

Data Size by  
number of  
rows

11M +



Kaggle

Data Source Used

Relational

Database Type Format

# Overview of Key Analytics & Insights

July 25, 2025 - Data Analytics Case Study BellaBeats

**Author** - Ruben D. Velasco

**Stakeholders:** Cofounder Sršen, Executive Mur, &  
Bellabeats Marketing Analytics Team

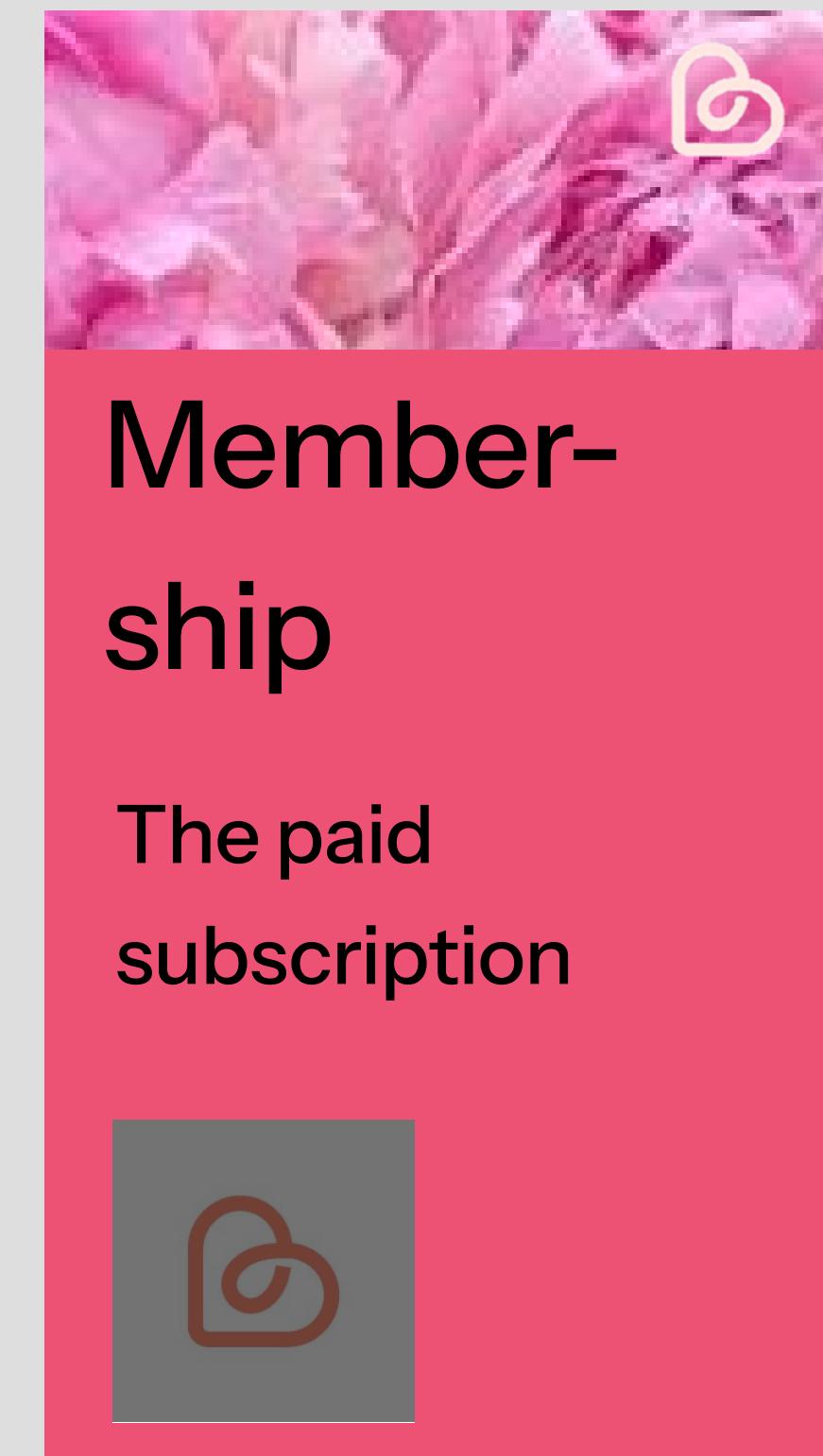
