

KIRTI  
KHATTAR

# ABOUT ME

---

Development researcher blending historical context and criminological theory with modern data analysis skills. Enthusiastic about supporting social programs through data-driven insights, and contextualized understanding of urban and rural challenges.

---

# ABOUT THE DATA ANALYSIS PROJECT

---

The data analysis project is about the health activity data that provides the basic information about gender, age, height, weight, BMI, daily steps, calories intake, hours of sleep, heart rate, blood pressure, exercise hour per week, smoker, alcohol consumption per week, diabetic and heart disease. This project is analysed through MYSQL queries based on questions to be analysed.

---

# What is the average age of all individuals in the dataset?

```
-- What is the average age of all  
-- individuals in the dataset?  
  
SELECT  
    TRUNCATE(AVG(age), 0) AS avg_age  
FROM  
    health.health_activity_data;
```

Result Grid	
	avg_age
▶	49

# How many individuals are males and how many are females?

```
-- How many individuals are males  
-- and how many are females?
```

```
SELECT  
    gender, COUNT(id)  
FROM  
    health.health_activity_data  
GROUP BY gender;
```

Result Grid				
	gender	COUNT(id)		
▶	Male	523		
	Female	477		

# How many individuals are males and how many are females?



```
-- How many individuals are males  
-- and how many are females?
```

```
SELECT  
    gender, COUNT(id)  
FROM  
    health.health_activity_data  
GROUP BY gender;
```




Result Grid				
	gender	COUNT(id)		
▶	Male	523		
	Female	477		

# Find the minimum, maximum and average of Daily Steps, Hours of Sleep and Calories Intake?




```
-- Find the minimum, maximum and average values for Daily_Steps
-- hours_of_sleep and calories intake.
Select min(Daily_Steps),
max(Daily_Steps),
truncate(avg(Daily_Steps),0) as avg_Daily_Steps
from health.health_activity_data;
```

Result Grid     Filter Rows: <input type="text"/>   Export			
	min(Daily_Steps)	max(Daily_Steps)	avg_Daily_Steps
▶	1016	19931	10717

```
-- Find the minimum, maximum and average values for Daily_Steps
-- hours_of_sleep and calories intake.
Select min(Hours_of_Sleep),
max(Hours_of_Sleep),
truncate(avg(Hours_of_Sleep),0) as avg_Hours_Of_Sleep
from health.health_activity_data;
```

Result Grid     Filter Rows: <input type="text"/>   Export:    Wr			
	min(Hours_of_Sleep)	max(Hours_of_Sleep)	avg_Hours_Of_Sleep
▶	4	10	6



```
-- Find the mininum, maximum and average values for Daily_Steps
-- hours_of_sleep and calories intake.
Select min(calories_intake),
max(calories_intake),
truncate(avg(calories_intake),0) as avg_Calories_Intake
from health.health_activity_data;
```

Result Grid     Filter Rows: <input type="text"/>   Export:    Wrap C			
	min(calories_intake)	max(calories_intake)	avg_Calories_Intake
▶	1201	3498	2327



# Count the number of individuals who are smokers and non-smokers

```
-- count the number of individuals who  
-- are smokers and non-smokers  
select smoker, count(id)  
from health.health_activity_data  
group by smoker;
```

Result Grid					Filter
	smoker	count(id)			
▶	No	809			
	Yes	191			

# What is the average BMI of all the individuals?

```
-- What is the average BMI of all individuals?  
SELECT TRUNCATE(AVG(BMI), 0) as avg_BMI  
FROM health.health_activity_data;
```

Result Grid	
	avg_BMI
▶	26

# Calculate the average daily steps of males and females separately.

```
-- Calculate the average daily steps
-- for males and females separately
SELECT gender, TRUNCATE(AVG(daily_steps), 0) as daily_steps
FROM health.health_activity_data group by gender;
```

Result Grid			Filter Row
	gender	daily_steps	
▶	Male	10960	
	Female	10450	

# Determine the average hours of sleep for individuals categorized into age groups.

```
-- Determine the average hours of sleep for
-- individuals categorized into age groups
SELECT
  CASE
    WHEN age BETWEEN 0 AND 12 THEN 'Children (0-12)'
    WHEN age BETWEEN 13 AND 17 THEN 'Teenagers (13-17)'
    WHEN age BETWEEN 18 AND 25 THEN 'Young Adults (18-25)'
    WHEN age BETWEEN 26 AND 64 THEN 'Adults (26-64)'
    WHEN age >= 65 THEN 'Seniors (65+)' END AS age_group, truncate( AVG(hours_of_sleep),0) AS average_sleep_hours
  FROM health.health_activity_data GROUP BY age_group;
```

Result Grid		Filter Rows:
	age_group	average_sleep_hours
	Adults (26-64)	6
	Seniors (65+)	6
	Young Adults (18-25)	6



# Find the average BMI of individuals based on their daily steps activity levels.

```
-- Find the average BMI for individuals based on their Daily_Steps
-- activity levels (e.g., less than 5000 steps = 'Low',
-- 5000-10000 = 'Medium', over 10000 = 'High')
SELECT CASE WHEN Daily_Steps < 5000 THEN 'Low' WHEN Daily_Steps
BETWEEN 5000 AND 10000 THEN 'Medium' ELSE 'High' END AS ActivityLevel,
round(AVG(BMI), 2) as avg_BMI FROM health.health_activity_data GROUP BY ActivityLevel;
```

Result Grid		Filter Rows:
	ActivityLevel	avg_BMI
▶	Medium	26.78
	High	26.69
	Low	26.78

# Calculate the average heart rate of individuals who have heart disease and those who do not have.

```
-- Calculate the average Heart_Rate for individuals
-- who have heart disease and those who do not
Select heart_disease, round(avg(heart_rate), 2) as avg_heart_rate
from health.health_activity_data group by Heart_Disease;
```

Result Grid     Filter Rows: <input type="text"/>		
	heart_disease	avg_heart_rate
▶	No	84.82
	Yes	83.68

# Determine the percentage of smokers who have heart disease and non-smokers who have heart disease.

```
-- Determine the percentage of smokers who have Heart_Disease
-- versus non-smokers who have Heart_Disease
SELECT
    Smoker AS Smoking_Status,
    (SUM(CASE WHEN Heart_Disease = 'Yes' THEN 1 ELSE 0 END) * 100.0) / COUNT(*) AS Percentage_With_Heart_Disease
FROM
    health.health_activity_data
GROUP BY
    Smoking_Status;
```

Result Grid			Filter Rows:
	Smoking_Status	Percentage_With_Heart_Disease	
	No	9.64153	
	Yes	7.85340	



# Identify the top 10 individuals by ID, Age, Gender with the highest daily steps who also have a BMI greater than 25.

```
-- Identify the top 10 individuals (by ID, Age, Gender)
-- with the highest Daily_Steps who also have a BMI greater than 25.
SELECT ID, Age, Gender, Daily_Steps, BMI
FROM health.health_activity_data WHERE BMI > 25
ORDER BY Daily_Steps DESC LIMIT 10;
```

	ID	Age	Gender	Daily_Steps	BMI
▶	113	79	Male	19931	25.64
	430	76	Female	19925	30.07
	325	33	Female	19816	32.8
	707	46	Female	19748	27.42
	649	43	Male	19734	27
	270	66	Male	19725	34.63
	521	67	Male	19714	28.46
	93	59	Female	19678	26.65
	729	24	Male	19621	31.58
	460	34	Male	19581	32.78