"Smart Device Insights Report"

Alexander Rouse

5/12/2021

```
library(tidyverse)
library(knitr)

daily_activities_merged <- read_csv("Fitabase_working_data/daily_activities_merged.csv")
hourly_data_merged <- read_csv("Fitabase_working_data/hourly_data_merged.csv")
```

Smart Device Insights Report

The purpose of this report is to gain insights into the use of non-Bellabeat smart devices. To do this three questions were looked at: 'When are people using their smart devices?', 'How are devices being used for sleep monitoring?', and 'How long are devices being worn throughout a day and how does that affect activity?' Then, using insights gathered from answering these questions, this report aims to help guide Bellabeat's marketing strategy for the Bellabeat Time smart device.

This report will cover the following:
1. Data sources used
2. Data limitations
3. Data credibility
4. Data cleaning and manipulation techniques
5. Summary of analysis
6. High Level Recommendations
7. Additional data that could be used to enhance findings
8. Appendix

1. Data Sources

The primary data source for this report is daily and hourly data from a group of 33 FitBit users between 4/12/16 and 5/12/16. For further insights and comparison, additional data from the FitBit website was also used. Lastly, for sleep data and recommendations, the CDC website's sleep section was used. For links to data sources, see Appendix.

2. Data Limitations

The following limitations are important to consider when applying insights from the primary data to Bellabeat's customer base.

- Bellabeat's user base is primarily women and the primary data does not list the sex of users.
- Although the sample size of 33 is large enough to be statistically able to represent a population, it is still a very small sample size when comparing to all smart device users
- The primary source of data was collected in 2016, making it five years old at the time of analysis.

- Not all users used the same device, which resulted in some differences in data collected. For instance, very few sleep records were recorded for people who wore their device for the entirety of a day.
- The data does not list what activities are being done by users. This could result in misleading data. A bike sprint on a stationary bike could produce the same data as having a panic attack, in terms of steps, heart rate, or intensity.

3. Data Credibility

Given the small sample size it was important find more data to help support the findings of the analysis. FitBit publicly shares some of its data for users to compare their usage with other users. Our primary source of data and the FitBit data for all FitBit users in the USA share two categories in common: Daily Steps and Sleep Averages.

- Primary Data Daily Average Steps: 8,329
- Fitbit Website's Published Daily Average Steps: 8,170
- Primary Data Daily Average Sleep Time: 419 minutes
- Fitbit Website's Published Daily Average Sleep Time: 436 minutes

For our data, the two categories are within 2% and 4%, respectively, of the average for all FitBit users in the USA. This suggests credibility when comparing our primary data and analysis to all FitBit users.

4. Data cleaning and manipulation techniques

This analysis used three main tools: Google Sheets, SQL (BigQuery), and R (RStudio Cloud). Google Sheets was used for the daily data. The small size of the dataset made a spreadsheet the ideal tool for analysis. SQL was used to analyze the hourly data. The dataset was large enough to create a lag for even simple changes in a spreadsheet, so SQL was a much more efficient tool for analysis. Lastly, Studio was used to help create visual representations of the findings and to create this document using R Markdown. For more detailed cleaning and manipulation processes, see change logs in Appendix.

5. Summary of analysis

As previously stated, this analysis looks at 'When are people using their smart devices?', 'How are devices being used for sleep monitoring?', and 'How long are devices being worn throughout the day and how does that affect activity?'
Note: All statistics and insights were derived from the primary data source unless otherwise noted

When are people using their smart devices?

This question analyzed data by average day, by day of the week, by hour of the day, and by hour of the week.

- The five most active hours across all days are from 12pm-2pm and 5pm-8pm, with the evening more popular on weekdays and the afternoon more popular on weekends.
- Saturday, Monday, and Tuesday all have above average (8,329) steps and the most active minutes. As Tables 1 and 2 show, steps and active minutes tend to be correlated.

```
daily_steps <- daily_activities_merged %>%
  group_by("Day" = DayOfWeek) %>%
  arrange(TotalSteps) %>%
  summarise(Average_Steps = mean(TotalSteps)) %>%
  arrange(desc(Average_Steps), .by_group = TRUE)

kable(daily_steps, caption = "Average Steps by Day", digits = 0)
```

