DISCOVERING OPPORTUNITES

MARKETING STRATEGIES

Bellabeat | Marketing Analytics Team

SUMMARY

Bellabeat, a successful company that manufactures high-tech health focused products for women has a great potential to become a global player in the smart-device market. We believe that analyzing the authentic and publicly available smart-device usage data will reveal ideas that can unlock growth opportunities for the company. Marketing Analytics Team has gathered, processed and analyzed smart-device usage data to highlight such potential opportunities of growth and vital insights have been presented in this report.

Depending on our various products, there are numerous opportunities for our company to grow by making data-driven decisions.

INTRODUCTION

BELLABEAT

Urška Sršen and Sando Mur have founded Bellabeat in 2013. Since its foundation, Bellabeat has quickly placed itself in the tech-driven wellness company for women. After successful launch of offices around the world and simultaneously empowering women with beautifully designed technology, Bellabeat is interested to improve marketing strategies for their products.

BUSINESS TASK

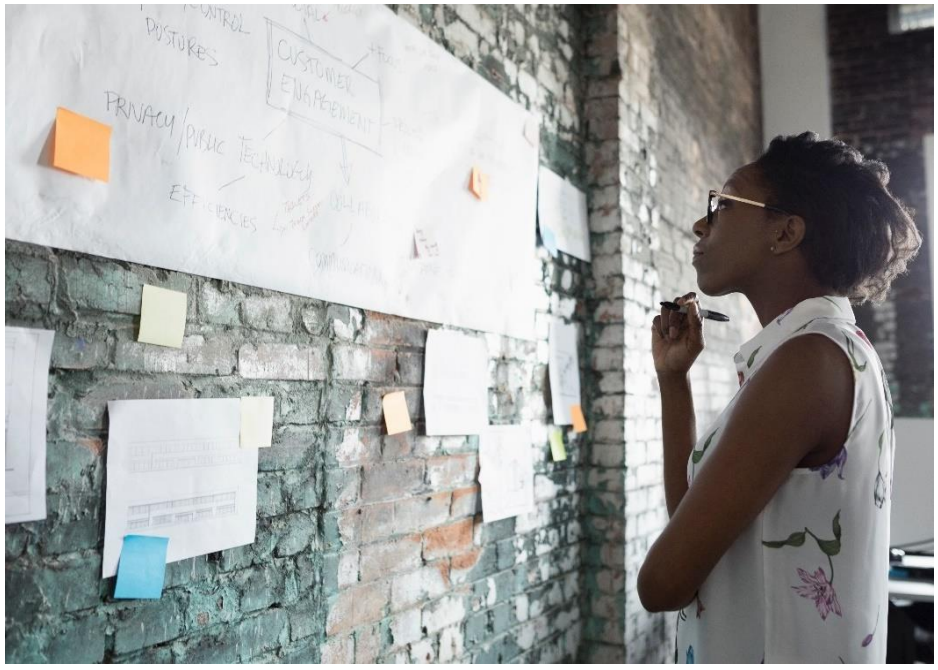
To analyze publicly available smart device usage data, gain vital insights on how consumers use smart devices and finally make data-driven decisions that will create growth opportunities for Bellabeats' products and the business.

KEY STAKEHOLDERS

Urška Sršen

Sando Mur

Marketing Analytics Team



THE ANALYSIS

BUSINESS PRIORITIES

- Maintain growth
- Value our stakeholder's needs

ADDED PRIORITIES

- Maintain service quality standards
- Explore product-wise growth opportunities
- Make data-drive decisions

DATA GATHERING & STORAGE

Users' smart-device usage data has been acquired from a public domain through Mobius. This Kaggle dataset contains personal fitness tracker from 31 Fitbit users who consented to the submission of personal tracker data, including minute-level output for physical activity, heart-rate and sleep monitoring.

Data has been downloaded to a secure drive of a server with encryption to avoid any breaches. The data is available for key personnel only with proper access credentials. The dataset has been stored in the format of ".csv (comma separated values) file" and ".xlsx (spreadsheets) files".

TRENDS: HOW THESE CAN HELP?

Analysing trends in this data can help us create new themes, applications and even create new ways to help users monitor their health easily.

Analysing such trends helps us to improve our products that we already offer; such as new visualization capabilities help users to understand their own data easily, we could devise new products specific for certain category of consumers (not everything is needed is one product). Understanding what user needs is important so that our product just delivers that.

These trends can help marketing strategy team to accurately target customers who need our products.

PROCESSING AND ANALYSIS

The team has been presented with the dataset and communicated with our stakeholders about the time frame of this task. Proper means of communication and frequent meetings with stakeholders has ensured that data is clean and ready to be analyzed.

The primary cleaning and processing method used here on this dataset is through MySQL server and Excel software which helps the team to collaboratively work on more than one data simultaneously. While keeping a track of sources of data, we are ensured that data is Reliable, Original, Current, Comprehensible and Cited properly (ROCCC). Highlighting sources of data and their reliability is vital to remove any sort of biases while analyzing the dataset.

