Bellabeat Case Study

Faroye Olumide Segun

13/03/2023

INTRODUCTION

Bellabeat is a high-tech company founded in 2013 by Urška Sršen and Sando Mur. They are a manufacturer of health-focused smart products for women. By collecting data on activity, sleep, stress and reproductive health, Bellabeat empowered women with knowledge about their health and habits towards it. Bellabeat is a successful small company , a high-tech manufacturer of health-focused
products for women.
As a junior data analyst working on the marketing analyst team at Bellabeat I was ask to analyse smart device usage data in order to gain insight into how consumers use non-Bellabeat smart devices. I was also asked to select one Bellabeat product to apply these insights to in my presentation.

Bellabeat Products are:

- 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.

Ask

The task is to analyse smart device usage data in order to gain insight into how consumers use non-Bellabeat smart devices, then wants me to select one Bellabeat product to apply these insights into it in my presentation.

Guiding questions for 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?

Key Stakeholders

- **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, analysing, and reporting data that helps guide Bellabeat's marketing strategy.

PREPARE

The public data that explores smart device users' daily habits used for this analysis is FitBit Fitness Tracker Data .
The data set contains personal fitness tracker from thirty fitbit users. Thirty-three 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. The dataset contained 18 CSV documents including minutes-level output for physical activity, heart rate, and sleep monitoring which are stored in long formats. The limitations of the data are its sample size and the absence of descriptions of the users, such as age and gender.

PROCESS

Questions:

(1) Where is your data stored?
The data was stored in a CSV file.

(2) How is the data organized? Is it in long or wide format?

The data was organized in a long format

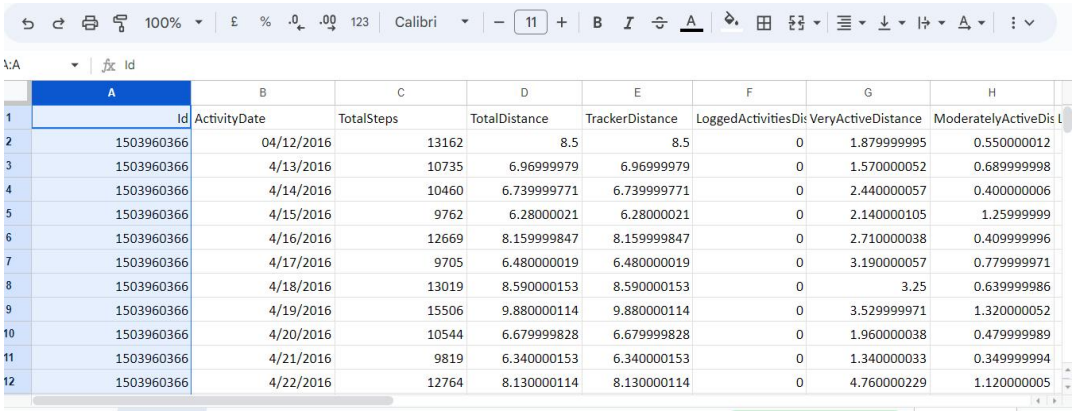
(3) Are there issues with bias or credibility in this data?

Yes the data is biased, all necessary data were not collected from all the thirty fitbit users on SleepDay, MinuteStepsNarrow, minuteMENTsNarrow

Loading packages in google sheets and cleaning data.

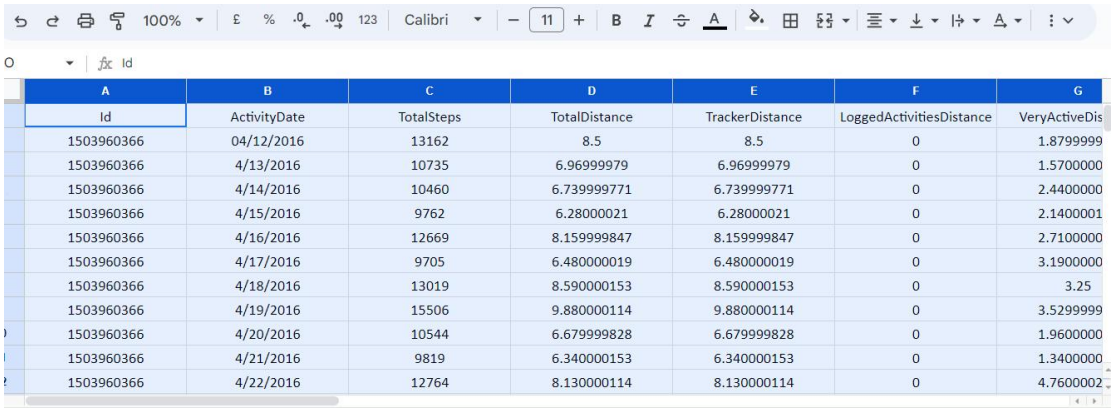
I changed data that ought to be numbers but stored in text to numbers and make sure the datas in each cells are accurate using CountIF, LEN e.t.c.

Before



	A	B	C	D	E	F	G	H
1	Id	ActivityDate	TotalSteps	TotalDistance	TrackerDistance	LoggedActivitiesDis	VeryActiveDistance	ModeratelyActiveDis
2	1503960366	04/12/2016	13162	8.5	8.5	0	1.879999995	0.550000012
3	1503960366	4/13/2016	10735	6.96999979	6.96999979	0	1.570000052	0.689999998
4	1503960366	4/14/2016	10460	6.739999771	6.739999771	0	2.440000057	0.400000006
5	1503960366	4/15/2016	9762	6.28000021	6.28000021	0	2.140000105	1.25999999
6	1503960366	4/16/2016	12669	8.159999847	8.159999847	0	2.710000038	0.409999996
7	1503960366	4/17/2016	9705	6.480000019	6.480000019	0	3.190000057	0.779999971
8	1503960366	4/18/2016	13019	8.590000153	8.590000153	0	3.25	0.639999986
9	1503960366	4/19/2016	15506	9.880000114	9.880000114	0	3.529999971	1.320000052
10	1503960366	4/20/2016	10544	6.679999828	6.679999828	0	1.960000038	0.479999989
11	1503960366	4/21/2016	9819	6.340000153	6.340000153	0	1.340000033	0.349999994
12	1503960366	4/22/2016	12764	8.130000114	8.130000114	0	4.760000229	1.120000005

After



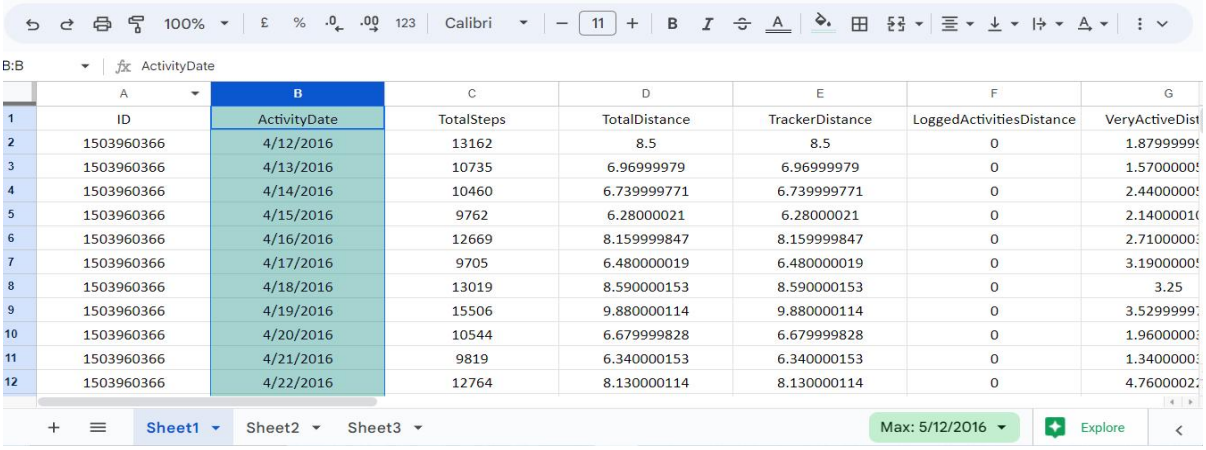
	A	B	C	D	E	F	G
	Id	ActivityDate	TotalSteps	TotalDistance	TrackerDistance	LoggedActivitiesDistance	VeryActiveDis
	1503960366	04/12/2016	13162	8.5	8.5	0	1.8799999
	1503960366	4/13/2016	10735	6.96999979	6.96999979	0	1.5700000
	1503960366	4/14/2016	10460	6.739999771	6.739999771	0	2.4400000
	1503960366	4/15/2016	9762	6.28000021	6.28000021	0	2.1400001
	1503960366	4/16/2016	12669	8.159999847	8.159999847	0	2.7100000
	1503960366	4/17/2016	9705	6.480000019	6.480000019	0	3.1900000
	1503960366	4/18/2016	13019	8.590000153	8.590000153	0	3.25
	1503960366	4/19/2016	15506	9.880000114	9.880000114	0	3.5299999
	1503960366	4/20/2016	10544	6.679999828	6.679999828	0	1.9600000
	1503960366	4/21/2016	9819	6.340000153	6.340000153	0	1.3400000
	1503960366	4/22/2016	12764	8.130000114	8.130000114	0	4.7600002

The COUNTUNIQUE function was used to determine the number of unique user IDs in the datasets.

