Train In Data

Mean Normalisation

Mean Normalisation

- Centres the variable at 0 and re-scales the variable to the value range.

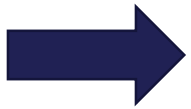
$$X\_scaled = \frac{X - mean(X)}{max(X) - \min(X)}$$

Mean normalisation: example

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

Mean = 79

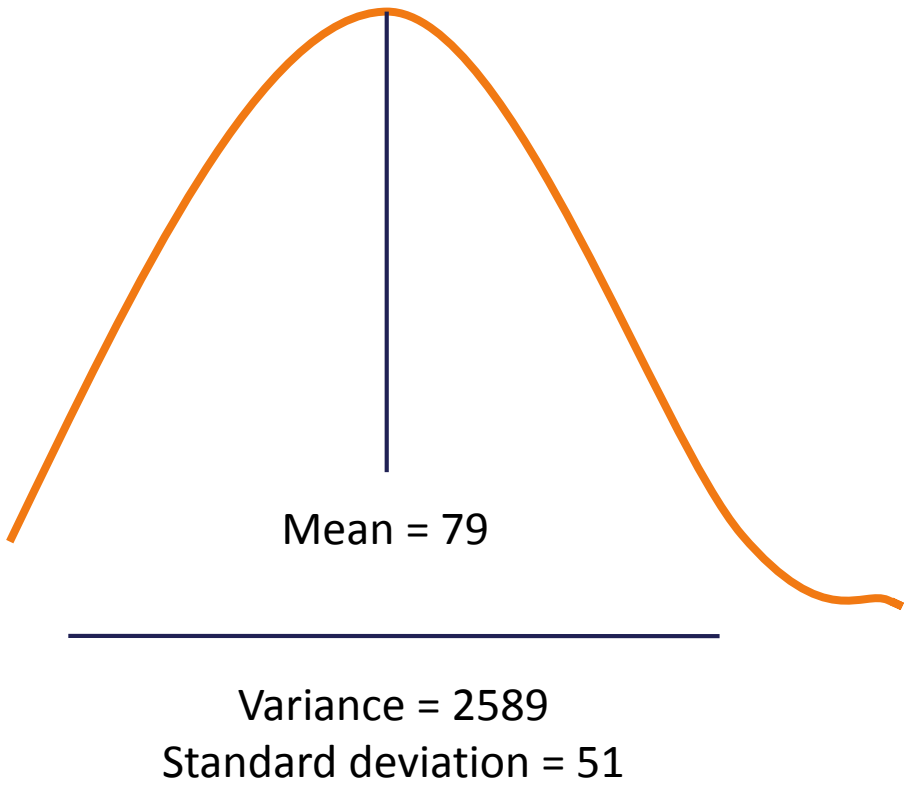
Max = 200
Min = 20
Range = 200 – 20 = 180