The following sources suffices our dataset's ROCCC :

A) Fitbit: www.fitbit.com
B) Kaggle-Mobius User Dataset: FitBit Fitness Tracker Data | Kaggle
C) MySQL workbench Server and MS-OFFICE: Excel

MySQL :: Developer Zone

D) R : publicly available free data visualization software
www.r-project.org

The dataset gathered was almost ready-to-use except for situations where the dataset was not in a comprehensible format. For example, dates recorded in each dataset had different format revealing the possibility of users belonging from various regions across the world. Appropriate date formatting techniques were applied to ensure data consistency.

Example: the below code runs in MySQL helps to maintain dates in 'yyyy/mm/dd' format

$$\textbf{mysql} > \text{DATE\_FORMAT}(\text{Dataset}, '\%Y/\%M/\%D')$$

Moreover, most datasets have both date and time recorded into one cell. This makes analyzing a bit tough. To simplify, we have used pivot tables and the following commands in MS-Excel.

$$\text{RIGHT}(\text{number\_of\_characters\_to\_be\_separated\_from}, \text{cell\_name})$$

This command successfully separates date and time, then places them in separate columns. Also, the below command formats all the time records in the form of 'hh:mm:ss AM/PM':

$$\text{TEXT}(\text{cell\_name}, \text{"hh: mm: ss AM/PM"})$$

Rest of the processing and analysis has been performed in R. Following sections of the report will highlight key trends and will reveal opportunities that can enhance our business performance.

Also, there is a key observation that can help us improve our standards. Not all users tracked their activity. We narrowed down users who tracked their daily steps, calories, minutes, distance and sleep by using the following set of commands in the MySQL server:

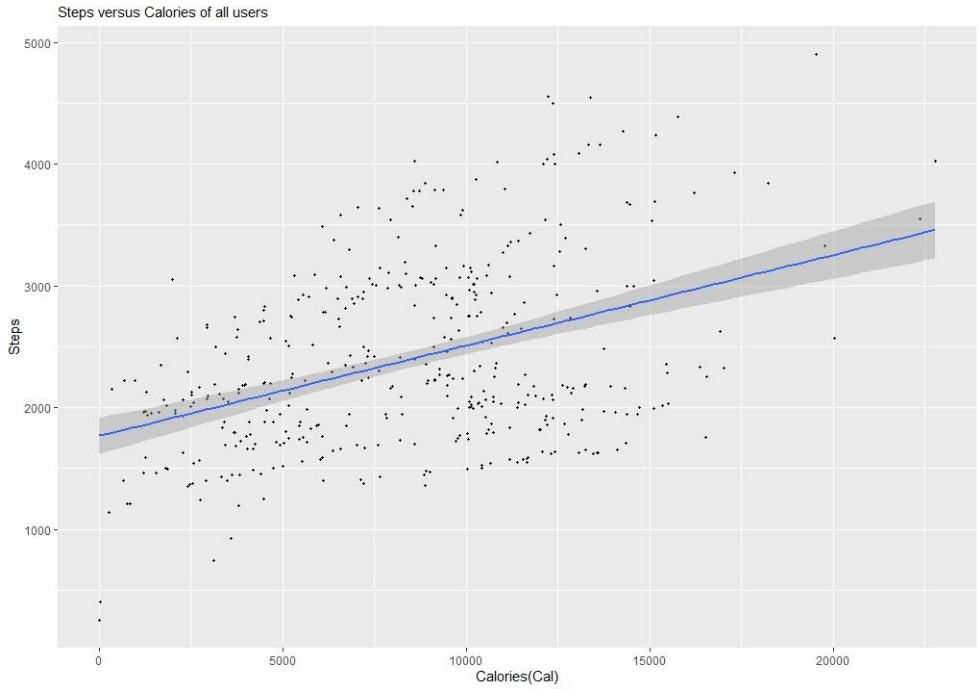
```
USE Bellabeat_DB
SELECT
    *
FROM
    dailyactivity AS A
INNER JOIN
    sleep AS S ON A.Id = S.Id
WHERE
    A.Date = S.DateLogged
```

WHAT DATA IS TELLING US?

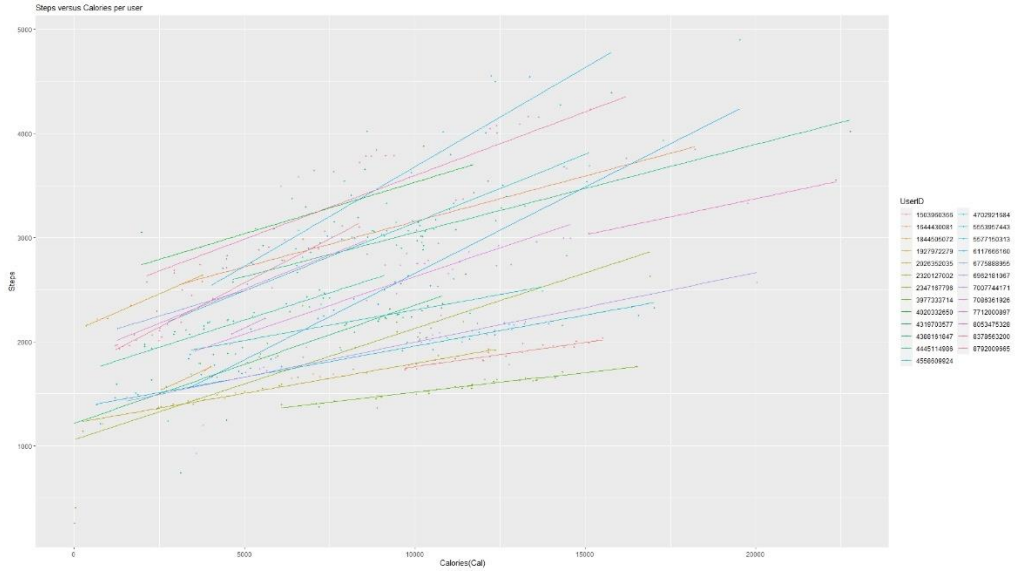
The user data has lot of trends and can help us to strategize our approach towards growth of Bellabeat.