Table 1: Average Steps by Day

Day	Average_Steps
Sat	8979
Tue	8927
Mon	8488
Wed	8191
Thu	8185
Fri	7821
Sun	7669

```
activity_by_day <- daily_activities_merged %>%
  group_by(Day = DayOfWeek) %>%
  summarise("Very_Active" = mean(VeryActiveMinutes),
  "Fairly_Active" = mean(FairlyActiveMinutes),
  "Total_Active" = mean(VeryActiveMinutes + FairlyActiveMinutes)) %>%
  arrange(desc(Total_Active), .by_group = TRUE)

kable(activity_by_day, caption = "Activity by Day in Minutes", digits = 0)
```

Table 2: Activity by Day in Minutes

Day	Very_Active	Fairly_Active	Total_Active
Tue	26	16	42
Mon	25	15	41
Sat	24	17	41
Sun	22	16	38
Wed	23	14	37
Thu	21	13	35
Fri	21	13	34

- For every day, except Saturday, the hours between 5-8pm are among the highest for average intensity.
- Saturday from 11am-3pm is the most active block of time in the week. Saturday from 1-2pm averages the most steps and the highest average intensity for all hours in the week. It averaged more than 10% more steps than the second most active hour in the week, Saturday from 2-3pm. All four hours in the Saturday from 11am-3pm block are in the top 8 hours of the week with the most average steps.

How are devices being used for sleep monitoring?

- For the available sleep records, the average user slept 419 minutes in a day. Table 3 shows that on Sunday the average user is sleeping ~50 minutes more than on Tuesday, Thursday, or Friday.

```
  sleep_per_day_average <- daily_activities_merged %>%
  filter(TotalMinutesAsleep != 0) %>%
  group_by(DayOfWeek) %>%
  summarise(tma = mean(TotalMinutesAsleep)) %>%
  arrange(desc(tma), .by_group = TRUE)

kable(sleep_per_day_average, col.names = c("Day", "Average Sleep (min)"), digits = 0,
caption = "Average Minutes of Sleep per Day")
```

Table 3: Average Minutes of Sleep per Day

Day	Average Sleep (min)
Sun	453
Wed	435
Mon	420
Sat	419
Fri	405
Tue	405
Thu	401

- The CDC recommends 7+ hours of sleep each night for adults and going to bed and waking up at the same time each day, including on weekends. They also state that 34.8% of women get less than seven hours of sleep each night.

How long are devices being worn throughout the day and how does that affect activity?

Splitting the users into quartiles for time wearing their smart device yielded some interesting results.

- Users who wore their device between 12 and 18 hours of the day spent the highest percentage of time wearing their device being either fairly active or very active. More than 30% more than the average for all users.

