edexcel

Guide to Maths for Geographers

AS and A Level Geography

Pearson Edexcel Level 3 Advanced GCE in Geography (9GE0)

Pearson Edexcel Level 3 Advanced Subsidiary GCE in Geography (8GE0)

ALWAYS LEARNING

PEARSON

Guide to Maths for Geographers

Contents

Introduction 3

1. Statistical graphs, charts and tables: data, bar charts, frequency tables and diagrams, pie charts, histograms 4

2. Graphs including lines of best fit, proportionality, gradients, relationships and correlations 20

3. Percentages and ratios 33

4. Visualise and represent 2-D and 3-D forms, including two-dimensional representations of 3-D objects 41

5. Calculate areas of triangles and rectangles, surface areas and volumes of cubes 44

6. Units and compound units, including conversion between units 47

7. Significant figures, decimal places, accuracy 52

8. Standard form 55

9. Descriptive statistics – average, mean, quartiles/percentiles and standard deviation 59

10. Inferential statistical tests: t-test, Spearman's Rank, Chi-squared 69

11. Sampling and data 79

Acknowledgements 82

Introduction

This guide to maths for geographers outlines the content that students will have covered in their maths lessons throughout KS3 and KS4. You can use this guide to help you understand how different areas are approached in maths, and therefore support your teaching of mathematical content in geography lessons.

The guide also provides support for AS and GCE geography in terms of maths skills and other quantitative skills. It should be used alongside the Fieldwork support materials from Pearson and other publishers.

The content is split into distinct mathematical concepts. Each chapter takes you through the terminology used in that area, as well as examples taken from Pearson maths textbooks to show you the methods students should be familiar with when solving mathematical problems.

Green highlighted sections/tables move beyond what is taught in Maths. They have a particular connection or reference to geography and may include geographical examples.

Sections which are highlighted in purple refer to inconsistencies between what is typically taught in Maths and what is taught in Geography, as well as common student errors.

1. Statistical graphs, charts and tables: data, bar charts, frequency tables and diagrams, pie charts, histograms

1.1 Data

Demand

All students learn the difference between discrete and continuous data in KS3. They also come across categorical data.

Terminology

Data is either qualitative (descriptive) or quantitative (numerical) and either discrete or continuous data.

Inconsistency: In geography, discrete data may be called discontinuous data.

- Discrete data can only take certain values, e.g. whole numbers, or shoe sizes. Continuous data is measured, e.g. length, time, and can take any value.
- Categorical data is where there is no numerical value, but data can still be sorted into groups, e.g. preferences from an interview.

- Quantitative data is data which has numerical values and is collected through measurement, e.g. recording slope angles.
- Qualitative data is non-numerical data that is used in a relatively unstructured and open-ended way. It is descriptive information, which often comes from interviews, focus groups or artistic depictions such as photographs.
- Primary data is new data, which has not previously been collected or processed. All data collected in person is primary data. Secondary data is often obtained through research of what has typically already been published.

The table below gives some examples of other specific data types, specific to AS/A level geography, from the specification skills lists.

Example data type	Comment
Big data	Several ideas exist including: large volume of data, complex, high velocity (quick accumulation rates), high levels of variation and variety. Examples might include transport (London Underground movements) as well as how companies such as Amazon handle customer logistics.
Crowd sourced data	Individuals and organisations use contributions from internet users to obtain insight into required services or ideas. Crowd sourced information could be used to obtain interview and questionnaire data, for instance, e.g. Twitter and blogs.

Discursive and creative material	Information which might be in a written format, or contain a combination of pictures and text. Discursive is sometimes presented as a report or brochure and contains persuasive information, whereas creative is more personal and artistic. This could be music , poetry, twitter feeds or blogs for instance.
Geospatial	Data and information which has geographical positioning information included within it, such as a road network in GIS form, or a geo-referenced satellite image. OS map co-ordinates also provide a geo-spatial reference.

1.2 Bar charts

Demand

All students learn to draw and interpret bar charts for discrete and continuous data in KS3.

Lower ability maths students (KS3 and KS4) may need help with interpreting scales on axes given in, for example, thousands (i.e. 2.2 thousand = 2200) or millions.

Approach

- Can show qualitative or quantitative discrete or continuous data.
- One axis is usually labelled 'Number of …' or Frequency.
- Frequency is usually shown on the vertical axis (but can be on the horizontal axis with the bars in the chart shown horizontally).
- Bars should be of equal width.
- For discrete or qualitative data there are gaps between the bars.
- A bar-line graph, for discrete or qualitative data, uses lines instead of bars. It can be used to save time drawing the bars.
- For continuous data there are no gaps between the bars.
- In questions on interpreting proportions from bar charts, ask for the fraction or percentage of students with brown eyes, not 'the proportion of students'.

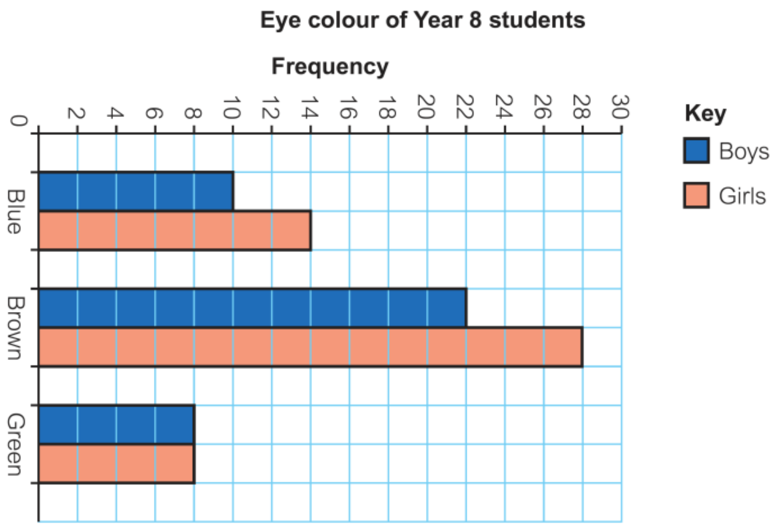


Figure 1 Horizontal bar chart (discrete, qualitative data)

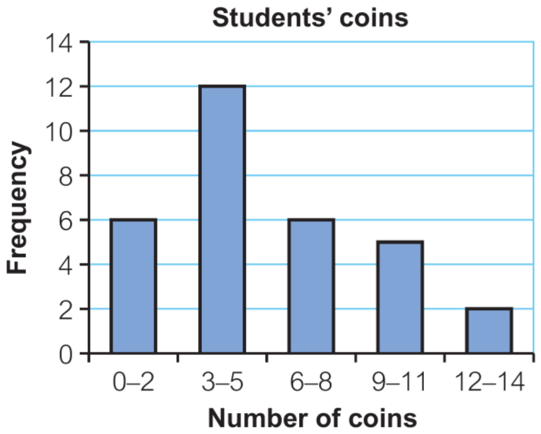


Figure 2 Bar chart (discrete, quantitative data): gaps between the bars

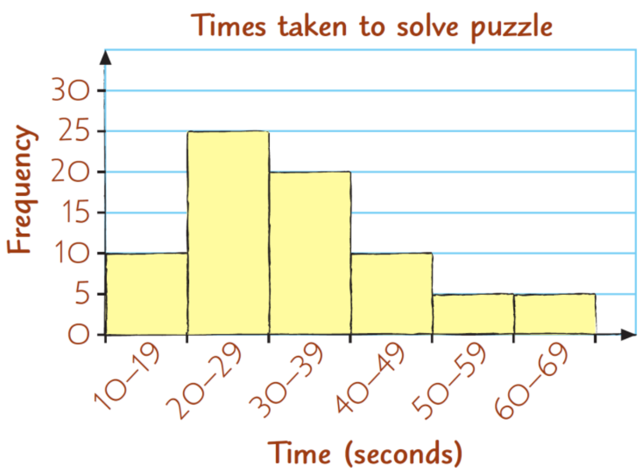


Figure 3 Bar chart (continuous, quantitative data): no gaps between the bars

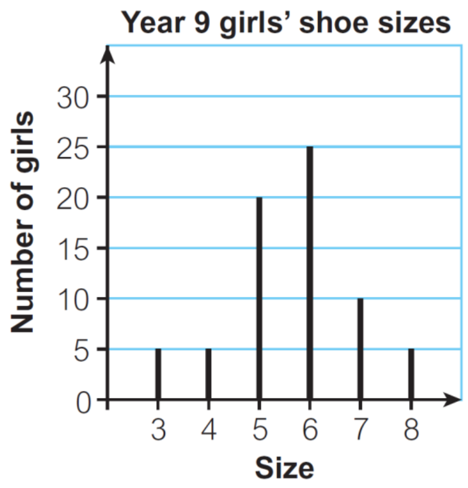


Figure 4 Bar-line chart (discrete, quantitative data)

In geography, there is a potential for more sophisticated bar charts to be used, including those that can include two types of data. A climate graph is a good example (Figure 5). These can cause confusion, so careful explanation is required of both the line and the bars, as well as the scales either side of the graph.

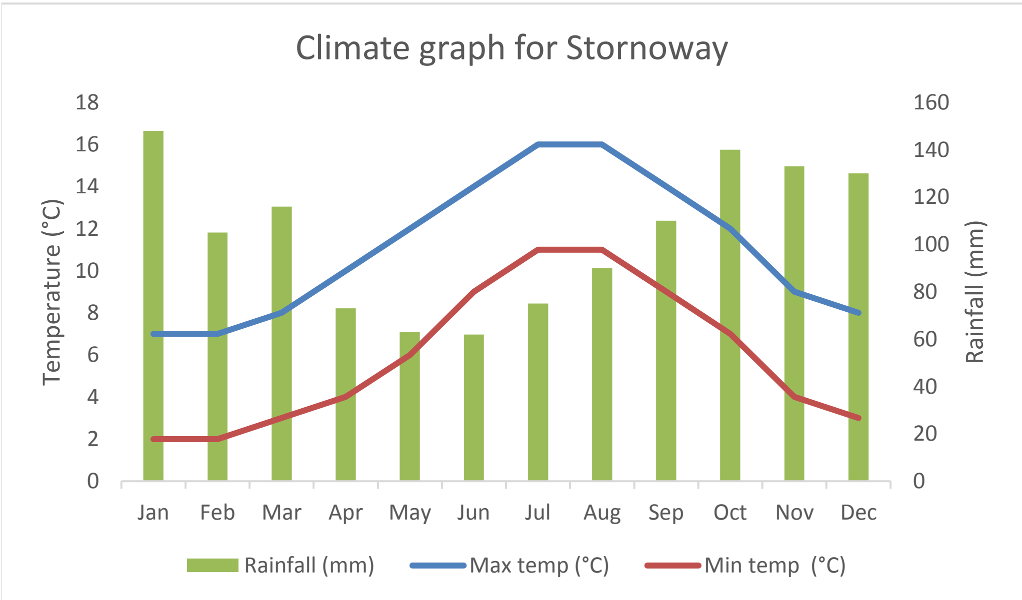


Figure 5 A climate graph showing the average monthly temperature (°C) and rainfall (mm) for Stornoway.

1.3 Frequency table

Demand

All students learn to draw and interpret frequency tables for discrete and continuous data in KS3.

Approach

- A table of data that shows the number of items, or frequency of each data value or each data group.
- Data can also be grouped. For discrete data use groups such as 0–5, 6–10, etc.
  For continuous data use groups such as $0 \leq t < 10$, $10 \leq t < 20$.
  The groups must not overlap.
- In maths, students learn that it is best to group numerical data into a maximum of 6 groups. If you need them to group data differently, tell them how many groups of equal width to group it into.

Shoe size	Frequency
3	3
4	5
5	7
6	10
7	10
8	6
9	1

Figure 6 Frequency table (ungrouped discrete, quantitative data)

Geography mark	0–10	11–20	21–30	31–40	41–50
Frequency	4	13	17	19	7

Figure 7 Frequency table (grouped discrete, quantitative data)

Distance (d metres)	Frequency
$10 \leq d < 20$	2
$20 \leq d < 30$	6
$30 \leq d < 40$	15
$40 \leq d < 50$	20
$50 \leq d < 60$	4

Figure 8 Frequency table (grouped continuous data)

1.4 Frequency diagram

Demand

All students learn to draw and interpret bar charts for discrete and continuous data in KS3.

Lower ability maths students (KS3 and KS4) may need help with interpreting scales on axes given in, for example, thousands (i.e. 2.2 thousand = 2200) or millions.

Approach

Inconsistency: In KS3 maths, the name 'Frequency diagram' is not used.

- Frequency diagram is another name for a bar chart where the vertical axis is labelled Frequency.
- They can be used to show discrete or continuous data.

These two bar charts could also be called frequency diagrams:

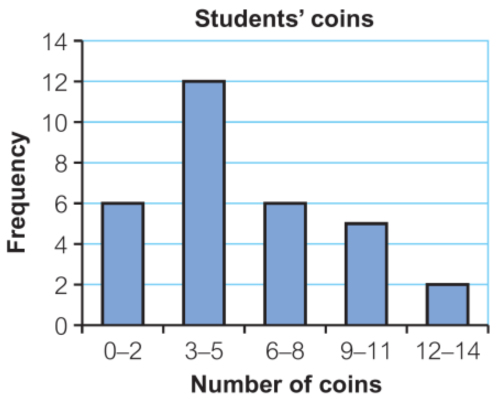


Figure 9 Frequency diagram (discrete, quantitative data)

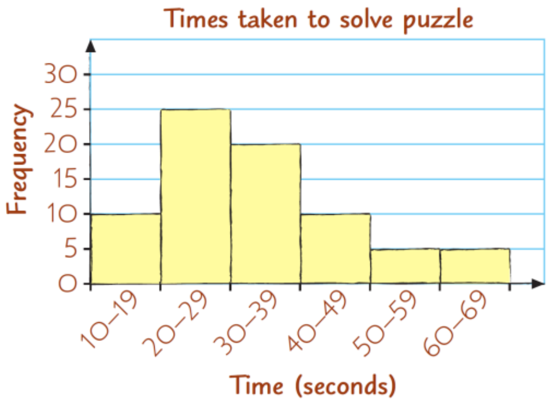


Figure 10 Frequency diagram (continuous quantitative data)

1.5 Comparative bar chart

Demand

All KS3 students learn to draw and interpret comparative bar charts in KS3.

Approach

- Compares two or more sets of data.
- Uses different coloured bars for each set of data.
- Needs a key to show what each colour bar represents.

