# M1.7 – Use a scatter diagram to identify a correlation between two variables

### Tutorials

Learners may be tested on their ability to:

* Interpret a scatter diagram, e.g. the effect of lifestyle factors on health.

### Scatter diagrams

Correlation is a very useful statistical technique that looks at relationships between two variables. For example, taller people tend to have larger shoe sizes than shorter people, therefore we can say that height is correlated with shoe size.

This relationship isn’t perfect, two people of the same height can have different shoe sizes, and a shorter person could have a larger shoe size than someone who is taller than them. The **correlation coefficient** tells us how close that relationship is to being perfect (see M1.9 for details on how to calculate and interpret this), but we can plot a range of people’s height measurements and shoe sizes on graphs called ‘scatter diagrams’ or ‘scatterplots’ to easily visualise this relationship.

|  |  |  |
| --- | --- | --- |
|  | **UK Shoe size** | **Height (inches)** |
| 1 | 4 | 57 |
| 2 | 4 | 62 |
| 3 | 6 | 60 |
| 4 | 7 | 68 |
| 5 | 8 | 71 |
| 6 | 9 | 73 |
| 7 | 11 | 72 |
| 8 | 11 | 74 |
| 9 | 12 | 70 |
| 10 | 12 | 79 |
| 11 | 5 | 58 |
| 12 | 5 | 54 |
| 13 | 8 | 68 |
| 14 | 8 | 69 |
| 15 | 9 | 71 |
| 16 | 9 | 68 |
| 17 | 7 | 65 |
| 18 | 7 | 64 |

In this example we can see that our graph shows a strong positive linear correlation. A strong correlation is one where we can draw a trend line through the data points, and they are all very close to the line. Although it is a strong correlation, you can see that few of the data points sit perfectly on the trend line, instead there are just as many above and below it. A data point that sits below the trendline describes an individual who is shorter than their shoe size would predict, while a data point above the line describes someone who is taller than their shoe size predicts. The closer the data points cluster to the line, the more accurately the trend line describes individuals in the dataset.

*Perfect positive correlation Weak positive correlation No correlation*

Our example shows positive correlation because as height increases so does shoe size.

Negative correlation would imply the opposite: as height increased, shoe size decreased

*An example of negative correlation*

Our example is linear because the relationship does not change (we do not reach a certain point where very tall people start to have smaller feet).

*Example of quadratic relationship*

It is important to remember that *correlation does NOT imply causation.* While it might seem intuitive to say that being taller **causes** feet to generally be bigger, we cannot conclude this simply from the evidence of correlation. It would be just as consistent with the observed correlation to suggest that the larger feet cause the greater height or that there is a third variable (such as genetics or diet) that affects both height and shoe size in the same way.

**OCR Resources**: *the small print*OCR’s resources are provided to support the delivery of OCR qualifications, but in no way constitute an endorsed teaching method that is required by the Board, and the decision to use them lies with the individual teacher. Whilst every effort is made to ensure the accuracy of the content, OCR cannot be held responsible for any errors or omissions within these resources.   
© OCR 2017 - This resource may be freely copied and distributed, as long as the OCR logo and this message remain intact and OCR is acknowledged as the originator of this work.

OCR acknowledges the use of the following content: n/a

Please get in touch if you want to discuss the accessibility of resources we offer to support delivery of our qualifications: [resources.feedback@ocr.org.uk](mailto:resources.feedback@ocr.org.uk)

We’d like to know your view on the resources we produce. By clicking on ‘[Like](mailto:resources.feedback@ocr.org.uk?subject=I%20liked%20the%20A%20Level%20Biology%20A%20and%20B%20M1.7%20–%20Use%20a%20scatter%20diagram%20to%20identify%20a%20correlation%20between%20two%20variables)’ or ‘[Dislike](mailto:resources.feedback@ocr.org.uk?subject=I%20disliked%20the%20A%20Level%20Biology%20A%20and%20B%20M1.7%20–%20Use%20a%20scatter%20diagram%20to%20identify%20a%20correlation%20between%20two%20variables)’ you can help us to ensure that our resources work for you. When the email template pops up please add additional comments if you wish and then just click ‘Send’. Thank you.

If you do not currently offer this OCR qualification but would like to do so, please complete the Expression of Interest Form which can be found here: [www.ocr.org.uk/expression-of-interest](http://www.ocr.org.uk/expression-of-interest)

Looking for a resource? There is now a quick and easy search tool to help find free resources for your qualification:   
[www.ocr.org.uk/i-want-to/find-resources/](http://www.ocr.org.uk/i-want-to/find-resources/)

This resource has been produced as part of our free A Level teaching and learning support package. All the A Level teaching and learning resources, including delivery guides, topic exploration packs, lesson elements and more are available on the qualification webpages.

If you are looking for examination practice materials, you can find the Sample Assessment Materials (SAMs) on the qualification webpages: [Biology A](http://www.ocr.org.uk/qualifications/as-a-level-gce-biology-a-h020-h420-from-2015/) / [Biology B](http://www.ocr.org.uk/qualifications/as-a-level-gce-biology-b-advancing-biology-h022-h422-from-2015/)

### Produced in collaboration with the University of East Anglia