Module: BENV0119 MLiSB

Boston Housing Data Readme

The data

This dataset contains information collected by the US Census service concerning housing in the area of Boston, Massachusetts. The data was originally published by Harrison, D. and Rubinfeld, D.L. `*Hedonic prices and the demand for clean air*', J. Environ. Economics & Management, vol.5, 81-102, 1978.

Input variables:

The input variables varied within the simulation are provided in table 1.

Table 1. Input parameters

#	Column title	Parameter description
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 oxide 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 bytown
12	B	$1000(Bk - 0.63)^2$ where Bk is the proportion of blacks by town
13	LSTAT	% lower status of the population

Output variables:

The output variables we will typically try to predict.

Table 2. Output parameters

#	Column title	Parameter description
1	MV	Median value of owner-occupied homes in $1000s