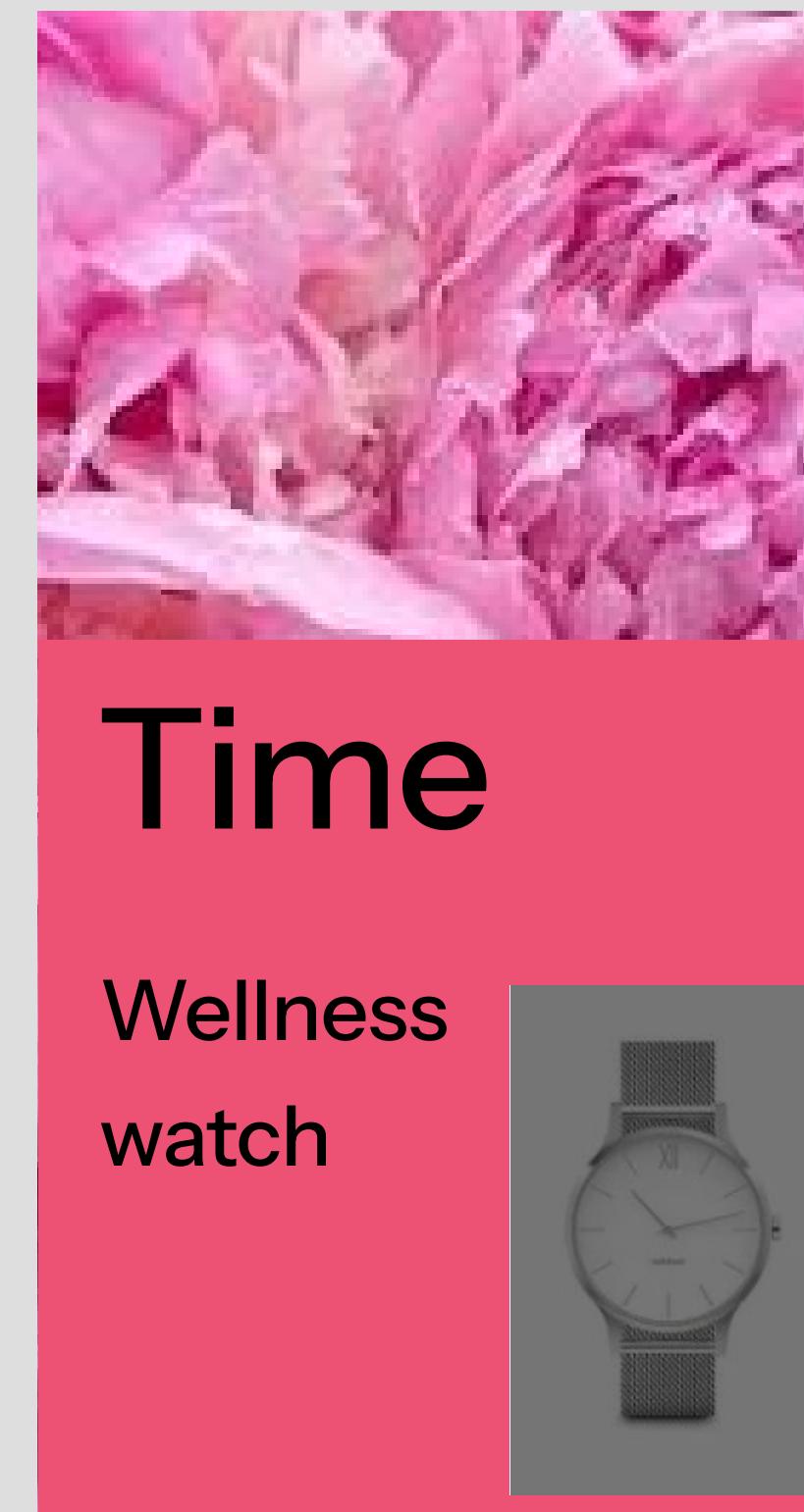
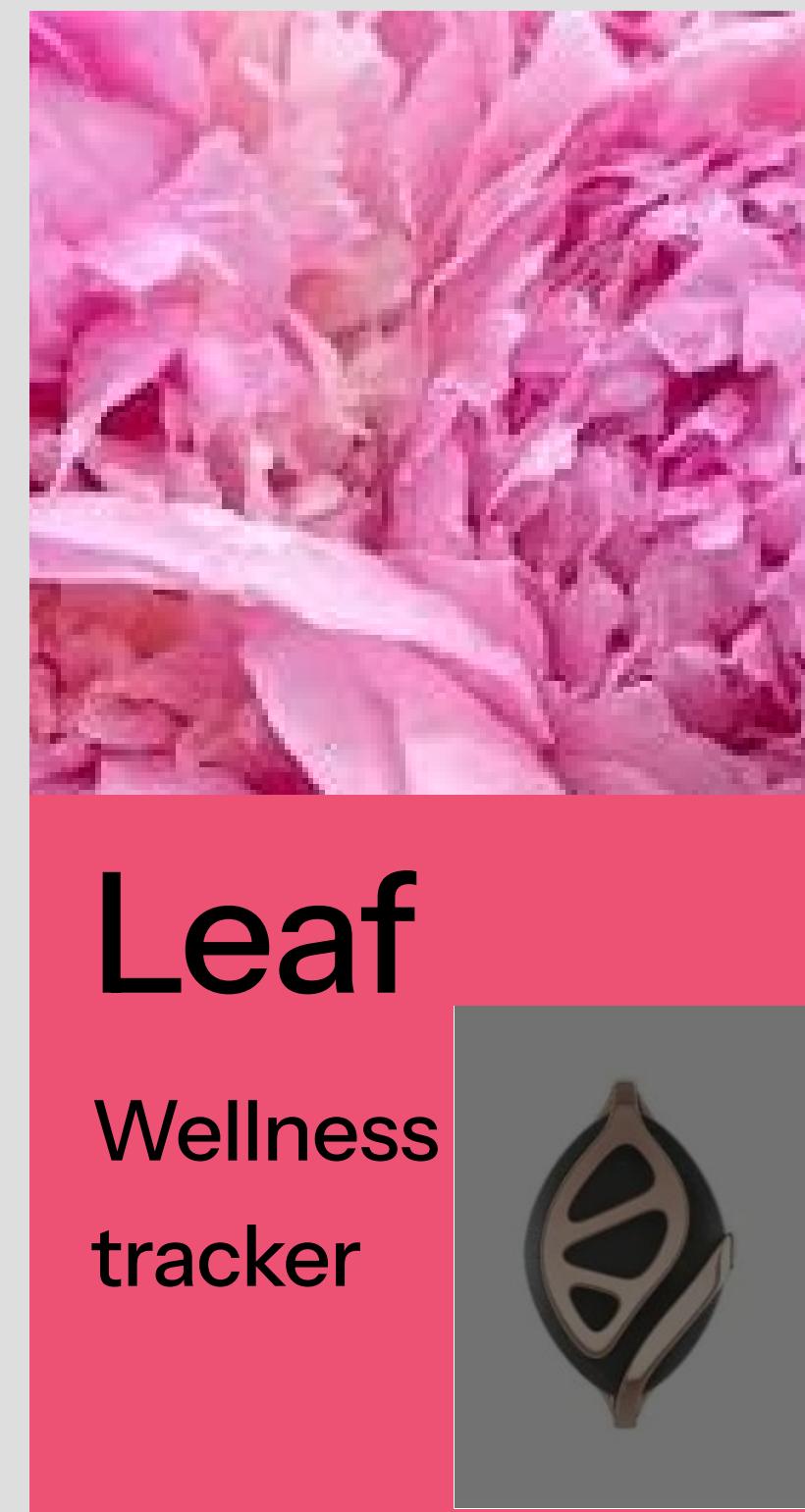
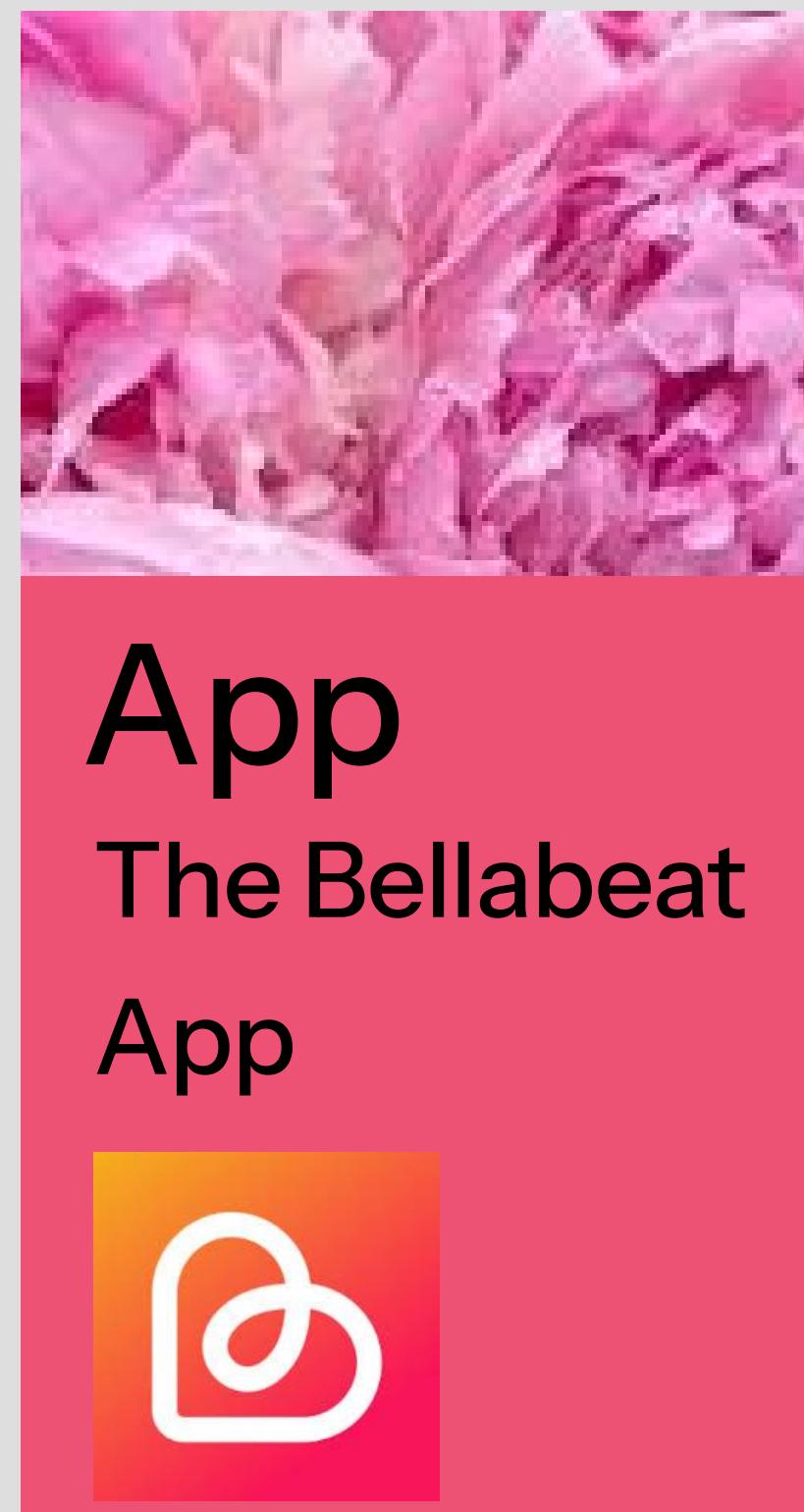
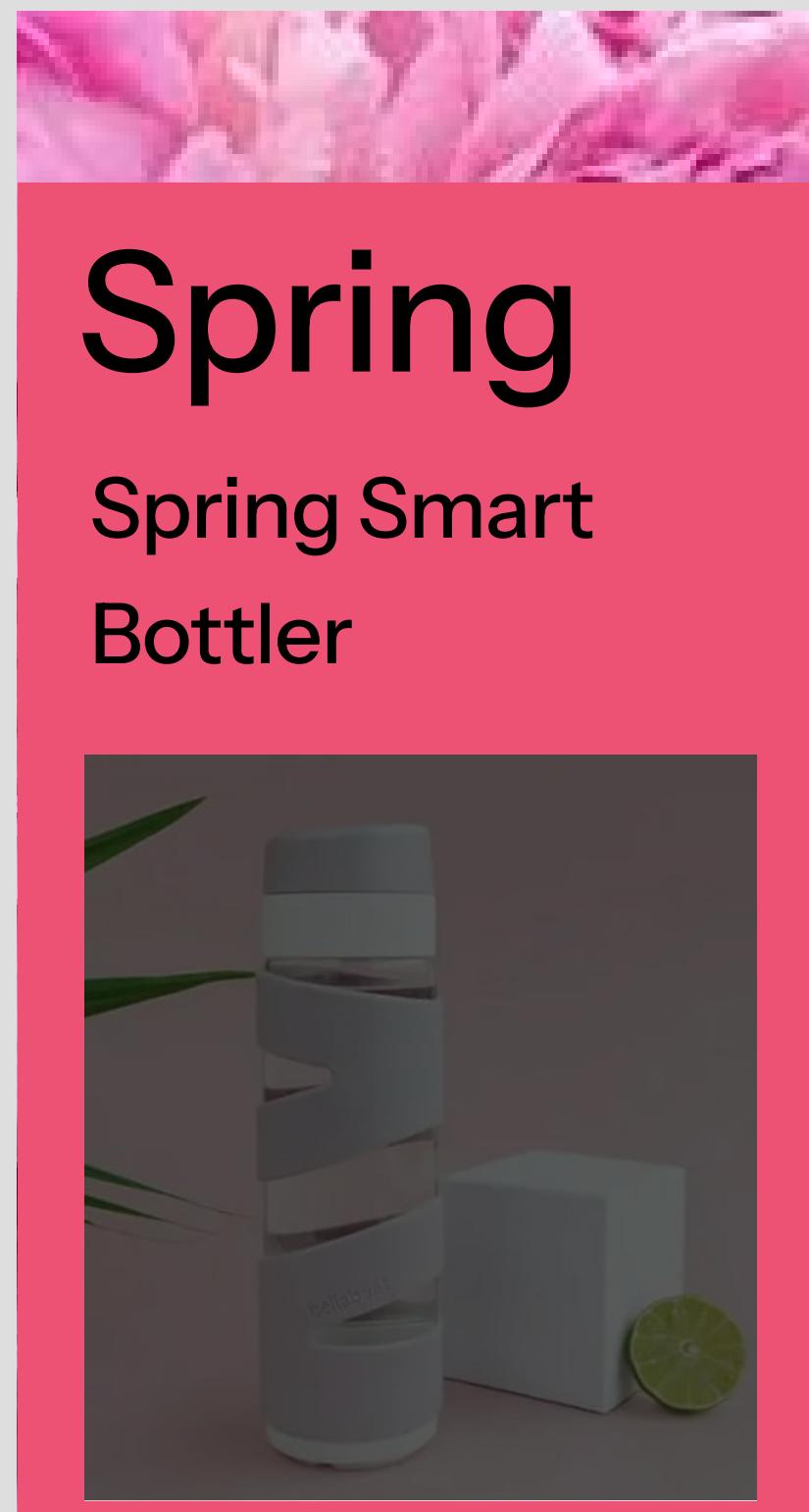
With focus on one product :

