Heart Disease Prediction Dataset — Project Description

Dataset Overview

This dataset contains **270 patient records** with **14 clinical attributes** used to analyze and predict the presence of heart disease. Each row represents a patient, and the target variable indicates whether heart disease is **Present** or **Absent**.

Features (Columns)

- **Age** – Patient age (years)
- **Sex** – Gender indicator (binary encoded)
- **Chest pain type** – Categorized chest pain levels
- **BP** – Resting blood pressure
- **Cholesterol** – Serum cholesterol level
- **FBS over 120** – Fasting blood sugar > 120 mg/dl (binary)
- **EKG results** – Resting electrocardiographic results
- **Max HR** – Maximum heart rate achieved
- **Exercise angina** – Exercise-induced angina (binary)
- **ST depression** – ST depression induced by exercise
- **Slope of ST** – Slope of the peak exercise ST segment
- **Number of vessels fluro** – Number of major vessels colored by fluoroscopy
- **Thallium** – Thallium stress test result
- **Heart Disease** – **Target variable** (Presence / Absence)

Objective

The objective of this dataset is to **predict the presence of heart disease** using patient clinical indicators. This supports binary classification modeling and exploratory analysis for healthcare risk assessment.

Data Characteristics

- **Total records:** 270
- **Total features:** 13 input features + 1 target
- **Target type:** Binary categorical
- **Data type:** Structured, tabular clinical data

Potential Use Cases

- Exploratory Data Analysis (EDA) of cardiac risk factors
- Feature importance analysis for medical indicators
- Binary classification using machine learning models
- Healthcare analytics and predictive modeling demos