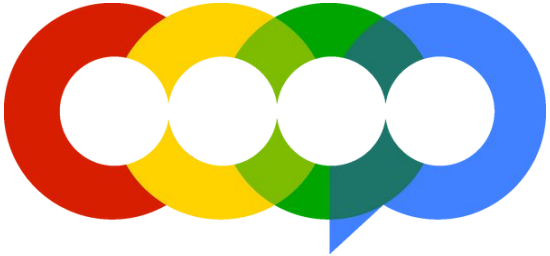
# Bellabeat

Google Data Analytics Capstone Project

Deyvis Mejia Zambrana

# About Me

- Graduated CUNY Hunter College in 2023 with a BA in biological science and a minor in geology
- 2+ years of work experience
- Graduated from COOP in the data analytics track in June 2024 (C395)
- Experience with R/Python programming, classical statistics, relational database SQL language, and Tableau


# Stakeholders

**Urška Sršen:** Bellabeat's cofounder and Chief Creative Officer

**Sando Mur:** Mathematician and Bellabeat's cofounder

Bellabeat marketing analytics team: A team of data analysts

Business Task: Analyze smart device usage to better inform Bellabeat marketing.

# Executive Summary

- ➔ Competitor Fitbit gathers nonsensical user data
- ➔ Fitbit data collection can be improved
- ➔ Fitbit users vary in the amount of consistency for Fitbit use
- ➔ Bellabeat can fill in the data gaps to be a stronger competitor, may include subscription-based information reporting for users depending on the user preference for information

# Methods, Data, and Limitations

**Limitations:** Fitbit data is incomplete and user use of Fitbit is inconsistent between subjects, there is also inconsistency within subject Fitbit use. Fitbit users are not all women.

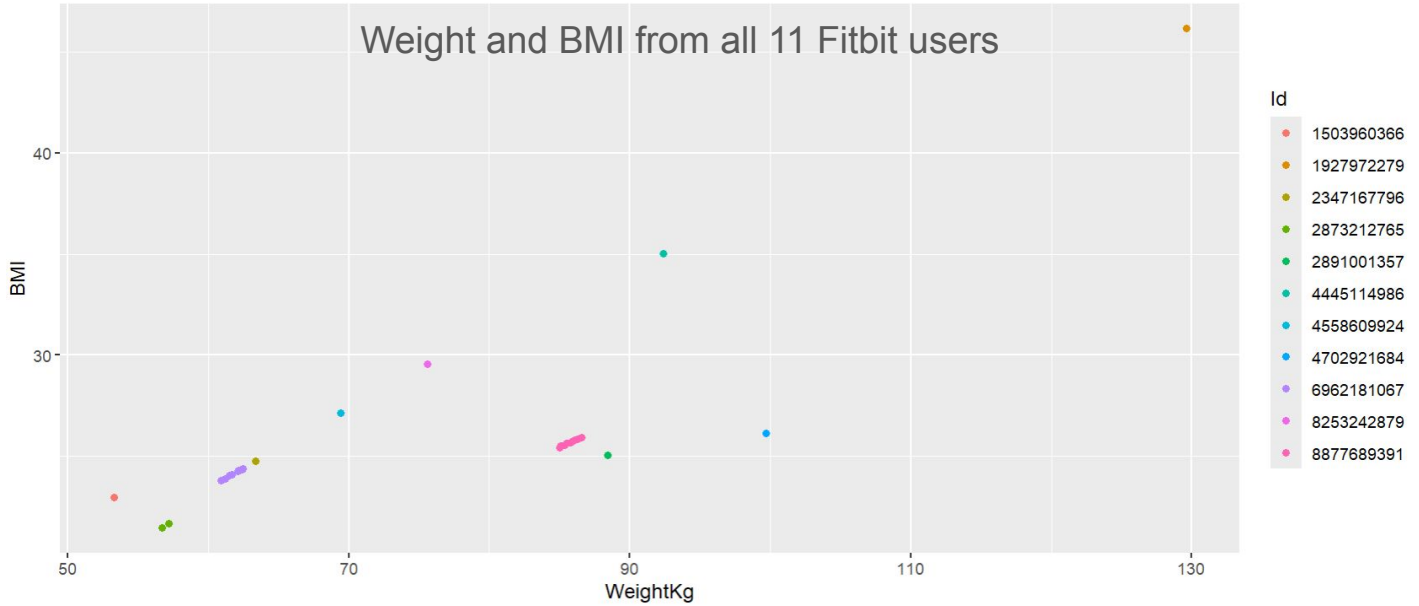
**Data:** Four data sources from Kaggle.

**Methods:** R programming for data cleaning, formatting, analysis, and visualization. Statistical analysis was also done in R language within RStudio.

# Insights From: *Weight Log*

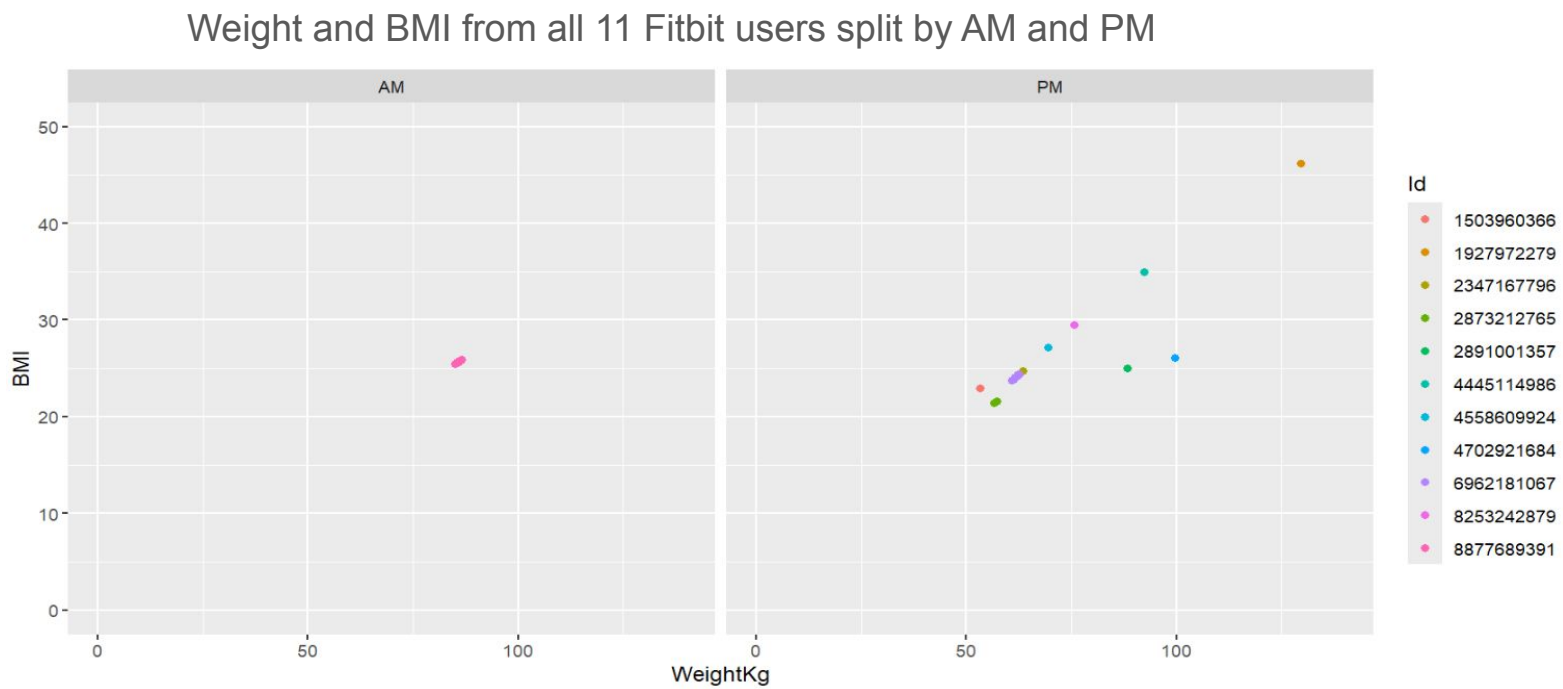
# BMI and Weight are Canonically Related

