Icon

Description automatically generated with medium confidence

Google Data Analytics Certificate: Capstone Project

**How Can a Wellness Technology Company Play It Smart?**

[**Bellabeat**](https://bellabeat.com/)

* This is an optional capstone project from the Google Data Analytics Course 8: [Capstone Project](https://www.coursera.org/learn/google-data-analytics-capstone?specialization=google-data-analytics)
* For this case study, I analyzed data from 33 Fitbit users over the course of 31-day period.
* The goal is to analyze smart device usage data in order to gain insight into how consumers use smart devices and apply this analysis to Bellabeat’s marketing strategy

**Step 1: Ask**

In this step, we define the objective and problem of our case study.

**Company Background**

* Founded by Urška Sršen and Sando Mur.
* Positioned as a high-tech company that manufacturers health-focused smart products for women.
* Bellabeat has experienced rapid growth since opening in 2013.
* Sršen started this company around the idea of informing and inspiring women worldwide about their health and fitness.
* The two founders are confident that an analysis of non-Bellabeat consumer data (ie. FitBit fitness tracker usage data) could reveal more growth opportunities for Bellbeat.

**Stakeholders**

* + **Urška Sršen**: Bellabeat's cofounder and Chief Creative Oﬃcer
  + **Sando Mur**: Mathematician and Bellabeat's cofounder; a 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

**Products Included**

* **Leaf:** Stylish and feminine wearable technology with a 6-month battery life. Can be worn as a bracelet, necklace or clip which makes it easier to consistently wear the device. The device is made of wood-composite material and hypoallergenic stainless-steel slip. Leaf is affordable ($89-149)
* **Bellabeat App:** Provides users with health data related to their activity, sleep, stress, menstrual cycle and mindfulness activities. Connects to their line of smart wellness products.
  + **Spring:** Water bottle that tracts daily water intake and connects to the Bellabeat app.
* **Bellabeat Membership:** Membership gives users 24/7 access to fully personalized guidance on nutrition, health and beauty.

**What are my project deliverables?**

* A clear summary of business task
* A description of all data sources used
* Documentation of any cleaning or manipulation of data
* A summary of analysis
* Supporting visualizations and key findings
* Marketing recommendations for a selected product based on the analysis

**Step 2: Prepare**

In the prepare phase we identify the data being used.

Data is publicly available on [Kaggle: FitBit Fitness Tracker Data](https://www.kaggle.com/arashnic/fitbit).

* There are 18 datasets available
* These datasets date from 03/12/2016 – 05/12/216.
* These datasets tracked 33 users on:
  + Daily Activity
  + Daily Sleep
  + Daily Steps
  + Heartrate
  + Weight and BMI

In order to get accurate analysis, validate and make sure the dataset does not include any bias, incorrect data and duplicates I am using **Microsoft SQL Server Management Studio**.

**Data Selection**  
The tables I am using are:

* [dbo].[dailyActivity\_merged$]
* [dbo].[dailySteps\_merged$]
* [dbo].[heartrate\_seconds\_merged$]
* [dbo].[hourlySteps\_merged$]
* [dbo].[sleepDay\_merged$]
* [dbo].[weightLogInfo\_merged$]

\*Note that Daily Steps and Daily Intensities are included in Daily Activities.

***My SQL queries are available on*** [***Github***](https://raw.githubusercontent.com/hillaryfuselier/Google-Capstone-Case-Studty-Bellabeat/main/README.md)

**Data Limitations**Insignificant Sample Size

* User sample: There are only 33 users participated in this dataset which makes the result could be far from statistical significance to represent the original population, especially for some functions that are utilized by even less than 33 users.
* Time frame: Data was collected from a period of only 31 days which could largely decrease the possibility of finding some significant insights.