=COUNTUNIQUE(A1:A941) in (dailyActivityId) = 33,
=COUNTUNIQUE(A1:A414) in (SleepdayId) = 24,
=COUNTUNIQUE(A1:A68) in (WeightLogId) = 8,
=COUNTUNIQUE(A1:A68) in (dailyCaloriesId) = 33,
=SUM (C2:C941) in (dailyCalories) to add all the Calories

DailyActivity, DailyCalories and MinuteMETsNarrow will be use later for the analysis, with these we can get the answer to the **smart device usage data we need in order to gain insight into how consumers use non-Bellabeat smart devices.**

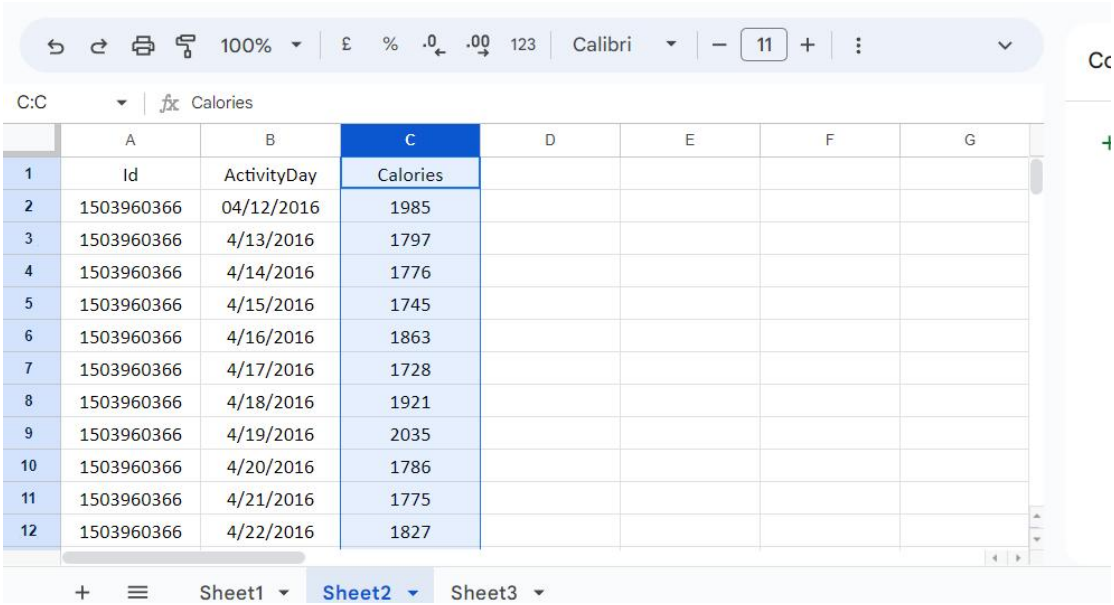
ANALYZE AND SHARE



	A	B	C	D	E	F	G
1	ID	ActivityDate	TotalSteps	TotalDistance	TrackerDistance	LoggedActivitiesDistance	VeryActiveDist
2	1503960366	4/12/2016	13162	8.5	8.5	0	1.8799999
3	1503960366	4/13/2016	10735	6.96999979	6.96999979	0	1.5700000
4	1503960366	4/14/2016	10460	6.739999771	6.739999771	0	2.4400000
5	1503960366	4/15/2016	9762	6.28000021	6.28000021	0	2.1400001
6	1503960366	4/16/2016	12669	8.159999847	8.159999847	0	2.7100000
7	1503960366	4/17/2016	9705	6.480000019	6.480000019	0	3.1900000
8	1503960366	4/18/2016	13019	8.590000153	8.590000153	0	3.25
9	1503960366	4/19/2016	15506	9.880000114	9.880000114	0	3.5299999
10	1503960366	4/20/2016	10544	6.679999828	6.679999828	0	1.9600000
11	1503960366	4/21/2016	9819	6.340000153	6.340000153	0	1.3400000
12	1503960366	4/22/2016	12764	8.130000114	8.130000114	0	4.7600002

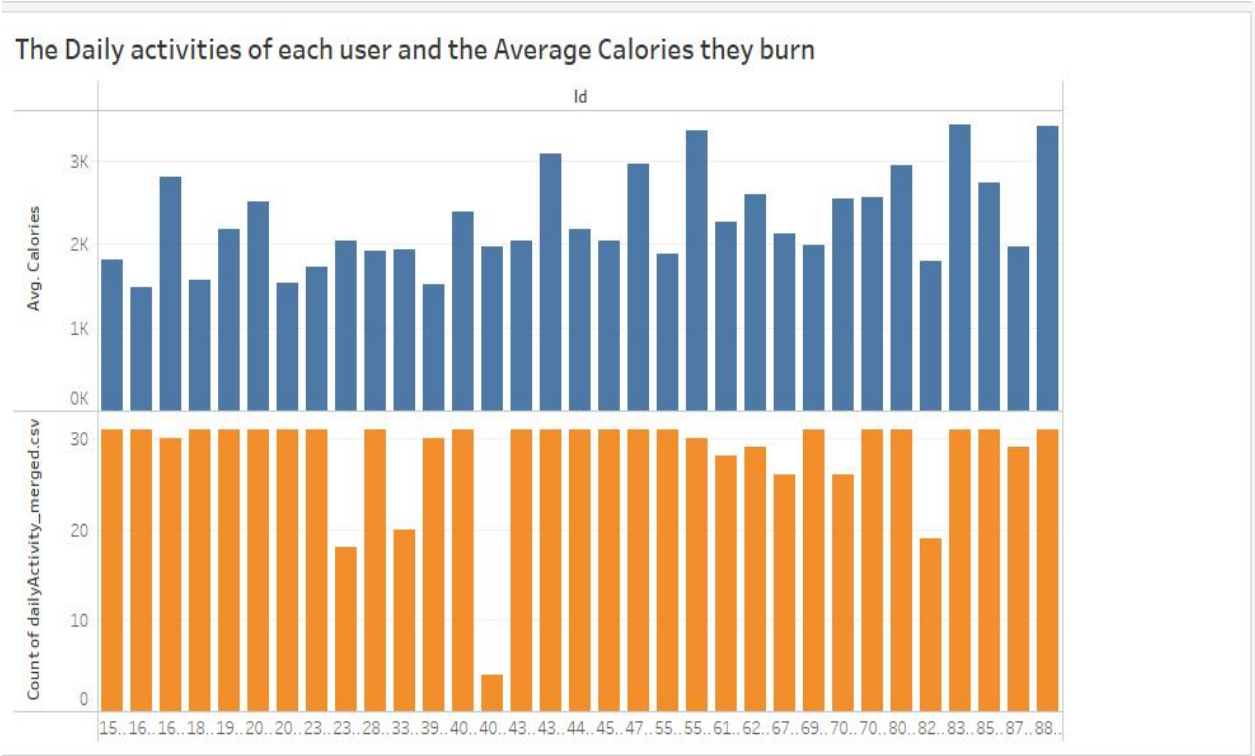
Sheet1 Sheet2 Sheet3 Max: 5/12/2016 Explore

Knowing the start date and the last date which is 4/12/2016 and ends on 5/12/2016