App

# The Ask - Business Task & Problem Statement

**The Business Task** -- analyze smart device data to provide insights to help Bellabeat grow into a larger player in the global smart device market

**Problem Statement** -- how are users immersing with their smart devices, and which opportunities can Bellabeats capitalize on to enhance their product & services based on this smart device data. Focus place on one from their suite of products, in this case the app. **Background at glance:**



# The Prep - Limitations & Data Detail Review

There are 29 dataset csv files (11 Million Rows of Data) pulled from Kaggle,  
Original Data Source -- <https://www.kaggle.com/datasets/arashnic/fitbit>

Data is outdated with  
entries: 3.12.16 - 4.11.16 &  
4.12.16 - 5.12.16

\*For this study, analysis is  
applicable for year 2016.

## Privacy

Only data provided  
with consent of  
users is used. IDs  
found in datasets  
are noted by  
anonymous  
numbers and not  
names.



## Format

The data format  
used here is Long



## ROCCC

The data does meet  
reliability, original, and  
cited. The only  
exceptions are  
comprehensive and  
current. This is due to  
the diminished unique  
user count in some  
datasets and outdated  
entries which are no  
longer relevant/current

# The Process - Tool Sets & Cleaning

Datasets uploaded to several tools - BigQuery, Excel, Google Sheets, R Project, Tableau, and Gemini Ai. Each tool was chosen to avoid dependency on any given one tool and increased redundancy.

