Boston Housing

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 centers

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 Black people by town

LSTAT: % lower status of the population

MEDV: Median value of owner-occupied homes in $1000s

<<<<<Data Dictionary

	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD	TAX	PTRATIO	B	LSTAT	MEDV
0	0.00632	18.0	2.31	0.0	0.538	6.575	65.2	4.0900	1	296	15.3	396.90	4.98	24.0
1	0.02731	0.0	7.07	0.0	0.469	6.421	78.9	4.9671	2	242	17.8	396.90	9.14	21.6
2	0.02729	0.0	7.07	0.0	0.469	7.185	61.1	4.9671	2	242	17.8	392.83	4.03	34.7
3	0.03237	0.0	2.18	0.0	0.458	6.998	45.8	6.0622	3	222	18.7	394.63	2.94	33.4
4	0.06905	0.0	2.18	0.0	0.458	7.147	54.2	6.0622	3	222	18.7	396.90	NaN	36.2
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
501	0.06263	0.0	11.93	0.0	0.573	6.593	69.1	2.4786	1	273	21.0	391.99	NaN	22.4
502	0.04527	0.0	11.93	0.0	0.573	6.120	76.7	2.2875	1	273	21.0	396.90	9.08	20.6
503	0.06076	0.0	11.93	0.0	0.573	6.976	91.0	2.1675	1	273	21.0	396.90	5.64	23.9
504	0.10959	0.0	11.93	0.0	0.573	6.794	89.3	2.3889	1	273	21.0	393.45	6.48	22.0
505	0.04741	0.0	11.93	0.0	0.573	6.030	NaN	2.5050	1	273	21.0	396.90	7.88	11.9

506 rows × 14 columns

```
NOX            0
RM             0
DIS            0
RAD            0
TAX            0
PTRATIO        0
B              0
MEDV           0
dtype: int64
```

```
Outliers (Z-score method):
 Empty DataFrame
Columns: [NOX, RM, DIS, RAD, TAX, PTRATIO, B, MEDV]
Index: []

Outliers (IQR method):
 Empty DataFrame
Columns: [NOX, RM, DIS, RAD, TAX, PTRATIO, B, MEDV]
Index: []
```

- Checking missing values and outliers detection.

```
Predicted MEDV results:
 [25.77262372 29.02603644 15.13877775 22.082231   18.42432142]

Actual MEDV:
     MEDV
0   24.0
1   21.6
2   34.7
3   33.4
4   36.2
```
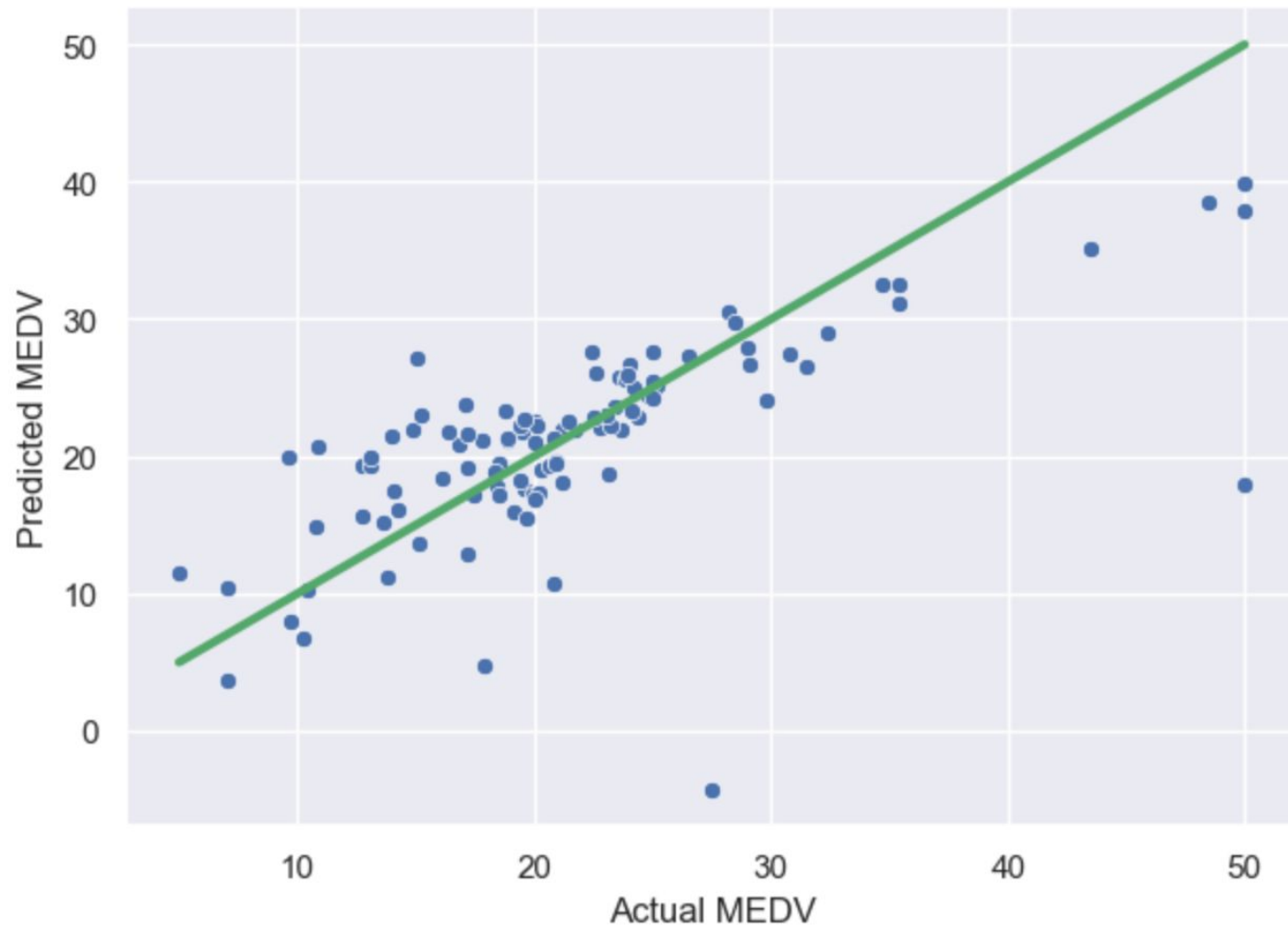
Just selected 5 array.

Actual vs Predicted MEDV

Mean Squared Error (MSE): 39.41540959794716
Mean Absolute Error (MAE): 3.8877568146294346