



Standardisation

Standardisation

- Centres the variable at zero and sets the variance to 1.

$$\text{Z-score} = \frac{X - \text{mean}(X)}{\text{Std}(X)}$$

Standardisation: example

Price
100
90
50
40
20
100
50
60
120
40
200

Mean = 79
Standard dev = 51

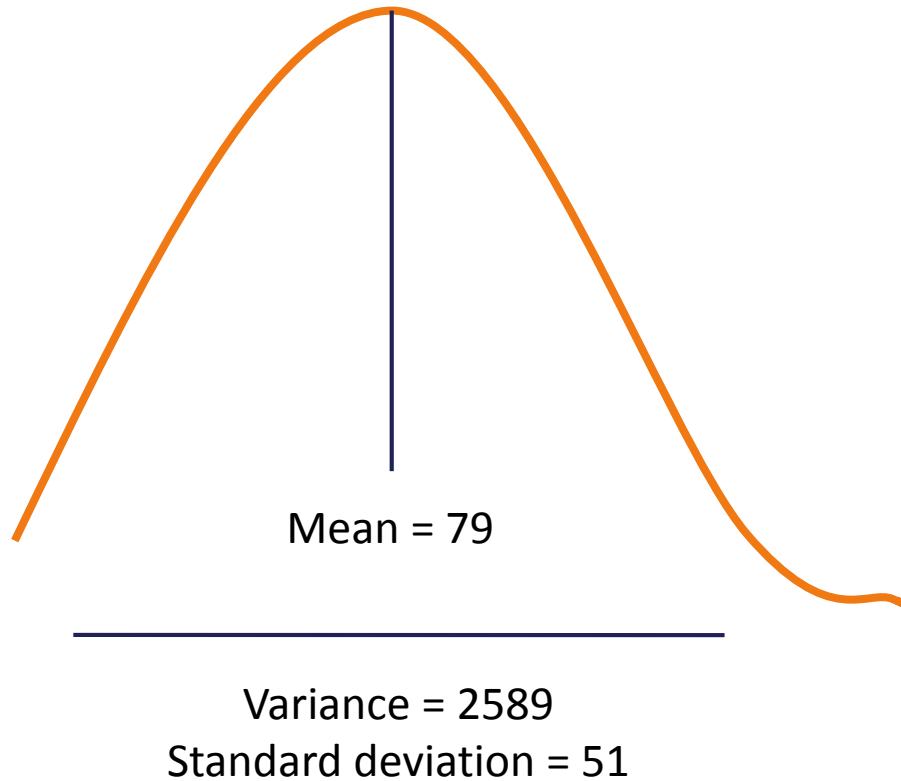


Obs. - Mean

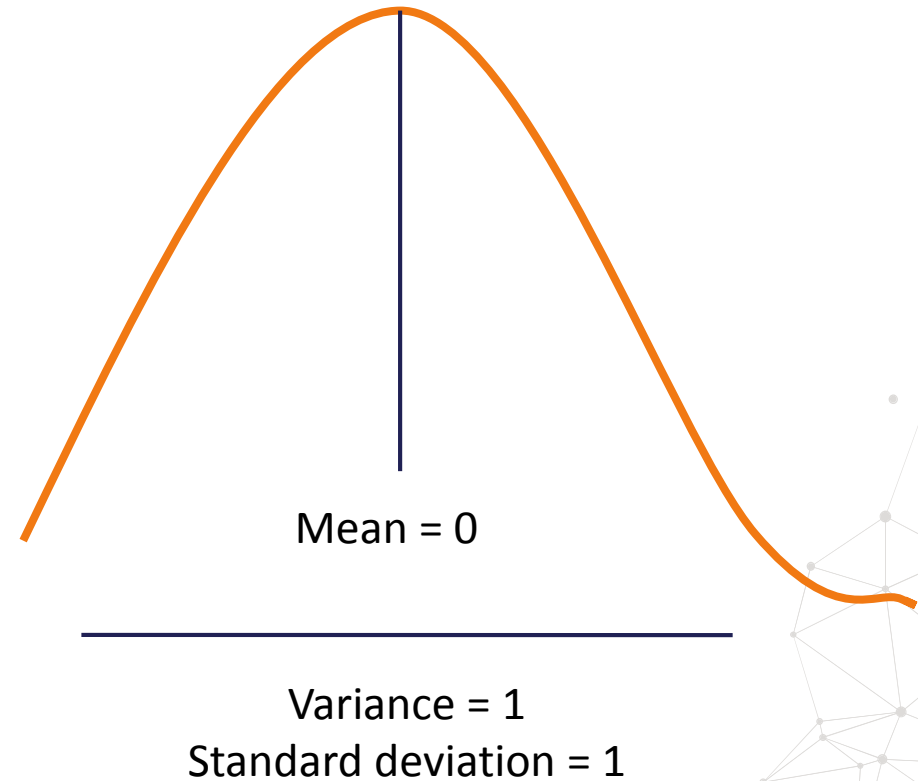
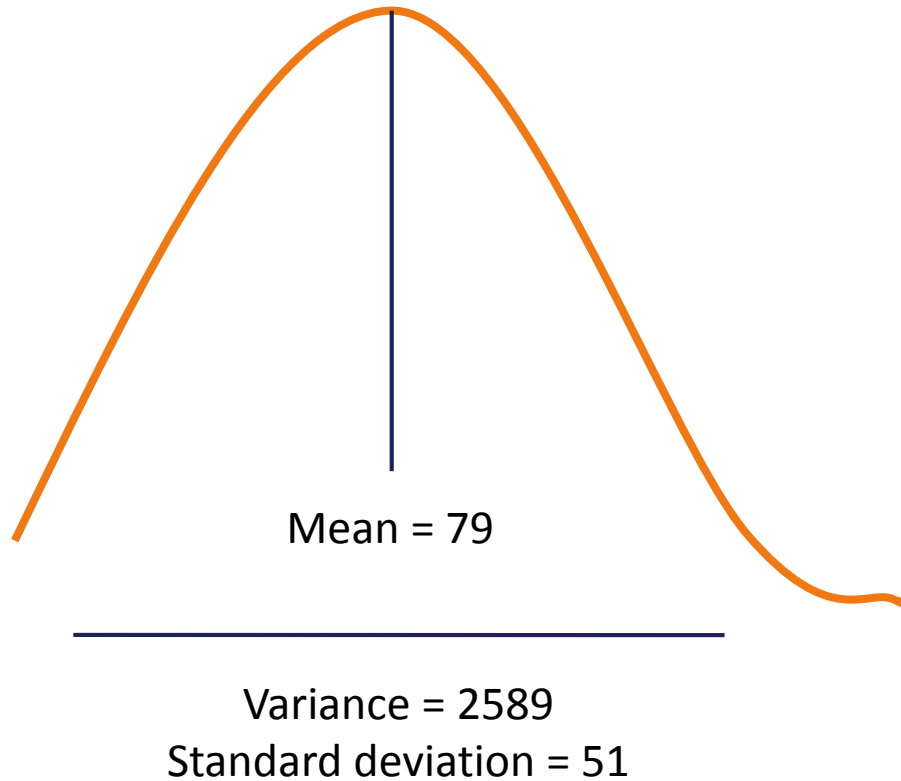
Standard dev

