

Heart Disease (UCI) Dataset - Data Dictionary

■ Target Variable:

num — Target variable representing presence of heart disease.

Values:

- 0 = No heart disease

- 1, 2, 3, 4 = Presence of heart disease (often converted to binary: 0 = healthy, 1 = disease)

■ Input Features:

age — Age in years

sex — Sex (1 = male, 0 = female)

cp — Chest pain type (values 1–4)

trestbps — Resting blood pressure (mm Hg)

chol — Serum cholesterol in mg/dl

fbs — Fasting blood sugar > 120 mg/dl (1 = true, 0 = false)

restecg — Resting electrocardiographic results (0, 1, 2)

thalach — Maximum heart rate achieved

exang — Exercise induced angina (1 = yes, 0 = no)

oldpeak — ST depression induced by exercise relative to rest

slope — Slope of the peak exercise ST segment (1 = upsloping, 2 = flat, 3 = downsloping)

ca — Number of major vessels (0–3) colored by fluoroscopy

thal — Thalassemia (3 = normal, 6 = fixed defect, 7 = reversible defect)

Source: UCI Machine Learning Repository / Kaggle Heart Disease Dataset