

**About**

Bella beat is a technology-driven women’s wellness brand offering various products and services focused on women's health. By gathering data on activity, sleep, stress, and reproductive health, Bella beat empowers women with insights into their own health and habits.

Founded in 2013, Bella beat has swiftly expanded, establishing itself as a leading tech-driven wellness company for women. Urška Sršen, co-founder and Chief Creative Officer of Bella beat, believes that analyzing fitness data from smart devices can help uncover new growth opportunities for the company.

**The key steps followed in the analysis:**

* Ask
* Prepare
* Process
* Analyze

**ASK**

In the ask phase of the data analysis process, the current problem is identified, the business task is defined, key stakeholders are considered, and the potential benefits of insights and findings for these stakeholders are outlined.

**Business Task** Analyze smart device fitness data to gain insights into consumer behavior and identify trends that can inform Bellabeat's marketing strategy for their app.

**Key Stakeholders** The key stakeholders for this project include:

* **Urška Sršen**: Bellabeat’s co-founder and Chief Creative Officer.
* **Sando Mur**: Mathematician, Bellabeat’s co-founder, and key member of the Bellabeat executive team.
* **Bella beat Marketing Analytics Team**: A team of data analysts responsible for collecting, analyzing, and reporting data to guide Bellabeat’s marketing strategy.

**PREPARE**

* We have opted for my sql workbench and excel for our data analysis due to their strong focus on data integrity, which guarantees the reliability of our analysis results. MySQL Workbench is an excellent choice for database design and management due to its user-friendly interface and comprehensive features. It offers a visual tool for database modeling, enabling users to design, model, and maintain databases with ease. The SQL development environment is robust, providing advanced code editing, syntax highlighting, and query execution capabilities. Additionally, MySQL Workbench supports database administration tasks, including user management, server configuration, and performance monitoring, making it a versatile tool for developers and DBAs alike. Its integration with MySQL Server ensures seamless workflow and enhanced productivity.
* For visualization we are using power BI, another powerful tool for data visualization and reporting because it provides powerful, user-friendly tools to create interactive and insightful visualizations, helping us convey our data analysis results effectively to both internal and external stakeholders.

## Data Sources

We will be using public data that explores smart device users’ daily habits:

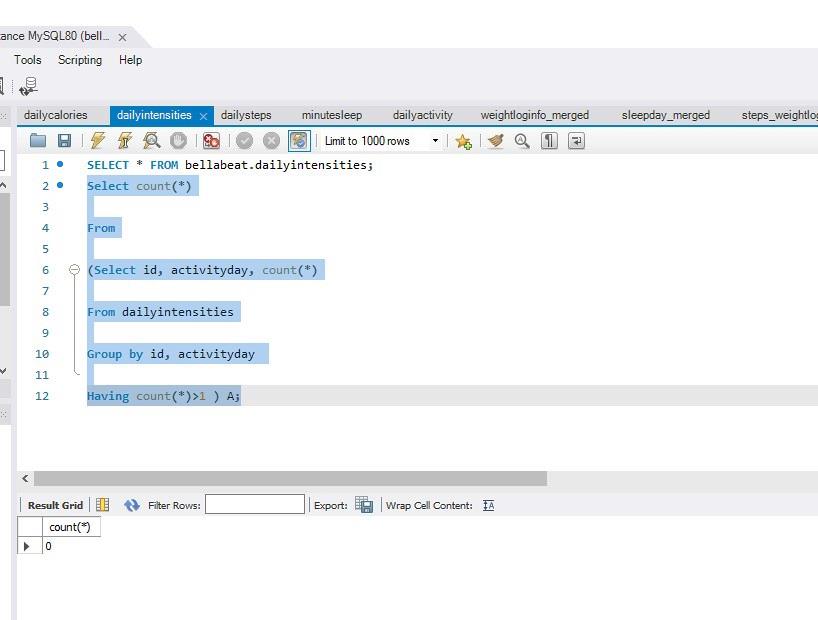
* [**Fit Bit Fitness Tracker Data**](https://www.kaggle.com/arashnic/fitbit)(CC0: Public Domain, dataset made available through [Mobius](https://www.kaggle.com/arashnic)): This Kaggle data set contains personal tracker data from thirty Fit bit users. These Fit bit users consented to submission of personal tracker data, including minute-level output for physical activity, heart rate, and sleep monitoring.