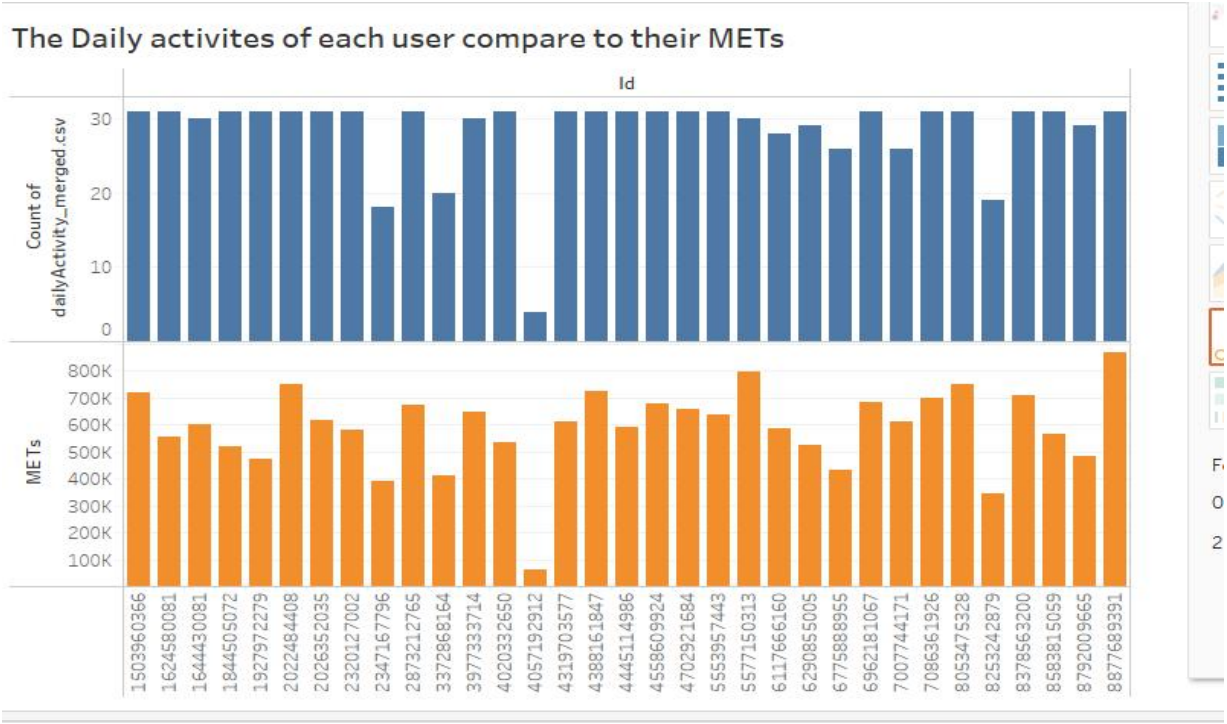


	A	B	C	D	E	F	G
1	Id	ActivityDay	Calories				
2	1503960366	04/12/2016	1985				
3	1503960366	4/13/2016	1797				
4	1503960366	4/14/2016	1776				
5	1503960366	4/15/2016	1745				
6	1503960366	4/16/2016	1863				
7	1503960366	4/17/2016	1728				
8	1503960366	4/18/2016	1921				
9	1503960366	4/19/2016	2035				
10	1503960366	4/20/2016	1786				
11	1503960366	4/21/2016	1775				
12	1503960366	4/22/2016	1827				

Comparing the daily activities of each user with the average Calories been burned

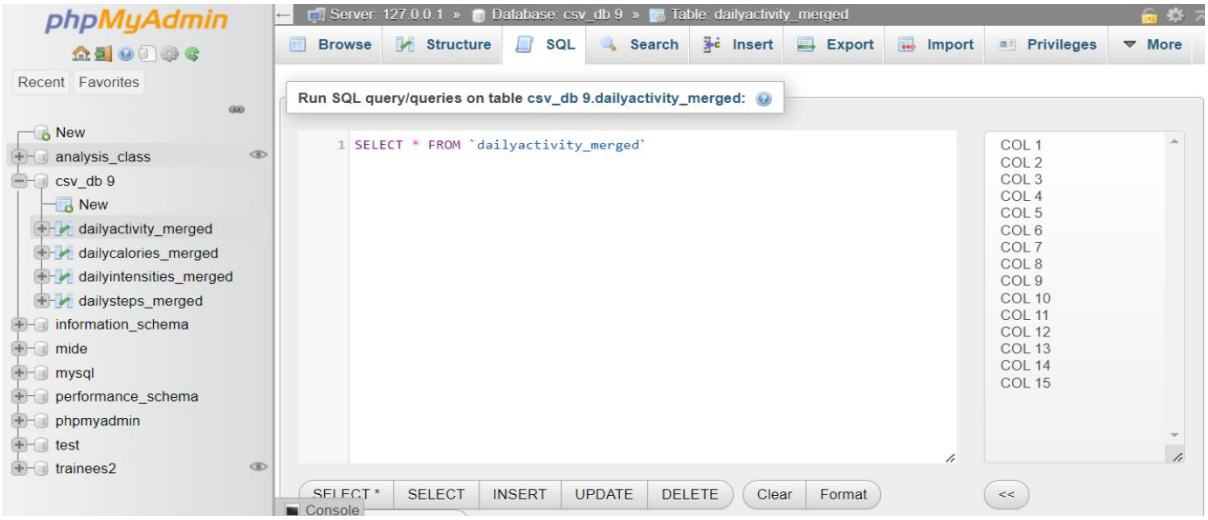


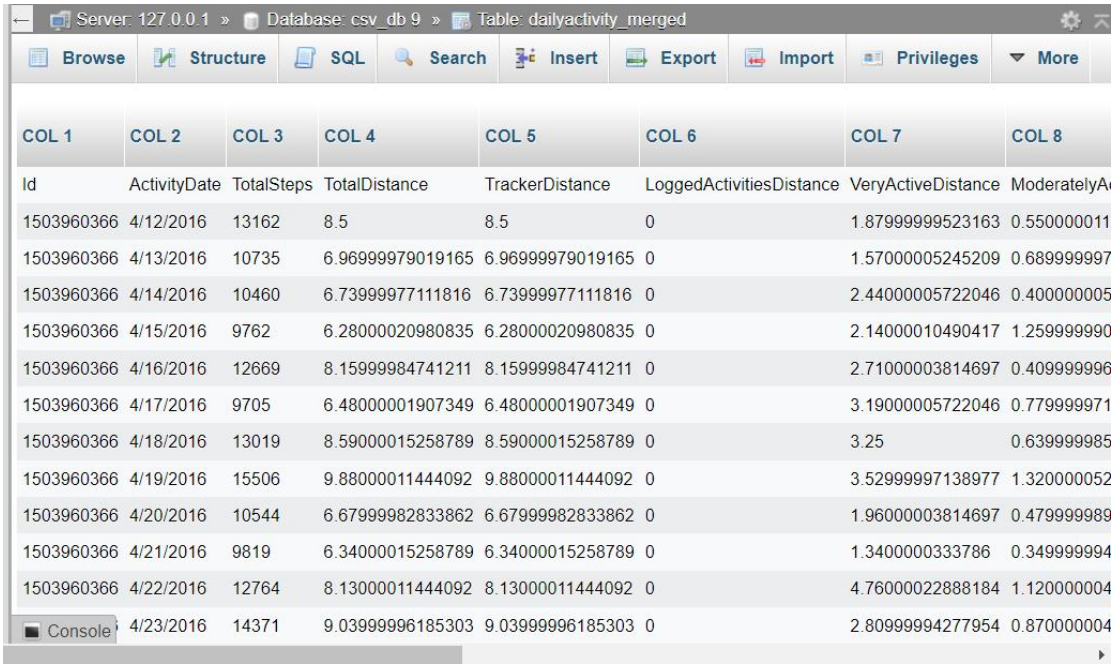
Comparing the daily activities of each user with the total METs perminutes.



