

a cloud of observations, and the line or curve drawn through the middle of it

What a scatterplot and its model are

- each dot is one observation with two measurements, one read off each axis. The plot shows individuals, not totals
- a MODEL is the line or curve drawn to summarise the cloud, so almost no dot sits exactly on it. The **line of best fit** is the straight one that comes closest to the whole cloud at once, with roughly half the dots above it and half below
- the direction of the cloud is its association: rising left to right is positive, falling is negative, and no straight tilt at all means no linear association to describe

Reading a straight model *

the model, always in this form

$$y = mx + b$$

m, the change in y for one more unit of x

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

b, the value the model gives at $x=0$

$$b = y - mx$$

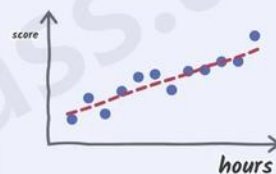
how far one observation sits off the model

residual = actual - predicted

every interpretation question is one of two: what does m mean in the story, or what does b mean in the story. m is a rate per unit, b is a starting amount

Which way does the cloud lean

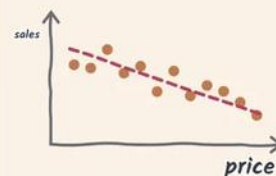
Positive association



the cloud rises left to right, with the dashed line of best fit running through it

- as x increases, y tends to increase as well
- the slope m of the line of best fit is POSITIVE
- the dots still scatter about the line, they never sit on it

Negative association



the cloud falls left to right, and the dashed line of best fit falls with it

- as x increases, y tends to decrease
- the slope m of the line of best fit is NEGATIVE
- a falling line is still a good model, it is not a worse one

a strong association means the dots sit CLOSE to the line, and that is a separate question from which way the line tilts

Worked: what m and b mean

Q. a club's membership x weeks after it opened is modelled by $y = 2.4x + 18$. What do 2.4 and 18 each stand for, and what does the model predict at week 25?

- 1 b first: put $x=0$ into the model, which gives $y=18$, the membership when the club opened
- 2 m next: every extra week adds 2.4 to y, so the club gains about 2.4 members a week
- 3 predict at $x=25$: $2.4 \times 25 = 60$, then $60 + 18 = 78$

Answer

18 is the starting membership, 2.4 is the weekly gain, and the model predicts 78 members at week 25

2.4 members a week is not a whole number of people, and that is fine: a model describes a trend