Bellabeat Google Capstone Report

Ajinkya Desai

2023-02-01

Introduction

The Bellabeat app delivers health-related information on various aspects such as activity, sleep, stress, menstrual cycle, and mindfulness habits, which empowers users to make informed decisions toward leading a healthier lifestyle. In addition, Bellabeat offers smart wellness products such as the Leaf tracker, which can be worn as a bracelet, necklace, or clip and monitors activity, sleep, and stress, and the Time watch, which combines classic design with smart technology to track activity, sleep, and stress. Furthermore, the Spring water bottle, another product from Bellabeat, utilizes smart technology to track daily water intake and can be synced with the Bellabeat app.

Furthermore, Bellabeat's membership program provides users with personalized guidance on nutrition, activity, sleep, health and beauty, and mindfulness based on their individual goals and lifestyle. The company is committed to gaining insights into smart device usage by analyzing FitBit Fitness Tracker Data, identifying growth opportunities, and developing innovative products accordingly.

Bellabeat has prioritized digital marketing, leveraging platforms such as Google Search and social media to reach a larger audience. The co-founder Sršen recognizes the potential for growth that analyzing consumer data holds and has made efforts to harness this information to enhance the company's strategy.

1)Ask: Questions for the analysis:

1. What are some current patterns in the utilization of intelligent devices?
2. In what ways could these patterns be relevant to Bellabeat's clientele?
3. How might these patterns contribute to shaping Bellabeat's marketing approach?

2)Prepare: Installing and loading of packages:

```
install.packages("tidyverse")
install.packages("here")
install.packages("skimr")
install.packages("janitor")
install.packages("dylyr")
install.packages("ggplot2")
```

```
library("tidyverse")
library("here")
library("skimr")
library("janitor")
library("dplyr")
library("ggplot2")
```

Importing datasets:

