Objectives

1. What are the risk-factors associated with the presence of a heart disease?
2. Can you provide to me a way to predict easily the risk of having a heart disease? I would like to be able on my own to know this risk for future patients.

Data

This database contains 14 variables for 303 randomly selected patients.

The target is a binary variable: 1 for presence of a heart disease and 0 otherwise.

The other 13 variables are attributes of the patients and here is a description.

Variables description

1. age: age in years
2. sex: sex (1 = male; 0 = female)
3. cp: chest pain type
-- Value 0: typical angina
-- Value 1: atypical angina
-- Value 2: non-anginal pain
-- Value 3: asymptomatic
4. trestbps: resting blood pressure (in mm Hg on admission to the hospital)
5. chol: serum cholestoral in mg/dl
6. fbs: (fasting blood sugar > 120 mg/dl) (1 = true; 0 = false)
7. restecg: resting electrocardiographic results
-- Value 0: normal
-- Value 1: having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of > 0.05 mV)
-- Value 2: showing probable or definite left ventricular hypertrophy by Estes' criteria
8. thalach: maximum heart rate achieved
9. exang: exercise induced angina (1 = yes; 0 = no)
10. oldpeak = ST depression induced by exercise relative to rest
11. slope: the slope of the peak exercise ST segment
-- Value 0: downsloping
-- Value 1: upsloping
-- Value 2: flat
12. ca: number of major vessels (0-4) colored by fluoroscopy
13. thal: 1 = normal; 2 = fixed defect; 3 = reversible defect