```
import numpy as np
import pandas as pd
%matplotlib inline
%load_ext autotime
```

Loading the Boston House Pricing Dataset

```
from sklearn.datasets import load_boston
```

```
time: 1.1e-05s
```

```
%ls ./data/boston*
```

```
./data/boston_corrected.txt  ./data/boston.csv  ./data/boston.txt
```

```
time: 0.0s
```

```
boston = load_boston()
print(boston.keys())
print(boston.DESCR)
```

```
dict_keys(['data', 'target', 'feature_names', 'DESCR', 'filename', 'data_module'])
.. _boston_dataset:

Boston house prices dataset
---------------------------

**Data Set Characteristics:**

    :Number of Instances: 506

    :Number of Attributes:: 13 numeric/categorical predictive. Median Value (attribute 14) is usually the target

    :Attribute Information (in order):
        - 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      nitric 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 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(Bk - 0.63)^2 where Bk is the proportion of blacks by town
        - LSTAT    % lower status of the population
        - MEDV     Median value of owner-occupied homes in $1000's

    :Missing Attribute Values: None

    :Creator: Harrison, D. and Rubinfeld, D.L.

This is a copy of UCI ML Housing dataset.
http://archive.ics.uci.edu/ml/datasets/Housing


This dataset was taken from the StatLib library and is maintained by Carnegie Mellon University.

The Boston Housing dataset of Harrison, D. and Rubinfeld, D.L. 'Hedonic
prices and the demand for clean air', J. Environ. Economics & Management,
vol.5, p.81-102, 1978.  Used in Belsley, Kuh & Welsch, 'Regression diagnostics
...Identifying Influential data and sources of collinearity', Wiley, 1980.
244-261.  Used in 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.

.. topic:: References

   - Belsley, Kuh, Welsch, 'Regression diagnostics ...Identifying Influential data and sources of collinearity', Wiley, 1980. 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.
```

```
print(boston.data)
```

```
[[ 6.32000  18.00000   2.31000   0.00000   0.53800   6.57500  65.20000
   4.09000   1.00000 296.00000  15.30000 396.90000   4.98000]
 [ 0.27310   0.00000   7.07000   0.00000   0.46900   6.42100  78.90000
   4.96710   2.00000 242.00000  17.80000 396.90000   9.14000]
 [ 0.27290   0.00000   7.07000   0.00000   0.46900   7.18500  61.10000
   4.96710   2.00000 242.00000  17.80000 392.83000   4.03000]
 [ 0.06905   0.00000   2.18000   0.00000   0.45800   6.87500  45.80000
   6.06470   3.00000 222.00000  18.70000 396.90000   5.33000]
 [ 0.02985   0.00000   2.18000   0.00000   0.45800   6.43000  58.70000
   6.06470   3.00000 222.00000  18.70000 394.12000   5.21000]
 [ 0.08829  12.50000   7.87000   0.00000   0.52400   6.01200  66.60000
   5.56050   5.00000 311.00000  15.20000 395.60000  12.43000]
 [ 0.14455  12.50000   7.87000   0.00000   0.52400   6.172000 96.10000
   5.96150   5.00000 311.00000  15.20000 396.90000  19.15000]
 [ 0.21124  12.50000   7.87000   0.00000   0.52400   5.63100  100.00000
   6.08210   5.00000 311.00000  15.20000 386.63000  29.93000]
 [ 0.17004  12.50000   7.87000   0.00000   0.52400   6.004000 85.90000
   6.59210   5.00000 311.00000  15.20000 386.71000  17.10000]
 [ 0.22489  12.50000   7.87000   0.00000   0.52400   6.377000 94.30000