First, observations reveal that not all users monitored daily activity, heart rate, sleep and weight. This tells us that users specifically use each service as per their need. Most users tracked their daily activity, while some users tracked only heart rate and some tracked multiple activities such as both sleep and weight or calories, sleep and weight.

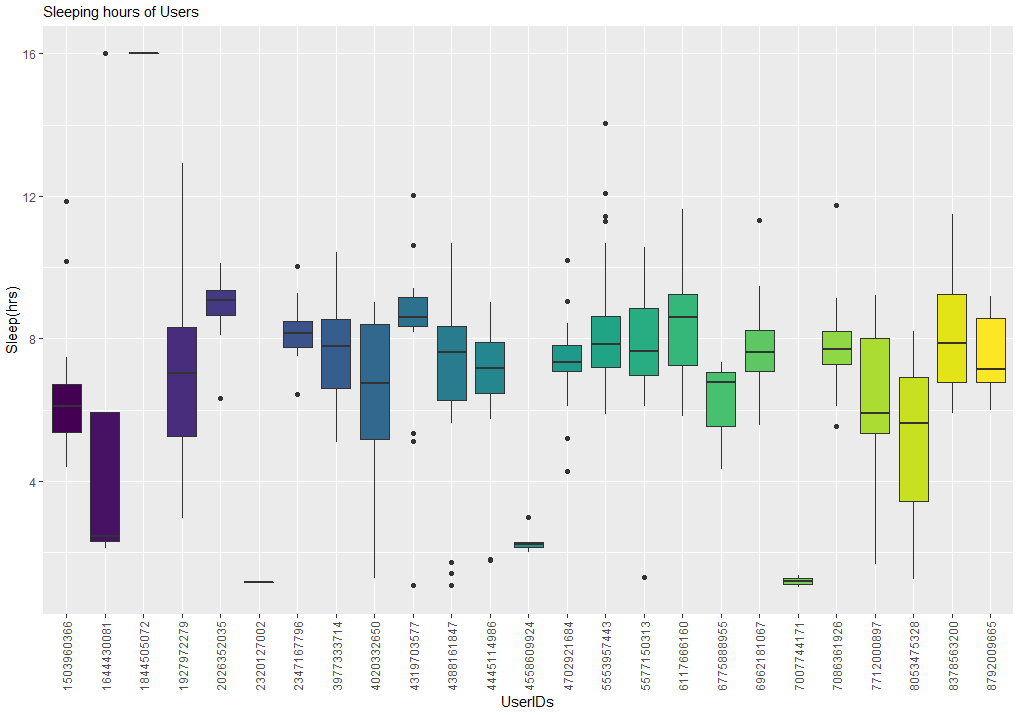
Diving deep into calories burnt versus steps taken by each user reveals a strong relationship between the variables.



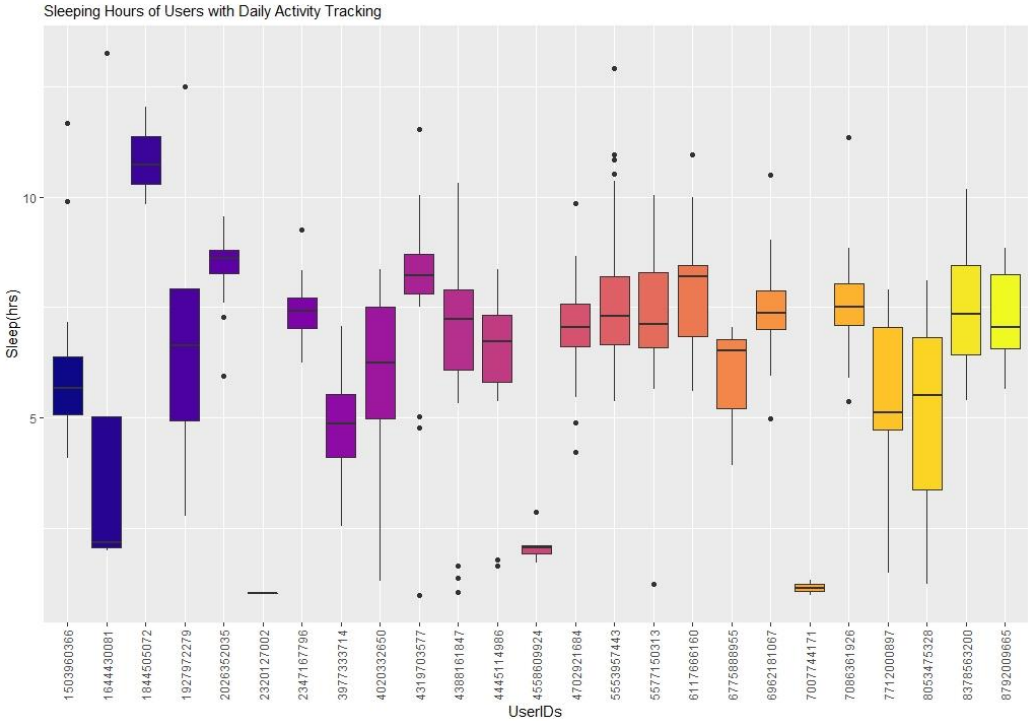
As a user takes more steps, user burns more calories. From below graph we can see that users are insterested to track their steps and distnce more rigorously than other parameters.