BMI uses weight in it's calculation as well as age and height, it assumes older and taller people will have larger weights but also that fat increases as weight increases. BMI can change because weight can change.

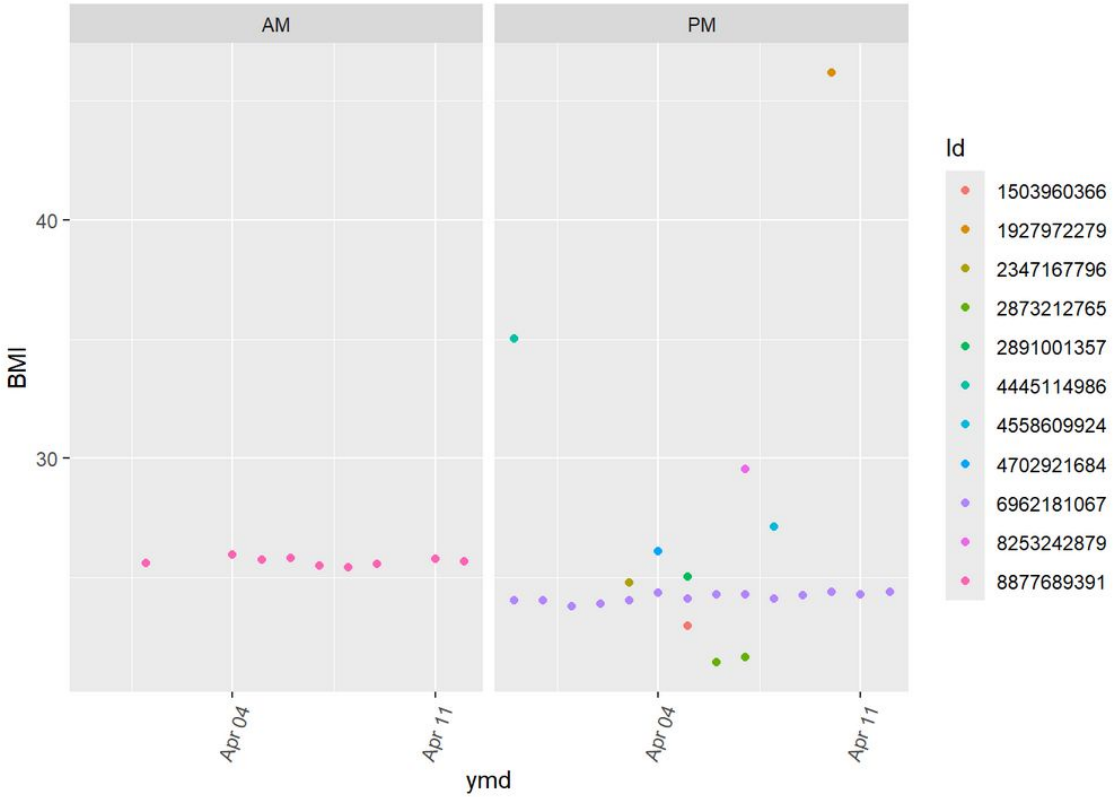



Bellabeat can help better inform user health beyond BMI. Engaging more users could depend on user education.