   6.34670   5.00000 311.00000  15.20000 392.52000  20.45000]
 [ 0.09178   0.00000   4.05000   0.00000   0.51000   6.416000 84.10000
   5.21440   5.00000 296.00000  16.60000 396.90000  15.39000]
 [ 0.05644   0.00000   4.05000   0.00000   0.51000   6.048000 28.90000
   5.87890   5.00000 296.00000  16.60000 396.90000   8.01000]
 [ 0.10574   0.00000   4.05000   0.00000   0.51000   5.572000 77.80000
   5.29500   5.00000 296.00000  16.60000 394.12000   5.68000]
 [ 0.06888   0.00000   2.46000   0.00000   0.48800   6.144000 62.20000
   2.25740   3.00000 193.00000  17.80000 396.90000   9.18000]
 [ 0.09164   0.00000   2.46000   0.00000   0.48800   5.709000 54.30000
   2.18240   3.00000 193.00000  17.80000 396.90000  10.83000]
 [ 0.10469   0.00000   2.46000   0.00000   0.48800   6.021000 40.10000
   2.19100   3.00000 193.00000  17.80000 396.90000  16.22000]
 [ 0.06664   0.00000   2.46000   0.00000   0.48800   5.710000 18.50000
   2.19800   3.00000 193.00000  17.80000 394.67000  13.16000]
 [ 0.07022   0.00000   2.46000   0.00000   0.48800   6.021000 19.10000
   2.22400   3.00000 193.00000  17.80000 396.90000   9.53000]
 [ 0.05011   0.00000   2.46000   0.00000   0.48800   5.872000 75.10000
   2.24750   3.00000 193.00000  17.80000 396.90000   8.05000]
 [ 0.05360  21.00000   5.64000   0.00000   0.43900   5.870000 69.10000
   2.47860   4.00000 243.00000  16.80000 396.90000   8.26000]
 [ 0.07151   0.00000   5.64000   0.00000   0.43900   6.004000 79.90000
   2.42780   4.00000 243.00000  16.80000 396.90000   8.80000]
 [ 0.07244   0.00000   5.64000   0.00000   0.43900   5.709000 70.30000
   2.75220   4.00000 243.00000  16.80000 396.90000   7.39000]
 [ 0.10050   0.00000   5.64000   0.00000   0.43900   5.390000 79.10000
   2.77820   4.00000 243.00000  16.80000 396.90000   7.99000]
 [ 0.12744   0.00000   5.64000   0.00000   0.43900   5.990000 77.80000
   2.72930   4.00000 243.00000  16.80000 396.90000   7.70000]
 [ 0.06129   0.00000   5.64000   0.00000   0.43900   5.949000 30.60000
   2.72930   4.00000 243.00000  16.80000 396.90000   6.27000]
 [ 0.06076   0.00000   5.64000   0.00000   0.43900   6.096000 28.30000
   2.83830   4.00000 243.00000  16.80000 396.90000   7.19000]
 [ 0.10959   0.00000   5.64000   0.00000   0.43900   5.834000 56.50000
   2.83830   4.00000 243.00000  16.80000 396.90000   8.81000]
 [ 0.04741   0.00000   5.64000   0.00000   0.43900   5.935000 70.10000
   2.82260   4.00000 243.00000  16.80000 396.90000   8.16000]
 [ 0.02763  75.00000   2.95000   0.00000   0.42800   6.595000 21.80000
   5.40110   3.00000 252.00000  18.30000 395.63000   4.45000]
 [ 0.02731   0.00000   2.95000   0.00000   0.42800   6.830000 35.60000
   5.40110   3.00000 252.00000  18.30000 395.62000   1.98000]
 [ 0.02544   0.00000   2.95000   0.00000   0.42800   6.429000 40.10000
   5.40110   3.00000 252.00000  18.30000 395.62000   4.84000]
 [ 0.02875  75.00000   2.95000   0.00000   0.42800   6.373000 56.10000
   5.49640   3.00000 252.00000  18.30000 395.62000   8.44000]
 [ 0.03548  80.00000   3.64000   0.00000   0.39200   6.416000 84.10000
   5.48160   1.00000 315.00000  16.40000 395.18000   8.80000]
 [ 0.03274  34.40000   3.09000   0.00000   0.44400   6.482000 68.20000
   3.37510   4.00000 398.00000  18.20000 396.90000   4.86000]
 [ 0.03465  34.40000   3.09000   0.00000   0.44400   6.454000 48.50000