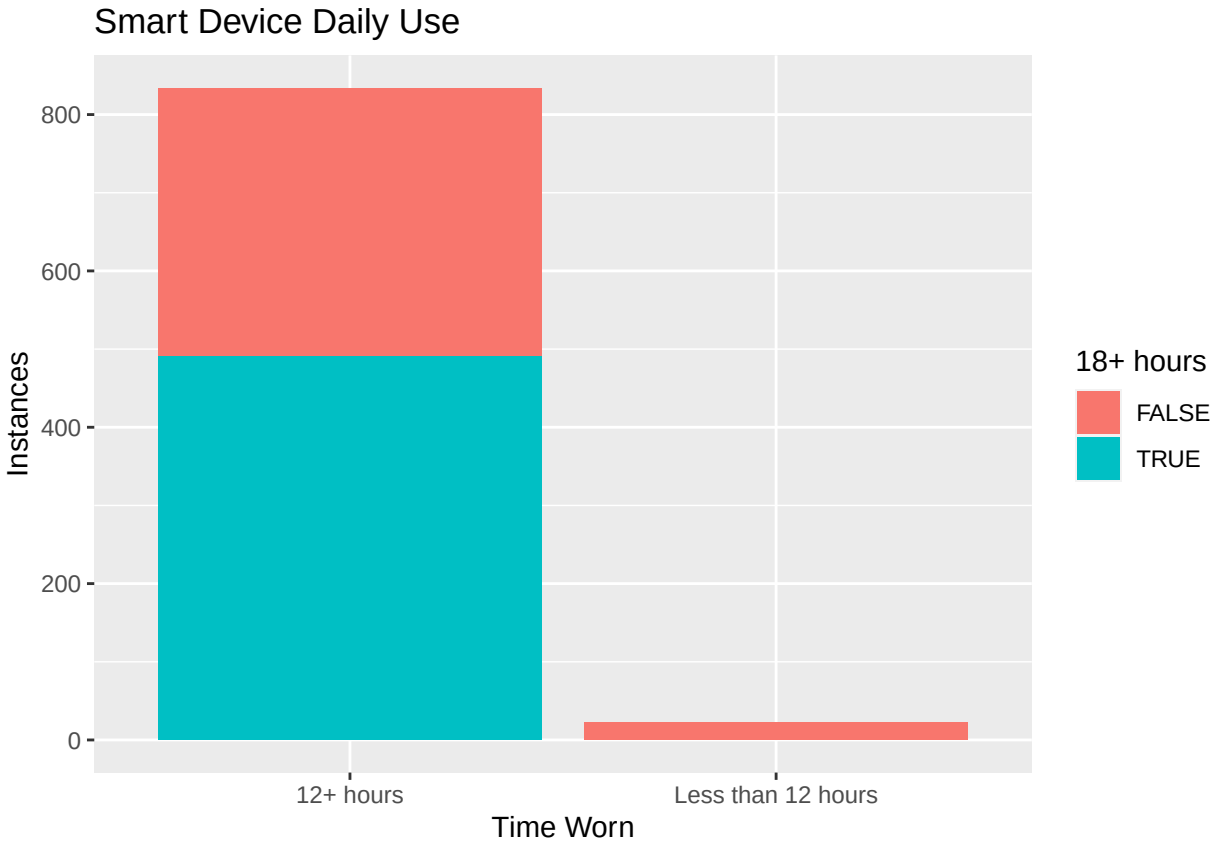
- The 12-18 hours worn quartile was more than 95% days with sleep records and also accounted for 81% of all days with available sleep records.

- Over 97% of all data is for users who wore their device for 12 or more hours of the day and over half is users who wore their device for 18 or more hours of the day.

```
Time_Worn <- daily_activities_merged %>%
  select(TotalMinutesWorn)
Time_Worn$day_use <- "Less than 12 hours"
Time_Worn$day_use[Time_Worn$TotalMinutesWorn >= 720] <- "12+ hours"

# most_worn identifies users with 1080+ minutes of wear
# it is stored as a boolean.  TRUE is 1080+ minutes worn
Time_Worn$most_worn <- TRUE
Time_Worn$most_worn[Time_Worn$TotalMinutesWorn < 1080] <- FALSE

ggplot(data=Time_Worn, aes(x= day_use, fill= most_worn)) +
  geom_bar() +
  labs(title = "Smart Device Daily Use", x = "Time Worn", y = "Instances", fill = "18+ hours" )
```



Other noteworthy findings

- Only four users of the 33 in the primary data source logged activities. Of those four, only two did so more than three times.
- The average FitBit user is overweight, according to the FitBit website.
- The most common activities done by female FitBit users are running, elliptical workouts, and biking, according to the FitBit website.
- According to the FitBit website, common user activities tend to mirror trends in the wider fitness community/popular culture, such as Crossfit in 2013 or Pokemon Go in 2016.

6. High Level Recommendations

Given the findings of this analysis I would recommend focusing on the following ideas for future Bellabeat Time marketing campaigns:

1. Bellabeat Time can be used to enhance overall wellbeing throughout the day.

- A large majority of users are wearing their smart devices for extended periods, so it is important to show users that they can get value from owning a smart device throughout the day, not just during workouts. This is especially true for sleep. Most users who wore their device all day had no sleep data. For those users, upgrading to a smart device that tracks sleep can provide added value without changing their routine.

2. Bellabeat Time is not just an activity tracker, but a stylish accessory.

- The highest levels of activity and intensity are recorded on Saturday afternoon and surrounding typical working hours (before, during lunch, and after). Show potential customers that their Bellabeat Time is just equally at home when running a 10k as when getting drinks with friends after, and that it can transition from the office to the gym as fast as they can.

3. Bellabeat Time can improve your health!

- Smart device users, and Americans as a whole, tend to be overweight and not get enough sleep. Bellabeat Time can offer insights into aspects of the wearers life that may be harming their mental or physical health. Having a device that can make potential customers aware of these problem areas can be the first step to a healthier, happier life.