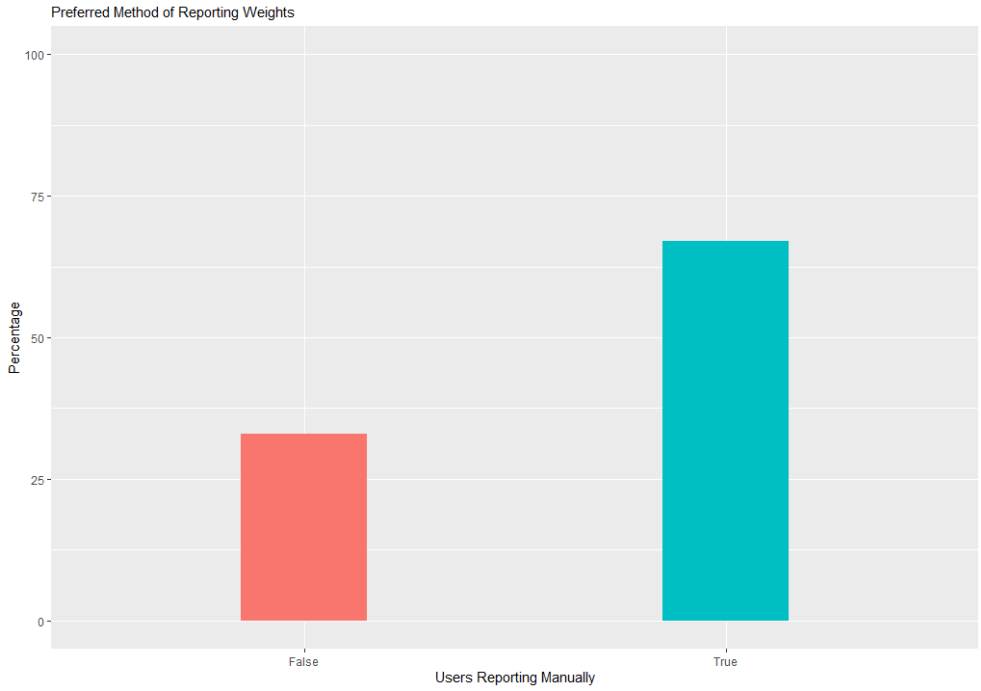
Furthermore, sleep patterns reveal that irrespective of how rigorous daily activity can be, sleeping hours seems to be ranging between 5 hours and 10 hours on average. There are exceptions however, maybe due to working hours or for some unknown reasons, sleep lasted for longer periods and shorter periods for some users.



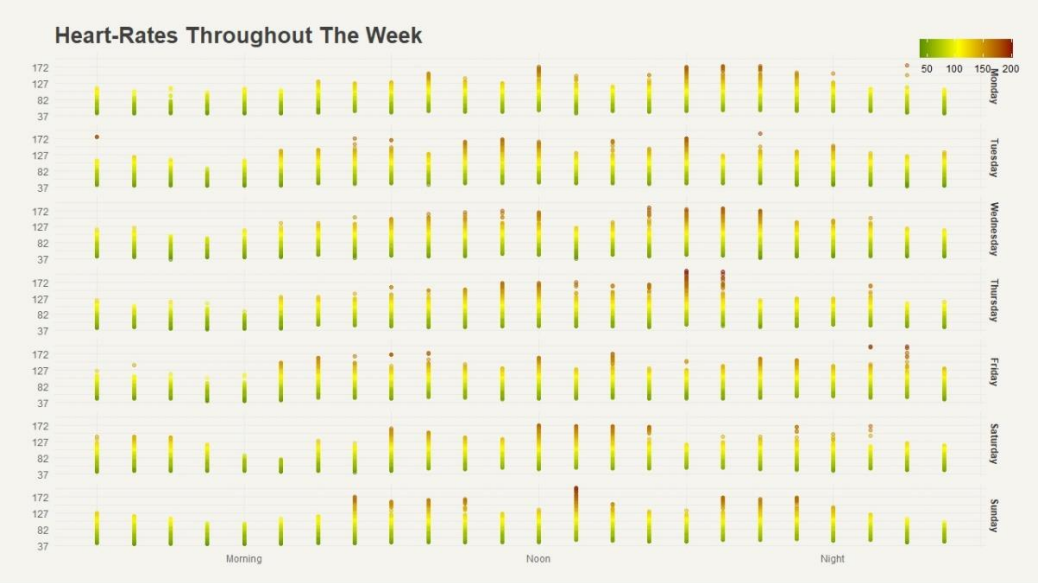
This trend appears to be same for users who tracked their daily activity, steps, minutes, calories as well as sleep consecutive dates.



