## horizontal line



Wellness Technology Company Play It Smart

05.09.2023

**─**

Adimulam Hima Sankar

Case Study

# Overview

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. 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 displays ads on the Google Display Network to support campaigns around key marketing dates.

# Goals

1. 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.
2. Then, using this information, we would like high-level recommendations for how these trends can inform Bellabeat marketing strategy.

# PHASE 1 : ASK

**Key points:**

1. **Urška Sršen:** Bellabeat’s co-founder and Chief Creative Officer.
2. **Sando Mur**: Mathematician and Bellabeat cofounder; key member of the Bellabeat executive team
3. **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.

# PHASE 2 : 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 (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. 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.

From that data i choosen as dailyActivity\_merged.

It was organised by

Id,

ActivityDate,

TotalSteps,

TotalDistance,

TrackerDistance,

LoggedActivitiesDistance,

VeryActiveDistance,

ModeratelyActiveDistance,

LightActiveDistance,

SedentaryActiveDistance,

VeryActiveMinutes,

FairlyActiveMinutes,

LightlyActiveMinutes,

SedentaryMinutes,

Calories.

#Check it first is there any blanks on it. To check the blanks use the function as

=IS.FILTER()

#Sort the ActivityDate by using the date format

SELECT (ActivityDate) - DATA - Text to columns - Delimited - Tab - Format : MDY - Finish.

#Check all the columns is there any blanks otherwise go to another step

# PHASE 3 : PROCESS

Find out the outliers

* To find the outliers we need to know about the quartiles.
* To find the quartile the formula as =QUARTILE(array,Quart).
* Find quartile one and quartile three, then subtract quartile three from quartile two.
* Then multiply 1.5 with (Quartile3 - Quartile1)
* Find the lower limits and upper limits

After all the findings,

|  | **1st Quartile** | **3rd Quartile** | **Q3 - Q1** | **1.5\*(Q3-Q1)** | **Lower Limit =Q1-1.5IQR** | **Upper Limit = Q3+1.5IQR** |
| --- | --- | --- | --- | --- | --- | --- |
| **TotalSteps** | 3789.75 | 10727 | 6937.25 | 10405.875 | -6616.125 | 21132.875 |
| **TotalDistance** | 2.61999989 | 7.712499976 | 5.0925 | 7.63875014 | -5.01875025 | 15.35125011 |
| **TrackerDistance** | 2.61999989 | 7.710000038 | 5.09 | 7.63500023 | -5.015000343 | 15.34500027 |
| **Logged Activities Distance** | 0 | 0 | 0 | 0 | 0 | 0 |
| **VeryActiveDistance** | 0 | 2.05249995 | 2.0525 | 3.07874992 | -3.078749925 | 5.131249875 |
| **ModeratelyActiveDistance** | 0 | 0.800000012 | 0.8 | 1.20000002 | -1.200000018 | 2.00000003 |
| **LightActiveDistance** | 1.94500002 | 4.782500148 | 2.8375 | 4.25625019 | -2.311250165 | 9.038750336 |
| **SedentaryActiveDistance** | 0 | 0 | 0 | 0 | 0 | 0 |
| **VeryActiveMinutes** | 0 | 32 | 32 | 48 | -48 | 80 |
| **FairlyActiveMinutes** | 0 | 19 | 19 | 28.5 | -28.5 | 47.5 |
| **LightlyActiveMinutes** | 127 | 264 | 137 | 205.5 | -78.5 | 469.5 |
| **SedentaryMinutes** | 729.75 | 1229.5 | 499.75 | 749.625 | -19.875 | 1979.125 |
| **Calories** | 1828.5 | 2793.25 | 964.75 | 1447.125 | 381.375 | 4240.375 |

# PHASE 4 : ANALYZE

#In this phase i used pivot tables for clear idea about the data.

