



Equal-width discretisation

Equal-width discretisation: definition

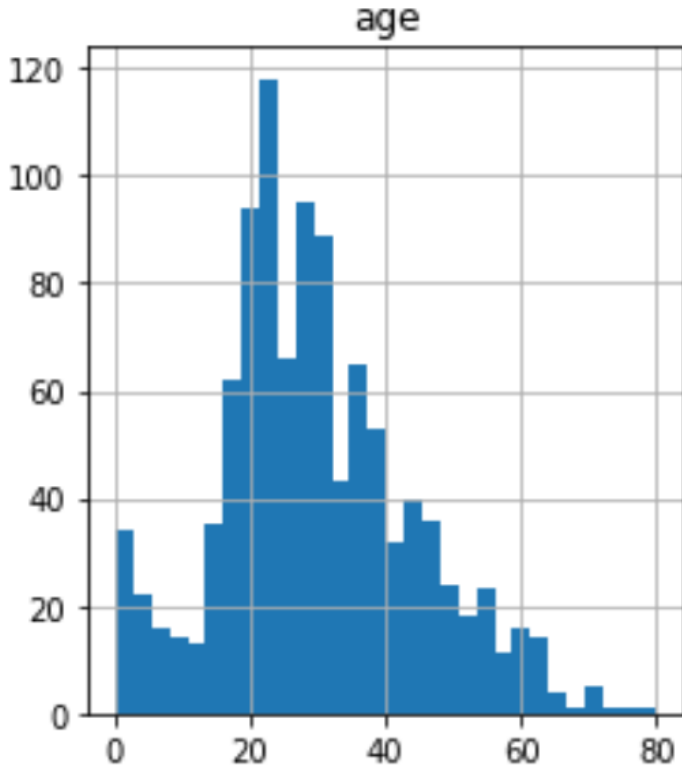
Equal width discretisation divides the scope of possible values into N bins of the same width.

The interval width is determined by:

$$\text{width} = (\text{max value} - \text{min value}) / N$$

where N is the number of bins or intervals.

Equal-width discretisation: calculation



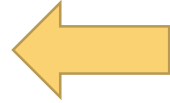
Min age = 0
Max age = 73
Intervals = 10

$$\text{width} = (73 - 0) / 10 = 7.3$$

Intervals:
0-7; 7-14; 14-21 ... 70-77

Equal-width discretisation: calculation

age	Age_disc
13.0	(7.0, 14.0]
4.0	(-0.001, 7.0]
30.0	(28.0, 35.0]
21.0	(14.0, 21.0]
22.0	(21.0, 28.0]
16.0	(14.0, 21.0]
30.0	(28.0, 35.0]
2.0	(-0.001, 7.0]
49.0	(42.0, 49.0]
35.0	(28.0, 35.0]



