```

	Id	ActivityDate	TotalSteps	TotalDistance	TrackerDistance
## 1	1503960366	4/12/2016	13162	8.50	8.50
## 2	1503960366	4/13/2016	10735	6.97	6.97
## 3	1503960366	4/14/2016	10460	6.74	6.74
## 4	1503960366	4/15/2016	9762	6.28	6.28
## 5	1503960366	4/16/2016	12669	8.16	8.16
## 6	1503960366	4/17/2016	9705	6.48	6.48
	LoggedActivitiesDistance	VeryActiveDistance	ModeratelyActiveDistance		
## 1	0	1.88	0.55		
## 2	0	1.57	0.69		
## 3	0	2.44	0.40		
## 4	0	2.14	1.26		
## 5	0	2.71	0.41		
## 6	0	3.19	0.78		
	LightActiveDistance	SedentaryActiveDistance	VeryActiveMinutes		
## 1	6.06	0	25		
## 2	4.71	0	21		
## 3	3.91	0	30		
## 4	2.83	0	29		
## 5	5.04	0	36		
## 6	2.51	0	38		
	FairlyActiveMinutes	LightlyActiveMinutes	SedentaryMinutes	Calories	
## 1	13	328	728	1985	
## 2	19	217	776	1797	
## 3	11	181	1218	1776	
## 4	34	209	726	1745	
## 5	10	221	773	1863	

```

## 6          20        164       539      1728

head(calories)

##           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

head(intensities)

##           Id ActivityDay SedentaryMinutes LightlyActiveMinutes
## 1 1503960366 4/12/2016            728            328
## 2 1503960366 4/13/2016            776            217
## 3 1503960366 4/14/2016           1218            181
## 4 1503960366 4/15/2016            726            209
## 5 1503960366 4/16/2016            773            221
## 6 1503960366 4/17/2016            539            164
##   FairlyActiveMinutes VeryActiveMinutes SedentaryActiveDistance
## 1                  13             25                   0
## 2                  19             21                   0
## 3                  11             30                   0
## 4                  34             29                   0
## 5                  10             36                   0
## 6                  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

head(sleep)

##           Id SleepDay TotalSleepRecords TotalMinutesAsleep
## 1 1503960366 4/12/2016            1                327
## 2 1503960366 4/13/2016            2                384
## 3 1503960366 4/15/2016            1                412
## 4 1503960366 4/16/2016            2                340
## 5 1503960366 4/17/2016            1                700
## 6 1503960366 4/19/2016            1                304
##   TotalTimeInBed
## 1                 346
## 2                 407
## 3                 442
## 4                 367
## 5                 712
## 6                 320

head(steps)

##           Id ActivityDay StepTotal
## 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

##Use glimpse() to get the count of rows and columns and data type of each column
glimpse(activity)

## Rows: 940
## Columns: 15
## $ Id <dbl> 1503960366, 1503960366, 1503960366, 150396036~  

## $ ActivityDate <chr> "4/12/2016", "4/13/2016", "4/14/2016", "4/15/~  

## $ TotalSteps <int> 13162, 10735, 10460, 9762, 12669, 9705, 13019~  

## $ TotalDistance <dbl> 8.50, 6.97, 6.74, 6.28, 8.16, 6.48, 8.59, 9.8~  

## $ TrackerDistance <dbl> 8.50, 6.97, 6.74, 6.28, 8.16, 6.48, 8.59, 9.8~  

## $ LoggedActivitiesDistance <dbl> 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, 3.5~  

## $ ModeratelyActiveDistance <dbl> 0.55, 0.69, 0.40, 1.26, 0.41, 0.78, 0.64, 1.3~  

## $ LightActiveDistance <dbl> 6.06, 4.71, 3.91, 2.83, 5.04, 2.51, 4.71, 5.0~  

## $ SedentaryActiveDistance <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ~  

## $ VeryActiveMinutes <int> 25, 21, 30, 29, 36, 38, 42, 50, 28, 19, 66, 4~  

## $ FairlyActiveMinutes <int> 13, 19, 11, 34, 10, 20, 16, 31, 12, 8, 27, 21~  

## $ LightlyActiveMinutes <int> 328, 217, 181, 209, 221, 164, 233, 264, 205, ~  

## $ SedentaryMinutes <int> 728, 776, 1218, 726, 773, 539, 1149, 775, 818~  

## $ Calories <int> 1985, 1797, 1776, 1745, 1863, 1728, 1921, 203~  

glimpse(calories)

## Rows: 940
## Columns: 3
## $ Id <dbl> 1503960366, 1503960366, 1503960366, 1503960366, 1503960366~  

## $ ActivityDay <chr> "4/12/2016", "4/13/2016", "4/14/2016", "4/15/2016", "4/16/~  

## $ Calories <int> 1985, 1797, 1776, 1745, 1863, 1728, 1921, 2035, 1786, 1775~  

glimpse(intensities)

## Rows: 940
## Columns: 10
## $ Id <dbl> 1503960366, 1503960366, 1503960366, 150396036~  

## $ ActivityDay <chr> "4/12/2016", "4/13/2016", "4/14/2016", "4/15/~  

## $ SedentaryMinutes <int> 728, 776, 1218, 726, 773, 539, 1149, 775, 818~  

## $ LightlyActiveMinutes <int> 328, 217, 181, 209, 221, 164, 233, 264, 205, ~  

## $ FairlyActiveMinutes <int> 13, 19, 11, 34, 10, 20, 16, 31, 12, 8, 27, 21~  

## $ VeryActiveMinutes <int> 25, 21, 30, 29, 36, 38, 42, 50, 28, 19, 66, 4~  

## $ SedentaryActiveDistance <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ~  

## $ LightActiveDistance <dbl> 6.06, 4.71, 3.91, 2.83, 5.04, 2.51, 4.71, 5.0~  

## $ ModeratelyActiveDistance <dbl> 0.55, 0.69, 0.40, 1.26, 0.41, 0.78, 0.64, 1.3~  

## $ VeryActiveDistance <dbl> 1.88, 1.57, 2.44, 2.14, 2.71, 3.19, 3.25, 3.5~  

glimpse(sleep)

## Rows: 413
## Columns: 5
## $ Id <dbl> 1503960366, 1503960366, 1503960366, 1503960366, 150~  

## $ SleepDay <chr> "4/12/2016 12:00:00 AM", "4/13/2016 12:00:00 AM", "~  

## $ TotalSleepRecords <int> 1, 2, 1, 2, 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, 2~
## $ TotalTimeInBed      <int> 346, 407, 442, 367, 712, 320, 377, 364, 384, 449, 3~
glimpse(steps)

## Rows: 940
## Columns: 3
## $ Id          <dbl> 1503960366, 1503960366, 1503960366, 1503960366, 1503960366~
## $ ActivityDay <chr> "4/12/2016", "4/13/2016", "4/14/2016", "4/15/2016", "4/16/~
## $ StepTotal   <int> 13162, 10735, 10460, 9762, 12669, 9705, 13019, 15506, 1054~

#Clean Data
#look for duplicates
sum(duplicated(activity))

## [1] 0
sum(duplicated(calories))

## [1] 0
sum(duplicated(intensities))

## [1] 0
sum(duplicated(steps))

## [1] 0
sum(duplicated(sleep))

## [1] 3
#found 3 duplicates in sleep and removed:
sleep<-sleep %>% distinct(.keep_all = TRUE)
sum(duplicated(sleep))

## [1] 0
#clean data set find unique id and verify by the count
# use n_distinct to count the number of distinct values
n_distinct(activity$Id)

## [1] 33
n_distinct(calories$Id)

## [1] 33
n_distinct(sleep$Id)

## [1] 24
n_distinct(steps$Id)

## [1] 33
###total distinct id: -activity 33 -calories 33 -intensities 33 -sleep 24 -steps 33 -sleep data set does not have
information for all 33 participants
#look for missing values:
sum(is.na(activity))

```

```

## [1] 0
sum(is.na(calories))

## [1] 0
sum(is.na(intensities))

## [1] 0
sum(is.na(sleep))

## [1] 0
head(activity)

##           Id ActivityDate TotalSteps TotalDistance TrackerDistance
## 1 1503960366 4/12/2016      13162        8.50          8.50
## 2 1503960366 4/13/2016      10735        6.97          6.97
## 3 1503960366 4/14/2016      10460        6.74          6.74
## 4 1503960366 4/15/2016      9762         6.28          6.28
## 5 1503960366 4/16/2016     12669        8.16          8.16
## 6 1503960366 4/17/2016      9705         6.48          6.48
##   LoggedActivitiesDistance VeryActiveDistance ModeratelyActiveDistance
## 1                      0            1.88                0.55
## 2                      0            1.57                0.69
## 3                      0            2.44                0.40
## 4                      0            2.14                1.26
## 5                      0            2.71                0.41
## 6                      0            3.19                0.78
##   LightActiveDistance SedentaryActiveDistance VeryActiveMinutes
## 1        6.06                  0                 25
## 2        4.71                  0                 21
## 3        3.91                  0                 30
## 4        2.83                  0                 29
## 5        5.04                  0                 36
## 6        2.51                  0                 38
##   FairlyActiveMinutes LightlyActiveMinutes SedentaryMinutes Calories
## 1             13                  328                 728    1985
## 2             19                  217                 776    1797
## 3             11                  181                1218    1776
## 4             34                  209                 726    1745
## 5             10                  221                 773    1863
## 6             20                  164                 539    1728

head(calories)