7. Additional data that could be used to enhance findings

For future analysis collecting the following data can lead to new insights, as well as expand upon the insights in this report:

- Data from Bellabeat users
- Data that all comes from users of one smart device, ie. Bellabeat Time
- A larger dataset
- Demographic data for users
- More data that smart devices don't automatically collect, ie. What activities are being done at what time

8. Appendix

Sources

Primary data source: FitBit user data

FitBit Activity Index

CDC Sleep Website

Change and Analysis Logs

Google Sheets Change log

Daily Files

Created fitabase_daily_data spreadsheet with the individual sheets comprised of daily data in the "Fitabase Data 4.12.16-5.12.16" folder. Raw Data copies remain in project folder

Renamed each individual sheet to have consistent naming style and for clarity

Changed "ActivityDate" format in daily_sleep to match rest of sheets

Sorted ActivityDate column in both directions to ensure all dates fell within the correct range and they all did

Checked length of all values in Id column, in all sheets, to be sure all Id's are 10 characters. Used LEN function and conditional formatting to highlight any values != 10

Switched the dates in daily_sleep and daily_activities_merged to strings, to better work with CONCATENATE function

Concatenated "Id" and ActivityDate columns to become Unique_Id (in column A) in daily_sleep and daily_activities_merged with the following function:

=CONCATENATE(B2,C2)
* Deleted Id and ActivityDate in daily_sleep

Used Unique_Id to bring the sleep data into the daily_activities_merged sheet with the following VLOOKUPS:
=IFERROR(VLOOKUP(A2, daily_sleep!$A2 :D, 2, false), 0)
=IFERROR(VLOOKUP(A2, daily_sleep!$A2 :D, 3, false), 0)
=IFERROR(VLOOKUP(A2, daily_sleep!$A2 :D, 4, false), 0)
* the IFERROR function is to prevent #N/A outputs and replaces them with 0
* 0's must be accounted for in analysis

Copied and special pasted values only for TotalSleepRecords, TotalMinutesAsleep, and TotalTimeInBed in daily_activities_merged

Used Unique_Id column in daily_sleep to recreate Id (column A) and ActivityDate (column B) with the following functions:
=LEFT(C2, 10) # 10 being the number of characters in each Id
=RIGHT(C2, (LEN(C2) - 10)) # It takes the length of the Unique_Id and subtracts the amount of characters in each Id to account for dates having different lengths (e.g. 4/30/2016 vs. 5/1/2016)
* Unique_Id was kept in daily_activities_merged to keep a primary key in case it is needed later.

Copy and special pasted values only for Id and ActivityDate in daily_sleep and deleted Unique_Id column

Switched the ActivityDate columns in daily_sleep and daily_activities_merged, from being formatted as strings, back to dates

Added DayOfWeek column next to date to be able to compare FitBit use by day of week with:
=CHOOSE(weekday(C2), "Sun", "Mon", "Tue", "Wed", "Thu", "Fri", "Sat")

Copy and Special pasted values in DayOfWeek column

Formatted distance, time, step, and calorie measurements in daily_activities_merged for improved readability

Cleared excess whitespace in all sheets

Created pivot table to show number of data points per user (COUNTA).
* There are 33 users and the Kaggle dataset says the data is for a group of 30 users.
* After examining the Id column there are no obvious errors. Could be different devices for one user

Removed duplicate columns, columns with all 0 values, and columns with insufficient or unusable data in daily_activities_merged

Removed 79 rows with 1440 (total minutes in a day) sedentary minutes under the assumption that the device was turned on, but not being worn. Rows decreased from 940 to 861. Did not affect total number of users.

Removed 5 more rows with zero steps and 0 distance travelled under the assumption that the FitBit was not used by the user on that day. Rows decreased from 861 to 856. Did not affect total number of users.

Created TotalMinutesWorn by adding VeryActiveMinutes, FairlyActiveMinutes, LightlyActiveMinutes, and SedentaryMinutes. Will be used to gauge how often users wear their device for the entirety of a day

Copy and special pasted values for TotalMinutesWorn

Sorted TotalMinutes Worn in descending order to make sure no data was recorded for more minutes than there are in a day

Created a scatter plot comparing TotalSteps and TotalDistance to check for positive correlation and look for outliers in the dataset. Everything looked good.

