

## Scatter diagrams

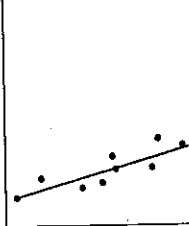
A scatter diagram is used to see whether there appears to be a relationship between two features, such as the handspan and height of the people in a particular group.

Each pair of values that relates to one member is plotted as a point.

Sometimes this gives a pattern where the points appear to lie clustered around a line – the 'line of best fit'. In such cases there is said to be correlation between the features.

These five examples are drawn to the same scale.

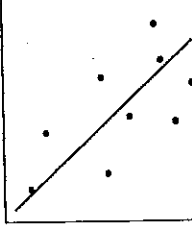
**Strong positive correlation**



Points appear to be clustered around a clear line sloping up to the right.

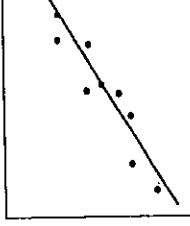
A positive correlation between the two sets of data means, roughly, 'the bigger the one, the bigger the other'.

**Some positive correlation**



Points roughly suggest a line sloping up to the right.

**Strong negative correlation**



Points strongly suggest a line sloping down to the right.

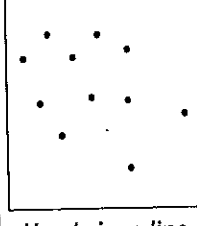
A negative correlation between the two sets of data means, roughly, 'the bigger the one, the smaller the other'.

**Some negative correlation**



Points appear to lie roughly on a line which slopes down to the right.

**No correlation**



No obvious line emerges from looking at the points.

Even where there is correlation, you cannot say that one set of values is determined by the other set.

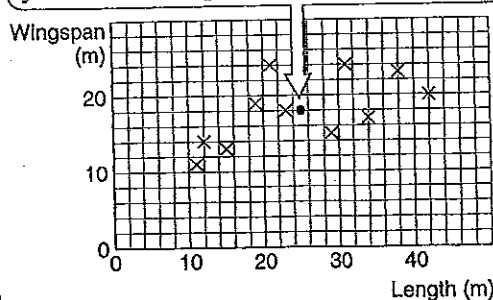
You may be asked to draw a line of best fit using the means of the values shown on the diagram.

Mean ► page 74

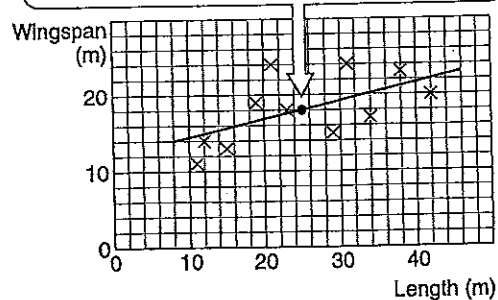
Suppose your diagram is about the dimensions of some aeroplanes. The mean length is 25 m and the mean wingspan is 18 m.

Plot a point at (25, 18).

Check that it does come in the middle of your cluster of points.



Now draw a line of best fit through it.



When you draw a line of best fit it helps to use a transparent ruler.

- 1 A group of pupils were tested on their multiplication tables.  
They then did some intensive tables practice.  
After the practice they were tested again.  
This table shows their test scores before and after the test.

Score before practice	6	3	2	7	4	8	4	7
Score after practice	8	7	6	7	8	9	6	10

On graph paper, draw a scatter diagram for the data and draw a line of best fit.  
Comment on how effective the practice was in improving their scores.

- 2 This table shows distances between pairs of cities, chosen at random from a map.  
One distance is 'as the crow flies' measured on the map.  
The other is the distance by a 'recommended' road route given in a mileage chart.

	Distance as the crow flies (miles)	Road distance (miles)
Nottingham to Lincoln	31	36
Birmingham to Leicester	34	42
Lincoln to Northampton	69	91
Nottingham to Oxford	82	103
Cambridge to Norwich	57	63
Oxford to Leicester	61	78
Colchester to Peterborough	67	89
Norwich to Lincoln	87	103
Coventry to Birmingham	17	19
Leicester to Coventry	22	26