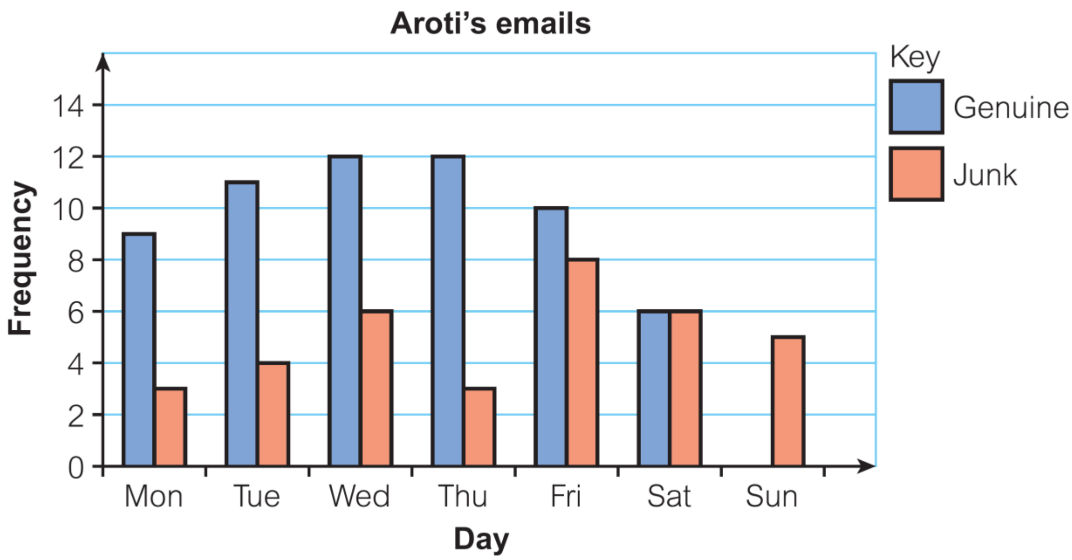


Figure 11 Comparative bar chart (discrete, qualitative data)

1.6 Compound bar chart

Demand

All students learn to draw and interpret compound bar charts in KS3.

Approach

- Combines different sets of data in one bar.
- Needs a key to show what each colour section represents.
- In questions on interpreting proportions in compound bar charts, ask for the fraction or percentage of chemistry students getting A*, not 'the proportion of students'.

Students often find it difficult to read the intermediate values on the graph i.e. students achieving an A grade. In geography, students may be required to compare the total frequency of subject choice here, as well as the relative fractions or percentages of students getting a particular grade. We sometimes see these graphs converted to a percentage on the Y-axis in geography, i.e. 0-100%, rather than as a frequency as in Figure 12.

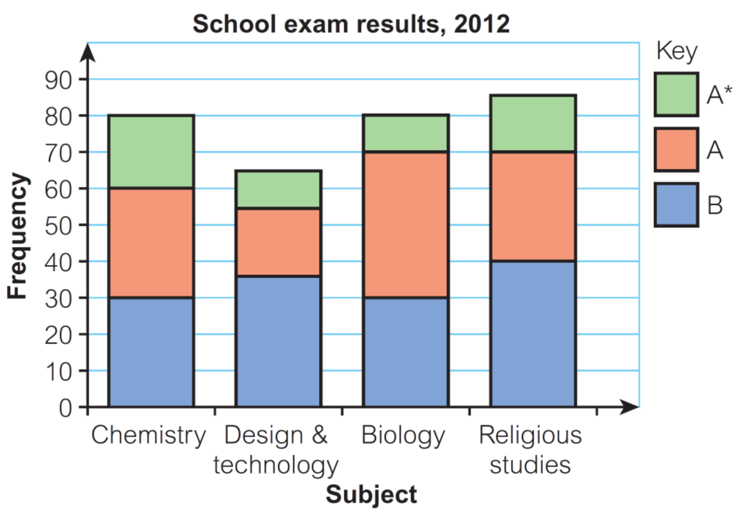


Figure 12 Compound bar chart (discrete, quantitative data)

In AS/A Level geography the demand may be increased, using a more complex scale or greater range of categories. Figure 13, from a legacy 2016 paper, shows economic losses caused by natural disasters 1998–2012.

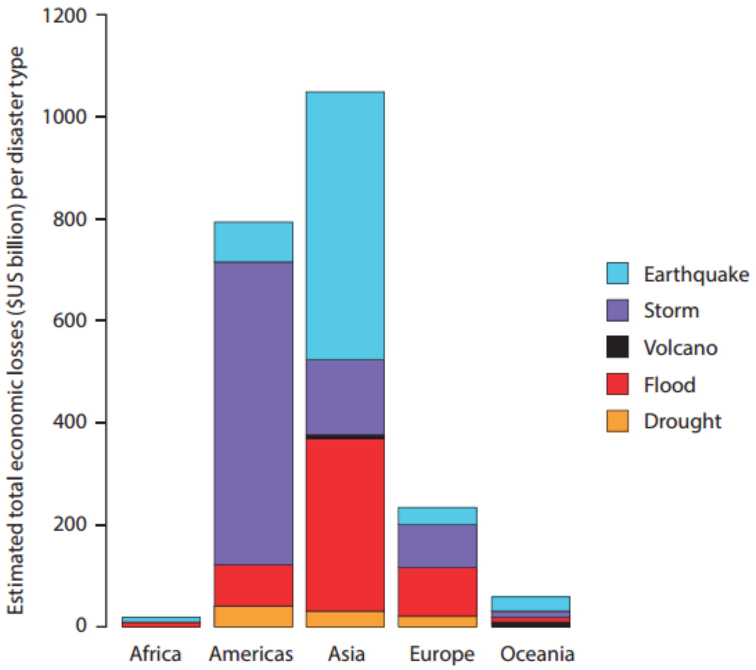


Figure 13 Compound bar chart (interpretation is more demanding)

1.7 Histogram

Demand

In maths, students do not meet histograms until KS4, although the bar charts they draw in KS3 for grouped continuous data could also be called histograms.

Histograms with unequal width bars/groups, where frequency density is plotted on the vertical axis, are only covered in Higher GCSE Maths, not Foundation.

Approach

- Can be drawn for grouped continuous data where groups/bars are of equal width.
- No gaps between the bars.
- If groups/bars are of unequal width, the vertical axis is labelled frequency density, which is calculated as:

$$\frac{\text{number in group}}{\text{group width}}$$

- The area of the bar is proportional to the number of items it represents (frequency).

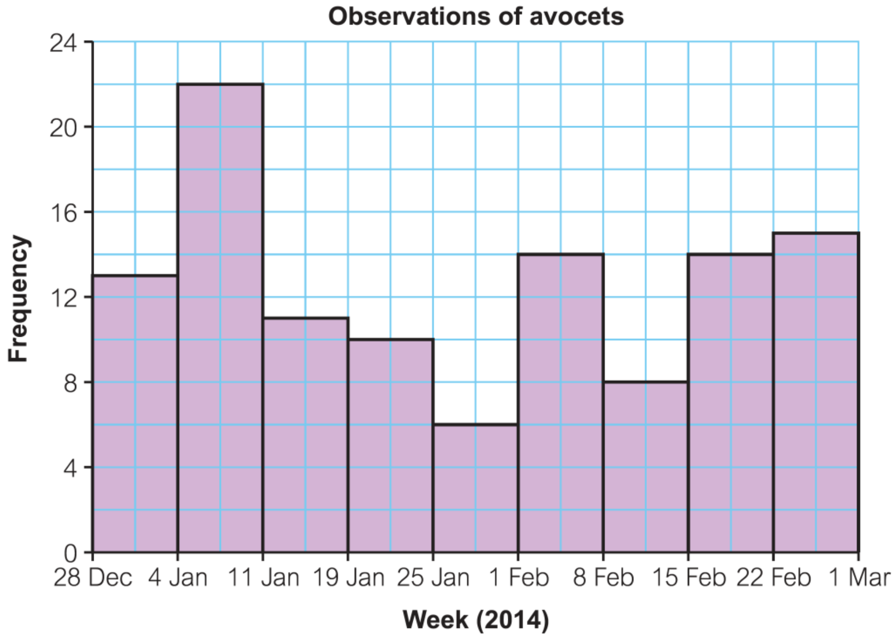


Figure 14 Histogram with equal width bars/groups

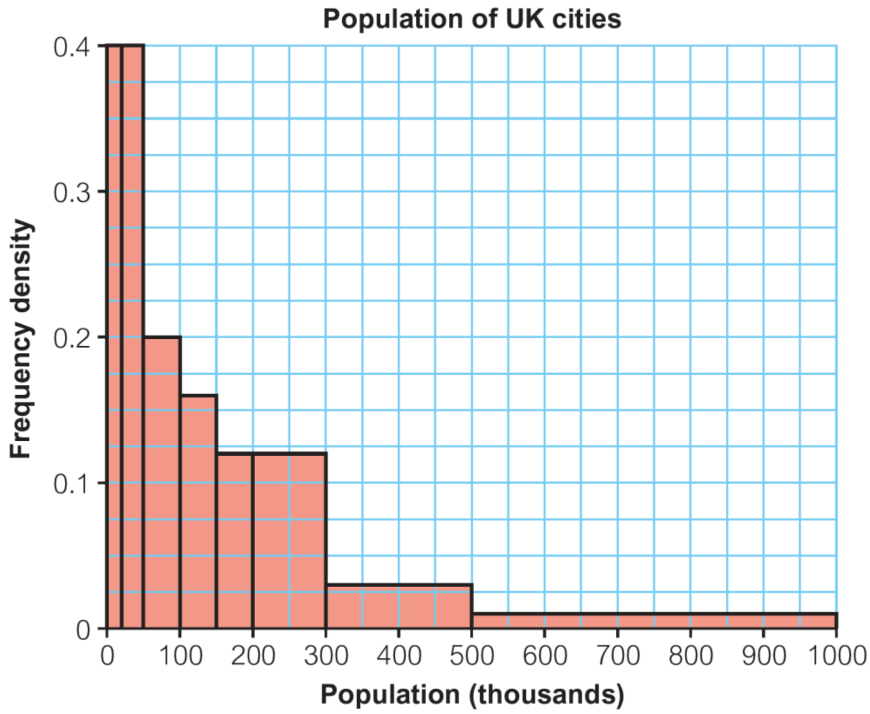


Figure 15 An example histogram with unequal width bars/groups: used when the data is grouped into classes of unequal width (i.e. based on the range of data within that group or class).

Drawing a histogram

From Pearson GCSE Mathematics Higher:

Key point 12

In a **histogram** the area of the bar represents the frequency. The height of each bar is the frequency density.

$$\text{Frequency density} = \frac{\text{frequency}}{\text{class width}}$$

Example 4

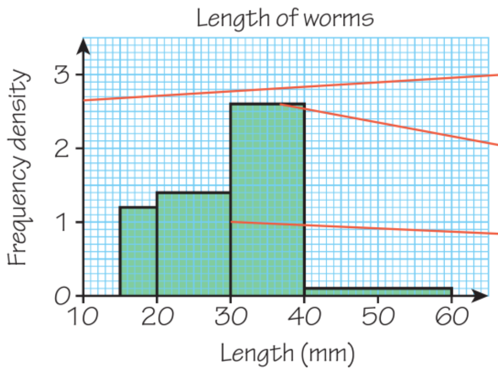
The lengths of 48 worms are recorded in this table.

Length, x (mm)	$15 < x \leqslant 20$	$20 < x \leqslant 30$	$30 < x \leqslant 40$	$40 < x \leqslant 60$
Frequency	6	14	26	2

Draw a histogram to display this data.

$6 \div 5 = 1.2,\ 14 \div 10 = 1.4,\ 26 \div 10 = 2.6,\ 2 \div 20 = 0.1$

Work out the frequency density for each class



Label the y-axis 'Frequency density'.

The height of each bar is the frequency density for each class.

Draw the bars with no gaps between them.

Source: *Edexcel GCSE (9-1) Mathematics Higher* student book

1.8 Pie charts

Demand

In KS3 all students should learn how to construct a simple pie chart.

Lower ability students will probably struggle with working out the angles as they will not have learned how to calculate percentages or fractions that are not nice round numbers.

All students draw and interpret pie charts in GCSE Maths.

Terminology

- A pie chart is a circle divided into **sectors**.
  NB: a 'slice' of a pie is a sector, not a **segment**.
- The angle of each sector is proportional to the number of items in that category
- Shows proportions of a set of data, e.g. fraction or percentage of waste recycled
- May need a key.

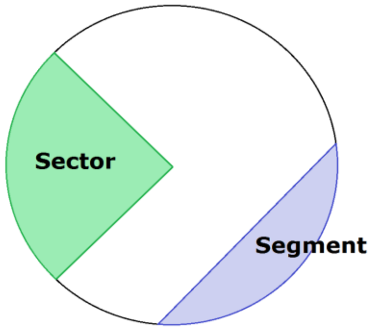


Figure 16 Pie chart terminology

In questions asking students to interpret a pie chart, ask for the fraction or percentage who learn Geography, not 'the proportion' who learn Geography.

Approach

Drawing a pie chart

Taken from Pearson GCSE Maths Foundation:

Example 5

The table shows the match results of a football team. Draw a pie chart to represent the data.

Result	Won	Drawn	Lost
Frequency	28	12	20

Total number of games = $28 + 12 + 20 = 60$ — The total number of games is the total frequency.

$\div 60$ 60 games : 360° $\div 60$
1 game : 6°

1 game = $360 \div 60 = 6°$ — Work out the angle for one game.

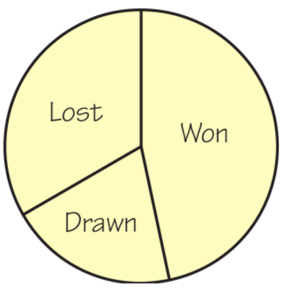
Won: $28 \times 6° = 168°$

Drawn: $12 \times 6° = 72°$ — Work out the angle for each result.

Lost: $20 \times 6° = 120°$

Check: $168 + 72 + 120 = 360$ — Check that your angles total 360°.

Team results



Draw the pie chart. Give it a title and label each section, or make a key.

Source: *Edexcel GCSE (9-1) Mathematics Foundation* student book

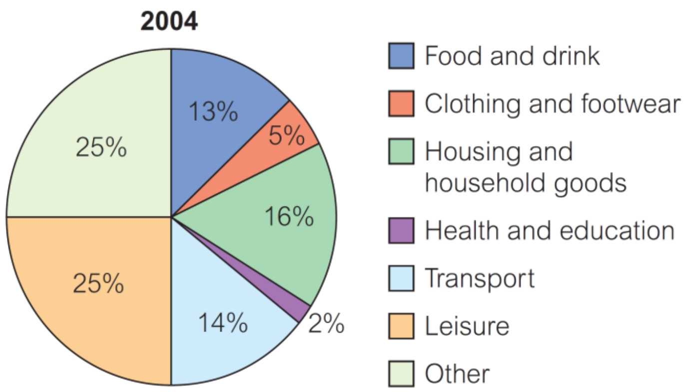


Figure 17 Pie chart showing average household spending (percentages)

1.9 Drawing graphs and charts to display data

Demand

Choosing a suitable graph or chart to draw to display data is above Level 6 in maths.

Lower ability students in KS3 and Foundation GCSE maths students may need some guidance on the type of chart to draw for given sets of data.

Lower ability maths students may also need help with choosing suitable scales for axes.

Geography students should think carefully about the number of categories that can be used so that the graph doesn't become too complex to interpret. It might be that an alternative is more appropriate.

Other graphs and charts relevant for AS/A level geography

The Pearson/Edexcel specification for AS/A level geography, which follows the DfE (2014) criteria, refers to additional types of graphs and charts including dot maps and kite diagrams.