##           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

head(intensities)

##           Id ActivityDay SedentaryMinutes LightlyActiveMinutes
## 1 1503960366 4/12/2016            728                 328

```

```

## 2 1503960366 4/13/2016 776 217
## 3 1503960366 4/14/2016 1218 181
## 4 1503960366 4/15/2016 726 209
## 5 1503960366 4/16/2016 773 221
## 6 1503960366 4/17/2016 539 164
## FairlyActiveMinutes VeryActiveMinutes SedentaryActiveDistance
## 1 13 25 0
## 2 19 21 0
## 3 11 30 0
## 4 34 29 0
## 5 10 36 0
## 6 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

head(sleep)

##           Id      SleepDay TotalSleepRecords TotalMinutesAsleep
## 1 1503960366 4/12/2016 12:00:00 AM          1            327
## 2 1503960366 4/13/2016 12:00:00 AM          2            384
## 3 1503960366 4/15/2016 12:00:00 AM          1            412
## 4 1503960366 4/16/2016 12:00:00 AM          2            340
## 5 1503960366 4/17/2016 12:00:00 AM          1            700
## 6 1503960366 4/19/2016 12:00:00 AM          1            304
## TotalTimeInBed
## 1 346
## 2 407
## 3 442
## 4 367
## 5 712
## 6 320

head(steps)

##           Id ActivityDay StepTotal
## 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

#renamed date columns to all match
activity <- activity %>% rename(date="ActivityDate") %>% mutate(date, format="%m/%d/%y")
calories <- calories %>% rename(date="ActivityDay")%>% mutate(date, format="%m/%d/%y")
intensities <- intensities %>% rename(date="ActivityDay") %>% mutate(date, format="%m/%d/%y")
sleep <- sleep %>% rename(date="SleepDay") %>% mutate(date, format="%m/%d/%y")
steps <- steps %>% rename(date="ActivityDay") %>% mutate(date, format="%m/%d/%y")

```

#Analyze the data to find insights

```

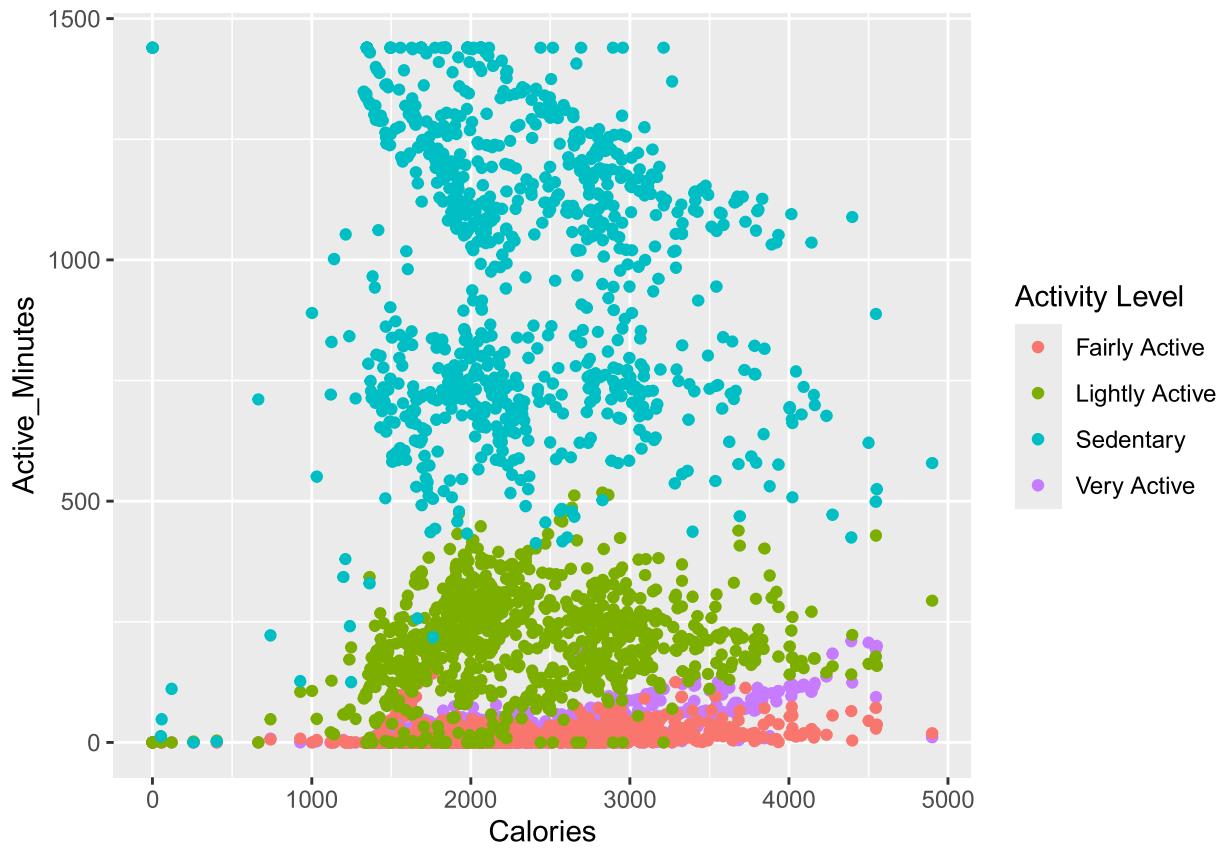
#activity_vs_calories:
#summary(activity and calories)
activity %>% select(VeryActiveMinutes,
                      FairlyActiveMinutes,
                      LightlyActiveMinutes,
                      SedentaryMinutes, Calories) %>% summary()

