Bellabeat Case Study

Jeffrey Cuevas

1/17/2022

Introduction

Bellabeat is a high-tech manufacturer of health-focused products for women. The company was founded in 2013 and has quickly positioned itself as a tech-driven wellness company for women. Currently, they are a successful small company with the potential to scale and become a larger player in the global smart device market.

They have three products that communicate data to the Bellabeat App, which provides users with health data related to their physical activity, sleep, stress, menstrual cycle, and mindfulness habits. This data helps users better understand their current habits and make healthier decisions.

1. The Leaf is Bellabeat's classic wellness tracker than can be worn as a bracelet, necklace, or clip. The Leaf tracker connects to the Bellabeat app to track activity, sleep, and stress.
2. The Time wellness watch combines a classic timepiece with smart technology to track user activity, sleep, and stress. The Time watch connects to the Bellabeat app to provide insights on daily wellness.
3. The Spring is a water bottle that tracks daily water intake using smart technology to ensure that the user is appropriately hydrated throughout the day. The Spring connects to the Bellabeat app to track hydration levels.

Bellabeat also offers a subscription-based membership program for users. Membership gives users 24/7 access to fully personalized guidance on nutrition, activity, sleep, health and beauty, and mindfulness based on their lifestyle and goals.

Ask

I've been asked to analyze smart devise usage from non-Bellabeat products in order to gain insights on how consumers are using competing products. Key questions that will guide the analysis include:

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?

Key stakeholders include:

- Urška Sršen - Bellabeat's cofounder and Chief Creative Officer.

- Sando Mur - Mathematician and Bellabeat's cofounder; key member of the Bellabeat executive team.
- Bellabeat Marketing Analytics Team - A team of data analysts responsible for collecting, analyzing, and reporting data that helps guide Bellabeat's marketing strategy.

Prepare

For this analysis, we are using a public dataset that explores 30 smart device users' daily habits. There are 18 total csv files that consist of attribute data for sleep, steps, activity, calories, intensity and others. I will focus on using the "DailyActivity" table that has variables compiled in one location.

Prior to analysis, I can identify limitations such as sample size and lack of user variables (age, height, weight, current health conditions, etc.). Data could be considered outdated being from 2016 because user habits may have changed.

Python and PySpark will be used for data cleaning and transformation. R Studio will be used for data visualization. R Markdown will be used for creating a final deliverable of results.

Process

Data was processed within a Jupyter Notebook utilizing PySpark. The Jupyter Notebook can be found in the "Bellabeat" repository on GitHub. These are the key tasks I focused on during data cleaning:

- Imported the data successfully into a data frame and previewed it.
- Printed the schema of the data frame to ensure data types were correct. In this case, they weren't correct because every field was classified as a string. I converted whole number fields to integers, decimal fields to floats, and the date field to date timestamp.
- Dropped duplicate records using the drop duplicates function (there were none).
- Used a function to count instances of null values in each field (there were none).
- Created a new column for day of the week using a date format function.
- Created a calculated field for total active minutes. Based on this new field, I created a second calculated field for total active hours.
- Printed summary statistics for total steps, total distance, total exercise hours, and calories to verify everything looked okay.
- Renamed columns for consistency and dropped unneeded columns.
- Exported data fram to a csv.

Analyze

Let's analyze summary statistics for the data:

```
##   total_steps      total_dist        calories
##  Min.   :    0   Min.   : 0.000   Min.   :   0
##  1st Qu.: 3790   1st Qu.: 2.620   1st Qu.:1828
##  Median : 7406   Median : 5.245   Median :2134
##  Mean   : 7638   Mean   : 5.490   Mean   :2304
##  3rd Qu.:10727   3rd Qu.: 7.713   3rd Qu.:2793
##  Max.   :36019   Max.   :28.030   Max.   :4900
```

