



Analysis Case Study: Bellabeat

Paulina Remis

Last Updated: September 2023

Table of Contents

- ▶ Introduction
- ▶ Project Overview
- ▶ Data and Analysis
- ▶ Visualizations
- ▶ Key Insights and Conclusions
- ▶ Recommendations and Actionable Steps



What is Bellabeat?



Bellabeat

Setting

- Founded in 2013
- Counted with several offices worldwide by 2016
- Their goal is to become a large player in the smart device market.

How it works

- Data collection on sleep, activity, stress, and reproductive health.
- Provide knowledge about users' health and habits

Types of Products

- **Bellabeat App** → provides users with their health data to help them understand their health and habits
- **Leaf** → wellness tracker; can be worn as bracelet, necklace, or clip; track activity, sleep, and stress
- **Time** → wellness watch; track activity, sleep, and stress
- **Spring** → water bottle that tracks daily water intake
- **Bellabeat Membership** → gives users 24/7 access to fully personalized guidance based on their lifestyle and goals

Where did the data
come from?



Database

Availability

- Available through [this link](#)
- It's a public domain dataset available through [Mobius](#) in Kaggle

Range

- April 12, 2016 - May 12, 2016

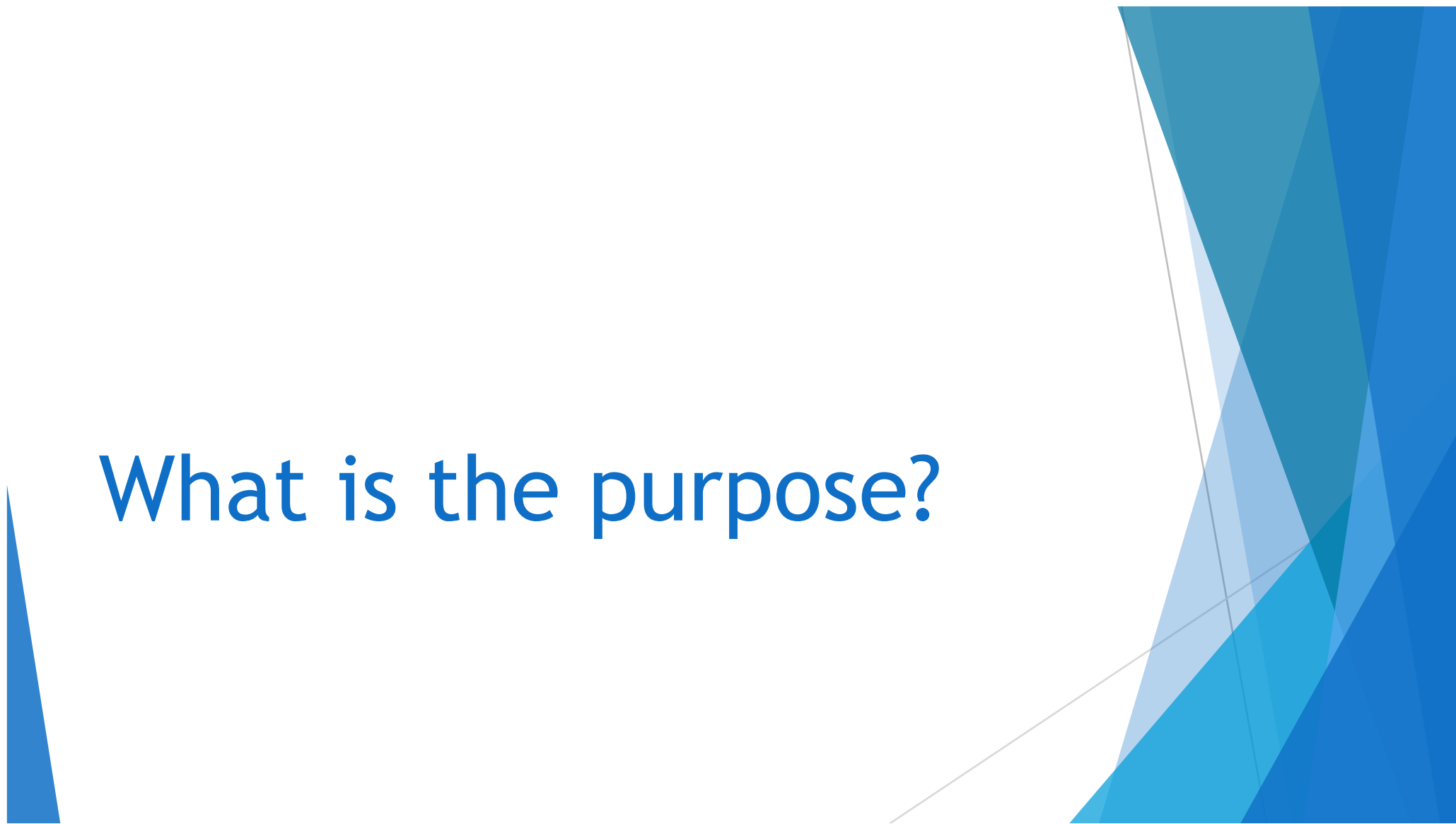
Summary

- 18 different .CSV files for different activities to explore users' habits.
- 33 FitBit Users

Limitations

- The actual time range is inconsistent with the one provided in the summary
- The actual number of users varies with the one provided in the summary