## #> #> VeryActiveMinutes FairlyActiveMinutes LightlyActiveMinutes SedentaryMinutes
## #> Min.    : 0.00      Min.    : 0.00      Min.    : 0.0      Min.    : 0.0
## #> 1st Qu.: 0.00      1st Qu.: 0.00      1st Qu.:127.0     1st Qu.: 729.8
## #> Median  : 4.00      Median  : 6.00      Median  :199.0     Median  :1057.5
## #> Mean    : 21.16     Mean    :13.56      Mean    :192.8     Mean    : 991.2
## #> 3rd Qu.: 32.00     3rd Qu.: 19.00     3rd Qu.:264.0     3rd Qu.:1229.5
## #> Max.    :210.00    Max.    :143.00     Max.    :518.0     Max.    :1440.0
## #> 
## #>   Calories
## #>   Min.    : 0
## #>   1st Qu.:1828
## #>   Median  :2134
## #>   Mean    :2304
## #>   3rd Qu.:2793
## #>   Max.    :4900

#As predicated, user that are very active spent more calories in less time. In comparison to those who

activity_levels <- activity %>%
  select(VeryActiveMinutes,
         FairlyActiveMinutes,
         LightlyActiveMinutes,
         SedentaryMinutes, Calories)
ggplot(activity_levels) +
  geom_point(aes(x = Calories, y = VeryActiveMinutes, color = "Very Active")) +
  geom_point(aes(x = Calories, y = FairlyActiveMinutes, color = "Fairly Active")) +
  geom_point(aes(x = Calories, y = LightlyActiveMinutes, color = "Lightly Active")) +
  geom_point(aes(x = Calories, y = SedentaryMinutes, color = "Sedentary")) +
  labs(y = "Active_Minutes", color = "Activity Level")

```



#In this group most users are sedentary with an average time of 991 minutes, followed by LightlyActive

#As predicated, user that are very active spent more calories in less time. In comparison to those wh

#Intensities

```
intensities %>% summary()
```

```
##      Id          date   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
## Min.   : 0.00      Min.   : 0.00      Min.   :0.000000
## 1st Qu.: 0.00      1st Qu.: 0.00      1st Qu.:0.000000
## Median : 6.00      Median : 4.00      Median :0.000000
## Mean   : 13.56     Mean   : 21.16     Mean   :0.001606
## 3rd Qu.: 19.00     3rd Qu.: 32.00     3rd Qu.:0.000000
## Max.   :143.00     Max.   :210.00     Max.   :0.110000
## LightActiveDistance ModeratelyActiveDistance VeryActiveDistance
## Min.   : 0.000      Min.   :0.0000      Min.   : 0.000
## 1st Qu.: 1.945      1st Qu.:0.0000      1st Qu.: 0.000
## Median : 3.365      Median :0.2400      Median : 0.210
## Mean   : 3.341      Mean   :0.5675      Mean   : 1.503
## 3rd Qu.: 4.782      3rd Qu.:0.8000      3rd Qu.: 2.053
## Max.   :10.710      Max.   :6.4800      Max.   :21.920
## format
```

```

##  Length:940
##  Class :character
##  Mode   :character
##
## 
## 

ggplot(intensities)+  

  geom_point(aes(x = VeryActiveMinutes, y = VeryActiveDistance, color = "Very Active")) +  