Before breaking data down to subcategories, we can observe the statistics for total daily calories, total daily distance, and total daily steps. On average, users took 7,638 steps per day, walked 5.5 kilometers per day, and burned 2,304 calories per day. According to the dietary guidelines published by the U.S. Health and Human Services, the average woman should burn 2,000 to 2,800 calories depending on an individual's body variables. 2,304 calories per day falls within that range but is not representative of weight loss. According to First Quote Health, depending on age and fitness goals, females should target anywhere from 6,000 to 13,500 steps per day. We can conclude our sample users lean toward the minimum steps per day, meaning they are not very active overall. These facts allow us to infer that these sample users are not using the app for tracking fitness or weight loss goals. We can further confirm this by looking at the summary statistics for the different classifications of activity distance:

```
##  very_act_dist    moderately_act_dist light_act_dist
##  Min.   : 0.000   Min.   :0.0000      Min.   : 0.000
##  1st Qu.: 0.000   1st Qu.:0.0000      1st Qu.: 1.945
##  Median : 0.210   Median :0.2400      Median : 3.365
##  Mean   : 1.503   Mean   :0.5675      Mean   : 3.341
##  3rd Qu.: 2.053   3rd Qu.:0.8000      3rd Qu.: 4.782
##  Max.   :21.920   Max.   :6.4800      Max.   :10.710
```

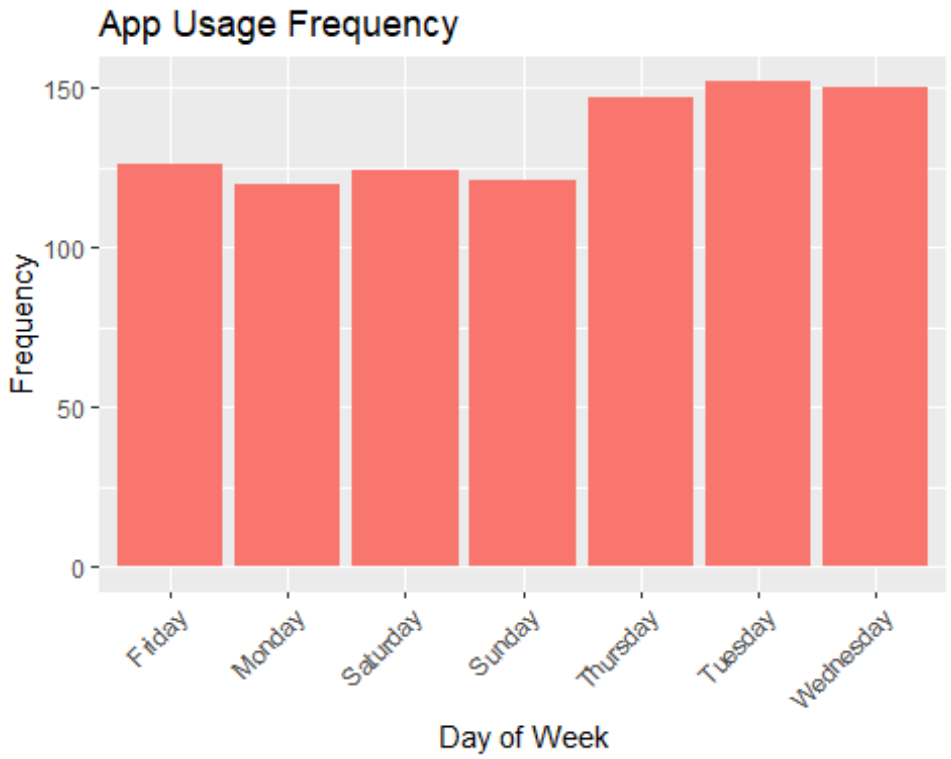
Average light active distance is the highest category for distance being logged on the app. Finally, we view the summary statistics for the different classifications of activity minutes:

```
##  very_act_mins    fairly_act_mins  light_act_mins  sedentary_mins
##  Min.   :  0.00   Min.   :  0.00   Min.   :  0.0   Min.   :   0.0
##  1st Qu.:  0.00   1st Qu.:  0.00   1st Qu.:127.0   1st Qu.: 729.8
##  Median :  4.00   Median :  6.00   Median :199.0   Median :1057.5
##  Mean   : 21.16   Mean   : 13.56   Mean   :192.8   Mean   : 991.2
##  3rd Qu.: 32.00   3rd Qu.: 19.00   3rd Qu.:264.0   3rd Qu.:1229.5
##  Max.   :210.00   Max.   :143.00   Max.   :518.0   Max.   :1440.0
```

