#The query is used to check for nulls in any column. The DailyActivity does not have any nulls

SELECT \*

FROM `bellabeat-99999.Case1.DailyActivity`

WHERE NOT (`bellabeat-99999.Case1.DailyActivity` IS NOT NULL);

#The query is used to check for nulls in any column. The sleepDay does not have any nulls

SELECT \*

FROM `bellabeat-99999.Case1.sleepDay`

WHERE NOT (`bellabeat-99999.Case1.sleepDay` IS NOT NULL);

#The query is used to check for nulls in any column. The WeightLogInfo does not have any nulls

SELECT \*

FROM `bellabeat-99999.Case1.weightLogInfo`

WHERE Id || Date || WeightKg || WeightPounds || Fat || BMI || IsManualReport || LogId IS NULL;

#Created a table with the day of the week added

SELECT

VeryActiveMinutes/60 as very\_active\_hours,

SedentaryMinutes/60 as sedentary\_hours,

Calories,

format\_date('%a', ActivityDate) as dayofweek

FROM `bellabeat-99999.Case1.DailyActivity`

#Join Activity and Sleep Activity

SELECT \*

FROM `bellabeat-99999.Case1.DailyActivity`dam

JOIN `bellabeat-99999.Case1.sleepDay` sd

ON dam.Id = sd.Id

#exported to table as Activity\_Merge\_Sleep

#Average amount of activity that a user is performing

SELECT

AVG(TotalSteps) as avg\_total\_steps,

AVG(TotalDistance) as avg\_total\_distance,

AVG(VeryActiveMinutes) as avg\_very\_active\_minutes,

AVG(FairlyActiveMinutes) as avg\_fairly\_active\_minutes,

AVG(LightlyActiveMinutes) as avg\_lightly\_active\_minutes,

AVG(SedentaryMinutes) as avg\_sedentary\_minutes,

AVG(Calories) as avg\_calories

FROM

`bellabeat-99999.Case1.DailyActivity`

/\*

Analysis:

- More minutes are spent being sedentary > lightly active > fairly active > very active

\*/

#Average amount of activity that is performed

SELECT

DISTINCT Id,

AVG(TotalSteps) as avg\_total\_steps,

AVG(TotalDistance) as avg\_total\_distance,

AVG(VeryActiveMinutes) as avg\_very\_active\_minutes,

AVG(FairlyActiveMinutes) as avg\_fairly\_active\_minutes,

AVG(LightlyActiveMinutes) as avg\_lightly\_active\_minutes,

AVG(SedentaryMinutes) as avg\_sedentary\_minutes,

AVG(Calories) as avg\_calories

FROM

`bellabeat-99999.Case1.DailyActivity`

GROUP BY

Id

ORDER BY

Id

#Average amount of user sleep in minutes and hours

WITH avg\_sleep\_hours as

(

  SELECT

  DISTINCT Id,

  AVG(TotalMinutesAsleep) as avg\_minutes\_asleep,

  AVG(TotalTimeInBed) as avg\_minutes\_in\_bed

  FROM `bellabeat-99999.Case1.sleepDay`

  GROUP BY

  Id

)

SELECT

\*,

avg\_sleep\_hours.avg\_minutes\_asleep/60 as avg\_sleep\_hours,

avg\_sleep\_hours.avg\_minutes\_in\_bed - avg\_sleep\_hours.avg\_minutes\_asleep as avg\_time\_awake

FROM avg\_sleep\_hours

ORDER BY

Id

/\*

Analysis:

- Looking at the average data only 12 out of 24 got the recommended 7 hours of sleep

- 12 other users got an avg of less than 7 hours of sleep

\*/

#combined tables AverageActivityOfUSer & AverageSleep\_MinutesHours

SELECT

\*

FROM `bellabeat-99999.Case1.AverageActivityOfUSer` aau

JOIN `bellabeat-99999.Case1.AverageSleep\_MinutesHours` asu

ON aau.Id = asu.Id

ORDER BY

avg\_very\_active\_minutes DESC

#The average amount of time a user is awake while in bed

SELECT

DISTINCT Id,

avg(TotalTimeInBed - TotalMinutesAsleep) as time\_spend\_awake\_inbed

FROM `bellabeat-99999.Case1.sleepDay`

GROUP BY

id

ORDER BY

time\_spend\_awake\_inbed

/\*

Analysis:

- All users have some time in which they are in bed but not sleeping

- 19 out of the 24 users have more than 18 minutes of time in bed not sleeping

- This can be further seen in detail by the query below that shows each days amounts

\*/

#The amount of time a user was awake in bed for each day

SELECT

DISTINCT Id,

SleepDay,

TotalTimeInBed - TotalMinutesAsleep as time\_spend\_awake\_inbed

FROM `bellabeat-99999.Case1.sleepDay`

ORDER BY

id, SleepDay

#Average users user information by day

WITH active\_time as

(

  SELECT

  format\_date('%a', ActivityDate) as dayofweek,

  AVG(TotalSteps) as avg\_total\_steps,

  AVG(TotalDistance) as avg\_total\_distance,

  AVG(VeryActiveMinutes) as avg\_very\_active\_minutes,

  AVG(FairlyActiveMinutes) as avg\_fairly\_active\_minutes,

  AVG(LightlyActiveMinutes) as avg\_lightly\_active\_minutes,

  AVG(SedentaryMinutes) as avg\_sedentary\_minutes,

  AVG(Calories) as avg\_calories

  FROM `bellabeat-99999.Case1.DailyActivity`

  group by

  dayofweek

  Order by

    CASE

    WHEN dayofweek = 'Sun' THEN 1

    WHEN dayofweek = 'Mon' THEN 2

    WHEN dayofweek = 'Tue' THEN 3

    WHEN dayofweek = 'Wed' THEN 4

    WHEN dayofweek = 'Thu' THEN 5

    WHEN dayofweek = 'Fri' THEN 6

    WHEN dayofweek = 'Sat' THEN 7

      END ASC

)

SELECT

dayofweek,

active\_time.avg\_total\_steps,

active\_time.avg\_total\_distance,

active\_time.avg\_very\_active\_minutes,

active\_time.avg\_fairly\_active\_minutes,

active\_time.avg\_lightly\_active\_minutes,

active\_time.avg\_sedentary\_minutes,

active\_time.avg\_calories,

active\_time.avg\_very\_active\_minutes + active\_time.avg\_fairly\_active\_minutes + active\_time.avg\_lightly\_active\_minutes as total\_active\_time

FROM active\_time

/\*

saved this query as a table called ActivityByDay

Analysis:

- The most steps happen on Tuesday and Saturday

- The most active day is Saturday with an average of 244 minutes of activity when combining very, faily, lightly

- The least active day is Sunday

- The most sedentary day is Monday with the least being Thursday

\*/

#Average users sleep information by day

WITH sleep\_time as

(

  SELECT

  format\_date('%a', SleepDay) as dayofweek,

  AVG(TotalMinutesAsleep) as avg\_minutes\_asleep,

  AVG(TotalTimeInBed) as avg\_minutes\_in\_bed

  FROM `bellabeat-99999.Case1.sleepDay`

  group by

  dayofweek

  Order by

    CASE

    WHEN dayofweek = 'Sun' THEN 1

    WHEN dayofweek = 'Mon' THEN 2

    WHEN dayofweek = 'Tue' THEN 3

    WHEN dayofweek = 'Wed' THEN 4

    WHEN dayofweek = 'Thu' THEN 5

    WHEN dayofweek = 'Fri' THEN 6

    WHEN dayofweek = 'Sat' THEN 7

      END ASC

)

SELECT

dayofweek,

sleep\_time.avg\_minutes\_asleep,

sleep\_time.avg\_minutes\_in\_bed,

avg\_minutes\_in\_bed - avg\_minutes\_asleep as average\_time\_awake,

sleep\_time.avg\_minutes\_asleep/60 as avg\_hour\_asleep,

sleep\_time.avg\_minutes\_in\_bed/60 as avg\_hour\_inbed

FROM sleep\_time

/\*

Saved this query as a table SleepByDay

Analysis:

- The most hours asleep is Sunday with 7.5 hours

- The most hours in bed is Sunday with 8.3 hours

\*/

#combine data form the activity by day and sleep by day tables

SELECT

\*

FROM `bellabeat-99999.Case1.ActivityByDay` abd

JOIN `bellabeat-99999.Case1.SleepByDay` sbd

ON abd.dayofweek = sbd.dayofweek

#This query gets us all the values together with total values.