Using SQL:

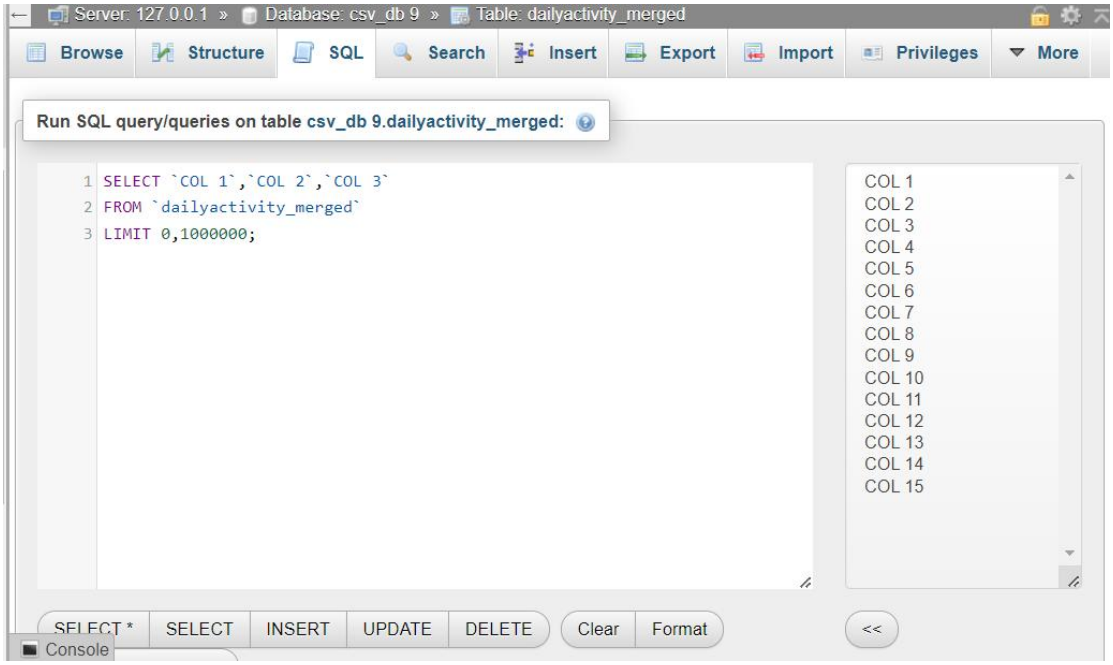
Importing data into SQL database

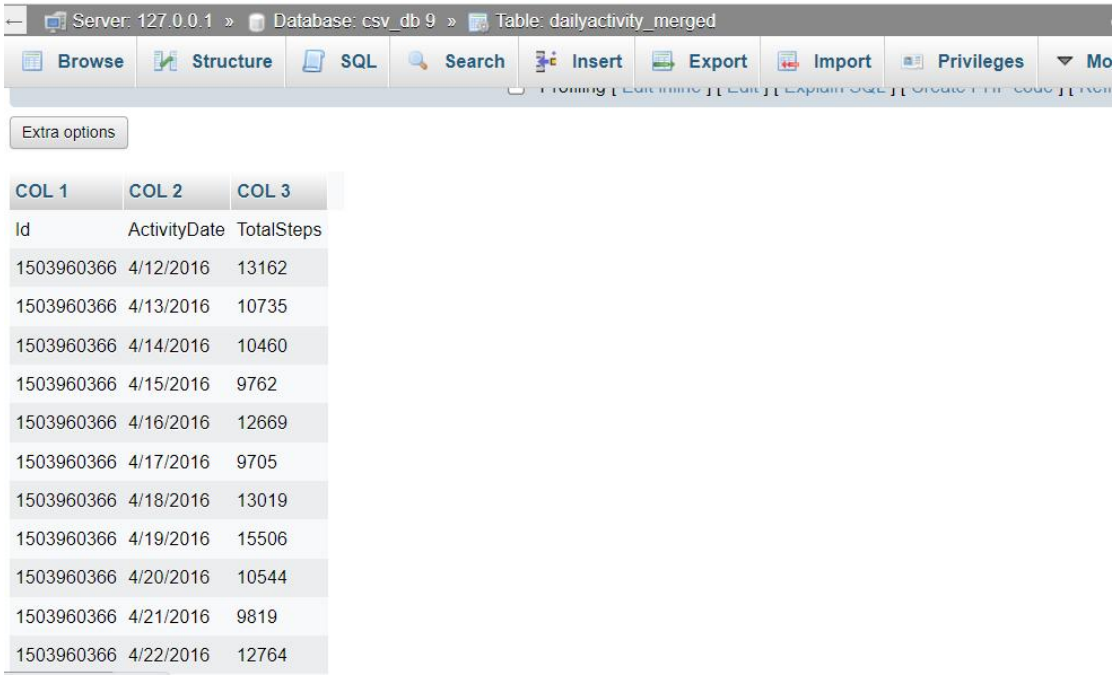




COL 1	COL 2	COL 3	COL 4	COL 5	COL 6	COL 7	COL 8
Id	ActivityDate	TotalSteps	TotalDistance	TrackerDistance	LoggedActivitiesDistance	VeryActiveDistance	ModeratelyA
1503960366	4/12/2016	13162	8.5	8.5	0	1.87999999523163	0.550000011
1503960366	4/13/2016	10735	6.96999979019165	6.96999979019165	0	1.57000005245209	0.689999997
1503960366	4/14/2016	10460	6.73999977111816	6.73999977111816	0	2.44000005722046	0.400000005
1503960366	4/15/2016	9762	6.28000020980835	6.28000020980835	0	2.14000010490417	1.259999990
1503960366	4/16/2016	12669	8.15999984741211	8.15999984741211	0	2.71000003814697	0.409999996
1503960366	4/17/2016	9705	6.48000001907349	6.48000001907349	0	3.19000005722046	0.779999971
1503960366	4/18/2016	13019	8.59000015258789	8.59000015258789	0	3.25	0.639999985
1503960366	4/19/2016	15506	9.88000011444092	9.88000011444092	0	3.52999997138977	1.320000052
1503960366	4/20/2016	10544	6.67999982833862	6.67999982833862	0	1.96000003814697	0.479999989
1503960366	4/21/2016	9819	6.34000015258789	6.34000015258789	0	1.3400000333786	0.349999994
1503960366	4/22/2016	12764	8.13000011444092	8.13000011444092	0	4.76000022888184	1.120000004
1503960366	4/23/2016	14371	9.03999996185303	9.03999996185303	0	2.80999994277954	0.870000004

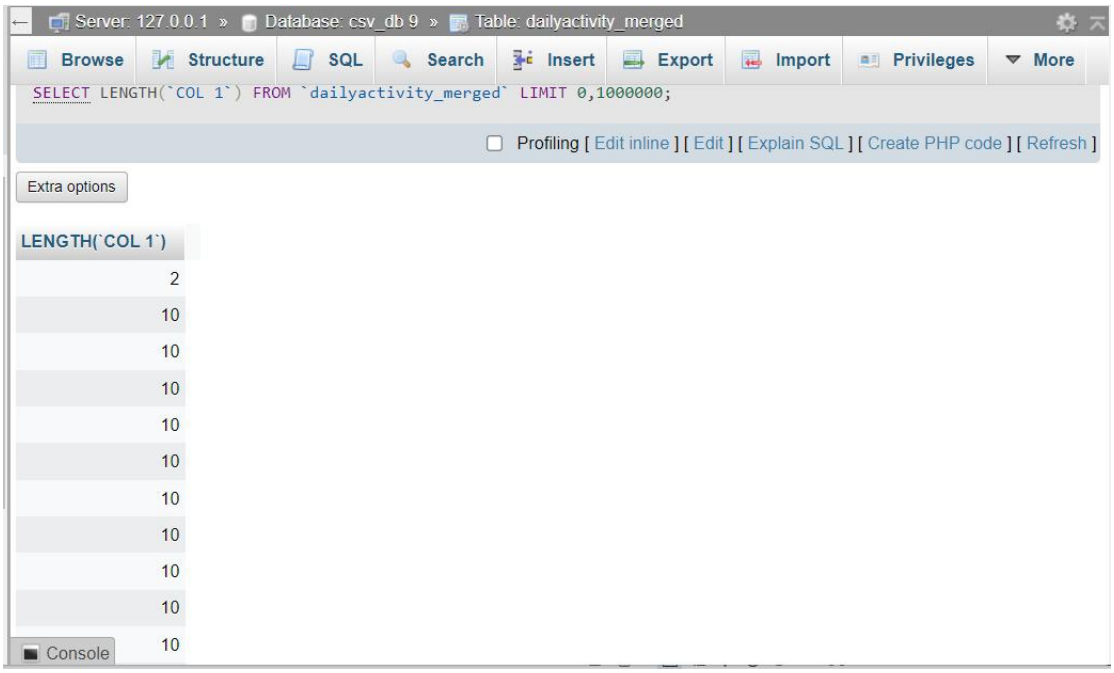
Selecting ID,ActivityDate and TotalSteps from DailyActivity datasets





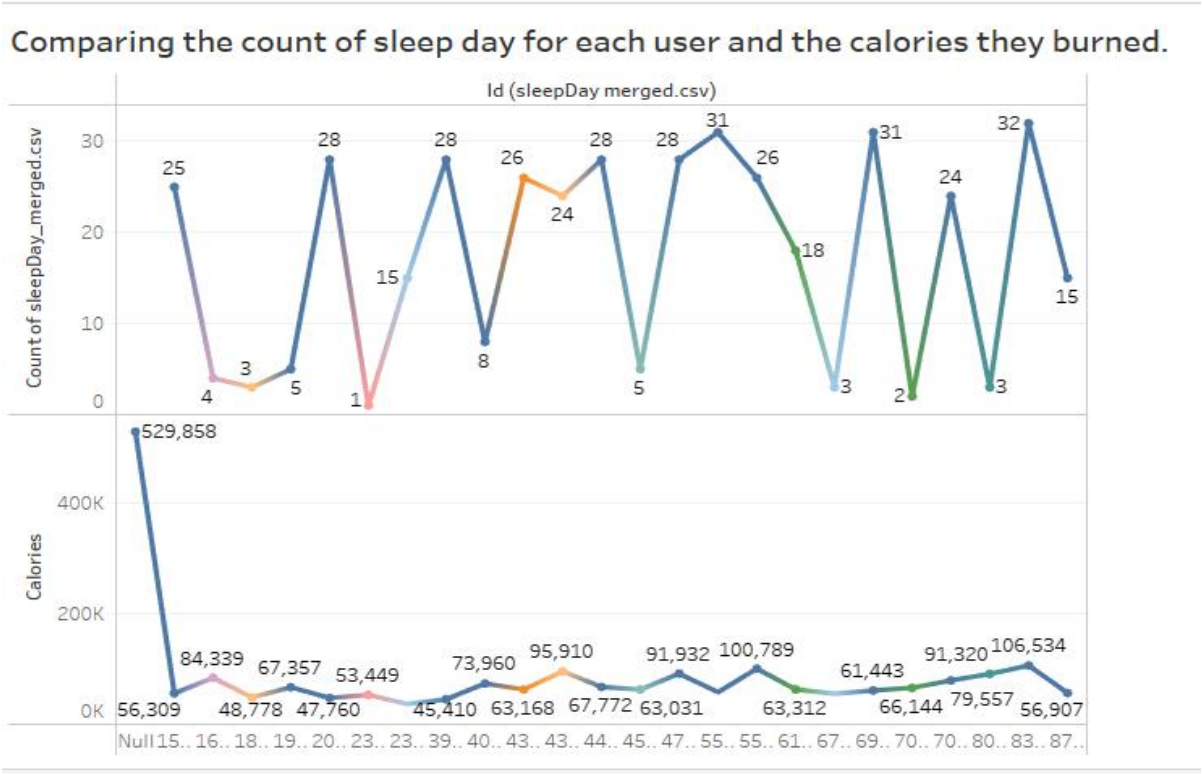
COL 1	COL 2	COL 3
Id	ActivityDate	TotalSteps
1503960366	4/12/2016	13162
1503960366	4/13/2016	10735
1503960366	4/14/2016	10460
1503960366	4/15/2016	9762
1503960366	4/16/2016	12669
1503960366	4/17/2016	9705
1503960366	4/18/2016	13019
1503960366	4/19/2016	15506
1503960366	4/20/2016	10544
1503960366	4/21/2016	9819
1503960366	4/22/2016	12764

SELECT LENGTH(`COL 1`)
FROM `dailyactivity_merged`
LIMIT 0,1000000; shows the length of each user ID



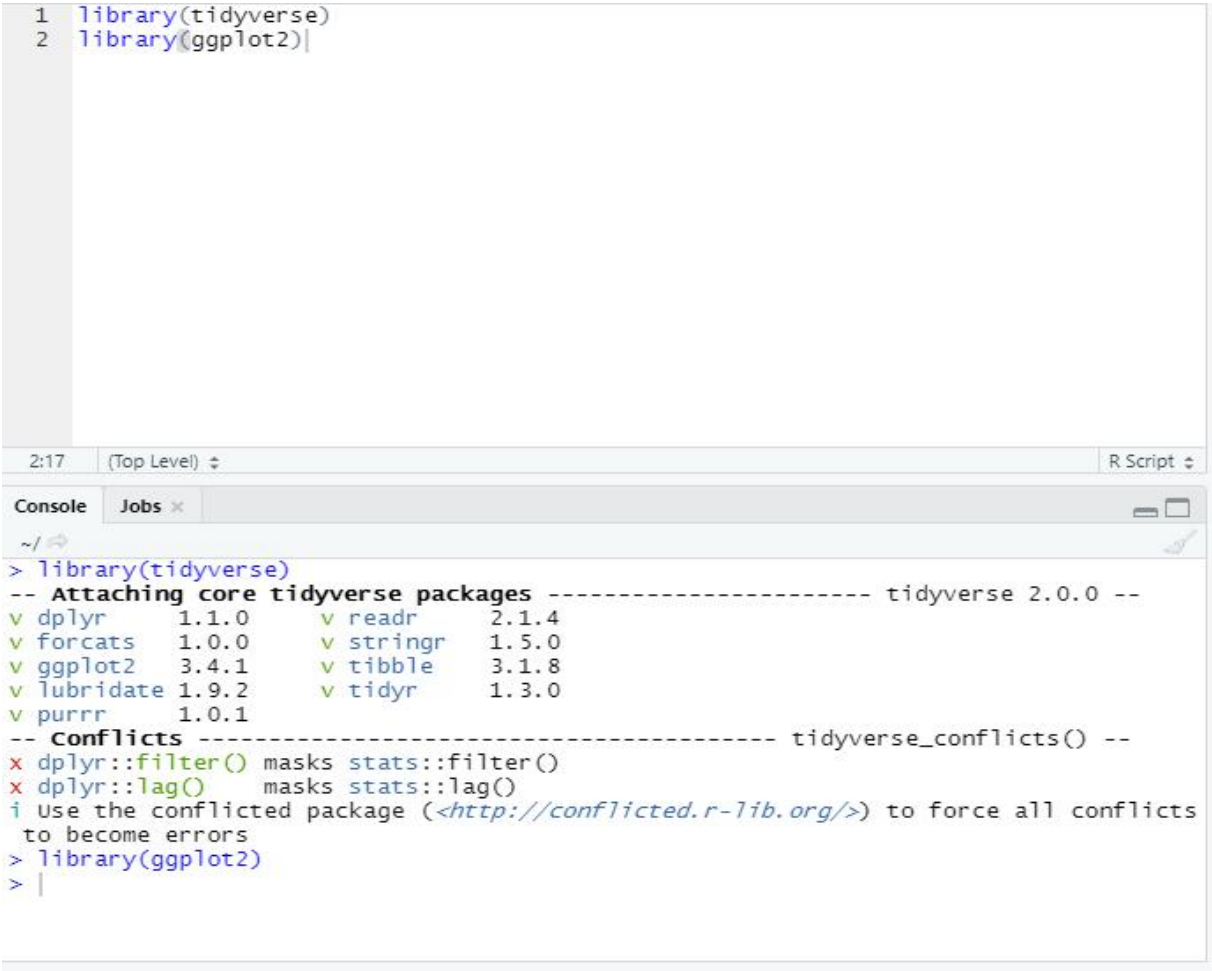
Used Cast() to covert text and string to numbers

Comparing the count of sleep day for each user and the Calories burned



USING R:

Loading R tidyverse and ggplots2 packages in Rstudio



Importing Datasets into R

```r
library(tidyverse)
library(ggplot2)
dailyActivity <- read.csv("dailyActivity_merged.csv")
dailyCalory <- read.csv("dailyCalories_merged.csv")
MinuteMETsNarrow <- read.csv("MinuteMETsNarrow_merged.csv")
```

5:60 (Top Level) R Script

Str(DailyActivity Datasets)

```r
library(tidyverse)
library(ggplot2)
view(dailyActivity_merged)
dailyActivity <- read_csv("wellness.csv/dailyActivity_merged.csv")
dailyCalories <- read_csv("dailyCalories_merged.csv.")
MinuteMETsNarrow <- read.csv("MinuteMETsNarrow_merged.csv")
dailyActivity
```