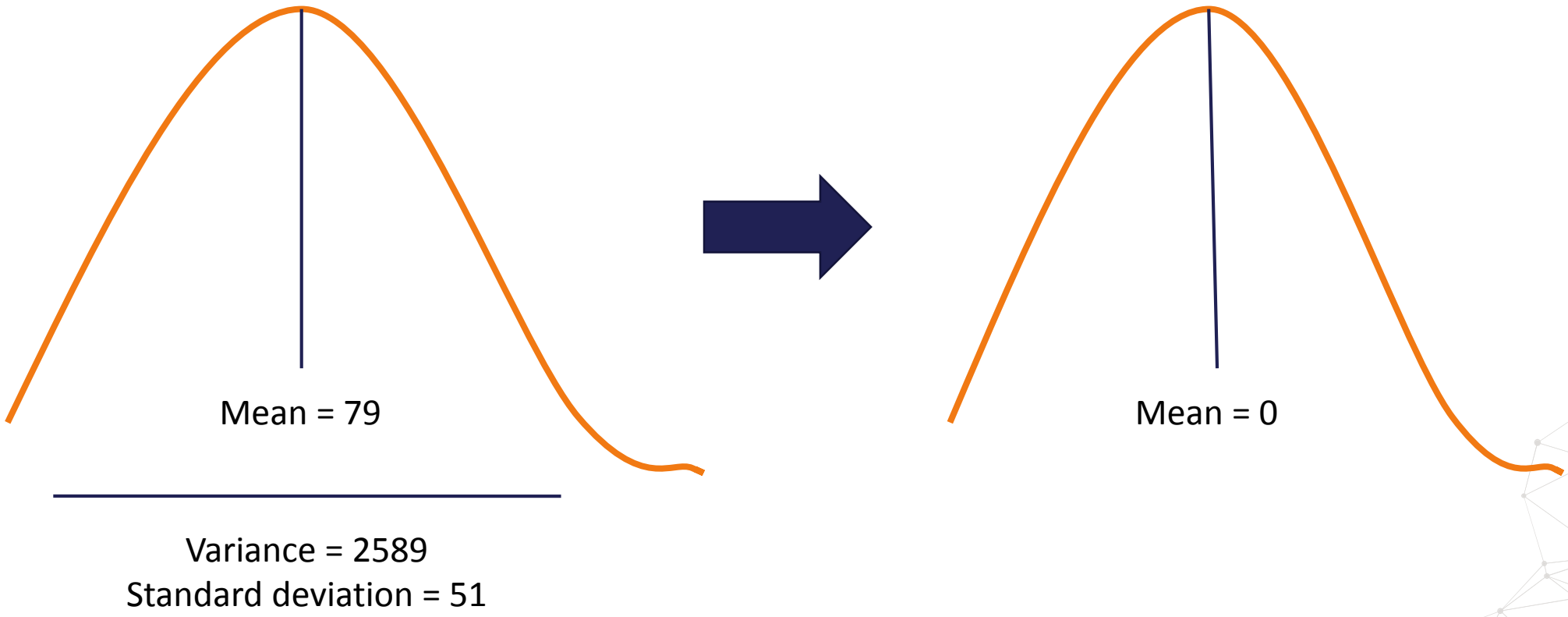
$$\frac{\text{Obs. - Mean}}{\text{Range}}$$

