# #Case Study: Bellabeat Analysis

Author: Prajana Luitel

Date: 6-17-2024

Introduction:

In this case study, I analyzed data from Bellabeat, a smart device company based in San Francisco. I looked at the data to answer key business questions and developed actionable insights.

Ask:

How can we better market our smart devices based on user trends?

What are some trends in smart device usage?

How could these trends apply to Bellabeat customers?

**Prepare:**

This data set does have its limitations. It was collected from 2016-03-12 to 2016-05-12 and had a participant base of 30. I looked at Four of the tables from this dataset, dailyActivity\_merged, hourlyIntensities\_merged, hourlySteps\_merged, and hourlyCalories\_Merged.

Process:

I choose to work in BigQuery.

I uploaded the datasets to BigQuery under the project “mturkfitbit\_export”.

I uploaded four datasets into BigQuery and renamed the dataset as:

dailyActivity\_merged as Dailyactivity\_3\_12,

hourlyCalories\_merged as hourly\_calories\_merged,

hourlyIntensities\_merged as hourly\_intensities,

hourlySteps\_merged as hourly\_steps\_merged

Analysis and Share

--Identifying distinct id in different dataset

SELECT count (DISTINCT Id)

FROM `mturkfitbit\_export.Dailyactivity\_3\_12`;

SELECT count (DISTINCT Id)

FROM `mturkfitbit\_export.hourly\_calories\_merged`;

SELECT count (DISTINCT Id)

FROM `mturkfitbit\_export.hourly\_intensities`;

SELECT count (DISTINCT Id)

FROM `mturkfitbit\_export.hourly\_steps\_merged`;

--transforming hourly data to daily data and creating a new table.

--from hourly\_step\_merged to Daily\_steps\_merged.

Create table `mturkfitbit\_export.Daily\_steps\_merged` as

SELECT

Id,

DATE(ActivityHour) AS date,

SUM(StepTotal) AS daily\_sum\_StepTotal,

AVG(StepTotal) AS daily\_avg\_StepTotal,

MAX(StepTotal) AS daily\_max\_StepTotal,

MIN(StepTotal) AS daily\_min\_StepTotal

FROM

`mturkfitbit\_export.hourly\_steps\_merged`

GROUP BY

Id, DATE(ActivityHour)

ORDER BY

date, Id;

--from hourly\_calories\_merged to daily\_calories\_merged.

Create table `mturkfitbit\_export.daily\_calories\_merged` as

SELECT

Id,

DATE(ActivityHour) AS date,

SUM(Calories) AS daily\_sum\_daily\_calories,

AVG(Calories) AS daily\_avg\_calories,

MAX(Calories) AS daily\_max\_calories,

MIN(Calories) AS daily\_min\_caloriesl

FROM

`mturkfitbit\_export.hourly\_calories\_merged`

GROUP BY

Id, DATE(ActivityHour)

ORDER BY

date, Id;

--from hourly\_intensities to daily\_intensities.

Create table `mturkfitbit\_export.daily\_intensities` as

SELECT

Id,

DATE(ActivityHour) AS date,

SUM(TotalIntensity) AS daily\_sum\_daily\_TotalIntensity,

AVG(TotalIntensity) AS daily\_avg\_TotalIntensity,

MAX(TotalIntensity) AS daily\_max\_TotalIntensity,

MIN(TotalIntensity) AS daily\_min\_TotalIntensity

FROM

`mturkfitbit\_export.hourly\_intensities`

GROUP BY

Id, DATE(ActivityHour)

ORDER BY

date, Id;

-- joining all four files together to get a new table with sum\_calories daily\_steps and sum of Intensities.

--using inner join and adding a new column day\_of\_week.

Create table `mturkfitbit\_export.combined table1` as

SELECT

t1.Id,TotalDistance,TrackerDistance,

t1.ActivityDate as date1,

FORMAT\_DATE('%A', t1.ActivityDate) AS day\_of\_week,

t2.daily\_sum\_daily\_calories as sum\_calories,

t3.daily\_sum\_StepTotal as daily\_steps,

t4.daily\_sum\_daily\_TotalIntensity as sum\_inten

FROM

`mturkfitbit\_export.Dailyactivity\_3\_12` AS t1

INNER JOIN

`mturkfitbit\_export.daily\_calories\_merged` AS t2

ON

t1.id = t2.id and t1.ActivityDate=t2.date

INNER JOIN

`mturkfitbit\_export.Daily\_steps\_merged` AS t3

ON

t1.id = t3.id and t1.ActivityDate=t3.date

INNER JOIN

`mturkfitbit\_export.daily\_intensities` AS t4

on

t1.id =t4.id and t1.ActivityDate=t4.date

how often individual user are using the tracker

--i wanted to see how many times each user uses the tracker.

SELECT

Id, COUNT (Id) AS Total\_Id

FROM

`mturkfitbit\_export.combined table1`

GROUP BY

Id

ORDER BY

Total\_Id DESC;

I used a case to find out the count of IDs into different intervals and count the number of IDs falling into each interval using the case.