SELECT

abd.dayofweek,

abd.avg\_very\_active\_minutes,

abd.avg\_fairly\_active\_minutes,

abd.avg\_lightly\_active\_minutes,

abd.avg\_sedentary\_minutes,

sbd.average\_time\_awake,

sbd.avg\_minutes\_asleep,

abd.total\_active\_time+ abd.avg\_sedentary\_minutes + sbd.avg\_minutes\_asleep + sbd.average\_time\_awake as all\_total\_time

FROM `bellabeat-99999.Case1.ActivityByDay` abd

JOIN `bellabeat-99999.Case1.SleepByDay` sbd

ON abd.dayofweek = sbd.dayofweek

Order by

  CASE

  WHEN dayofweek = 'Sun' THEN 1

  WHEN dayofweek = 'Mon' THEN 2

  WHEN dayofweek = 'Tue' THEN 3

  WHEN dayofweek = 'Wed' THEN 4

  WHEN dayofweek = 'Thu' THEN 5

  WHEN dayofweek = 'Fri' THEN 6

  WHEN dayofweek = 'Sat' THEN 7

    END ASC

#Find the SUM of active minutes and sedentary minutes

WITH active\_time as

(

  SELECT

  format\_date('%a', ActivityDate) as dayofweek,

  SUM(TotalSteps) as sum\_total\_steps,

  SUM(TotalDistance) as sum\_total\_distance,

  SUM(VeryActiveMinutes) as sum\_very\_active\_minutes,

  SUM(FairlyActiveMinutes) as sum\_fairly\_active\_minutes,

  SUM(LightlyActiveMinutes) as sum\_lightly\_active\_minutes,

  SUM(SedentaryMinutes) as sum\_sedentary\_minutes,

  SUM(Calories) as sum\_calories

  FROM `bellabeat-99999.Case1.DailyActivity`

  group by

  dayofweek

  Order by

    CASE

    WHEN dayofweek = 'Sun' THEN 1

    WHEN dayofweek = 'Mon' THEN 2

    WHEN dayofweek = 'Tue' THEN 3

    WHEN dayofweek = 'Wed' THEN 4

    WHEN dayofweek = 'Thu' THEN 5

    WHEN dayofweek = 'Fri' THEN 6

    WHEN dayofweek = 'Sat' THEN 7

      END ASC

)

SELECT

dayofweek,

active\_time.sum\_total\_steps,

active\_time.sum\_total\_distance,

active\_time.sum\_very\_active\_minutes as very\_active\_minutes,

active\_time.sum\_fairly\_active\_minutes as fairly\_active\_minutes,

active\_time.sum\_lightly\_active\_minutes as lightly\_active\_minutes,

active\_time.sum\_sedentary\_minutes as sedentary\_hours,

active\_time.sum\_calories,

active\_time.sum\_very\_active\_minutes + active\_time.sum\_fairly\_active\_minutes + active\_time.sum\_lightly\_active\_minutes as total\_active\_time,

active\_time.sum\_very\_active\_minutes + active\_time.sum\_fairly\_active\_minutes + active\_time.sum\_lightly\_active\_minutes + active\_time.sum\_sedentary\_minutes as total\_amount\_of\_time

FROM active\_time

#exported as SumTotalTimes

#Find Percentage of time spent Sedentary & active per type

SELECT

avg(very\_active\_minutes)/avg(total\_amount\_of\_time) \* 100 as very\_active\_percentage,

avg(fairly\_active\_minutes)/avg(total\_amount\_of\_time) \* 100 as fairly\_active\_percentage,

avg(lightly\_active\_minutes)/avg(total\_amount\_of\_time) \* 100 as lightly\_active\_percentage,

avg(total\_active\_time/total\_amount\_of\_time) \* 100 as total\_active\_time\_percentage,

avg(sedentary\_hours/total\_amount\_of\_time) \* 100 as total\_sedentary\_hours\_percentage

FROM `bellabeat-99999.Case1.SumTotalTimes`

#exported as ActivityPercentage