Select all the data and insert the pivot table and drag the Id to rows. They count as 33. Then select the ActivityDate and drag into values then fix it as count.

| **Row Labels** | **Count of ActivityDate** |
| --- | --- |
| 1503960366 | 31 |
| 1624580081 | 31 |
| 1644430081 | 30 |
| 1844505072 | 31 |
| 1927972279 | 31 |
| 2022484408 | 31 |
| 2026352035 | 31 |
| 2320127002 | 31 |
| 2347167796 | 18 |
| 2873212765 | 31 |
| 3372868164 | 20 |
| 3977333714 | 30 |
| 4020332650 | 31 |
| 4057192912 | 4 |
| 4319703577 | 31 |
| 4388161847 | 31 |
| 4445114986 | 31 |
| 4558609924 | 31 |
| 4702921684 | 31 |
| 5553957443 | 31 |
| 5577150313 | 30 |
| 6117666160 | 28 |
| 6290855005 | 29 |
| 6775888955 | 26 |
| 6962181067 | 31 |
| 7007744171 | 26 |
| 7086361926 | 31 |
| 8053475328 | 31 |
| 8253242879 | 19 |
| 8378563200 | 31 |
| 8583815059 | 31 |
| 8792009665 | 29 |
| 8877689391 | 31 |

Next, add the Total steps, Total distance, Very active distance, Moderately active distance, Calories, Light Active Distance and change it to average. Then result should be like this,

| **Row Labels** | **Count of ActivityDate** | **Average of TotalSteps** | **Average of TotalDistance** | **Average of VeryActiveDistance** | **Average of ModeratelyActiveDistance** | **Average of Calories** | **Average of LightActiveDistance** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 1503960366 | 31 | 12116.74194 | 7.809677386 | 2.858387113 | 0.79419354 | 1816.419355 | 4.152903241 |
| 1624580081 | 31 | 5743.903226 | 3.914838729 | 0.939354843 | 0.360645162 | 1483.354839 | 2.60677418 |
| 1644430081 | 30 | 7282.966667 | 5.295333354 | 0.730000001 | 0.951000002 | 2811.3 | 3.609000021 |
| 1844505072 | 31 | 2580.064516 | 1.706129037 | 0.008387097 | 0.049032256 | 1573.483871 | 1.647419339 |
| 1927972279 | 31 | 916.1290323 | 0.634516123 | 0.095806451 | 0.031290323 | 2172.806452 | 0.507096769 |
| 2022484408 | 31 | 11370.64516 | 8.084193491 | 2.421612884 | 0.720000003 | 2509.967742 | 4.942580638 |
| 2026352035 | 31 | 5566.870968 | 3.454838715 | 0.006129032 | 0.011290322 | 1540.645161 | 3.43612904 |
| 2320127002 | 31 | 4716.870968 | 3.187741904 | 0.106774193 | 0.097741935 | 1724.16129 | 2.980322567 |
| 2347167796 | 18 | 9519.666667 | 6.355555536 | 1.059444433 | 1.074999991 | 2043.444444 | 4.221666708 |
| 2873212765 | 31 | 7555.774194 | 5.10161286 | 0.676129031 | 0.276129032 | 1916.967742 | 4.143548414 |
| 3372868164 | 20 | 6861.65 | 4.707000041 | 0.629499997 | 0.152999998 | 1933.1 | 3.909999967 |
| 3977333714 | 30 | 10984.56667 | 7.516999944 | 1.614999982 | 2.750999979 | 1513.666667 | 3.134333344 |
| 4020332650 | 31 | 2267.225806 | 1.626129039 | 0.142258061 | 0.129677417 | 2385.806452 | 1.308387082 |
| 4057192912 | 4 | 3838 | 2.862500012 | 0.052499998 | 0.064999998 | 1973.75 | 2.6875 |
| 4319703577 | 31 | 7268.83871 | 4.892258047 | 0.278064516 | 0.502258065 | 2037.677419 | 3.768709638 |
| 4388161847 | 31 | 10813.93548 | 8.393225893 | 1.719354835 | 0.901935478 | 3093.870968 | 5.396129047 |
| 4445114986 | 31 | 4796.548387 | 3.24580644 | 0.523225811 | 0.07548387 | 2186.193548 | 2.64483871 |
| 4558609924 | 31 | 7685.129032 | 5.080645177 | 0.549354839 | 0.682258066 | 2033.258065 | 3.847741946 |
| 4702921684 | 31 | 8572.064516 | 6.955161283 | 0.417419355 | 1.304838709 | 2965.548387 | 5.225483921 |
| 5553957443 | 31 | 8612.580645 | 5.63967745 | 1.464193549 | 0.669032263 | 1875.677419 | 3.504516096 |
| 5577150313 | 30 | 8304.433333 | 6.213333305 | 3.113666656 | 0.658000005 | 3359.633333 | 2.427999989 |
| 6117666160 | 28 | 7046.714286 | 5.342142914 | 0.128214285 | 0.083928572 | 2261.142857 | 4.843214299 |
| 6290855005 | 29 | 5649.551724 | 4.272413805 | 0.085517242 | 0.128275863 | 2599.62069 | 4.048620709 |
| 6775888955 | 26 | 2519.692308 | 1.813461516 | 0.709230767 | 0.38423077 | 2131.769231 | 0.711538466 |
| 6962181067 | 31 | 9794.806452 | 6.585806477 | 1.616451619 | 0.959999993 | 1982.032258 | 4.001612925 |
| 7007744171 | 26 | 11323.42308 | 8.015384592 | 2.414999987 | 0.73846153 | 2544 | 4.861538447 |
| 7086361926 | 31 | 9371.774194 | 6.388064508 | 2.781290325 | 0.773225805 | 2566.354839 | 2.818709662 |
| 8053475328 | 31 | 14763.29032 | 11.4751612 | 8.514838742 | 0.423870965 | 2945.806452 | 2.533870955 |
| 8253242879 | 19 | 6482.157895 | 4.667368468 | 2.21421049 | 0.695789477 | 1788 | 1.754736845 |
| 8378563200 | 31 | 8717.709677 | 6.913548462 | 2.503548402 | 0.519032255 | 3436.580645 | 3.889354821 |
| 8583815059 | 31 | 7198.516129 | 5.615483822 | 0.798064526 | 1.020645153 | 2732.032258 | 2.617419347 |
| 8792009665 | 29 | 1853.724138 | 1.186551717 | 0.024827587 | 0.058275863 | 1962.310345 | 1.103448271 |
| 8877689391 | 31 | 16040.03226 | 13.21290314 | 6.637419362 | 0.337741935 | 3420.258065 | 6.188709674 |