- There is no way to determine how many of the users in the sample are women

What is the purpose?



Overview

Business Task

- ▶ Analyze smart device usage data in order to gain insight into how consumers use non-Bellabeat smart devices and how these patterns can provide valuable insights for shaping Bellabeat's 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
- ▶ Bellabeat marketing analytics team



Data and Analysis



Database Selection

- ▶ All 18 .CSV files were reviewed
- ▶ Due to redundancy in the information within the files, only 5 .CSV files were used for the analysis.
- ▶ Databases selected:
 - ▶ dailyActivity_merged_clean.csv
 - ▶ dailyCalories_merged_clean.csv
 - ▶ sleepDay_merged_clean.csv
 - ▶ dailySteps_merged_clean.csv
 - ▶ weightLogInfo_merged_clean.csv

Number of Unique IDs

- ▶ 3 of the databases contain tracking data from 33 unique IDs
- ▶ One database contains tracking data from 24 unique IDs
- ▶ One database contains data from only 8 unique IDs
- ▶ All of the IDs found in the Daily Sleep and Weight Log databases are also found in the other 3 databases

Distinct ID Values

Daily Steps	33.00
Daily Calories	33.00
Daily Activity	33.00
Daily Sleep	24.00
Weight Log	8.00

Use of Trackers

- ▶ A majority of the users has fairly consistent tracking habits
- ▶ Only 1 user logged data for less than half the month
- ▶ Approximately 64% of users logged data every single day
- ▶ About 88% of users logged data for more than $\frac{3}{4}$ of the days in the month

number_of_logged_data	number_of_users
31	21
30	3
29	2
28	1
26	2
20	1
19	1
18	1
4	1

