Objective

- Understand how consumers are using their smart devices.
- Provide recommendations for Bellabeat's marketing strategy.

Who

- Present to Bellabeat executive team which includes Sando Mur who is a **mathematician and Bellabeat's cofounder**.

Products

- **Bellabeat app**: track activity, sleep, stress, menstrual cycle, and mindfulness habits.
- **Leaf**: a bracelet that tracks activity, sleep, and stress on Bellabeat app.
- **Time**: tracks activity, sleep, and stress. Connect to Bellabeat app to provide insights into your **daily wellness**.
- **Spring**: water bottle to track daily water intake. Connect to Bellabeat app to **track hydration levels**.
- **Bellabeat membership**: Access to fully **personalized guidance** on nutrition, activity, sleep, health and beauty, and mindfulness based on their lifestyle and goals.

Guidance Questions

- What are some trends in smart device usage?
- How could these trends apply to Bellabeat customers?
- How could these trends help influence Bellabeat marketing strategy?

Ask

- What is the problem you are trying to solve?
- How can your insights drive business decisions?

Clear summary of the business task

The goal of this project is to analyze smart device fitness data to gain insight into how consumers are using their smart devices. We will focus on one of Bellabeat's products called *Leaf* because the data tracked by this wearable is similar to the smart device data that we will be using to perform our analysis. After analyzing the data, we will present our analysis on consumer trends to the Bellabeat executive team along with high-level recommendations for Bellabeat's marketing strategy.

Data Source

The public dataset that we will be working on is *FitBit Fitness Tracker Data* which is made by a data scientist working in the healthcare sector. This dataset is *CC0: Public Domain* licensed and was generated by respondents to a distributed survey via Amazon Mechanical Turk between 3 - 5 December 2016. Thirty eligible Fitbit users consented to the submission of personal tracker data, including minute-level output for physical activity, heart rate, and sleep monitoring. It includes information about daily activity, steps, and heart rate that can be used to explore users' habits.

Data Integrity Verification

Most of the datasets are sorted in long format except *minuteCaloriesWide_merged.csv*, *minuteIntensitiesWide_merged.csv*, and *minuteStepsWide_merged.csv*. Evident findings have shown that the data summary was not accurately reported. After data exploration, we have found out that there are 33 *FitBit* users instead of 30 who have submitted their personal tracker data. The dataset does not have clear descriptions of all the column names and there is no clear identification of the columns in units. For example, the *TotalDistance*, *Calories*, *TotalIntensity*, *METs*, *TotalTimeInBed*, and *Fat* are not labeled in units. Unable to derive the meaning of the *LoggedActivitiesDistance* column in *dailyActivity_merged.csv* dataset. Only one dataset called *weightLogInfo_merged.csv* has missing values. 65 out of 67 observations have missing values in the *Fat* column. It seems that users do not use the wearable to track body fat.

Datasets such as *dailyActivity_merged.csv*, *dailyCalories_merged.csv*, *dailyIntensities_merged.csv*, *dailySteps_merged.csv* have 940 observations. Datasets such as *hourlyCalories_merged.csv*, *hourlyIntensities_merged.csv*, and *hourlySteps_merged.csv* have 22099 observations. Datasets such as *minuteCaloriesNarrow_merged.csv*, *minuteIntensitiesNarrow_merged.csv*, *minuteMETsNarrow_merged.csv*, and *minuteStepsNarrow_merged.csv* have 1325580 observations.

Description of all data sources used

The following descriptions of the columns are based on my own interpretations according to the column names indicated.

Daily Activity (dailyActivity_merged.csv)

Column Name	Description
Id	A unique ID that identifies the Fitbit user
ActivityDate	Date that activity data was recorded
TotalSteps	Total steps recorded on that date

TotalDistance	Total distance in kilometers clocked on that date
TrackerDistance	Actual distance in kilometers clocked by the tracker
LoggedActivitiesDistance	?
VeryActiveDistance	Distance in kilometers clocked for very active activity
ModeratelyActiveDistance	Distance in kilometers clocked for moderately active activity
LightActiveDistance	Distance in kilometers clocked for light active activity
SedentaryActiveDistance	Distance in kilometers clocked for sedentary active activity
VeryActiveMinutes	Total duration in minutes clocked for very active activity
FairlyActiveMinutes	Total duration in minutes clocked for fairly active activity
LightActiveMinutes	Total duration in minutes clocked for light active activity
SedentaryMinutes	Total duration in minutes clocked for sedentary active activity
Calories	Total calories clocked for that date

Daily Calories (dailyCalories_merged.csv)

Column Name	Description
Id	A unique ID that identifies the Fitbit user
ActivityDay	Date that activity data was recorded
Calories	Total calories clocked for that date

Daily Intensity (dailyIntensities_merged.csv)

Column Name	Description
Id	A unique ID that identifies the Fitbit user
ActivityDay	Date that activity data was recorded
SedentaryMinutes	Total duration in minutes clocked for sedentary active activity
LightlyActiveMinutes	Total duration in minutes clocked for light active activity
FairlyActiveMinutes	Total duration in minutes clocked for fairly active activity

VeryActiveMinutes	Total duration in minutes clocked for very active activity
SedentaryActiveDistance	Distance in kilometers clocked for sedentary active activity
LightActiveDistance	Distance in kilometers clocked for light active activity
ModeratelyActiveDistance	Distance in kilometers clocked for moderately active activity
VeryActiveDistance	Distance in kilometers clocked for very active activity

Daily Steps (dailySteps_merged.csv)

Column Name	Description
Id	A unique ID that identifies the Fitbit user
ActivityDay	Date that activity data was recorded
StepTotal	Total steps recorded on that date