Price
0.12
0.06
-0.16
-0.22
-0.33
0.12
-0.16
-0.11
0.23
-0.22
0.67

Mean normalisation: effect



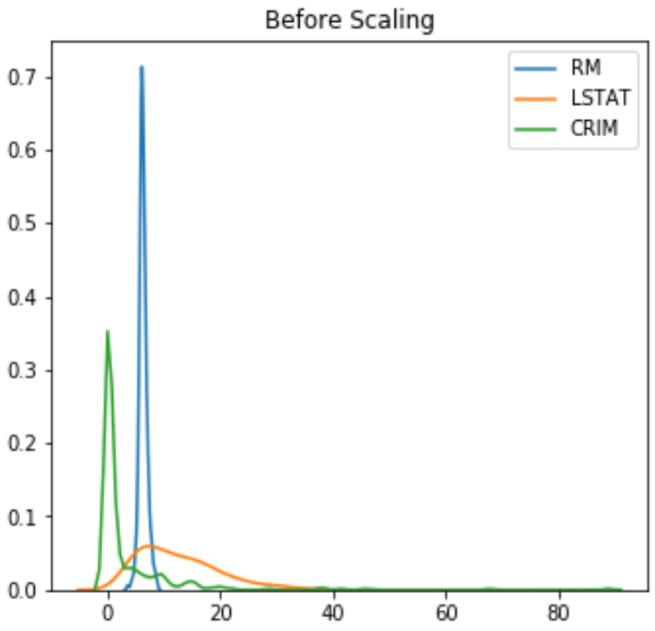
Mean normalisation: effect

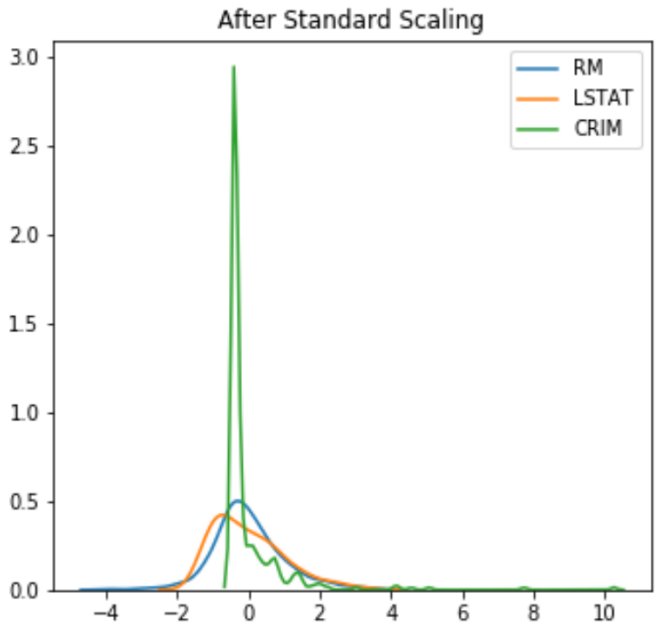


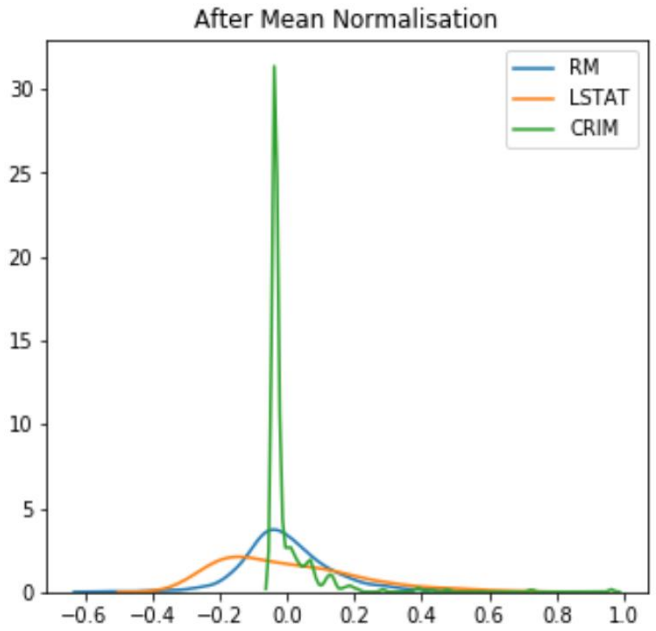
Mean normalisation: summary

- Centres the mean at 0
- Variance varies
- May alter shape of the original distribution
- Minimum and maximum values within [-1;1]
- Preserves outliers

