**Cardiovascular Disease dataset**

#### Data description

There are 3 types of input features:

* Objective: factual information;
* Examination: results of medical examination;
* Subjective: information given by the patient.

Features:

1. Age | Objective Feature | age | int (days)
2. Height | Objective Feature | height | int (cm) |
3. Weight | Objective Feature | weight | float (kg) |
4. Gender | Objective Feature | gender | categorical code |
5. Systolic blood pressure | Examination Feature | ap\_hi | int |
6. Diastolic blood pressure | Examination Feature | ap\_lo | int |
7. Cholesterol | Examination Feature | cholesterol | 1: normal, 2: above normal, 3: well above normal |
8. Glucose | Examination Feature | gluc | 1: normal, 2: above normal, 3: well above normal |
9. Smoking | Subjective Feature | smoke | binary |
10. Alcohol intake | Subjective Feature | alco | binary |
11. Physical activity | Subjective Feature | active | binary |
12. Presence or absence of cardiovascular disease | Target Variable | cardio | binary |

Q1 : What are the factors that are causing Cardiovascular diseases

Q2 : Age pattern for Cardio patients

Q3 :

**Understanding the domine concept**

**Cholesterol**

The levels of cholesterol are often divided into different categories:

**Normal**: This level indicates that the cholesterol levels are within the healthy range. It typically means that the levels of LDL cholesterol (often referred to as "bad" cholesterol) are lower while the levels of HDL cholesterol ("good" cholesterol) are higher.

**Above Normal**: This category suggests that the cholesterol levels, especially LDL cholesterol, are higher than the recommended range. Elevated LDL cholesterol can lead to the accumulation of cholesterol in the arteries, potentially increasing the risk of heart disease.

**Well Above Normal**: This level signifies significantly elevated cholesterol levels, especially LDL cholesterol, which substantially increases the risk of cardiovascular problems. Having well above normal cholesterol levels may indicate a greater risk of developing heart disease or experiencing related health issues.

**Bood pressure**

Blood pressure categories help determine the level of risk associated with a person's blood pressure readings. They are typically classified as follows:

**Normal**: Blood pressure falls within a healthy range. It usually indicates that your heart is functioning well, and your risk of heart disease or other health issues related to high blood pressure is lower.

**Elevated**: This category suggests that the blood pressure readings are higher than the optimal range but not high enough to be classified as hypertension. It's a warning sign that indicates a higher risk of developing hypertension in the future if preventive measures are not taken.

**Hypertension Stage 1**: This stage indicates moderately high blood pressure. It means that the systolic pressure (the pressure when the heart beats) ranges from 130 to 139 mm Hg or the diastolic pressure (the pressure when the heart is at rest between beats) ranges from 80 to 89 mm Hg. At this stage, lifestyle changes are often recommended to manage blood pressure and reduce the risk of complications.

**Hypertension Stage 2:** This stage indicates a more severe form of high blood pressure. It involves systolic pressure of 140 mm Hg or higher or diastolic pressure of 90 mm Hg or higher. At this stage, lifestyle modifications and often medications are recommended to control blood pressure and lower the risk of heart disease, stroke, and other health issues.

**Glucose**

**Normal**: Glucose levels within the healthy range indicate that the body is processing sugar properly. Typically, this means that the fasting blood sugar level is between 70 to 100 milligrams per deciliter (mg/dL).

**Above Normal**: This could imply that the glucose levels are higher than the optimal range but might not be classified as diabetic. It might indicate a borderline or prediabetic condition, where the fasting blood sugar levels are slightly elevated, usually between 100 to 125 mg/dL.

**Well Above Normal**: This category typically suggests significantly elevated glucose levels, often associated with diabetes or uncontrolled blood sugar. It indicates fasting blood sugar levels of 126 mg/dL or higher, which is considered diagnostic for diabetes.

Q. What the data is telling

"In examining the dataset provided, it encapsulates key patient details including age, gender, height, and weight, coupled with comprehensive medical reports concerning cardiovascular health. This combination of information offers a comprehensive view into the factors potentially influencing cardiovascular disease.

By meticulously examining this dataset, we aim to uncover correlations, trends, and predictive indicators that may contribute significantly to understanding the dynamics of cardiovascular health within this cohort. The integration of demographic factors with medical insights is poised to offer actionable intelligence, aiding healthcare professionals and policymakers in refining preventive strategies and enhancing patient care.