3)Process: Viewing the data frames:

```
head(daily_activity)
```

```
## # A tibble: 6 × 15
##       Id Activ…¹ Total…² Total…³ Track…⁴ Logge…⁵ VeryA…⁶ Moder…⁷ Light…⁸
Seden…⁹
##    <dbl> <chr>     <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
<dbl>
## 1 1.50e9 4/12/2…   13162    8.5     8.5        0    1.88   0.550    6.06
0
## 2 1.50e9 4/13/2…   10735    6.97    6.97       0    1.57   0.690    4.71
0
## 3 1.50e9 4/14/2…   10460    6.74    6.74       0    2.44   0.400    3.91
0
## 4 1.50e9 4/15/2…    9762    6.28    6.28       0    2.14   1.26     2.83
0
## 5 1.50e9 4/16/2…   12669    8.16    8.16       0    2.71   0.410    5.04
0
## 6 1.50e9 4/17/2…    9705    6.48    6.48       0    3.19   0.780    2.51
0
## # … with 5 more variables: VeryActiveMinutes <dbl>, FairlyActiveMinutes
<dbl>,
## #   LightlyActiveMinutes <dbl>, SedentaryMinutes <dbl>, Calories <dbl>,
and
## #   abbreviated variable names ¹ActivityDate, ²TotalSteps, ³TotalDistance,
## #   ⁴TrackerDistance, ⁵LoggedActivitiesDistance, ⁶VeryActiveDistance,
## #   ⁷ModeratelyActiveDistance, ⁸LightActiveDistance, ⁹
SedentaryActiveDistance
```

```
head(daily_calories)
```

```
## # A tibble: 6 × 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
```

```
head(daily_intensities)
```

```
## # A tibble: 6 × 10
##       Id Activ…¹ Seden…² Light…³ Fairl…⁴ VeryA…⁵ Seden…⁶ Light…⁷ Moder…⁸
VeryA…⁹
##    <dbl> <chr>     <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
<dbl>
## 1 1.50e9 4/12/2…     728     328      13      25       0    6.06   0.550
1.88
## 2 1.50e9 4/13/2…     776     217      19      21       0    4.71   0.690
1.57
## 3 1.50e9 4/14/2…    1218     181      11      30       0    3.91   0.400
2.44
## 4 1.50e9 4/15/2…     726     209      34      29       0    2.83   1.26
2.14
## 5 1.50e9 4/16/2…     773     221      10      36       0    5.04   0.410
2.71
## 6 1.50e9 4/17/2…     539     164      20      38       0    2.51   0.780
3.19
## # … with abbreviated variable names ¹ActivityDay, ²SedentaryMinutes,
## #   ³LightlyActiveMinutes, ⁴FairlyActiveMinutes, ⁵VeryActiveMinutes,
## #   ⁶SedentaryActiveDistance, ⁷LightActiveDistance, ⁸
ModeratelyActiveDistance,
## #   ⁹VeryActiveDistance
```

```
head(daily_steps)
```

```
## # A tibble: 6 × 3
##           Id ActivityDay StepTotal
##        <dbl> <chr>           <dbl>
## 1 1503960366 4/12/2016       13162
## 2 1503960366 4/13/2016       10735
## 3 1503960366 4/14/2016       10460
## 4 1503960366 4/15/2016        9762
## 5 1503960366 4/16/2016       12669
## 6 1503960366 4/17/2016        9705
```

```
head(heartrate_sec)
```

```
## # A tibble: 6 × 3
##           Id Time                 Value
##        <dbl> <chr>                <dbl>
## 1 2022484408 4/12/2016 7:21:00 AM    97
## 2 2022484408 4/12/2016 7:21:05 AM   102
## 3 2022484408 4/12/2016 7:21:10 AM   105
## 4 2022484408 4/12/2016 7:21:20 AM   103
## 5 2022484408 4/12/2016 7:21:25 AM   101
## 6 2022484408 4/12/2016 7:22:05 AM    95
```

```
head(minute_METs)
```

```
## # A tibble: 6 × 3
##           Id ActivityMinute         METs
##        <dbl> <chr>                 <dbl>
## 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
```

```
head(sleep_day)
```

```
## # A tibble: 6 × 5
##           Id SleepDay              TotalSleepRecords TotalMinutesAsleep
TotalT…¹
##        <dbl> <chr>                             <dbl>              <dbl>
<dbl>
## 1 1503960366 4/12/2016 12:00:00 AM                 1                327
346
## 2 1503960366 4/13/2016 12:00:00 AM                 2                384
407
## 3 1503960366 4/15/2016 12:00:00 AM                 1                412
442
## 4 1503960366 4/16/2016 12:00:00 AM                 2                340
367
## 5 1503960366 4/17/2016 12:00:00 AM                 1                700
712
## 6 1503960366 4/19/2016 12:00:00 AM                 1                304
320
## # … with abbreviated variable name ¹TotalTimeInBed
```

```
head(weight_log)
```

```
## # A tibble: 6 × 8
##           Id Date                   WeightKg Weight…¹   Fat   BMI IsMan…²
LogId
##        <dbl> <chr>                     <dbl>    <dbl> <dbl> <dbl> <lgl>
<dbl>
## 1 1503960366 5/2/2016 11:59:59 PM      52.6     116.    22  22.6 TRUE
1.46e12
## 2 1503960366 5/3/2016 11:59:59 PM      52.6     116.    NA  22.6 TRUE
1.46e12
## 3 1927972279 4/13/2016 1:08:52 AM     134.      294.    NA  47.5 FALSE
1.46e12
## 4 2873212765 4/21/2016 11:59:59 PM     56.7     125.    NA  21.5 TRUE
1.46e12
## 5 2873212765 5/12/2016 11:59:59 PM     57.3     126.    NA  21.7 TRUE
1.46e12
## 6 4319703577 4/17/2016 11:59:59 PM     72.4     160.    25  27.5 TRUE
```

```
1.46e12
## # … with abbreviated variable names ¹WeightPounds, ²IsManualReport
```

