Business Task Statement

Bellabeat wants to analyze smart device usage trends from non-Bellabeat devices to identify patterns in user engagement, activity tracking, and wellness habits. These insights will then be applied to a selected Bellabeat product to refine its marketing strategy, improve user engagement, and optimize customer acquisition.

Key Stakeholders

- Urška Sršen (Cofounder & CCO) – Wants data-driven insights to enhance product marketing.
- Sando Mur (Cofounder & Mathematician) – Key decision-maker on product positioning.
- Bellabeat Marketing Analytics Team – Responsible for extracting insights and making data-driven recommendations.

Data Source Overview: Fitbit Fitness Tracker Data

This dataset, titled `Fitbit Fitness Tracker Data`, is available on Kaggle as a CC0 Public Domain resource, meaning it is free for use without copyright restrictions. It was made available through Mobius and contains personal fitness tracker data from 30 eligible Fitbit users. Key features include:

- Activity Data**:** Daily activity metrics such as steps taken, distance covered, and activity levels.
- Heart Rate Data: Minute-level heart rate measurements.
- Sleep Data: Detailed sleep monitoring information, including duration and sleep stages.

The data is self-reported from a limited sample of thirty Fitbit users, which might not represent the broader population and can be biased.

Introduction of Data Analysis

In this report, we explore a dataset of user activity, sleep, and weight data to analyze patterns in physical activity and its relationships to health metrics like calories burned and weight. The primary aim is to understand the activity habits of users, categorize them based on their activity levels, and identify key factors that influence health behaviors. The analysis is performed using data from three datasets: `activity_data (dailyActivity_merged.csv)`, `sleep_data (sleepDay_merged.csv)`, and `weight_data`

(weightLogInfo_merged.csv) from Fitabase Data 4.12.16-5.12.16 folder.

Data Preprocessing and Cleaning

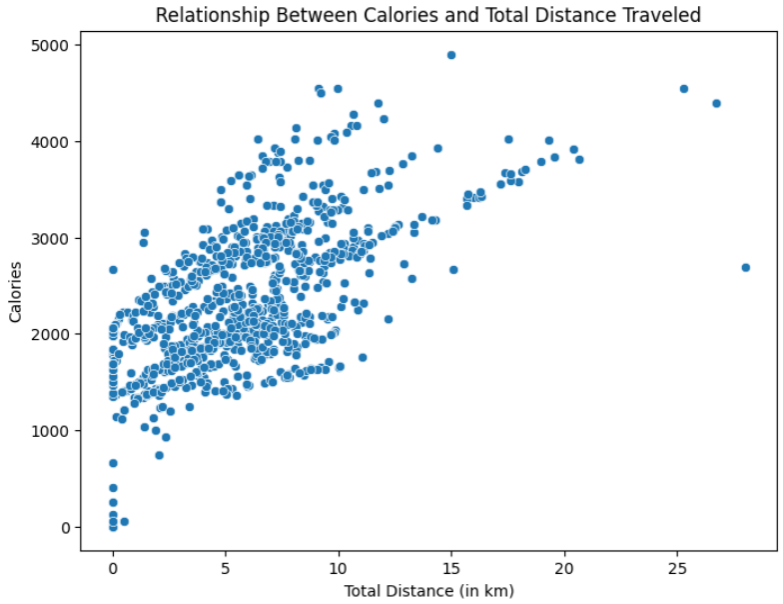
Upon inspecting the data, several unnecessary columns were removed, including SedentaryActive, TrackerDistance from activity data, TotalTimeInBed from sleep data and WeightPounds, Fat from weight data. Missing and duplicate data were handled. The data was merged when needed, based on user IDs and dates, to allow for comparison and finding correlation of activity, sleep, and weight data. The date columns were converted to datetime category for effective analysis.

Exploratory Data Analysis (EDA)

