Bellabeat Case Study

Scenario

Bellabeat is a small company that manufactures high tech products with a focus on women's health. They offer a variety of wearable tech that monitors health and activity, a smart bottle that tracks hydration and a subscription service that offer health and wellness programs customized for each user. Smart tech links to an app that acts as a hub for the user to explore their health and wellness goals.

Urška Sršen, cofounder and Chief Creative Officer of Bellabeat, believes that analyzing smart device fitness data could help unlock new growth opportunities for the company. Sršen knows that an analysis of Bellabeat's available consumer data would reveal more opportunities for growth. She has asked the marketing analytics team to focus on a Bellabeat product and analyze smart device usage data in order to gain insight into how people are already using their smart devices. Then, using this information, she would like high-level recommendations for how these trends can inform Bellabeat marketing strategy.

Business Task

Identify the following:

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

Finding answers to these questions will help Bellabeat create marketing strategies to reach more women and hopefully help them reach their health and wellness goals.

Key Stakeholders:

- Urška Sršen, cofounder and Chief Creative Officer of Bellabeat
- Sando Mur: Mathematician and Bellabeat's co-founder; key member of the Bellabeat executive team

Prepare

We will be using public data made available by FitBit and hosted on www.kaggle.com. The data is organized into tables by metric and duration, ex. heartrate_seconds_merged, hourlyCalories_merged, dailyActivity_merged. The data are in long format, there is a date and time for each participants ID number for everything that is tracked.

The data are nearly six years old and were collected via an opt-in scheme by Amazon Mechanical Turk of 30 eligible FitBit users for one month. Since they were collected by a third party and hosted by another third party the data's originality and reliability are not strong. They are well organized and well documented, published under the CC0: Public Domain License. We are not made aware of gender or age statistics of the volunteers, nor what constitutes an "eligible" user, or what units the active distance is tracked in.

Choosing the Data

We will use the following files to perform analysis on. Daily levels of activity should be useful in determining trends for wearable fitness tech.

- weightLogInfo_merged
- sleepDay_merged
- dailyActivity_merged
- dailyCalories_merged
- dailyIntensities_merged
- dailySteps_merged

Processing and Cleaning the Data

Google Sheets and BigQuery will be used to clean and process the data for analysis. daily_sleep-Sleep and BMI-weightLogInfo are uploaded to Google Sheets to clean. “SleepDay” and “Date” were not formatted in a way that would be helpful. The timestamp was removed entirely the “Date” as we are focusing on daily activities only. The files were then uploaded to BigQuery for SQL manipulation.

The following SQL queries were run to asses the number of participants tracking a daily metric.

```
SELECT COUNT(DISTINCT(Id)) FROM `bellabeats.daily_activity`
SELECT COUNT(DISTINCT(Id)) FROM `bellabeats.daily_calories`
SELECT COUNT(DISTINCT(Id)) FROM `bellabeats.daily_intensities`
SELECT COUNT(DISTINCT(Id)) FROM `bellabeats.daily_sleep`
SELECT COUNT(DISTINCT(Id)) FROM `bellabeats.daily_steps`
SELECT COUNT(DISTINCT(Id)) FROM `bellabeats.BMI`
```

Results:

1. 33 users tracked activity, calories, steps, and intensities.
2. 24 users tracked sleep
3. 8 users tracked weight

Discrepancies in the data appear with 33 total ID's with the information listing 30, and not all 33 users tracked sleep or weight.

This query will asses the number of manually tracked BMI entries and automatically tracked BMI entries by user ID.

```
SELECT
IsManualReport,
COUNT(*)
FROM
`bellabeats.BMI`
GROUP BY
IsManualReport
```

Results

1. Manual - 41
2. Automatic - 26

Analysis

We will continue to use SQL to identify trends in the data. The following query returns user ID, date, calories, and active distance and active minutes if they total more than zero.

```
SELECT
Id,ActivityDate,Calories,TotalSteps,TotalDistance,TrackerDistance,LoggedActivitiesDistance,
(VeryActiveDistance+ModeratelyActiveDistance+LightActiveDistance) AS TotalActiveDistance, SedentaryActiveDistance,
(VeryActiveMinutes+FairlyActiveMinutes+LightlyActiveMinutes) AS TotalActiveMinutes, SedentaryMinutes
FROM
`bellabeats.daily_activity` AS a
WHERE
(VeryActiveMinutes+FairlyActiveMinutes+LightlyActiveMinutes) > 0 AND
(VeryActiveDistance+ModeratelyActiveDistance+LightActiveDistance) > 0
ORDER BY
Id
```

This query returns maximum, average, and minimum values for steps, calories, and distance grouped by user ID.

```
SELECT
DISTINCT(Id),
MAX(TotalSteps) AS max_steps,
Round(AVG(TotalSteps),2) AS avg_step,
MIN(TotalSteps) AS min_steps,
MAX(Calories) AS max_cal,
Round(AVG(Calories),2) AS avg_cal,
MIN(Calories) AS min_cal,
MAX(TotalDistance) AS max_dist,
Round(AVG(TotalDistance),2) AS avd_dist,
MIN(TotalDistance) AS min_dist
FROM
`bellabeats.daily_activity`
GROUP BY
Id
```

These query results will be saved as .csv's and uploaded to Tableau for visualization.

Visualizations

https://public.tableau.com/app/profile/posuanecat/viz/BellabeatCaseStudy_16835690623610/Bellabeat?publish=yes

All the users tracked calories and burned a minimum of 1,500 calories. Because calories are a metric we interact with everyday it makes sense that is would be tracked more than nuanced or personal metrics like sleep quality and BMI.

The next visualization demonstrates a positive relationship between calories burned and total distance moved. The more active distance a user covers the more calories they will burn. R-Squared ~.375

The final visualization also demontrates a positive relationship between calories burned and number/intensity of active minutes. The more minutes you are active will increase your calories burned, with more intense minutes burning more calories.

Conclusions

Users who are more active are likely to burn more calories, a heavily tracked metric by individuals on their fitness or maintenance programs. Similarly, the level of activity has a positive relationship on calories burned. The more minutes a user is working hard, the more calories they will burn.

This relationship is not surprising but users who track such things would be more informed of their caloric needs and would make prime candidates for Bellabeat's subscription service in addition to a smart device. The subscription service could also encourage users to track the more private aspects of their health journey like weight.

Less users track these proactively likely due to their sensitive or long-term nature. Daily calories and steps are easy to track and notice how "well" you are doing. BMI is less likely to change in the short term and requires more active participation to track and change. Around twenty-five percent of FitBit participants tracked their weight and most of the logs were manual, indicating that weight was a primary goal for them. Bellabeat could attract more customers by advertising their subscriptions and wearable tech as beneficial to a weight loss journey. The more information in a customers' hands the better equipped they will be to improve their overall wellness.

Recommendations

1. Encourage Wearable owners to move/be more active through notifications on their tech or apps with milestones for all daily metrics. Ex. "X more steps needed before a user defined time milestone"
2. Market subscription services to users who track more sensitive aspects of their health in an encouraging manner. Emphasize success stories of other users' who have subscribed. Ex. "Our subs are X% more likely to lose and keep off X amount of weight"