7:1 (Top Level) R Script

Console Jobs

~/

```
> dailyActivity
# A tibble: 940 x 15
           Id Activit~1 Total~2 Total~3 Track~4 Logge~5 VeryA~6 Moder~7 Light~8
        <dbl> <chr>       <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
 1 1503960366 4/12/2016   13162     8.5     8.5       0    1.88   0.550    6.06
 2 1503960366 4/13/2016   10735    6.97    6.97       0    1.57   0.690    4.71
 3 1503960366 4/14/2016   10460    6.74    6.74       0    2.44   0.400    3.91
 4 1503960366 4/15/2016    9762    6.28    6.28       0    2.14   1.26     2.83
 5 1503960366 4/16/2016   12669    8.16    8.16       0    2.71   0.410    5.04
 6 1503960366 4/17/2016    9705    6.48    6.48       0    3.19   0.780    2.51
 7 1503960366 4/18/2016   13019    8.59    8.59       0    3.25   0.640    4.71
 8 1503960366 4/19/2016   15506    9.88    9.88       0    3.53   1.32     5.03
 9 1503960366 4/20/2016   10544    6.68    6.68       0    1.96   0.480    4.24
10 1503960366 4/21/2016    9819    6.34    6.34       0    1.34   0.350    4.65
# ... with 930 more rows, 6 more variables: SedentaryActiveDistance <dbl>,
#   VeryActiveMinutes <dbl>, FairlyActiveMinutes <dbl>,
#   LightlyActiveMinutes <dbl>, SedentaryMinutes <dbl>, Calories <dbl>, and
```

Str(DailyCalories)

```
1  library(tidyverse)
2  library(ggplot2)
3  view(dailyActivity_merged)
4  dailyActivity <- read_csv("wellness.csv/dailyActivity_merged.csv")
5  dailyCalories <- read_csv("wellness.csv/dailyCalories_merged.csv.")
6  MinuteMETsNarrow <- read.csv("wellness.csv/MinuteMETsNarrow_merged.csv")
7  dailyActivity
8  dailyCalories
```

```
# A tibble: 940 x 3
           Id ActivityDay Calories
        <dbl> <chr>           <dbl>
 1 1503960366 4/12/2016        1985
 2 1503960366 4/13/2016        1797
 3 1503960366 4/14/2016        1776
 4 1503960366 4/15/2016        1745
 5 1503960366 4/16/2016        1863
 6 1503960366 4/17/2016        1728
 7 1503960366 4/18/2016        1921
 8 1503960366 4/19/2016        2035
 9 1503960366 4/20/2016        1786
10 1503960366 4/21/2016        1775
# ... with 930 more rows
# i Use `print(n = ...)` to see more rows
> |
```

Str(MinutesMETsNarrow)

```
1  library(tidyverse)
2  library(ggplot2)
3  view(dailyActivity_merged)
4  dailyActivity <- read_csv("wellness.csv/dailyActivity_merged.csv")
5  dailyCalories <- read_csv("wellness.csv/dailyCalories_merged.csv.")
6  MinuteMETsNarrow <- read.csv("wellness.csv/MinuteMETsNarrow_merged.csv")
7  dailyActivity
8  dailyCalories
9  MinuteMETsNarrow
10
```

```
# i Use `print(n = ...)` to see more rows
> MinuteMETsNarrow <- read.csv("wellness.csv/MinuteMETsNarrow_merged.csv")
> MinuteMETsNarrow
           Id          ActivityMinute METs
1   1503960366 4/12/2016 12:00:00 AM   10
2   1503960366 4/12/2016 12:01:00 AM   10
3   1503960366 4/12/2016 12:02:00 AM   10
4   1503960366 4/12/2016 12:03:00 AM   10
5   1503960366 4/12/2016 12:04:00 AM   10
6   1503960366 4/12/2016 12:05:00 AM   12
7   1503960366 4/12/2016 12:06:00 AM   12
8   1503960366 4/12/2016 12:07:00 AM   12
9   1503960366 4/12/2016 12:08:00 AM   12
10  1503960366 4/12/2016 12:09:00 AM   12
11  1503960366 4/12/2016 12:10:00 AM   12
12  1503960366 4/12/2016 12:11:00 AM   12
```

The n_distinct() function was used to determine the number of unique user IDs in the three datasets

```
10
11  n_distinct(dailyActivity$Id)
12  n_distinct(dailyCalories$Id)
13  n_distinct(MinuteMETsNarrow$Id)
14
```

11:1 (Top Level) R Script

Console Jobs

```
> n_distinct(dailyActivity$Id)
[1] 33
> n_distinct(dailyCalories$Id)
[1] 33
> n_distinct(MinuteMETsNarrow$Id)
[1] 33
>
```

Changing of data type:

The date column in dailyActivity and dailyCalories datasets will be formatted from character data type to date, as they will be used later in the analysis.

```
15  dailyActivity <- dailyActivity%>%
16    mutate(ActivityDate = as.Date(ActivityDate,format="%m/%d/%y"))
17  str(dailyActivity)
18  dailyActivity
```

18:14 (Top Level) R Script

Console Jobs

```
> dailyActivity <- dailyActivity%>%
+   mutate(ActivityDate = as.Date(ActivityDate,format="%m/%d/%y"))
> str(dailyActivity)
tibble [940 x 15] (S3: tbl_df/tbl/data.frame)
 $ Id                       : num [1:940] 1.5e+09 1.5e+09 1.5e+09 1.5e+09 1.5e+09
 ...
 $ ActivityDate             : Date[1:940], format: "2020-12-04" NA ...
 $ TotalSteps               : num [1:940] 13162 10735 10460 9762 12669 ...
 $ TotalDistance            : num [1:940] 8.5 6.97 6.74 6.28 8.16 ...
 $ TrackerDistance          : num [1:940] 8.5 6.97 6.74 6.28 8.16 ...
 $ LoggedActivitiesDistance: num [1:940] 0 0 0 0 0 0 0 0 0 0 ...
 $ VeryActiveDistance       : num [1:940] 1.88 1.57 2.44 2.14 2.71 ...
 $ ModeratelyActiveDistance: num [1:940] 0.55 0.69 0.4 1.26 0.41 ...
 $ LightActiveDistance      : num [1:940] 6.06 4.71 3.91 2.83 5.04 ...
 $ SedentaryActiveDistance : num [1:940] 0 0 0 0 0 0 0 0 0 0 ...
 $ VeryActiveMinutes        : num [1:940] 25 21 30 29 36 38 42 50 28 19 ...
 $ FairlyActiveMinutes      : num [1:940] 13 19 11 34 10 20 16 31 12 8 ...
```

```
21  dailyCalories <- dailyActivity%>%
22    mutate(ActivityDay = as.Date(ActivityDate,format="%m/%d/%y"))
23  str(dailyCalories)
```

21:1 (Top Level) R Script

Console Jobs

```
 - attr(*, "problems")=<externalptr>
> dailyCalories <- dailyActivity%>%
+   mutate(ActivityDay = as.Date(ActivityDate,format="%m/%d/%y"))
> str(dailyCalories)
tibble [940 x 16] (S3: tbl_df/tbl/data.frame)
 $ Id                       : num [1:940] 1.5e+09 1.5e+09 1.5e+09 1.5e+09 1.5e+09
 ...
 $ ActivityDate             : Date[1:940], format: "2020-12-04" NA ...
 $ TotalSteps               : num [1:940] 13162 10735 10460 9762 12669 ...
 $ TotalDistance            : num [1:940] 8.5 6.97 6.74 6.28 8.16 ...
 $ TrackerDistance          : num [1:940] 8.5 6.97 6.74 6.28 8.16 ...
 $ LoggedActivitiesDistance: num [1:940] 0 0 0 0 0 0 0 0 0 0 ...
 $ VeryActiveDistance       : num [1:940] 1.88 1.57 2.44 2.14 2.71 ...
 $ ModeratelyActiveDistance: num [1:940] 0.55 0.69 0.4 1.26 0.41 ...
 $ LightActiveDistance      : num [1:940] 6.06 4.71 3.91 2.83 5.04 ...
 $ SedentaryActiveDistance : num [1:940] 0 0 0 0 0 0 0 0 0 0 ...
```

ANALYZE AND SHARE

Data summaries

Daily Activity

```
24
25  dailyActivity %>%
26    select(Id,TotalSteps,TotalDistance,VeryActiveMinutes,
27           FairlyActiveMinutes,LightlyActiveMinutes,Calories)%>%
28  summary()
29
```