Looking at the users' weight records revealed that users are interested in tracking their weights manually and quite frequently. About 67% logs in weights are manually reported.



The sleep patterns can be related to the user's heart rates per week. As shown below, It is observed that high heartrates are achieved from morning till evening hours. Also, comparing heartrates during weekdays and weekends (Saturdays & Sundays) revealed activity is more during the weekdays due to work. Nonetheless, it is clear that users are



keen on keeping track of their heartrates irrespective of their activity and day unlike weight logs. After analyzing the fitness data, vital insights can be drawn and these insights can be useful in improving our business.

WHAT'S NEXT

The following conclusions and insights will help Bellabeat, plan new methods that will help the company to deploy useful tools through Bellabeat's smart devices and help consumers to track their fitness in a simple yet comprehensive manner. In this busy world, simple and comprehensive tools are best for users to improve their fitness.

INSIGHTS

Setting 'Reminders' or 'Targets' can be a fruitful way to help users.

- **Bellabeat can motivate users to improve fitness by providing a flexible way of setting targets and reminders.** This is apt for 'sleep' to improve sleeping patterns and sleep quality, 'steps' to motivate users to exercise more and 'heartrates' to keep track of their heart's health by providing some facts and tips to improve fitness.
- **For users who are interested in loosing or gaining weight, Bellabeat can suggest plans or device tools that can help them track their calorie intake along with hydration.** Bellabeat can devise an enhancement in their smart devices where users can log their food and water intake. Motivating to keep a healthy diet can improve user's fitness.
- **Bellabeat can create a community pages or activities for like-minded users where healthy competition is encouraged towards common fitness goals.** Users can participate in virtual sessions – online or in-person at some sessions that offer fitness. In this way, Bellabeat can truly become a trustworthy company that leads in providing intelligent devices for fitness tracking.
- **There is a need of additional data such as users' geography and occupation (engineer or athlete) to help users improve their fitness by suggesting them exactly what they need.** More surveys must be performed and data must be collected in a diverse manner to remove any form of biases. This can be achieved via arranging campaigns, online surveys or forms and updating Bellabeat's user profile forms.

APPENDIX A

CODING : MYSQL

A1. Daily Activity – Data Base

```
-- MySQL dump 10.13  Distrib 8.0.25, for Win64 (x86_64)
--
-- Host: 127.0.0.1    Database: bellabeat_db
-- ------------------------------------------------------
-- Server version           8.0.26

/*!40101 SET @OLD_CHARACTER_SET_CLIENT=@@CHARACTER_SET_CLIENT */;
/*!40101 SET @OLD_CHARACTER_SET_RESULTS=@@CHARACTER_SET_RESULTS */;
/*!40101 SET @OLD_COLLATION_CONNECTION=@@COLLATION_CONNECTION */;
/*!50503 SET NAMES utf8 */;
/*!40103 SET @OLD_TIME_ZONE=@@TIME_ZONE */;
/*!40103 SET TIME_ZONE='+00:00' */;
/*!40014 SET @OLD_UNIQUE_CHECKS=@@UNIQUE_CHECKS, UNIQUE_CHECKS=0 */;
/*!40014 SET @OLD_FOREIGN_KEY_CHECKS=@@FOREIGN_KEY_CHECKS, FOREIGN_KEY_CHECKS=0 */;
/*!40101 SET @OLD_SQL_MODE=@@SQL_MODE, SQL_MODE='NO_AUTO_VALUE_ON_ZERO' */;
/*!40111 SET @OLD_SQL_NOTES=@@SQL_NOTES, SQL_NOTES=0 */;

--
-- Table structure for table `dailyactivity`
--

DROP TABLE IF EXISTS `dailyactivity`;
/*!40101 SET @saved_cs_client     = @@character_set_client */;
/*!50503 SET character_set_client = utf8mb4 */;
CREATE TABLE `dailyactivity` (
  `Id` INT NOT NULL,
  `ActivityDate` VARCHAR(45) NOT NULL,
  `Date` DATE NOT NULL,
  `TotalSteps` INT NOT NULL,
  `TotalDistance` INT NOT NULL,
  `TrackerDistance` INT NOT NULL,
  `LogActDistance` INT NOT NULL,
  `VeryActDistance` INT NOT NULL,
  `ModerateActDistance` INT NOT NULL,
  `LightActDistnace` INT NOT NULL,
  `SedentaryActDistance` INT NOT NULL,
  `VeryActiveMins` INT NOT NULL,
  `FairlyActiveMins` INT NOT NULL,
  `LightlyActiveMins` INT NOT NULL,
  `SedentaryActiveMins` INT NOT NULL,
  `Calories` INT NOT NULL,
```

```
  PRIMARY KEY (`Id`)
)  ENGINE=INNODB DEFAULT CHARSET=UTF8MB3;
/*!40101 SET character_set_client = @saved_cs_client */;

--
-- Dumping data for table `dailyactivity`
--

LOCK TABLES `dailyactivity` WRITE;
/*!40000 ALTER TABLE `dailyactivity` DISABLE KEYS */;
/*!40000 ALTER TABLE `dailyactivity` ENABLE KEYS */;
UNLOCK TABLES;
/*!40103 SET TIME_ZONE=@OLD_TIME_ZONE */;

/*!40101 SET SQL_MODE=@OLD_SQL_MODE */;
/*!40014 SET FOREIGN_KEY_CHECKS=@OLD_FOREIGN_KEY_CHECKS */;
/*!40014 SET UNIQUE_CHECKS=@OLD_UNIQUE_CHECKS */;
/*!40101 SET CHARACTER_SET_CLIENT=@OLD_CHARACTER_SET_CLIENT */;
/*!40101 SET CHARACTER_SET_RESULTS=@OLD_CHARACTER_SET_RESULTS */;
/*!40101 SET COLLATION_CONNECTION=@OLD_COLLATION_CONNECTION */;
/*!40111 SET SQL_NOTES=@OLD_SQL_NOTES */;

-- Dump completed on 2022-06-07 21:48:11
```