I am going to explore the Fit Bit Fitness Tracker data across ten CSV files. The data is organized in rows and columns. They are broadly classified based on time tracked into:

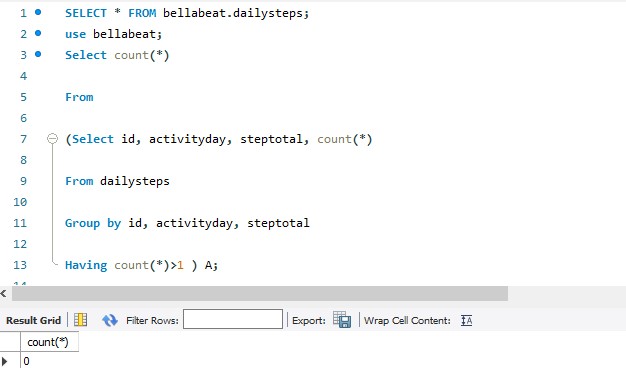
1. Daily activity
2. Daily steps
3. Weight log
4. Sleep day

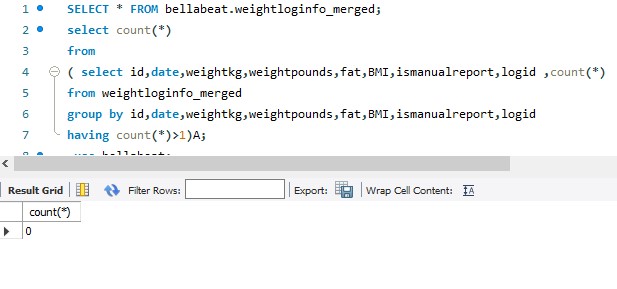
**Process**

In the third phase of data analysis, we focus on processing and cleaning the data to prepare it for analysis. This step involves identifying and correcting inaccuracies, errors, and inconsistencies in the data, ensuring that it is free from issues that could skew results or introduce discrepancies. Cleaning the data is crucial for maintaining the accuracy and credibility of the analysis.



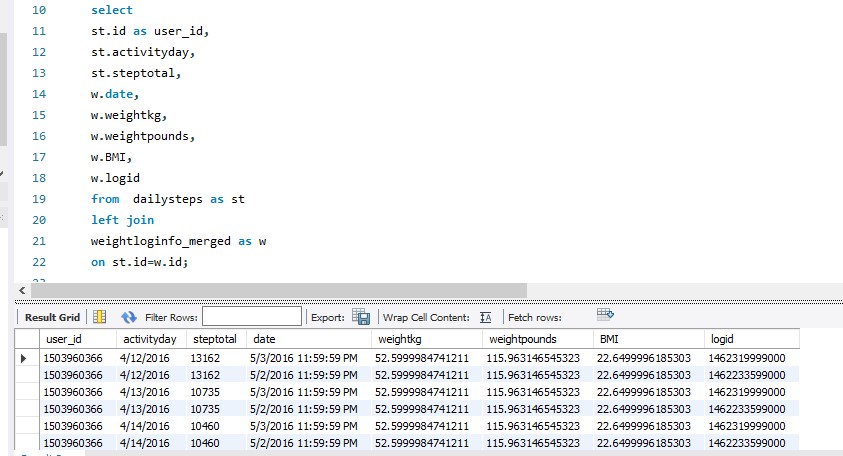
**Data Cleaning:** With the data imported, the next step is to ensure that the data within each table is clean and ready for analysis.