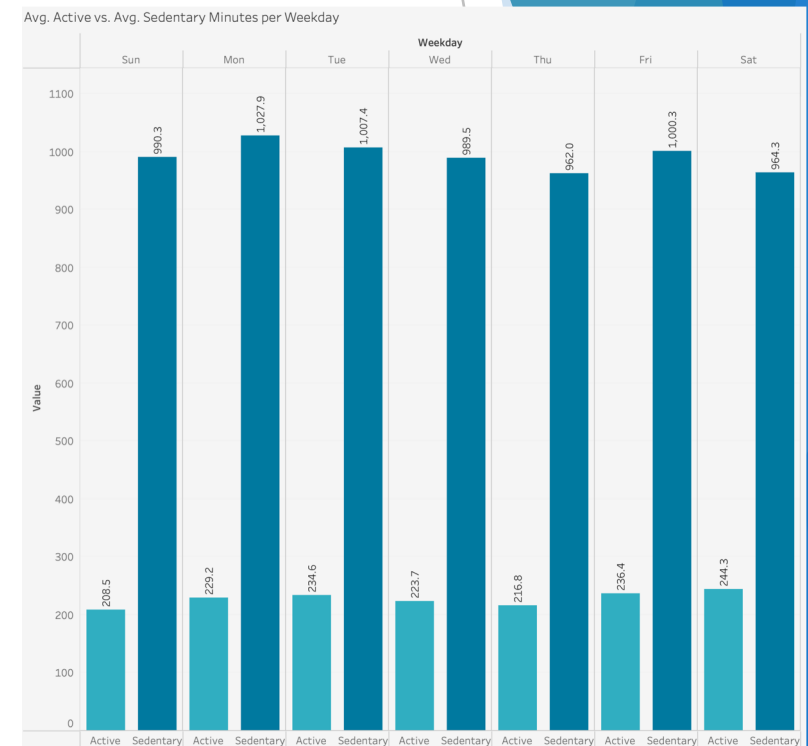
Type of User

- ▶ Categorized users based on the number of days they used the tracker
 - ▶ Heavy User: 29 to 31 days
 - ▶ Moderate User: 22 to 28 days
 - ▶ Light User: 16 to 21 days
 - ▶ Sporadic User: 0 to 15 days
- ▶ Most users are either heavy users or moderate users
- ▶ Users display tracking consistency

type_of_user	user_quantity
Heavy User	26
Moderate User	3
Light User	3
Sporadic User	1

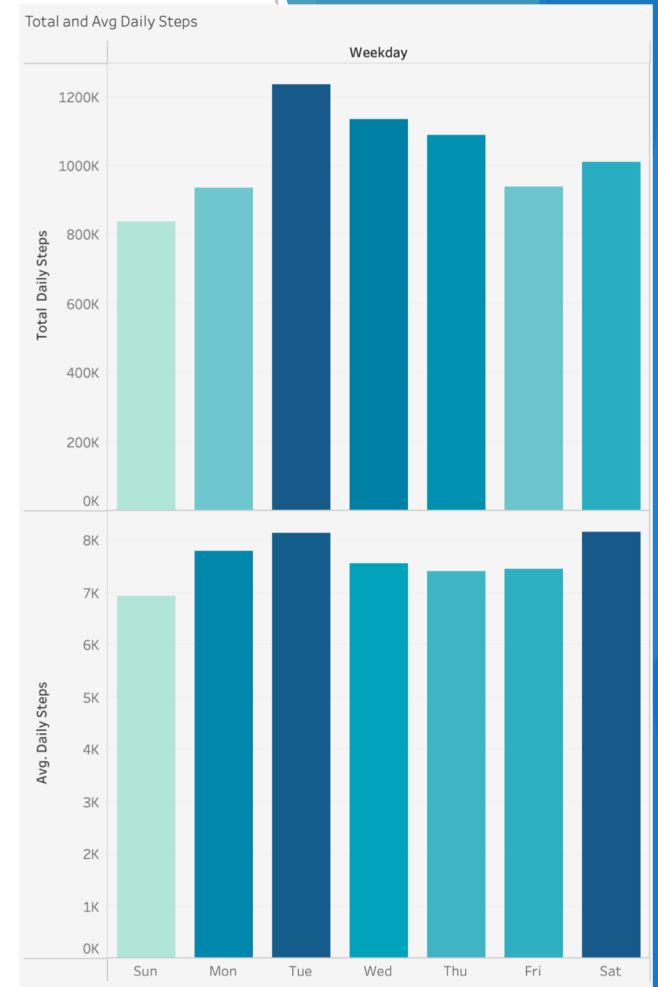
Active vs. Sedentary Minutes per Weekday

- ▶ Users spend a significantly higher amount of time being sedentary than active
- ▶ In general, when the value of sedentary minutes is higher, the number of active minutes also tends to rise
- ▶ The increase in both of these values may be indicative that users track during more extended periods of time during these days
- ▶ The days with more tracking activity are Tuesday and Wednesday



