

# Introduction

Welcome to the Bellabeat data analysis case study! In this case study, you will perform many real-world tasks of a junior data analyst. You will imagine you are working for Bellabeat, a high-tech manufacturer of health-focused products for women, and meet diﬀerent characters and team members. In order to answer the key business questions, you will follow the steps of the data analysis process: **ask**, **prepare**, **process**, **analyze**, **share**, and **act**. Along the way, the **Case Study Roadmap** tables —

including guiding questions and key tasks — will help you stay on the right path.

By the end of this lesson, you will have a portfolio-ready case study. Download the packet and reference the details of this case study anytime. Then, when you begin your job hunt, your case study will be a tangible way to demonstrate your knowledge and skills to potential employers.

**Scenario**

You are a junior data analyst working on the marketing analyst team at Bellabeat, a high-tech manufacturer of health-focused products for women. Bellabeat is a successful small company, but they have the potential to become a larger player in the

global [**smart device**](https://en.wikipedia.org/wiki/Smart_device) market. 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. You have been asked to focus on one of Bellabeat’s products and analyze smart device data to gain insight into how consumers are using their smart devices. The

insights you discover will then help guide marketing strategy for the company. You will present your analysis to the Bellabeat executive team along with your high-level recommendations for Bellabeat’s marketing strategy.

# Characters and products

## Characters

* + **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. You joined this team six months ago and have been busy learning about Bellabeat’s mission and business goals — as well as how you, as a junior data analyst, can help Bellabeat achieve them.

## Products

* + **Bellabeat app:** The Bellabeat app provides users with health data related to their activity, sleep, stress,

menstrual cycle, and mindfulness habits. This data can help users better understand their current habits and make healthy decisions. The Bellabeat app connects to their line of smart wellness products.

* + **Leaf:** Bellabeat’s classic wellness tracker can be worn as a bracelet, necklace, or clip. The Leaf tracker connects to the Bellabeat app to track activity, sleep, and stress.
  + **Time:** This wellness watch combines the timeless look of a classic timepiece with smart technology to track user activity, sleep, and stress. The Time watch connects to the Bellabeat app to provide you with insights into your daily wellness.
  + **Spring:** This is a water bottle that tracks daily water intake using smart technology to ensure that you are appropriately hydrated throughout the day. The Spring bottle connects to the Bellabeat app to track your hydration levels.
* **Bellabeat membership:** 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.

# About the company

Urška Sršen and Sando Mur founded Bellabeat, a high-tech company that manufactures health-focused smart products.

Sršen used her background as an artist to develop beautifully designed technology that informs and inspires women around the world. Collecting data on activity, sleep, stress, and reproductive health has allowed Bellabeat to empower women with knowledge about their own health and habits. Since it was founded in 2013, Bellabeat has grown rapidly and quickly positioned itself as a tech-driven wellness company for women.

By 2016, Bellabeat had opened oﬃces around the world and launched multiple products. Bellabeat products became available through a growing number of online retailers in addition to their own e-commerce channel on [**their website**](https://bellabeat.com/). The company has invested in traditional advertising media, such as radio, out-of-home billboards, print, and television, but focuses on digital marketing extensively. Bellabeat invests year-round in Google Search, maintaining active Facebook and Instagram pages, and consistently engages consumers on Twitter. Additionally, Bellabeat runs video ads on YouTube and display ads on the Google Display Network to support campaigns around key marketing dates.

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.

# Ask

Sršen asks you to analyze smart device usage data in order to gain insight into how consumers use non-Bellabeat smart

devices. She then wants you to select one Bellabeat product to apply these insights to in your presentation. These questions will guide your analysis:

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?**

You will produce a report with the following deliverables:

1. A clear summary of the business task
2. A description of all data sources used
3. Documentation of any cleaning or manipulation of data
4. A summary of your analysis
5. Supporting visualizations and key ﬁndings
6. Your top high-level content recommendations based on your analysis

**Guiding questions**

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

**Key tasks**

1. Identify the business task – ROCC of data, cleaning and manipulation, find the trends from the given source of data, suggest marketing strategy
2. Consider key stakeholders - Urška Sršen, Sando Mur, finally my Bellabeat marketing analytics team.

# Prepare

Sršen encourages you to use public data that explores smart device users’ daily habits. She points you to a speciﬁc data set:

* [**FitBit Fitness Tracker Data**](https://www.kaggle.com/arashnic/fitbit) (CC0: Public Domain, dataset made available through [Mobius](https://www.kaggle.com/arashnic)): This Kaggle data set

contains personal ﬁtness tracker from thirty ﬁtbit users. 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.

Sršen tells you that this data set might have some limitations, and encourages you to consider adding another data to help address those limitations as you begin to work more with this data.

Now, prepare your data for analysis using the following Case Study Roadmap as a guide:

**Guiding questions**

* Where is your data stored? Data are stored in PC for initial cleaning process.
* How is the data organized? Is it in long or wide format? The data is organized activity wise and also in merged format. The data is stored in both long and wide format.
* Are there issues with bias or credibility in this data? [Does your data ROCCC?](https://www.coursera.org/learn/data-preparation/lecture/lHirM/what-is-bad-data)

The dataset is publicly available in Kaggle. This data is not perfectly ROCCC complaint, this a periodic dataset from Apr 2016 to May 2016.

* How are you addressing licensing, privacy, security, and accessibility? Since this is a public dataset there is no issue of privacy, licensing and accessibility issue.
* How did you verify the data’s integrity?
* Are there any problems with the data? We have a problem of limited dataset of 30 customers so our solution also will be based on these data. For future scope we test the same analysis with larger dataset and verify the same.

**Key tasks**

1. Have downloaded the data and named them with proper conventions.
2. We have same data with both wide and long format. This need to be cross verified to ensure that there is no bias and finally the sorted individual table will be verified with data in main merged table to ensure data integrity

Delieverable:

The dataset is credible because it’s directed by the course and also available publicly in Kaggle.

Process

**Guiding questions**

* What tools are you choosing and why? Since this is a small dataset we are using Excel to sort and clean the data.
* Have you ensured your data’s integrity?
* What steps have you taken to ensure that your data is clean?
* How can you verify that your data is clean and ready to analyze?
* Have you documented your cleaning process so you can review and share those results?

**Key tasks**

1. Check the data for errors

* The dates in every tables weren’t formatted so we had to format the data in Indian standard dd/mm/yy.
* And the records in auxiliary tables are being verified with main table (dailyactivity\_merged.csv)
* Here for main analysis we will using six main table: 1) dailyactivity\_merged.csv 2) dailyCalories\_merged.csv 3) dailyIntensities\_merged.csv 4) dailysleep\_merged.csv 5) dailysteps\_merged.csv 6) weightLogInfo\_merged.csv
* Rest of the tables forms base for calculating the above-mentioned tables.

1. Choose your tools: Excel for cleaning, R for analysis and Tableau for visualization.
2. Document the cleaning process.

We have 18 tables in csv format. The cleaning documentation have been attached below



# Analyze

Now that your data is stored appropriately and has been prepared for analysis, start putting it to work. Use the following Case Study Roadmap as a guide:

**Guiding questions**

* How should you organize your data to perform analysis on it?
* Has your data been properly formatted?
* What surprises did you discover in the data?
* What trends or relationships did you ﬁnd in the data?
* How will these insights help answer your business questions?

The analysis have been documented and all the findings have been attached as document below:



This case study has been published / shared in Github platform.