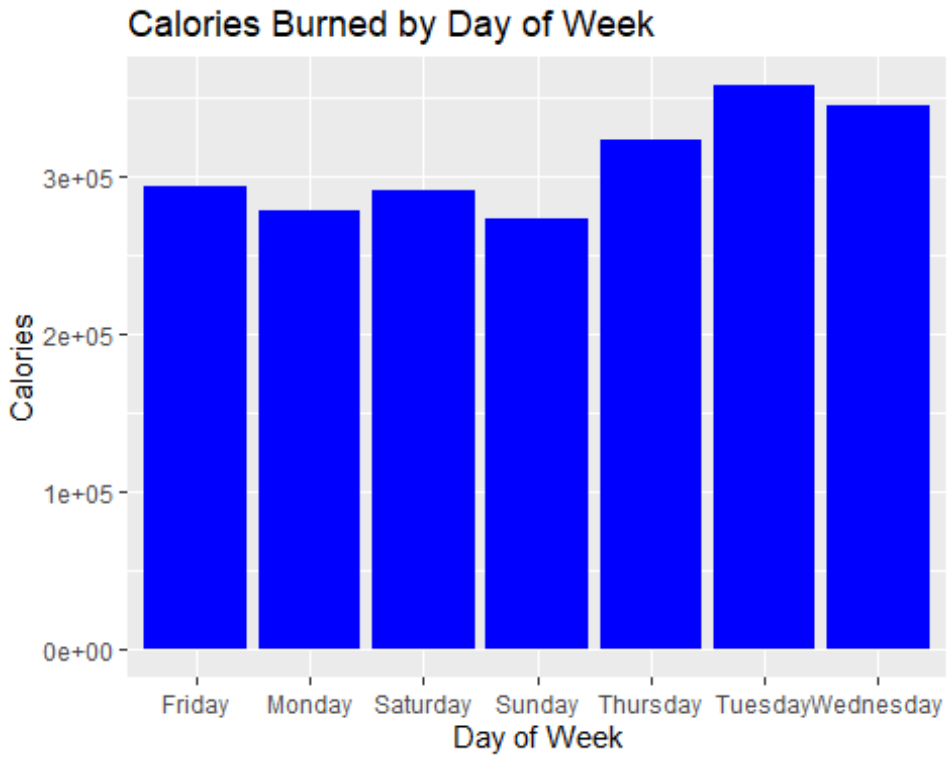
On average, users are primarily logging sedentary and light activity minutes. This further supports the theory that users are more interested in logging their normal day rather than tracking vigorous workouts.