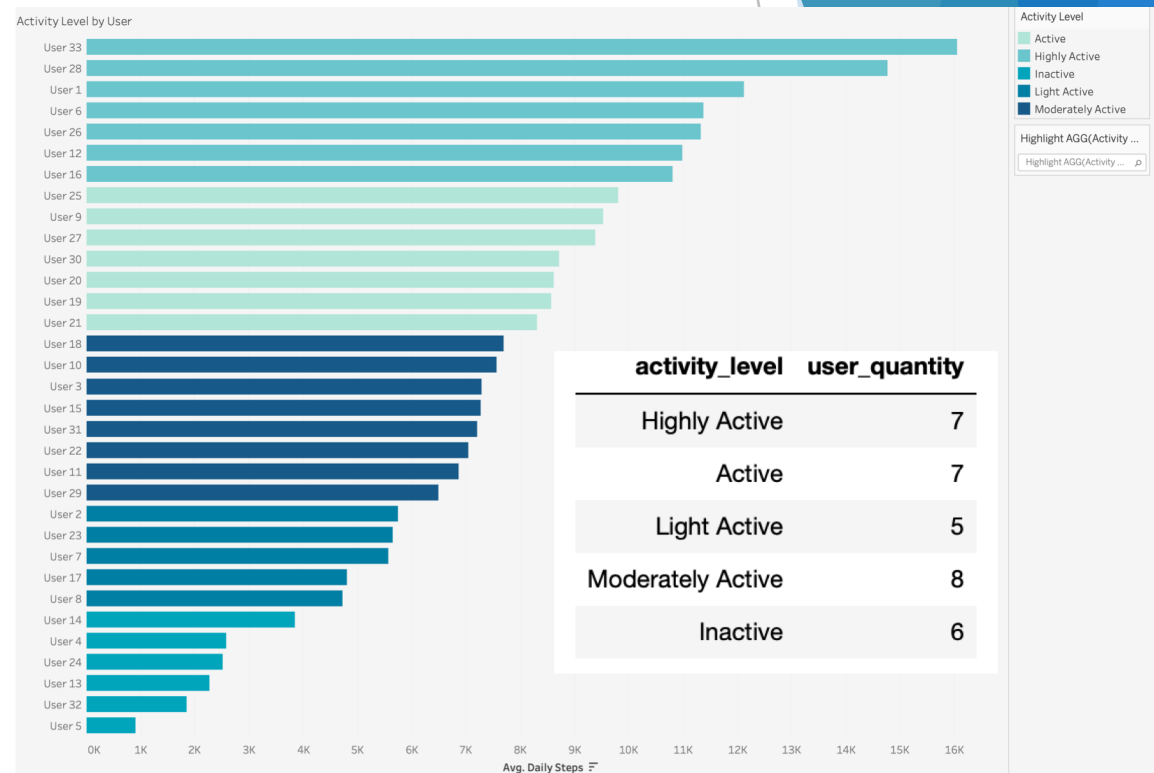
Total and Average Daily Steps

- ▶ When looking at the total the peak is on Tuesday, with all the other days having considerably lower values
- ▶ When looking at the averages, the peak is on Saturdays, closely followed by Tuesdays, and there's not as much variation between the values
- ▶ This is indicative that the records for Saturday have generally higher values but there may not be as many records on Saturday as there are on Tuesdays.



