R Notebook

Case Study -BellaBeat

Introduction : About the company

Bellabeat is a wellness Company ,that develops wearable computers for women.It was founded in 2013 by Sandro Mur and Urška Sršen. BellaBeat develop wearables and accompanying products that monitor biometric and lifestyle data. By collecting data on activity, sleep, stress, and reproductive health has allowed Bellabeat to empower women with knowledge about their own health and habits.Since 2013, Bellabeat has grown rapidly and quickly positioned itself as a tech-driven wellness company for women. Bellabeat empowers women to reconnect with themselves, unleash their inner strengths and be what they were meant to be via products like:

*Bellabeat app: The Bellabeat app provides users with health data related to their activity, sleep, stress, menstrual cycle, and mindfulness habits. This data can help users better understand their current habits and make healthy decisions. The Bellabeat app connects to their line of smart wellness products.

*Leaf: Bellabeat's classic wellness tracker can be worn as a bracelet, necklace, or clip. The Leaf tracker connects to the Bellabeat app to track activity, sleep, and stress.

*Time: This wellness watch combines the timeless look of a classic timepiece with smart technology to track user activity, sleep, and stress. The Time watch connects to the Bellabeat app to provide you with insights into your daily wellness.

*Spring: This is a water bottle that tracks daily water intake using smart technology to ensure that you are appropriately hydrated throughout the day. The Spring bottle connects to the Bellabeat app to track your hydration levels.

*Bellabeat membership: Bellabeat also offers a subscription-based membership program for users. Membership gives users 24/7 access to fully personalized guidance on nutrition, activity, sleep, health and beauty, and mindfulness based on their lifestyle and goals.

Ask Phase

Questions for Analysis. 1.What are some trends in the usage of non-bellabeat smart devices?

2.How could these trends apply tobellabeat customers?

3.how could these trends help influence bellabeat marketing strategy?

Business Task: To identify the frequency in the usage of smart devices among the users, further utilizing the insights to develop and improve the BellaBeat products thereby improving their marketing strategy.

Key Stakeholders: Urška Sršen: Bellabeat's cofounder and Chief Creative Officer.

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

Prepare Phase

DataSets Used:

FitBit Fitness Tracker Data (CC0: Public Domain, dataset made available through Mobius): This Kaggle data setcontains personal fitness tracker from thirty fitbit users. Thirty eligible Fitbit users consented to

the submission of personal tracker data, including minute-level output for physical activity, heart rate, and sleep monitoring. It includes information about daily activity, steps, and heart rate that can be used to explore users' habits. I chose the hourly_tracking datas for my analysis. the datasets are in narrow format. The datasets are organized by Activitydate and Activityhour. The datasets have thirty three unique Id's in common for all dataset.The dataset doesn't hold any details like height, age, location and other factors, so the bias in this dataset cannot be determined.The dataset has limited number of entries to analyse and also the data here is

1.reliable 2.open 3.comprehensive 4.current(data was collected on 2016). 5.cited.

Process Phase

The datasets have values for a month and are small.So sheets and SQL is used for data sorting and organizing.The data are collected for 30 users with distinct 33 Ids'. Not everyone in the group had their data fully logged also the documented properly ensuring its integrity.To avoid any errors, only data considered is formatted in the table. The datasets are named according to the data being considered.

```
install.packages('tidyverse')
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.2'
## (as 'lib' is unspecified)
```

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
## v ggplot2 3.3.6     v purrr   0.3.4
## v tibble  3.1.7     v dplyr   1.0.9
## v tidyr   1.2.0     v stringr 1.4.0
## v readr   2.1.2     v forcats 0.5.1
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
install.packages("ggplot2")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.2'
## (as 'lib' is unspecified)
```

```
library(ggplot2)
install.packages("tidyselect")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.2'
## (as 'lib' is unspecified)
```

```
library(tidyselect)
library(lubridate)
```

```
##
## Attaching package: 'lubridate'
## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```
install.packages("dplyr")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.2'
## (as 'lib' is unspecified)
```

```r
library(dplyr)
#import dataset

activity<- read.csv("dailyActivity.csv")


calorie<- read.csv("hourlyCalories_merged.csv")

intensity<-read.csv("hourlyIntensities_merged.csv")


step<-read.csv("hourlySteps_merged.csv")


weight<-read.csv("weightLogInfo_merged.csv")


 sleep<-read.csv("sleepDay_merged.csv")
```

```r
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 TotalMinutes
## 1                  13                  328              728         1094
## 2                  19                  217              776         1033
## 3                  11                  181             1218         1440
## 4                  34                  209              726          998
## 5                  10                  221              773         1040
## 6                  20                  164              539          761
##   Calories
## 1     1985
## 2     1797
## 3     1776
## 4     1745
```

