

Bellabeat Smart Device Usage Analysis

Introduction

In this case-study I am presenting the analysis on usage data of 'Bella beat'. Bella beat is a tech-driven wellness company aiming to create products for women to help encourage healthy habits. Their smart devices track activity, sleep, stress, reproductive health, and more. To help the business grow, Bellabeat management wish to know how people are using smart devices and would like high-level recommendations to guide marketing strategy. To answer the business questions, I will follow 6 data analysis phases: Ask, Prepare, Process, Analyse, Share and Act.

This case study is a capstone project done towards completion of [Google Data Analytics Professional Certification](#) course.

In further sections, I will introduce the problem statement, analysis steps and bring forwards various data-driven insights to improve marketing and business strategies.

Problem Statement

About Company

Urška Sršen and Sando Mur founded Bellabeat, a tech company that manufactures health-focused smart products focusing on Women health. Sršen used her background as an artist to develop beautifully designed technology that informs and inspires women around the world. Collecting data on activity, sleep, stress, and reproductive health has allowed Bellabeat to empower women with knowledge about their own health and habits. Since it was founded in 2013, Bellabeat has grown rapidly and quickly.

By 2016, Bellabeat had opened offices around the world and launched multiple products. Bellabeat products became available through a growing number of online retailers in addition to their own e-commerce channel on their website. The company has invested in traditional advertising media, such as radio, out-of-home billboards, print, and television, but focuses on digital marketing extensively. Bellabeat invests year-round in Google Search, maintaining active Facebook and Instagram pages, and consistently engages consumers on Twitter. Additionally, Bellabeat runs video ads on YouTube and display ads on the Google Display

Network to support campaigns around key marketing dates. Sršen knows that an analysis of Bellabeat available consumer data would reveal more opportunities for growth. She has asked the marketing analytics team to focus on a Bellabeat product and analyse smart device usage data in order to gain insight into how people are already using the smart devices. Then, using this information, she would like high-level recommendations for how these trends can improve Bellabeat marketing strategy.

Some High-level 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?

- **Project Objective:** Analyse smart device data to gain insights into how consumers use smart devices and provide recommendations for marketing strategy.

Case Study Roadmap: ASK

Identifying Key Business Task

- The director of marketing believes the company's future success depends on maximizing the marketing strategy.
- I am analysing smart device data to understand how customer is using Bellabeat smart device. Collecting data on activity, sleep, Steps, Calories Burn, Heart rate and weight observation health has allowed Bellabeat to empower women with knowledge about their own health and habits.

Key Stakeholders

- Urška Sršen: Bellabeat cofounder and Chief Creative Officer.
- Sando Mur: Mathematician and Bella beat cofounder; key member of the Bellabeat executive team.
- Bella beat marketing analytics team: A team of data analysts responsible for collecting, analysing, and reporting data that helps guide Bellabeat marketing strategy.

Case Study Roadmap: PREPARE

Download data and store it appropriately.

<https://www.kaggle.com/arashnic/fitbit>

The dataset was made available in Coursera's Google Analytics course. This is a publicly available dataset shared through [Kaggle](#).

Data Organization

Dataset has data for 2 months, April-2016 to May-2016 for analysis. Data for both months are available in Comma Separated values(.csv) file. These data are download and stored in local PC for further processing and analysis. Data is shared through 19 different .csv files. These files were downloaded and saved in local directory on PC.

Table below describes all .csv files.

Table-1 Details of Files

Table Name	Description
dailyActivity_merged	Daily Activity over 31 days of 33 users. Tracking daily: Steps, Distance, Intensities, Calories
dailyCalories_merged	Daily Calories over 31 days of 33 users
dailyIntensities_merged	Daily Intensity over 31 days of 33 users. Measured in Minutes and Distance, dividing groups in 4 categories: Sedentary, Lightly Active, Fairly Active, Very Active
dailySteps_merged	Daily Steps over 31 days of 33 users
heartrate_seconds_merged	Exact day and time heartrate logs for just 7 users
hourlyCalories_merged	Hourly Calories burned over 31 days of 33 users
hourlyIntensities_merged	Hourly total and average intensity over 31 days of 33 users
hourlySteps_merged	Hourly Steps over 31 days of 33 users
minuteCaloriesNarrow_merged	Calories burned every minute over 31 days of 33 users (Every minute in single row)
minuteCaloriesWide_merged	Calories burned every minute over 31 days of 33 users (Every minute in single column)
minuteIntensitiesNarrow_merged	Intensity counted by minute over 31 days of 33 users (Every minute in single row)
minuteIntensitiesWide_merged	Intensity counted by minute over 31 days of 33 users (Every minute in single column)
minuteMETsNarrow_merged	Ratio of the energy you are using in a physical activity compared to the energy you would use at rest. Counted in minutes
minuteSleep_merged	Log Sleep by Minute for 24 users over 31 days. Value column not specified
minuteStepsNarrow_merged	Steps tracked every minute over 31 days of 33 users (Every minute in single row)

minuteStepsWide_merged	Steps tracked every minute over 31 days of 33 users (Every minute in single column)
sleepDay_merged	Daily sleep logs, tracked by: Total count of sleeps a day, Total minutes, Total Time in Bed
weightLogInfo_merged	Weight track by day in Kg and Pounds over 30 days. Calculation of BMI.5 users report weight manually 3 users not. In total there are 8 users

Determine the credibility the data.

Data is made public by Bellabeat for public use. These data are from actual usage of smart device and made available by company, hence we can assume that the data is credible.

Case Study Roadmap: PROCESS

Check the data error:

The date and time were not properly formatted which was causing error during importing the data in SQL. Date was formatted in 24-hour DD-MM-YYYY HH:MM format by loading and transforming the data in power query tool.

Tools Used:

I have used:

1. power query – for formatting date and time column.
2. SQL – for data extraction.
3. Tableau – for visualization and dashboarding.

Transform the data:

As explained above, actual .csv file had issue with data-time format. Hence, the files was loaded in power query, transformed and saved as .csv file. It was then imported into PostgreSQL.

