December 7th, 2023

Merged all dataset into one workbook for ease of viewing

Established which datasets would be useful during this analysis removed datasets that we would not be using during analysis

Formatted data according to the following:

- Using narrow long data opposed to wide data
- Removing data that is too specific i.e; choosing to use hourly caloric data instead of minute data

Expanded all columns to be able to look at all data

Removed the “fat” column from the weight_log info due to lack of ample data for other rows

Separated data and time stamps from weight_log info table

Weight_log info; use INT function to separate date

Use subtraction formula cell to isolate time

Convert Log10 to whole numbers; changed formatted form general to numerical

Separate date and time from hour per second

Changed all decimals into number format

LogID in weight_info table formatted into whole numbers

Sleep day table formatted into date format

Daily intensities format decimals into number format

Delete tracker distance in daily_sleep as it was equal to total distance,

December 8th, 2023

Tables were all individually uploaded into BigQuery under the `bigquery-practice-101099.bellabeat_capstone_project` dataset

- weightLogInfo_merged.csv was renamed as weight_log
- sleepDay_merged.csv was renamed as daily_sleep
- heartRate_seconds_merged.csv was renamed as seconds_heartrate
- Daily_activity_merged.csv was renamed as daily_activity
- Daily_steps_merged.csv was renamed as daily_steps
- Daily_calories_merged.csv was renamed as daily_calories

```
Count the distinct number of ID's: SELECT DISTINCT ID FROM
`bigquery-practice-101099.bellabeat_capstone_project.daily_calories'

Count the distinct number of ID's: SELECT DISTINCT ID FROM
`bigquery-practice-101099.bellabeat_capstone_project.weight_log

Count the distinct number of ID's: SELECT DISTINCT ID FROM
`bigquery-practice-101099.bellabeat_capstone_project.daily_sleep

Count the distinct number of ID's: SELECT DISTINCT ID FROM
`bigquery-practice-101099.bellabeat_capstone_project.seconds_heartrate

Count the distinct number of ID's: SELECT DISTINCT ID FROM
`bigquery-practice-101099.bellabeat_capstone_project.daily_activity

Count the distinct number of ID's: SELECT DISTINCT ID FROM
`bigquery-practice-101099.bellabeat_capstone_project.daily_steps
```

The following tables were eliminated from analysis, due to inconsistencies found amount distinct ID's;

- weight_log
- daily_sleep
- seconds_heartRate

December 9th, 2023

Total number of rows:
```
SELECT COUNT (*) as count_daily_activity
FROM `bigquery-practice-101099.bellabeat_capstone_project.daily_activity `
```

```
SELECT MAX(totaldistance)

FROM `bigquery-practice-101099.bellabeat_capstone_project.daily_activity `
```

```
SELECT MIN(totaldistance)
FROM `bigquery-practice-101099.bellabeat_capstone_project.daily_activity
```

```
SELECT AVG (moderatelyactivedistance)
FROM `bigquery-practice-101099.bellabeat_capstone_project.daily_activity `
```

```
SELECT MIN(steptotal) FROM
`bigquery-practice-101099.bellabeat_capstone_project.daily_steps`
SELECT Max(steptotal) FROM
`bigquery-practice-101099.bellabeat_capstone_project.daily_steps`
SELECT AVG (steptotal)
FROM `bigquery-practice-101099.bellabeat_capstone_project.daily_steps`
```

```
SELECT MIN(calories) FROM
`bigquery-practice-101099.bellabeat_capstone_project.daily_calories `
SELECT MAX(calories) FROM
`bigquery-practice-101099.bellabeat_capstone_project.daily_calories `
SELECT AVG(calories) FROM
`bigquery-practice-101099.bellabeat_capstone_project.daily_calories `
```

December 10th, 2023

```
SELECT id, count(id) AS total_id_count

FROM `bigquery-practice-101099.bellabeat_capstone_project.daily_activity `

group by id
```

```
SELECT DISTINCT DA.ID, sum (DA.totaldistance) as sum_total_distance, sum(DS.steptotal)
as sum_step_total
FROM `bigquery-practice-101099.bellabeat_capstone_project.daily_activity ` as DA
left join bigquery-practice-101099.bellabeat_capstone_project.daily_steps as DS
ON DS.ID = DA.id
group by DA.ID
order by sum_total_distance desc
```

```sql
SELECT DISTINCT DA.ID, sum (DA.totaldistance) as sum_total_distance, sum(DC.calories)
as sum_calories
FROM `bigquery-practice-101099.bellabeat_capstone_project.daily_activity ` as DA
left join bigquery-practice-101099.bellabeat_capstone_project.daily_calories as DC
ON DA.ID = DC.id
group by DA.ID
order by sum_total_distance desc
```