A good data source is ROCCC which stands for Reliable, Original, Comprehensive, Current and Cited. **Is this data ROCCC?**

* Reliable – Low
* Original – Low
* Comprehensive – Medium
* Current – Low
* Citied - Low

**Step 3: Process**

**Processing the data involved cleaning and ensuring that it is accurate, relevant and free from error by:**

* Exploring and observing data
* Check for null values
* Transform and format data
* Perform statistical analysis

**After preparing the environment and cleaning the data I observed:**

* There are NULL values
* Data has 15 columns and 940 rows.
* There are 33 unique users, instead of 30, all of whom tracked their steps.
* *ActivityDate* is wrongly classified

***My SQL queries are available on*** [***Github***](https://raw.githubusercontent.com/hillaryfuselier/Google-Capstone-Case-Studty-Bellabeat/main/README.md)

**Step 4: Analysis of Data**

**Perform calculations:**

* Count – number of rows
* Mean (average)
* Std (standard deviation)
* Min and max
* Usage rate

**Statistical findings:**

* The usage rate of wearable technology is about **91%**
* Only 21% of users took more than 10,000 steps per day.
* An average of 83% of the user’s day us made up of sedentary minutes.
* An average of 20% of the user’s day is made up of fairly active minutes.
* 27% did not record any sleep
* Users are most active between **5pm – 7pm daily**

**Step 5: Share**

**Data Visualization and Findings**

The chart below shows the usage rate by date. We can clearly see that over the course of the 31-day period, user data decreased.

* + - **Usage rate = Active Users/Total Days**
    - Active users = SUM(total\_steps) > 0
    - Total days = SUM(total\_steps) <=0

Chart, line chart

Description automatically generated

And below we are looking at the Usage Rate in a Calendar View. (The darker the color the higher the usage rate.)

Chart, treemap chart

Description automatically generated

The chart below shows how each user is spending most of their time. It shows that users are mostly in Sedentary Activity compared to Very Active, Fairly Active and lightly active.

Chart, bar chart

Description automatically generated

**Now let’s look at user activity as a whole**

This chart below shows that **97% of users time are not related to performance workout.** Users spend majority of their time in sedentary activities

Chart, pie chart

Description automatically generated

This chart below shows that time of the day that users are most active by measuring their average intensity per hour. **Most users were more active between the hours of 5pm – 7pm**

Chart

Description automatically generated

The chart below shows the Average Calories burned by the day of the week. We can conclude that **there’s not much correlation between calories and the day of the week**.

Chart, bar chart

Description automatically generated

The chart below shows the Average total steps taken by the day of the week. We can conclude that **users are recording more steps on Tuesday and Saturday and the least amount on Sunday and Wednesday.**

**Chart, bar chart

Description automatically generated**

***Click the image below to view my Tableau Dashboard***

[Chart, treemap chart

Description automatically generated](https://public.tableau.com/app/profile/hillary.fuselier/viz/BellabeatGoogleCapstoneCaseStudy/Dashboard1?publish=yes)

**Step 6: Act**

In the final step, we provide our insights and recommendations.

**Recommendations**

* First, I would suggest focusing on “Leaf”, the wearable device. Users must be able to express themselves, **so Bellabeat should add more options for designs and related accessories**
* Efficiency and engagement. Increasing automation for tracking weight, by encouraging users to use Bluetooth scales that can connect to the Bellabeat app, collaborating with more wellness technologies that can connect to Bellabeat app to make the app **more efficient** and create brand awareness for Bellabeat.
* Since the users are spending 83% of their time in sedentary activity, we can use the Bellabeat app to **encourage women to get up and get a little more active**. Giving them incentives to track their steps, activities and calories with competitions, live workouts via social media, and community involvement events.

**Branding**

* Bellabeat’s target audience is women and as the data shows, the users are not using their wearable technologies to track extensive workouts and rarely anything fitness related so focusing on building a **fashion brand** rather than fitness can differentiate Bellabeat from other wearable technologies
* The brand image must be geared towards women and create a **community for women**. For example, embracing different types of beauty, sharing educational information on common female problems and providing a solution to those problems, being involved in women empowerment movements will inspire women and they will become a part of the community, and they are not just a customer. Community creates loyal customers.