Dot maps

Dot (distribution) maps (also known as dot density maps) use a dot symbol to show the presence of a feature at a location. Dot maps rely on a visual scatter to show spatial pattern. Figure 18 shows the sphere of influence of visitors to a honeypot site and accompanying isolines (which are lines connecting points of equal numbers of tourists in this instance).

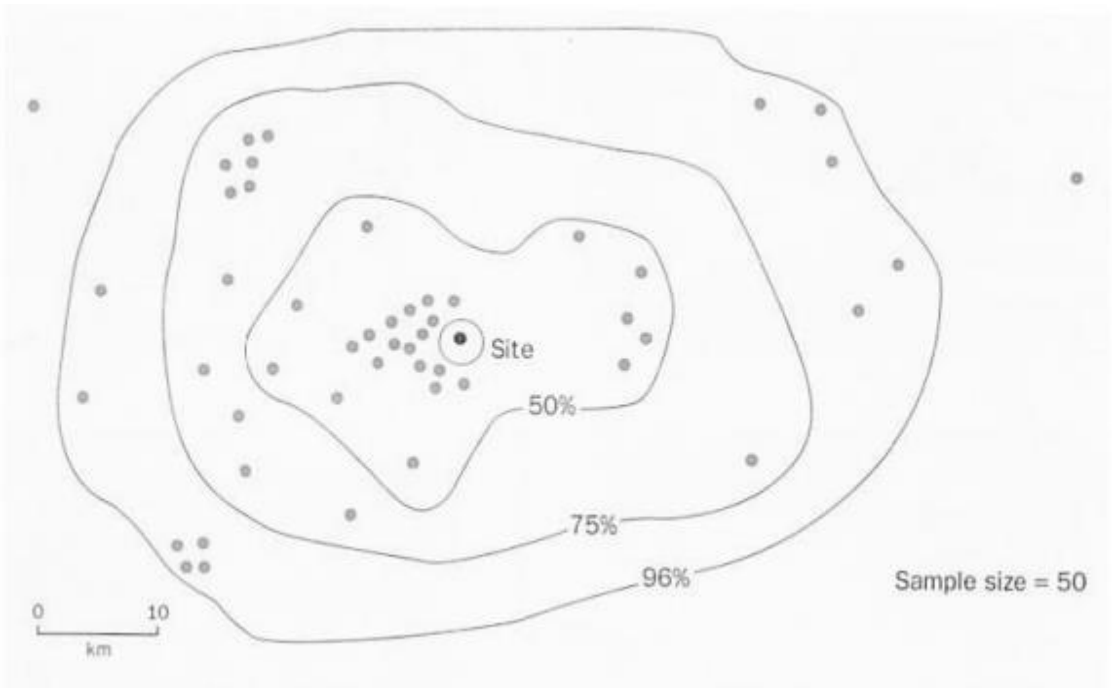
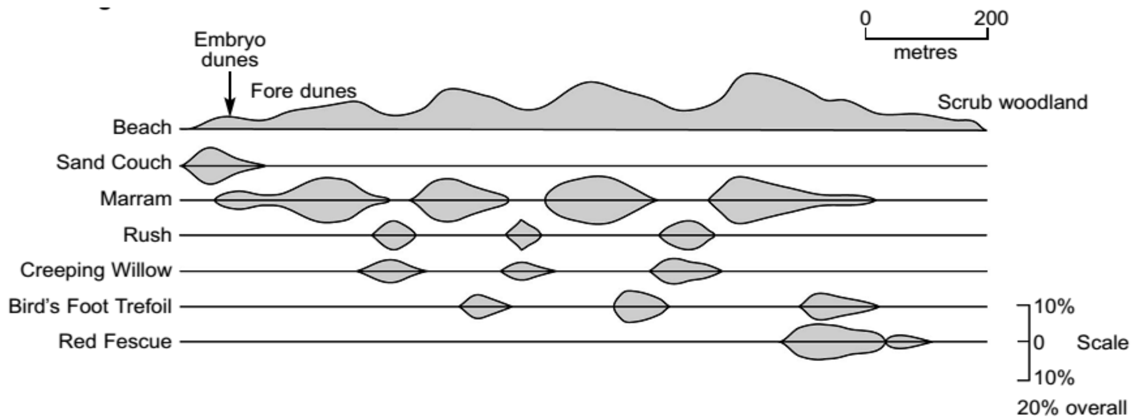


Figure 18 An example of a dot distribution map at a local scale

Kite diagrams

Kite diagrams (or graphs) typically show the changes in plant frequency along a line of sampling (a transect). The plot visually exaggerates the changing species, making it easier to spot patterns in the zonation (or evidence of succession). There are some good guides explaining how to make a kite graph on YouTube: search 'How to make a Kite Graph' to find an example from Darron Gedge's Geography Channel.

In an AS fieldwork exam, a kite diagram could be presented linked to 'changing coastal profiles'.



Students could be asked to:

Explain two reasons why this type of diagram was used to present the fieldwork data. (4 marks)

In this context, an AO3 response might be based around the following:
The kite diagram shows changes spatially, which allows the reader to compare frequency changes in individual species across the 1000+ distance.

It also makes it easier for the reader to compare changes in the frequencies of different plant species, identifying where one plant might be becoming more frequent (dominant) at the expense of another.

Other specialised charts

In AS/A level geography, specialised charts can also be used to display data in a comparative context, e.g. proportional circles and proportional flow lines. These types of graphs are used to help visualise the data more easily and then, in this instance, make comparisons between countries and cities.

Two examples are provided, showing:

- the population of selected African cities in 2010 and projected growth to 2025 (Figure 19), and
- 'Top ten' global remittance flows in 2010 (valued in $US billions) (Figure 20).

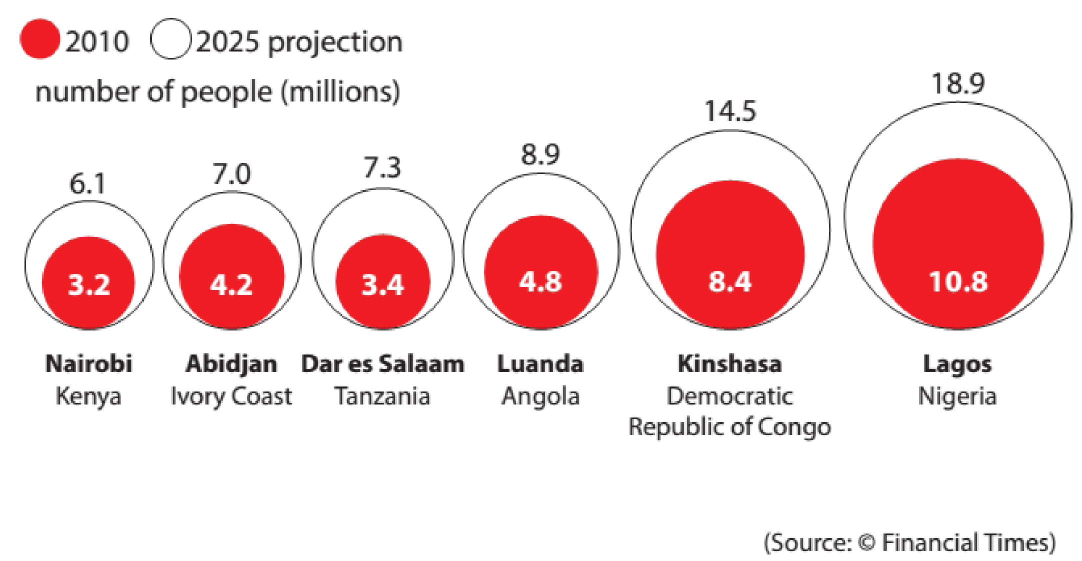


Figure 19 Using embedded proportional circles to allow comparison of data

(Source: © The Economist Newspaper Limited 2010)

Figure 20 Proportional arrows and flow lines are commonly used in exam stimulus material

In AS/A level geography, students are expected to be able to independently choose a suitable graph or chart. They should consider factors such as visual appeal, technical difficulty of construction, appropriateness of plot choice, etc.

Lorenz curves and Gini coefficient of inequality

The Gini coefficient was developed by the Italian statistician Corrado Gini (1912) as a summary measure of income inequality in society. It is usually associated with the plot of wealth concentration introduced a few years earlier by Max Lorenz (1905).

The Gini coefficient is based on a statistical technique. The coefficient ranges from a score of 0 (complete equality) to 1.0 (or sometimes 100%), which is complete inequality.

A global interpretation (in terms of personal wealth) of the Gini coefficient is as follows:

- Low inequality: under 0.299, e.g. Belarus and Hungary
- Relatively low inequality: 0.300–0.399, e.g. Nepal and Poland
- Relatively high inequality: 0.400–0.449, e.g. Russia and Nigeria
- High inequality: 0.450–0.499, e.g. China and Mexico
- Very high inequality: 0.500–0.599, e.g. Hong Kong and Colombia
- Extremely high inequality: 0.600 and over [many of the countries in this group are in southern Africa].

Source: https://en.wikipedia.org/wiki/List_of_countries_by_income_equality (CIA Gini %). There are lots of examples online, e.g. YouTube, which provides "walk-throughs" of the calculations and interpretations. S

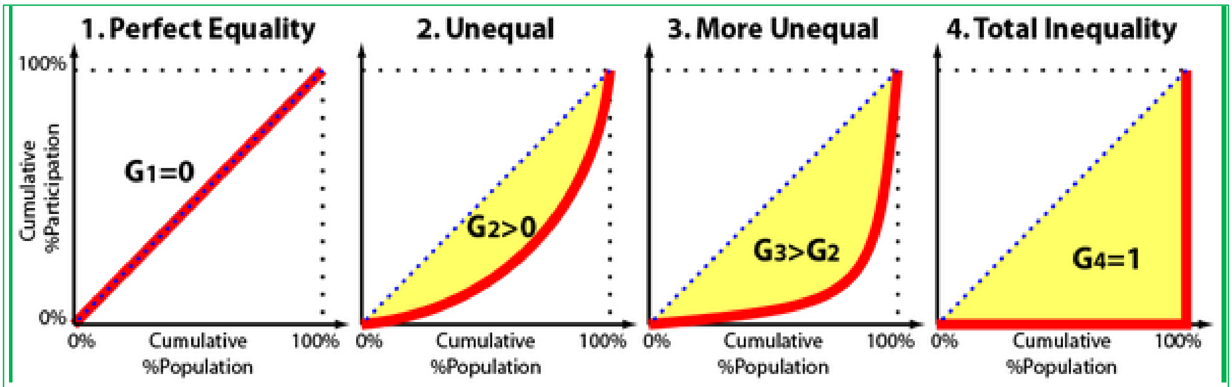


Figure 21 Comparing different Lorenz curves and the relationship to inequality

- A Lorenz curve (Figure 21) illustrates the degree of unevenness (or inequality) in a geographical distribution. It is drawn on graph paper and makes use of cumulative percentage data. The vertical axis carries the cumulative data, and points are plotted in the order of the largest first, to which is then added the second largest, then the third largest, and so on. The horizontal axis simply records the cumulative process.
- The plots are then connected by a line. If another line is drawn onto the graph to represent an even distribution (a diagonal line), then the degree of unevenness (or inequality) can be seen.
- The greater the deviation the plotted line has from the line of even distribution, the greater the degree of unevenness. A highly concave Lorenz curve represents a high level of unevenness (inequality), and therefore a high level of concentration.

In the context of an exam for AS and AS/A level Geography, students will be provided with a formula.

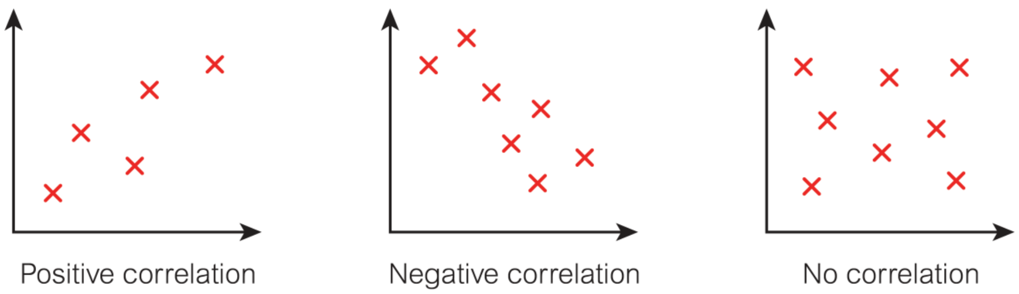
2. Graphs including lines of best fit, proportionality, gradients, relationships and correlations

2.1 Scatter graphs

Terminology

Inconsistency: In maths, these are called scatter graphs, not scatter diagrams or scatter plots. In geography we may unknowingly mix up the terminology and therefore cause confusion.

- A scatter graph plots two sets of data on the same graph to see if there is a relationship or correlation between them.
- Scatter graphs can show positive, negative or no correlation.



- Correlation is when two sets of data are linked. For example, when one value increases as the other decreases, or when one value decreases as the other increases.
- In maths, points on scatter graphs are plotted with crosses.
- The line of best fit follows the shape of the data and has roughly the same number of crosses above and below the line. There may also be crosses on the line.
- In geography, we may be asked to describe the strength of the correlation. Figure 22 has a stronger (observed) correlation than Figure 23, because of the distance of the crosses (points) away from the line of best fit.