Create table `mturkfitbit\_export.use distribution`as

SELECT

CASE

WHEN Total\_Id < 5 THEN 'uses less than 5'

WHEN Total\_Id >= 5 AND Total\_Id < 10 THEN 'uses between 5 and 10'

WHEN Total\_Id >= 10 AND Total\_Id < 15 THEN 'uses between 10 and 15'

WHEN Total\_Id >= 15 AND Total\_Id < 20 THEN 'uses between 15 and 20'

WHEN Total\_Id >= 20 AND Total\_Id < 25 THEN 'uses between 20 and 25'

ELSE 'uses greater than or equal to 25'

END AS interval1,

COUNT(Total\_Id) AS uses\_count

FROM (

SELECT

Id,

COUNT(Id) AS Total\_Id

FROM

`mturkfitbit\_export.combined table1`

GROUP BY

Id

)

GROUP BY

interval1

ORDER BY

interval1;

------line chart time series data everyday min max step count

----------------average, min , max steps of individual person

select

ActivityDate,

min(TotalSteps) as min\_step,

max(TotalSteps) as max\_step,

avg(TotalSteps) as avg\_steps

from `mturkfitbit\_export.Dailyactivity\_3\_12`

group by

ActivityDate;

--extracting day-of week from date 1 column

SELECT

FORMAT\_DATE('%A', date1) AS day\_of\_week,

COUNT(\*) AS total\_count

FROM

`mturkfitbit\_export.combined table1`

GROUP BY

day\_of\_week

ORDER BY

day\_of\_week;

--i wanted to see what is daily average intensities

----bar chart for total intensity used by day of week.

select

FORMAT\_DATE('%A', date1) AS day\_of\_week,

avg(sum\_Inten) as Total\_intensity

FROM

`mturkfitbit\_export.combined table1`

GROUP BY

day\_of\_week

order by

Total\_intensity desc;

-------------------------

–created new table adding days as days of week in dailyactivity.

Create table `mturkfitbit\_export.Dailyactivity\_3\_12\_days`as

SELECT \*, FORMAT\_DATE('%A', ActivityDate) AS day\_of\_week,

FROM

`mturkfitbit\_export.Dailyactivity\_3\_12`

----------------------------------------I wanted to determine if people were more active on a certain day of the week.

---bar chart average activity level by day of week.

SELECT day\_of\_week,

ROUND (avg(VeryActiveMinutes), 2) AS Avg\_Very\_Active\_Minutes,

ROUND (avg(FairlyActiveMinutes), 2) AS Avg\_Fairly\_Active\_Minutes,

ROUND (avg(LightlyActiveMinutes), 2) AS Avg\_Lightly\_Active\_Minutes,

ROUND (avg(SedentaryMinutes), 2) AS Avg\_Sedentary\_Minutes,

FROM

`mturkfitbit\_export.Dailyactivity\_3\_12\_days`

GROUP BY

day\_of\_week;

-------relation ship between steps and calories burned

---time series plot.

SELECT

date1,

AVG(daily\_Steps) AS Avg\_Total\_Steps,

FROM

`mturkfitbit\_export.combined table1`

group by

date1;

-------relation ship between total distance and calories burned

---time series plot.

SELECT

date1,

AVG(TotalDistance) AS Avg\_Total\_Distance,

FROM

`mturkfitbit\_export.combined table1`

group by

date1;

Share:

– add tableau link here

**Act**

**Through my analysis i was able to uncover the following:**

Sedentary Behavior Dominance:

The data reveals that users predominantly spend most of their time in the sedentary category, indicating a potential lack of physical activity among the user population.

Daily Usage Frequency:

Only two out of 34 users are consistently using the device every day, suggesting that daily usage is not common among the majority of users.

Monthly Usage Duration:

A significant portion, 85% of users, utilize the device for less than 15 days in a month. This indicates that a large proportion of users may not be consistently engaging with the device or maintaining a regular exercise routine.

Steps and Calorie Burn:

There is a positive correlation between the number of steps taken and the calories burned. This emphasizes the importance of physical activity in calorie expenditure and overall health.

Weekday vs. Weekend Activity Patterns:

Weekdays, from Monday to Friday, exhibit lower levels of very active, fairly active, and lightly active minutes compared to weekends (Saturday and Sunday). This shift suggests that users may have more leisure time and engage in higher levels of physical activity during weekends.

#Recommendations:

**Promoting Physical Activity:** Encouraging users to reduce sedentary behavior and increase physical activity levels can have significant health benefits. Implementing reminders or challenges to increase movement throughout the day may be beneficial.

**Enhancing User Engagement:** Strategies to increase daily usage frequency and promote consistent use of the device should be explored. Providing incentives or rewards for regular use may help in sustaining user engagement.

**Targeted Interventions:** Understanding user demographics and preferences can inform targeted interventions aimed at improving device usage patterns. Tailored approaches to address specific user needs and preferences may lead to better outcomes.