Document Cleaning Process:

All csv files were imported into PostgreSQL. SQL was used to create queries to extract meaningful and required data based on analysis requirements.

Case Study Roadmap: ANALYSE

We will use the daily activity table. I first run query to check the number of rows and columns. As shown in Query-1, there are 940 rows and 15 columns.

Rows are record of each user for every day for 31 days.

Available columns: ID, activity date, total distance, tracker distance, very active distance, logged activities distance, moderately active distance, light active distance, secondary active distance, very active minute, fairly active minute, lightly active minute, secondary active minute, calories.

Query: 1: Analysis of Daily Activity

```
select *  
from daily_activity  
output (There are 940 rows and 15 columns)
```

Using Distinct, we were able to get that we have data for 33 unique users for which data is collected.

Query: 2: Analysis of Daily Activity For users

```
Select distinct(id)  
from daily activity  
output (There are 33 users )
```

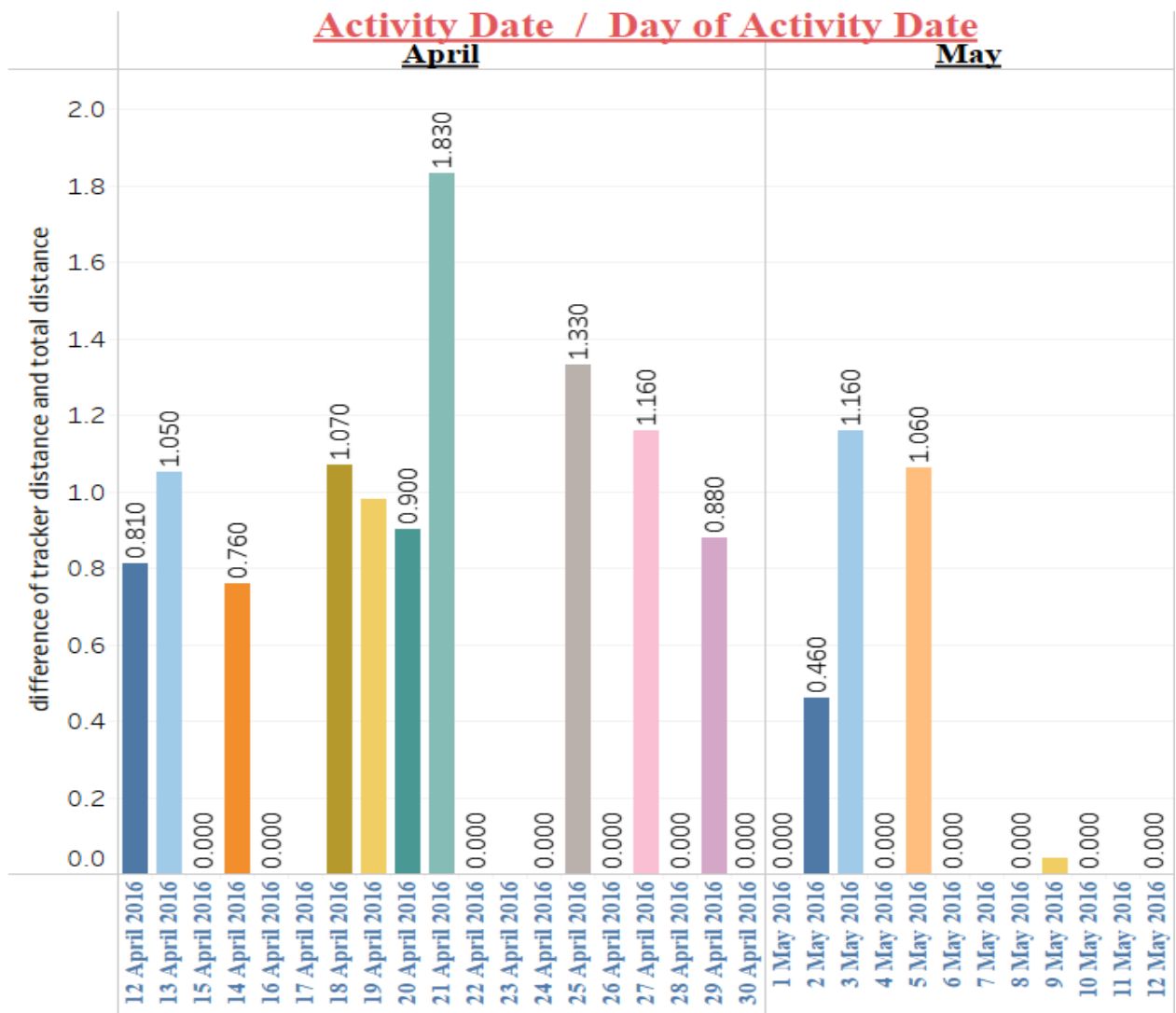
Identify trends and relationships.

1. Tracker distance vs Total Distance

Query: 3: To find difference between total distance vs tracker distance

```
select activity_date, round(sum(abs(total_distance- tracker_distance)),2) as distance  
from daily_activity  
group by activity_date  
order by activity_date
```