```
## 5       1863
## 6       1728
```

```
colnames(activity)
```

```
##  [1] "Id"                   "ActivityDate"
##  [3] "TotalSteps"           "TotalDistance"
##  [5] "TrackerDistance"      "LoggedActivitiesDistance"
##  [7] "VeryActiveDistance"   "ModeratelyActiveDistance"
##  [9] "LightActiveDistance"  "SedentaryActiveDistance"
## [11] "VeryActiveMinutes"    "FairlyActiveMinutes"
## [13] "LightlyActiveMinutes" "SedentaryMinutes"
## [15] "TotalMinutes"         "Calories"
```

```
nrow(activity)
```

```
## [1] 940
```

```
nrow(calorie)
```

```
## [1] 22099
```

```
nrow(step)
```

```
## [1] 22099
```

```
nrow(sleep)
```

```
## [1] 413
```

```
nrow(weight)
```

```
## [1] 67
```

```
activity %>%
  select(TotalSteps,
         TotalDistance,
         SedentaryMinutes,
         Calories) %>%
  summary()
```

```
##    TotalSteps     TotalDistance    SedentaryMinutes    Calories
##  Min.   :    0   Min.   : 0.000   Min.   :    0.0   Min.   :   0
##  1st Qu.: 3790   1st Qu.: 2.620   1st Qu.: 729.8   1st Qu.:1828
##  Median : 7406   Median : 5.245   Median :1057.5   Median :2134
##  Mean   : 7638   Mean   : 5.490   Mean   : 991.2   Mean   :2304
##  3rd Qu.:10727   3rd Qu.: 7.713   3rd Qu.:1229.5   3rd Qu.:2793
##  Max.   :36019   Max.   :28.030   Max.   :1440.0   Max.   :4900
```
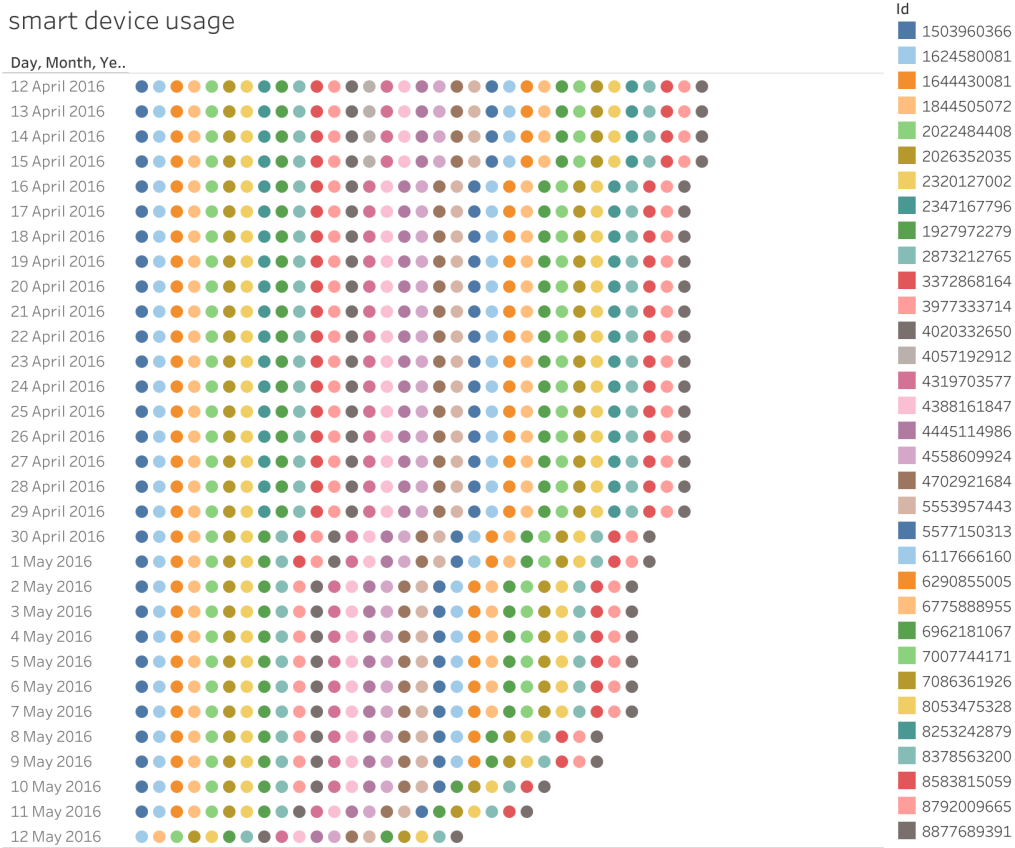
```
sleep %>%
  select(TotalMinutesAsleep) %>%
  summary()
```