   3.37510   4.00000 398.00000  18.20000 396.90000   5.64000]
 [ 0.06162   0.00000   4.15000   0.00000   0.42900   5.304000 77.30000
   3.32440   4.00000 390.00000  18.40000 396.90000   7.26000]
 [ 0.07013   0.00000   4.15000   0.00000   0.42900   5.957000 62.00000
   3.10750   4.00000 390.00000  18.40000 396.90000   8.47000]
 [ 0.10000   0.00000   4.15000   0.00000   0.42900   6.826000 74.40000
   3.08190   4.00000 390.00000  18.40000 396.90000   8.67000]
 [ 0.11744   0.00000   4.15000   0.00000   0.42900   6.480000 77.10000
   3.26340   4.00000 390.00000  18.40000 396.90000   9.29000]
 [ 0.09378  12.50000   7.87000   0.00000   0.52400   6.012000 66.10000
   5.56050   5.00000 311.00000  15.20000 395.60000  12.43000]
 [ 0.62976   0.00000   8.14000   0.00000   0.53800   5.949000 61.80000
   4.70710   4.00000 307.00000  21.00000 396.90000   8.26000]
 [ 0.63796   0.00000   8.14000   0.00000   0.53800   6.096000 84.50000
   4.48670   4.00000 307.00000  21.00000 380.02000  10.26000]
 [ 0.62739   0.00000   8.14000   0.00000   0.53800   5.834000 56.50000
   4.48670   4.00000 307.00000  21.00000 395.62000   8.47000]
 [ 1.05393   0.00000   8.14000   0.00000   0.53800   5.935000 29.30000
   4.49860   4.00000 307.00000  21.00000 386.85000   6.58000]
 [ 0.78420   0.00000   8.14000   0.00000   0.53800   5.990000 81.70000
   4.26580   4.00000 307.00000  21.00000 386.75000  14.67000]
 [ 0.80271   0.00000   8.14000   0.00000   0.53800   5.456000 36.60000
   3.79965   4.00000 307.00000  21.00000 288.99000  11.69000]
 [ 0.77011   0.00000   8.14000   0.00000   0.53800   5.727000 69.50000
   3.79965   4.00000 307.00000  21.00000 390.95000  11.28000]
 [ 0.18836   0.00000   6.91000   0.00000   0.44800   5.786000 54.40000
   5.72321   3.00000 233.00000  17.90000 396.90000   7.18000]
 [ 0.35114   0.00000   6.91000   0.00000   0.44800   6.381000 62.60000
   5.72321   3.00000 233.00000  17.90000 390.07000   7.43000]
 [ 0.26838   0.00000   6.91000   0.00000   0.44800   6.267000 68.10000
   5.72321   3.00000 233.00000  17.90000 396.90000   7.67000]
 [ 0.23912   0.00000   6.91000   0.00000   0.44800   6.369000 67.20000
   5.72321   3.00000 233.00000  17.90000 390.39000   9.48000]
 [ 0.17783   0.00000   6.91000   0.00000   0.44800   5.745000 51.10000
   5.72321   3.00000 233.00000  17.90000 393.74000   8.65000]
 [ 0.22438   0.00000   6.91000   0.00000   0.44800   6.854000 69.30000
   5.72321   3.00000 233.00000  17.90000 396.90000   8.10000]
 [ 0.06263   0.00000   6.91000   0.00000   0.44800   6.590000 51.90000
   5.72321   3.00000 233.00000  17.90000 390.94000   5.52000]
 [ 0.04527   0.00000   6.91000   0.00000   0.44800   6.120000 79.30000
   5.72321   3.00000 233.00000  17.90000 396.90000   5.29000]
 [ 0.06076   0.00000   6.91000   0.00000   0.44800   6.976000 51.10000
   5.72321   3.00000 233.00000  17.90000 396.90000   7.22000]
 [ 0.10959   0.00000   6.91000   0.00000   0.44800   6.794000 80.80000
   5.72321   3.00000 233.00000  17.90000 396.90000   6.72000]
 [ 0.04741   0.00000   6.91000   0.00000   0.44800   6.030000 74.40000
   5.72321   3.00000 233.00000  17.90000 391.98000   7.51000]
 [ 0.26838   0.00000   6.91000   0.00000   0.44800   6.267000 68.10000
   5.72321   3.00000 233.00000  17.90000 396.90000   7.67000]
 [ 0.23912   0.00000   6.91000   0.00000   0.44800   6.369000 67.20000