## BigQuery

Data Cleaning: All datasets containing “date/time” were separated for BigQuery upload

## Google Sheets + Excel

Inspected contents to verify data for columns and rows made sense, no string text where values should be

## Gemini Ai & R Project

Ran analysis through Ai & R Project which performed calculations to find mean, min, max, sums, quartiles. Duplicates removed here as well

# Data Manipulations - Calculations

Examples: **BigQuery (SQL)** for # Unique IDs in Datasets

**R Project** for quartiles, mean, median, etc

**Tableau** for new calculated totals such as Overall Daily Activity

Calculations used to find # of Unique IDs in weightLogInfo and other datasets - BigQuery (SQL)

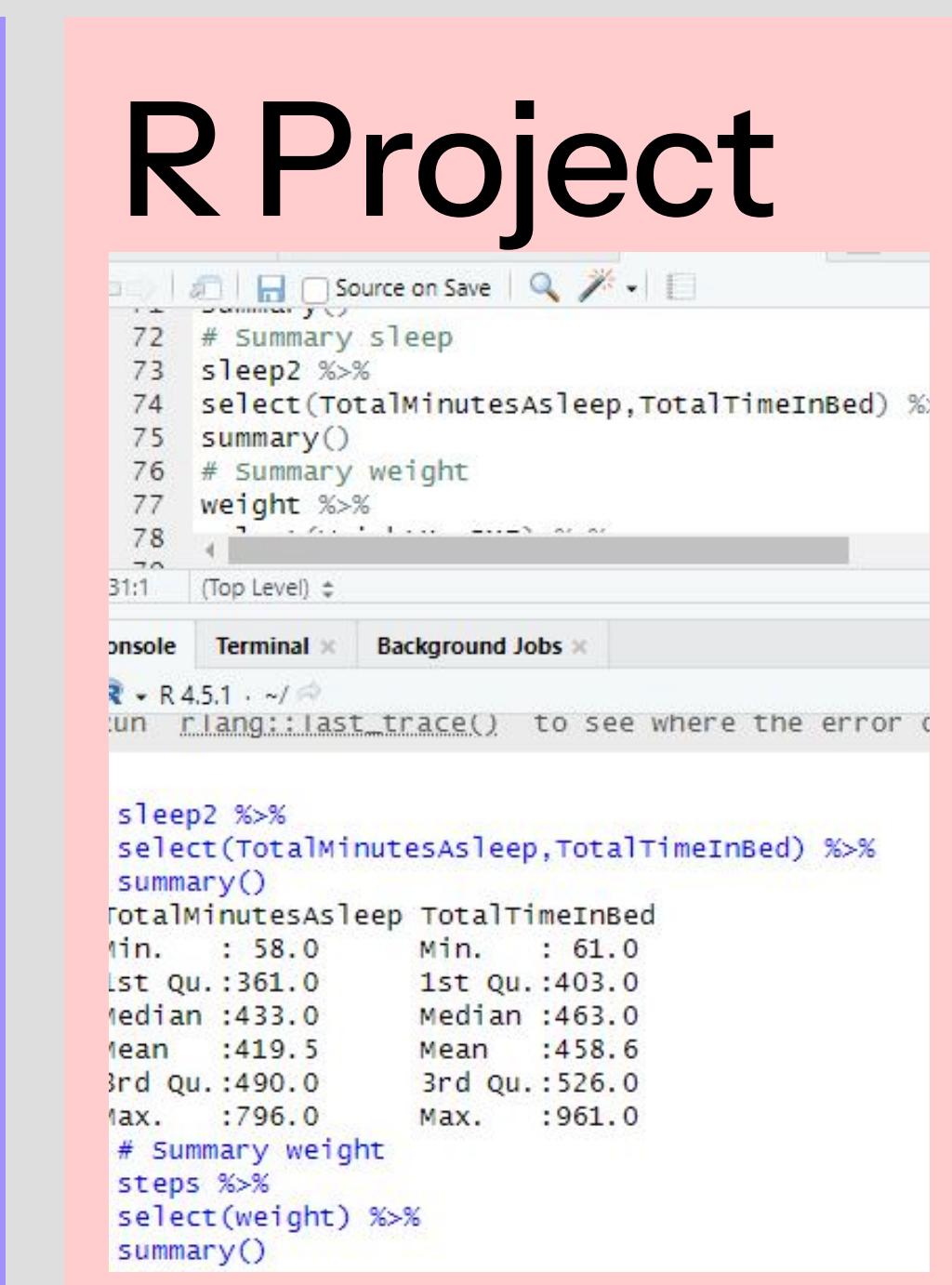
Big Query - SQL Left Join Weig...  Run  Save query  Download

```
1 SELECT Distinct T1.ID FROM
2 `woven-environs-460417-p8.PathtoPower.weightLogInfo_merged` AS T1
3 Left JOIN `woven-environs-460417-p8.PathtoPower2ndSet.weightLogInfo_merged` AS T2
4 ON T1.Id = T2.Id
5 LIMIT 1000;
```

This query will process 264 B when run.