# More Users Track Weight and BMI During AM Hours



# AM Times are More Common for Data Tracking

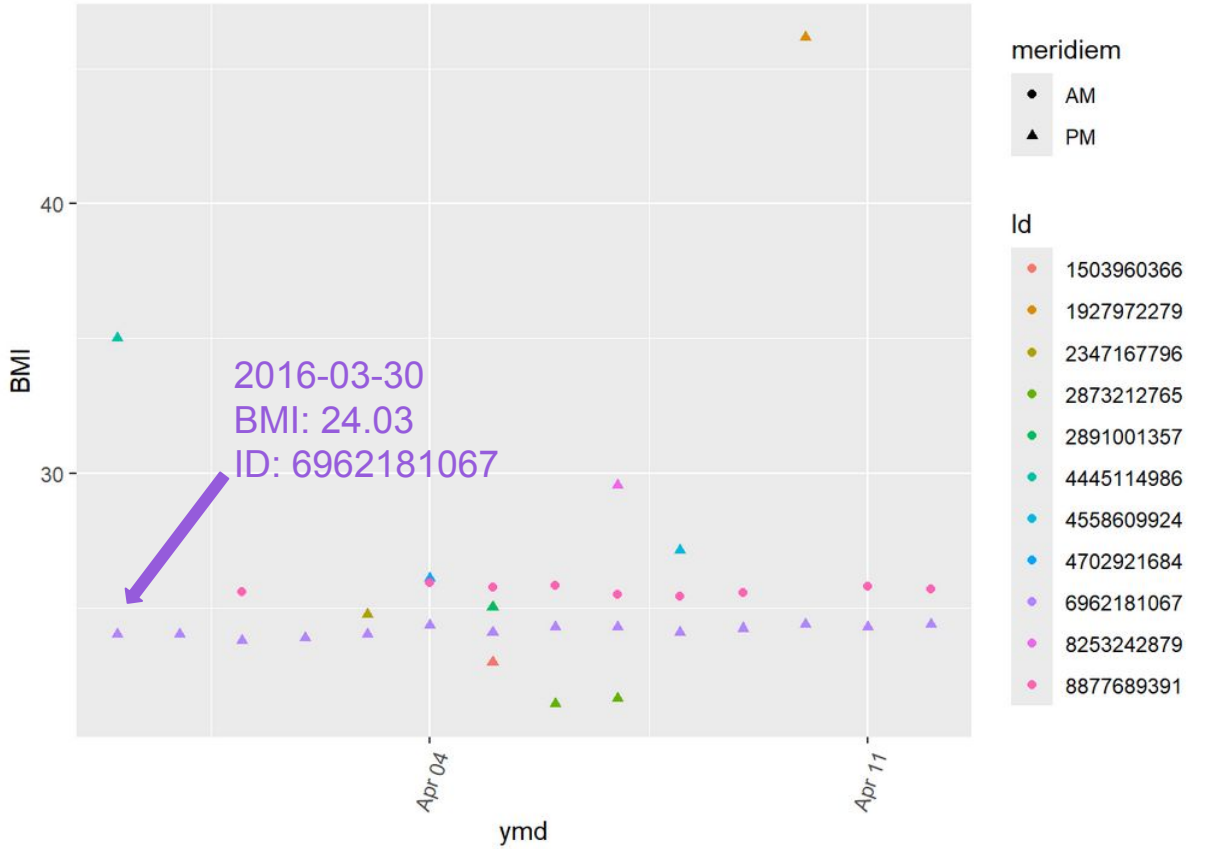



Most users do not track their BMI consistently.

Bellabeat can time workout notifications during active hours.

BMI calculator

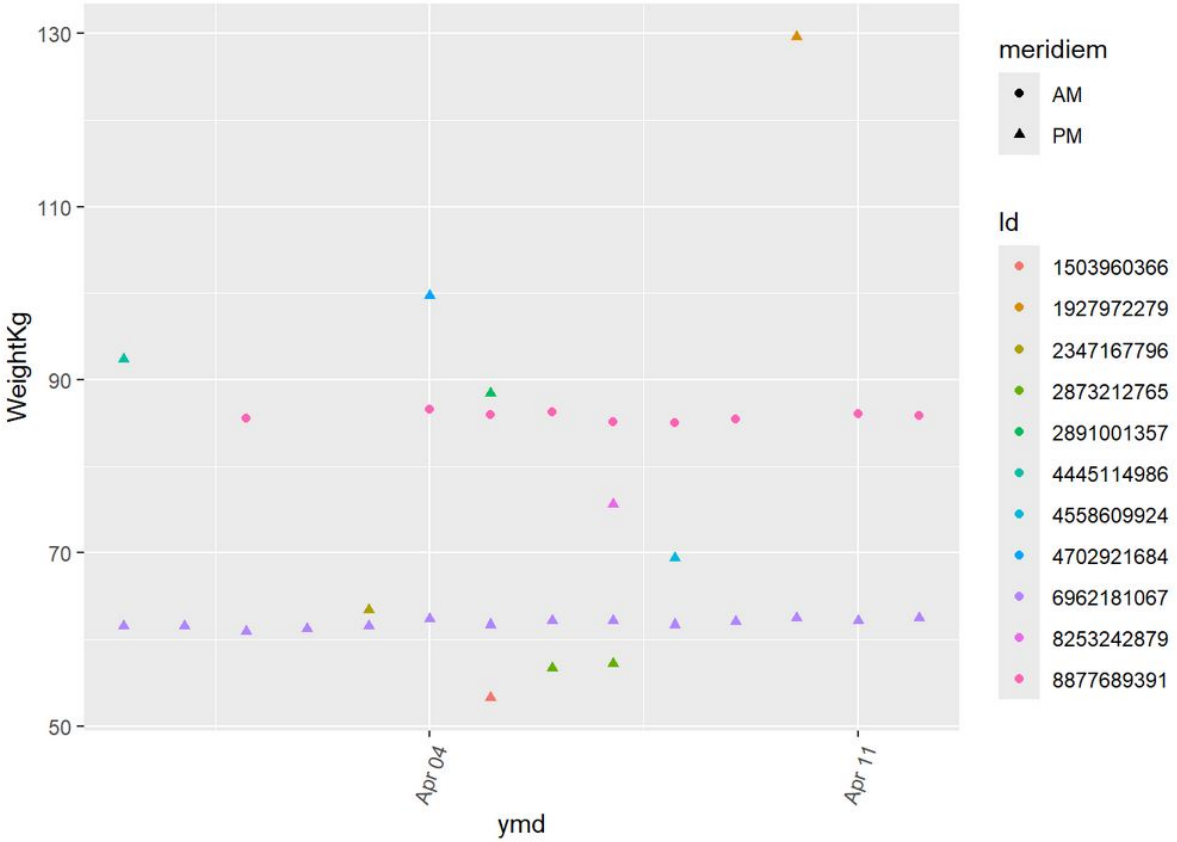
# Some Users Track Their BMI More Consistently




Mosy BMI for users remain under 30 through time.

Bellabeat could implement a data dashboard for BMI.

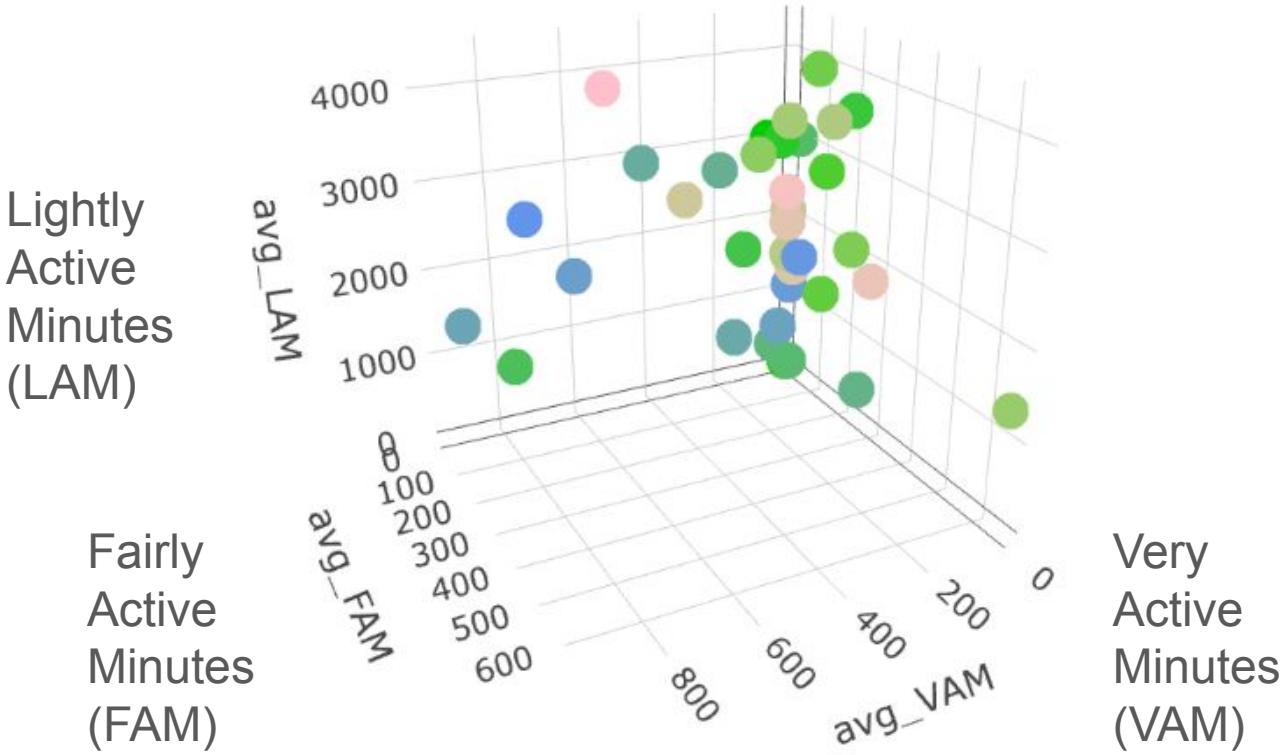
# Weights tracked through time are inconsistent