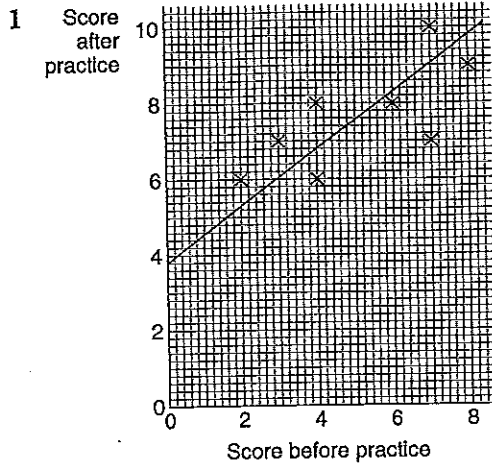
- (a) On graph paper, draw a scatter diagram for this information.  
(b) Draw a line of best fit.  
(c) Use what you have done to get a 'rule of thumb' for estimating the road distance if you know the distance between two cities 'as the crow flies'.
- 3 The table shows the scores of 12 pupils in maths and French tests.

Maths score	5	10	14	24	28	35	43	47	52	57	62	70
French score	12	6	12	24	18	24	33	23	29	35	28	35

- (a) (i) Calculate the mean of the maths scores.  
(ii) Calculate the mean of the French scores.  
(b) (i) Plot the scores of these twelve pupils on a scatter diagram.  
(ii) On the scatter diagram, plot the point whose coordinates are the mean scores you have calculated in part (a).  
(iii) Draw the line of best fit on the scatter diagram.  
(c) Vina scored 42 in the maths test but was absent for the French test.  
Use your line of best fit to state a score she might have had for her French test if she had been present.  
(d) Comment on this method.

## Scatter diagrams (page 72)

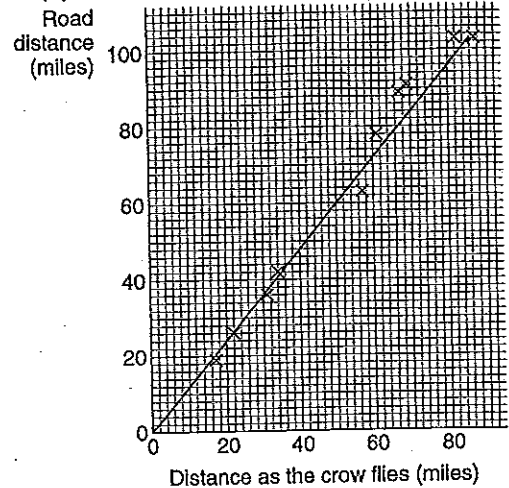
Normally in an exam you will be given the grid on which to plot the scatter diagrams. The normal convention is for the axis for the first set of data to be along the bottom of the grid and the axis for the second set to be up the side of the grid. If your diagrams look different from those shown here, first check that you have kept to this convention.



You could work out the means to help you draw your line of best fit but you do not have to unless the question tells you to.

The practice seems to have improved pupils' scores, though those who were poorer with their tables to start with tended to improve more than those who were better to start with. One pupil's score stayed the same. In this case the moderate correlation shown by the scatter diagram tells us that the better pupils generally stayed better – in other words, not much 'overtaking' occurred.

2 (a) and (b)

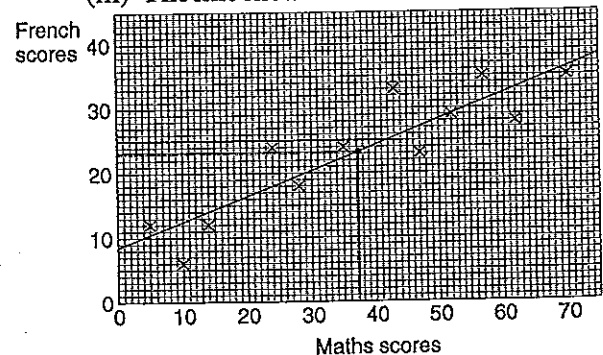


(c) Work out the gradient of the line to find a rule of thumb.

A rough rule is "Take the "as the crow flies distance" and add on 20% (one-fifth) or 'Multiply the "as the crow flies distance" by 1.2'. Your rule might be slightly different from this.

3 (a) (i) The mean of the maths scores is 37.25.  
(ii) The mean of the French scores is 23.25.

(b) (i) The crosses show the scores.  
(ii) The dot shows the mean scores.  
(iii) The line shown is the line of best fit.



(c) The line of best fit suggests a French score of 25.

(d) Other pupils' French scores were as much as 8 marks above or below the line. So we cannot be sure that she would have got a score of 25.

### More help or practice

Scatter diagrams – Book Y4 pages 120 to 121

Correlation – Kids like us pages 10 to 11