-- count # of IDs in dailyActivity\_Merged

SELECT COUNT(DISTINCT ID)

FROM `bellabeat-case-study-410302.activity.activity\_daily`

-- count # of IDs in hourlyCalories\_Merged

SELECT COUNT(DISTINCT ID)

FROM `bellabeat-case-study-410302.calories.calories\_hourly`

-- count # of IDs in hourlyIntensities\_Merged

SELECT COUNT(DISTINCT ID)

FROM `bellabeat-case-study-410302.intensities.intensities\_hourly`

-- count # of IDs in hourlySteps\_Merged

SELECT COUNT(DISTINCT ID)

FROM `bellabeat-case-study-410302.steps.steps\_hourly`

-- count # of IDs in sleepDay\_Merged

SELECT COUNT(DISTINCT ID)

FROM `bellabeat-case-study-410302.sleep.sleep\_daily`

-- check for duplicate entries in dailyActivity\_Merged

SELECT Id, ActivityDate, COUNT(\*) AS Occurrences

FROM `bellabeat-case-study-410302.activity.activity\_daily`

GROUP BY Id, ActivityDate

HAVING COUNT(\*) > 1;

-- check for duplicate entries in hourlyCalories\_Merged

SELECT Id, ActivityHour\_Date, ActivityHour\_Time, COUNT(\*) AS Occurrences

FROM `bellabeat-case-study-410302.calories.calories\_hourly`

GROUP BY Id, ActivityHour\_Date, ActivityHour\_Time

HAVING COUNT(\*) > 1;

-- check for duplicate entries in hourlyIntensities\_Merged

SELECT Id, ActivityHour\_Date, ActivityHour\_Time, COUNT(\*) AS Occurrences

FROM `bellabeat-case-study-410302.intensities.intensities\_hourly`

GROUP BY Id, ActivityHour\_Date, ActivityHour\_Time

HAVING COUNT(\*) > 1;

-- check for duplicate entries in hourlySteps\_Merged

SELECT Id, ActivityHour\_Date, ActivityHour\_Time, COUNT(\*) AS Occurrences

FROM `bellabeat-case-study-410302.steps.steps\_hourly`

GROUP BY Id, ActivityHour\_Date, ActivityHour\_Time

HAVING COUNT(\*) > 1;

-- check for duplicate entries in sleepDay\_Merged

SELECT Id, SleepDay, COUNT(\*) AS Occurrences

 FROM `bellabeat-case-study-410302.sleep.sleep\_daily`

 GROUP BY Id, SleepDay

 HAVING COUNT(\*) >1;

-- create new table merged.Hourly

CREATE TABLE merged.Hourly (

  Id numeric (10,0),

  dates string (50),

  times int,

  calories numeric (5,0),

  total\_intensity numeric (5,0),

  avg\_intensity float64,

  total\_steps numeric (5,0)

);

-- join individual hourly data tables together and map columns to the right columns in merged.Hourly

INSERT INTO merged.Hourly(

  Id, dates, times, calories, total\_intensity, avg\_intensity, total\_steps)

  (SELECT cal.Id, cal.dates, cal.times, cal.calories, int.total\_intensity, int.avg\_intensity, step.total\_steps

  FROM calories\_hourly AS cal

  INNER JOIN intensities\_hourly AS int

  ON cal.Id = int.Id AND cal.dates = int.dates AND cal.times = int.times

  INNER JOIN steps\_hourly AS step

  ON cal.Id = step.Id AND cal.dates = step.dates AND cal.times = step.times);

-- merge daily data sets into one table for more efficient queries

SELECT \*

FROM `bellabeat-case-study-410302.activity.activity\_daily` AS activity\_daily

LEFT JOIN `bellabeat-case-study-410302.sleep.sleep\_daily` AS sleep\_daily

ON activity\_daily.ActivityDate = sleep\_daily.SleepDay AND activity\_daily.Id = sleep\_daily.Id

-- average time spent in minutes for each activity level

SELECT ROUND(AVG(activity\_daily.VeryActiveMinutes),2) AS Avg\_Very\_Active\_Minutes,

ROUND(AVG(activity\_daily.FairlyActiveMinutes),2) AS Avg\_Fairly\_Active\_Minutes,

ROUND(AVG(activity\_daily.LightlyActiveMinutes),2)AS Avg\_Lightly\_Active\_Minutes,

ROUND(AVG(activity\_daily.SedentaryMinutes),2) AS Avg\_Sedentary\_Minutes

FROM `bellabeat-case-study-410302.activity.activity\_daily` AS activity\_daily

LEFT JOIN `bellabeat-case-study-410302.sleep.sleep\_daily` AS sleep\_daily

ON activity\_daily.ActivityDate = sleep\_daily.SleepDay AND activity\_daily.Id = sleep\_daily.Id

