

Case Study : How Can a Wellness Technology Company Play It Smart?

Analysis by Sajith M

INTRODUCTION AND MOTIVATION

My name is Sajith M and I am a graduate in computer science from University Of Calicut and currently I am seeking opportunities to show case my skills at reputed organizations. During my college days I became very Interested in Data science and Data Analysis and decided to make a career in these fields. I really like the concept of data and how much of today's world revolve around them, when I more looked into it I saw these fields of Data Science and Data Analysis as the major areas that studies data.

So, I took upon a course on Data science about 6 months and completed it and also took on a Google certified course at Coursera to further increase my knowledge, I make these case studies to further increase my knowledge as self-improvement is one way of achieving my dreams.

BUSINESS PROBLEM AND DATA

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 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

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 (CC0: Public Domain, dataset made available through Mobius): 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.

So from this dataset we need to analyze and make recommendations to improve our product

DATA FEATURE ENGINEERING

We are only importing the required datasets as there are many type of data sets that have same columns so it does not hinder our analysis process

So, after importing the datasets, I cleaned the datasets and I saw there are null values and I decide to drop those null values

The next step is to verify the number of unique users as the ID column acts as a foreign key across the whole dataset, therefore I could merge the whole datasets using ID as it is shared by each dataframe

I noticed that daily_sleep dataset has 3 duplicated rows where as none of other has duplicated rows. Let's merge the duplicated rows in daily_sleep

So, I decided to combine the data_sleep_1 and daily_activity but I saw the in daily_activity it is activitydate col but in daily_sleep it is sleepdate col lets rename it

There is discrepancies in all of the date formats in all of these dataframes lets change into same format using lubridate package In hourly dataframes the There is discrepancies in the timestamps so lets also correct that

ANALYZE PHASE

I merge the dailyActivity_merged table and daily_sleep_1 into a new data frame called "daily_activity_and_sleep"

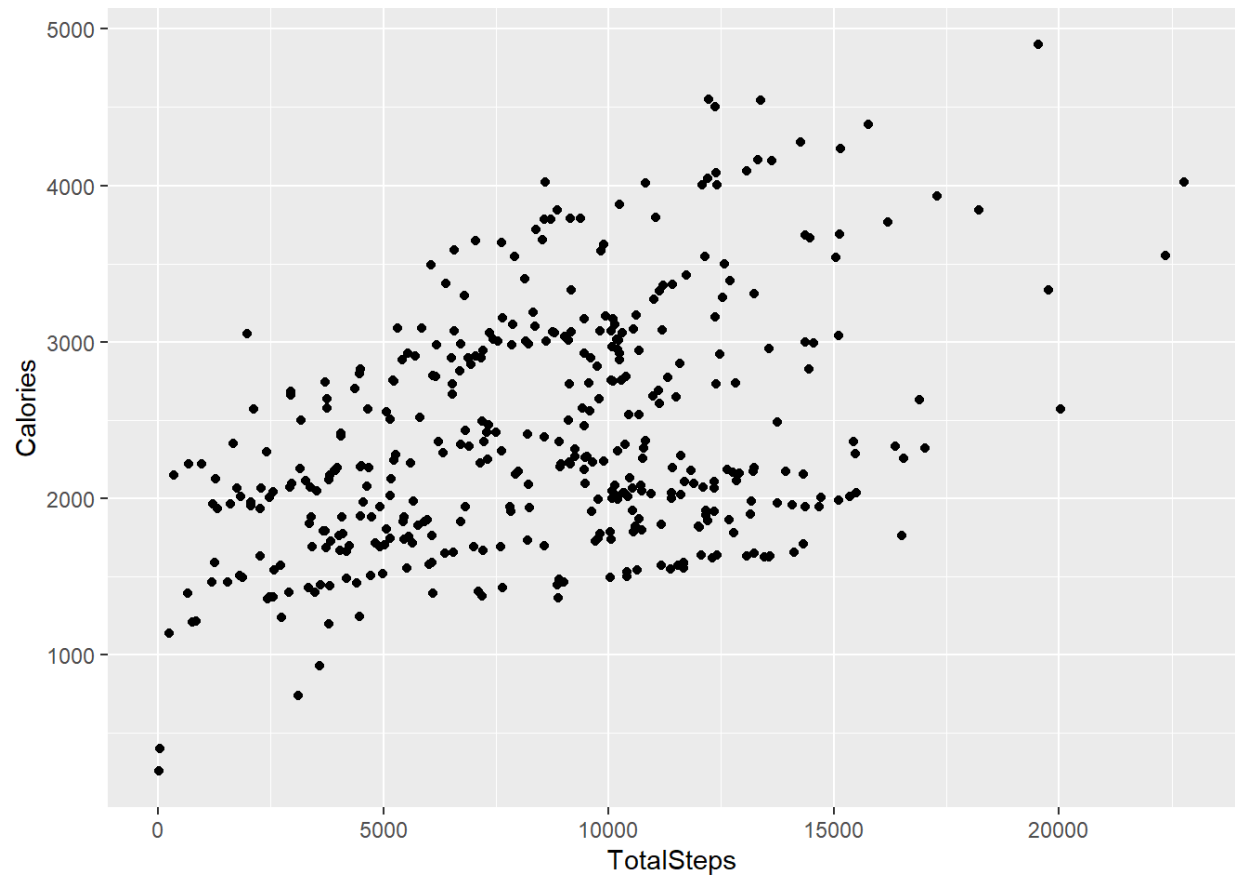
I merge the hourlyCalories, hourlyIntensities these 2 data frames into a single data frame called "hourly_activity". This is done via "Id" and "ActivityHour"