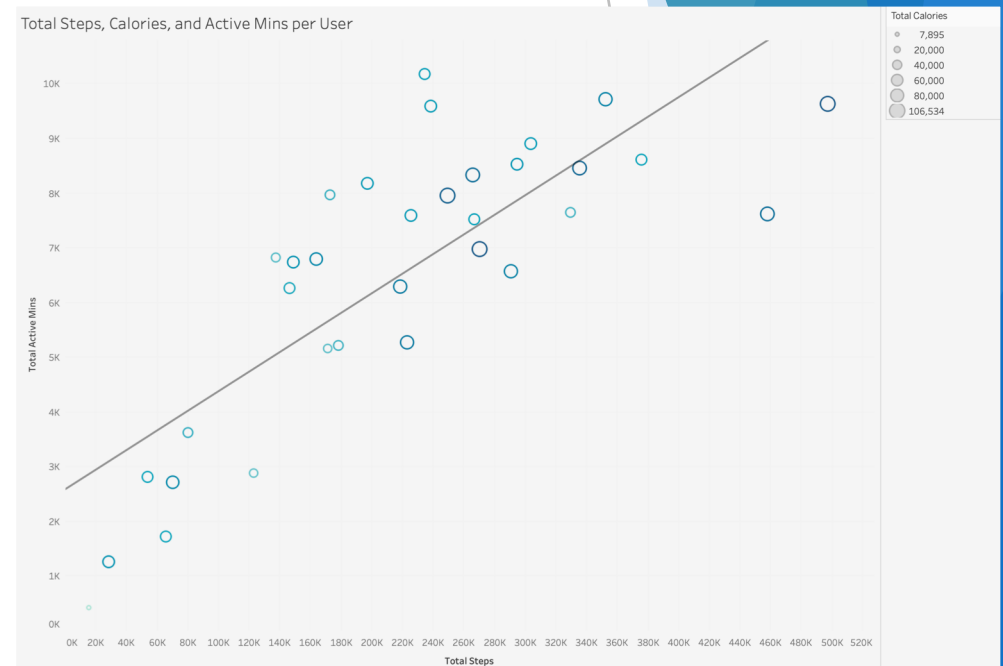
Activity Level per User

- ▶ Each user was assigned to one of five different activity levels based on their average steps:
 - ▶ Inactive -> average between 0 and 4,500 daily steps
 - ▶ Lightly Active -> average between 4,501 and 6,000 daily steps
 - ▶ Moderately Active -> average between 6,001 and 8,000 daily steps
 - ▶ Active -> average between 8,001 and 10,000 daily steps
 - ▶ Highly Active -> average over 10,000 daily steps
- ▶ There's an even distribution between all types of users



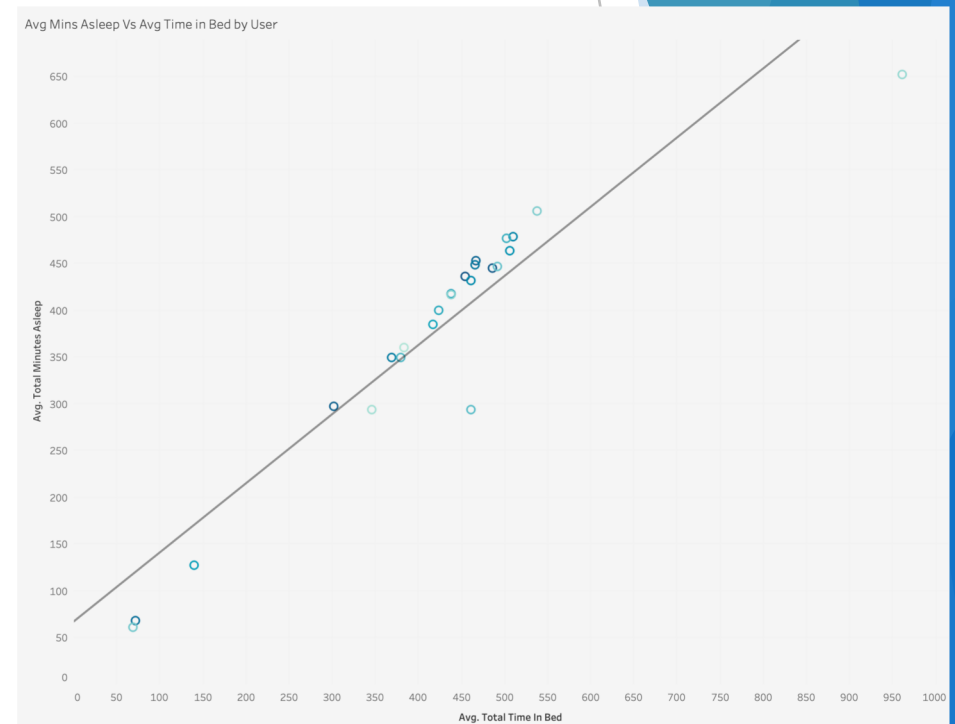
Scatter Plot Total Steps, Calories, and Active Mins per User

- ▶ Positive correlation between total steps and active minutes
- ▶ No clear pattern or trend observed between total steps and active minutes with the total calories burned