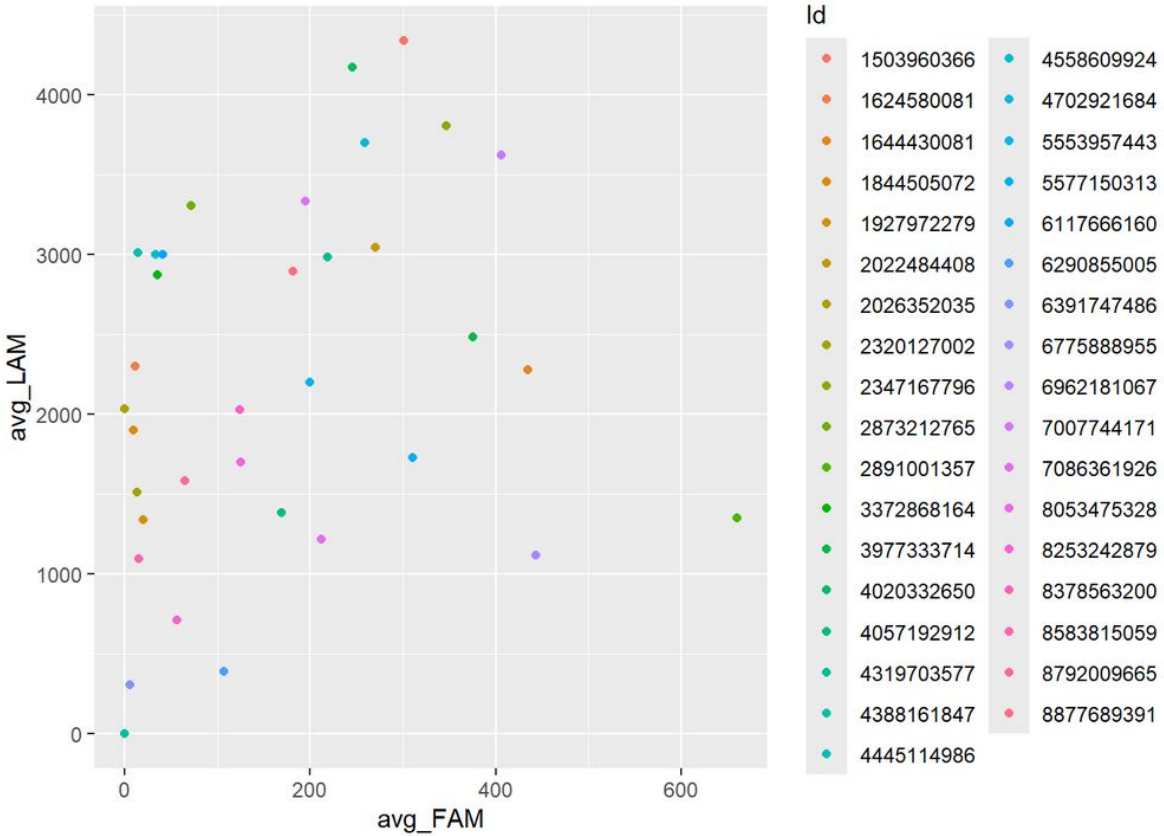

For some weight gain (or loss) may be a goal.