   5.72321   3.00000 233.00000  17.90000 390.39000   9.48000]
 [ 0.17783   0.00000   6.91000   0.00000   0.44800   5.745000 51.10000
   5.72321   3.00000 233.00000  17.90000 393.74000   8.65000]
 [ 0.22438   0.00000   6.91000   0.00000   0.44800   6.854000 69.30000
   5.72321   3.00000 233.00000  17.90000 396.90000   8.10000]
 [ 0.06263   0.00000   6.91000   0.00000   0.44800   6.590000 51.90000
   5.72321   3.00000 233.00000  17.90000 390.94000   5.52000]
 [ 0.04527   0.00000   6.91000   0.00000   0.44800   6.120000 79.30000
   5.72321   3.00000 233.00000  17.90000 396.90000   5.29000]
 [ 0.06076   0.00000   6.91000   0.00000   0.44800   6.976000 51.10000
   5.72321   3.00000 233.00000  17.90000 396.90000   7.22000]
 [ 0.10959   0.00000   6.91000   0.00000   0.44800   6.794000 80.80000
   5.72321   3.00000 233.00000  17.90000 396.90000   6.72000]
 [ 0.04741   0.00000   6.91000   0.00000   0.44800   6.030000 74.40000
   5.72321   3.00000 233.00000  17.90000 391.98000   7.51000]
```

```
print(boston.target)
```

```
[ 24.00000  21.60000  34.70000  33.40000  36.20000  28.70000  22.90000
  16.10000  18.90000  15.00000  18.90000  21.70000  20.40000  18.20000
  19.90000  23.10000  17.50000  20.20000  18.20000  13.60000  19.60000
  15.20000  14.50000  15.60000  13.90000  16.80000  21.70000  18.30000
  20.10000  29.00000  23.10000  19.70000  18.30000  21.60000  20.00000
  16.30000  14.60000  13.40000  15.60000  11.90000  20.80000  16.80000
  14.90000  21.90000  27.50000  21.90000  23.10000  50.00000  50.00000
  50.00000  22.60000  24.80000  22.90000  23.90000  26.60000  26.40000
  29.60000  26.50000  29.00000  33.10000  23.20000  28.70000  18.70000
  15.00000  16.10000  20.80000  23.80000  31.00000  35.10000  44.80000
  50.00000  37.00000  60.00000  54.40000  50.00000  32.00000  29.60000
  28.40000  44.80000  50.00000  29.80000  34.90000  37.60000  40.50000
  33.10000  29.10000  35.40000  32.40000  28.32000  28.80000  26.40000
  26.60000  26.43000  24.48000  24.00000  22.90000  22.00000  20.10000
  18.50000  17.50000  20.10000  18.50000  20.60000  23.10000  19.40000
  18.70000  18.30000  17.80000  18.20000  16.80000  19.90000  19.30000
  19.90000  21.70000  22.80000  18.80000  18.70000  18.50000  18.30000
  20.10000  23.10000  19.40000  18.70000  18.30000  17.80000  18.20000
  16.80000  19.90000  19.30000  19.90000  21.70000  22.80000  18.80000
  18.70000  18.50000  18.30000  20.10000  23.10000  19.40000  18.70000
  18.30000  17.80000  18.20000  16.80000  19.90000  19.30000  19.90000
  21.70000  22.80000  18.80000  18.70000  18.50000  18.30000  20.10000
  23.10000  19.40000  18.70000  18.30000  17.80000  18.20000  16.80000
  19.90000  19.30000  19.90000  21.70000  22.80000  18.80000  18.70000
  18.50000  18.30000  20.10000  23.10000  19.40000  18.70000  18.30000
  17.80000  18.20000  16.80000  19.90000  19.30000  19.90000  21.70000
  22.80000  18.80000  18.70000  18.50000  18.30000  20.10000  23.10000
  19.40000  18.70000  18.30000  17.80000  18.20000  16.80000  19.90000
  19.30000  19.90000  21.70000  22.80000  18.80000  18.70000  18.50000
  18.30000  20.10000  23.10000  19.40000  18.70000  18.30000  17.80000
  18.20000  16.80000  19.90000  19.30000  19.90000  21.70000  22.80000
  18.80000  18.70000  18.50000  18.30000  20.10
```

```
Text(0.5,0.1,'Adjusted R-Squared Value for Model Selection')
```