Min age = 0

Max age = 73

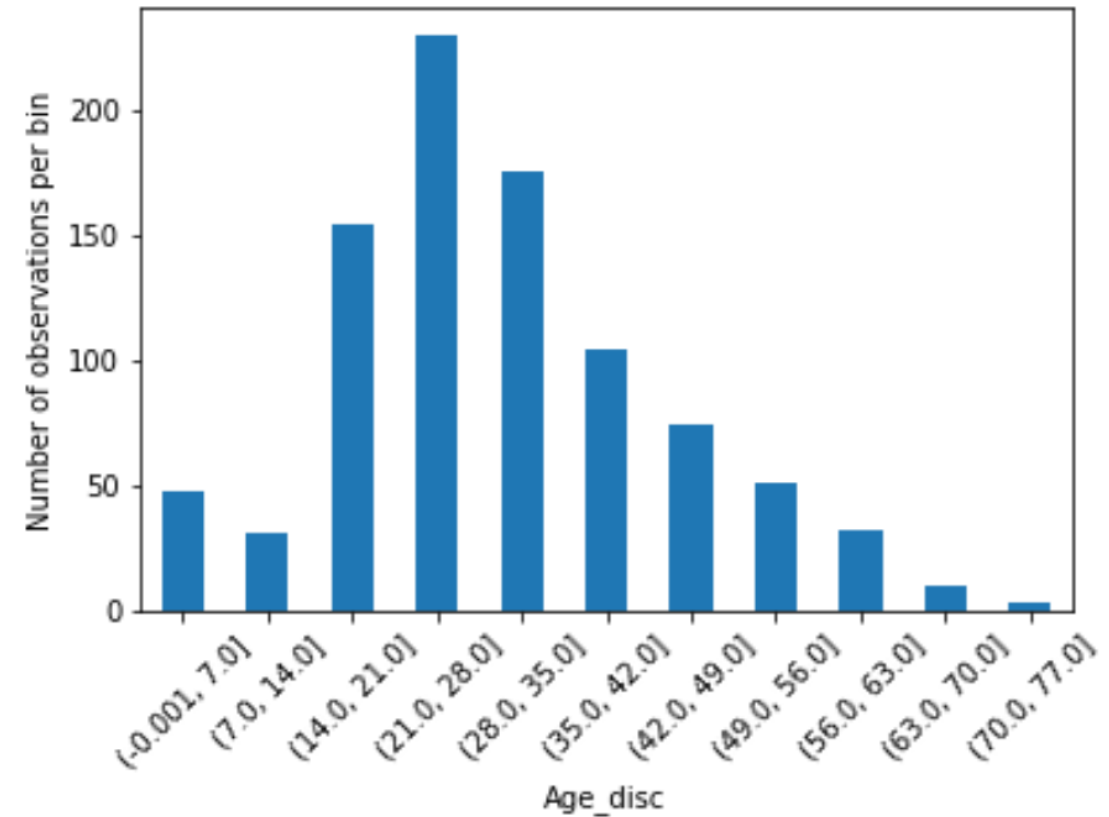
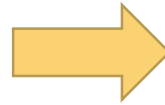
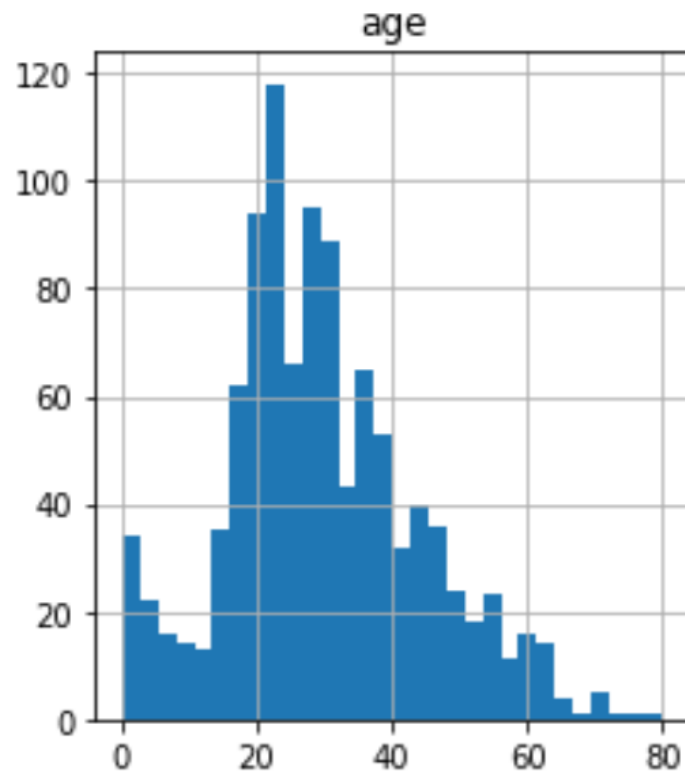
Intervals = 10

$\text{width} = (73 - 0) / 10 = 7.3$

Intervals:

0-7; 7-14; 14-21 ... 70-77

Equal-width discretisation: calculation





Equal-width discretisation: summary

- Does not improve value spread
- Handles outliers
- Creates discrete variable
- Good to combine with categorical encodings



THANK YOU

www.trainindata.com