

If you have two variables and suspect that a relationship exists between them, you can draw a scatter diagram (also known as a scattergraph) which will demonstrate if it is worth pursuing the matter further.

Detailed study of such diagrams leads to discussions about correlation and lines of regression, which are not part of this course as they are too specialised. However, the scatter diagram is interesting enough in its own right.

The variables are plotted along a pair of axes in the same way as you would for a line graph, the only difference being that you don't (in fact can't) join them up with a line. As usual, the variable which, it is suspected, is dependent on the other is plotted vertically and the independent variable is plotted horizontally.

It is important to note, however, that the appearance of a relationship does not imply that one depends on the other. It simply shows a connection. For all you know, both of the variables could be dependent on a third, unknown, one.

Example 11a

This table shows the diameters (in inches) of a variety of cylindrical components produced on a lathe and the time taken (in seconds) by the machinist to make them.

Draw a scatter diagram to illustrate this.

Solution:

If anything, the time taken to make the component will depend on the diameter of the component (although they might both depend on something else).

So we plot time vertically and diameter horizontally.

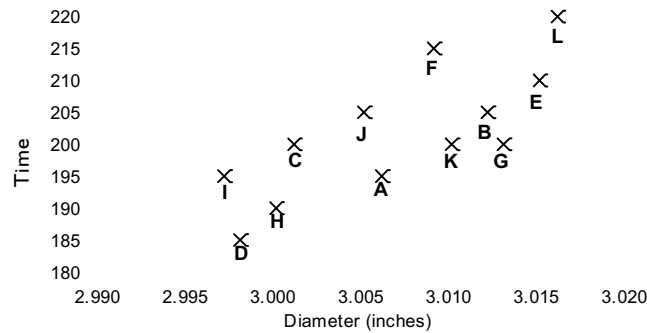
Component	Diameter	Time
A	3.006	195
B	3.012	205
C	3.001	200
D	2.998	185
E	3.015	210
F	3.009	215
G	3.013	200
H	3.000	190
I	2.997	195
J	3.005	205
K	3.010	200
L	3.016	220

First we need to decide on a scale. The lowest diameter is 2.997 and the highest is 3.016, so a horizontal scale from 2.990 to 3.020 will suffice.

The lowest time is 185 and the highest is 220 so a vertical scale from 180 to 220 will do.

The resulting scatter diagram is shown on the next page.

Scatter Diagram showing Diameters and Production Times

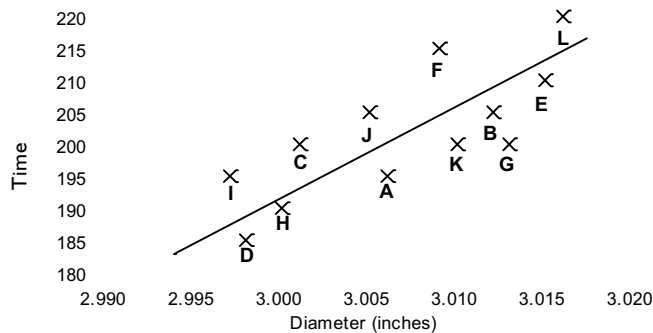


Each point is plotted in its correct position relative to the axes. I have also typed the reference letter beside each one, though this is not usually done. As you see, joining the points up with a line would make no sense.

The scattergraph shows a reasonably strong relationship between the two variables. In general, the larger the diameter of the component, the longer it takes to make.

We say that the correlation is reasonably strong and that it is a linear correlation. That is because we could draw a straight line through all the plotted points and use it to give approximate values of time for diameters which were not actually included in the diagram:

Scatter Diagram showing Diameters and Production Times



A component of diameter 3.002 inches would take roughly 195 seconds to produce.

The line has been drawn by eye; there are mathematical ways to ensure that it is the 'best fitting' line, but that's another story.

The correlation is **positive**, because the line slopes *up* from left to right.

Example 11b

This table shows the pupil:staff ratio at fourteen secondary schools and the percentage of pupils who passed five or more certificate exams. Draw a scatter diagram to illustrate this.

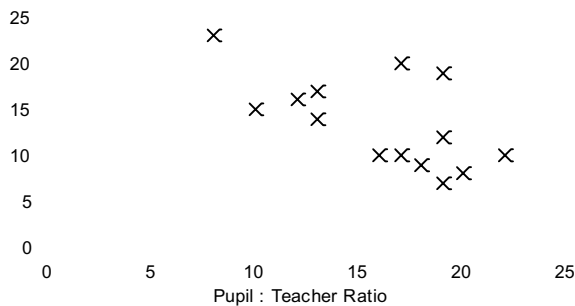
School	Pupils per teacher	Percentage
A	17	10
B	19	12
C	12	16
D	10	15
E	16	10
F	13	14
G	18	9
H	20	8
I	22	10
J	13	17
K	19	7
L	8	23
M	19	19
N	17	20

Solution:

The resulting scatter diagram is shown below.

I haven't identified the schools by letter on the diagram, but it's easy to see which is which.

Scatter Diagram of Teaching Ratio and Pass Rate



Here we see a relatively strong **negative** (because the general slope of the 'line' is *down left to right*) correlation between the two variables. In general, the more pupils there are per teacher, the fewer of the pupils do well in their exams. The strength of the relationship is clearly not as high as it was in Example 11a.



Draw a scatter diagram to illustrate each of the following tables.

In Table 1 an advertising executive is investigating the relationship between the size of the budget for certain advertising campaigns and the resulting sales.

Table 1

Advertising Expenditure (£m)	Sales (£m)
3.4	12.0
4.0	15.8
5.2	19.7
6.5	19.2
7.9	26.0
8.3	24.7
9.2	29.2
10.0	29.7

In Table 2 a town planner is looking at the population density of certain areas of the town and their distance from the town centre.

Table 2

Persons per Hectare	Distance from Centre (km)
65	0.8
21	3.7
78	1.6
48	2.7
56	1.5
53	2.6
31	3.4
34	1.9

In Table 3 an economist from an energy company is looking at the demand from a power station and how it is affected by the average daily temperature.

Table 3

Temperature (°C)	Demand (MWh)
-2.0	60.5
1.0	58.9
2.3	56.4
4.1	51.3
4.9	48.0
8.3	42.6
6.5	43.2
3.8	53.1