Price
0.41
0.22
-0.57
-0.76
-1.16
0.41
-0.57
-0.37
0.80
-0.76
2.37

Standardisation: effect

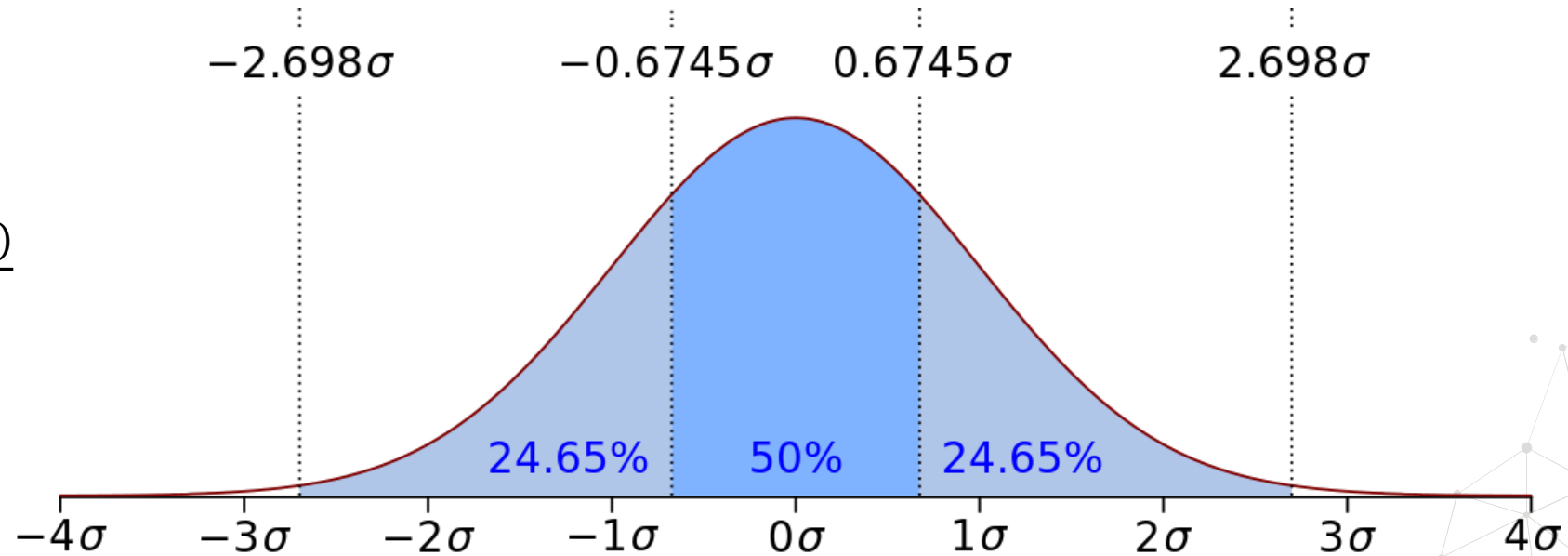


Standardisation: effect