4)Analyze: Summarizing the data:

```
n_distinct(daily_activity$Id)

## [1] 33

n_distinct(daily_calories$Id)

## [1] 33

n_distinct(daily_intensities$Id)

## [1] 33

n_distinct(daily_steps$Id)

## [1] 33

n_distinct(heartrate_sec$Id)

## [1] 14

n_distinct(minute_METs$Id)

## [1] 33

n_distinct(sleep_day$Id)

## [1] 24

n_distinct(weight_log$Id)

## [1] 8

nrow(daily_activity)

## [1] 940

nrow(daily_calories)

## [1] 940

nrow(daily_intensities)

## [1] 940

nrow(daily_steps)

## [1] 940

nrow(heartrate_sec)
```

```
## [1] 2483658
```

```
nrow(minute_METs)
```

```
## [1] 1325580
```

```
nrow(sleep_day)
```

```
## [1] 413
```

```
nrow(weight_log)
```

```
## [1] 67
```

```
daily_activity %>%
  select(TotalSteps,TotalDistance,SedentaryMinutes,LightlyActiveMinutes,
         FairlyActiveMinutes,VeryActiveMinutes,Calories) %>%
  summary()
```

```
##    TotalSteps    TotalDistance    SedentaryMinutes LightlyActiveMinutes
##  Min.   :    0   Min.   : 0.000   Min.   :   0.0   Min.   :  0.0
##  1st Qu.: 3790   1st Qu.: 2.620   1st Qu.: 729.8   1st Qu.:127.0
##  Median : 7406   Median : 5.245   Median :1057.5   Median :199.0
##  Mean   : 7638   Mean   : 5.490   Mean   : 991.2   Mean   :192.8
##  3rd Qu.:10727   3rd Qu.: 7.713   3rd Qu.:1229.5   3rd Qu.:264.0
##  Max.   :36019   Max.   :28.030   Max.   :1440.0   Max.   :518.0
##  FairlyActiveMinutes VeryActiveMinutes    Calories
##  Min.   :  0.00      Min.   :  0.00    Min.   :   0
##  1st Qu.:  0.00      1st Qu.:  0.00    1st Qu.:1828
##  Median :  6.00      Median :  4.00    Median :2134
##  Mean   : 13.56      Mean   : 21.16    Mean   :2304
##  3rd Qu.: 19.00      3rd Qu.: 32.00    3rd Qu.:2793
##  Max.   :143.00      Max.   :210.00    Max.   :4900
```

```
heartrate_sec %>%
  select(Value) %>%
  summary()
```

```
##      Value
##  Min.   : 36.00
##  1st Qu.: 63.00
##  Median : 73.00
##  Mean   : 77.33
##  3rd Qu.: 88.00
##  Max.   :203.00
```

```
minute_METs %>%
  select(METs) %>%
  summary()
```

```
##       METs
##  Min.   :  0.00
```

```
##  1st Qu.: 10.00
##  Median : 10.00
##  Mean   : 14.69
##  3rd Qu.: 11.00
##  Max.   :157.00
```

```
sleep_day %>%
  select(TotalTimeInBed, TotalMinutesAsleep,TotalSleepRecords) %>%
  summary()
```

```
##  TotalTimeInBed  TotalMinutesAsleep TotalSleepRecords
##  Min.   : 61.0   Min.   : 58.0      Min.   :1.000
##  1st Qu.:403.0   1st Qu.:361.0      1st Qu.:1.000
##  Median :463.0   Median :433.0      Median :1.000
##  Mean   :458.6   Mean   :419.5      Mean   :1.119
##  3rd Qu.:526.0   3rd Qu.:490.0      3rd Qu.:1.000
##  Max.   :961.0   Max.   :796.0      Max.   :3.000
```