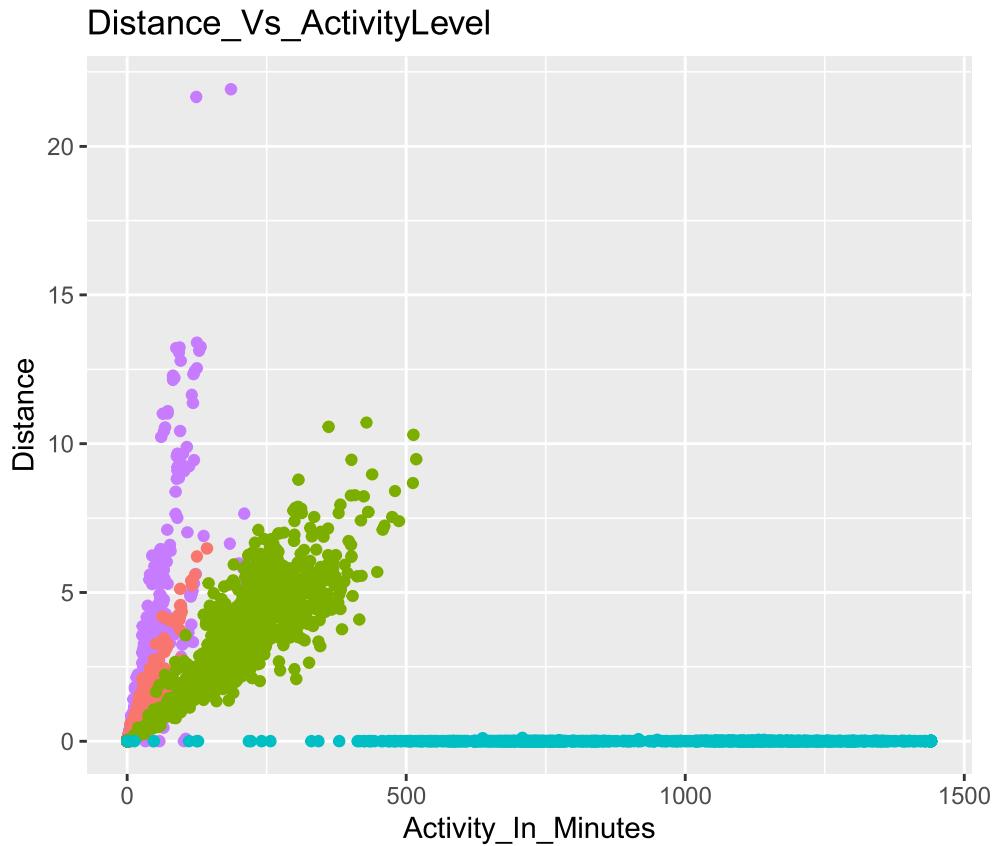
  geom_point(aes(x = FairlyActiveMinutes, y = ModeratelyActiveDistance, color = "Fairly Active")) +  

  geom_point(aes(x = LightlyActiveMinutes, y = LightActiveDistance, color = "Lightly Active")) +  

  geom_point(aes(x = SedentaryMinutes, y = SedentaryActiveDistance, color = "Sedentary")) +  

  labs(title = "Distance_Vs_ActivityLevel", x="Activity_In_Minutes", y="Distance")

```



#Shows the very activity covered more distance in fewer minutes. Those that where Lightly active were
#Per Cleveland Clinic Health Essentials, ""Researchers have done studies that say it's sometimes even b

```

#activity_vs_sleep
#merge activity and sleep
activityLevel_vs_sleep <- merge(activity, sleep, by = "Id")

# Create the new variable for Total Hours Asleep
activityLevel_vs_sleep <- activityLevel_vs_sleep %>%
  mutate(TotalHoursAsleep = TotalMinutesAsleep / 60)