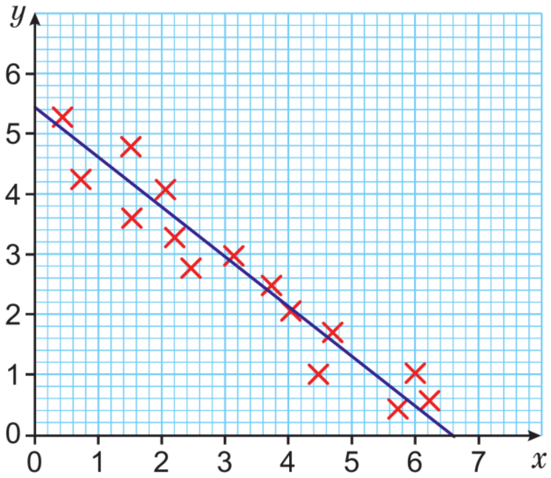


Figure 22 Strong correlation

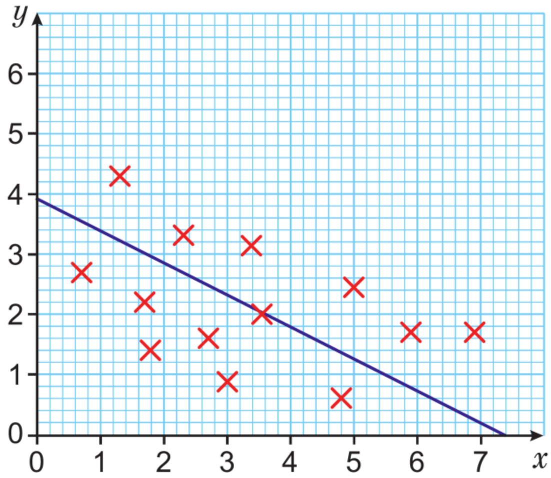
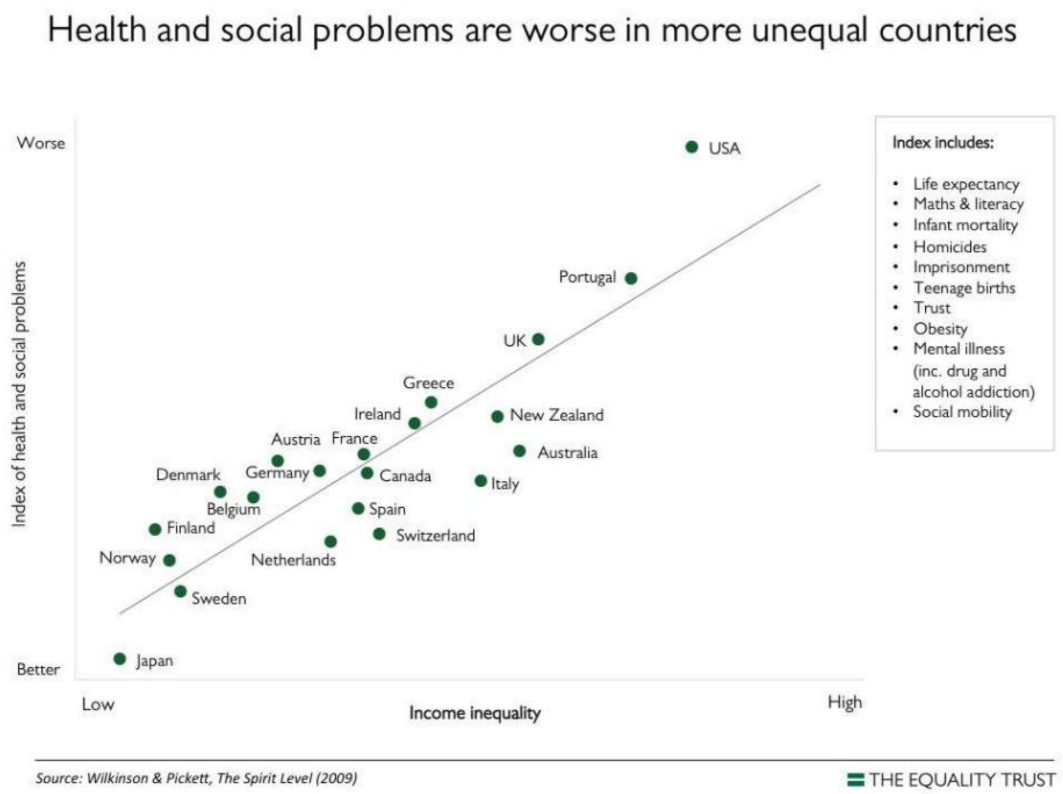


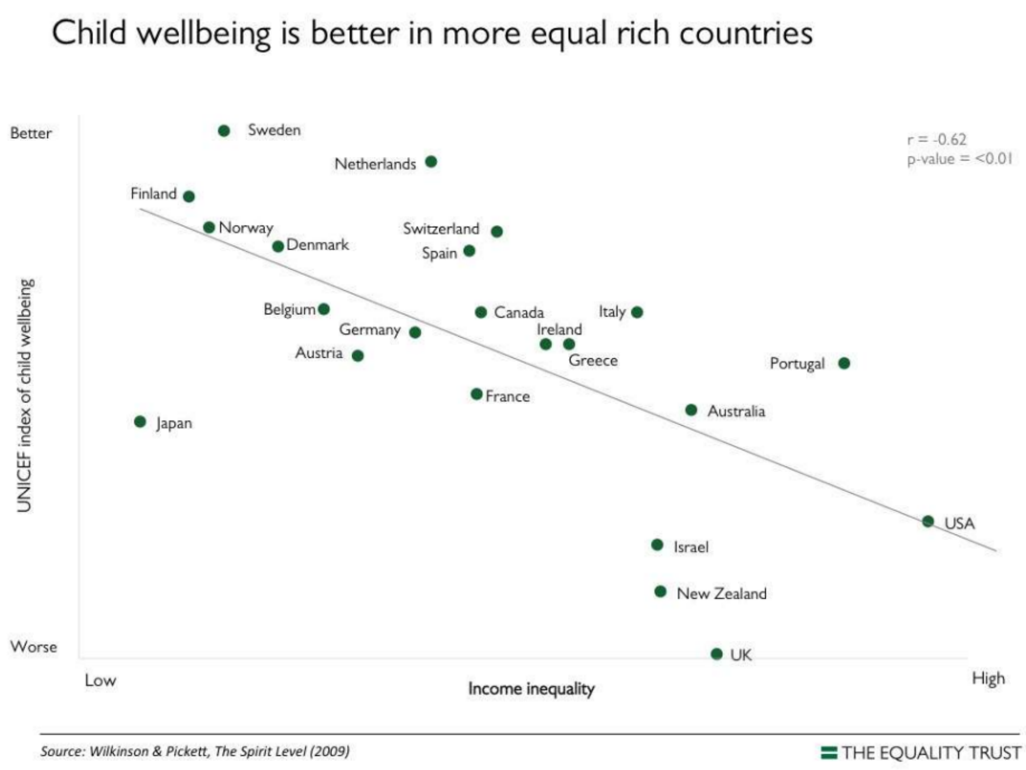
Figure 23 Weak correlation

In AS/A level geography, scatter graphs can be used to 'tell geographical stories', for example the relationship between 'health and social problems' and 'income inequality'.



(Source: www.equalitytrust.org)

Figure 24 Strong positive correlation using real and complex data



(Source: www.equalitytrust.org)

Figure 25 Strong negative correlation of data.
The UK is outside of the expected pattern, i.e. an outlier.

Inconsistency: In maths, when interpreting a scatter graph, an acceptable answer is 'shows positive correlation', unless the question explicitly asks for this to be explained in context. We should use the same language in geography and then comment on the geographical meaning or relevance of such a relationship.

- The line of best fit shows a relationship between two sets of data.
- When the points on a scatter graph are on or close to a straight line:
  - There is strong correlation between the variables.
  - There could be a linear relationship between the variables, e.g. $y = mx$ or $y = mx + c$. The equation of the line of best fit describes this relationship.
- When the points on a scatter graph are not close to a straight line there may be another relationship between the variables.

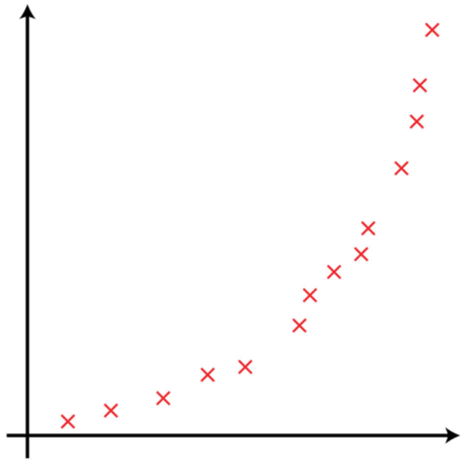
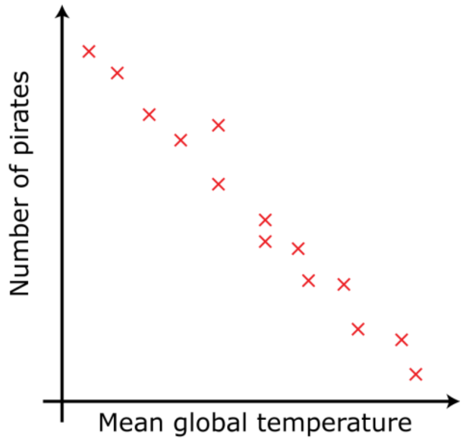
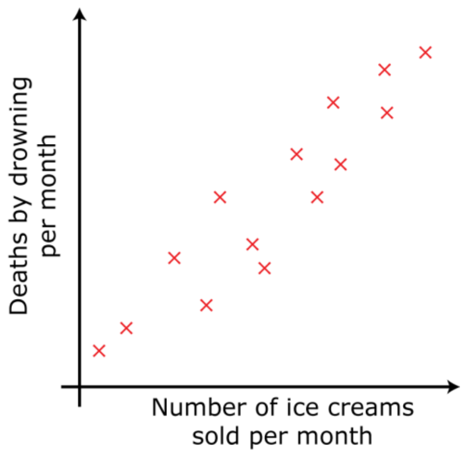


Figure 26 Scatter graph showing a possible non-linear relationship

- Correlation does not imply causation. Sometimes there may be another factor that affects both variables, or there may be no connection between them at all.



There is a negative correlation between number of pirates and mean global temperature, but it is unlikely that one causes the other!



There is positive correlation between number of ice creams sold and death by drowning, but it is unlikely that one causes the other (i.e. directly causal). A more likely explanation is a third factor – temperature. On hot days more people buy ice creams and more people swim, leading to increased numbers of drownings.

Many areas of geography do not show clear causal linkages.

There are numerous examples online of discussions around causal and non-causal correlations. One example is here: http://data-analysis.org/tag/correlation/. But there are many other references that could be explored.

In geography we also consider the strength of relationships as well as direction.

Approach

Drawing scatter graphs

Lower ability maths students would not be expected to know which variables to put on which axis for a scatter graph. They may need help with deciding which is the independent variable, and a reminder that this goes on the horizontal axis.

Drawing a line of best fit

Place your ruler on the graph, on its edge. Move the ruler until it is following the shape of the data, with roughly the same number of points above or below it. Ignore any points on the line.

Common error

Students often try to make their lines of best fit go through (0, 0) or the origin. A line of best fit does not necessarily pass through the origin. It should stop at the first or last plot point or cross, as it is not possible to extrapolate data beyond these reference points.

In geography, data (especially fieldwork data) often doesn't go to zero as it cannot be measured e.g. river velocity, so students need to be made aware of this.

2.2 Interpreting scatter graphs

Demand

All students should learn to interpret scatter graphs in KS3. They will have learned about correlation and causation in KS3 maths.

Lower ability students at GCSE would not be expected to know which variables to put on which axis for a scatter graph.

At GCSE Higher level, they would be expected to state that there is a possible non-linear relationship if the points on a scatter graph closely follow a smooth curve.

Common error

Students often find interpreting scatter graphs difficult as they do not know how to put into words what the graph shows, so it is good to give them examples of graphical interpretations, or at least sentences to copy and complete. For example:

- The steeper the gradient, the ____________ the river's energy.

You can also use statements like this:

- The greater the __________, the longer the ___________.

Anomalies and outliers

In AS/A level geography, students should be expected to explain the significance of anomalies and outliers on scatter graphs. In a statistical sense these two terms are essentially identical and refer to points in a pattern that are outside the normal range (i.e. well away from the line of best fit when considering a scatter plot).

However, in the context of experimental design and investigation, an outlier should be regarded as a legitimate data point that's far away from the mean or median in a distribution. Whereas an anomaly might be considered 'noisy', i.e. an untrustworthy data point that's generated by a different process than whatever generated the rest of the data.

2.3 Line graphs

Demand

Lower ability students at KS3 would not be expected to know which variables to put on which axis for a scatter graph. They may need help with deciding which is the independent variable, and a reminder that this goes on the horizontal axis.

In geography, the independent variable is often time.

Terminology

- In maths, a line graph that shows how a variable changes over time (i.e. with time on the horizontal axis) is often called a time-series graph.
- Line graphs can show trends in data. The trend is the general direction of change, ignoring individual ups and downs.

In geography, we are frequently exposed to climate change graphs, surface temperature graphs, with trend lines, and population prediction graphs.

This website has climate-related graphs (from the IPCC) which show clear trends: www.carbonbrief.org/ipcc-six-graphs-that-explain-how-the-climate-is-changing

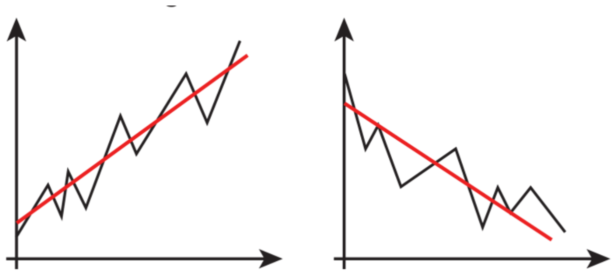


Figure 27 Examples of graphs showing increasing (left) and decreasing (right) trends

2.3.1 Drawing line graphs

Demand

Lower ability students at KS3 would not be expected to know which variables to put on which axis for a line graph. They may need help with deciding which is the independent variable, and a reminder that this goes on the horizontal axis.

In geography, the independent variable is often time.

Choosing a suitable graph or chart to draw to display data is above Level 6 in maths.

Lower ability students in KS3 and Foundation GCSE maths students may need some guidance on the type of chart to draw for given sets of data.

Lower ability maths students may also need help with choosing suitable scales for axes.

Approach

- In maths, students draw graphs on squared or graph paper.
- They plot points with crosses (×).
- They join points with straight lines or a smooth curve – the question needs to tell them which.
- All graphs should have labels on the axes and a title.
- When more than one data set is shown, the lines could be, for example, one solid and one dashed. The graph will need a key to explain solid/dashes. Alternatively, colour-code the lines according to category.

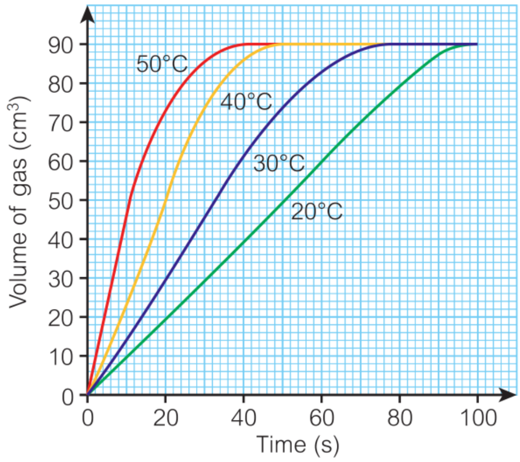


Figure 28 Line graph showing more than one data set

2.3.2 Interpreting line graphs

Demand

All students interpret line graphs in KS3.

Lower ability maths students (KS3 and KS4) may need help with interpreting scales on axes given in, for example, thousands (i.e. 2.2 thousand = 2200) or millions.

At KS4, all students will have limited experience of interpreting real-life graphs that dip below zero.

Only higher ability maths students are likely to have seen graphs with two different vertical scales to read from.

Approach

Students interpret 'real-life' graphs in maths, in a variety of contexts.