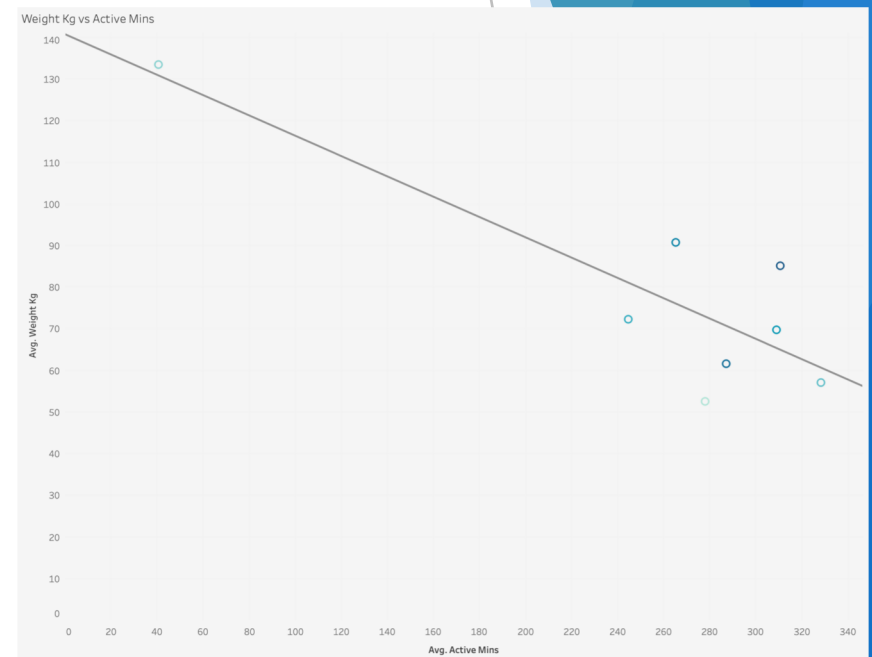
Scatter Plot Avg. Time in Bed and Avg. Mins Asleep

- ▶ Very strong correlation between average time in bed and average minutes asleep
- ▶ R-Squared of 0.884
- ▶ The trend shows that as the total time in bed increases, the minutes asleep also tends to rise.



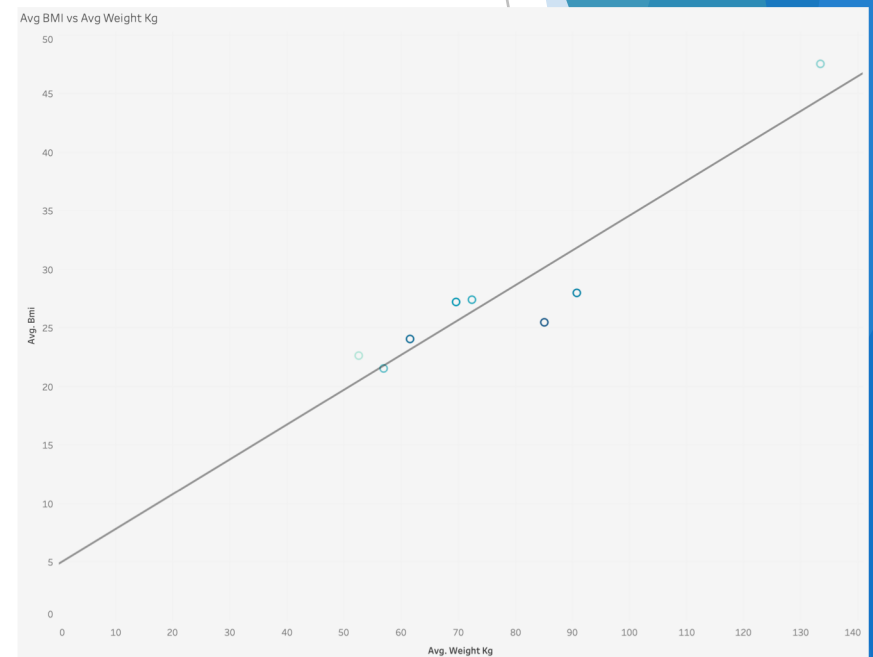
Scatter Plot Weight in Kg vs Active Mins

