***What are some trends in smart device usage?***

* *What are the measurable features of the Fitbit data?*

Steps/day/hour/minute

Distance/day/hour/minute

Activity Levels/minute

Intensity Levels/minute

Calories burned/day/hour/minute

METs/minute

Minutes of sleep/restlessness/awake both per day and per minute

Heartrate/second

Weight tracking

* *How are they using their devices?*

They’re “using” (perhaps complacently) the features that collect data automatically more readily than the features that require manual interaction/input, such as weight logging.

* *What are they measuring themselves doing (or being measured doing automatically)?*

They’re measuring how many steps they take per day/hour.

They’re measuring distance traveled per day/hour.

They’re measuring their travelled distance at different levels of activity.

They’re measuring how many calories they are burning per day/hour.

They’re predominately **NOT** measuring their weight.

They are **PARTLY** measuring their quality of sleep. It is unclear from this study whether users are actively monitoring their sleep habits.

They are **NOT** using the device to keep track of ovulation or menstruation because it is not a feature of Fitbit products.

Fitbit does offer a stress tracking feature on their devices, however there was no data included from the test pool that applied to that feature. Therefore, no insight can be gained from this case study for the purpose of benefiting Bellabeat customers or marketing strategies.

***How could these trends be applied to Bellabeat customers?***

* Bellabeat could add a weight tracking (or other fitness goal tracking) function to the App.
* Use the Bellabeat App to inform users about the importance of proper sleep quality.
* The average user is most likely to not focus on activity intensity or higher-performance metrics such as MET’s. Use the Bellabeat App to inform users about which metrics are important or how they could be better used to reach fitness goals.
* The App could issue reminders to users about their fitness goals, as well as any progress that they are making. Make the user feel like Bellabeat is their fitness “partner”. (Steps or distance/day/hour – cumulative steps/distance, weight loss/gain, water consumption, calorie tracking/intake)

***How could these trends help influence the Bellabeat marketing strategy?***

* The Bellabeat leaf, combined with the Bellabeat app, already covers most of the health/fitness-related functionality of any wearable Fitbit product.
* Fitbit’s primary device focus is on smartwatches - Bellabeat’s focus is on discrete and stylish wearable trackers that resemble jewelry. Easier to “set and forget”, as well as MUCH longer battery life.
* Make the user feel like Bellabeat is their fitness “partner” through interactions with the Bellabeat App. Reminders, encouragement with goal setting and achievement, etc.
* Make data collection as automated as possible (if it’s not already) so that the user requires minimal intervention other than observing results and insights.

**For the purposes of this study, I don’t actually care about the fitness activity of the users. I care about how the users interact with their devices.**

**From a marketing standpoint there should be a goal to encourage users to be more active and healthier through the use of their devices because this would potentially lead to more sales and higher revenues.**

**The overall goal is to profile the customers and their level of engagement with their devices more so than their fitness activities.**

Counting steps appears to be the layman’s way of measuring fitness/activity levels.

Does a higher average step count coincide with longer device wear times?

* At first glance of the figures, the answer appears to be ‘No’.

Does a higher average step count coincide with higher rates of overnight device wearing? Weight tracking?

* There does not appear to be any significant correlation between either how long someone wears their fitness tracker, or the average of how many steps they take in a day, against the number of weight records they log.
* Perhaps these results are both positive and negative from a marketing perspective. Negative because there is a lack of correlation which means a lack of a connection between the features. You can’t promote one to automatically boost the use of the other. However, it might be positive because it could afford the opportunity to create and market a new product that could generate additional revenue for the company – and could further aid the users in achieving their fitness goals.

What percentage of users are wearing their devices all day?

What percentage of users are NOT wearing their devices at all?

* What daily percentage wear time would be a suitable cutoff?
* What would be the minimum acceptable device wear time for consideration?

***This requires an alteration to the data cleaning process to NOT omit records showing low device usage. That data is important to answering these questions. The only data that potentially needs to be removed/altered is data that is deemed to be corrupted or “bad”/inaccurate, or duplicate records.***

What does an average hourly user device wear time distribution look like?

What day of the week do users wear their devices most? Least?

What percentage of users are wearing their devices while they sleep?

What percentage of users are logging their weight?

What percentage of users are logging their weight automatically vs manually reporting?

What percentage of weight logs are automatically vs manually reported?

What are the feature participation rates for the user pool for different trackable functions?

* Weight logging
* Sleep tracking
* Step counting

What granularity of time-based data is going to be the most useful?

* Day/hour/second

Is the time-based data granularity going to be circumstantial or situationally dependent on the data table context? (i.e., steps/day vs steps/hour vs steps/minute)

Which tables are going to matter the most?

* Daily\_activity
  + Has the most complete picture with Id’s, dates, step counts, distance, activity level for distance and time, calories burned.
  + Has the data that is automatically collected.
  + This table could show how many people only wear their devices while they are being active.
* Daily\_sleep
  + Has data that is automatically collected.
  + This table could show how many people wear their devices during the night and find value in learning about their sleep quality.
* Weight\_log
  + Has data that is both automatically collected and manually reported.
  + This table could show the potential success of linking the App to a Bluetooth scale.
* Seconds\_heartrate
  + Has data that is automatically collected.
  + This table could show or verify how long per day users are wearing their devices.

**Notes from cleaning to include in analysis**

There are 79 records in the daily\_activity table that show being sedentary for 1440 minutes (24 hours). Perhaps want to check how many a zero across the board – calories, as an example.

A day with 0 calorie burn in the daily\_activity table would probably mean that the user did not wear their device that day and be a better indicator than the sedentary minutes figure. There are only 4 records in the table that fit this criteria.

A mismatch between distances and step counts in the daily\_activity table.

* 25 records over 10 user Id’s

Step counts > 0 with distances, but with no active minutes – There are 7 records in the daily\_activity table with steps > 0, 0’s for all activity distances and activity minutes, sedentary for 1440 minutes, calories > 0, but where the total distance = the tracker distance.

In the data dictionary for the daily\_activity table:

* TotalDistance as “Total kilometers tracked.”
* TrackerDistance as “Total kilometers tracked by Fitbit device.”
* LoggedActivitiesDistance as “Total kilometers from logged activities.”

The data dictionary is unfortunately vague on their definition of the word, ‘log’ or ‘logged’. It is used both with auto detected records and manually entered records. In this case it is not entirely clear what these three figures represent. A problem arises when there is a discrepancy between these figures due to ambiguity of data collection methods. This limits the extent and value of the insight that can be gained from an analysis of this data.

**What’s the story I want to tell with the data?**

Some users wear it all day and some users don’t.

More users wear it for a full day than those who wear it for a partial day. ***Is this true?***

* Why might a user choose full day wear?
  + Collect more data.
  + Stay connected to their phone (Fitbit is also a smartwatch).
* Why might a user choose partial day wear?
  + Not want to wear it to sleep.
  + Might be inconvenient/inappropriate for type of dress or outfit.