**Bellabeat Case Study With SQL & PowerBI**

--Data Cleaning

--Step-1 Checking for unique number of users in each table

--Hourly Calories Table

select count(distinct (id))

from Bellabeat..hourlyCalories\_merged;

--Hourly Intensity Table

select count(distinct (id))

from Bellabeat..hourlyIntensities\_merged;

--Hourly Steps Table

select count(distinct (id))

from Bellabeat..hourlySteps\_merged;

--Daily Activity Table

select count(distinct (id))

from Bellabeat..dailyActivity;

--Daily Sleep Table

select count(distinct (id))

from Bellabeat..sleepDay;

--Step-2 Duplicates record Check

--Hourly Calories Table

select id,ActivityHour,Calories,count(\*)

from Bellabeat..hourlyCalories\_merged

group by id,ActivityHour,Calories

having count(\*)>1;

--Hourly Intensity Table

select id,ActivityHour,TotalIntensity,AverageIntensity, Count(\*)

from Bellabeat..hourlyIntensities\_merged

group by id,ActivityHour,TotalIntensity,AverageIntensity

having count(\*)>1;

--Hourly Steps Table

select id,ActivityHour,StepTotal,count(\*)

from Bellabeat..hourlySteps\_merged

group by id,ActivityHour,StepTotal

having count(\*)>1;

--Activity Table

Select id, activitydate, totalsteps, count(\*)

from Bellabeat..dailyActivity

Group by id, activitydate, totalsteps

Having count(\*)>1 ;

--SleepDay Table

Select id, sleepday, TotalMinutesAsleep,TotalSleepRecords,TotalTimeInBed, count(\*)

From Bellabeat..sleepDay

Group by id, sleepday, TotalMinutesAsleep,TotalSleepRecords,TotalTimeInBed

Having count(\*)>1;

--Deleting common data from Sleep table and creating new sleep table

SELECT DISTINCT id, sleepday, totalsleeprecords, totalminutesasleep, totaltimeinbed

INTO Bellabeat..NewSleep

FROM Bellabeat..sleepDay;

--Check new table if duplicate records are removed from Sleepday\_New table

SELECT id, sleepday, TotalMinutesAsleep, COUNT(\*)

FROM Bellabeat..NewSleep

GROUP BY id , sleepday , TotalMinutesAsleep

HAVING COUNT(\*) > 1;

---Marging Three Hour tables

select \*

from Bellabeat..hourlyCalories\_merged c

left join Bellabeat..hourlyIntensities\_merged i

on c.id = i.id and c.ActivityHour = i.ActivityHour

left join Bellabeat..hourlySteps\_merged s

on i.id = s.id and i.ActivityHour = s.ActivityHour;

--Selecting Relevant Column

select c.id as id,c.ActivityHour as ActivityHour,Calories,TotalIntensity,AverageIntensity,StepTotal

from Bellabeat..hourlyCalories\_merged c

left join Bellabeat..hourlyIntensities\_merged i

on c.id = i.id and c.ActivityHour = i.ActivityHour

left join Bellabeat..hourlySteps\_merged s

on i.id = s.id and i.ActivityHour = s.ActivityHour;

--DataType Changing for Mathematical Operation

select cast(c.id as float) as id,cast(c.ActivityHour as datetime) as ActivityHour,

cast(Calories as int) as Calories,cast(TotalIntensity as int) as TotalIntensity,cast(AverageIntensity as float) as AverageIntensity,

cast(StepTotal as int) as StepTotal

from Bellabeat..hourlyCalories\_merged c

left join Bellabeat..hourlyIntensities\_merged i

on c.id = i.id and c.ActivityHour = i.ActivityHour

left join Bellabeat..hourlySteps\_merged s

on i.id = s.id and i.ActivityHour = s.ActivityHour

--Creating New Table by Marzing three hourly table

-- First, create the new table structure

CREATE TABLE Bellabeat..new\_table (

id FLOAT,

ActivityHour DATETIME,

Calories INT,

TotalIntensity INT,

AverageIntensity FLOAT,

StepTotal INT

);

-- Then, insert the data into the new table

INSERT INTO Bellabeat..new\_table (id, ActivityHour, Calories, TotalIntensity, AverageIntensity, StepTotal)

SELECT

CAST(c.id AS FLOAT) AS id,

CAST(c.ActivityHour AS DATETIME) AS ActivityHour,

CAST(c.Calories AS INT) AS Calories,

CAST(i.TotalIntensity AS INT) AS TotalIntensity,

CAST(i.AverageIntensity AS FLOAT) AS AverageIntensity,

CAST(s.StepTotal AS INT) AS StepTotal

FROM Bellabeat..hourlyCalories\_merged c

LEFT JOIN Bellabeat..hourlyIntensities\_merged i

ON c.id = i.id AND c.ActivityHour = i.ActivityHour

LEFT JOIN Bellabeat..hourlySteps\_merged s

ON i.id = s.id AND i.ActivityHour = s.ActivityHour;

select \*

from Bellabeat..new\_table;

--Top 5 id by AverageCalories

SELECT TOP 5 id, AVG(Calories) AS AverageCalories

FROM Bellabeat..new\_table

GROUP BY id

ORDER BY AverageCalories DESC;

--Busy Hours of for Use of product

with CTE as (

select \*, DATEPART(HOUR, ActivityHour) AS Hour\_part

from new\_table)

select Hour\_part ,avg(StepTotal) AverageSteps

from CTE

group by Hour\_part

order by AverageSteps desc;

select \*,DATEPART(HOUR, ActivityHour) AS Hour\_part,DATEPART(, ActivityHour) AS Hour\_part

from new\_table;

--Mearzing Activity and Sleep table and cast data type for mathematical operation

SELECT A.\*, B.totalsleeprecords, B.totalminutesasleep, B.totaltimeinbed

INTO Bellabeat..DailyActivitySleep

FROM (

SELECT

Id,

CAST(ActivityDate AS DATE) AS Date\_1,

DATENAME(WEEKDAY, CAST(ActivityDate AS DATETIME)) AS Day\_Week,

cast(TotalSteps as int) as TotalSteps,

cast(veryactivedistance as float) as veryactivedistance,

cast(moderatelyactivedistance as float) as moderatelyactivedistance,

cast(lightactivedistance as float) as lightactivedistance,

cast(veryactiveminutes as int) as veryactiveminutes,

cast(fairlyactiveminutes as int) as fairlyactiveminutes,

cast(lightlyactiveminutes as int) as lightlyactiveminutes,

cast(sedentaryminutes as int) as sedentaryminutes,

cast(calories as int) as calories

FROM Bellabeat..dailyActivity

) A

LEFT JOIN (

SELECT

id,

CAST(Sleepday AS DATE) AS Date\_1,

totalsleeprecords,

totalminutesasleep,

totaltimeinbed

FROM Bellabeat..NewSleep

) B

ON A.id = B.id AND A.Date\_1 = B.Date\_1;

**Bellabeat Dashboard**