Standardised variable: meaning

$$\text{Z-score} = \frac{X - \text{mean}(X)}{\text{Std}(X)}$$

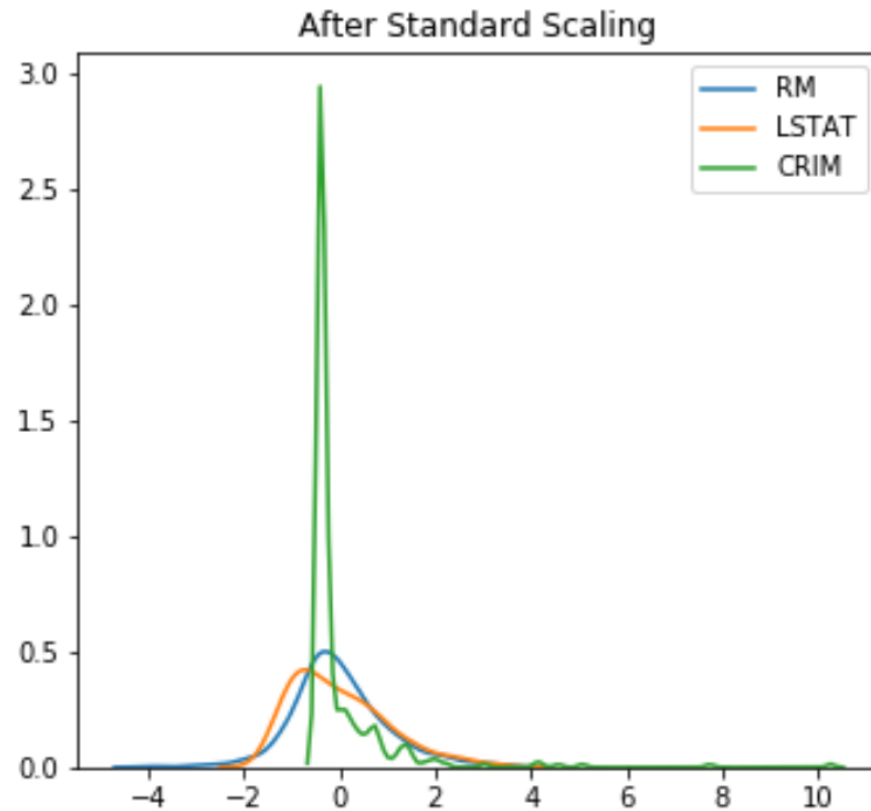
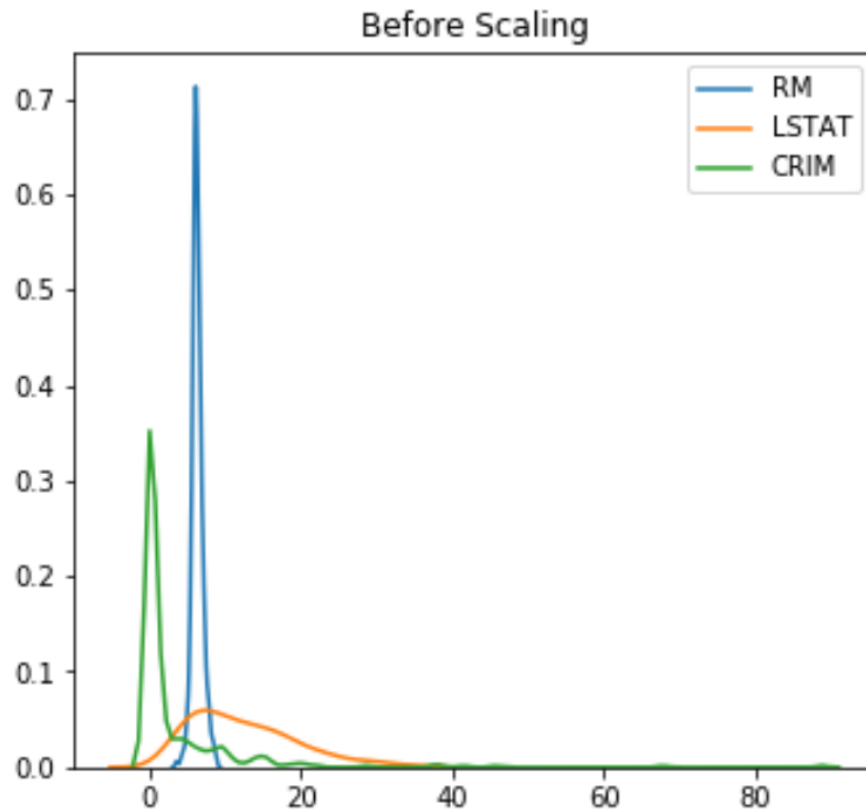