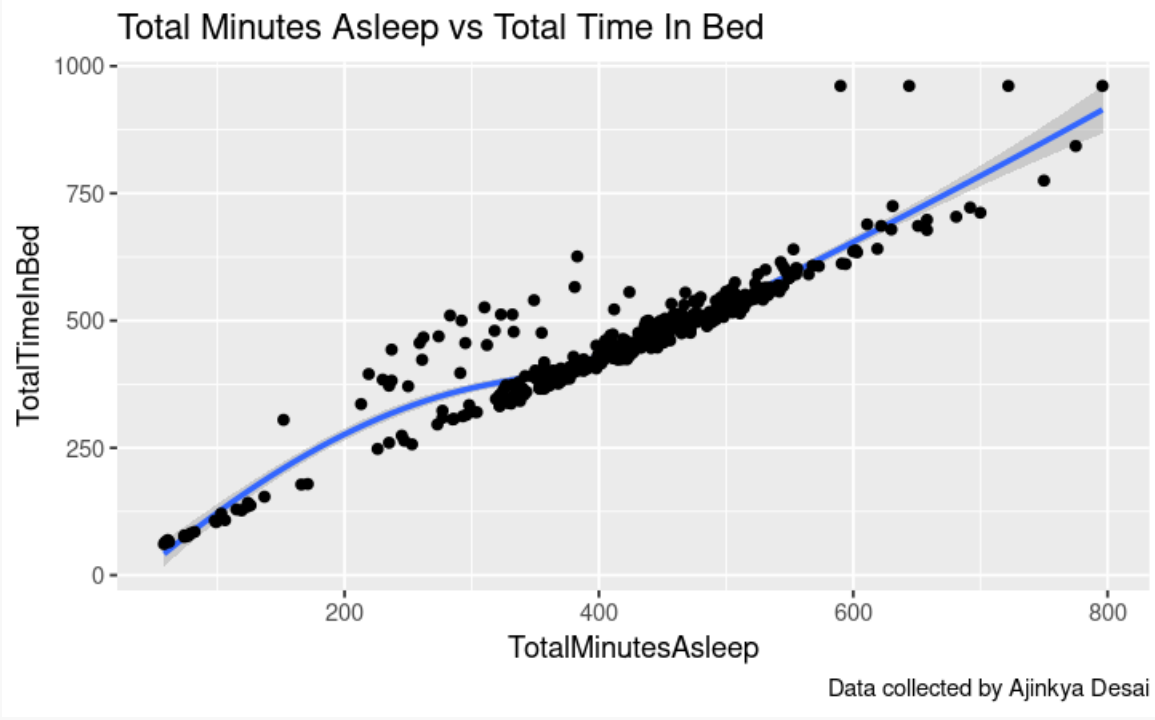
```
weight_log %>%
  select(WeightPounds, BMI) %>%
  summary()
```