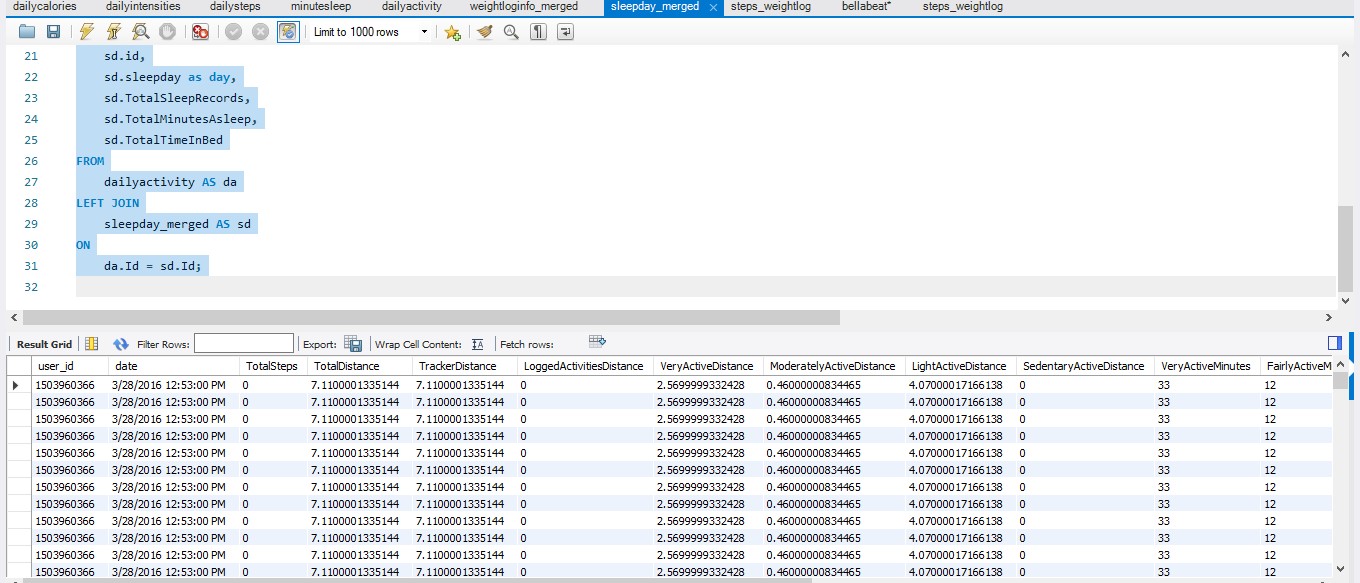
**Data Integration:** To facilitate comprehensive analysis, I have joined relevant tables to create a unified dataset. This involved merging tables based on common keys to ensure that all necessary information is consolidated, allowing for more accurate and insightful analysis.

**Joined two tables daily steps and weight log**



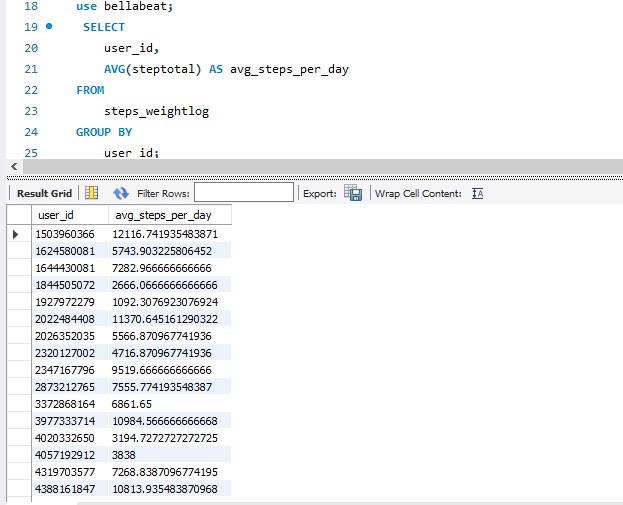
**Another tables joined are daily activity and sleep day**



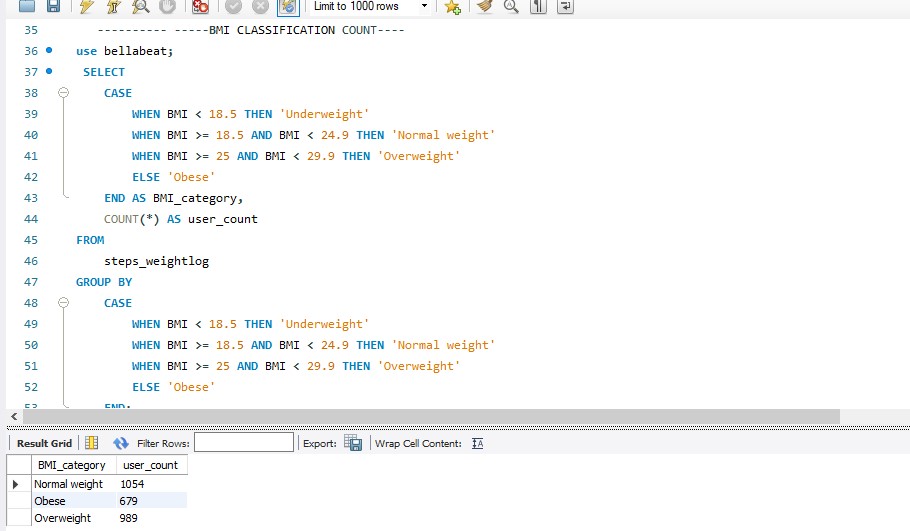


From above joined tables I have analyzed few KPI metrics:

Average steps per user



BMI classification:

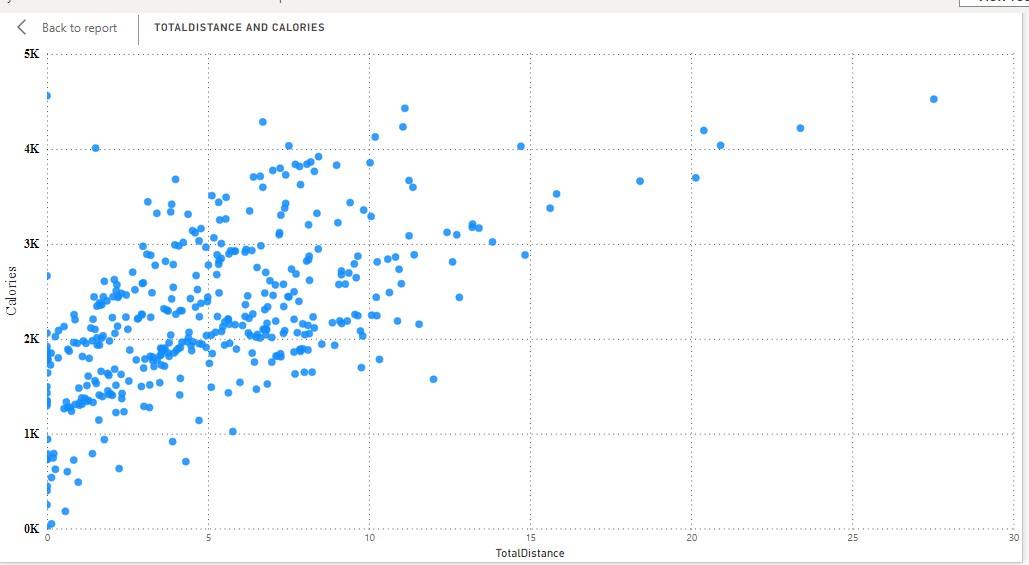


Distance count based on active distance

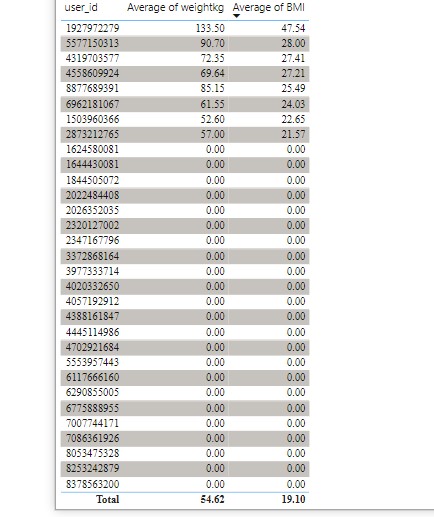


# ANALYZE

The fourth phase is to analyze the data. The main goal in this phase is to find trends, patterns and relationship between data that will help solve the main business goal at hand. Here, the data is further sorted and formatted to understand how different groups work individually and in relation to one another and powerful insights are drawn. This phase is all about finding what story the data is trying to tell.



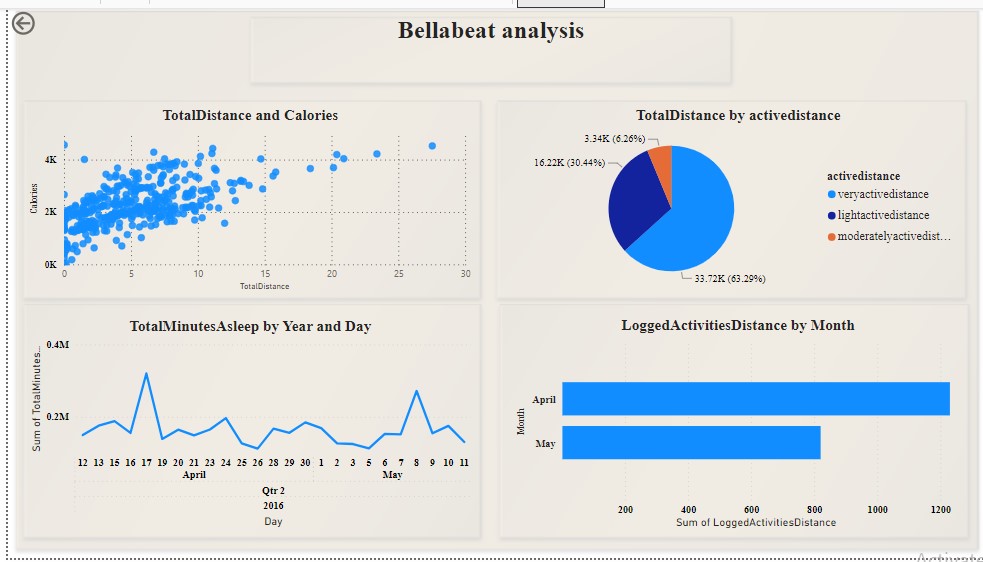
The table displays the average weight and BMI for a list of users identified by their user IDs. Among the users, the highest recorded average weight is 133.50 kg with a corresponding BMI of 47.54. There are several entries where both weight and BMI are recorded as zero, indicating missing or unrecorded data for those users. The overall averages at the bottom of the table indicate an average weight of 54.62 kg and an average BMI of 19.10 across the recorded entries. This detailed data can help in analyzing user health metrics and identifying trends or anomalies within the dataset.