Created pivot table to show COUNTA for Unique_Id. The table showed that no Id and ActivityDate were combined more than once and there are not duplicate rows.

daily_activities_merged Spreadsheet Analysis Created a pivot table with average steps taken by day of the week and overall
* Average steps of all data is 8,329
* In descending order Sat(8,979), Tue(8,927), Mon(8,488), AVERAGE(8,329) Wed(8,191), Thurs(8,185), Fri(7,821), Sun(7,669)

Edited pivot table to show COUNTA of Unique_Id. Since the data begins on a Tuesday and ends on a Thursday, I filtered out the last three days of data to have an equal number of every day of the week. Shows which days of the week had the largest number of FitBit uses.
* Friday had most uses and Sunday and Monday tied for the least uses

Edited pivot table to show Average of TotalMinutesAsleep for users with 1 sleep record for the day, to get average amount of sleep recorded overnight. 0 records would indicate the user didn't wear the device while sleeping. 2 or more would skew TotalMinutesAsleep average because of naps.
* Sunday had most sleep. Saturday had least sleep.
* Average sleep for days with sleep data was 413 minutes, indicating the average user does not get the recommended amount of sleep.

Edited pivot table to check DayOfWeek vs average of TotalMinutesWorn. Nothing stands out as useful.

Edited pivot table to look compare Id vs TotalSleepRecords and average of TotalMinutesAsleep on days with 1 sleep record.
* 24 of 33 users recorded sleep data at least once
* 13 of 33 users recorded sleep data 22 or more times, 11 of 33 had between 1 and 15 sleep records, and 9 of 33 with 0 sleep records
* Minor positive correlation between tracking sleep and average TotalMinutesAsleep

Created a chart comparing VeryActiveMinutes and TotalMinutesAsleep for users on days with 1 sleep record recorded
* The users with the most and fewest minutes of sleep tended to have the least amount of minutes asleep * The users who recorded the most VeryActiveMinutes tended to average around 400 minutes asleep

Created the same chart as above with FairlyActiveMinutes instead of VeryActiveMinutes
* Again users with the most and least sleep had among the least amount of FailyActiveMinutes

Hourly Data Spreadsheet Imported hourlyIntensities_merged into Google Sheets and renamed to hourly_intensities to match format of other sheets

Changed ActivityHour, in hourly_intensities, date format to: 9/26/2008 15:59:00
* The date time format in the raw data could not be imported into Big Query

Imported hourlySteps_merged into Google Sheets and renamed to hourly_steps to match format of other sheets

Changed ActivityHour, in hourlySteps_merged, date format to: 9/26/2008 15:59:00
* The date time format in the raw data could not be imported into Big Query

Downloaded the updated hourly_intensities and hourly_steps to the "Fitabase Working data" folder locally for uploading, cleaning, and preparing data in SQL

Google Sheets Analysis

daily_activities_merged Analysis Created a pivot table with average steps taken by day of the week and overall
* Average steps of all data is 8,329
* In descending order Sat(8,979), Tue(8,927), Mon(8,488), AVERAGE(8,329) Wed(8,191), Thurs(8,185), Fri(7,821), Sun(7,669)

Edited pivot table to show COUNTA of Unique_Id. Since the data begins on a Tuesday and ends on a Thursday, I filtered out the last three days of data to have an equal number of every day of the week. Shows which days of the week had the largest number of FitBit uses.
* Friday had most uses and Sunday and Monday tied for the least uses

Edited pivot table to show Average of TotalMinutesAsleep by the day of the week.
* Sunday had most sleep. Tuesday, Thursday, and Friday are all close at the bottom ~50 minutes less than Sunday.
* Average sleep for days with sleep data was 419 minutes, indicating the average user does not get the

recommended amount of sleep.

Edited pivot table to check DayOfWeek vs average of TotalMinutesWorn. Nothing stands out as useful.