**SQL**

**R Project**



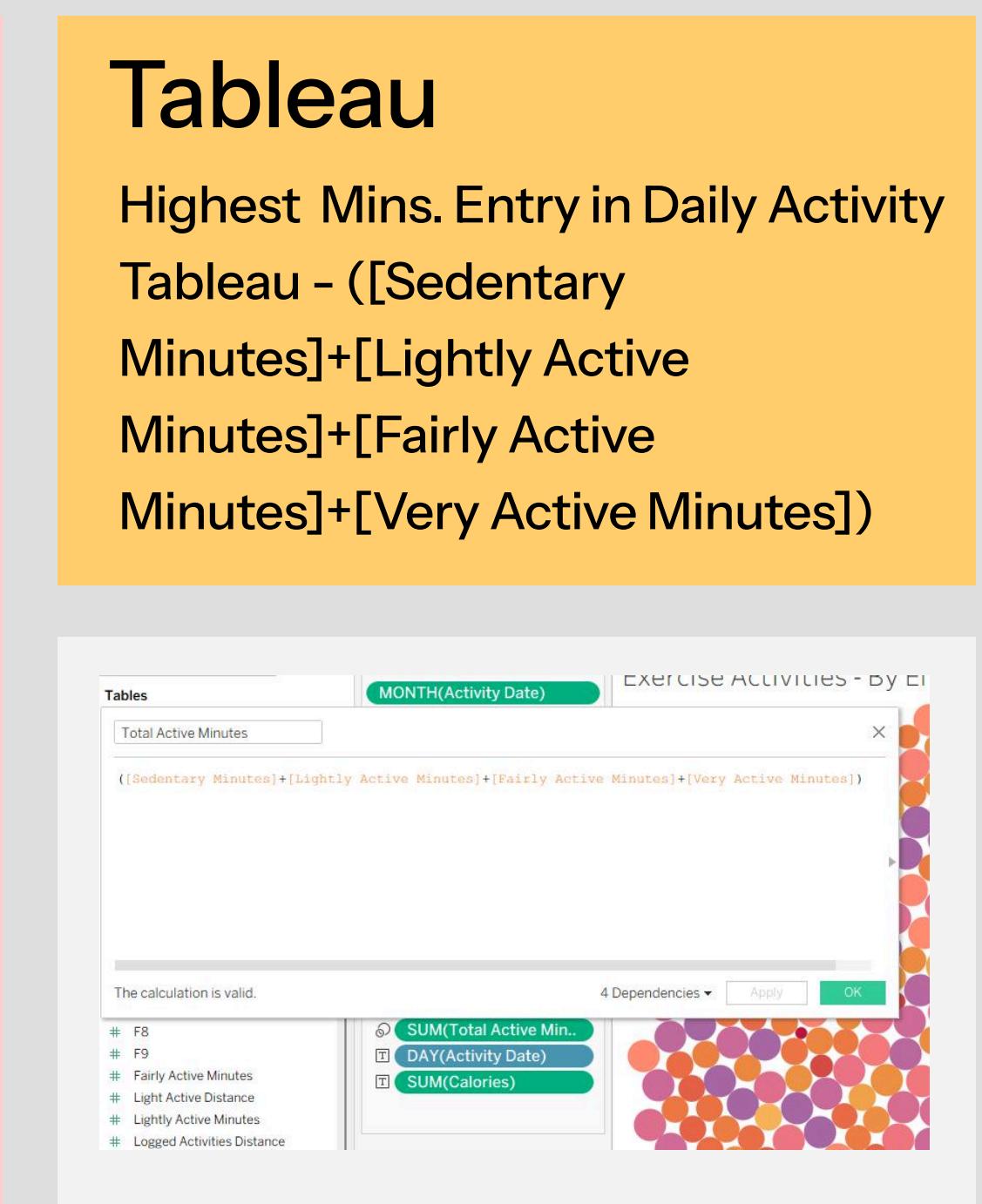
```
# Summary sleep
sleep2 %>%
  select(TotalMinutesAsleep, TotalTimeInBed) %>%
  summary()
# Summary weight
weight %>%
  select(steps) %>%
  summary()

# Summary sleep
sleep2 %>%
  select(TotalMinutesAsleep, TotalTimeInBed) %>%
  summary()
totalMinutesAsleep TotalTimeInBed
Min. : 58.0      Min. : 61.0
1st Qu.:361.0    1st Qu.:403.0
Median :433.0    Median :463.0
Mean   :419.5    Mean   :458.6
3rd Qu.:490.0    3rd Qu.:526.0
Max.   :796.0    Max.   :961.0

# Summary weight
steps %>%
  select(weight) %>%
  summary()
```

**Tableau**

Highest Mins. Entry in Daily Activity  
Tableau - ([Sedentary Minutes]+[Lightly Active Minutes]+[Fairly Active Minutes]+[Very Active Minutes])



Tables MONTH(Activity Date) exercise activities - by day

Total Active Minutes

([Sedentary Minutes]+[Lightly Active Minutes]+[Fairly Active Minutes]+[Very Active Minutes])

The calculation is valid.

F8 SUM(Total Active Min...) OK

F9 DAY(Activity Date)

Fairly Active Minutes

Light Active Distance

Lightly Active Minutes

Logged Activities Distance

SUM(Calories)

# Before & After

After saving csv datasets locally, data time column needed to be formatted to be uploaded successfully in BigQuery

B1	A	B	C	D
1	Id	Time	Value	
2	2022484408	4/1/2016 7:54	93	
3	2022484408	4/1/2016 7:54	91	
4	2022484408	4/1/2016 7:54	96	
5	2022484408	4/1/2016 7:54	98	
6	2022484408	4/1/2016 7:54	100	
7	2022484408	4/1/2016 7:54	101	
8	2022484408	4/1/2016 7:54	104	
9	2022484408	4/1/2016 7:54	105	
10	2022484408	4/1/2016 7:54	100	

heartrate\_seconds\_merged

Ready Accessibility: Unavailable

# Split Text to Column

123	Date	Time	Value
6391747486	4/5/2016	4:00:50 AM	
6391747486	4/5/2016	4:01:00 AM	
6391747486	4/5/2016	4:01:10 AM	
6391747486	4/5/2016	4:01:20 AM	
6391747486	4/5/2016	4:01:30 AM	
6391747486	4/5/2016	4:01:35 AM	

heartrate\_seconds\_merged

# The Analysis - Focus on Trends, Gaps, & Insights

Surprisingly, the most notable trend is the high amount of sedantry minutes (idle) activity across all users. Analyzed using Tableau. For a fitness focused device having this being the single largest outlier metric is alarming.