#create a ggplot for each activity level vs TotalHoursAsleep

#activity_Level Very Active

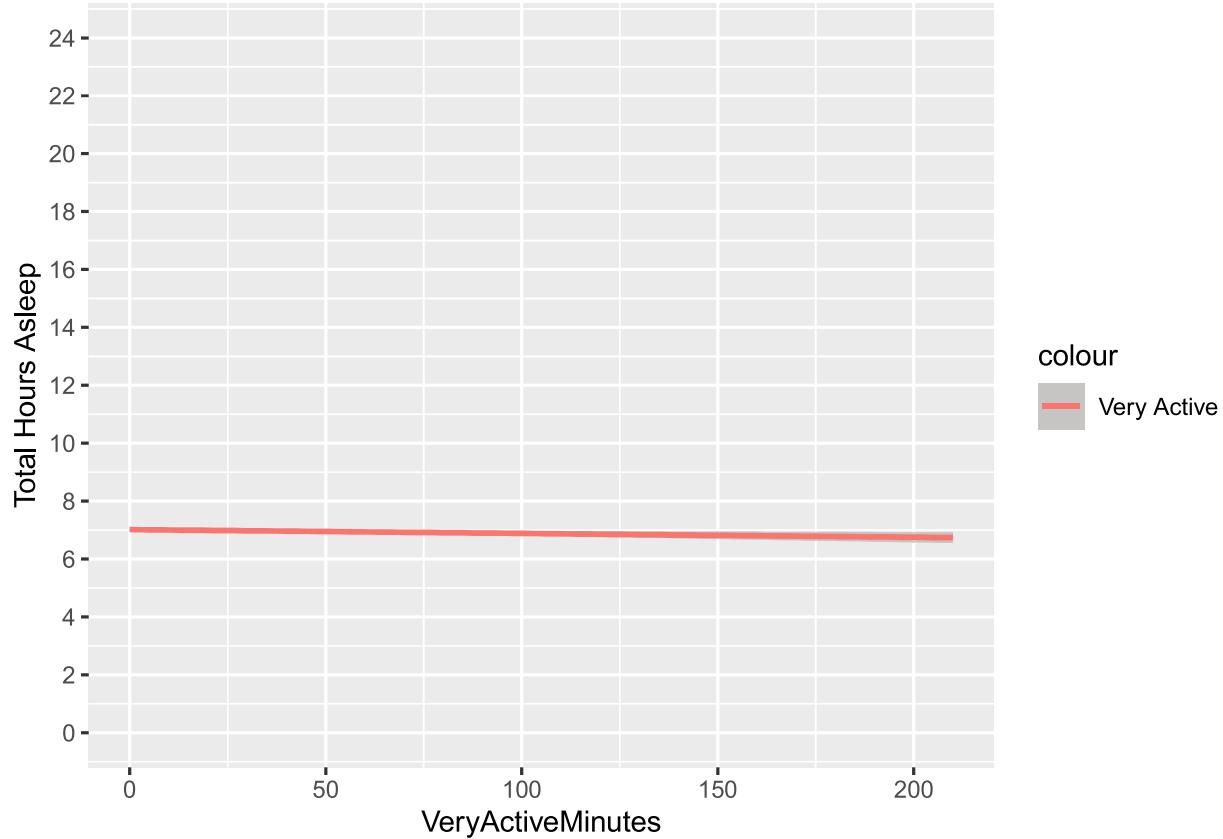
```

```

ggplot(activityLevel_vs_sleep) +
  geom_smooth(aes(x = VeryActiveMinutes, y = TotalHoursAsleep, color = "Very Active"), method = "lm") +
  scale_y_continuous(labels = comma, limits = c(0, 24), breaks = seq(0, 24, by = 2)) +
  labs(y = "Total Hours Asleep")

```

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

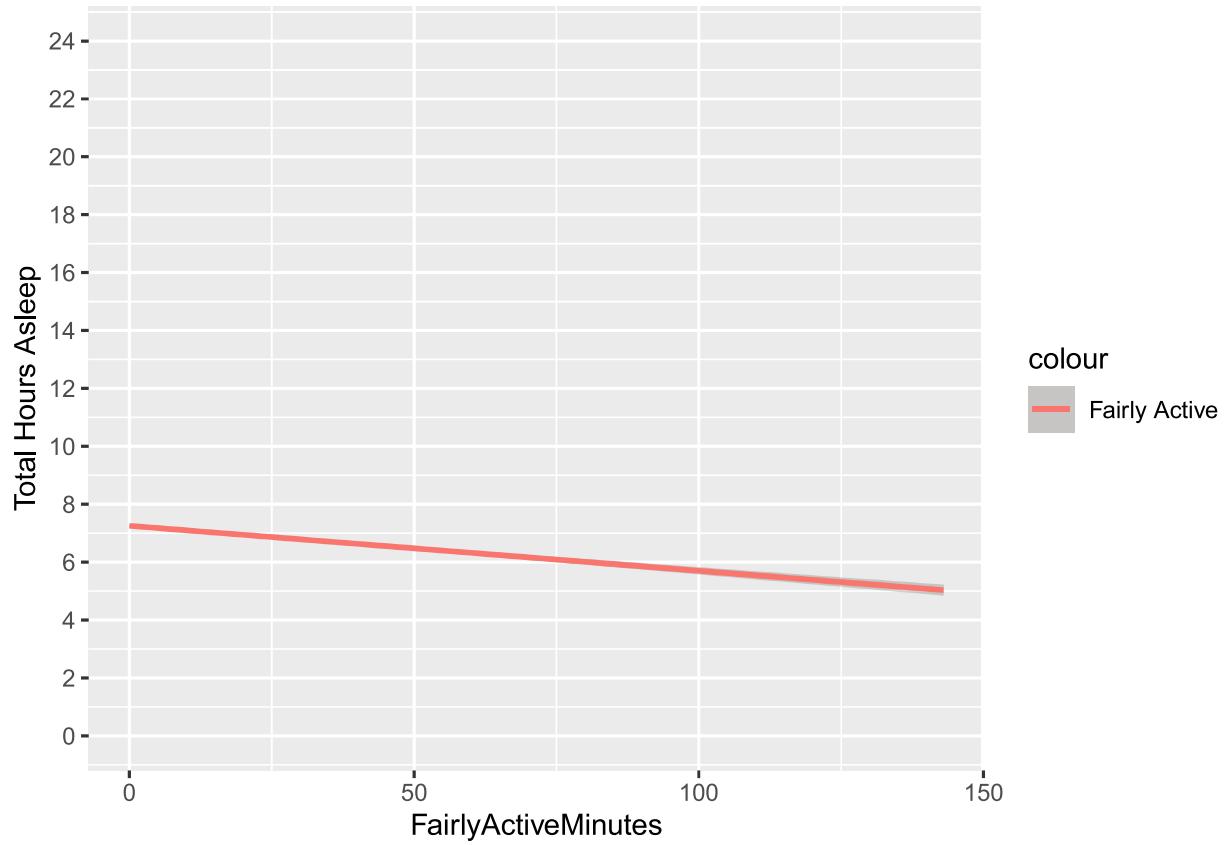