Next, sort the data as activity date and largest to smallest, calories as largest to smallest and total steps as largest to smallest. Then we get this,

| **Row Labels** | **Count of ActivityDate** | **Average of Calories** | **Average of TotalSteps** |
| --- | --- | --- | --- |
| 8792009665 | 29 | 1982.032258 | 8304.433333 |
| 1624580081 | 31 | 1875.677419 | 11323.42308 |
| 8053475328 | 31 | 1573.483871 | 8717.709677 |
| 1844505072 | 31 | 2965.548387 | 916.1290323 |
| 4445114986 | 31 | 1916.967742 | 9794.806452 |
| 7007744171 | 26 | 2945.806452 | 11370.64516 |
| 8583815059 | 31 | 1483.354839 | 16040.03226 |
| 1503960366 | 31 | 3359.633333 | 7282.966667 |
| 2347167796 | 18 | 1962.310345 | 5649.551724 |
| 4702921684 | 31 | 1724.16129 | 6482.157895 |
| 7086361926 | 31 | 2172.806452 | 10813.93548 |
| 3977333714 | 30 | 2599.62069 | 9519.666667 |
| 4558609924 | 31 | 2043.444444 | 7685.129032 |
| 4388161847 | 31 | 1933.1 | 2519.692308 |
| 5577150313 | 30 | 2261.142857 | 3838 |
| 2320127002 | 31 | 2037.677419 | 8572.064516 |
| 6117666160 | 28 | 2566.354839 | 7555.774194 |
| 6775888955 | 26 | 1788 | 14763.29032 |
| 8378563200 | 31 | 2811.3 | 5566.870968 |
| 8253242879 | 19 | 2732.032258 | 4716.870968 |
| 8877689391 | 31 | 1816.419355 | 9371.774194 |
| 3372868164 | 20 | 3436.580645 | 12116.74194 |
| 2022484408 | 31 | 2186.193548 | 7268.83871 |
| 4020332650 | 31 | 2385.806452 | 2267.225806 |
| 6290855005 | 29 | 2544 | 6861.65 |
| 2026352035 | 31 | 3093.870968 | 2580.064516 |
| 5553957443 | 31 | 1540.645161 | 7198.516129 |
| 1644430081 | 30 | 2131.769231 | 4796.548387 |
| 4057192912 | 4 | 3420.258065 | 5743.903226 |
| 2873212765 | 31 | 1973.75 | 7046.714286 |
| 4319703577 | 31 | 1513.666667 | 1853.724138 |
| 6962181067 | 31 | 2509.967742 | 10984.56667 |
| 1927972279 | 31 | 2033.258065 | 8612.580645 |