## Sedantary Levels by Hue

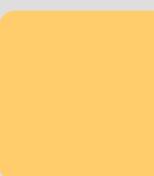
The glaring metric is there's no big red circles, only small ones. This means most of the large amounts consist of idle time



Highest Sedantary



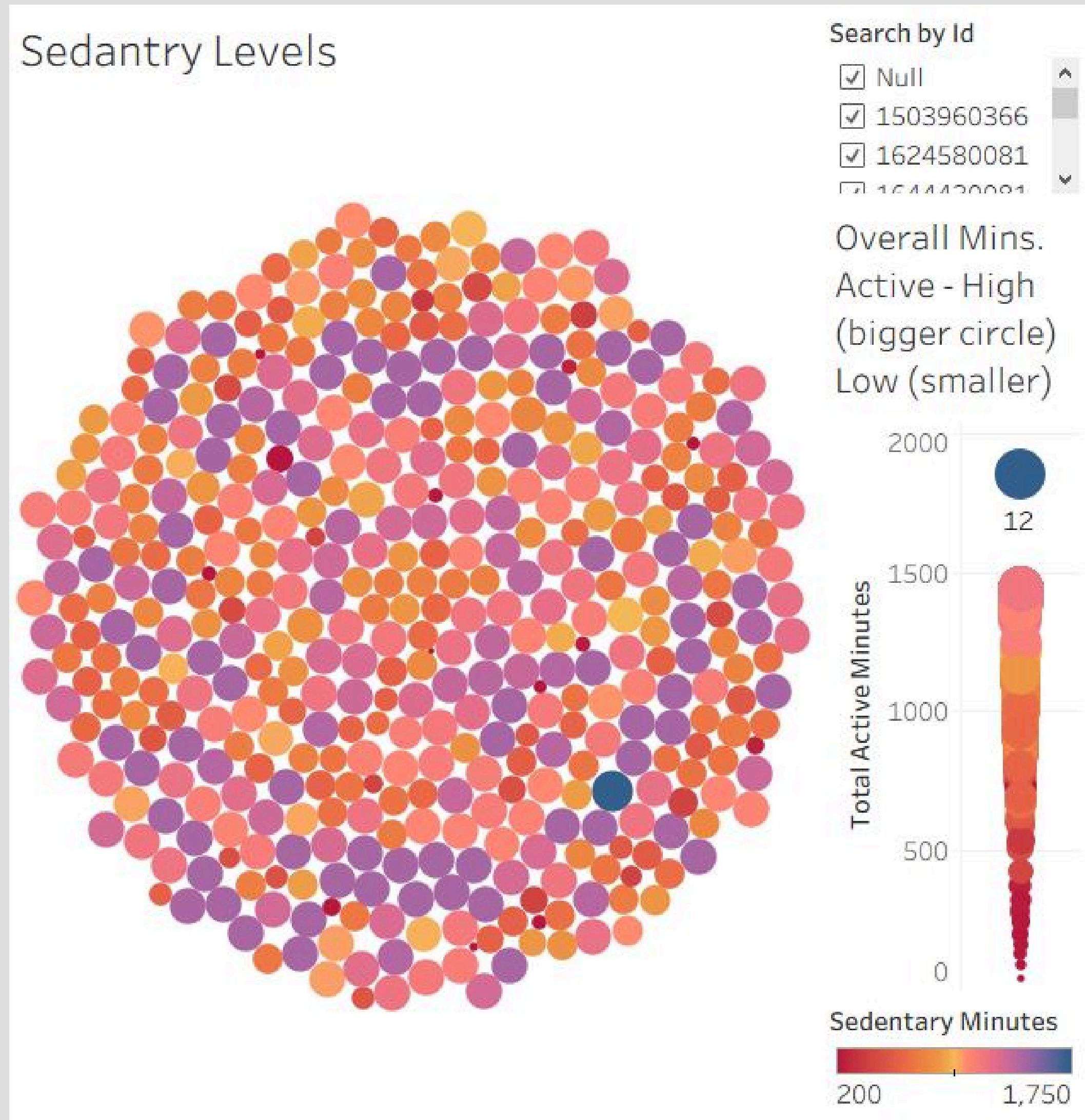
High Sedantary



Moderate

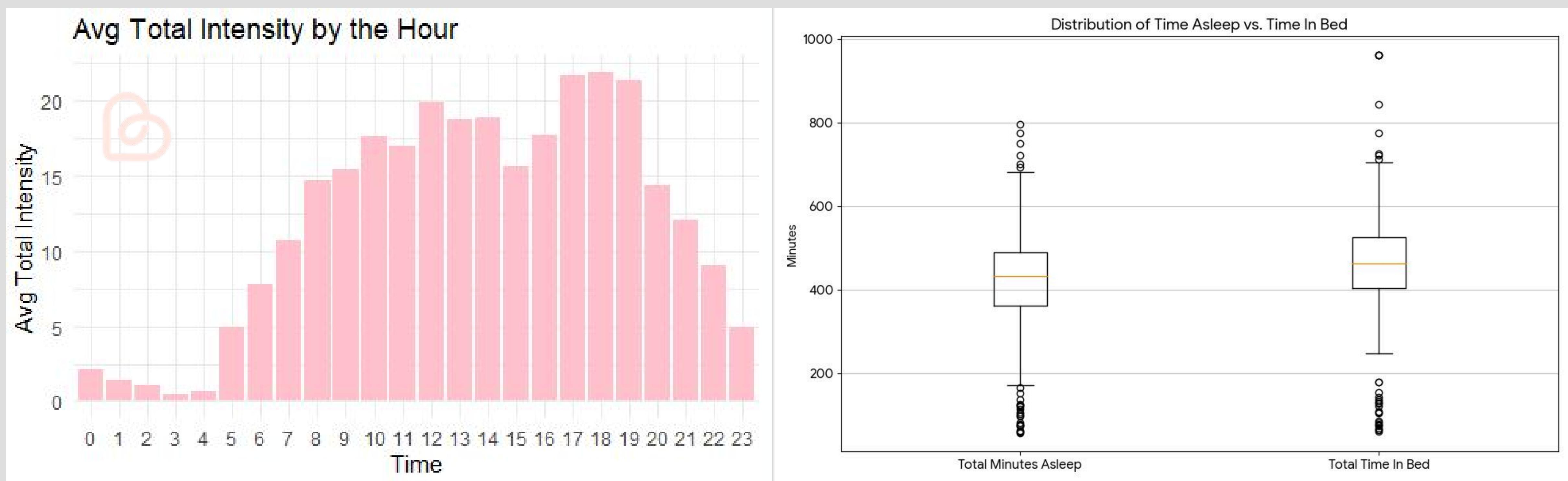


Low



# The Analysis Continued - Gaps, Timings, and Insights

For Peak Intensity, relationship exists where consistent user higher activity exists and peak times between 5PM to 7PM. We also see an insight of people on average 7.6 hrs (460 Mins.) in bed, yet only asleep for 7 hrs (420 Mins.) of that time. Surprisingly, there is a high "Time Awake In Bed." Analyzed using R Project and Gemini Ai, respectively.



# Share

Highest Recorded  
Metrics & Outliers  
Reveal

990~  
Sedentary  
Minutes daily

Significant Opportunity  
here. Have the app and  
smart devices use gentle  
reminders to move during  
prolonged idle periods

Peak Usage Trend:  
Most in 5 - 7PM window



39~  
Average Minutes  
Spent Awake in Bed

The app can use these  
metrics for personalized  
sleep guidance

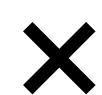
Thursday &  
Saturdays

Relationships exists where  
Users are most active on  
Thursdays and Saturdays;  
Sunday and Monday being  
lower activity days. Valuable  
weekly routine insight

# Share - The Story: How Success in People, Transforms Success in Bellabeats



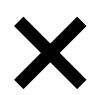
The key is in the timing. A person has so many things in a day, yet what if they saw the most critical moments, and jumped...



Rather than seeing the moments pass by,  
Bellabeats too can make the leap forward...



The data does show there critical times for success, that cannot afford to be missed...

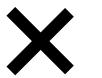


# A C T

Bellabeats can grow using the data, users are need help in...



What trends & insights does the data show to help market Bellabeats



Current usage reveal what the company can do next



## Recommendations: how Bellabeats can leverage their app & marketing

# Launch Marketing Campaign around “Prime Time”

Here the app can be used to target reminders, workout routines, and motivational content around the peak periods 5PM to 7PM. As well as Thursdays and Saturdays.