Difference of Total and Tracker Distance

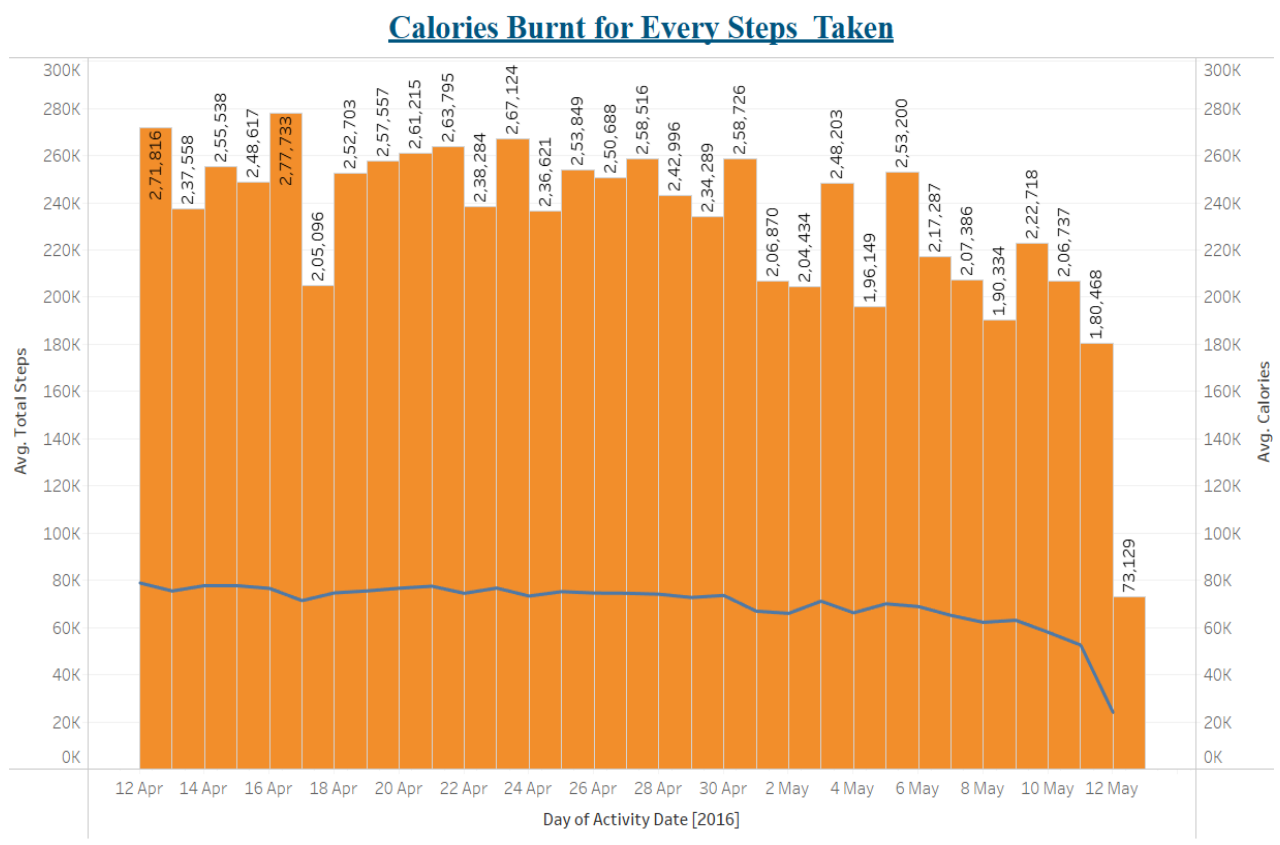


Summary of analysis:

Dataset has separate columns for total distance and tracker distance. It was not very clear how both distances were measured. I made an intuitive assumption that tracker distance is distance travelled with tracking enabled and in most cases they should be same. In such a case tracker distance should always be less than or equal to total distance. However, the result does not match the assumption. On Analysis, this was not the case. Out of 31 days, 13 days we can see the variation in reading which means in around 42% cases we noticed deviation in tracker and total distance. Also, maximum deviation is around 1.8Km and minimum deviation is 400m.

Query: 4: To analyse daily activity for calories burnt for every steps taken

```
select activity_date, sum(total_steps)as total_steps,sum(calories)as calories
from daily_activity
group by activity_date
order by activity_date
```



Summary of analysis:

Calories burned for every step taken

From the column cluster chart plot, we discovered that:
It is a positive correlation.

We observed that intensity of calories burned increase when users are at the range of 2,71,816 steps with calories burn rate cooling down from 78,893 steps onwards.

Noted a few outliers:

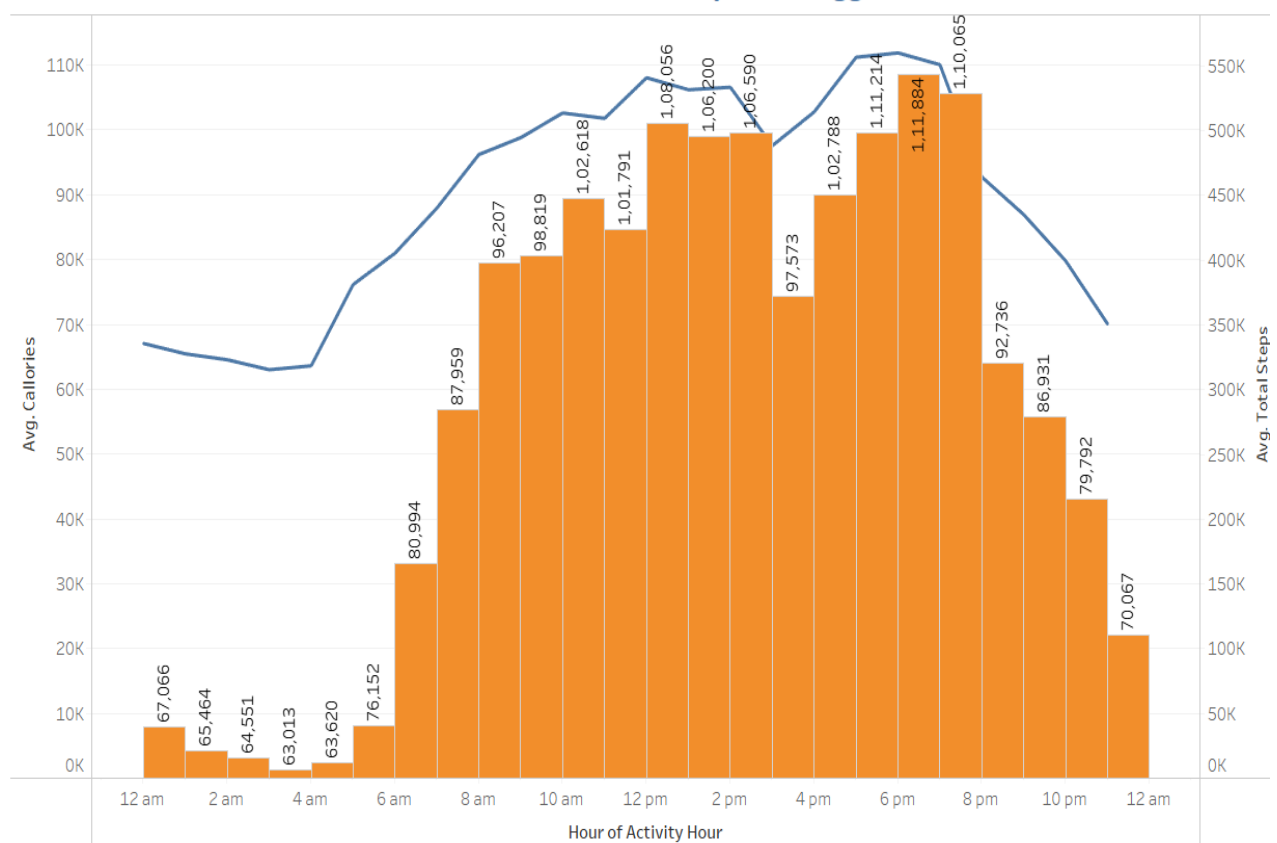
Zero steps with zero to minimal calories burned.