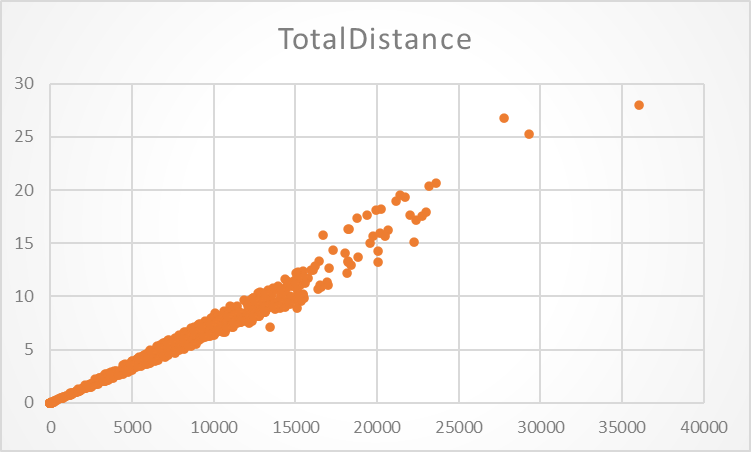
Here, we can see the average total steps as 8304 then the average calories as 1982 for the first id. It be different for all the persons who walk more than a distance between morning and afternoon. The calories are not equal to the steps of the same person because its depends on the body weight and energy.

| **Row Labels** | **Count of ActivityDate** | **Max of TotalSteps** | **Max of TotalDistance** |
| --- | --- | --- | --- |
| 1503960366 | 31 | 18134 | 12.21000004 |
| 1624580081 | 31 | 36019 | 28.03000069 |
| 1644430081 | 30 | 18213 | 13.23999977 |
| 1844505072 | 31 | 8054 | 5.320000172 |
| 1927972279 | 31 | 3790 | 2.619999886 |
| 2022484408 | 31 | 18387 | 12.90999985 |
| 2026352035 | 31 | 12357 | 7.710000038 |
| 2320127002 | 31 | 10725 | 7.489999771 |
| 2347167796 | 18 | 22244 | 15.07999992 |
| 2873212765 | 31 | 9685 | 6.650000095 |
| 3372868164 | 20 | 9715 | 6.630000114 |
| 3977333714 | 30 | 16520 | 11.05000019 |
| 4020332650 | 31 | 11728 | 8.430000305 |
| 4057192912 | 4 | 5974 | 4.46999979 |
| 4319703577 | 31 | 13658 | 9.489999771 |
| 4388161847 | 31 | 22770 | 17.54000092 |
| 4445114986 | 31 | 9105 | 6.110000134 |
| 4558609924 | 31 | 13743 | 9.079999924 |
| 4702921684 | 31 | 15126 | 12.27000046 |
| 5553957443 | 31 | 17022 | 11.11999989 |
| 5577150313 | 30 | 15764 | 11.77999973 |
| 6117666160 | 28 | 19542 | 15.01000023 |
| 6290855005 | 29 | 9837 | 7.440000057 |
| 6775888955 | 26 | 10771 | 7.71999979 |
| 6962181067 | 31 | 20031 | 13.23999977 |
| 7007744171 | 26 | 20067 | 14.30000019 |
| 7086361926 | 31 | 14560 | 10.78999996 |
| 8053475328 | 31 | 22988 | 17.95000076 |
| 8253242879 | 19 | 11268 | 8.56000042 |
| 8378563200 | 31 | 16208 | 12.85000038 |
| 8583815059 | 31 | 15168 | 11.82999992 |
| 8792009665 | 29 | 8360 | 5.349999905 |
| 8877689391 | 31 | 29326 | 26.71999931 |