```
##  TotalMinutesAsleep
##  Min.   : 58.0
##  1st Qu.:361.0
##  Median :433.0
##  Mean   :419.5
##  3rd Qu.:490.0
##  Max.   :796.0
```
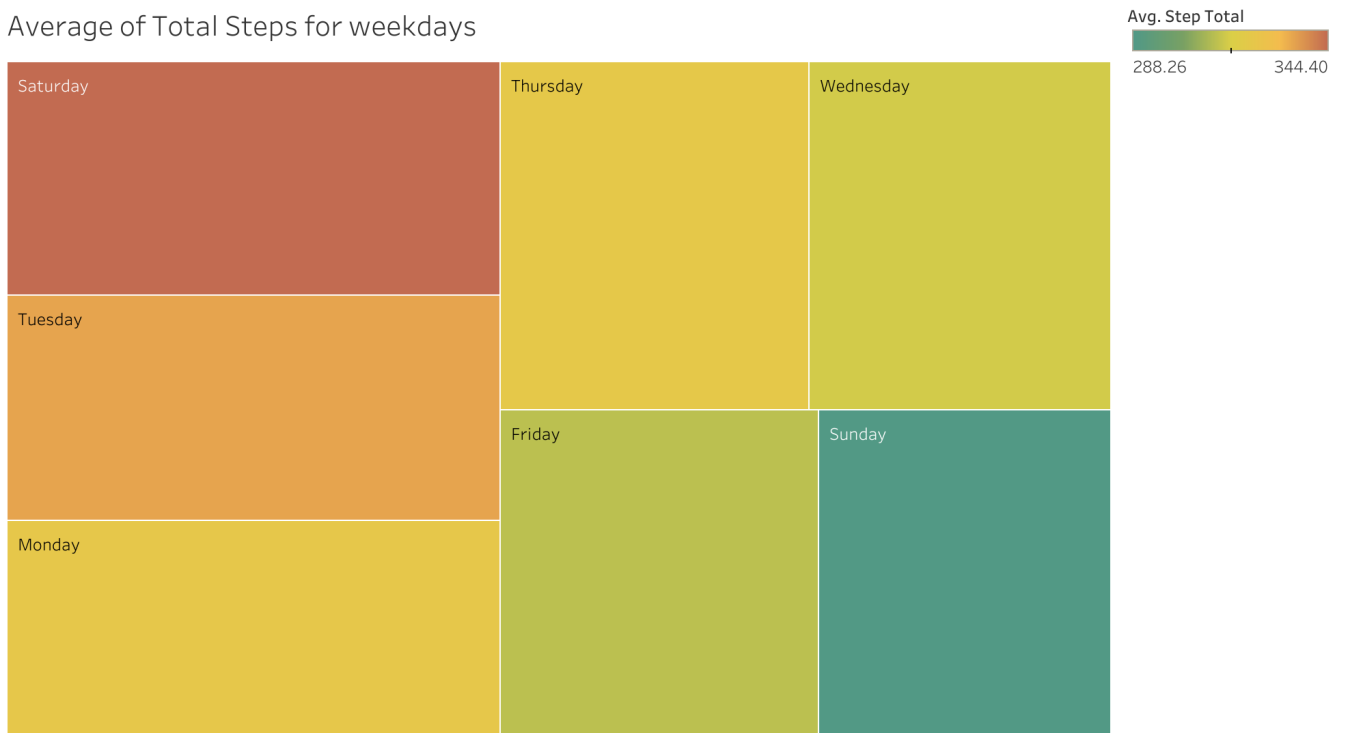
Analyse Phase

Usage of smart device

The users were all active for a week and their usage decline as days move on.The below plot shows their decline in usage of non-BellaBeat smart device

Average of Total Steps for weekdays

Avg. Step Total
288.26 — 344.40

Saturday	Thursday	Wednesday
Tuesday		
	Friday	Sunday
Monday		

Organizing data From the given dataset, mean steps, mean calorie, mean distance and mean sleep are calculated.