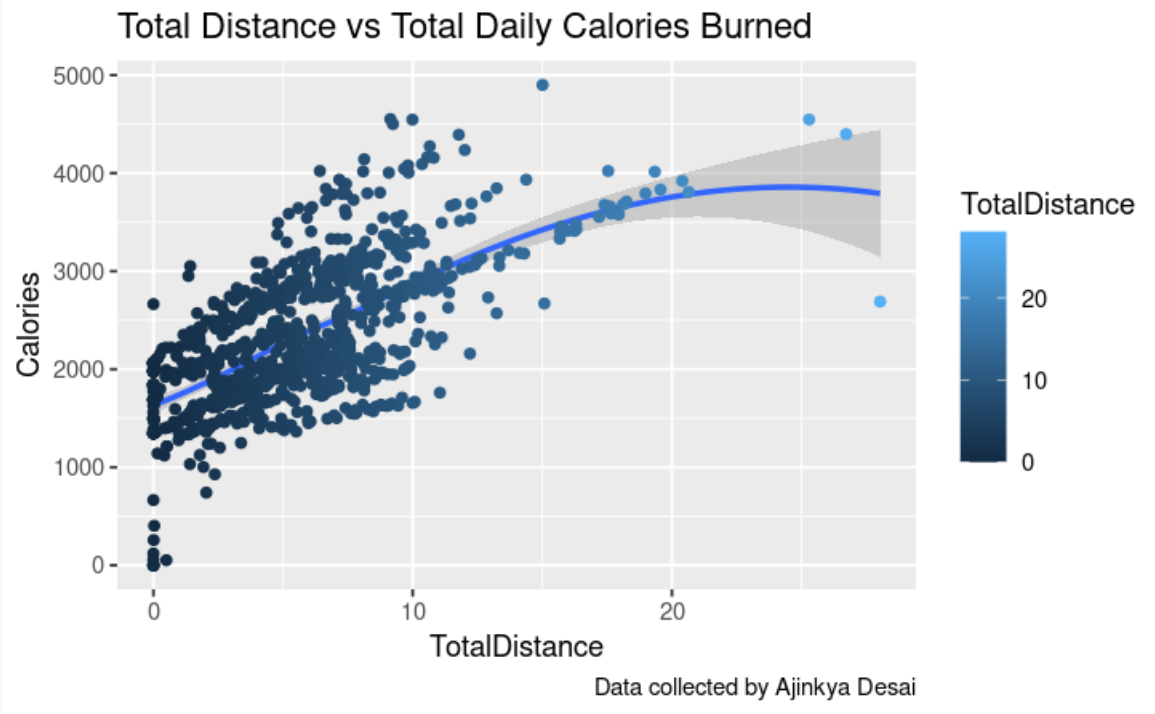
```
##   WeightPounds        BMI
##  Min.   :116.0   Min.   :21.45
##  1st Qu.:135.4   1st Qu.:23.96
##  Median :137.8   Median :24.39
##  Mean   :158.8   Mean   :25.19
##  3rd Qu.:187.5   3rd Qu.:25.56
##  Max.   :294.3   Max.   :47.54
```

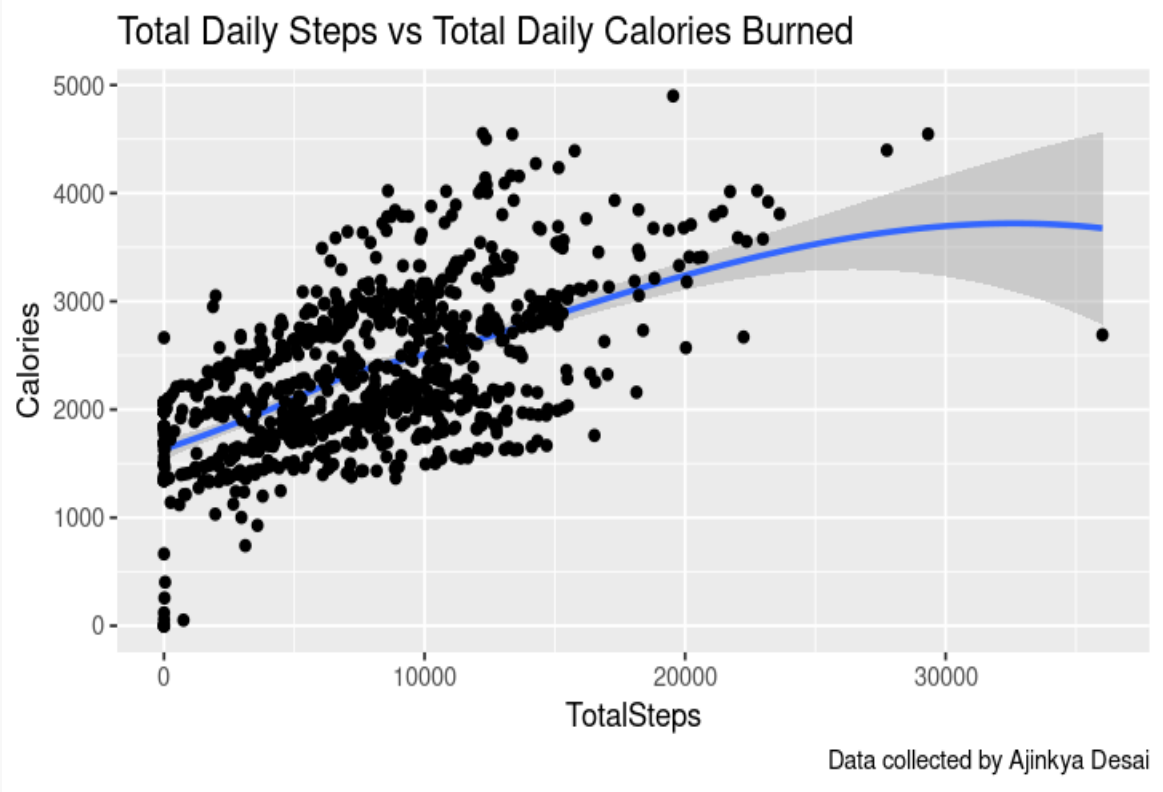
5)Share:

