Scaling

Linear Scaling

With any data set you can simplify by adding or subtracting any number, then calculating the mean – or median – and then adding back that number.

$\sum$ (12 + 13 + 14 + 15 + 16) = 70

$\sum$ (2 + 3 + 4 + 5 + 6) = 20

20 + 50 = 70

because we subtracted 5 lots f ten.

So

$\sum$ (12 + 13 + 14 + 15 + 16) / 5 = 14

$\sum$ (2 + 3 + 4 + 5 + 6) / 5 = 4

and 4 + 10 = 14

Proof

$\{ \sum (x_i - a) \} / n + a =$

$\{ \sum x_i \} / n - \{ \sum a \} / n + a$

Now $\sum a = na$

So $\{ \sum x_i \} / n - an / n + a$

$= \{ \sum x_i \} / n - a + a$

$= \{ \sum x_i \} / n$

Proportionate Scaling

120 + 130 + 140 + 150 + 160 = 700

divide through by ten.

12 + 13 + 14 + 15 + 16 = 70

Multiply sum by ten = 700

In short we're saying $\sum a x_i = a \sum x_i$

So we can reduce the data set by any proportion, calculate the mean and then multiply the answer by that proportion.

That is $\sum x_i = p \sum (x_i / p)$

Proof

Let a = 1 / p

So from $\sum a x_i = a \sum x_i$

we get $\sum (x_i / p) = 1 / p \sum x_i$

Switch around

$1 / p \sum x_i = \sum (x_i / p)$

Multiply through by p

$p / p \sum x_i = p \sum (x_i / p)$

$\sum x_i = p \sum (x_i / p)$

However it is not recommended to use this particularly when dealing with histograms. It just gets confusing when you start changing the vertical frequency density axis.

rg

Scaling

Linear Scaling

With any data set you can simplify by adding or subtracting any number, then calculating the mean – or median – and then adding back that number.

$\sum$ (12 + 13 + 14 + 15 + 16) = 70

$\sum$ (2 + 3 + 4 + 5 + 6) = 20

20 + 50 = 70

because we subtracted 5 lots f ten.

So

$\sum$ (12 + 13 + 14 + 15 + 16) / 5 = 14

$\sum$ (2 + 3 + 4 + 5 + 6) / 5 = 4

and 4 + 10 = 14

Proof

$\{ \sum (x_i - a) \} / n + a =$

$\{ \sum x_i \} / n - \{ \sum a \} / n + a$

Now $\sum a = na$

So $\{ \sum x_i \} / n - an / n + a$

$= \{ \sum x_i \} / n - a + a$

$= \{ \sum x_i \} / n$

Proportionate Scaling

120 + 130 + 140 + 150 + 160 = 700

divide through by ten.

12 + 13 + 14 + 15 + 16 = 70

Multiply sum by ten = 700

In short we're saying $\sum a x_i = a \sum x_i$

So we can reduce the data set by any proportion, calculate the mean and then multiply the answer by that proportion.

That is $\sum x_i = p \sum (x_i / p)$

Proof

Let a = 1 / p

So from $\sum a x_i = a \sum x_i$

we get $\sum (x_i / p) = 1 / p \sum x_i$

Switch around

$1 / p \sum x_i = \sum (x_i / p)$

Multiply through by p

$p / p \sum x_i = p \sum (x_i / p)$

$\sum x_i = p \sum (x_i / p)$

However it is not recommended to use this particularly when dealing with histograms. It just gets confusing when you start changing the vertical frequency density axis.

rg