## Building up user base, while encouraging success in powerful moments.

# Sleep

Better rest

In a holistic partnership with not only exercise, the app can remedy and improve sleeping habits. Promoting “Ready in Bed” Mode. Promoting mindful habits and relaxation exercise in key moments of the night.



Prime  
Time

2024



Gamify

2025

Act now

2026

# Gamify Wrap: Yearly Wins

In a similar way Spotify Wrap summarizes achievements. Have users gamify their progress to include “sleep” streak” or “sleep score” with overall annual review of significant strides made by the end of the year.

Another action: Bellabeat should integrate heart rate monitoring with stress and mindfulness features too. Notifications can be sent with guided breathing exercises during elevated periods of inactivity.

**Position tool for holistic wellness,  
not just fitness**

# User Base

Bellabeats users a highly active group with a clear, cyclical routine, make the app part of the routine

## Overview of the App - the new hub

Making the app central for personalized guidance. Providing also Sleep Quality Score and actionable tips “wind-down routine” as well.

## BellaBeat Prime

Market a campaign on BellaBeat Prime to capitalize on the key windows (5-7PM)



## Revise Product Descriptions

Highlight holistic wellness tracking capabilities. Capitalize on stress management and sleep health into fitness

# Closing Remarks: Looking Forward Together



Sharing the appendix of where R Project, Tableau, Excel, and Google Sheets are all stored related to the research conducted

[View More](#)

Case Study Prompt & Background attached in google docs -  
<https://drive.google.com/file/d/1mUEUCi8u5Bj2bj5sQRnJN7uU34mA129r/view?usp=sharing>

## Tableau

[https://public.tableau.com/app/profile/ruben.velasco/viz/Capstone\\_17535696/193400/ExerciseActivities-ByEntrying](https://public.tableau.com/app/profile/ruben.velasco/viz/Capstone_17535696/193400/ExerciseActivities-ByEntrying)

## Github

<https://github.com/rubeskadoobs/CapStoneDataAnalyticsProject/blob/main/RProjectCapstone.R>

## Google Sheets

<https://docs.google.com/spreadsheets/d/e/2PACX-1vSktKtNgDqy়umH1Es-Wp7GjOFplooS4yQ8tnFDcSVbEwPw2udKWJ2T1Z0u-i4DuKqjvNbs3gXgs6Oi5/pubhtml?gid=1732160294&single=true>