Edited pivot table to look compare Id vs TotalSleepRecords and average of TotalMinutesAsleep on days with 1 sleep record.
* 24 of 33 users recorded sleep data at least once
* 13 of 33 users recorded sleep data 22 or more times, 11 of 33 had between 1 and 15 sleep records, and 9 of 33 with 0 sleep records
* Minor positive correlation between tracking sleep and average TotalMinutesAsleep

Created a chart comparing VeryActiveMinutes and TotalMinutesAsleep for users on days with 1 sleep record recorded
* The users with the most and fewest minutes of sleep tended to have the least amount of minutes asleep
* The users who recorded the most VeryActiveMinutes tended to average around 400 minutes asleep

Created a chart comparing FairlyActiveMinutes and TotalMinutesAsleep for users on days with 1 sleep record recorded
* Again users with the most and least sleep had among the least amount of FairlyActiveMinutes

Created 4 columns to measure PercentMoreActive (VeryActiveMinutes + FairlyActiveMinutes divided by TotalMinutesWorn) , PercentLessActive (LightlyActiveMinutes + SedentaryMinutes divided by TotalMinutesWorn), PercentVeryActive (VeryActiveMinutes divided by TotalMinutesWorn), and PercentSedentary (SedentaryMinutes divided by TotalMinutesWorn)

Copy and Special Pasted values only for the four columns and then formatted them as percents

Created pivot table to compare TotalMinutesWorn to levels of activity while wearing FitBit
* Users who wore their FitBit between half and three quarters of the day had the highest percent of MoreActive time. Over 30% more time than the next closest quartile.
* When compared to which users had sleep records, users in the half to three quarters quartile for TotalMinutesWorn recorded 81% of all sleep records in the study
* Users who wore their device between 1/4 and 1/2 a day recorded the lowest PercentMoreActive percentage. Less than half the average for the whole dataset.
* Users wearing their device for less than 1/4 of the day recorded the lowest percentage for PercentSedentary, while being average for PercentMoreActive.

hourly_data_merged Analysis Instances of activity by hour of day top 5
1. 12pm-1pm
2. 7pm-8pm
3. 5pm-6pm
4. 6pm-7pm
5. 1pm-2pm

1am-6am has by far the least activity recorded

Highest average intensity sorted by hour and day
Sunday- 10am, 5pm, 2pm, 7pm, 6pm
Monday- 5-8pm
Tuesday- 5am, 5-7pm, 12pm

Wednesday- 5-8pm
Thursday- 5am, 4-8pm
Friday- 6-8pm, 5-7am
Saturday- 11am-3pm, 1-2pm highest

Hours of week with highest intensity
1. Saturday 1-2pm
2. Tuesday 5-6am
3. Monday 6-7pm
4. Wednesday 5-6pm
5. Saturday 2-3pm

Most Average Steps sorted by hour and day
1. Saturday 1-2pm
2. Saturday 2-3pm
3. Wednesday 6-7pm
4. Wednesday 5-6pm
5. Monday 6-7pm
6. Saturday 11am-12pm
7. Sunday 10-11am
8. Saturday 12-1pm

SQL Hourly Analysis Queries SQL Queries: Hourly Data

/* This query joins the hourly_intensities and hourly_steps tables by concatenating the Id and ActivityHour to make a UniqueId column used for the join */

```
SELECT
hourly_intensities.Id,
hourly_intensities.ActivityHour,
TotalIntensity,
AverageIntensity,
StepTotal
FROM
fitness_tracker_data.hourly_intensities AS hourly_intensities
FULL OUTER JOIN
fitness_tracker_data.hourly_steps AS hourly_steps ON hourly_intensities.Id = hourly_steps.Id AND
hourly_intensities.ActivityHour = hourly_steps.ActivityHour
ORDER BY
hourly_intensities.Id, hourly_intensities.ActivityHour
```

–Checking that all AverageIntensity values fall in expected range

```
SELECT
MAX(AverageIntensity) AS max_instensity,
MIN(AverageIntensity) AS min_intensity
FROM
fitness_tracker_data.hourly_intensities
```

– Looking at highest AverageIntensity values in the hourly_intensities data

```
SELECT
AverageIntensity
```

```
FROM
fitness_tracker_data.hourly_intensities
ORDER BY
AverageIntensity DESC
LIMIT
10
```

– Looking at the highest StepTotal's in the data

```
SELECT
StepTotal
FROM
fitness_tracker_data.hourly_steps
ORDER BY
StepTotal DESC
LIMIT
25
```