A2. Sleep Time – Data Base

```
-- MySQL dump 10.13  Distrib 8.0.25, for Win64 (x86_64)
--
-- Host: 127.0.0.1    Database: bellabeat_db
-- ------------------------------------------------------
-- Server version           8.0.26

/*!40101 SET @OLD_CHARACTER_SET_CLIENT=@@CHARACTER_SET_CLIENT */;
/*!40101 SET @OLD_CHARACTER_SET_RESULTS=@@CHARACTER_SET_RESULTS */;
/*!40101 SET @OLD_COLLATION_CONNECTION=@@COLLATION_CONNECTION */;
/*!50503 SET NAMES utf8 */;
/*!40103 SET @OLD_TIME_ZONE=@@TIME_ZONE */;
/*!40103 SET TIME_ZONE='+00:00' */;
/*!40014 SET @OLD_UNIQUE_CHECKS=@@UNIQUE_CHECKS, UNIQUE_CHECKS=0 */;
/*!40014 SET @OLD_FOREIGN_KEY_CHECKS=@@FOREIGN_KEY_CHECKS, FOREIGN_KEY_CHECKS=0 */;
/*!40101 SET @OLD_SQL_MODE=@@SQL_MODE, SQL_MODE='NO_AUTO_VALUE_ON_ZERO' */;
/*!40111 SET @OLD_SQL_NOTES=@@SQL_NOTES, SQL_NOTES=0 */;

--
-- Table structure for table `sleep`
--

DROP TABLE IF EXISTS `sleep`;
/*!40101 SET @saved_cs_client     = @@character_set_client */;
/*!50503 SET character_set_client = utf8mb4 */;
CREATE TABLE `sleep` (
  `Id` int NOT NULL,
```

```
  `SleepDay` date NOT NULL,
  `SleepRecords` int NOT NULL,
  `TotminsAsleep` int NOT NULL,
  `TotTimeBed` int NOT NULL,
  PRIMARY KEY (`Id`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb3;
/*!40101 SET character_set_client = @saved_cs_client */;

--
-- Dumping data for table `sleep`
--

LOCK TABLES `sleep` WRITE;
/*!40000 ALTER TABLE `sleep` DISABLE KEYS */;
INSERT INTO `sleep` VALUES (1503960366,'2016-04-12',1,327,346),(1644430081,'2016-04-
29',1,119,127),(1844505072,'2016-04-15',1,644,961),(1927972279,'2016-04-12',3,750,775),(2026352035,'2016-04-
12',1,503,546);
/*!40000 ALTER TABLE `sleep` ENABLE KEYS */;
UNLOCK TABLES;
/*!40103 SET TIME_ZONE=@OLD_TIME_ZONE */;

/*!40101 SET SQL_MODE=@OLD_SQL_MODE */;
/*!40014 SET FOREIGN_KEY_CHECKS=@OLD_FOREIGN_KEY_CHECKS */;
/*!40014 SET UNIQUE_CHECKS=@OLD_UNIQUE_CHECKS */;
/*!40101 SET CHARACTER_SET_CLIENT=@OLD_CHARACTER_SET_CLIENT */;
/*!40101 SET CHARACTER_SET_RESULTS=@OLD_CHARACTER_SET_RESULTS */;
/*!40101 SET COLLATION_CONNECTION=@OLD_COLLATION_CONNECTION */;
/*!40111 SET SQL_NOTES=@OLD_SQL_NOTES */;

-- Dump completed on 2022-06-07 21:48:12
```