They may be less familiar with:

- graphs showing negative values, such as the one below

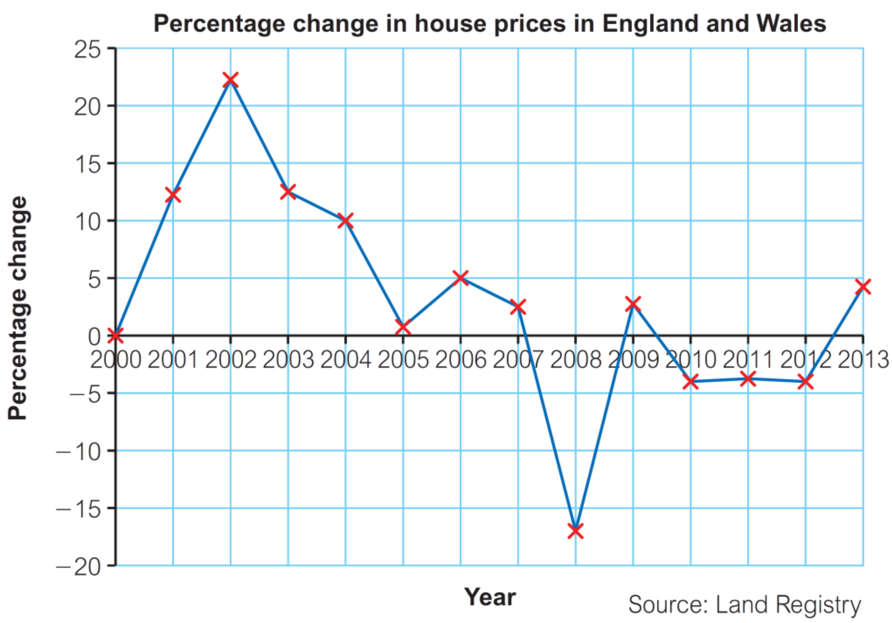


Figure 29 Graph showing negative values

- two types of graph on one set of axes, and two different vertical axes for the same graph.

In geography, an example of this is a climate graph, similar to the one presented earlier. It is normal to show temperature as a continuous line and rainfall using discrete bars (for its monthly average totals).

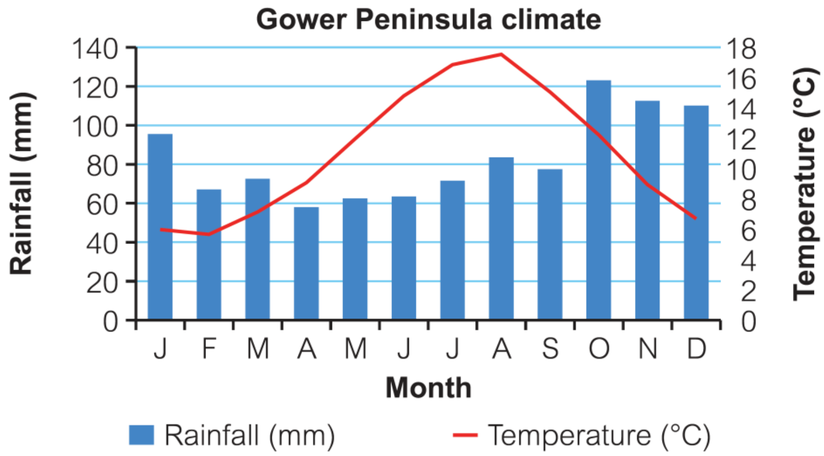


Figure 30 Graph showing two types of graph on one set of axes, including two different vertical axes

NB: Ensure the terms used in any questions match the labels on the graph.

2.4 Recognising proportional relationships from graphs

Demand

In year 9, the majority of maths students should know that a straight line graph shows the two variables are in direct proportion (please use direct proportion, not just proportion). Only top set maths students will have met graphs showing inverse proportion.

In KS4 all students should learn that the origin is the point (0, 0).

In KS4 all students will meet graphs showing inverse proportion.

Terminology

- A straight line graph through the origin (0, 0) shows that the two variables are in **direct proportion**.
  When one variable doubles, so does the other. When one halves, so does the other.
  The relationship is of the form $y = mx$, where m is the gradient of the graph.
- A straight line graph not through the origin shows a **linear relationship**.
  The relationship is of the form $y = mx + c$, where m is the gradient of the graph and c is the y-intercept (where the graph crosses the y-axis).

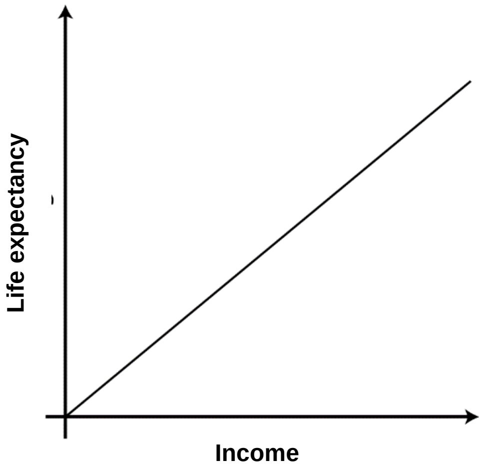


Figure 31 Graph showing two variables in direct proportion to each other

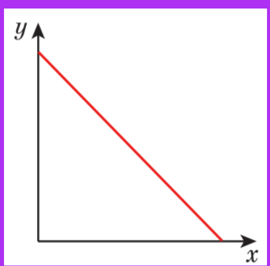
The graph shows a directly proportional relationship (theoretical) between a country's income and life expectancy. The gradient is $m \approx 10$.

You can find more relationships such as this on the Gapminder website: www.gapminder.org/.world/

This type of data is relevant to development, inequality and globalisation topics.

Common error

A straight line graph with negative gradient shows a linear relationship – not an inverse proportion. This is a fairly common misconception.



Logarithmic scales in geography

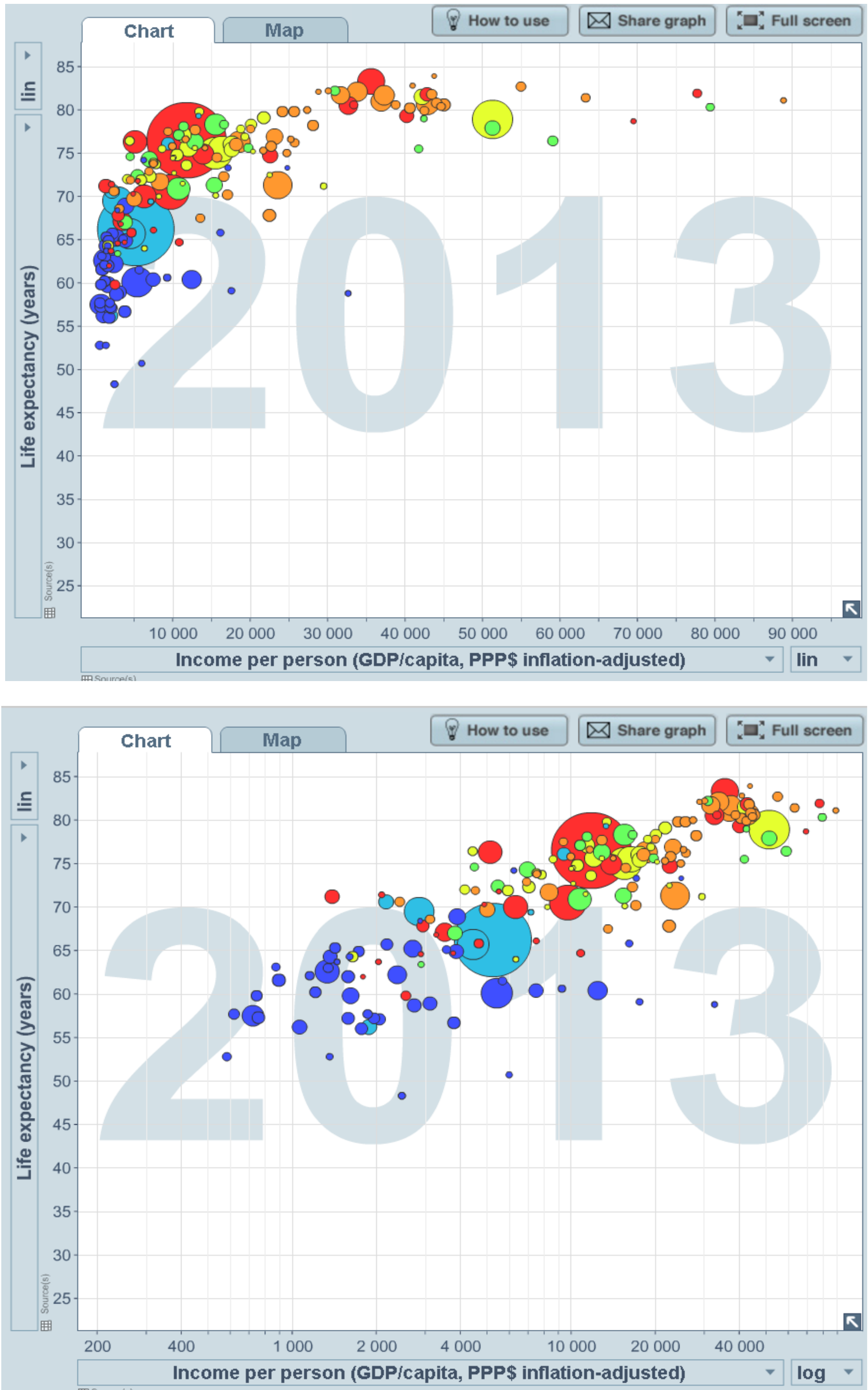
There are numerous examples of the use of logarithmic scales in geography, both in physical and human aspects of the discipline, e.g. theoretical process relationships in a river channel (Hjulström curve), earthquakes (for both the MMS and Richter scales), noise (dB), incomes (e.g. Gapminder website), etc.

- Logarithmic scales allow you to work with a large range of numbers and thereby identify patterns and trends more easily.
- Logarithmic scales are divided into a number of cycles, each representing a 10-fold change in the range of values, e.g. 0.01 to 0.1; 0.1 to 1.0; 1.0 to 10; 10 to 100, etc.

Students should be aware of the disadvantages and limitations of such graphs. These include the fact that zero cannot be plotted and it's not possible to have both positive and negative values on the same graph.

The lines are spaced in each cycle according to the logarithms of the numbers 1 to 10 – these can be difficult to plot without a computer or specialist graph paper. It can also lead to very different conclusions, based on whether the reader interpreting the plot knows the data is in fact plotted logarithmically.

Figure 33, from the Gapminder website, illustrates the same data being plotted: the one on the top uses a logarithmic- (or log)-scale (the default), and the one at the bottom uses a linear scale. Log scales are used by default for many of the Gapminder graphs, since they allow data with a wide range to be plotted on to the same graph.



Source: www.gapminder.org/world/

Figure 32 The default scale (log-scale) versus the linear scale – for the same data set. Circles are sized proportionally to population.

Dispersion diagrams

These specialised diagrams are especially useful when comparing differences between areas or localities (Figure 34 compares deprivation data between two areas). One or more variables can be plotted so that a distribution pattern is revealed.

Critically, it is useful for seeing the spread of data (e.g. the standard deviation), as well as the range.

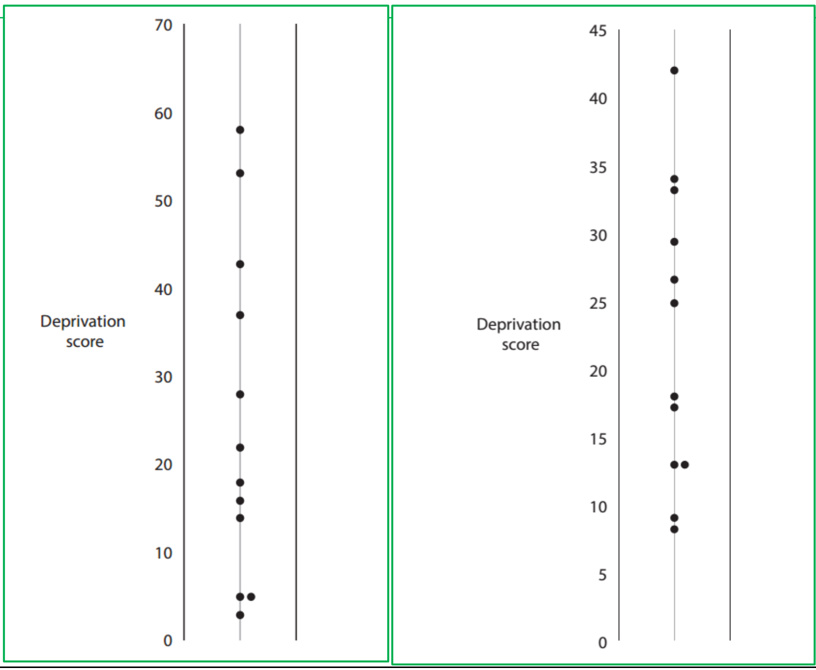


Figure 33 Example of dispersion diagrams

Dispersion diagrams are useful in presenting where the Upper Quartile and Lower Quartile can be found as well as the mean, median, mode and extreme values and interquartile range.

This would be a useful approach in coursework, for example, when looking for differences in a sample between two locations. This approach might work well for sediment sizes on two different parts of a beach, for instance.

Other similar diagrams, for example Box and Whisker, can show both spread of data and the statistical spread.

3. Percentages and ratios

3.1 Percentages

Demand

In KS3, students in lower maths sets will only have written one number of a percentage of another where the larger number is a multiple or factor of 100, so may find this difficult. They would not be expected to do this calculation in maths without a calculator.

Percentage change

Foundation level students learn this in the Autumn term of year 11.

Approach

Convert a percentage to a fraction

Worked example

Write 20% as a fraction.

$$20\% = \frac{20}{100} \overset{\div 10}{=} \frac{2}{10} \overset{\div 2}{=} \frac{1}{5}$$

First write 20% as a fraction of 100.

Then simplify the fraction by dividing the numerator and denominator by the same number. Keep doing this until the fraction is in its simplest form.

Convert a percentage to a decimal

Worked example

Write 35% as a decimal.

$$35\% = \frac{35}{100} = 0.35$$

Write 35% as a fraction out of 100. Then divide 35 by 100 to write it as a decimal.

Convert a fraction to a percentage

Convert the fraction to a decimal, then convert the decimal to a percentage.

For example: $\frac{34}{80} = 0.425 = 42.5\%$

Students can input $\frac{34}{80}$ as a fraction into a scientific calculator and press = (or the S-D button on some calculators) to get the equivalent decimal.

Write one number as a percentage of another

Write as a fraction, then convert to a percentage.

For example, in a class of 28 students, 13 are boys. What percentage are boys?

$\frac{13}{28} = 0.4642\ldots = 46.4$ (1 dp)

Without a calculator:

For example, in a high street with 40 shops, 24 are charity shops. What percentage are charity shops?

$$\text{Percentage of charity shops} = \frac{\text{number of charity shops}}{\text{total number of shops}} \times 100$$

$$= \frac{\overset{3}{\cancel{24}}}{\underset{\underset{1}{\cancel{8}}}{\cancel{40}}} \times \overset{20}{\cancel{100}}$$

$$= 60\%$$

Calculating a percentage of an amount

50% is the same as $\frac{1}{2}$, so to find 50% divide by 2.

10% is the same as $\frac{1}{10}$, so to find 10% divide by 10.

To calculate, for example, 30% mentally, you can find 10% and multiply by 3.

To calculate 5% mentally, find 10% and halve it.

Calculating percentages using a calculator

Input the percentage as a fraction

For example, to calculate 30% of 20 m, input $\frac{30}{100} \times 20$ and press = to get 6 m.

Input the percentage using a decimal multiplier

65% = 0.65

So to calculate 65% of 80, input 0.65 × 80 and press = to get 52.