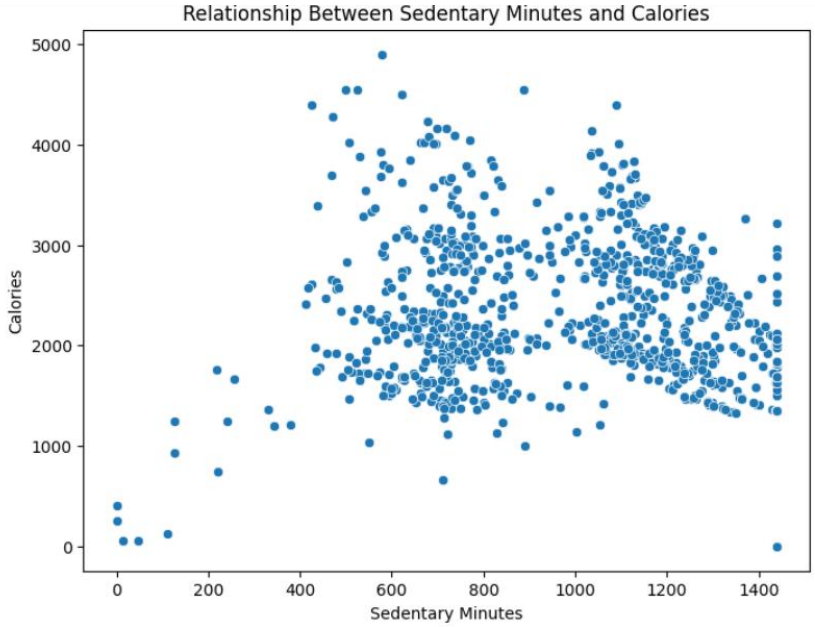
To better understand the data, we calculated descriptive statistics and visualized the key variables.

- **Calories vs Steps**: A positive correlation was observed between calories burned and number of steps taken.

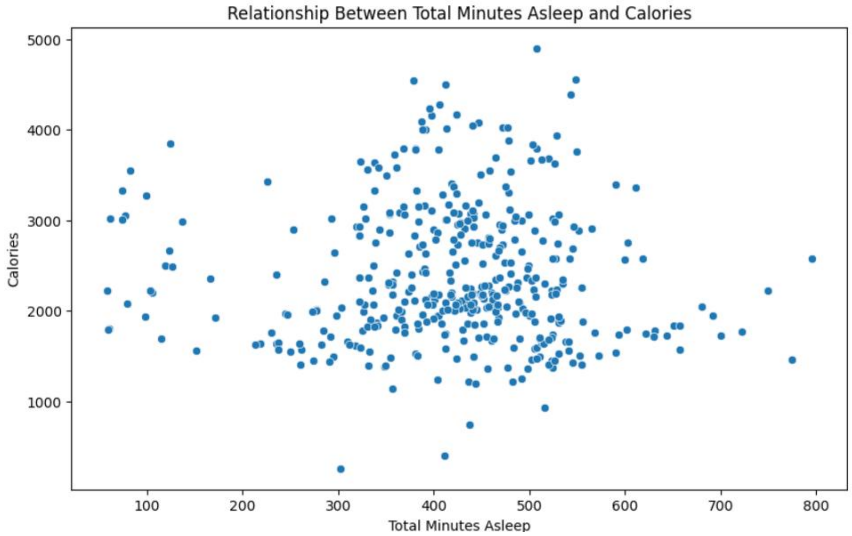
- **Calories vs Distance**: A positive correlation was observed between calories burned and total distance travelled.



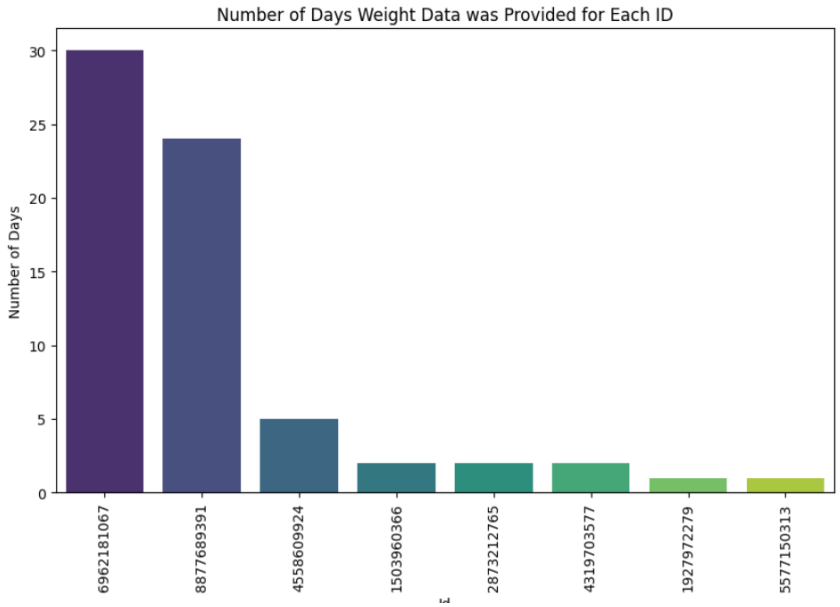
- **Calories vs Sedentary Minutes:** A negative correlation was observed between calories burned and sedentary minutes.



- **Calories vs Total Minutes Asleep:** A negative correlation was observed between calories burned and total minutes asleep.