Here, we can see the maximum steps and maximun distance by each person.

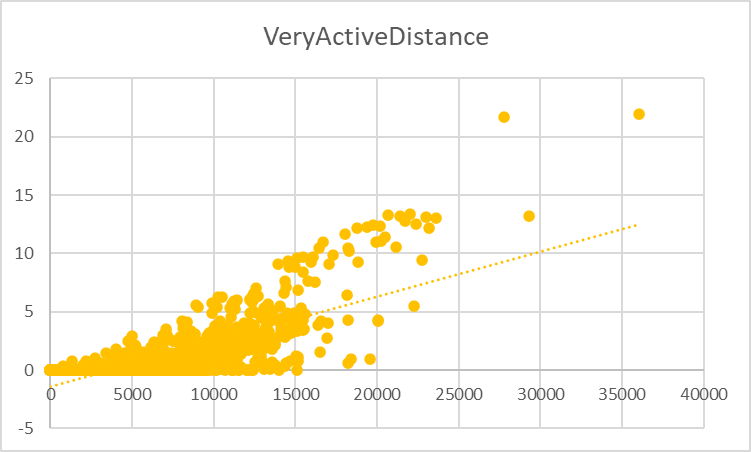
# PHASE 5 : SHARE

#Using the total steps and total distance by scatter plot I found this.



In the above scatter plot it showing that the number of steps is correlated to the number of distance. It seems that every person reaching the 10 to 15 kilometers.

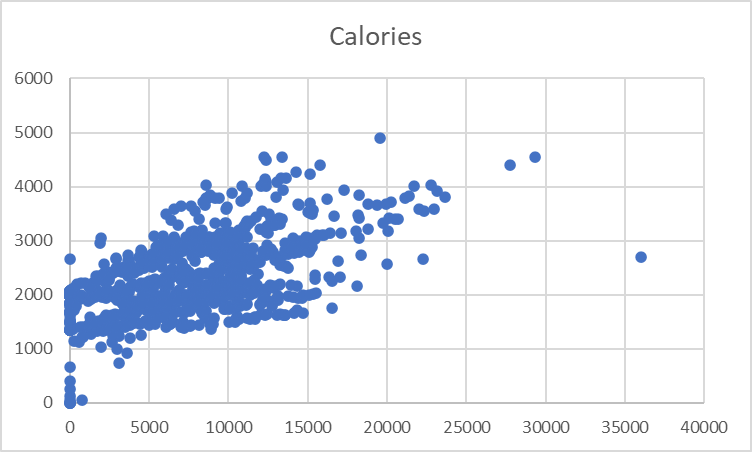
#Using the total steps and very higly active distance by scatter plot, I found this.



Here, we can see the total steps and total active distance is not much corelated by the scatter diagram.

#Using the total steps and calories

After the scatter plot visualization, I found that there is no corelated for the total steps and calories.



# PHASE 6 : ACT

My top three recomendations are:

1. The app should encourage to notify upto 9000 steps a day, then it be good for their health.
2. Remember the user at holidays to walk more and drink water to get more energy.
3. Remember users to sleep atleast 6 hours a day.

**THANK YOU**