

House Prices in Suburbs of Boston

Objective

The **medv** variable is the target variable, which reflects the median prices of the houses. A prediction of prices will be by analysing the relationship between crim (crime) and the **medv** (prices) variable.

Data Description

Dataset has 506 rows and 14 columns and 506 columns

This dataset contains the following columns:

crim

per capita crime rate by town.

zn

proportion of residential land zoned for lots over 25,000 sq.ft.

indus

proportion of non-retail business acres per town.

chas

Charles River dummy variable (= 1 if tract bounds river; 0 otherwise).

nox

nitrogen oxides concentration (parts per 10 million).

rm

average number of rooms per dwelling.

age

proportion of owner-occupied units built prior to 1940.

dis

weighted mean of distances to five Boston employment centres.

rad

index of accessibility to radial highways.

tax

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

ptratio

pupil-teacher ratio by town.

b

$1000(B_k - 0.63)^2$ where B_k is the proportion of blacks by town.

lstat

lower status of the population (percent).

medv

median value of owner-occupied homes in \$1000s.

Data Source

Harrison, D. and Rubinfeld, D.L. (1978) Hedonic prices and the demand for clean air. *J. Environ. Economics and Management* 5, 81–102.

Belsley D.A., Kuh, E. and Welsch, R.E. (1980) *Regression Diagnostics. Identifying Influential Data and Sources of Collinearity*. New York: Wiley.