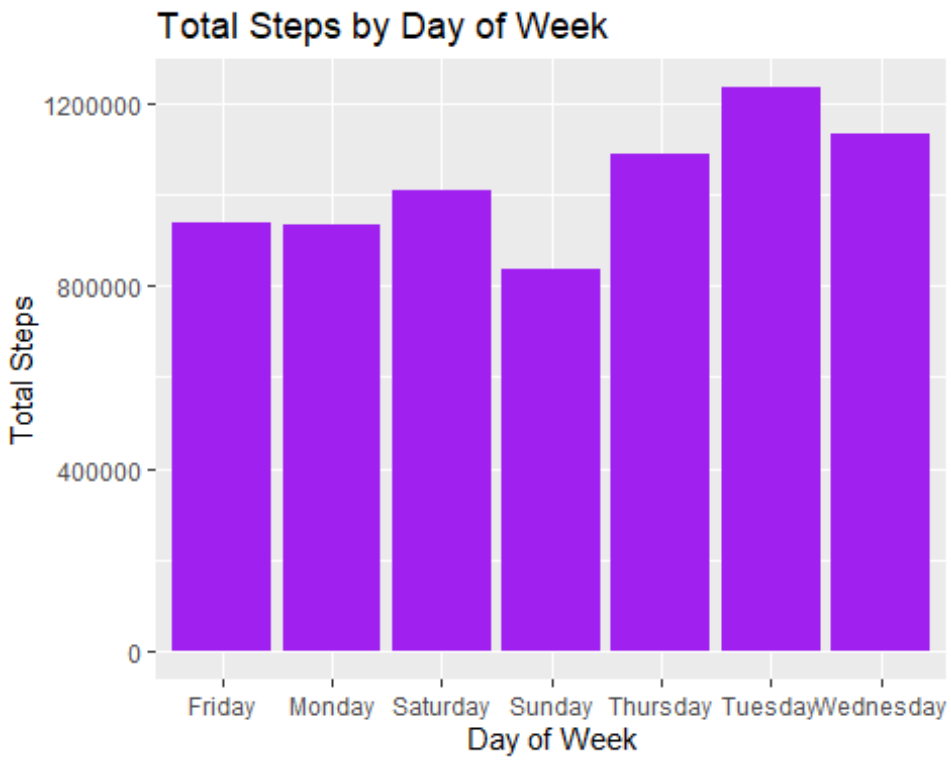
Share

First, let's look at the frequency of app usage across the week:



App usage is stronger Tuesday through Thursday and significantly falls off Friday through Monday. We can infer users care more about tracking their activity during the middle of the work week and not so much over the weekend. Let's also take a look at total calories burned throughout the week and total steps taken throughout the week to gain more insights:





We can see that both trends match the app usage frequency trend. Users care more about tracking their total steps and calories burned during the middle of the work week. Let's compare total calories burned with total steps taken:

There is a direct correlation between total steps taken and total calories burned. This could be a result of users being primarily sedentary while using the app, meaning they aren't just using it to track workout activity. This could be a crucial data insight for the Bellabeat Marketing Team.

We can also see a direct correlation between daily total calories burned and daily total steps logged on the app. The longer distance a user logs on the app, the more likely they are to burn a higher amount of calories.

Act

Our business task is to analyze smart device usage from non-Bellabeat products in order to gain insights on how consumers are using competing products. We can conclude our analysis by revisiting the key questions from our business task:

1. What are some trends in smart device usage?
   - Smart device usage is higher Tuesday through Thursday and lower Friday through Monday. This could be a result of users only wanting to track their health data during their work week routine and not so much on the weekend.
   - The majority of smart device usage comes from sedentary and light activity usage. Combined, both categories account for 97.1% of total minutes logged.
2. How could these trends apply to Bellabeat customers?
   - Competing smart device app usage is most likely similar to Bellabeat's app usage. We can infer that competing brands are attracting users on the lower spectrum of activity level based on the analysis. Bellabeat could use this insight to create some kind of fitness challenge or workout incentive program that offers rewards and discounts to Bellabeat products.
   - For example, completing a fitness challenge results in a 50% off Bellabeat subscription-based membership.

- Complete 7 consecutive daily workout challenges and receive a Bellabeat t-shirt.

3. How could these trends help influence Bellabeat marketing strategy?
   - The marketing team could use these trends to recommend different ways of utilizing Bellabeat products, especially on the weekends. This analysis has identified areas of weakness in app usage, specifically with vigorous activity and on the weekends. Creating a marketing campaign that boosts app usage on the weekends and promotes increased activity levels will benefit the company.
   - Run an optional survey for new and existing Bellabeat users to identify their primary fitness goals. Try to pinpoint user's exact intentions with tracking their health data. Currently, it seems like users only want to monitor their normal routine. Allowing them to set goals through the app could incentivize them to increase their activity level, leading to increased app usage.
   - Depending on age, weight, height, and other variables that we don't currently know, Bellabeat could run targeted social media ads to a wider population that match the demographics of our sample users.