# Insights From:
# *Daily Activity*

# What are the relationships between active minutes categories? Are any predictable from the other?

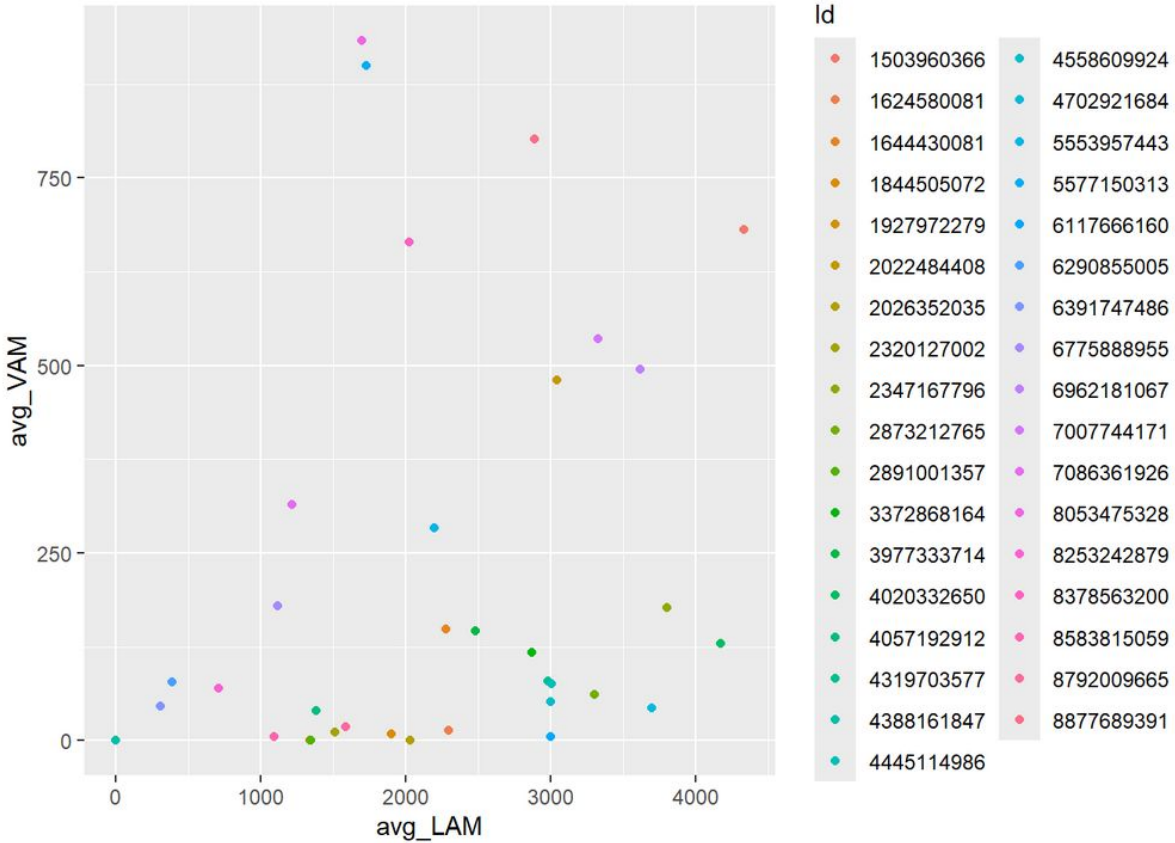


# FAM Is Not a Great Predictor for LAM



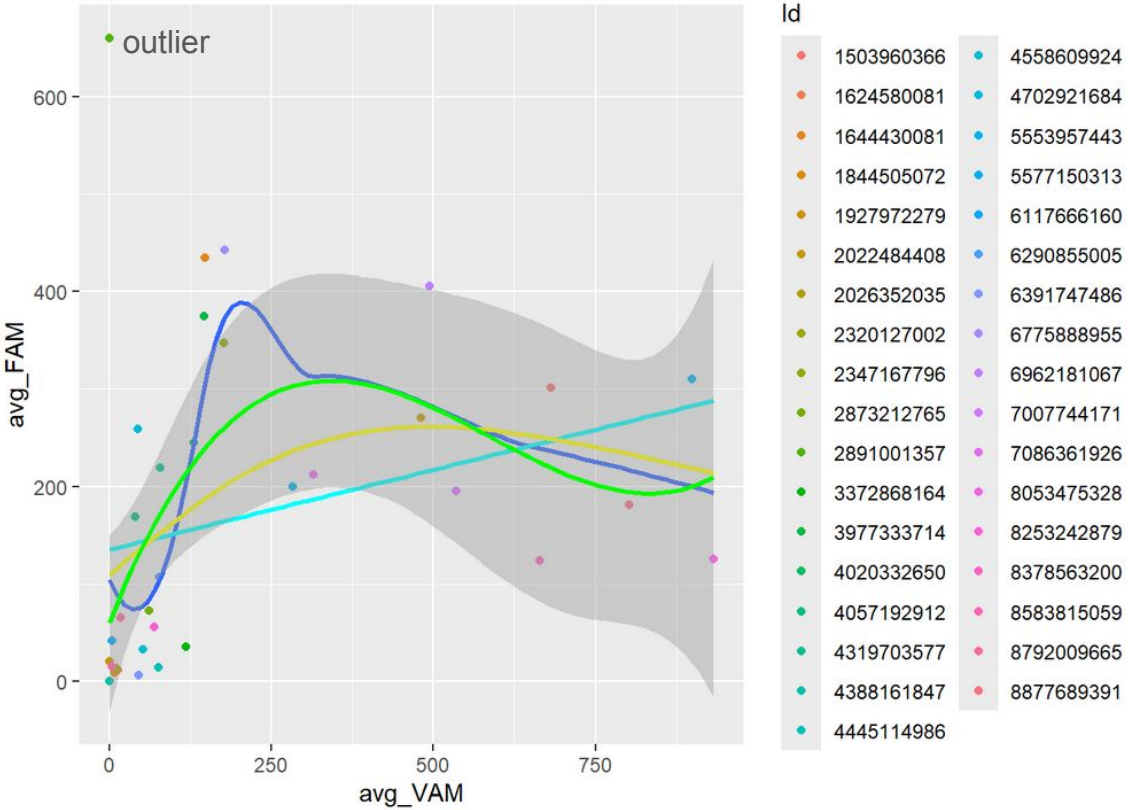

Generally, there is a slight increase for a user's average lightly active minutes for their average fairly active minutes.

# LAM Is Not a Great Predictor for VAM




Most of the points remain under 250 very active minutes irrespective of lightly active minutes.

# VAM Can Predict FAM: Exploring Different Models

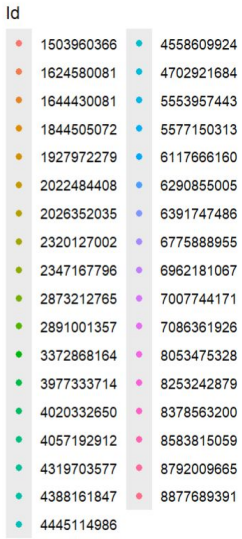



The blue LOESS model can show the FAM relationship is *not* linear.

The green polynomial model represents a polynomial relationship.

The model output is affected by the labeled outlier.