Percentage increase/decrease

Work out the increase and add it on/subtract it

Examples

To increase 45 by 20%

20% of 45 = 9

45 + 9 = 54

To decrease 220 by 5%

5% of 220 = 11

220 – 11 = 209

Using a multiplier

Examples

To increase 45 by 20%

After the increase you will have 100% + 20% = 120% = 1.2

1.2 × 45 = 54

To decrease 220 by 5%

After the decrease you will have 100% – 5% = 95%

$0.95 \times 220 = 209$

Finding the original amount

Using arrow diagrams

Worked example

20% of an amount is £40.
Work out the original amount.

$$20\% = £40 \xrightarrow{\div 20} 1\% = £2 \xrightarrow{\times 100} 100\% = £200$$

Example

When studying ecosystems, for example, consider the area of a woodland that has been reduced by 15% because of deforestation.

The final area is 320 km^2.

Calculate the original area.

To calculate the area after a 15% decrease, you would multiply by 0.85:

$$\text{original} \longrightarrow \boxed{\times\ 0.85} \longrightarrow 320\ \text{km}^2$$

$$376.5\ \text{km}^2 \longleftarrow \boxed{\div\ 0.85} \longleftarrow 320\ \text{km}^2$$

Percentage change

$$\text{percentage change} = \frac{\text{actual change}}{\text{original amount}} \times 100$$

Example

When studying development, for instance:

In 2010, a region’s annual GDP was £80m.

In 2014, the same region has increased its economic output to £120m.

The actual increase in GDP is 120 – 80 = £40m.

The fractional increase is $\frac{\text{actual increase}}{\text{original price}} = \frac{40}{80}$

$\frac{40}{80}$ as decimal is 0.5

Percentage increase is $0.5 \times 100 = 50\%$

Below is an example of a Pearson/Edexcel sample question based around this skill.

(ii) Calculate the percentage growth in GDP (PPP) for China from 2000 to 2015.

You must show your working.

(2)

Terminology

- Percent means 'out of 100'. A percentage is a fraction with denominator 100.
- You can calculate percentages of amounts, e.g. 20% of $500.
- You can write one number as a percentage of another, e.g. write $\frac{7}{50}$ as a percentage.

3.2 Ratios

Demand

Students learn to simplify ratios and write them in the form 1 : *n* or *n* : 1 in KS3.

Students learn to relate ratios to fractions in KS3, but many continue to make errors with this type of calculation.

Approach

Simplifying ratios

A ratio in its simplest form only contains whole number values.

Divide all the numbers in the ratio by the highest common factor:

2 : 6

÷2 ÷2

1 : 3

6 : 15

÷3 ÷3

2 : 5

This ratio is not in its simplest form, because the two numbers both still have a common factor, 2:

8 : 20

÷2 ÷2

4 : 10

Writing in the form 1 : n (sometimes called a unit ratio)

Divide both numbers by the first number in the ratio:

$$5 : 7 \xrightarrow{\div 5} 1 : 1.4$$

Writing in the form n : 1

Divide both numbers by the second number in the ratio:

$$20 : 12 \xrightarrow{\div 12} 1.67 : 1$$

Comparing ratios

Write both ratios in the form 1 : *n* or *n* : 1.

Example

When studying development, for instance:

In Country A there are 20 people unemployed for every 120 people who are economically active.

In Country B there are 15 people unemployed for every 85 people who are economically active.

Which country has more unemployment?

Country A

$$20 : 120 \xrightarrow{\div 20} 1 : 6$$

Country B

$$15 : 85 \xrightarrow{\div 15} 1 : 5.7$$

Country B has a lower ratio, so as a proportion there are more unemployed people.

Ratio and proportion

When studying coastal landscapes, for instance:

A cliff is composed of two different rock types (geologies), A and B, in the ratio 2 : 3.

What fraction of the cliff is:

a) Geology A?

b) Geology B?

Draw a bar model to illustrate the mixture:

A	**A**	**B**	**B**	**B**

$\frac{2}{5}$ is A and $\frac{3}{5}$ is B

Terminology

'Write the ratio of *A* to *B*' means write *A* : *B*. If you want students to write the ratio as $\frac{A}{B}$ you need to say 'write the ratio as $\frac{A}{B}$'.

To simplify a ratio, divide all the numbers in the ratio by their highest common factor.

To compare ratios, write them in the form 1 : *n*, or n : 1. This is sometimes called a unit ratio.

A **ratio** compares two quantities, and translates into a statement such as 'for every 3 black there are 2 red'.

A **proportion** compares a part with a whole. A proportion can be given as a fraction or a percentage.

Common error

Students look at 2 : 3 and think the fraction is $\frac{2}{3}$.

3.3 Applying percentage and ratio skills in AS/A level geography

The procedural skills of calculating percentage change and ratios are expected to be embedded throughout topics in both AS and A level. They can also be tested in the exam.

Knowledge of these skills allow students to be more confident with the AO3 skills, which require them to interpret, analyse and evaluate. These could be tested using a resource where students have the opportunity to demonstrate deeper analysis skills.

AO3	Use a variety of relevant quantitative, qualitative and fieldwork skills to: • investigate geographical questions and issues • interpret, analyse and evaluate data and evidence • construct arguments and draw conclusions

Figure 34 shows the results of a global attitude survey about different energy resources.

Statement \ Energy Source	Hydroelectric power (HEP)	Nuclear power	Wind power	Bio-fuels (bioethanol / biodiesel)
1 It is reliable	55	55	40	42
2 It is environmentally friendly	73	22	81	44
3 It is a long-term solution to future energy demands	64	41	68	43

Key: data shown are percentage of people who agreed with the statement

- 66% or more of people agree with the statement
- 46–65% of people agree with the statement
- 26–45% of people agree with the statement
- 25% or fewer people agree with the statement

Figure 34 Results of a global survey of attitudes about different energy sources

Source: Pearson Edexcel A level Geography Unit 3 June 2016

A candidate taking Paper 3 of A level geography will need to interpret unseen data and evidence. This could be tested with an 'analyse' or 'evaluate' question.

Analysis of the attitude survey could apply procedural maths in the following context:

- The ratios about reliability (1) show HEP and Nuclear as even (1:1), but Nuclear vs Wind is in favour of Nuclear (11:8).
- In terms of environmental friendliness (2), there is a 2:1 split with Bio-fuels vs Nuclear.
- There is a close ratio between HEP and Wind (16:17) as long-term solutions to future energy demands (Q3).

Example – drumlin shape survey

These maths application skills might also be useful in the context of the NEA (Non-Examined Assessment). For instance, a student considering a glaciation focus for their Independent Investigation could use the abstract below to devise a topic focus. The relevance of maths and ratios is indicated in blue.

From a systematic programme of drumlin mapping from digital elevation models and satellite images of Britain and Ireland, we used a geographic information system to compile a range of statistics on length L, width W, and elongation ratio E (where E = L/W) for a large sample. Mean L, is found to be 629 m (n = 58,983), mean W is 209 m and mean E is 2.9 (n = 37,043). Most drumlins are between 250 and 1000 metres in length; between 120 and 300 metres in width; and between 1.7 and 4.1 times as long as they are wide. Analysis of such data and plots of drumlin width against length reveals some new insights. All frequency distributions are unimodal from which we infer that the geomorphological label of ‘drumlin’ is fair in that this is a true single population of landforms, rather than an amalgam of different landform types. Drumlin size shows a clear minimum bound of around 100 m (horizontal). Maybe drumlins are generated at many scales and this is the minimum, or this value may be an indication of the fundamental scale of bump generation (‘proto-drumlins’) prior to them growing and elongating. A relationship between drumlin width and length is found (with r2 = 0.48) and that is approximately W = 7 L 1/2 when measured in metres. A surprising and sharply-defined line bounds the data cloud plotted in E–W–L space, and records a scale-dependent maximum elongation limit (approximated by Emax = L1/3, when L measured in metres). For a given length, for some reason as yet unknown, drumlins do not exceed the elongation ratio defined by this scaling law.

Source: http://nora.nerc.ac.uk/6960/

This model could be simplified and a smaller, localised study undertaken using a combination of primary field surveys (e.g. small scale- for one area / drumlin), detailed OS mapping and GIS to establish whether the ratios for example (W, L and E) can be seen in an area that is under investigation (research).

4. Visualise and represent 2-D and 3-D forms, including two-dimensional representations of 3-D objects

Demand

All students learn about simple nets in KS3.

Terminology

- 2-D representations of 3-D shapes include 3-D sketches, accurate 3-D drawings on isometric paper, nets, and plans and elevations.
- A net is the 2-D shape that can be folded up to make a 3-D shape.
- The plan view is the view from above an object (like an OS map). The side elevation is the view from one side and the front elevation is the view from the front. GIS, for example, could then be used to create a cross-section.

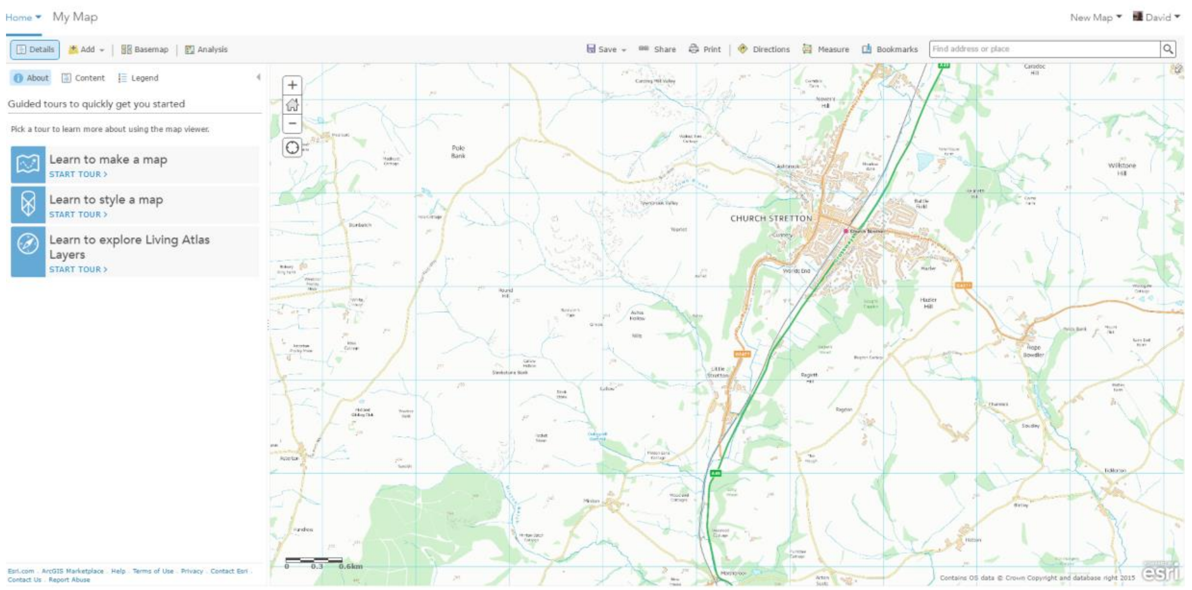
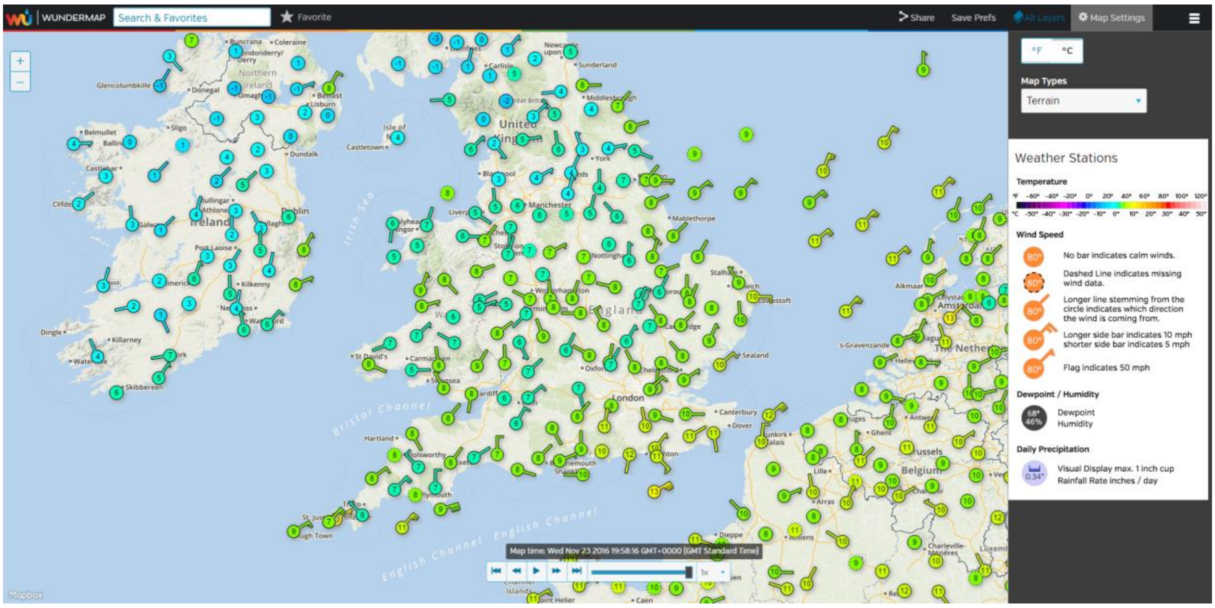


Figure 35 Plan view of an area (from a GIS program – ArcGIS Online)



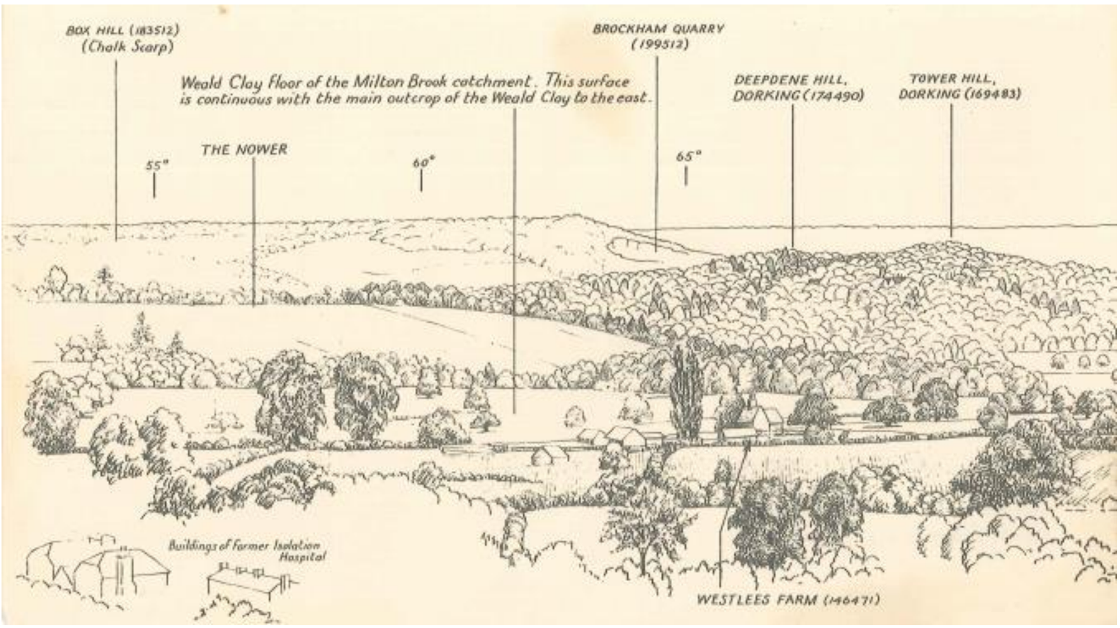
Source: www.wunderground.com/wundermap/

Figure 36 Using GIS to live weather data feeds. This is a number-rich source with historic time series that can be analysed and interpreted.