28:10 (Top Level) R Script

Console Jobs

```
+ summary()
       Id              TotalSteps    TotalDistance    VeryActiveMinutes
 Min.   :1.504e+09   Min.   :    0   Min.   : 0.000   Min.   :  0.00
 1st Qu.:2.320e+09   1st Qu.: 3790   1st Qu.: 2.620   1st Qu.:  0.00
 Median :4.445e+09   Median : 7406   Median : 5.245   Median :  4.00
 Mean   :4.855e+09   Mean   : 7638   Mean   : 5.490   Mean   : 21.16
 3rd Qu.:6.962e+09   3rd Qu.:10727   3rd Qu.: 7.713   3rd Qu.: 32.00
 Max.   :8.878e+09   Max.   :36019   Max.   :28.030   Max.   :210.00
 FairlyActiveMinutes LightlyActiveMinutes    Calories
 Min.   :  0.00      Min.   :  0.0        Min.   :   0
 1st Qu.:  0.00      1st Qu.:127.0        1st Qu.:1828
 Median :  6.00      Median :199.0        Median :2134
 Mean   : 13.56      Mean   :192.8        Mean   :2304
 3rd Qu.: 19.00      3rd Qu.:264.0        3rd Qu.:2793
 Max.   :143.00      Max.   :518.0        Max.   :4900
>
```

The summary shows the max totalSteps,max Distance covered and the max activeMinutes.

```
14
15  dailyCalories %>%
16    select(Id,ActivityDay,Calories)%>%
17    summary()
```

15:1 (Top Level) R Script

Console Jobs

~/

```
Error in `select()`:
! Can't subset columns with `dailyActivity`.
x `dailyActivity` must be numeric or character, not a <spec_tbl_df/tbl_df/tbl/dat
a.frame> object.
Run `rlang::last_error()` to see where the error occurred.
> dailyCalories %>%
+   select(Id,ActivityDay,Calories)%>%
+   summary()
       Id              ActivityDay          Calories
 Min.   :1.504e+09   Length:940         Min.   :   0
 1st Qu.:2.320e+09   Class :character   1st Qu.:1828
 Median :4.445e+09   Mode  :character   Median :2134
 Mean   :4.855e+09                      Mean   :2304
 3rd Qu.:6.962e+09                      3rd Qu.:2793
 Max.   :8.878e+09                      Max.   :4900
> |
```

The summary shows the max of calories, and the length of activityDay which is 940

```
19
20  MinuteMETsNarrow %>%
21    select(Id,ActivityMinute,METs)%>%
22    summary()
```

18:1 (Top Level) R Script

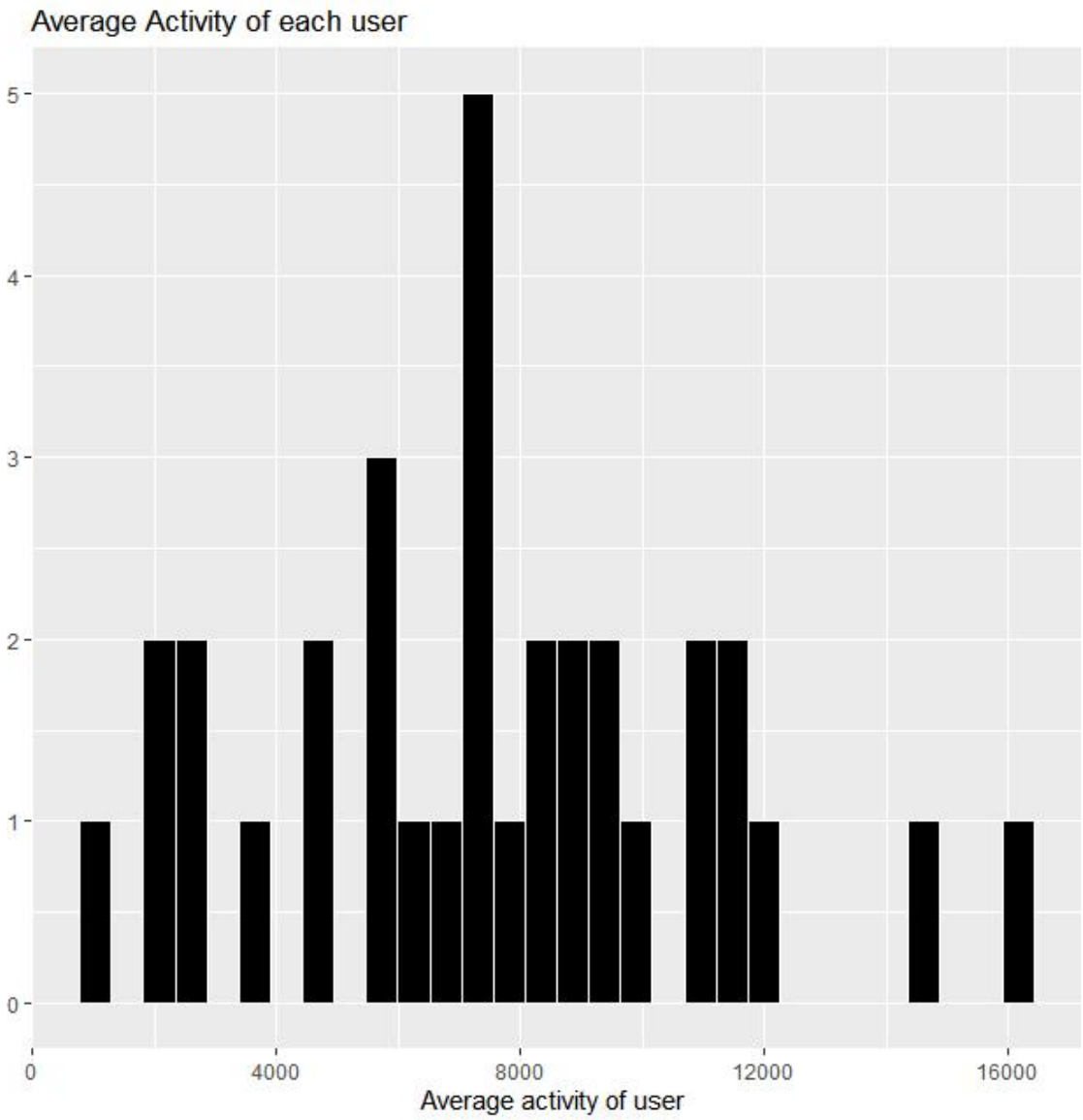
Console Jobs

~/

```
> MinuteMETsNarrow %>%
+   select(Id,ActivityMinute,METs)%>%
+   summary()
       Id              ActivityMinute          METs
 Min.   :1.504e+09   Length:1325580     Min.   :  0.00
 1st Qu.:2.320e+09   Class :character   1st Qu.: 10.00
 Median :4.445e+09   Mode  :character   Median : 10.00
 Mean   :4.848e+09                      Mean   : 14.69
 3rd Qu.:6.962e+09                      3rd Qu.: 11.00
 Max.   :8.878e+09                      Max.   :157.00
> |
```