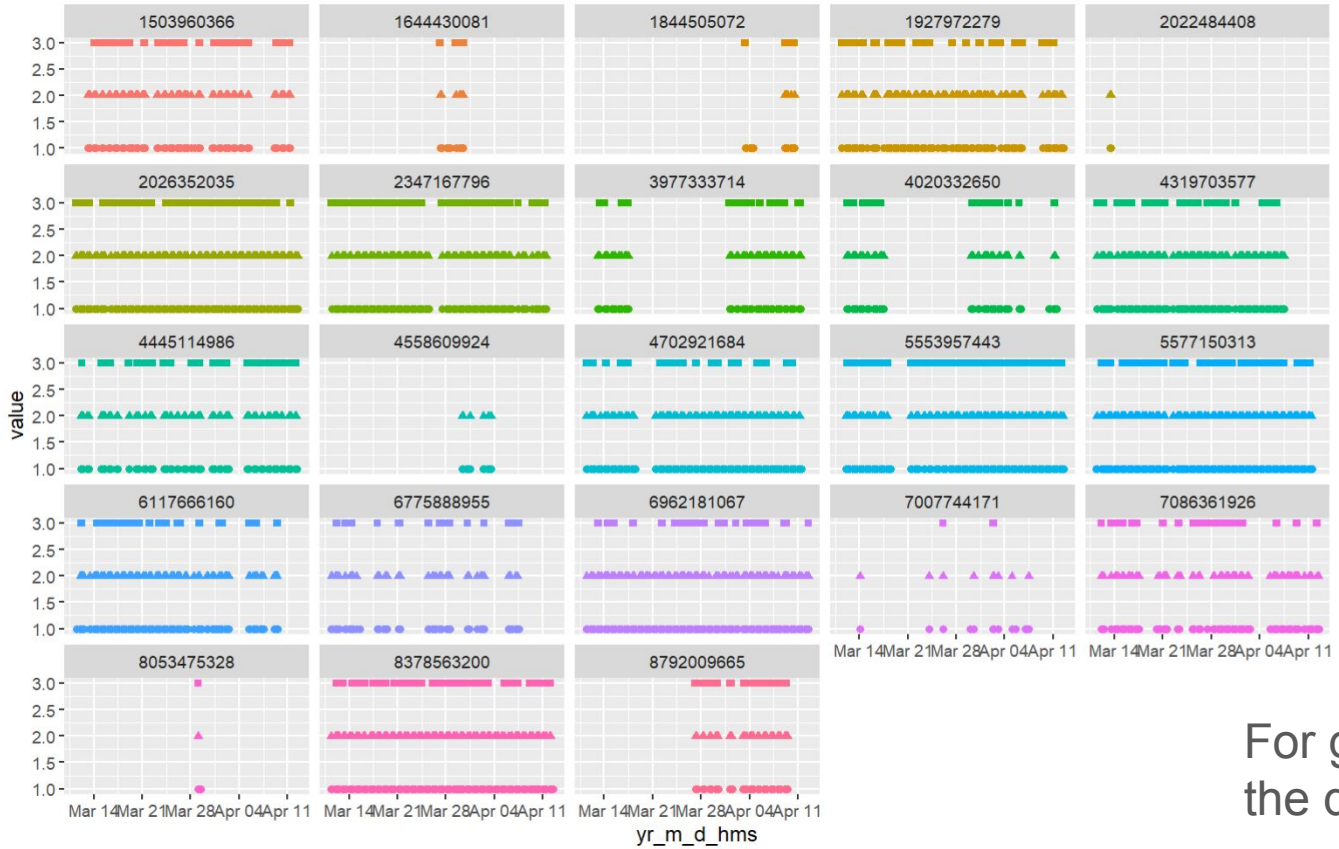
# Piecing Together a FAM Model from VAM


Fairly active minutes goes up quickly as VAM increases, peaks at around 450 with very active minutes between 125-250 then comes back down as very active minutes increases past 300.

# Insights From: *Sleep Minutes*

# **Overview:** Fitbit Sleep Tracking Is the Most Consistent

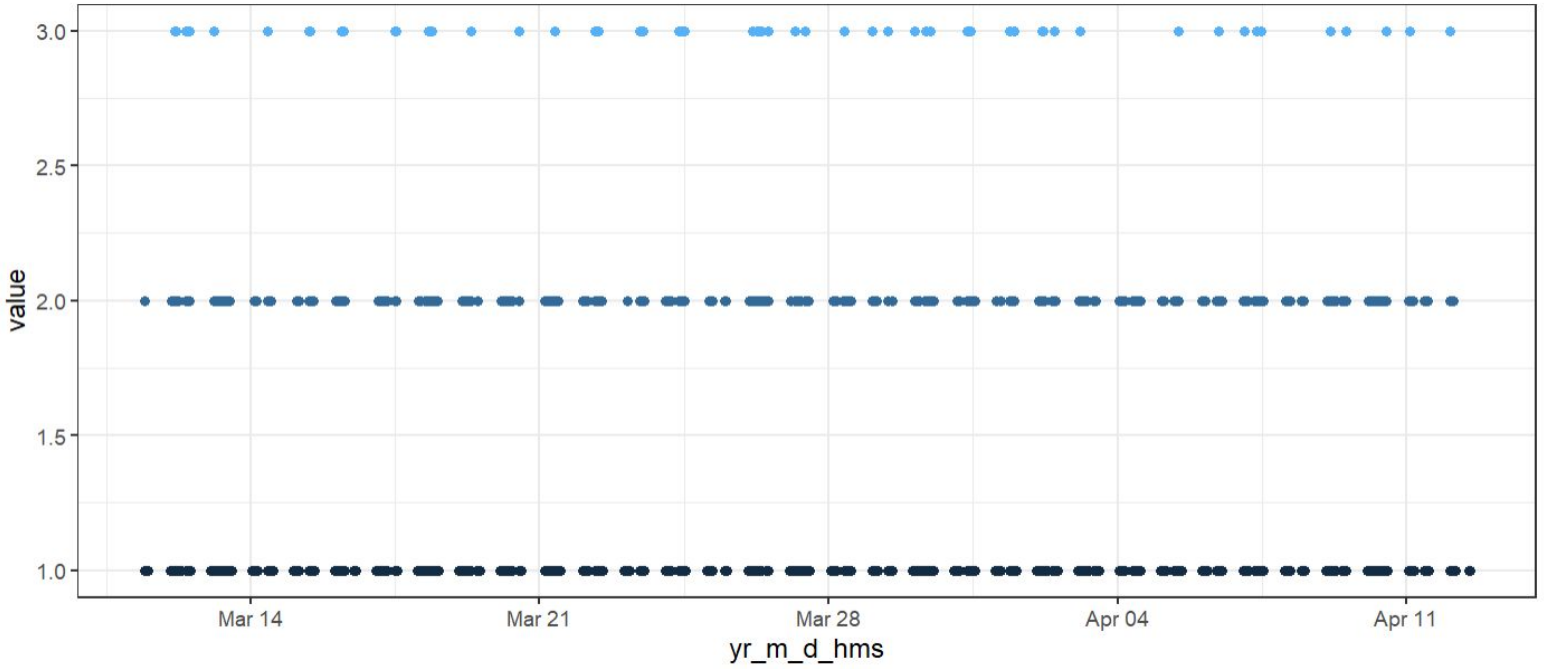



Values:
1 = asleep
2 = restless
3 = awake

Restlessness is ubiquitous! More sleep coverage is healthier.