A3. Weight Logs – Data Base

```
-- MySQL dump 10.13  Distrib 8.0.25, for Win64 (x86_64)
--
-- Host: 127.0.0.1    Database: bellabeat_db
-- ------------------------------------------------------
-- Server version           8.0.26

/*!40101 SET @OLD_CHARACTER_SET_CLIENT=@@CHARACTER_SET_CLIENT */;
/*!40101 SET @OLD_CHARACTER_SET_RESULTS=@@CHARACTER_SET_RESULTS */;
/*!40101 SET @OLD_COLLATION_CONNECTION=@@COLLATION_CONNECTION */;
/*!50503 SET NAMES utf8 */;
/*!40103 SET @OLD_TIME_ZONE=@@TIME_ZONE */;
/*!40103 SET TIME_ZONE='+00:00' */;
/*!40014 SET @OLD_UNIQUE_CHECKS=@@UNIQUE_CHECKS, UNIQUE_CHECKS=0 */;
/*!40014 SET @OLD_FOREIGN_KEY_CHECKS=@@FOREIGN_KEY_CHECKS, FOREIGN_KEY_CHECKS=0 */;
/*!40101 SET @OLD_SQL_MODE=@@SQL_MODE, SQL_MODE='NO_AUTO_VALUE_ON_ZERO' */;
/*!40111 SET @OLD_SQL_NOTES=@@SQL_NOTES, SQL_NOTES=0 */;

--
-- Table structure for table `weightlogs`
--
```

```
DROP TABLE IF EXISTS `weightlogs`;
/*!40101 SET @saved_cs_client     = @@character_set_client */;
/*!50503 SET character_set_client = utf8mb4 */;
CREATE TABLE `weightlogs` (
  `Id` int NOT NULL,
  `Date` date NOT NULL,
  `Time` varchar(45) NOT NULL,
  `WeightLogKg` int NOT NULL,
  `WeightLoglbs` int NOT NULL,
  `Fat` int DEFAULT NULL,
  `BMI` int NOT NULL,
  `ManualReport` varchar(45) NOT NULL,
  `WeightLogsID` int NOT NULL,
  PRIMARY KEY (`Id`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb3;
/*!40101 SET character_set_client = @saved_cs_client */;

--
-- Dumping data for table `weightlogs`
--

LOCK TABLES `weightlogs` WRITE;
/*!40000 ALTER TABLE `weightlogs` DISABLE KEYS */;
/*!40000 ALTER TABLE `weightlogs` ENABLE KEYS */;
UNLOCK TABLES;
/*!40103 SET TIME_ZONE=@OLD_TIME_ZONE */;

/*!40101 SET SQL_MODE=@OLD_SQL_MODE */;
/*!40014 SET FOREIGN_KEY_CHECKS=@OLD_FOREIGN_KEY_CHECKS */;
/*!40014 SET UNIQUE_CHECKS=@OLD_UNIQUE_CHECKS */;
/*!40101 SET CHARACTER_SET_CLIENT=@OLD_CHARACTER_SET_CLIENT */;
/*!40101 SET CHARACTER_SET_RESULTS=@OLD_CHARACTER_SET_RESULTS */;
/*!40101 SET COLLATION_CONNECTION=@OLD_COLLATION_CONNECTION */;
/*!40111 SET SQL_NOTES=@OLD_SQL_NOTES */;

-- Dump completed on 2022-06-07 21:48:12
```

A4. Merging – Data Base

```
----Merging required data----------------------

USE Bellabeat_DB
SELECT
        *
FROM
        dailyactivity AS A
INNER JOIN
        sleep AS S ON A.Id = S.Id
WHERE
        A.   Date = S.DateLogged
```