```
#organize
 MeanSteps<- activity %>% group_by(Id) %>%
  summarise(mean_steps=mean(TotalSteps))

 MeanCalorie<- activity %>% group_by(Id) %>%
   summarise(mean_calorie=mean(Calories))

 MeanDistance<- activity %>% group_by(Id) %>%
   summarise(mean_distance=mean(TotalDistance))

 MeanSleep<- sleep %>% group_by(Id) %>%
   summarise(mean_sleep=mean(TotalMinutesAsleep))

 ActivityStatus<-MeanSteps %>% mutate(activityStatus=case_when(mean_steps>10000 ~ "Active user", mean_s
```
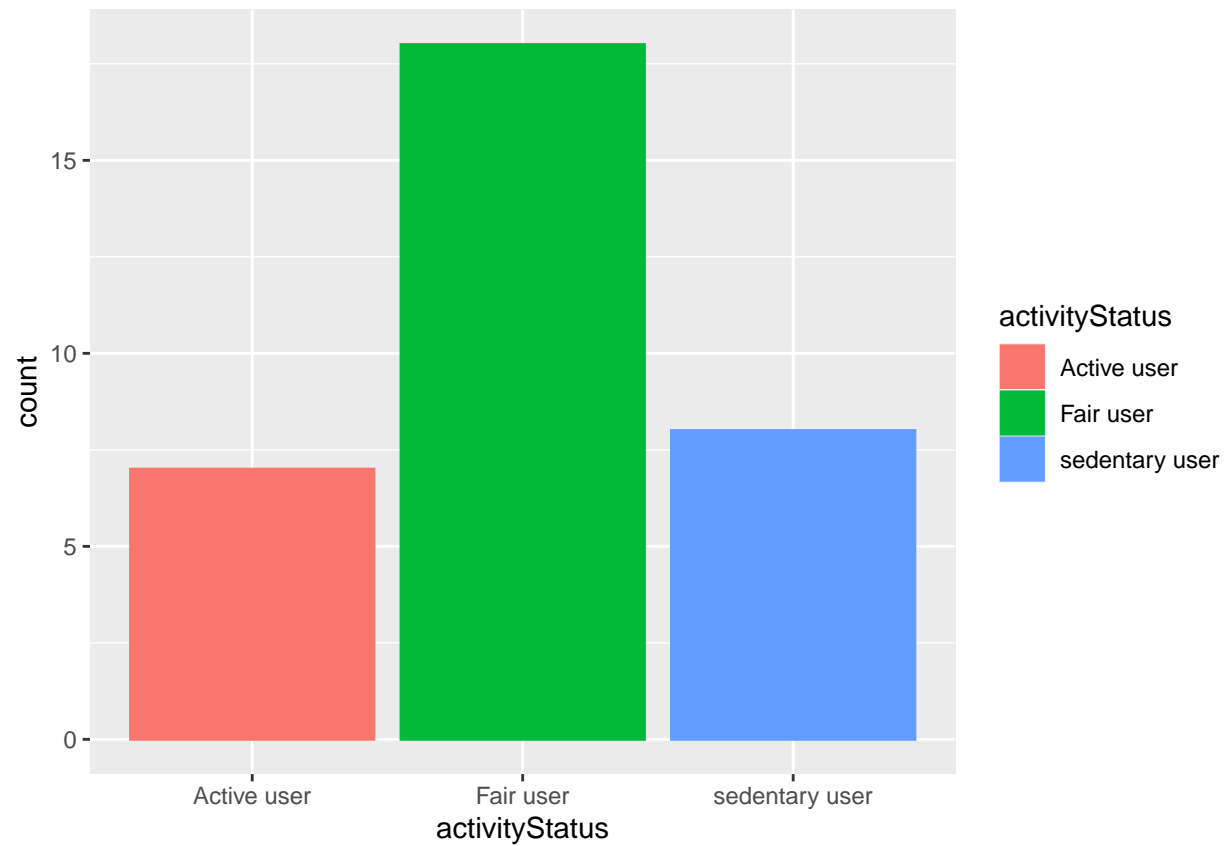
plotting based on activity status