For gaps it's assumed the device was not used

# 8378563200 Id: A Detailed View

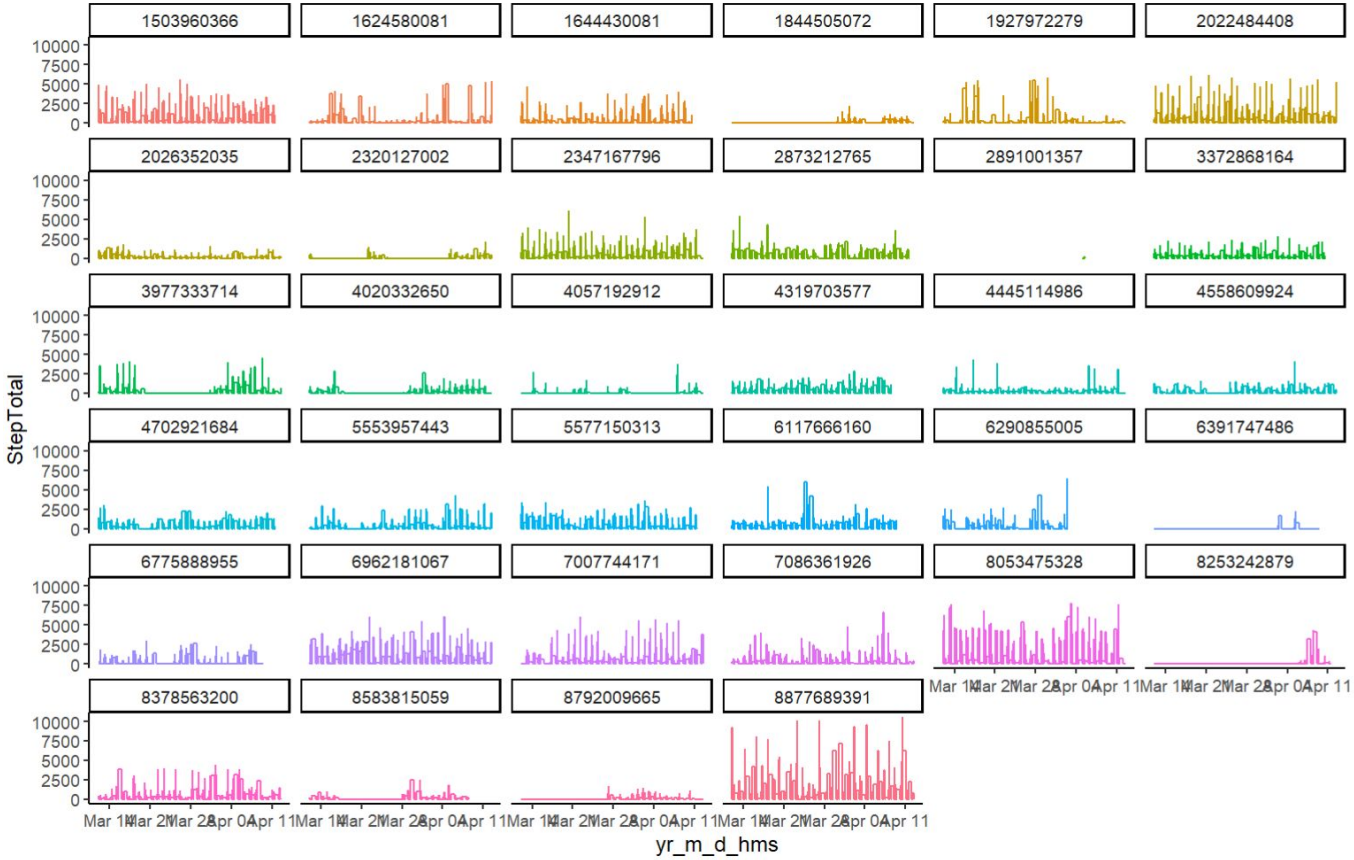



Values:
1 = asleep
2 = restless
3 = awake

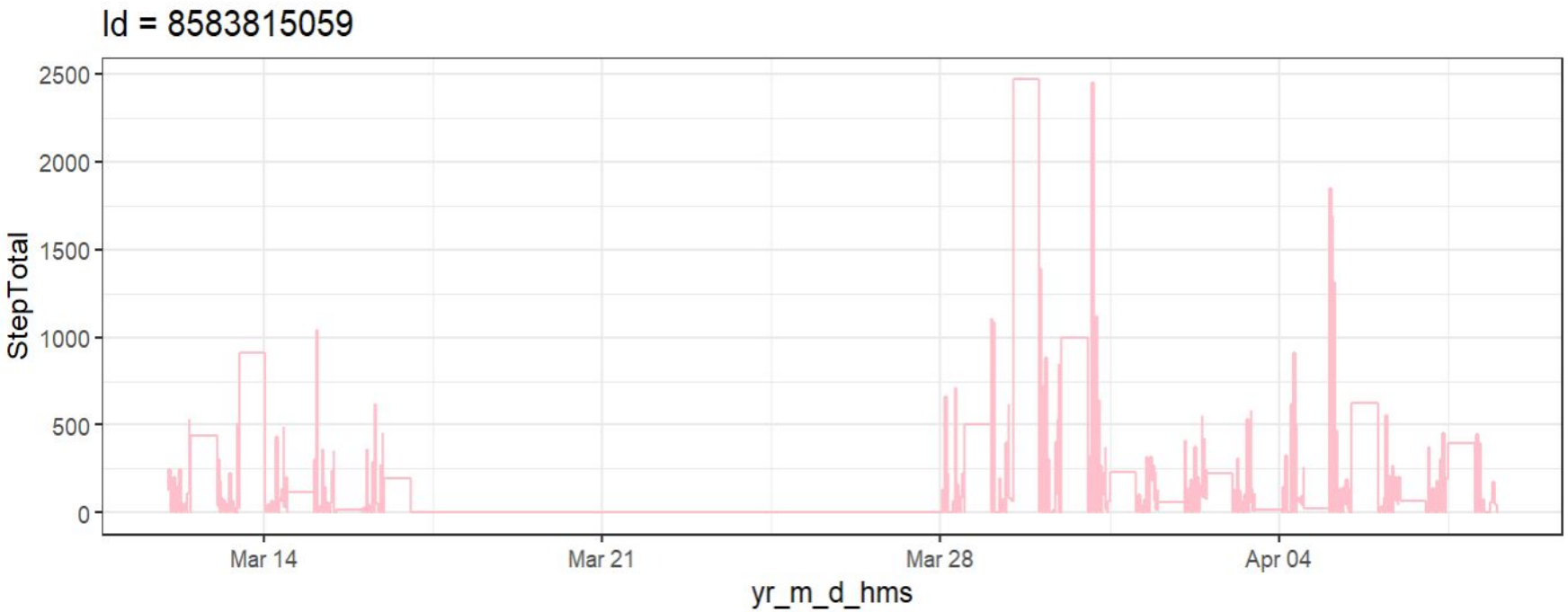
# Insights From: *Hourly Steps*

# **Overview:** Fitbit Step Tracking Is Widely Used

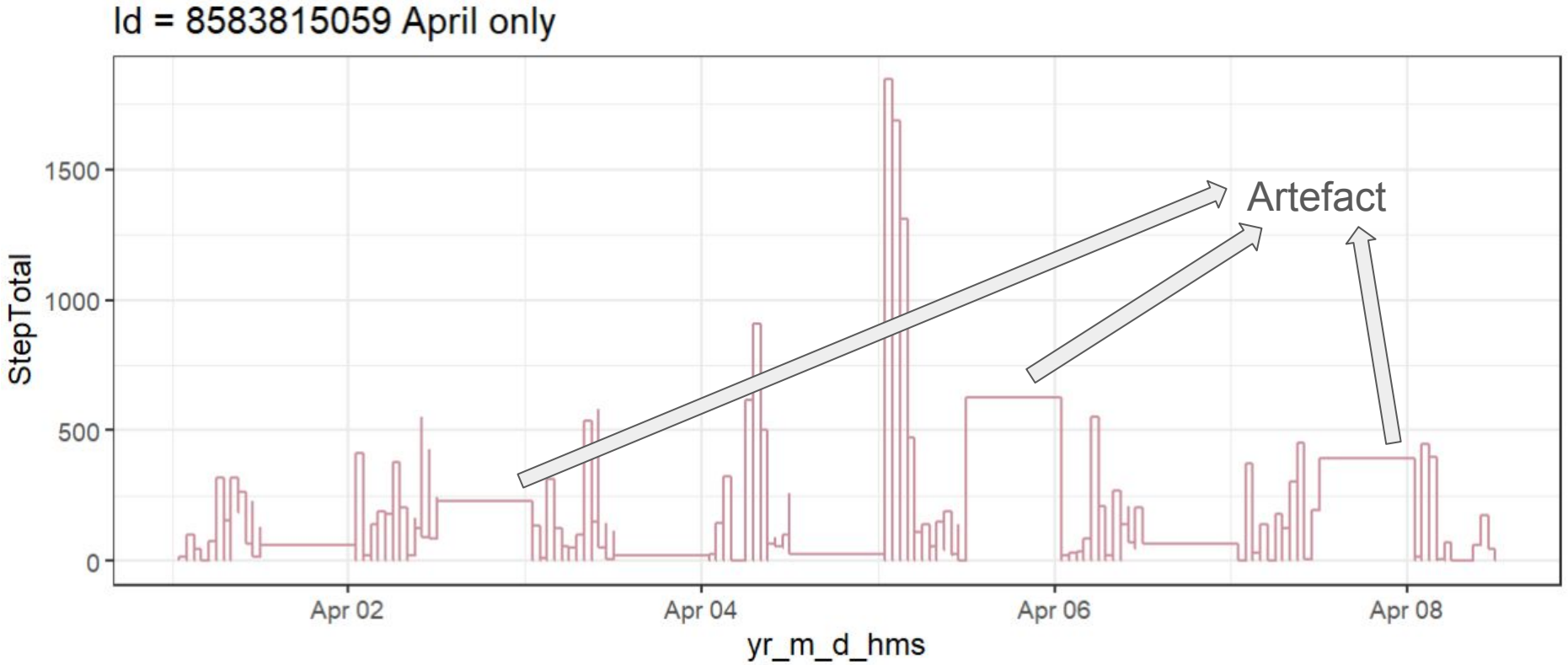
For most users, maximum hourly total steps taken remains consistent over time.



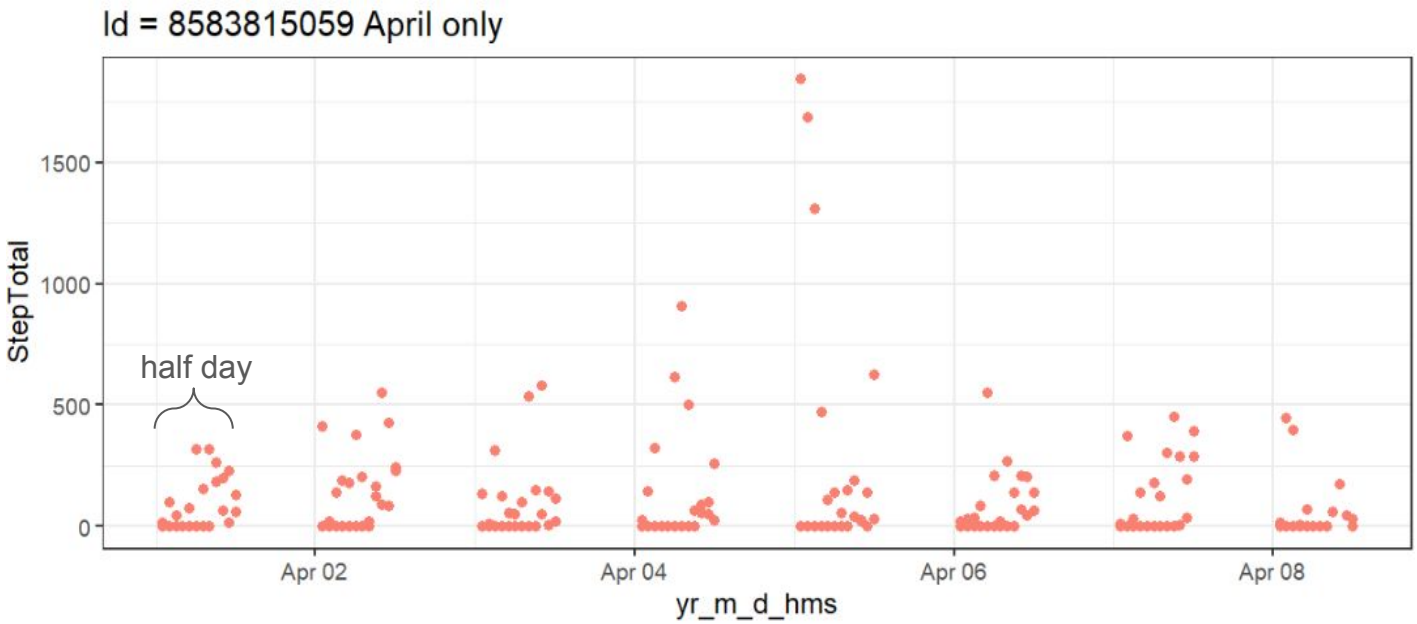
# Some Users Take Breaks from Step Tracking or Forget



# Total Steps Seem to Follow a Daily Pattern of Activity



# Granular View: Users use Fitbits During Certain Times




For some users, Fitbit use may be predominantly during the first half of the day.

Data can help Bellabeat understand when users are active.

# Recommendations

1. Bellabeat provide recommendations for healthy workout timed to user activity and data output via monthly or yearly subscription for user-level health information reporting.
2. Bellabeat can also market more health indicators beyond BMI which rely on self-reported data.
3. Targeted marketing to those that are consistently active, reduce marketing costs.

# References and Resources

- A complete guide on using R programming for data analysis with the data can be found at https://rpubs.com/deyvis305/bellabeat2
- Illustrated images extracted from https://bellabeat.com/fitness-exercise/running-during-period/
- Code dump in the Rmarkdown format can be found at https://github.com/deyvis305/rpubs_deyvis305/blob/main/Bellabeat%20capstone%20GoogleDA.Rmd
- Metadata can be found at https://www.fitabase.com/media/1930/fitabasedatadictionary102320.pdf

A special thank you to Kaggle users for their insightful questions on data validity.