

# Business Intelligence cis-5270

**Team Members:** Nagender Chauhan & Lekha Ajit

**Dataset link:** <https://www.kaggle.com/datasets/rahulsharma51/heart-failure-lagacy-dataset>

## **Dataset description:**

The data analysis project focuses on exploring the relationships between various factors related to cardiovascular health. The dataset includes information on age, sex, chest pain type, resting blood pressure, cholesterol levels, fasting blood sugar levels, resting electrocardiogram results, maximum heart rate achieved during exercise, presence of exercise-induced angina, and the oldpeak value (ST depression). The objective of the study is to gain insights into how these factors are related to each other and how they impact overall cardiovascular health. The analysis will involve statistical methods to identify patterns and correlations among the variables, as well as machine learning algorithms to build predictive models. The results of the project will provide valuable information for healthcare professionals and policymakers to improve preventive measures and treatment options for cardiovascular diseases.

## **Attributes**

Age: age of the patient [years]

Sex: sex of the patient [M: Male, F: Female]

Height: height of the patient

Smoke: does the person smoke[Y: Yes, N: No]

ChestPainType: chest pain type [TA: Typical Angina, ATA: Atypical Angina, NAP: Non-Anginal Pain, ASY: Asymptomatic]

RestingBP: resting blood pressure [mm Hg]

Cholesterol: serum cholesterol [mm/dl]

FastingBS: fasting blood sugar

ATA: maximum heart ATA

RestingECG: resting electrocardiogram results [Normal: Normal, ST:abnormality ]

MaxHR: maximum heart rate

ExerciseAngina: exercise-induced angina [Y: Yes, N: No]

Oldpeak: oldpeak = ST [Numeric value measured in depression]

ST\_Slope: the slope of the peak exercise ST segment [Up: upsloping, Flat: flat, Down: downsloping]

HeartDisease: output class [1: heart disease, 0: Normal]

## **Questions:**

1. Is there a significant difference in the prevalence of chest pain types among different age groups and sexes?
2. How does serum cholesterol level relate to resting blood pressure, and are these factors significant predictors of heart disease risk?
3. Is there a correlation between maximum heart rate achieved during exercise and the occurrence of exercise-induced angina?
4. How does the presence of ST segment depression (as indicated by old peak values) on electrocardiogram results relate to fasting blood sugar levels, and does this relationship vary by sex or age?
5. What factors are most strongly associated with the occurrence of exercise-induced angina, and how do these factors differ by sex and age?