```
ggplot(data=ActivityStatus)+geom_bar(mapping=aes(x=activityStatus,color=activityStatus,fill=activitySta
```
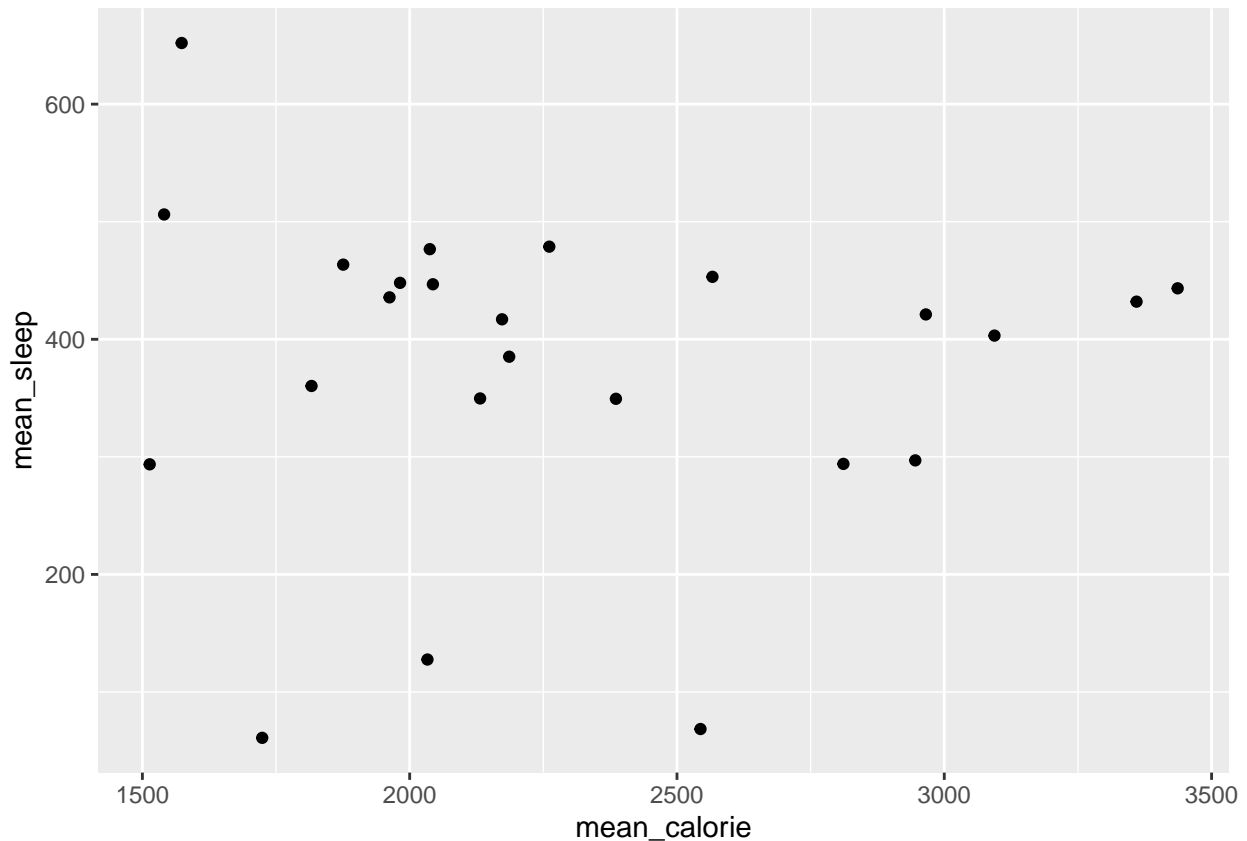


```
Mean_table_1 <- merge(ActivityStatus,MeanDistance, by="Id")


Mean_table_2 <- merge(MeanSleep,MeanCalorie, by="Id")


FinalTable <- merge(Mean_table_1,Mean_table_2, by="Id")


ggplot(data = FinalTable)+geom_point(mapping = aes(x=mean_calorie,y=mean_sleep))
```

Share & Act Phase

Key Findings

1.The Usage of smart devices reduces gradually as days passes.

2.From the analysis, it is found that out of 33 users only 24 users had their data logged.

Recommendations:

The BellaBeat products must be designed in a way that it can be worn everyday also must send notifications to its user via the BellaBeat app about the poor battery or any defect in case of failure to record the data.

The app must ask its user for their age, heigth, weight and their daily active status and must give them various plans accordingly.

the app also should let its user to set their daily step target , daily sleep minutes and also their calorie intake.

It should also send prior notifications to its user if they fail to meet their daily step ,sleep,calorie target also a weekly report of their past week activity status should be given to improve their involvement.