**Visualization**

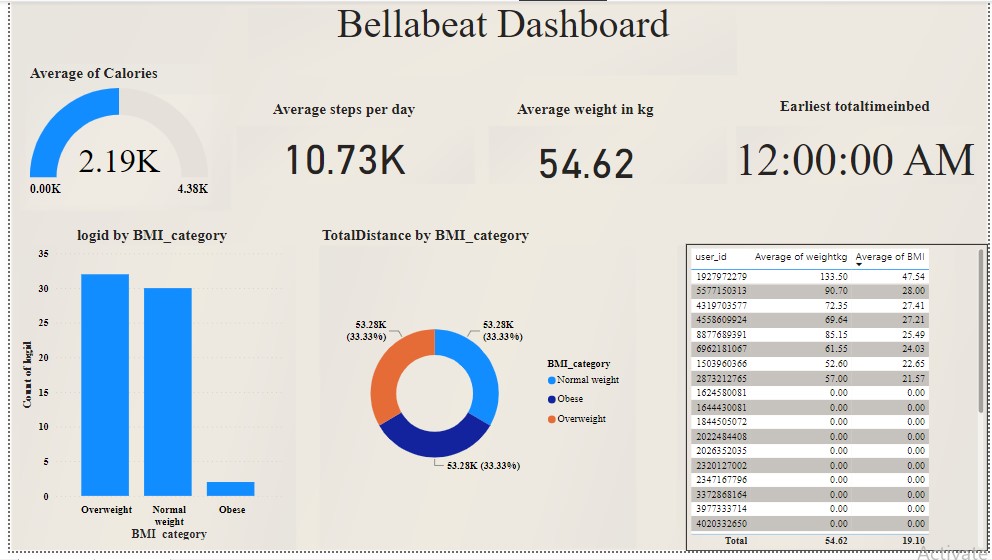
1. **Total minutes asleep by year, day and month** Created a line graph in spreadsheet to visually represent how the women in this sample spend their sleep time in the day.  
   **key findings**: April was a month with highest rating of minutes asleep.
2. **Pie chart :**used to represent active distance by total distance covered.
3. **Bar chart:** around 1200 ids were logged in month of April which is highest throughout year.

The Bella beat analysis dashboard offers detailed insights into user activity and sleep patterns. The scatter plot titled "Total Distance and Calories" reveals a positive correlation between the distance traveled and calories burned. The pie chart "Total Distance by Active Distance" categorizes total distance into very active, lightly active, and moderately active segments, showing that the majority (63.29%) falls under very active distance. The line graph "Total Minutes Asleep by Year and Day" tracks sleep patterns over time, highlighting fluctuations in total minutes asleep across different days from April to May 2016. Additionally, the bar chart "Logged Activities Distance by Month" compares the sum of logged activity distances between April and May, indicating higher activity levels in April. Together, these visualizations provide a comprehensive view of user behaviors, supporting Bellabeat's data-driven approach to enhancing women's wellness.



**Key findings**

* 7628 - Average total steps in a day which is below the recommended 10,000 steps per day.
* 12:00AM is earliest time in bed
* 2190 total average calories burnet in a day



The Bella beat Dashboard provides a comprehensive overview of key health metrics derived from smart device data. The dashboard showcases various aggregated data points such as the average calories burned (2.19K), average steps per day (10.73K), and average weight in kilograms (54.62). It also highlights the earliest recorded total time in bed, marked as 12:00:00 AM. Visual representations, such as the bar chart categorizing logins by BMI category and the pie chart illustrating total distance covered by BMI category, offer insights into user behavior and health trends. Additionally, a detailed table lists user IDs alongside their average weight and BMI, providing granular data that can inform targeted health interventions and personalized marketing strategies for Bella beat.