- ▶ Negative correlation between average active minutes and average weight in kg
- ▶ R-Squared 0.74
- ▶ As the average active minutes increases, the average weight tends to decrease



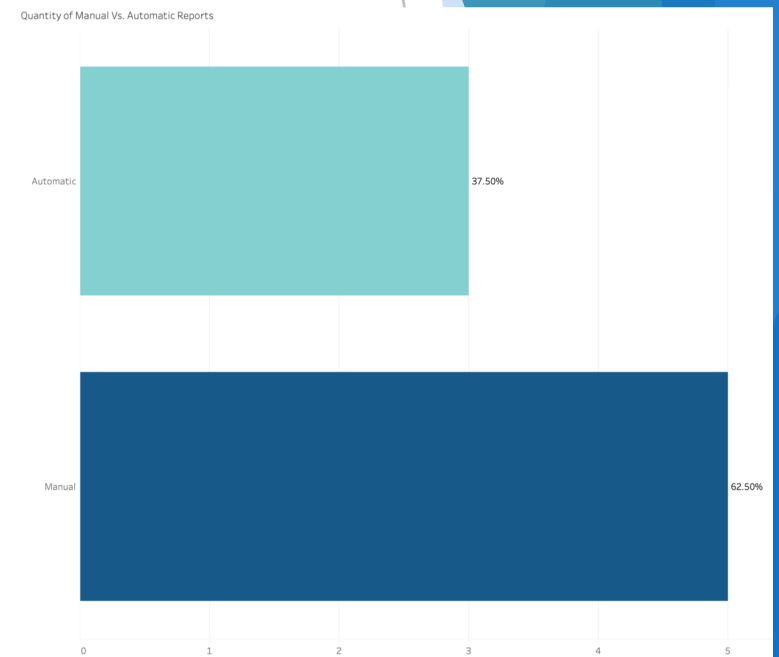
Scatter Plot Average BMI and Average Weight in Kg

- ▶ Strong positive correlation between average BMI and average weight in kg.
- ▶ R-Squared 0.88
- ▶ As the average weight trends upwards, so does the average BMI



Type of Reports Generated for Weight DB

- ▶ Most of the users report weight data manually
- ▶ The added effort to add weight reports may account for the decreased sample size of the weight database. Further analysis necessary to confirm



What did we learn?



Key Insights & Conclusions

- ▶ Users are more likely to track their activities during weekdays
- ▶ A lot of the activity being tracked corresponds to sedentary data
- ▶ While all users used the FitBit pretty consistently to track daily steps, calories, and activities, only about 73% of users tracked their sleep data, and only around 24% of users tracked their weight data.
- ▶ The majority of the users only log in weight data manually
- ▶ The average daily steps per weekday was around or below 8,000 steps, which is below the recommended 10,000 daily steps.
- ▶ Only 21% of users belong to the “Highly Active” group, which means that, on average, the other 79% of users don’t reach the recommended 10,000 steps.
- ▶ 57% of users don’t even reach an average of 8,000 daily steps.
- ▶ Most of the time spent in bed is used to sleep
- ▶ There is insufficient demographic data to draw accurate conclusions.

What can I do with this information?



Actionable Steps

- ▶ Bellabeat can include a notification function in the App, Leaf, and Time devices to notify users when they have been sedentary for too long to encourage users to reach the recommended daily steps
- ▶ Include a reward based system for users to encourage them to use the devices on a daily basis. For example, if a user tracks data for one month in a row, they may get a free trial of the Bellabeat membership for a week, which will also expose users to the benefits of becoming a member.
- ▶ Encourage users to log in sleep data by providing detailed insights into their sleeping habits and educating them over the overall impact that it has over their health and habits.

Recommendations

- ▶ Further analysis to determine why tracking weight data is not popular amongst users (survey with questions about functionality, use difficulty, etc)
- ▶ Further analysis to determine why sleep wasn't tracked on a daily basis (survey with questions about why the device is not used at night, functionality, battery durability, etc)