We can do this in hours too:

-- average time spent in hours for each activity level

SELECT ROUND(AVG(activity\_daily.VeryActiveMinutes)/60, 2) AS Avg\_Very\_Active\_Hours,

ROUND(AVG(activity\_daily.FairlyActiveMinutes)/60,2) AS Avg\_Fairly\_Active\_Hours,

ROUND(AVG(activity\_daily.LightlyActiveMinutes)/60,2) AS Avg\_Lightly\_Active\_Hours,

ROUND(AVG(activity\_daily.SedentaryMinutes)/60,2) AS Avg\_Sedentary\_Hours

FROM `bellabeat-case-study-410302.activity.activity\_daily` AS activity\_daily

LEFT JOIN `bellabeat-case-study-410302.sleep.sleep\_daily` AS sleep\_daily

ON activity\_daily.ActivityDate = sleep\_daily.SleepDay AND activity\_daily.Id = sleep\_daily.Id

-- average calories burned and average total steps taken

SELECT ROUND(AVG(activity\_daily.Calories), 2) AS Avg\_Calories\_Burned,

ROUND(AVG(activity\_daily.TotalSteps),2) AS Avg\_Total\_Steps

FROM `bellabeat-case-study-410302.activity.activity\_daily` AS activity\_daily

LEFT JOIN `bellabeat-case-study-410302.sleep.sleep\_daily` AS sleep\_daily

ON activity\_daily.ActivityDate = sleep\_daily.SleepDay AND activity\_daily.Id = sleep\_daily.Id

-- Add columns to the table

ALTER TABLE activity.activity\_daily

ADD COLUMN VAD\_Ratio DECIMAL (10,2),

ADD COLUMN MAD\_Ratio DECIMAL (10,2),

ADD COLUMN LAD\_Ratio DECIMAL (10,2),

ADD COLUMN SAD\_Ratio DECIMAL (10,2);

-- Update the values in the new columns

UPDATE activity.activity\_daily

SET

  VAD\_Ratio = activity\_daily.VeryActiveDistance / activity\_daily.TotalDistance,

  MAD\_Ratio = activity\_daily.ModeratelyActiveDistance / activity\_daily.TotalDistance,

  LAD\_Ratio = activity\_daily.LightActiveDistance / activity\_daily.TotalDistance,

  SAD\_Ratio = activity\_daily.SedentaryActiveDistance / activity\_daily.TotalDistance;

-- Find the average active distance ratio

SELECT ROUND(AVG(VAD\_Ratio), 2) AS Avg\_VAD\_Ratio,

ROUND(AVG(MAD\_Ratio), 2) AS Avg\_MAD\_Ratio,

ROUND(AVG(LAD\_Ratio), 2) AS Avg\_LAD\_Ratio,

ROUND(AVG(SAD\_Ratio), 2) AS Avg\_SAD\_Ratio

FROM `bellabeat-case-study-410302.activity.activity\_daily`

-- Avg. time sleeping and spent in bed

SELECT ROUND(AVG(TotalMinutesAsleep)/60,2) AS Avg\_Sleep\_Hours,

ROUND(AVG(TotalTimeInBed)/60,2) AS Avg\_Hours\_in\_Bed,

ROUND((ROUND(AVG(TotalTimeInBed)/60,2) - ROUND(AVG(TotalMinutesAsleep)/60,2)),2) AS Avg\_Hours\_in\_Bed\_Not\_Sleeping

 FROM `bellabeat-case-study-410302.sleep.sleep\_daily`

-- What time of day are people taking the most steps?

SELECT ActivityHour\_2, SUM(StepTotal) AS Total\_Steps\_Taken

FROM `bellabeat-case-study-410302.steps.steps\_hourly`

GROUP BY ActivityHour\_2

ORDER BY Total\_Steps\_Taken DESC

-- What day of the week are people most/least active?

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

ROUND(AVG(VeryActiveMinutes) + AVG(FairlyActiveMinutes) + AVG(LightlyActiveMinutes),2)AS Active\_Minutes,

FROM `bellabeat-case-study-410302.activity.activity\_daily`

GROUP BY day\_of\_week

ORDER BY Active\_Minutes DESC

-- What day of the week do people sleep the most?

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

ROUND(AVG(TotalMinutesAsleep)/60,2) AS Avg\_Sleep

 FROM `bellabeat-case-study-410302.sleep.sleep\_daily`

 GROUP BY day\_of\_week

 ORDER BY Avg\_Sleep DESC

-- What time of day do people burn the most calories?

SELECT ActivityHour\_Time,

SUM(Calories) AS Calories\_Burned

FROM `bellabeat-case-study-410302.calories.calories\_hourly`

GROUP BY ActivityHour\_Time

ORDER BY Calories\_Burned DESC