1. Title: Boston Housing Data

2. Sources:

(a) Origin: This dataset was taken from the StatLib library which is

maintained at Carnegie Mellon University.

(b) Creator: Harrison, D. and Rubinfeld, D.L. 'Hedonic prices and the

demand for clean air', J. Environ. Economics & Management,

vol.5, 81-102, 1978.

(c) Date: July 7, 1993

3. Past Usage:

- Used in Belsley, Kuh & Welsch, 'Regression diagnostics ...', Wiley,

1980. N.B. Various transformations are used in the table on

pages 244-261.

- Quinlan,R. (1993). Combining Instance-Based and Model-Based Learning.

In Proceedings on the Tenth International Conference of Machine

Learning, 236-243, University of Massachusetts, Amherst. Morgan

Kaufmann.

4. Relevant Information:

Concerns housing values in suburbs of Boston.

5. Number of Instances: 506

6. Number of Attributes: 13 continuous attributes (including "class"

attribute "MEDV"), 1 binary-valued attribute.

7. Attribute Information:

1. CRIM per capita crime rate by town

2. ZN proportion of residential land zoned for lots over

25,000 sq.ft.

3. INDUS proportion of non-retail business acres per town

4. CHAS Charles River dummy variable (= 1 if tract bounds

river; 0 otherwise)

5. NOX nitric oxides concentration (parts per 10 million)

6. RM average number of rooms per dwelling

7. AGE proportion of owner-occupied units built prior to 1940

8. DIS weighted distances to five Boston employment centres

9. RAD index of accessibility to radial highways

10. TAX full-value property-tax rate per $10,000

11. PTRATIO pupil-teacher ratio by town

12. B 1000(Bk - 0.63)^2 where Bk is the proportion of blacks

by town

13. LSTAT % lower status of the population

14. MEDV Median value of owner-occupied homes in $1000's

8. Missing Attribute Values: None.

1. Con variables con correlación > |0.4|
2. Con variables con correlación > |0.4|

pero quitar colineales

1. Con variables con correlación > |0.3|