Histogram plot of average activity for each user

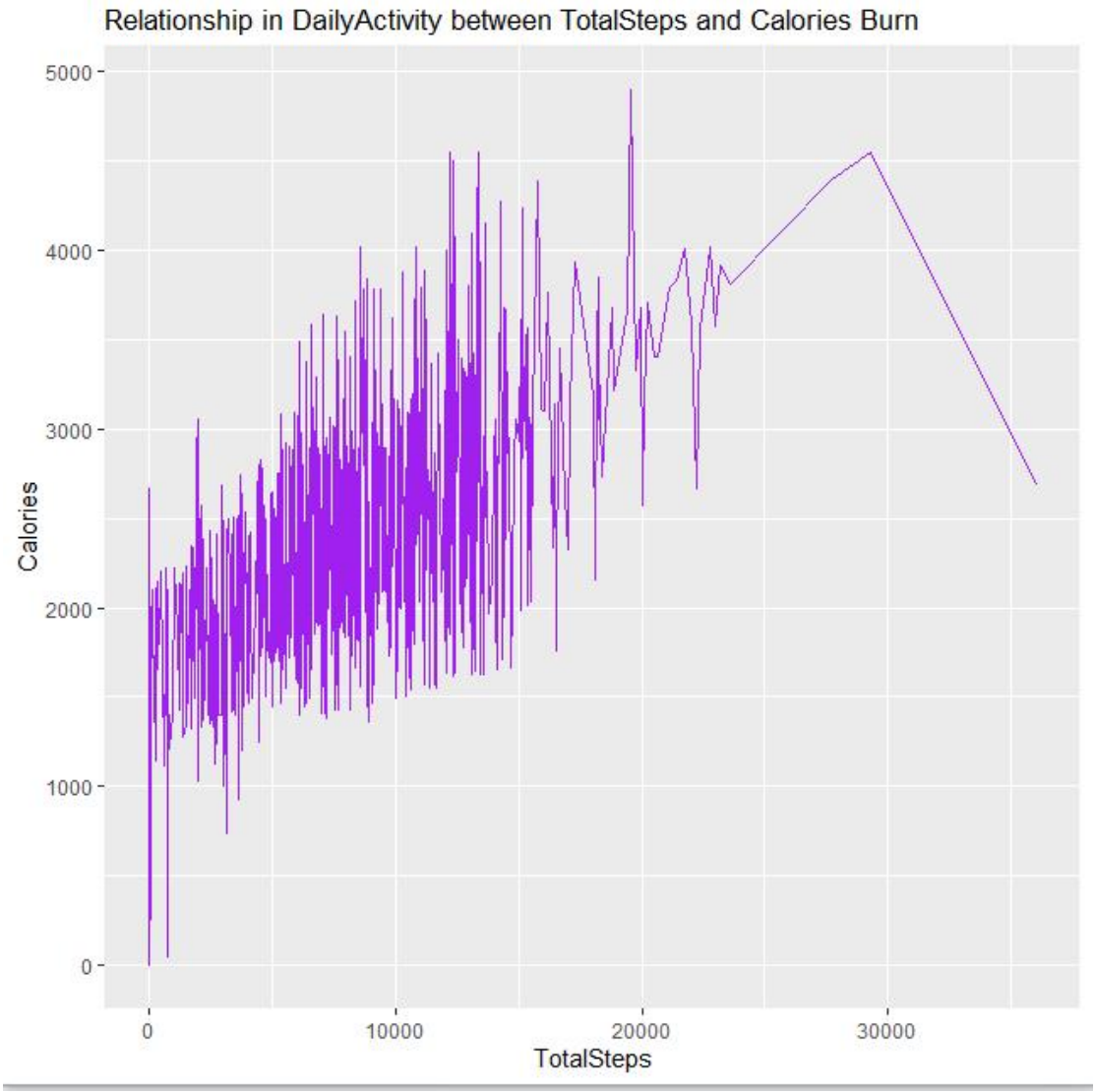
```
userAvgDailyActivity <-dailyActivity %>%
  group_by(Id) %>%
  summarise(AverageDailyActivity = mean(TotalSteps))

ggplot(data=userAvgDailyActivity)+
  geom_histogram(mapping = aes(x=AverageDailyActivity),color='white', fill='b
            labs(title = 'Average Activity of each user',
                 x='Average activity of user ')
```



The plot shows that six users have an average activity of less than 4000 while seven users have an average activity of more than 10000. A total of 100000 activities is recommended in a month.

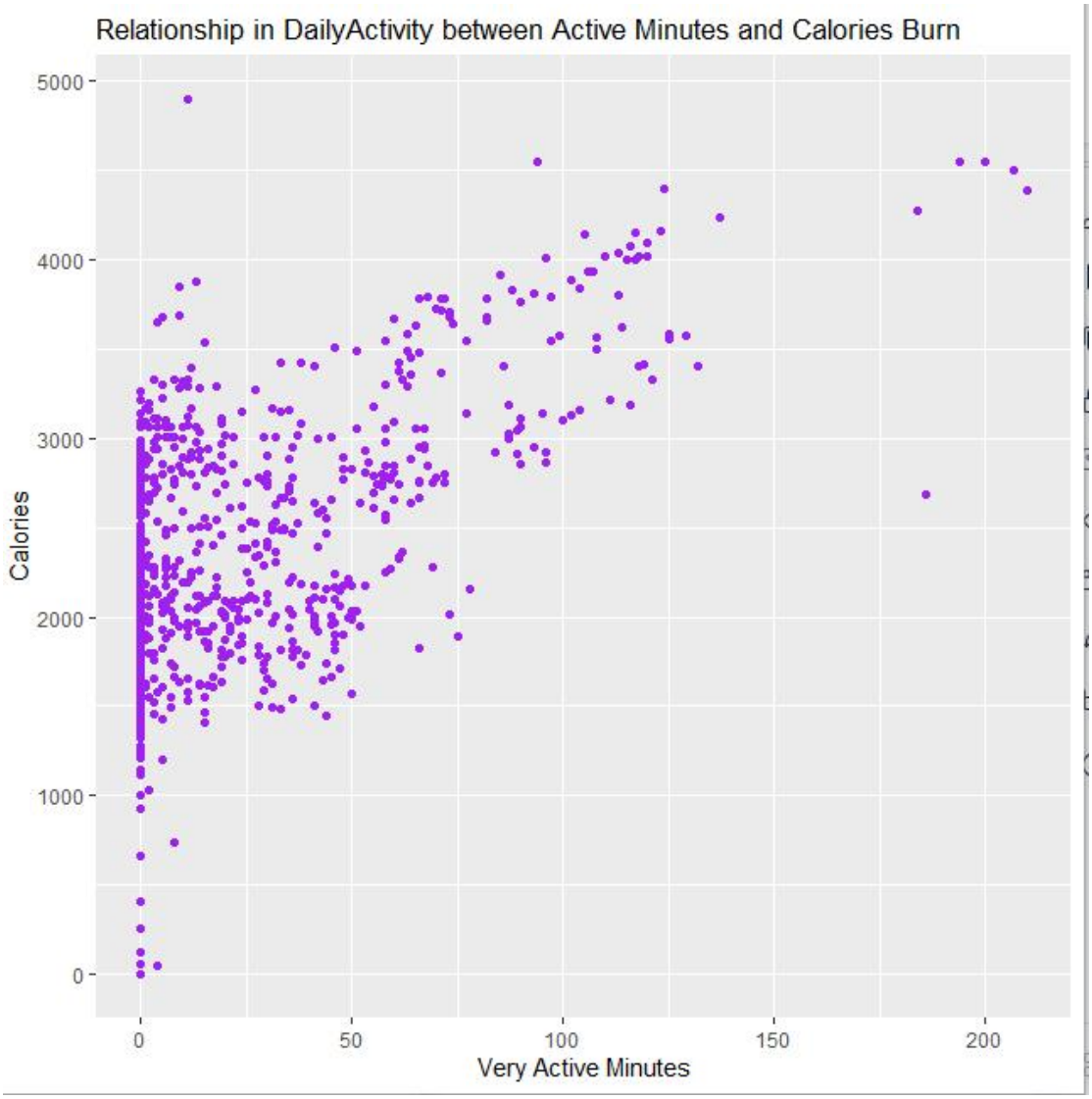
Relationship between total DailyActivities and Calories burned

```
ggplot(data=dailyActivity)+
  geom_line(mapping = aes(x= TotalSteps, y= Calories), color= "purple")+
  labs(title = 'Relationship in DailyActivity between TotalSteps and Calories
       x= 'TotalSteps',
       y= 'Calories')
```



The plot shows a positive relationship between total steps and the calories burned. As the number of steps increased the calories burned increased as well. However there are records of zero step with positive values for calories burned.

```
ggplot(data=dailyActivity)+
  geom_point(mapping = aes(x= VeryActiveMinutes, y= Calories), color= "purple
  labs(title = 'Relationship in DailyActivity between Active Minutes and Calo
       x= 'Very Active Minutes',
       y= 'Calories')
```



There is a visible linear dependency between the total active minutes and calories burned. That is as total active minutes of users increased the calories burned increased as well. However, there were records of calories spent with no active minutes.

ACT

Summary of key findings and recommendations

1. Analysis of the data revealed that users walked an average of fewer than 10000 steps ,I suggest a timed notification for each user encouraging them to increase their daily activity to meet the recommendations of at least 30 minutes of moderate physical activity and 10000 steps every day. Rewards should be

given to user who reach their daily goals, such points can be used to unlock or get discount on premium features on Bellabeat app, such as workout routines and customized meal plans.

2. The analysis carried out showed that users tend to have insufficient sleep,especially on Friday and Tuesdays. I recommend allowing users to set their preferred bedtime with the option of a 30 minutes before that time and to set a preferred time to disable specific notifications. For users whose data show that they often spend more than 30 minutes to 1 hour in bed before sleeping,the Bellabeat app could provide recommendations on good bedtime routines such as light meals, warm baths, listening to music or journalling.

3. Based on the insights from this analysis, digital marketing for Bellabeat can make use of Time Product in order to gain insight into how consumers use non-Bellabeat smart devices. Using DailyActivity of each user the digital marketing can get all the activities of each user which are calories burned,TotalSteps, Activeminutes e.t.c which can be used to track the user activities per day and get insight into consumers who use non-Bellabeat smart devices.

REFERENCES

1. How much sleep do I need?Sleep and sleep disorders. Centres for diseases control and prevention, 14 Sept.2022

2. How many steps should people take per day? Medical News today,20 Jan. 2021

3. How much physical activity do adults need? Centres for disease control and prevention, 2 June 2022.