```

#activity_Level Fairly Active
ggplot(activityLevel_vs_sleep) +
  geom_smooth(aes(x = FairlyActiveMinutes, y = TotalHoursAsleep, color = "Fairly Active"), method = "lm") +
  scale_y_continuous(labels = comma, limits = c(0, 24), breaks=seq(0,24,by=2)) +
  labs(y = "Total Hours Asleep")

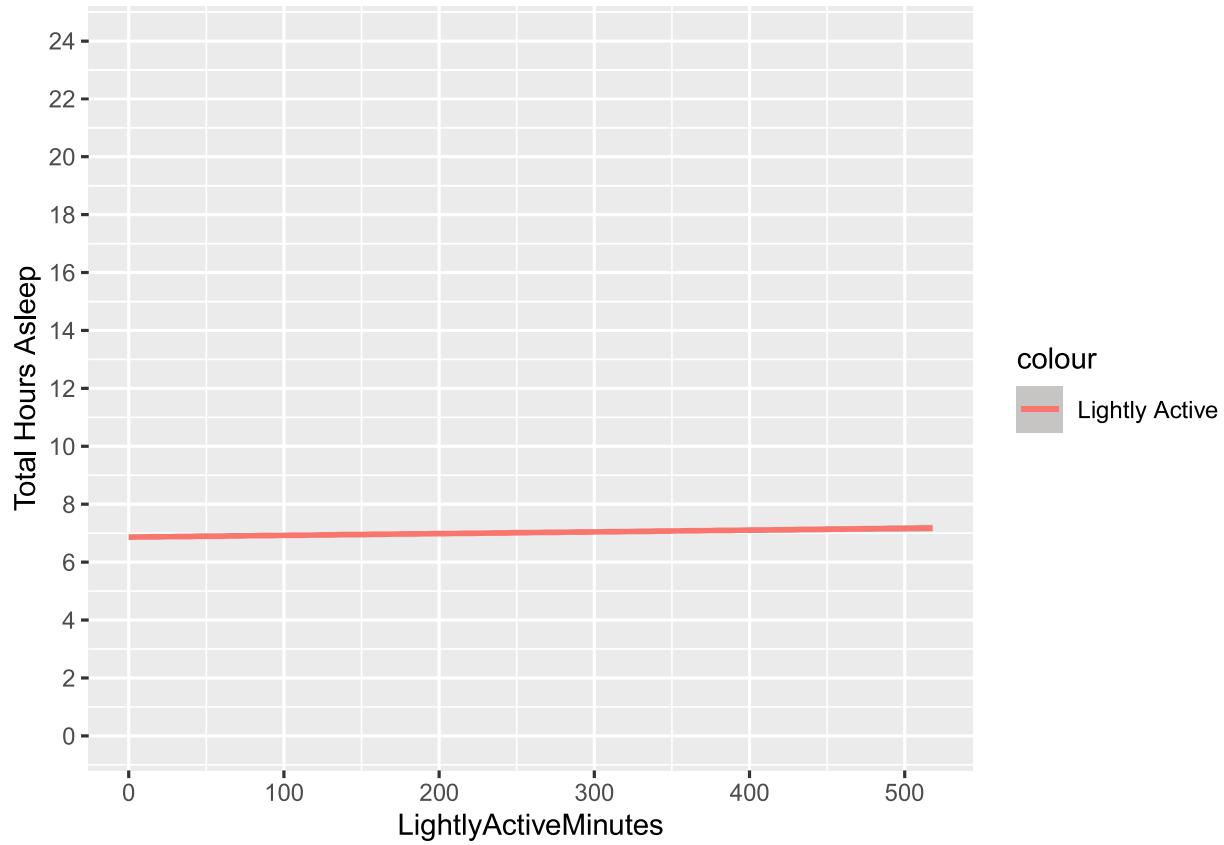
```

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



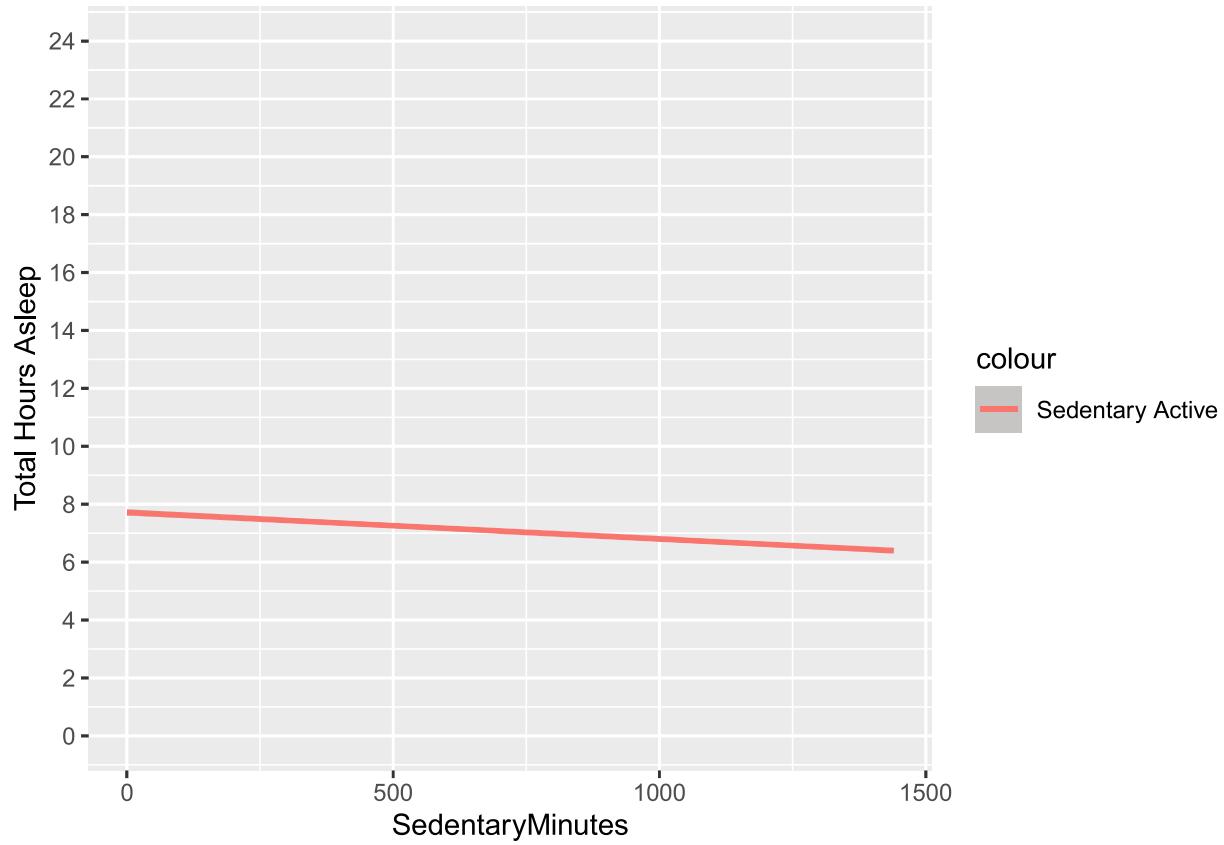
```
#activity_Level Lightly Active
ggplot(activityLevel_vs_sleep) +
  geom_smooth(aes(x = FairlyActiveMinutes, y = TotalHoursAsleep, color = "Fairly Active"), method =
  scale_y_continuous(labels = comma, limits = c(0, 24), breaks=seq(0,24,by=2)) +
  labs(y = "Total Hours Asleep")

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



```
#activity_level Sedentary Active
ggplot(activityLevel_vs_sleep) +
  geom_smooth(aes(x = SedentaryMinutes, y = TotalHoursAsleep, color = "Sedentary Active"), method = "loess",
  scale_y_continuous(labels = comma, limits = c(0, 24), breaks=seq(0,24,by=2)) +
  labs(y = "Total Hours Asleep")

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



#TotalHoursAsleep appears to have not positive affect on activity level, with the exception of LightA

Recommendations:

I recommend that Bellabeat launch a campaign targeting the sedentary group, which is the largest segment of participants. The campaign should encourage them to start with light, sustainable activities and gradually progress to higher-intensity workouts in shorter durations. Additionally, the app should include daily motivational points to keep users engaged.

Regarding bed versus sleep time, it's observed that people often spend more time in bed than actually sleeping, likely due to activities like watching TV or browsing social media. Implementing daily reminders to turn off media and TV could help users achieve better sleep quality.