Sketches and drawings are particular to geography. Sketches might cover field equipment, but are more likely to be at 'landscape' scale.

An historic example is provided in Figure 38, based around the Surrey Hills.

Sketches and drawings should be seen as a vital skill-set, allowing observation, recording and interpretation skills to be nurtured.



Source: Field Studies Council. Hutchings (1955) An Introduction to Geographical Landscape Drawing.

Figure 37 An example of a traditional field sketch

Drawings in geography may also be used to make a political or socio-economic representation. They could be useful in the context of individual literature research and often provide a focus for an idea that can be investigated.

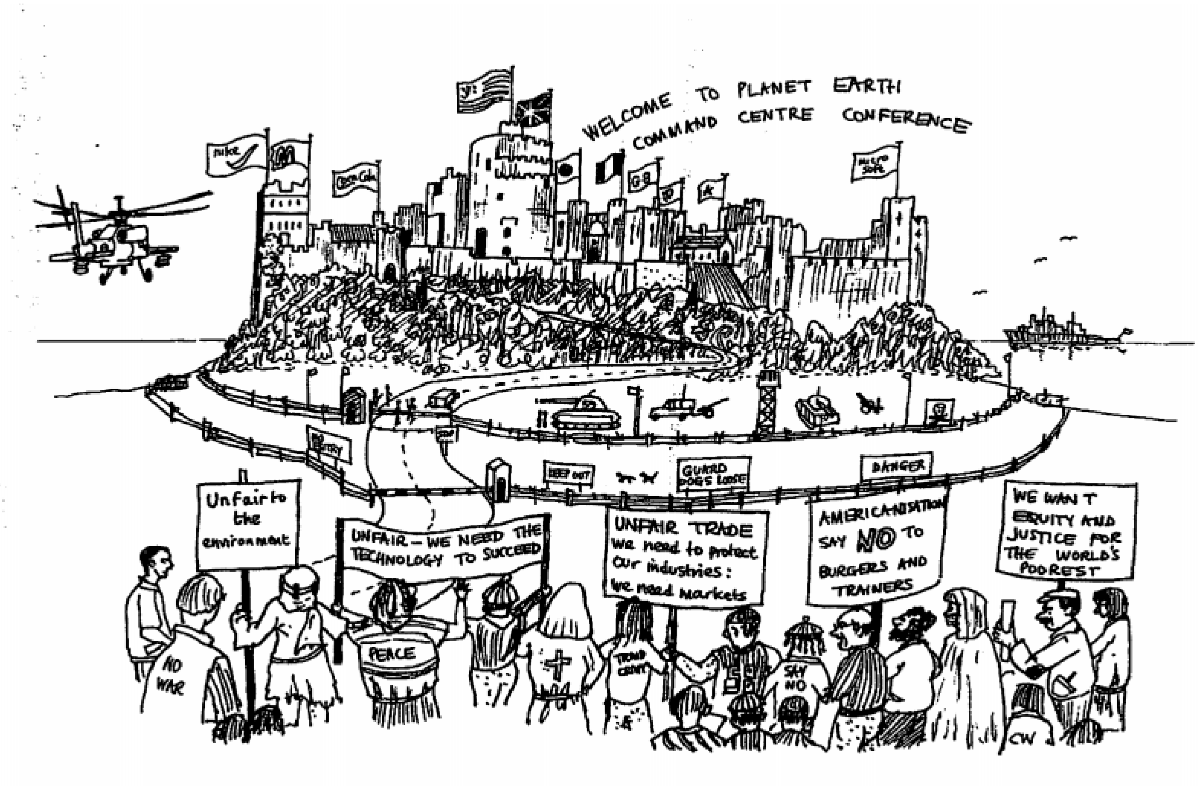


Figure 38 An example of a geographical drawing. Source: Chris Warn.

5. Calculate areas of triangles and rectangles, surface areas and volumes of cubes

Demand

All students learn to calculate the area of a rectangle and triangle in KS3.

Top and middle set maths students use hectares in KS3.

All students calculate surface area and volume of cubes and cuboids in KS3.

Approach

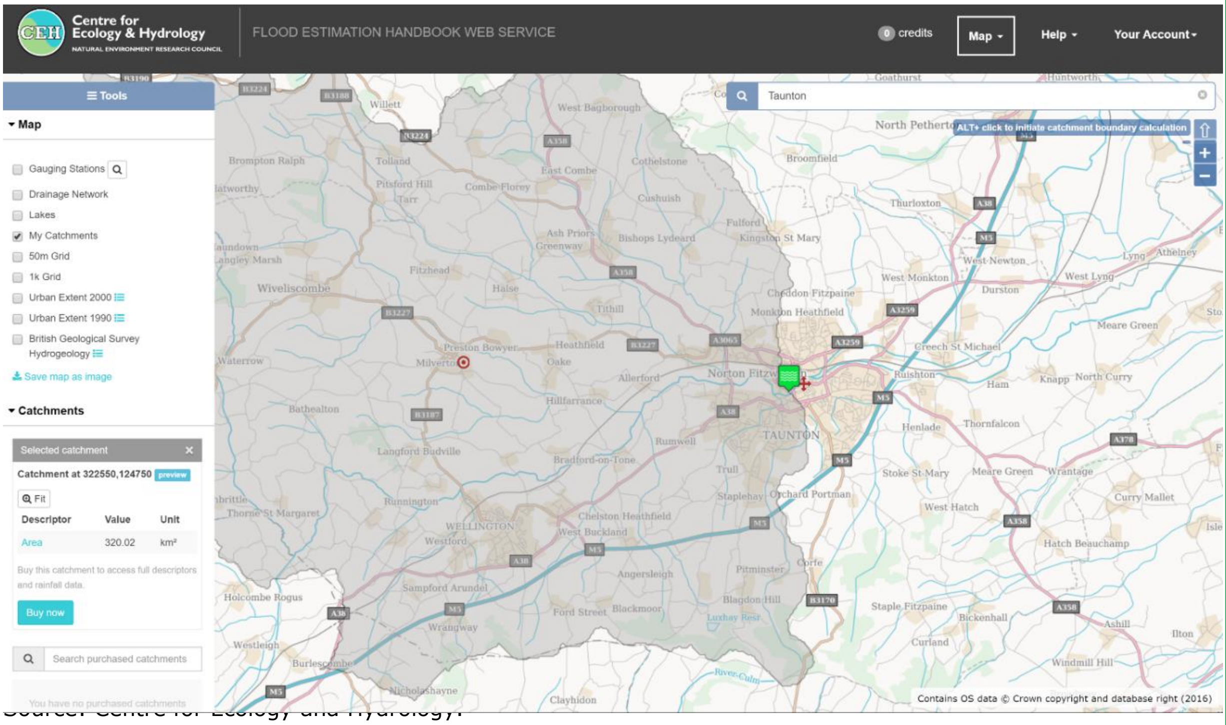
Estimating the area of an irregular shape – counting squares

To calculate the area of an irregular shape, such as a river catchment area, draw round the shape on graph paper that has small squares. Then count the squares inside the area. For squares that cross the perimeter, count those that are more than half-in as whole ones, and don't count those that are more than half-out. This will give you an **estimate** of the area. The smaller the graph squares, the more accurate the estimate.

Remember that area is measured in square units, such as km^2, m^2 or cm^2.

You can use an online free GIS resource to verify and check your answer, such as the Flood Estimation Handbook Service. A free account is required to register: https://fehweb.ceh.ac.uk/GB/map

Of course, other GIS programs could also do this, e.g. ArcGIS Online.



Source: Centre for Ecology and Hydrology.

Figure 39 An example of a river catchment area. Note the area is shown in the bottom left of the web page – 320km^2

Calculate the area of a rectangle (i.e. cross sectional area)

Area of a rectangle = length × width

$A = l \times w$ or $A = lw$

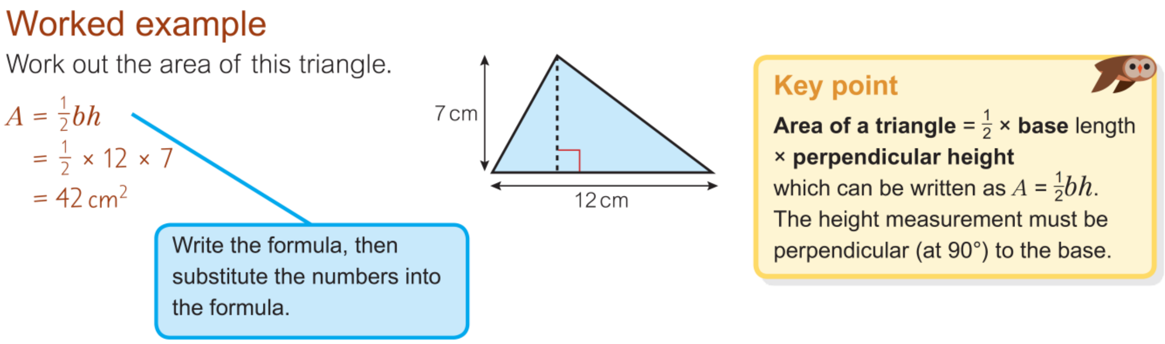
Calculate the area of a triangle

This formula works for all triangles, not just right-angled ones.

h is the perpendicular height.

Area of triangle = $\frac{1}{2}$ × base × height

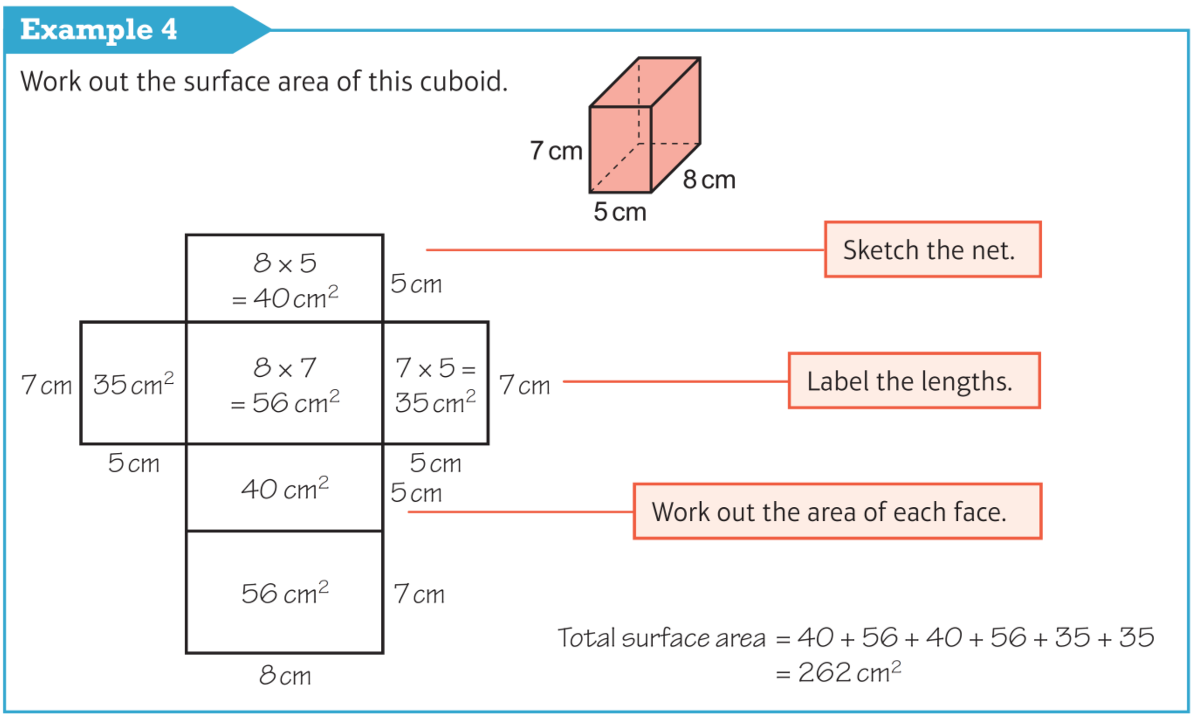
$A = \frac{1}{2}bh$



Source: *KS3 Maths Progress*

In a right-angled triangle, the two sides that meet at the right angle are the base and the height.

Calculate surface area by drawing the net



Source: *Edexcel GCSE (9-1) Mathematics*

Terminology

- Use the mathematical terms rectangle and rectangular (instead of oblong) and rhombus (instead of diamond).
- Use the mathematical term cuboid (instead of box-like).
- Specify the shape of an object when you ask students to calculate the area. If it is a rectangle, say so clearly. Otherwise this reinforces a common misconception that the area of any shape is length × width. For example, ‘Estimate the abundance of foxgloves in an area 60 m long and 10 m wide’ is not accurate enough. Tell them it is a rectangular area.
- The perimeter of a 2-D shape is the distance all around the outside.
- The area of a 2-D shape is the amount of space inside the shape. It is measured in squared units, e.g. mm^2, cm^2, m^2, hectares (1 ha = 10 000 m^2) and km^2. Shifting between scales and areas is important in geography. See the next section.
- You can estimate the area of an irregular shape by drawing around it on cm squared paper and counting the squares.
- If a shape is close to a rectangle, you can estimate the area by approximating it to a rectangle.
- The surface area of a 3-D shape is the total area of all the surfaces added together.
- In maths we use 'area' for 2-D shapes (e.g. a rectangular section of beach) and 'surface area' for 3-D shapes, because it is the areas of all the surfaces added together.
- To calculate the surface area of a cuboid, find the areas of all the faces and add them together.
- The volume of a 3-D shape is the amount of space it takes up. It is measured in cubed units: mm^3, cm^3, m^3.
- Capacity is the amount of liquid a 3-D solid can hold. It is measured in ml or litres.

6. Units and compound units, including conversion between units

Demand

All students meet the prefixes for metric units in GCSE maths. The only ones they are likely to use frequently in maths are kg, km, cm, ml, mm.

All students learn to convert between metric units of area and volume in GCSE maths.

All students learn to convert speeds in m/s to km/h and vice versa in GCSE maths.

Students are not expected to know metric/imperial conversions, or imperial-to-imperial conversions such as lbs to stones, or feet to yards.

Approach

Use arrow diagrams and function machines for simple conversions

'By counting the number in 15 seconds and multiplying by 4'

Rather than present this as a 'magic' formula, it would be good to get students to work out how many 15 seconds there are in a minute, and so what to multiply by.

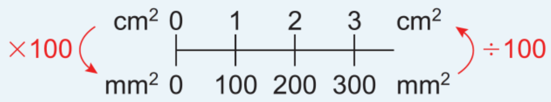
Area conversions

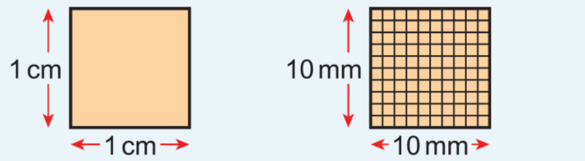
Students may not remember area conversion factors, but will learn in GCSE maths how to work them out, as follows.