1 observation of > 73,129 steps with < 29,925 calories burned.

Query: 5: Analysis of hourly Calories Burnt

```
select activity_hour, sum(total_step)as total_steps, sum(total_intensity)as total_intensity, sum(calories)as callories
from hourly_activity
group by activity_hour
order by activity_hour
```

Calories Burnt for Every hour logged



Summary of analysis:

We will check the users hourly calories burn.

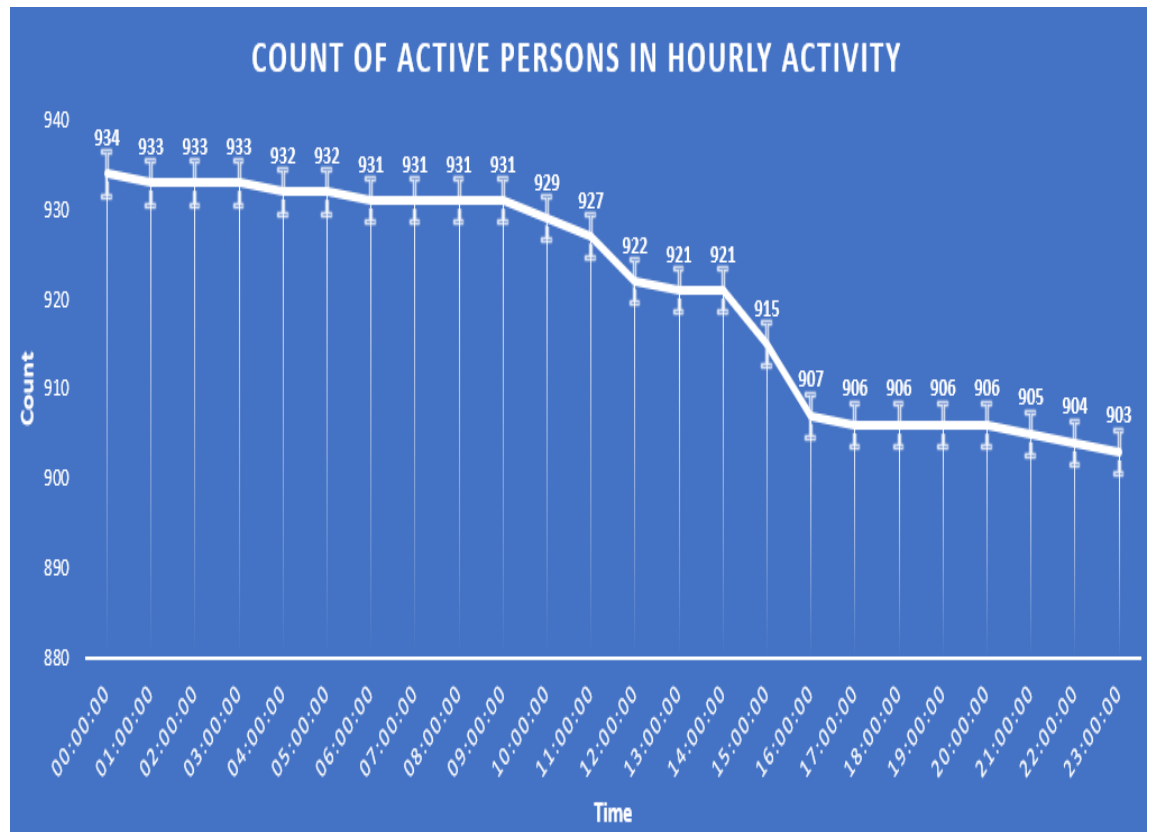
on 12 am to 7 am highly increase calories burnt.

on 8 am to 11 pm average calories burnt.

Smart device tracker worked in day and night properly.

Query: 6: Count of Active person in hourly activity

```
select count(id), activity_hour
from hourly_activity
group by activity_hour
order by activity_hour
```



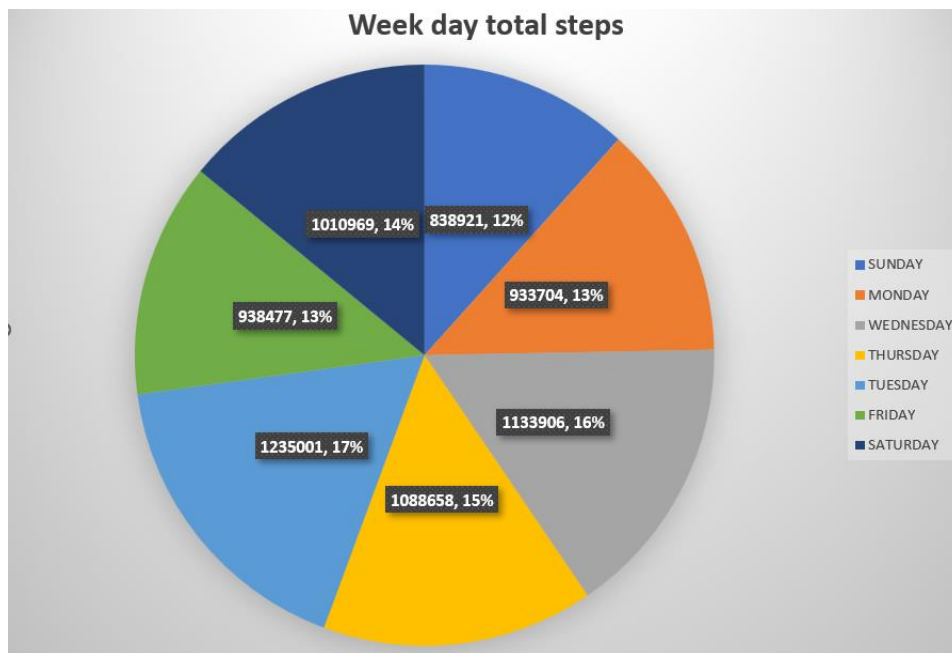
Summary of analysis:

In this analysis it is also found that using smart device of Bellabeat at night will more as compared to day time.

Highly increased user uses this device 12 am to 12 pm

Query: 7: Weekly analysis of total steps and calories Burnt

```
select sum(total_steps), To_CHAR(activity_date, 'DAY')as "day"
from daily_activity
group by day
```

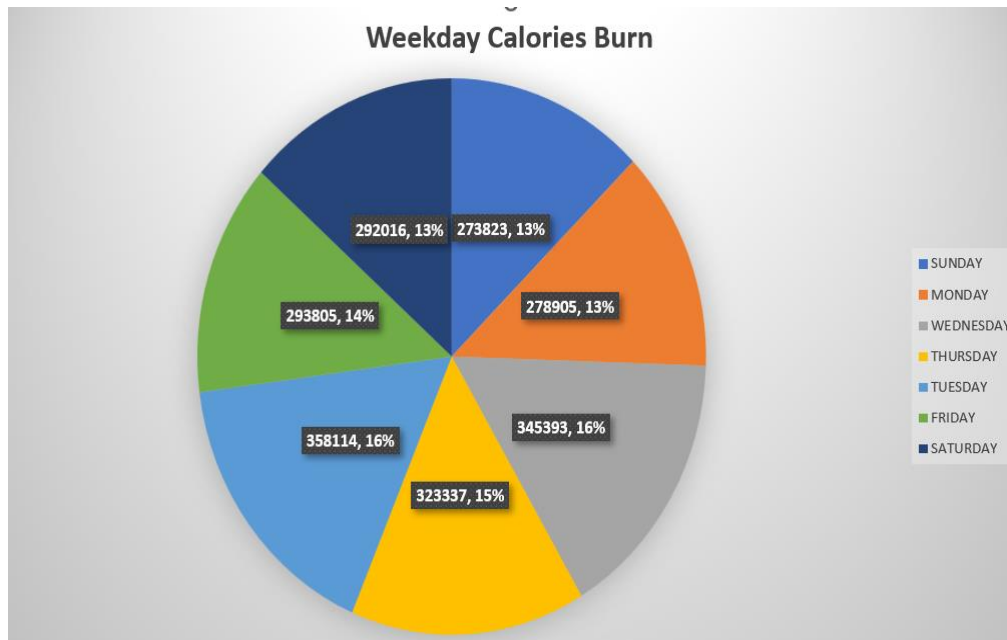


Summary of analysis:

I am analysing total steps for week day users are using this device frequently. Majority of customer are using the device between Monday to Friday.

Query : 8

```
select sum(calories)as calories, To_CHAR(activity_date, 'DAY')as"day"
from daily_activity
group by day
```



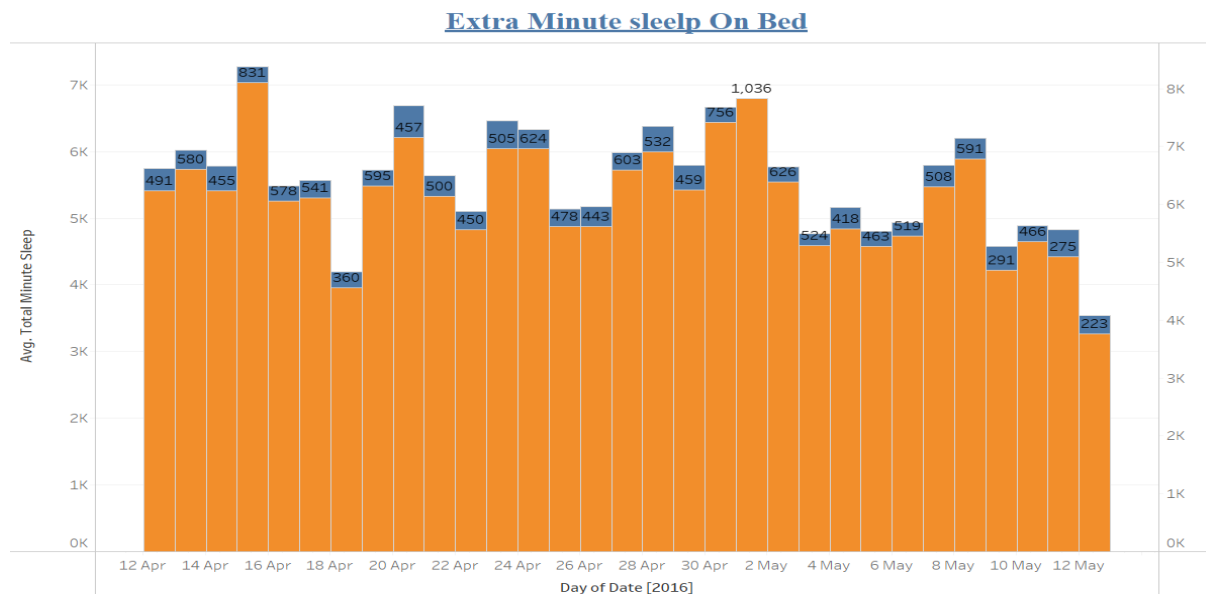
Summary of analysis:

I am analysing total calories burnt for week days users are using this device frequently. Average customer using Fitbit device Monday to Friday more.