Heart Rate (heartrate_seconds_merged.csv)

Column Name	Description
Id	A unique ID that identifies the Fitbit user
Time	Date and time that heart rate data was recorded
Value	Heart rate

Hourly Calories (hourlyCalories_merged.csv)

Column Name	Description
Id	A unique ID that identifies the Fitbit user
ActivityHour	Date and time that calories was recorded
Calories	Amount of calories

Hourly Intensity (hourlyIntensities_merged.csv)

Column Name	Description
Id	A unique ID that identifies the Fitbit user
ActivityHour	Date and time that intensity was recorded
TotalIntensity	Total intensity clocked per hour
AverageIntensity	Average intensity per minute in one hour

Hourly Steps (hourlySteps_merged.csv)

Column Name	Description
Id	A unique ID that identifies the Fitbit user
ActivityHour	Date and time that number of steps was recorded
StepTotal	Number of steps taken

Calories by minutes in long format (minuteCaloriesNarrow_merged.csv)

Column Name	Description
Id	A unique ID that identifies the Fitbit user
ActivityMinute	Date and time that calories was recorded
Calories	Amount of calories

Calories by minutes in wide format (minuteCaloriesWide_merged.csv)

Column Name	Description
Id	A unique ID that identifies the Fitbit user
ActivityHour	Date and time that calories was recorded
Calories00 to 59	Amount of calories recorded per minute

Intensity by minutes in long format (minuteIntensitiesNarrow_merged.csv)

Column Name	Description
Id	A unique ID that identifies the Fitbit user
ActivityMinute	Date and time that intensity was recorded
Intensity	Minutes

Intensity by minutes in wide format (minuteIntensitiesWide_merged.csv)

Column Name	Description
Id	A unique ID that identifies the Fitbit user
ActivityHour	Date and time that intensity was recorded
Intensity00 to 59	Intensity recorded per minute

METs by minutes in long format (minuteMETsNarrow_merged.csv)

Column Name	Description
Id	A unique ID that identifies the Fitbit user
ActivityMinute	Date and time that intensity was recorded
METs	Metabolic equivalents

Amount of sleep in minutes (minuteSleep_merged.csv)

Column Name	Description
Id	A unique ID that identifies the Fitbit user
date	Date and time that sleep was recorded
value	Deep Sleep = 1 Light Sleep = 2 Rapid Eye Movement Sleep = 3
logId	A unique log ID that identifies the log record

Number of steps by minutes in long format (minuteStepsNarrow_merged.csv)

Column Name	Description
Id	A unique ID that identifies the Fitbit user
ActivityMinute	Date and time that number of steps was recorded
Steps	Number of steps taken per minute

Number of steps by minutes in wide format (minuteStepsWide_merged.csv)

Column Name	Description
Id	A unique ID that identifies the Fitbit user
ActivityHour	Date and time that steps was recorded
Steps00 to 59	Number of steps recorded per minute

Number of sleeps by day (sleepDay_merged.csv)

Column Name	Description
Id	A unique ID that identifies the Fitbit user
SleepDay	Date and time that sleep was recorded
TotalSleepRecords	Number of sleep per day
TotalMinutesAsleep	Amount of time spend sleeping per day
TotalTimeInBed	Amount of time spend on bed per day

Weight Log (weightLogInfo_merged.csv)

Column Name	Description
Id	A unique ID that identifies the Fitbit user
Date	Date and time that weight was recorded
WeightKg	Weight in kilograms

WeightPounds	Weight in pounds
Fat	Body fat
BMI	Body Mass Index
IsManualReport	True if weight is recorded manually
LogId	A unique log ID that identifies the log record

Analysis on trends and relationships

1. Fitbit users tend to use the smart device on weekdays especially Tuesday to Thursday. The distribution of average time spent by all users for sedentary activity, light activity, fairly active activity, and very active activity is 81%, 16%, 1%, and 2% of total time, respectively.
2. Among the 33 Fitbit users, 21 of them use the smart device daily for an average of 20 hours. 11 out of 21 daily users used more than 20 hours a day on average.
3. Fitbit users who use the smart device frequently tend to be very active. 11 Fitbit users spent more than 2% of the time on very active activity with an average of 53 minutes per day. Majority of such users used the smart device for more than 25 days.
4. More than 70% of Fitbit users track their sleeping habits to detect sleep abnormality. Among the 24 Fitbit users who track their sleeping habits, 12 users have more than one sleeping record per day within a month. One particular user id '***4445114986***' had 11 occurrences of more than 1 sleeping record within 28 days.
5. More than 90% of the total 940 daily records have tracker distance. This shows that users are generally satisfied with the tracking accuracy of the smart device. In fact, the smart device has good tracking accuracy of 98%.
6. Daily users tend to keep track of their body weights compared to non-daily users. Among the 7 daily users, 3 of them have lost 1 to 2 kg in a month by doing very active activity for at least 22 days with at least 14 minutes per day on average.

Recommendations

1. Our findings have shown that frequent users tend to put on the smart device for around 20 hours a day. Therefore, I recommend the *Leaf* to be lightweight, water resistant, and have a long battery life span.

2. We can make the *Leaf* to have the necessary features for sports since 50% of the daily users wear it for high intensity activities.
3. Our *Leaf* must have a distance tracking accuracy of close to 100% so that we can attract consumers, who are very active in running long distances, to purchase our products.
4. Our *Leaf* must have the capability to track users' sleeping patterns and provide insights on their sleeping habits (i.e. frequency of waking up in the middle of sleep).
5. We can promote *Leaf* to consumers who are very active in sports and care about their daily wellness because they are more likely to be motivated to use the product and achieve their health goals.