I merge the hourly activity dataframe and hourly steps in to a new dataframe called "hourly_act"

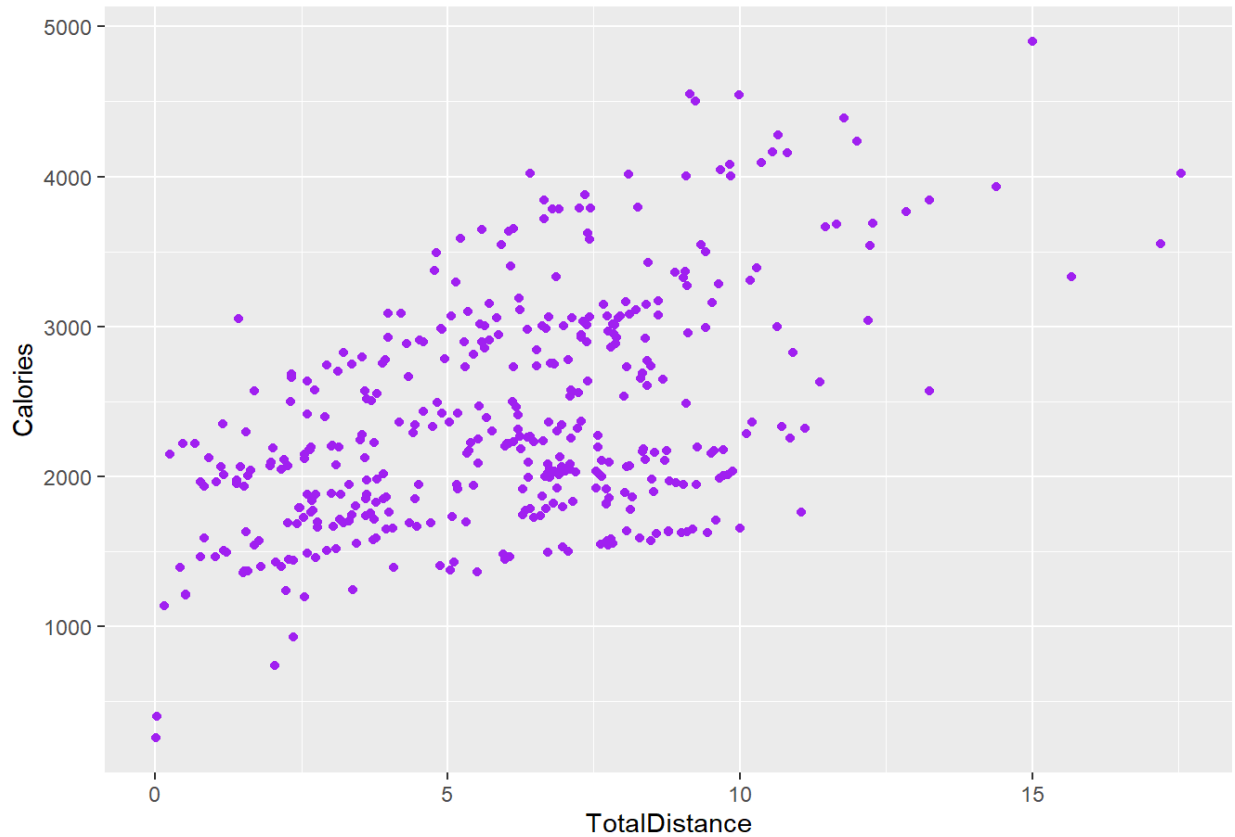
Then I checked each of these 3 newly created dataframes and proceeded to visualization phase

DATA VISUALIZATION PHASE

Determination of relationship between TotalSteps and Calories of each users using the daily_activity_with_sleep dataframe



Determining the relationship between the distance and the calories burnt



- We can see a positive correlation between the Total Steps and calories burnt. Therefore, users who are using the Bellabeat smart technologies could try to integrate a healthier lifestyle by performing daily exercise such as jogging. Bellabeat can create an app to track numbers of steps and calories burnt from users daily.
- There is also a positive correlation between the number of sleeps with the calories burnt. Bellabeat can showcase customers the benefits of sufficient amount of sleep (6-8 hours) to attain weight loss.
- A marketing strategy should be themed around promoting Bellabeat products that emphasise on providing daily encouragement and motivation to its user base. Especially for those aiming to lose weight, the # The calories one burns in a day is a key metric.