Daily Analysis of Minute Activity Total extra time in bed

Query : 9

```
select To_CHAR(sleep_day, 'day')as day, sum(total_time_in_bed- total_minute_sleep)as extra_time_in_bed
from sleep_day
group by day
```



Summary of analysis:

Analysing of Relation between Total time in bed and total minute sleep

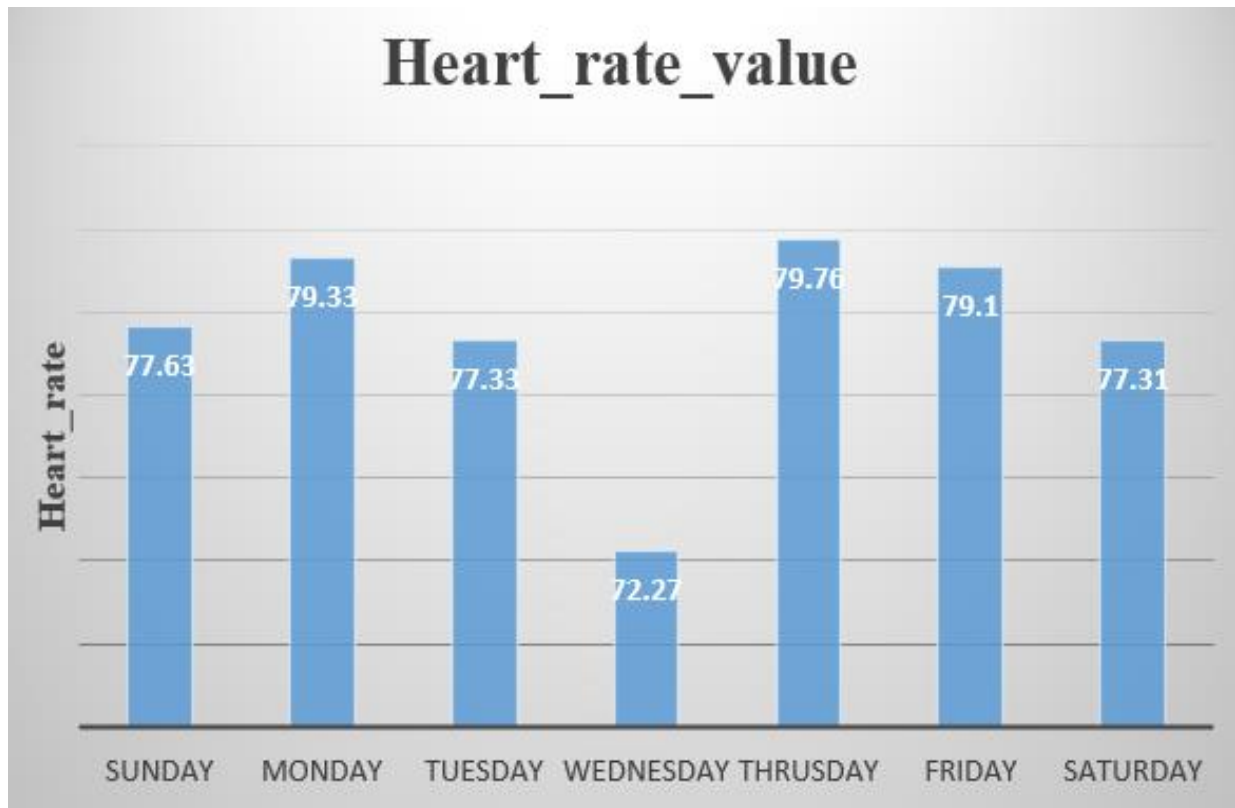
From this graph, we can clearly see that there is a positive trend correlating the total time users spent sleeping and the total time users spent being in bed. However, we can also see that there are several cases where users are spending considerably more time simply lying awake in bed rather than sleeping. This may indicate that users had found some trouble in falling asleep.

Weakly Analysis of Heart Rate

Query : 10

```
select round(Avg(value),2), TO_CHAR(date, 'day') AS "day"
from heart_rate
group by day
```

Day	Heart_rate_value
Sunday	77.63
Monday	79.33
Tuesday	77.33
Wednesday	72.27
Thursday	79.76
Friday	79.1
Saturday	77.31



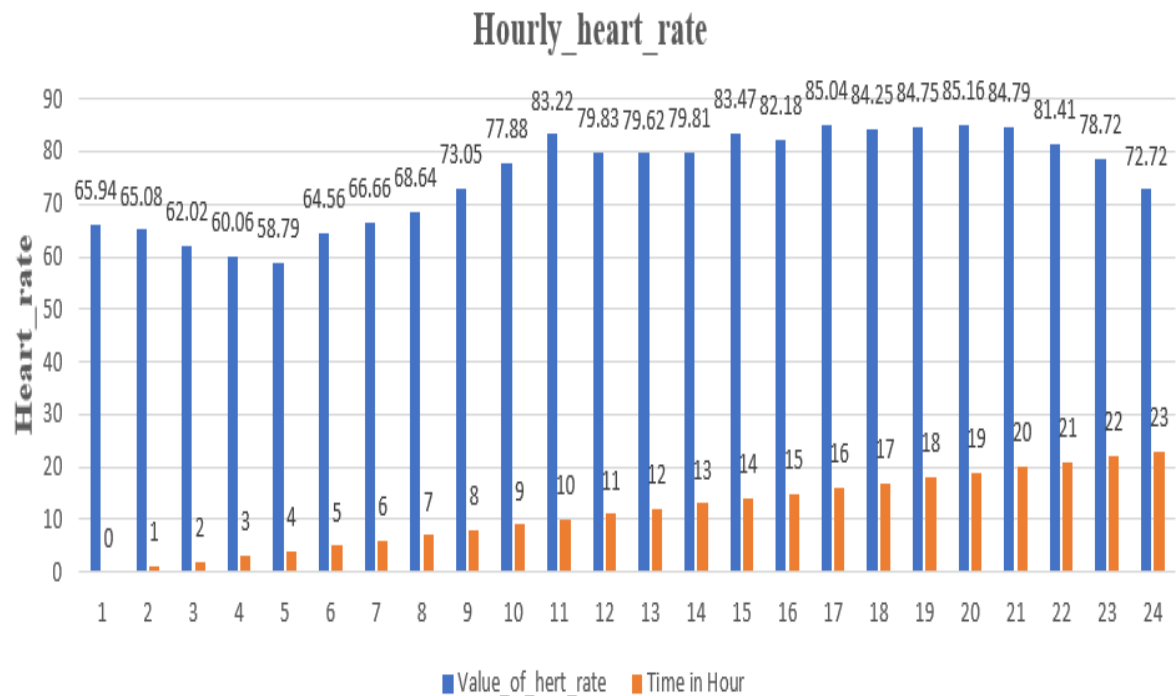
Summary of analysis:

This visualization below was done in Tableau by setting the row and column to heart_rate and week day respectively.

Analysis of hourly Heartrate

Query : 11

```
select count(id), round(avg(value),2), TO_CHAR(time, 'HH24') AS "hour"
from heart_rate
group by hour
order by hour
```



Summary of analysis:

This visualization below was done in Tableau by setting the row and column to very active hours and heart_rate respectively.

From this observation, there is no correlation between the heart_rate and very active hours. However, this might be due to the small sample of heart_rate data available.

Analysis of Daily Weight Observation

Query : 12

```
select date, weight_kg, weight_pound, TO_CHAR(date, 'day') AS "day", TO_CHAR(date, 'HH24') AS "hour", TO_CHAR(date, 'DD-MM-YYYY')  
AS "Date"  
from total_weight
```



Summary of analysis:

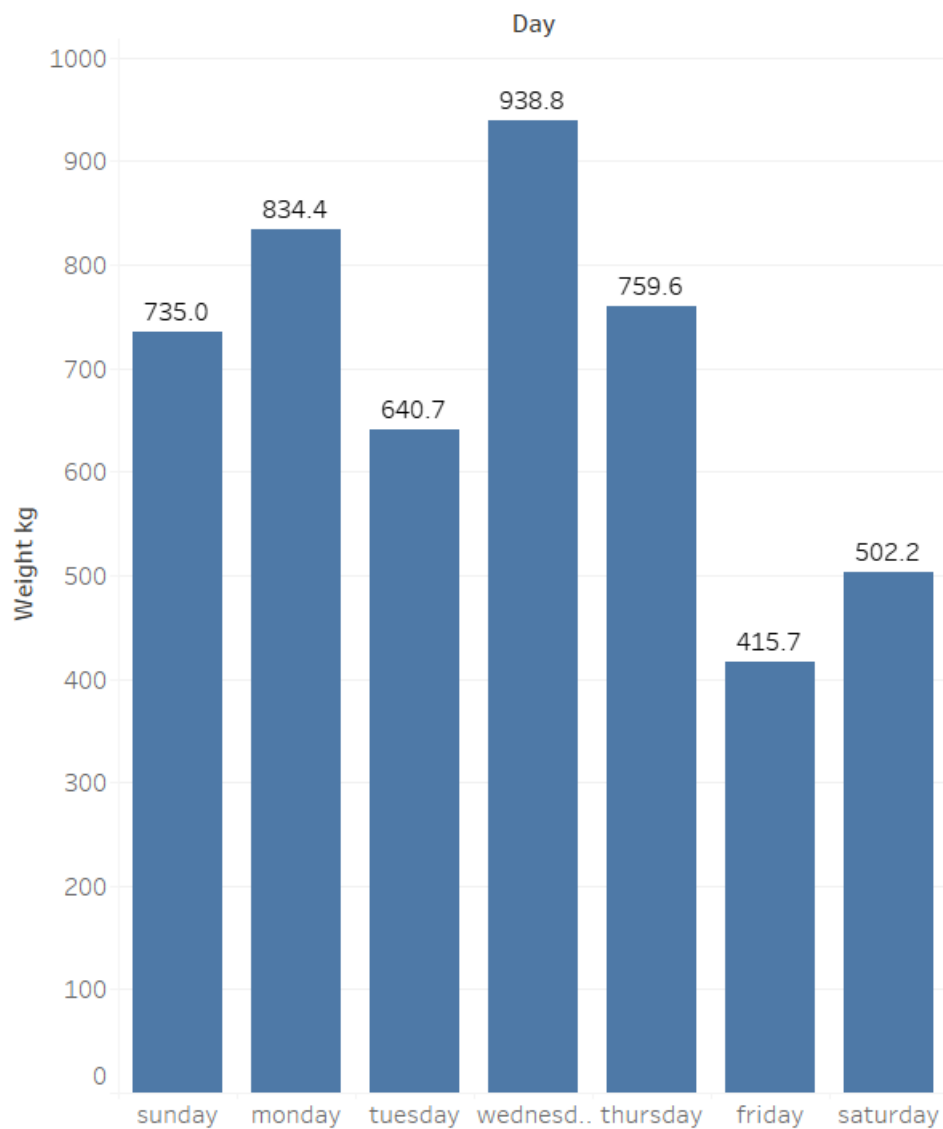
This visualization below was done in Tableau by setting the row and column to daily activity and weight(lbs) respectively.