APPENDIX B

CODING : R

```
library(readxl)
library(crayon)
library(stringr)
library(stringi)
library(openxlsx)
library(janitor)
library(tidyverse)
library(lubridate)
library(dplyr)
library(ggplot2)
library(plotly)
library(plotrix)
library(viridisLite, lib.loc="~/R/win-library/3.5")
library(viridis, lib.loc="~/R/win-library/3.5")
library(hrbrthemes)

# Now that we have loaded the packages, we will import our data:----------
# Load User's daily activity data:
DailyActs <- read.csv(file.choose(), header = TRUE, sep = ",")
summary(DailyActs)

# The below data being loaded belongs to users who tracked their daily activity and sleep on the respective dates:
Data <- read.csv(file.choose(), header = TRUE, sep = ",")
summary(Data)

# Plots for users who recorded activity and sleep together
Data$TotalTimeMinsBed <- Data$TotalMinsBed/60
Data$TotalMinsSlept <- Data$TotalMinsSlept/60

SleepRecord <-
  ggplot(Data, aes(x=factor(Id), y=TotalMinsSlept, fill=factor(Id))) +
  geom_boxplot() +
  scale_fill_viridis(discrete = TRUE, option = "C") +
  theme(legend.position = "none", plot.title = element_text(size=11)) +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=0.5)) +
  ggtitle("Sleeping Hours of Users with Daily Activity Tracking") +
  xlab("UserIDs") +
  ylab("Sleep(hrs)")

# Lets See how users are tracking their heartrates:
HeartData <- read.csv(file.choose(), header = TRUE, sep = ",")
summary(HeartData)
```

```
#Clean-up process:
HeartData<- HeartData %>%
 clean_names() %>%
 mutate(time = mdy_hms(time), weekday = weekdays(time)) %>%
 mutate(date = date(time), hour = hour(time), minute = minute(time), second = second(time))%>%
 select(-time)
HeartData<-HeartData[,c(1,4,3,5,6,7,2)]%>%
 group_by(id, date, weekday, hour, minute)%>% #remember that group_by() calculates the average of values(heart
rates) during hh:mm:00 to hh:mm:59.
 summarise(value=round(mean(value),1)) #this average of values (heart rates) during seconds is reported along hh:mm
head(HeartData,3)
# Not all users have their HR tracked.
USERS <- unique(HeartData$id)

HeartData$weekday<-factor(HeartData$weekday, levels = c("Monday","Tuesday", "Wednesday", "Thursday", "Friday",
"Saturday", "Sunday"))

#To resize graph and plot heart rate through the week

options(repr.plot.width=20, repr.plot.height = 10)

plt <- ggplot(HeartData, aes(x=HeartData$hour, y=value, color=value)) +
 geom_point(alpha=.5) +
 facet_grid(rows=vars(HeartData$weekday)) +
 scale_colour_gradient2(low="darkgreen", high="darkred", mid="yellow", midpoint=105) +
 labs(title="Heart-Rates Throughout The Week") +
 scale_x_continuous(breaks = c(4, 12, 20), label = c("Morning", "Noon", "Night")) +
 scale_y_continuous(breaks = round(seq(min(HeartData$value), max(HeartData$value), by = 45),.4)) +
 theme(text= element_text(size=20))

plt <- plt +
 # A theme with no background annotations
 theme_minimal(base_family = "Fira Sans Compressed") +
 theme(
  # Top-right position
  legend.position = c(0.975, 0.995),
  # Elements within a guide are placed one next to the other in the same row
  legend.direction = "horizontal",
  # Different guides are stacked vertically
  legend.box = "vertical",
  # No legend title
  legend.title = element_blank(),
  # Light background color
  plot.background = element_rect(fill = "#F5F4EF", color = NA),
  plot.margin = margin(20, 30, 20, 30),
  # Customize the title. Note the new font family and its larger size.
  plot.title = element_text(
   margin = margin(0, 0, 10, 0),
   size = 20,
   family = "KyivType Sans",
   face = "bold",
   vjust = 0,
   color = "grey25"
```

```
  ),
  plot.caption = element_text(size = 11),
  # Remove titles for x and y axes.
  axis.title = element_blank(),
  # Specify color for the tick labels along both axes
  axis.text = element_text(color = "grey40"),
  # Specify face and color for the text on top of each panel/facet
  strip.text = element_text(face = "bold", color = "grey20")
 )


# Now let's look at their daily activity
stepup <- ggplot(Data, aes(x=TotalSteps, y=Calories, size=factor(Id))) +
 geom_jitter(size = 0.7, alpha = 0.9) +
 geom_smooth(se = TRUE, method = lm, size = 1) +
 scale_fill_viridis(discrete = TRUE) +
 theme(legend.position = "None", plot.title = element_text(size=11)) +
 ggtitle("Steps versus Calories of all users") +
 xlab("Calories(Cal)") +
 ylab("Steps")



# Is the above observation true for each user? lets have a look at it:
Data$UserID <- factor(Data$Id)
UserStepup <- ggplot(Data, aes(x=TotalSteps, y=Calories, size=UserID, color=UserID)) +
 geom_jitter(size = 0.8, alpha = 0.9) +
 geom_smooth(se = FALSE, method = lm, size = 0.5) +
 scale_fill_viridis(discrete = TRUE) +
 theme(legend.position = "right", plot.title = element_text(size=11)) +
 ggtitle("Steps versus Calories per user") +
 xlab("Calories(Cal)") +
 ylab("Steps")

# Thus we can come to a conclusion that more steps they take, more calories they burn.



# Now let's load data containing records of Weights from different users:----------
Weight_Data <- read.csv(file.choose(), header = TRUE, sep = ",")
Weight_Data$Id <- factor(Weight_Data$Id)
# Grouped scatter plot highlighting user's preferred method of tracking weights:
T_entries <- length(Weight_Data$ManualReport[Weight_Data$ManualReport == TRUE]) #users preferring manual
entries
F_entries <- length(Weight_Data$ManualReport[Weight_Data$ManualReport == FALSE])
cent_True <- (T_entries/(T_entries+F_entries)) * 100
cent_False <- (F_entries/(T_entries+F_entries)) * 100
Cents <- data.frame(percentage = c(cent_True , cent_False), Report = c("True", "False"))

Preference <- ggplot(Cents, aes(x=Report , y=percentage, fill = Report)) +
 geom_bar(stat = "identity", width = 0.3) +
 scale_color_viridis(discrete = TRUE) +
 theme(legend.position = "none", plot.title = element_text(size=11)) +
 ggtitle("Preferred Method of Reporting Weights") +
```

```
  ylim(0,100) +
  xlab("Users Reporting Manually") +
  ylab("Percentage")

# Now let's load data containing records of sleep from different users:----------
Sleep_Data <- read.csv(file.choose(), header = TRUE, sep = ",")
Sleep_Data$TotalTimeInBed <- Sleep_Data$TotalTimeInBed/60
Sleep_Data$TotalMinutesAsleep <- Sleep_Data$TotalMinutesAsleep/60


dataSleep%>%
  ggplot(Sleep_Data, aes(x=factor(Id), y=TotalTimeInBed, fill=factor(Id))) +
  geom_boxplot() +
  scale_fill_viridis(discrete = TRUE) +
  theme(legend.position = "none", plot.title = element_text(size=11)) +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=0.5)) +
  ggtitle("Sleeping hours of Users") +
  xlab("UserIDs") +
  ylab("Sleep(hrs)")



# This confidence interval for all the users who kept a track of their daily activity
# will justify our theories
cor.test(DailyActs$TotalSteps, DailyActs$Calories, method = 'pearson', conf.level = 0.95)
```