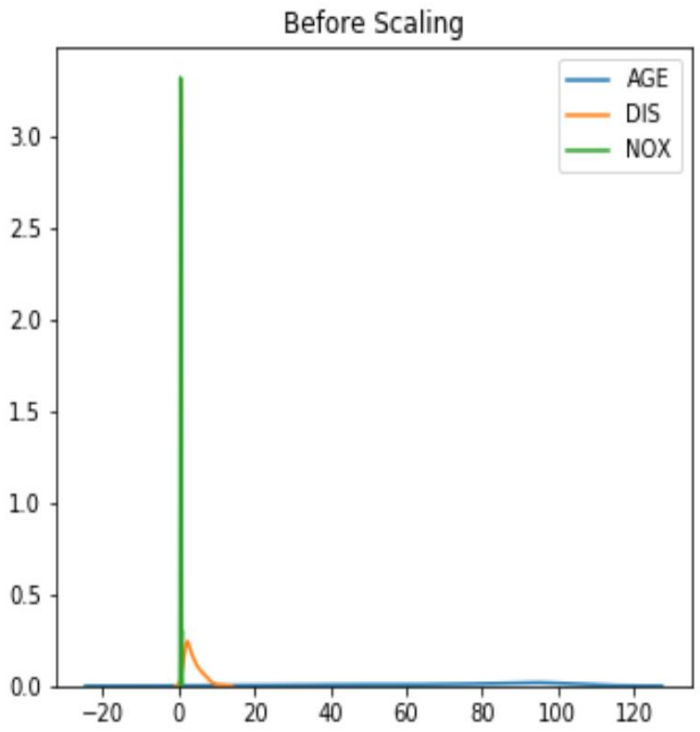
Mean normalisation: Notebook

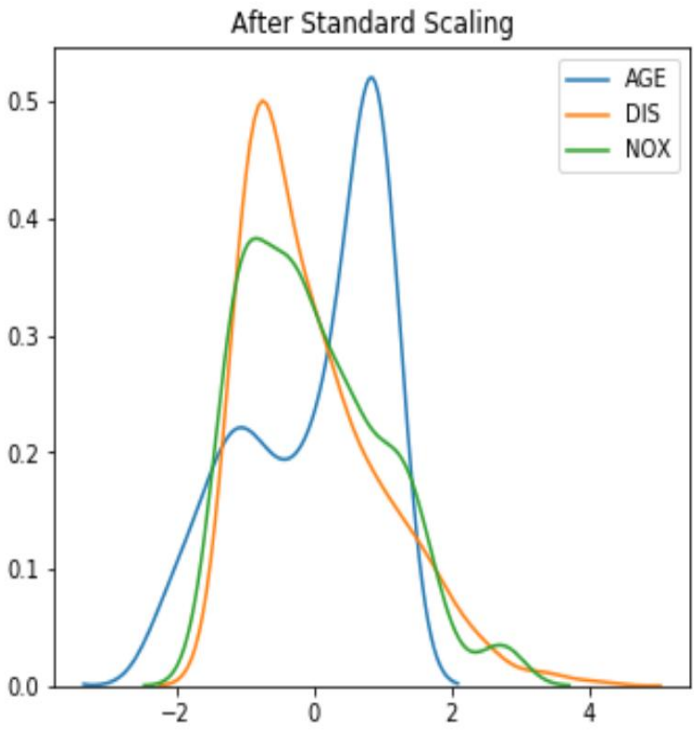


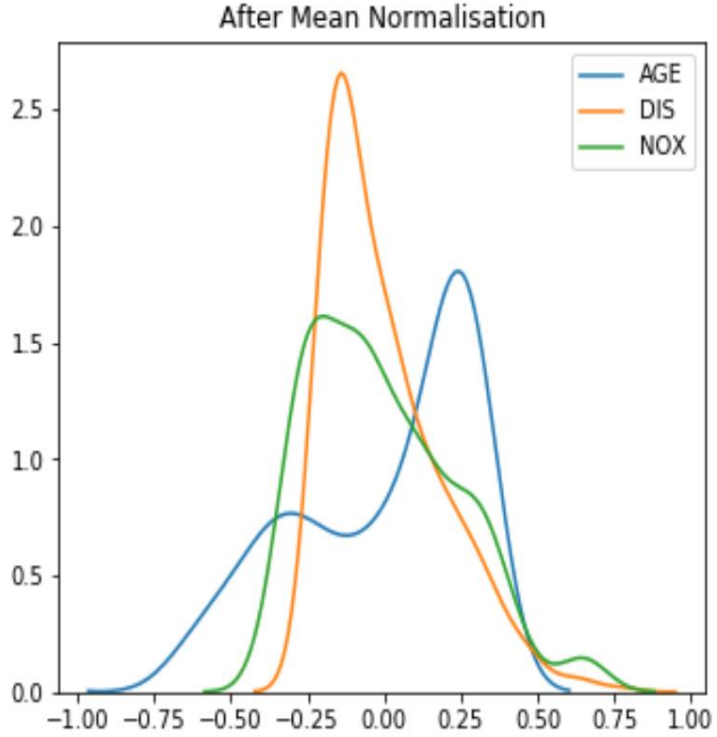




Mean normalisation: Notebook







Accompanying Jupyter Notebook

- Read the accompanying Jupyter Notebook
- Mean normalisation with pandas and a work-around with Scikit-learn

Train In Data

THANK YOU

www.trainindata.com