/* This query joins the hourly_intensities and hourly_steps tables by concatenating the Id and ActivityHour to make a UniqueId column used for the join */

```
SELECT
hourly_intensities.Id,
hourly_intensities.ActivityHour,
TotalIntensity,
AverageIntensity,
StepTotal
FROM
fitness_tracker_data.hourly_intensities AS hourly_intensities
FULL OUTER JOIN
fitness_tracker_data.hourly_steps AS hourly_steps ON hourly_intensities.Id = hourly_steps.Id AND
hourly_intensities.ActivityHour = hourly_steps.ActivityHour
ORDER BY
hourly_intensities.Id, hourly_intensities.ActivityHour
```

– Compares AverageIntensity and StepTotal by the hour of day across all participants – Also calls highest and lowest AverageIntensity and StepTotal

```
SELECT
EXTRACT(TIME FROM ActivityHour) AS HourOfDay,
MAX(AverageIntensity) AS HighestIntensity,
MIN(AverageIntensity) AS LowestIntensity,
MAX(StepTotal) AS HighestSteps,
MIN(StepTotal) AS LowestSteps,
AVG(AverageIntensity) AS AverageIntensityByHour,
AVG(StepTotal) AS AverageStepsByHour
FROM
fitness_tracker_data.hourly_data_merged
GROUP BY
HourOfDay
```

– Finds all instances where a subject had no recorded activity for an entire day

```
SELECT
```

```
Id,
EXTRACT(DAYOFYEAR FROM ActivityHour) AS Day,
AVG(TotalIntensity) AS TotalIntensityPerDay,
AVG(AverageIntensity) AS AverageIntensityPerDay,
AVG(StepTotal) AS StepTotalPerDay,


FROM
fitness_tracker_data.hourly_data_merged
WHERE
TotalIntensity = 0 AND AverageIntensity = 0 AND StepTotal = 0
GROUP BY
Day, Id


– Queries the hours of the day that had the most instances of any activity across all days and users

SELECT
EXTRACT(HOUR FROM ActivityHour) AS HourOfDay,
COUNT(TotalIntensity) AS IntensityCount
FROM
fitness_tracker_data.hourly_data_merged
WHERE
TotalIntensity > 0
GROUP BY
HourOfDay
ORDER BY
IntensityCount DESC


– Queries the hours of the day that had the most instances of any activity – Further seperates by day of week

SELECT
EXTRACT(DAYOFWEEK FROM ActivityHOur) AS DayOfWeek,
EXTRACT(HOUR FROM ActivityHour) AS HourOfDay,
COUNT(TotalIntensity) AS IntensityCount
FROM
fitness_tracker_data.hourly_data_merged
WHERE
TotalIntensity > 0
GROUP BY
HourOfDay, DayOfWeek
ORDER BY
IntensityCount DESC


/* Queries the hours of week for each week day and returns in order the hours throughout with highest
average intensity, where any level of intensity was recorded. This sorts by day of week then by hour to show
which hours are most active, for each individual day */

SELECT
EXTRACT(DAYOFWEEK FROM ActivityHOur) AS DayOfWeek,
EXTRACT(HOUR FROM ActivityHour) AS HourOfDay,
AVG(AverageIntensity) AS IntensityCount
FROM
fitness_tracker_data.hourly_data_merged
WHERE
```

```
AverageIntensity > 0
GROUP BY
HourOfDay, DayOfWeek
ORDER BY
DayOfWeek, IntensityCount DESC
```

/* Queries the hours of week where the most steps were taken on average. Only hours when any steps are taken are accounted for */

```
SELECT
EXTRACT(DAYOFWEEK FROM ActivityHOur) AS DayOfWeek,
EXTRACT(HOUR FROM ActivityHour) AS HourOfDay,
AVG(StepTotal) AS StepCount
FROM
fitness_tracker_data.hourly_data_merged
WHERE
StepTotal > 0
GROUP BY
HourOfDay, DayOfWeek
ORDER BY
StepCount DESC
```

/* Checking to see if there are any hours in the week with high intensity and low steps, signaling common times for activities other than walking or running */

```
SELECT
EXTRACT(DAYOFWEEK FROM ActivityHOur) AS DayOfWeek,
EXTRACT(HOUR FROM ActivityHour) AS HourOfDay,
AVG(StepTotal) AS StepCount,
AVG(AverageIntensity) AS IntensityAverage
FROM
fitness_tracker_data.hourly_data_merged
WHERE  StepTotal > 0
GROUP BY
HourOfDay, DayOfWeek
ORDER BY
StepCount, IntensityAverage DESC
```