In [1]:

```
from sklearn.datasets import fetch_california_housing
```

In [2]:

```
from sklearn.model_selection import train_test_split
```

In [3]:

```
from sklearn.linear_model import LinearRegression
```

In [4]:

```
from sklearn.metrics import mean_absolute_error,mean_squared_error,r2_score
```

In [5]:

```
data = fetch_california_housing()  # data = pd.read_csv('houseprice.csv')
```

In [11]:
```
print(data)
```

```
{'data': array([[   8.3252    ,   41.        ,    6.98412698, ...,    2.5555
5556,
          37.88      , -122.23      ],
       [   8.3014    ,   21.        ,    6.23813708, ...,    2.10984183,
          37.86      , -122.22      ],
       [   7.2574    ,   52.        ,    8.28813559, ...,    2.80225989,
          37.85      , -122.24      ],
       ...,
       [   1.7       ,   17.        ,    5.20554273, ...,    2.3256351 ,
          39.43      , -121.22      ],
       [   1.8672    ,   18.        ,    5.32951289, ...,    2.12320917,
          39.43      , -121.32      ],
       [   2.3886    ,   16.        ,    5.25471698, ...,    2.61698113,
          39.37      , -121.24      ]]), 'target': array([4.526, 3.585, 3.52
1, ..., 0.923, 0.847, 0.894]), 'frame': None, 'target_names': ['MedHouseVa
l'], 'feature_names': ['MedInc', 'HouseAge', 'AveRooms', 'AveBedrms', 'Popul
ation', 'AveOccup', 'Latitude', 'Longitude'], 'DESCR': '.. _california_housi
ng_dataset:\n\nCalifornia Housing dataset\n--------------------------\n\n**D
ata Set Characteristics:**\n\n    :Number of Instances: 20640\n\n    :Number
of Attributes: 8 numeric, predictive attributes and the target\n\n    :Attri
bute Information:\n        - MedInc        median income in block group\n
- HouseAge      median house age in block group\n        - AveRooms      ave
rage number of rooms per household\n        - AveBedrms     average number o
f bedrooms per household\n        - Population    block group population\n
- AveOccup      average number of household members\n        - Latitude
block group latitude\n        - Longitude     block group longitude\n\n    :
Missing Attribute Values: None\n\nThis dataset was obtained from the StatLib
repository.\nhttps://www.dcc.fc.up.pt/~ltorgo/Regression/cal_housing.html\n
\nThe target variable is the median house value for California districts,\ne
xpressed in hundreds of thousands of dollars ($100,000).\n\nThis dataset was
derived from the 1990 U.S. census, using one row per census\nblock group. A
block group is the smallest geographical unit for which the U.S.\nCensus Bur
eau publishes sample data (a block group typically has a population\nof 600
to 3,000 people).\n\nA household is a group of people residing within a hom
e. Since the average\nnumber of rooms and bedrooms in this dataset are provi
ded per household, these\ncolumns may take surprisingly large values for blo
ck groups with few households\nand many empty houses, such as vacation resor
ts.\n\nIt can be downloaded/loaded using the\n:func:`sklearn.datasets.fetch_
california_housing` function.\n\n.. topic:: References\n\n    - Pace, R. Kel
ley and Ronald Barry, Sparse Spatial Autoregressions,\n      Statistics and
Probability Letters, 33 (1997) 291-297\n'}
```

In [12]:
```
x = data.data
```

In [14]:
```
y = data.target
```

In [15]:
```
x_train, x_test, y_train, y_test = train_test_split(x,y,test_size=0.2)
```

In [16]:

```
model = LinearRegression()
```

In [17]:

```
model.fit(x_train,y_train)
```

Out[17]:

```
LinearRegression()
```

In a Jupyter environment, please rerun this cell to show the HTML representation or trust the notebook.
On GitHub, the HTML representation is unable to render, please try loading this page with nbviewer.org.

In [18]:

```
y_pred = model.predict(x_test)
```

In [19]:

```
y_pred
```

Out[19]:

```
array([1.02221485, 1.48808491, 1.75052818, ..., 3.20591363, 1.07120023,
       2.94545239])
```

In [20]:

```
mae = mean_absolute_error(y_test,y_pred)
```

In [21]:

```
mse = mean_squared_error(y_test,y_pred)
```

In [22]:

```
r2_score = r2_score(y_test,y_pred)
```

In [23]:

```
print('Mean Absolute Error', mae)
```

```
Mean Absolute Error 0.5345160416296426
```

In [24]:

```
print('Mean Square Error',mse)
```

```
Mean Square Error 0.5206071326972168
```

In [25]:

```
print('R2 Score',r2_score)
```

```
R2 Score 0.6074818254825243
```

In [ ]: