GCSE Maths – Statistics

Scatter Graphs

Worksheet

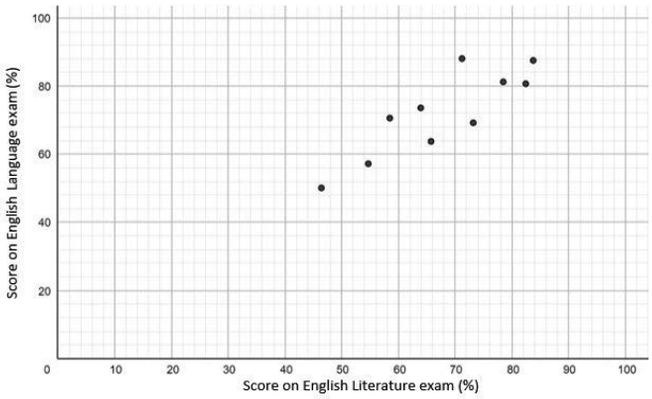
WORKED SOLUTIONS

This worksheet will show you how to work out different types of scatter graph questions. Each section contains a worked example, a question with hints and then questions for you to work through on your own.

Section A

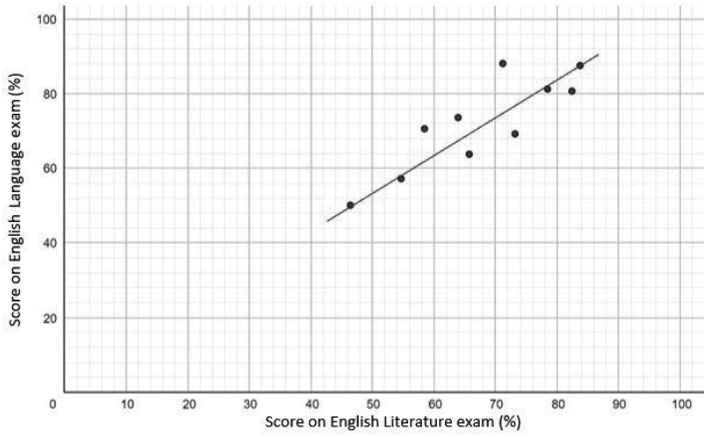
Worked Example

Below is a graph of the scores of a class on an English Literature and English Language exam. Draw a line of best fit and identify the type of correlation.



Step 1: Draw a line between the points, so that there are approximately the same number of points lying either side of the line.

The line of best fit should only span the data range and ignore any anomalies.



Step 2: Identify the type and strength of correlation by looking at the gradient of the line of best fit, and how many points lie on it.

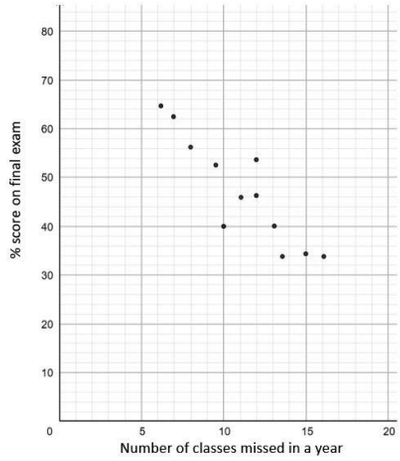
The gradient of the line of best fit is positive, so the correlation is ***positive****.*

Most of the points are not on the line of best fit, so the correlation can be described as ***weak****.*

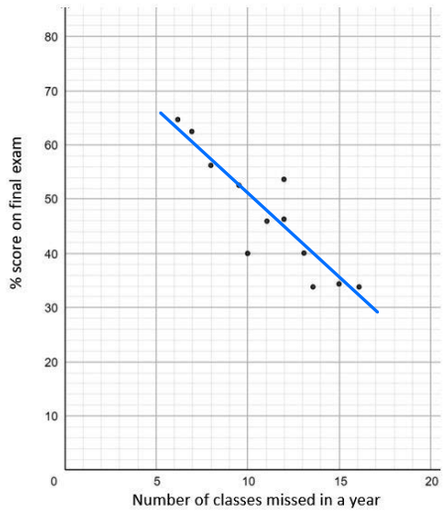
The graph shows a weak positive correlation between the scores on the English Language exam and the scores on the English Literature exam.

Guided Example

A college begins to track the attendance of its students; below is a graph of the results against the score they achieved on their final exam. Draw a line of best fit on the graph and describe the correlation.



Step 1: Draw a line between the points, which goes through as many of the points as possible and has roughly the same number of points either side of it. Ignore anomalies.



Step 2: Identify the type of correlation by looking at the gradient of the line of best fit and identify the strength of the correlation by looking at how many points lie on the line.

The gradient of the line of best fit is negative so the correlation is negative

Most points are close to/on the line so the correlation can be described as strong.

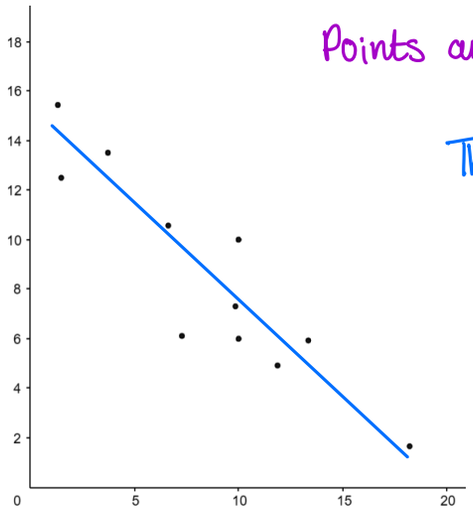
The graph shows a strong negative correlation between the number of classes missed and the % score of the final exam.

* As the correlation is not too strong, we could say it is moderately strong.

Now it's your turn!
If you get stuck, look back at the worked and guided examples.

1. On the following scatter graphs, draw the line of best fit and describe the correlation. Circle any anomalies.

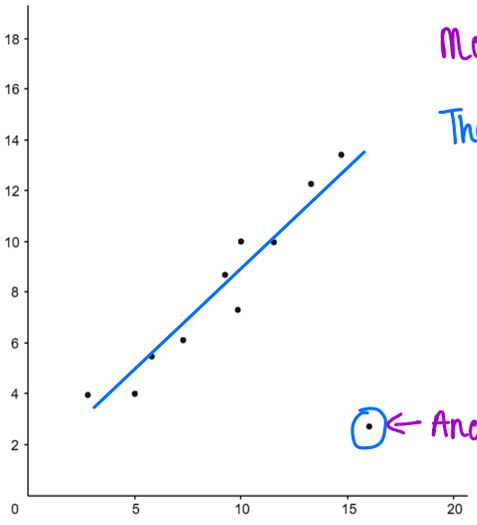
a)



Points are not close to the line of best fit.

The correlation can be described as weak negative.

b)

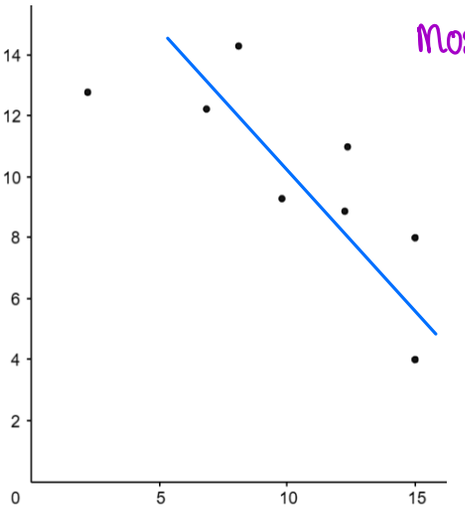


Most points are close to the line

The correlation can be described as strong positive.

Anomaly

c)



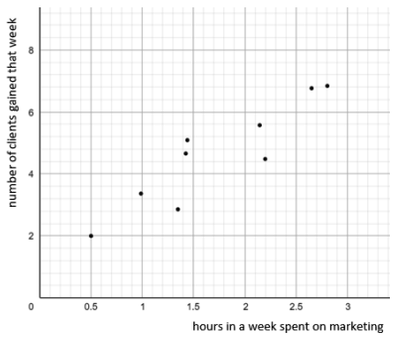
Most points are far away from the line

The correlation can be described as very weak negative.

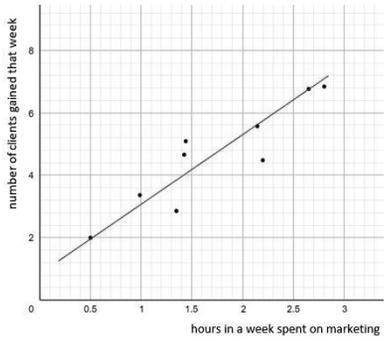
Section B

Worked Example

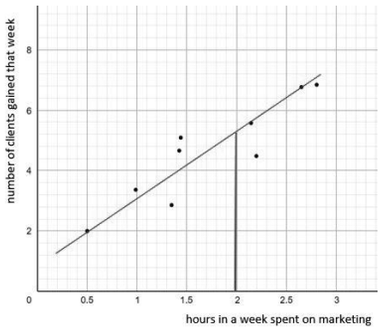
A small business has been tracking how effective its marketing strategy is. They have made a graph of the number of hours spent on marketing and the number of clients gained, which is pictured below. If the employees spend a total of two hours in a week on marketing, how many clients can they expect to gain?



Step 1: Draw a line of best fit on the graph.

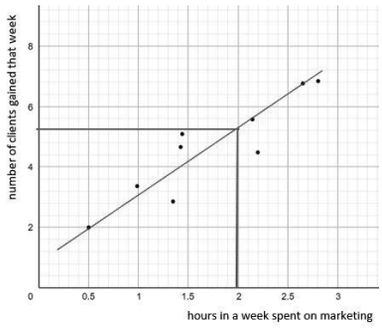


Step 2: Draw a vertical line from 2 hours on the x-axis to the line of best fit.



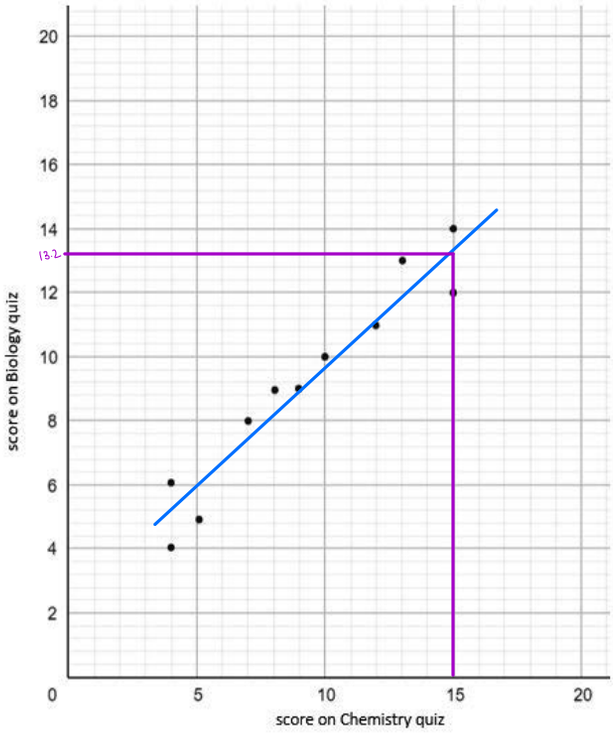
Step 3: Draw a horizontal line across to the y-axis to estimate how many clients gained from two hours of marketing.

*The value read from the y-axis rounds to **5 clients** gained in a week.*



Guided Example

Below is a graph of students' chemistry and biology scores on a quiz. Using the graph, predict what Clara will score on the biology quiz if she scores 15 on the chemistry quiz?



Step 1: Draw a line of best fit on the graph. ✓

Step 2: Draw a line from Clara's score on the x axis to the line of best fit.

15

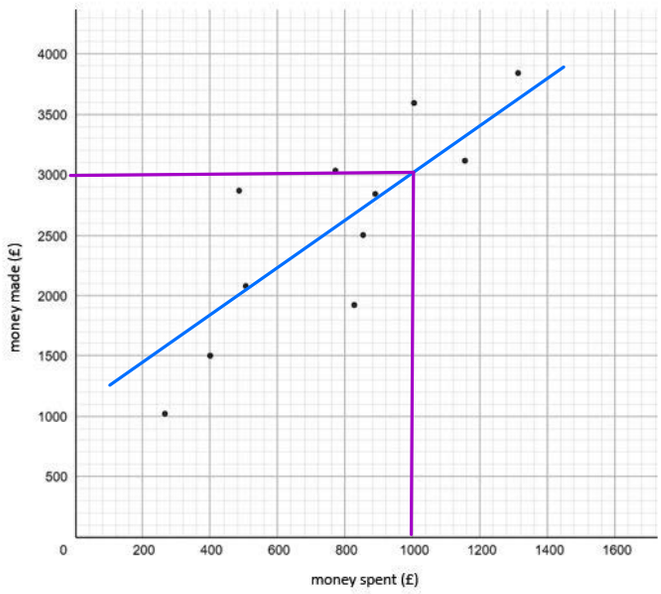
Step 3: Draw a line across from where the two drawn lines intersect to the y-axis, and proceed to read off the Biology score.

Clara's Biology score is 13.2

* Value can range between 12 and 14 depending on line drawn.

Now it's your turn!
If you get stuck, look back at the worked and guided examples.

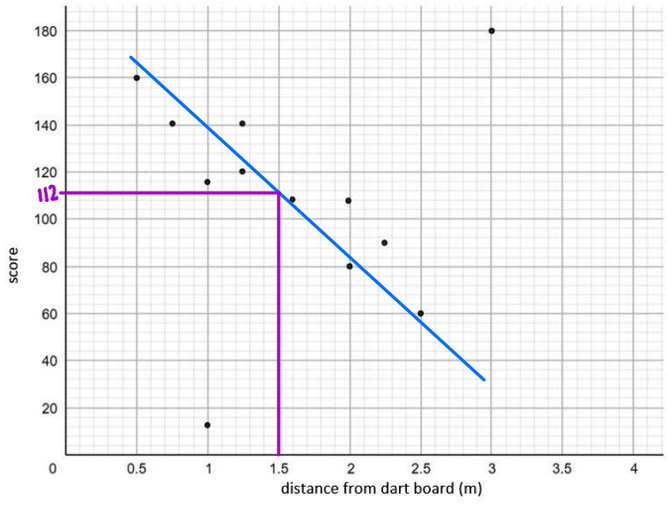
2. A small film company makes a graph of the money spent and made from each feature film it has produced. If the company spends £1000 making a film, how much can they expect to make from it?



They can expect to make £3000.

* Value can range between 2750 and 3250.

3. In an amateur darts competition, during the practice run, the organisers record the distance each turn is taken, and the score achieved. If a person stands 1.5m from the dart board, estimate what they will score.



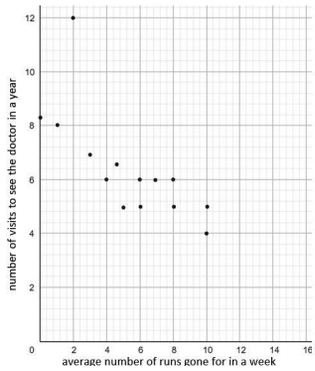
They should score 112

* Value can range between 104 and 124

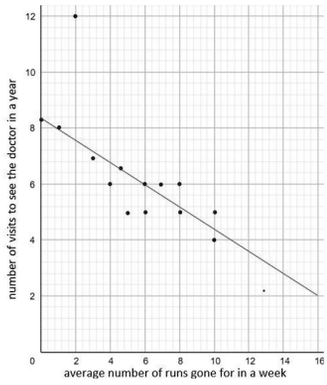
Section C

Worked Example

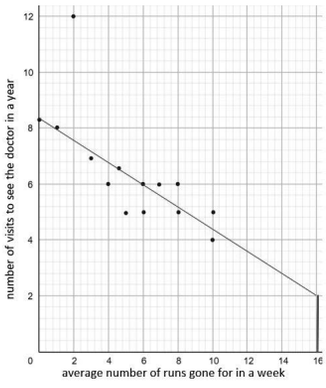
The graph shows the number of times a person has gone running in a week and the number of times they visited the doctor that year. If I go running 16 times in a week, how many times can I expect to see the doctor over a year? Comment on the reliability of your estimate.

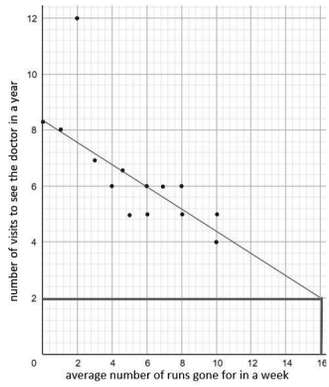


Step 1: Draw a line of best fit on the graph, ignoring the anomaly at $(2,12)$.



Step 2: Draw a line at $x = 16$ for the average number of runs on the x axis to the line of best fit. Draw a line across to the y –axis, to estimate how many visits to the doctors I can expect.





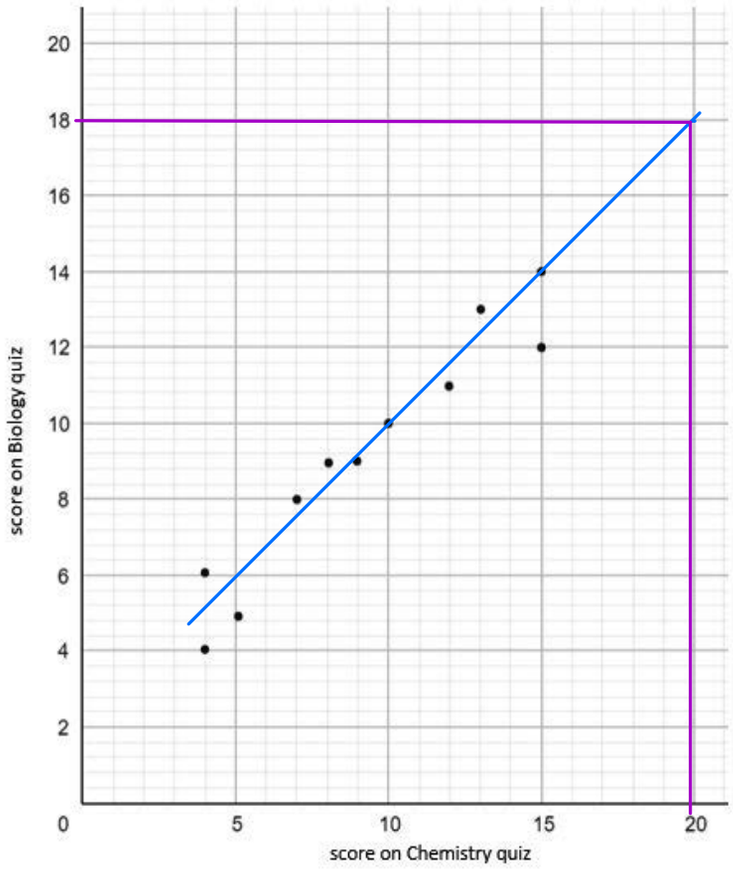
If I go on 16 runs weekly, I can expect to see the doctor twice.

Step 3: Decide whether you have used interpolation or extrapolation to evaluate the reliability of the estimate.

The value of 16 *runs is beyond the data range, which, as we can tell from the graph, is* 0 *to 10 runs. So, we have* ***extrapolated****, which means the* ***estimate is not very reliable.***

Guided Example

Below is a graph of students' chemistry and biology scores on a quiz. Using the graph, predict what Peter will score on the chemistry quiz if she scores 18 on the biology quiz? How reliable is your prediction?



Step 1: Draw a line of best fit on the graph. ✓

Since Peter's score exceeds the data range, make sure to extend the line beyond the data points.

Step 2: Draw a line from Peter's score on the y-axis to the line of best fit. ✓

Step 3: Draw a line across from where the two drawn lines intersect to the x-axis and proceed to read off the chemistry score. ✓

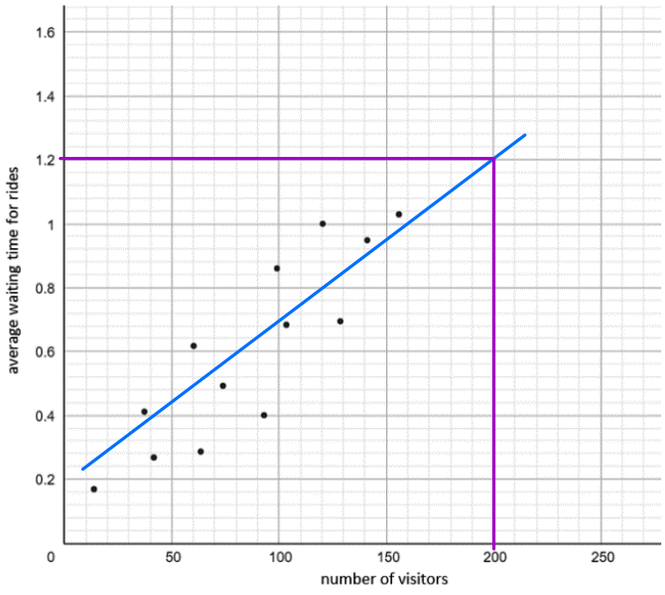
Step 4: Decide whether you have used interpolation or extrapolation to evaluate the reliability of the estimate.

An estimate for Peter's score is 20.

However, this isn't reliable as we have used extrapolation.

Now it's your turn!
If you get stuck, look back at the worked and guided examples.

4. This scatter graph shows the number of visitors to a theme park and the average wait time for rides per day. Predict how long the wait time will be when a record number of 200 people visit the theme park in one day.



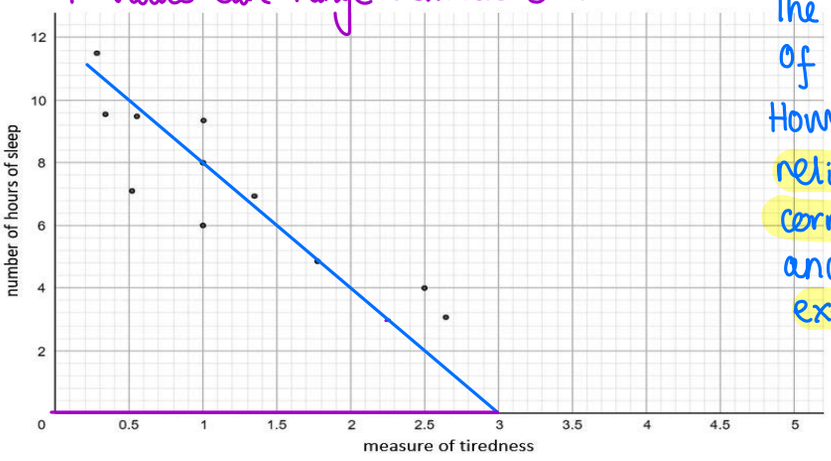
The predicted wait time is 1.2

* Value can range between 1.1 and 1.3.

5. A scientist has devised a quantitative method of measuring tiredness which gives subjects a score out of 5.

   The scientist tests his measure out and creates a scatter graph of participants' scores against how many hours of sleep they have per night. Using the graph, predict how many hours of sleep someone with a score of 3 gets per night. Comment on the reliability of the prediction.

* Value can range between 0 and 1



The predicted number of hour of sleep is 0. However, this is not reliable between the correlation is weak and we have used extrapolation.