**Additional Information**

This database contains 76 attributes, but all published experiments refer to using a subset of 14 of them. In particular, the Cleveland database is the only one that has been used by ML researchers to date. The "goal" field refers to the presence of heart disease in the patient. It is integer valued from 0 (no presence) to 4. Experiments with the Cleveland database have concentrated on simply attempting to distinguish presence (values 1,2,3,4) from absence (value 0). The names and social security numbers of the patients were recently removed from the database, replaced with dummy values. One file has been "processed", that one containing the Cleveland database. All four unprocessed files also exist in this directory.

**Variables Table**

| **Variable Name** | **Role** | **Type** | **Demographic** | **Description** | **Units** | **Missing Values** |
| --- | --- | --- | --- | --- | --- | --- |
| age | Feature | Integer | Age | Age in years | years | no |
| sex | Feature | Categorical | Sex | Sex (1 = male, 0 = female) |  | no |
| cp | Feature | Categorical |  | cp: chest pain type  -- Value 1: typical angina  -- Value 2: atypical angina  -- Value 3: non-anginal pain  -- Value 4: asymptomatic |  | no |
| trestbps | Feature | Integer |  | resting blood pressure (in mm Hg on admission to the hospital) | mm Hg | no |
| chol | Feature | Integer |  | serum cholestoral in mg/dl | mg/dl | no |
| fbs | Feature | Categorical |  | (fasting blood sugar > 120 mg/dl) (1 = true; 0 = false) |  | no |
| restecg | Feature | Categorical |  | 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 |  | no |
| thalach | Feature | Integer |  | maximum heart rate achieved |  | no |
| exang | Feature | Categorical |  | exercise induced angina (1 = yes; 0 = no) |  | no |
| oldpeak | Feature | Integer |  | ST depression induced by exercise relative to rest |  | no |
| slope | Feature | Categorical |  | the slope of the peak exercise ST segment  -- Value 1: up sloping  -- Value 2: flat  -- Value 3: down sloping |  | no |
| ca | Feature | Integer |  | number of major vessels (0-3) colored by fluoroscopy |  | yes |
| Thal | Feature | Categorical |  | Thalassemia: 3 = normal; 6 = fixed defect; 7 = reversable defect |  | yes |
| num | Target | Integer |  | diagnosis of heart disease (angiographic disease status)  -- Value 0: < 50% diameter narrowing  -- Value 1: > 50% diameter narrowing  (in any major vessel: attributes 59 through 68 are vessels)  (values 1,2,3,4) from absence (value 0). |  | no |

Additional Variable Information

Only 14 attributes used:

1. #3 (age)

2. #4 (sex)

3. #9 (cp)

4. #10 (trestbps)

5. #12 (chol)

6. #16 (fbs)

7. #19 (restecg)

8. #32 (thalach)

9. #38 (exang)

10. #40 (oldpeak)

11. #41 (slope)

12. #44 (ca)

13. #51 (thal)

14. #58 (num) (the predicted attribute)