Analysis of Weekly Weight Observation

Query : 13

```
select date, sum(weight_kg), sum(weight_pound), TO_CHAR(date, 'day') AS "day"  
from total_weight  
group by day
```

Week_Day_Weight_Observation



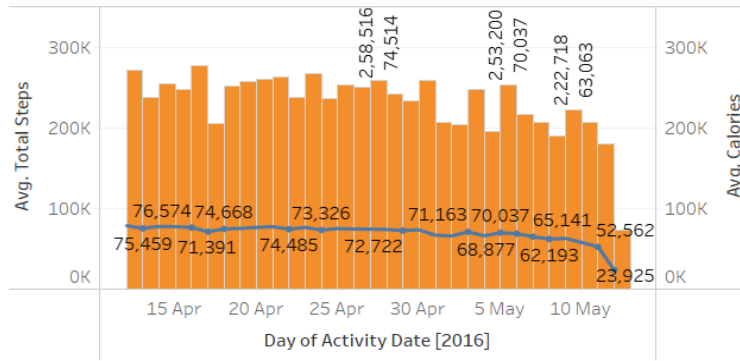
Summary of analysis:

This visualization below was done in Tableau by setting the row and column to very active in a day and weight(lbs) respectively. From this observation, there is no correlation between the weight and very active day. However, this might be due to the small sample of weight data available. Therefore, this is not enough data to conclude on this relationship.

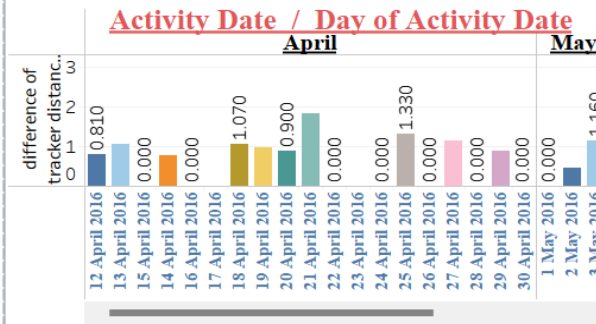
Case Study Roadmap: SHARE

Bellabeat Smart Device Dashboard

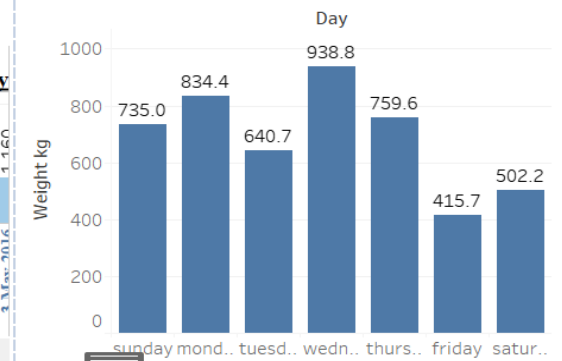
Calories Burnt for Every Steps Taken



Difference of Total and Tracker Distance

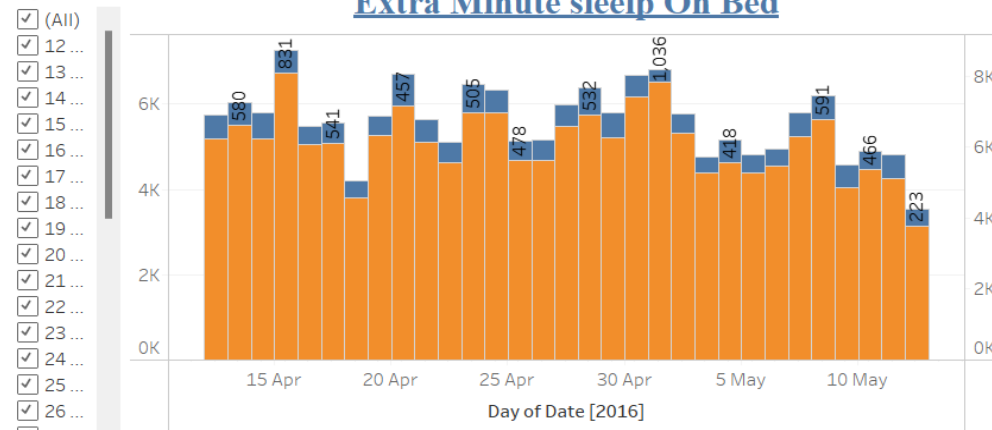


Week_Day_Weight_Observation

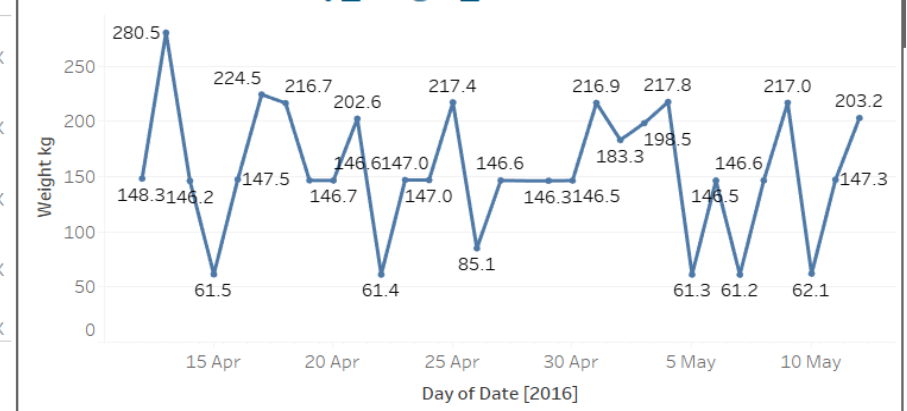


Day of Date

Extra Minute sleep On Bed



Daily_weight_observation



Case Study Roadmap: ACT

- Each user takes 8,515 steps per day on average, which can be considered to moderately active.
- The average sleep time is about 7 hrs per day.
- Having more step could be one of the primary reason for the loss in calories.
- Sleep Time doesn't have a strong effect on how active users are during the day.
- Users might be able to have more sleep if they're more active during the day.
- Users are most active between 12:00 am - 2:00 pm and 5:00 pm - 7:00 pm.

Recommendations

- Users prefer to track their activities during weekdays as compared to weekends - perhaps because they spend more time outside on weekdays and stay in on weekends.
- Develop products focused on providing women with their health, habit and fitness data and encouraging them to understand their current habits and make healthy decisions.
- Bellabeat marketing team can encourage users by educating and equipping them with knowledge about fitness benefits, suggest different types of exercise (i.e. simple 10 minutes exercise on weekday and a more intense exercise on weekends) and calories intake and burnt rate information on the Bellabeat app.
- On weekends, Bellabeat app can also prompt notification to encourage users to exercise.