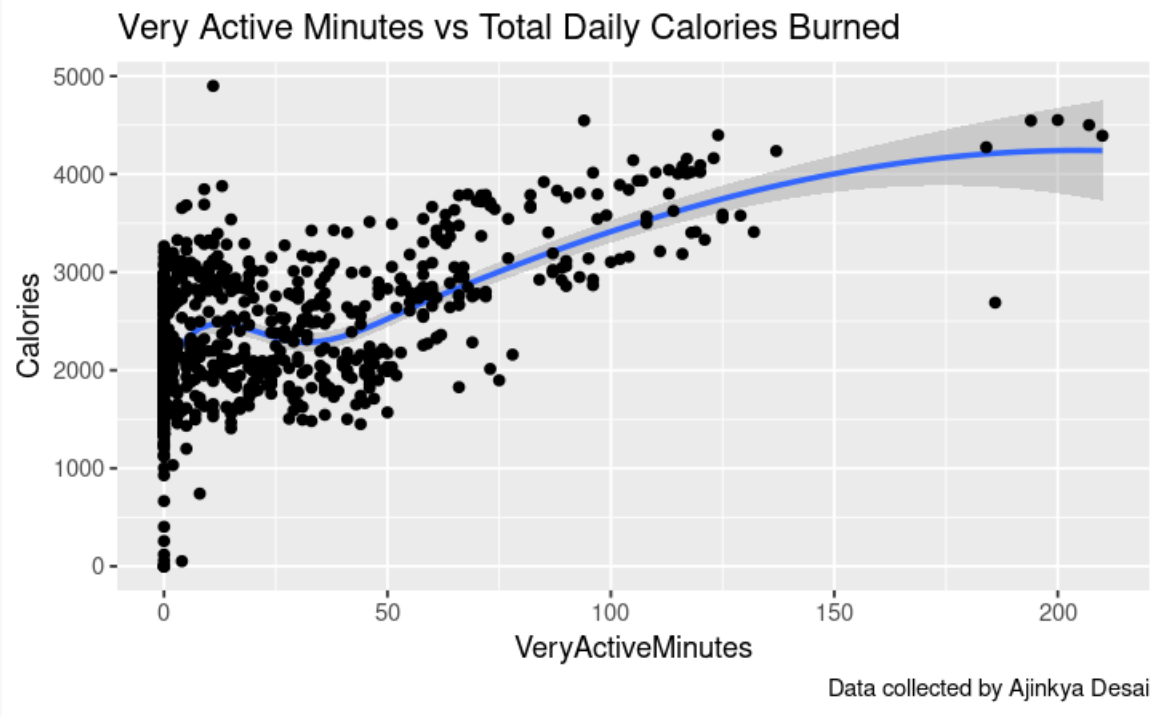
```
##The Relationship between Very Active Minutes vs Total Daily Calories Burned

ggplot(data=daily_activity)+

      geom_smooth(mapping=aes(x=VeryActiveMinutes, y=Calories))+

      geom_point(mapping=aes(x=VeryActiveMinutes, y=Calories))+

      labs(title = "The Relationship between Very Active Minutes vs Total Daily Calories Burned", caption = "Data collected by Ajinkya Desai")



##The Relationship between Total Daily Steps vs Total Daily Calories Burned

ggplot(data=daily_activity)+

 geom_smooth(mapping=aes(x=TotalSteps, y=Calories))+
```

```
  geom_point(mapping=aes(x=TotalSteps, y=Calories))+

  labs(title = "The Relationship between Total Daily Stepss vs Total Daily Calories Burned",
caption = "Data collected by Ajinkya Desai")


##The Relationship between Total  Distance vs Total Daily Calories Burned

ggplot(data=daily_activity)+

  geom_smooth(mapping=aes(x=TotalDistance, y=Calories))+

  geom_point(mapping=aes(x=TotalDistance, y=Calories, color=TotalDistance))+

  labs(title = "The Relationship between Total Distance vs Total Daily Calories Burned",
caption = "Data collected by Ajinkya Desai")


##The Relationship between Total Minutes Asleep vs Total Time In Bed

ggplot(data=sleep_day)+

  geom_smooth(mapping=aes(x=TotalMinutesAsleep, y=TotalTimeInBed))+

  geom_point(mapping=aes(x=TotalMinutesAsleep, y=TotalTimeInBed))+

  labs(title = "The Relationship between Total Minutes Asleep vs Total Time In Bed", caption
= "Data collected by Ajinkya Desai")
```