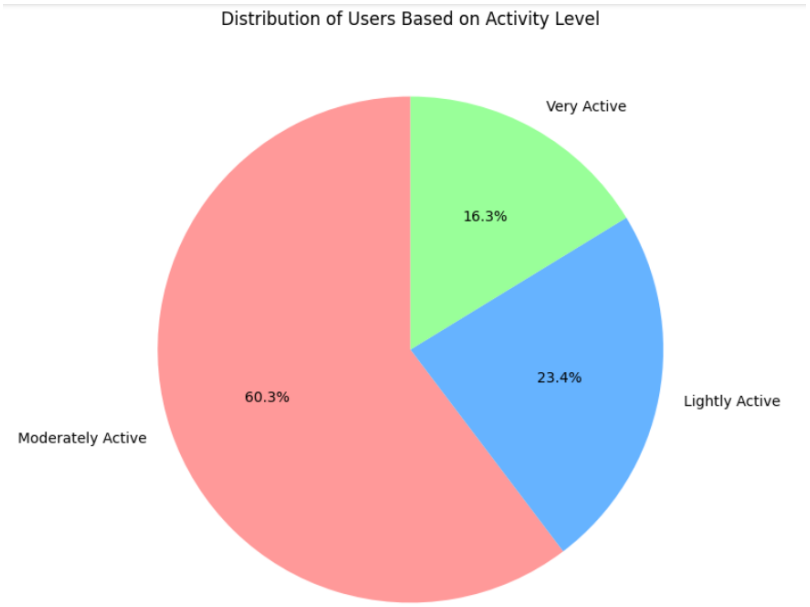


- **Shortage of weight data**: Only 8 users recorded their weight over the time period compared to 33 for activity data and 24 for activity data and most of them did not have weight records for majority of days.

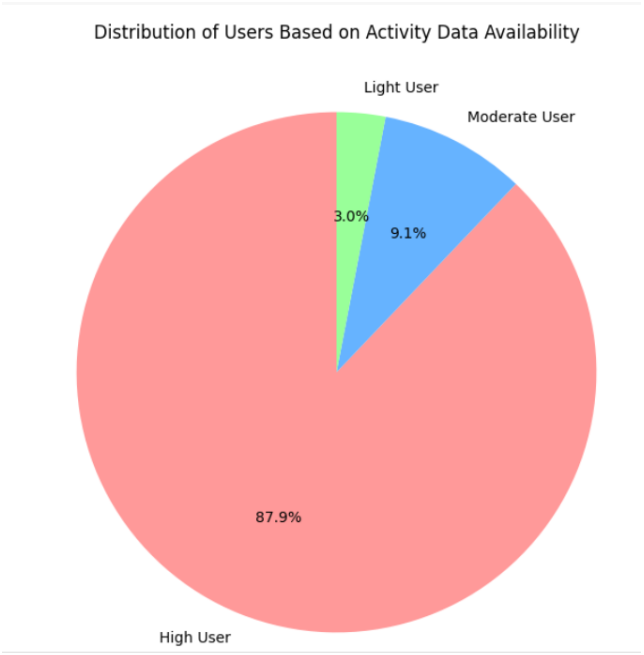


- **User Activity Categories:**

Users were categorized into three groups based on calories burned: ‘Very Active’ (3000+), ‘Moderately Active’ (1800+) and ‘Lightly Active’ (1800-). Majority fell into the ‘Moderately Active’ category as seen in the pie chart below.



Also, based on the number of days of activity data, users were categorized into three groups: ‘High’ (21+ days), ‘Moderate’ (11+ days) and ‘Light’ (11- days) users of the smart device. A pie chart visualized the distribution of users across these categories, with most users falling into the ‘Moderate User’ category.



Recommendations (for Bellabeat app and the company as a whole)

- **Daily Notifications/Reminders**

Implement daily reminders for users to log their activity, sleep, and health metrics. Personalized notifications will encourage consistent engagement and help users form habits, increasing data input and overall app usage.

- **Reward System**

Introduce a reward system to incentivize regular usage of the app, especially for tasks like weight and sleep tracking. Users could earn points or badges for achieving milestones, with rewards redeemable for discounts on Bellabeat products or services.

- **Organize Promotional Events**

Host events to promote the benefits of regular physical activity and the Bellabeat app. These events can raise brand awareness, educate the public, and encourage more users to engage in health challenges, thus increasing app adoption and usage.

- **Further Analysis**

Expand the app's tracking capabilities to include additional metrics like blood pressure, blood sugar and body fat percentage. This will provide users with a more comprehensive view of their health, fostering better decisions and deeper engagement.

- **Enhanced Personalization Through AI/ML-Driven Insights**

Utilize AI and ML to offer personalized health insights and recommendations based on individual user data. This will improve user experience and retention by providing actionable, tailored suggestions for better health outcomes.