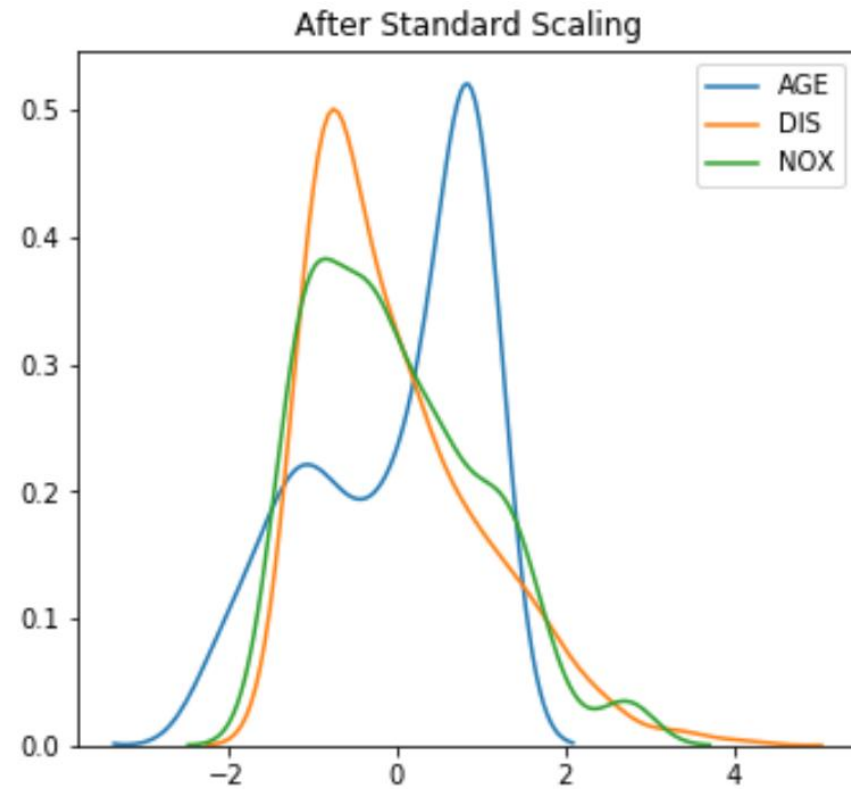
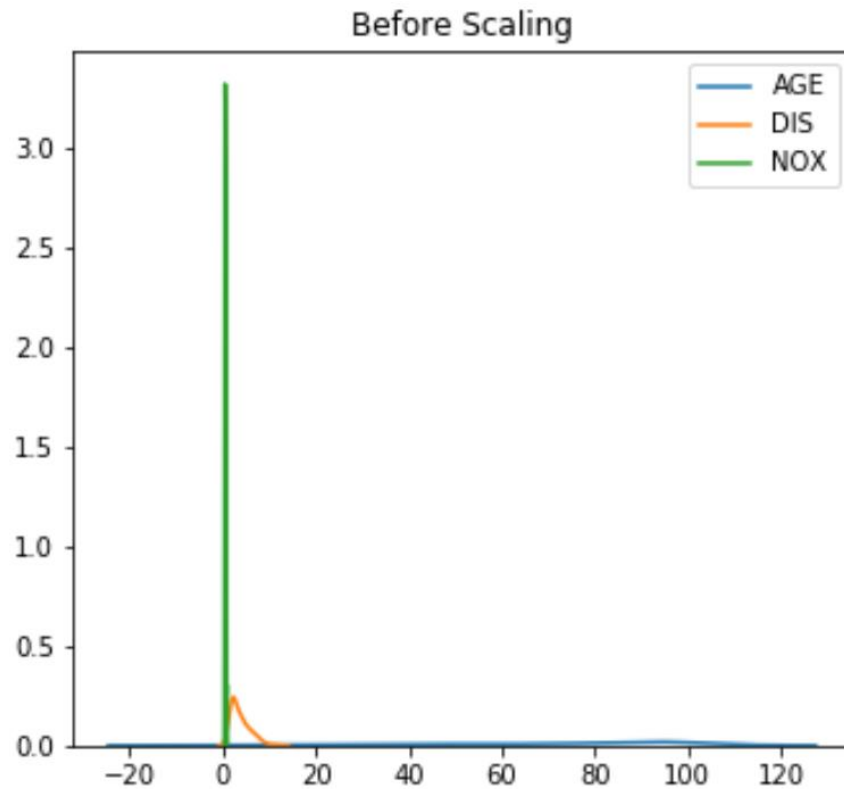
Standardisation: summary

- Centres the mean at 0
- Scales the variance at 1
- Preserves the shape of the original distribution
- Minimum and maximum values vary
- Preserves outliers

Standardisation: Notebook examples



Standardisation: Notebook examples



Accompanying Jupyter Notebook



- Read the accompanying Jupyter Notebook
- Standardisation with Scikit-learn

THANK YOU

www.trainindata.com