The information contained within this dataset holds the promise of revealing valuable connections between patient attributes and cardiovascular health markers, fostering a deeper understanding of this critical health concern."

**Process**

**Modifying the columns**

* Changing the Gender
* =IF(C2=1, "F", IF(C2=2, "M", ""))
* Changing the Cholesterol
* =IF(I2=1,"Normal",IF(I2=2,"Above Normal",(IF(I2=3,"Well Above Normal"))))
* Changing Glucose
* =IF(K2=1,"Normal",IF(K2=2,"Above Normal",(IF(K2=3,"Well Above Normal"))))
* Changing Smoking, Alcohol intake, Physical activity by the below sample formula
* =IF(Q2=0,"No",IF(Q2=1,"Yes",""))
* Changing Presence or absence of cardiovascular disease
* =IF(S2=0,"Absent",IF(S2=1,"Present",""))

**SQL Analysis**

--checking duplicate id's

select id, count(\*) from [Cardiovascular Disease]

group by id

having count(\*) > 1

-- inserting new​ column for BMI category

ALTER TABLE [Cardiovascular Disease]

ADD bmi\_category NVARCHAR(100);

-- inserting values

UPDATE [Cardiovascular Disease]

SET bmi\_category =

CASE

WHEN bmi < 18.5 THEN 'Underweight'

WHEN bmi BETWEEN 18.5 AND 24.99999 THEN 'Normal weight'

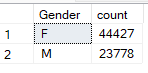
WHEN bmi BETWEEN 25 AND 29.99999 THEN 'Overweight'

WHEN bmi >= 30 THEN 'Obesity'

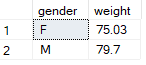
ELSE 'Unknown' -- or any other

END;

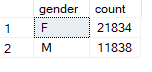
* count of female and male



* Average **Weight by Gender** where cardiovascular disease present



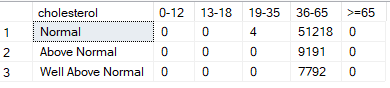
* **Gender count** by cardiovascular disease present



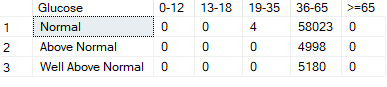
* count the occurrences of individuals with **cardiovascular disease** in different age groups



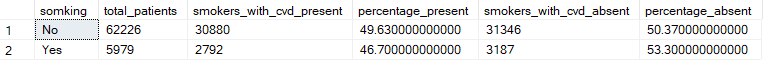
* count the occurrences of individuals with **cholesterol level** in different age groups



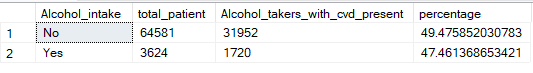
* count the occurrences of individuals with **glucose level** in different age groups



* Smoking Percentage in Both present and absent patient



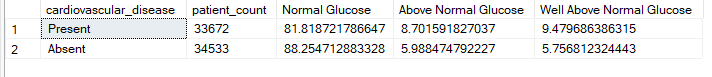
* Alcohol taker percentage where cardiovascular disease is present



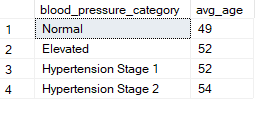
* cholesterol percentage of patient by cardiovascular disease



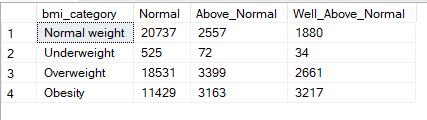
* Glucose percentage of patient by cardiovascular disease



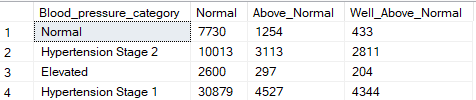
* Average age of patients by blood pressure category



* Count of patient’s cholesterol by BMI category



* Count of patient’s cholesterol by blood pressure



-- Risk Stratification analysis

-- patients may need to change their lifestyle and the patients have higher risk of getting cardiovascular disease

select id, cardiovascular\_disease, age\_years from [Cardiovascular Disease]

where cholesterol = 'Above Normal'

and Blood\_pressure\_category = 'Hypertension Stage 1' and bmi\_category in ('Overweight','Obesity')