Total Minutes Asleep vs Total Time In Bed
TotalTimeInBed
1000
750
500
250
0
200
400
600
800
TotalMinutesAsleep
Data collected by Ajinkya Desai


Total Distance vs Total Daily Calories Burned
Calories
5000
4000
3000
2000
1000
0
0
10
20
TotalDistance
Data collected by Ajinkya Desai
TotalDistance
20
10
0


Total Daily Steps vs Total Daily Calories Burned
Calories
5000
4000
3000
2000
1000
0
0
10000
20000
30000
TotalSteps
Data collected by Ajinkya Desai


Very Active Minutes vs Total Daily Calories Burned
Calories
5000
4000
3000
2000
1000
0
0
50
100
150
200
VeryActiveMinutes
Data collected by Ajinkya Desai

5)Act:

Bellabeat has succeeded by empowering women by providing data on their activity, sleep, stress, hydration levels, and reproductive health. By analyzing how Fitbit users engage with features, recommendations can be made to enhance Bellabeat's growth.

To increase its reach, Bellabeat should significantly transform and upgrade its app. Rather than simply offering health data, the app should encourage users to reach their fitness goals and become a social media platform.

The Centers for Disease Control and Prevention (CDC) recommends working out with friends to boost motivation, experiment with workouts, and maintain consistency. The CDC even suggests using social media workout apps to connect with friends and achieve fitness goals. Bellabeat's app could serve as a social media workout app for women, creating a supportive community of women dedicated to prioritizing their health.

Recommendations for Bellabeat's app include:

1. Adding social networking features to enable users to share their favorite workouts, wellness tips, healthy meals, etc.
2. Enabling users to add friends and view each other's activity.
3. Creating weekly fitness and wellness challenges to promote usage.
4. Allowing health and fitness companies to advertise on the app.
5. Setting a daily goal of 10,000 steps and sending alert notifications to encourage users to meet the goal.
6. Recommending users to get at least 7 hours of sleep per night and enabling alert notifications to help them meet the goal.
7. Suggesting users to engage in 75 minutes of vigorous activity per week and enabling alert notifications to encourage them to meet this goal.
8. Encouraging users to input their weight and height to track their BMI.
9. Send notifications to keep users on track to burn the necessary calories if they express interest in losing weight.
10. Enabling alert notifications if a user's resting heart rate deviates significantly from their normal range.
11. Send notifications to encourage activity if a user has been awake in bed for an hour.
12. Sending notifications to encourage activity if a user has been sedentary for an extended period of time.