You can use a double number line to convert between area measures.

Key point 4

These two squares have the same area.
To convert from cm^2 to mm^2, multiply by 100.
To convert from mm^2 to cm^2, divide by 100.





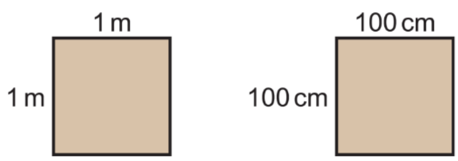
Area = 1 cm × 1 cm
= 1 cm^2

Area = 10 mm × 10 mm
= 100 mm^2

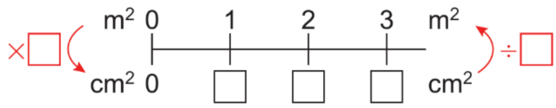
Source: *Edexcel GCSE (9-1) Mathematics*

The question below asks students to convert m^2 to cm^2 and vice versa.

Use these diagrams to help you work out the number of cm^2 in 1 m^2.



Copy and complete the double number line.



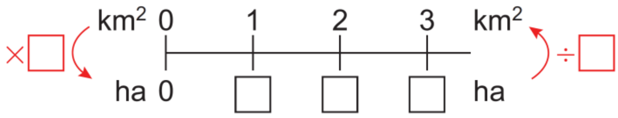
This question shows how to convert hectares to km² and vice versa, something important in the use of OS maps in geography.

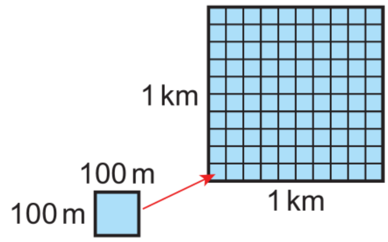
The diagram shows 1 km² divided into 100 m squares.

a What is the area of each 100 m square?

b How many hectares are there in 1 km²?

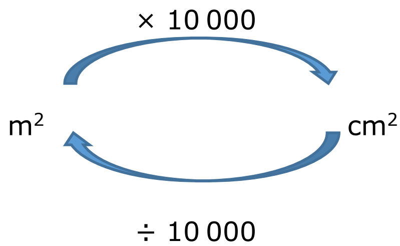
c Copy and complete the double number line.





Note a hectare is a measurement of a unit of area which is 100x100m. There are 100 ha in 1km²

You can use arrow diagrams to help convert between area measures.



In this geography example, the blue grid lines on the maps are 1km apart. Lizard Wood (in the figure) is approximately 1 km², or 100 ha. Dog Wood takes up about 20% of a single square, so would be 20 ha or 200 000 m².

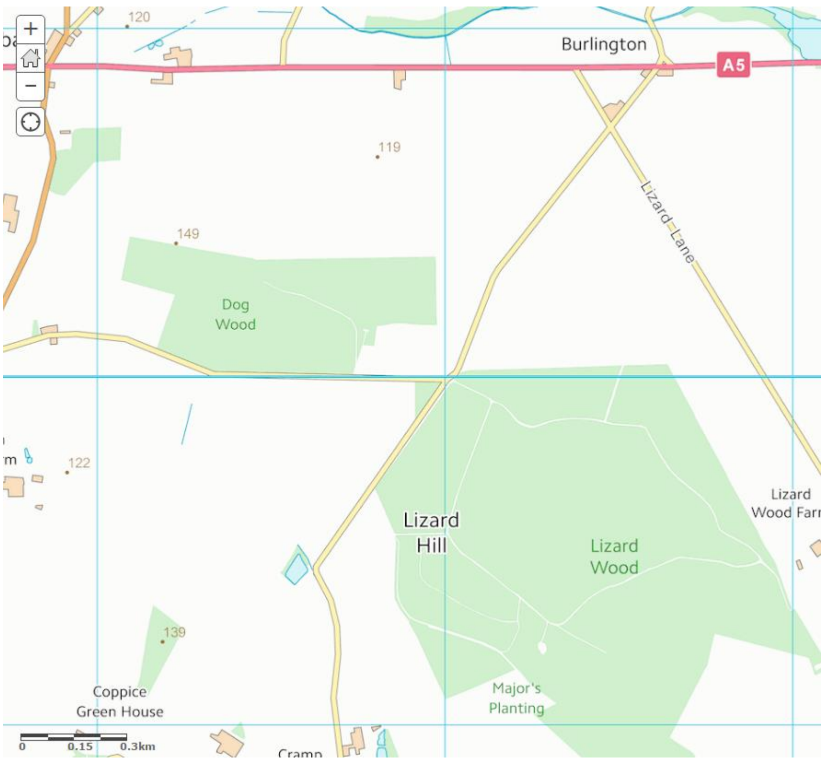


Figure 40 A map showing areas of different land uses. Source: Ordnance Survey

Terminology

- **Inconsistency:** In maths, science and geography we use cm, mm, kg, etc. These are abbreviations, not symbols.

- We use the abbreviations min (not m) for minutes.
- For compound units we use m/s, km/h, g/cm^3, rather than ms^{-1} etc.
- In maths books, where we use 'per' we give a literacy hint, e.g. 8 g/cm^2 means 8 grams in every cm^2.
- Area is measured in squared units: mm^2, cm^2, m^2, hectares (1 ha = 10 000 m^2) and km^2.
- Volume is measured in cubed units: mm^3, cm^3, m^3.
- Capacity is measured in litres and ml. Students do not use dl in maths. Some students may use cl.
- 1 cm^3 = 1 ml. This might be relevant in geography when dealing with fieldwork data on rivers, for example. Conventionally, discharge is described as m^3/sec (m^{3}s^{-1}), but for smaller rivers it may be better to express it as litres/second, or l/sec.
- Some students meet the prefixes for metric units in KS3 maths, e.g. M stands for Mega and means 10^6.
- It is easier to convert measures to the units required before doing a calculation, than to convert the answer into the units required.
- If students are required to convert between metric and imperial units, they should be given the conversion factor. In GCSE maths they are not expected to know metric/imperial equivalents.

Worked example for calculating the dissolved organic carbon in rivers

The following example shows how, through a series of linked calculations and stages, the dissolved organic carbon (DOC) can be estimated as a flux (low) from a catchment system. This exercise illustrates conversion between units.

This example could be used in the context of a piece of coursework which was investigating the daily fluxes.

This is the starting information.

Example data	Value	Units
Catchment area	5	km^{2}
Mean stream velocity	0.7	m s^{-1}
Stream cross-sectional area	0.8	m^{2}
DOC concentration	15	mg l^{-1}

This data comes from an online website to calculate area

Fieldwork is used to measure these variables

This figure is estimated based on the colour 'brownness' of the water

The next parts of the calculation describe how unit transformations are achieved.

Stage	Calculation	Mathematical transformation	Value	Units
1	Discharge (Q)	V (m s^{-1}) × A (m^{2})	0.56	m^{3} s^{-1}
2	DOC concentration	[DOC] (mg l^{-1}) × 1000	15 000	mg m^{3}
3	DOC flux	[DOC] (mg m^{-3}) × Q (m^{3} s^{-1})	8400	mg s^{-1}
4	DOC flux (g)	DOC flux (mg s^{-1}) / 1000	8.4	g s^{-1}
5	DOC flux daily	DOC flux (g s^{-1}) × 86 400	725760	g
6	Daily DOC flux areal	DOC flux daily (g) / catchment area (km^{2})	145152	g km^{2}
7	DOC flux areal	Daily DOC areal (g km^{-2}) / 1 000 000	0.145	t km^{2} d^{-1}

Although this is complex, it is largely "common-sense" and some learners will be able to access these procedural calculations.

The stages are more fully described below.

Stage 1	Calculation of river discharge – velocity (V) × cross-sectional area (A)
Stage 2	DOC – convert from mg per litre to mg per cubic metre (to match the same units as discharge)
Stage 3	DOC flow rate (flux) – multiply the calculation in (1) by (2)
Stage 4	Conversion of flux from mg to g, i.e. 8.4 g of **carbon per second**
Stage 5	Use the figure from (4) and multiply by 86 400, which is the number of seconds in a day. This is the amount of **carbon per day**
Stage 6	This calculation works out the amount of DOC in a geographical area, i.e. g per km^2. So, 725760/ 5 (area) = 145152 g km^2 or145.152 kg per km^2
Stage 7	This final calculation is simply to convert the Stage 6 answer into a more universal figure: tonnes per square km a day. The answer is still the same as in (6), but it is in tonnes, rather than grams

7. Significant figures, decimal places, accuracy

Demand

All students will have learned to round to the nearest whole number and 1, 2 or 3 decimal places (dp) by the end of KS3. They should be able to cope with rounding to more dp as an extension of rounding to 3 dp.

Significant figures

Only Higher tier students learn about upper and lower bounds, and percentage error.

For percentage error they answer questions such as: Given a percentage error of ± 10%, what is the largest/smallest possible value?

Answering questions such as 'What is the percentage error?' for a given value is not covered in the GCSE Maths specification.

Approach

Look at the digit *after the* last one you want to keep. Round up if this digit is 5 or more; round down if it is 4 or less.

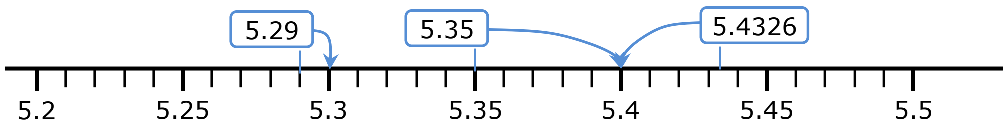
Rounding to 1 dp

5.4|326
less than 5
round down
5.4 (1 d.p.)

5.2|91
5 or more
round up
5.3 (1 d.p.)

5.3|5
5 or more
round up
5.4 (1 d.p.)

On a number line, round to the nearest value with 1 decimal place:



35.42|9
5 or more
round up
35.43 (2 d.p.)

126.37|2
less than 5
round down
126.37 (2 d.p.)

Rounding to 3 dp

0.053|21
less than 5
round down
0.053 (3 d.p.)

11.291|5
5 or more
round up
11.292 (3 d.p.)

Rounding to significant figures

Small numbers

1st significant figure
= 4 ten thousandths

0.000 483

Large numbers

1st significant figure
= 5 billion

518 376 000

Round to 2 significant figures (2 s.f.)

0.00048|3

less than 5
round down

0.00048

51|8 376 000

5 or more
round up

520 000 000

Add zeroes so the 5 is still
in the 'billions' position

Upper and lower bounds calculations

Find the upper and lower bounds of the given values, before doing the calculation.

Terminology

- The number of **decimal places** is the number of digits after the decimal point. So, 10.5219 has 4 decimal places, and 10 has no decimal places.
- In any number the first **significant figure** is the one with the highest place value. It is the first non-zero digit counting from the left.

- **Inconsistency:** Zero is counted as a significant figure if it is between two other non-zero significant figures. Other zeros are place holders – if you took them out, the place value of the other digits would change.

place holders

0 . 0 0 5 0 7 3

1st 2nd 3rd 4th

significant figures

place holders

3 9 0 4 1 2 0 0 0

1st 2nd 3rd 4th 5th 6th

significant figures

- To round a number to a given number of significant figures or decimal places, look at the digit after the last one you need. Round up if the digit is 5 or more, and round down if the digit is 4 or less.

- **Inconsistency:** Rounding numbers reduces accuracy. Your results cannot be more accurate than your starting values. In geography, in calculations your answer cannot have more significant figures than the numbers in the calculation. In maths we tell students not to give more decimal places in the answer than in the calculation, and also to consider if their answers are practical – e.g. could you measure 4.321 cm to that level of accuracy? In geography, where fieldwork equipment *may be* more accurate, the answer to this could be 'yes'.

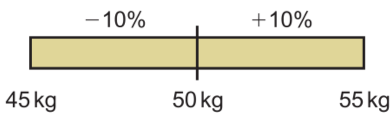
- 8.95 rounded to 1 decimal place is 9.0. You must write the '.0' to show the value in the decimal place.
- When a value is rounded, its true value lies within half a unit either side of the rounded value.
  The values of x that rounds 5.2 to 1 dp are:
  $5.15 \leq x < 5.25$
- Note, this means that the highest value that rounds 5.2 to 1 dp is 5.2499999…
- The upper bound is half a unit greater than the rounded measurement.
  The lower bound is half a unit less than the rounded measurement.

$$\underbrace{12.5}_{\text{lower bound}} \leq x < \underbrace{13.5}_{\text{upper bound}}$$

NB Only Higher tier maths students learn this.

- To determine an appropriate level of accuracy for an answer to a calculation, you can find the upper and lower bounds of the calculation.
  E.g. if upper bound is 28.42896
  and lower bound is 28.42712
  then 28.42 is a suitable level of accuracy.

  NB Only Higher tier maths students learn this.
- A 10% error interval means that a value could be up to 10% larger or smaller than the value given.



- You can write an error interval as an inequality: $45 \leq m \leq 55$ kg.

8. Standard form

Demand

All KS3 students learn to write numbers in index form and use the index laws for multiplication and division.

All students learn the positive and negative powers of 10 in Unit 1 of GCSE maths. Foundation students often find the negative and zero powers difficult to understand/ remember, as they are the only negative and zero powers they use. Higher students use negative and zero indices with a range of numbers so are likely to have a better understanding.

All GCSE students learn to read and write very small and very large numbers in standard form.

In Pearson maths:

- Higher students learn this in the first unit of the GCSE textbook
- Foundation students learn it in Unit 18 of the GCSE textbook, in the spring term of Year 11.

Approach

Calculating powers of 10

Follow a pattern:

$$10^1 = 10$$
$$10^2 = 10 \times 10 = \square$$
$$10^3 = 10 \times 10 \times 10 = \square$$
$$10^4 = 10 \times 10 \times 10 \times 10 = 10\,000$$
$$10^5 = \ldots$$
$$10^6 = \ldots$$

$$10^3 = 1000$$

$$10^2 = 100$$

$$10^1 = 10$$

$$10^{\square} = 1$$

$$10^{-1} = \frac{1}{10} = 0.1$$

$$10^{\square} = \frac{1}{100} = \frac{1}{10^2} = 0.01$$

$$10^{\square} = \square = \square = \square$$

Writing large numbers in standard form

These examples are taken from the *Edexcel GCSE (9-1) Mathematics Foundation* student book.

Example 4

Write 4000 in standard form.

$4000 = 4 \times 1000$

$= 4 \times 10^3$

Write the number as a number between 1 and 10 multiplied by a power of 10.

Write the power of 10 using indices.

Example 5

Write 45 600 in standard form.

$45\,600 = 4.56 \times 10^4$

4.56 lies between 1 and 10.
Multiply by the power of 10 needed to give the original number.
4 5 6 0 0

Writing small numbers in standard form

Example 6

Write 0.000 05 in standard form.

$0.000\,05 = 5 \times 0.000\,01$

$= 5 \times 10^{-5}$

Write the number as a number between 1 and 9 multiplied by a power of 10.

Key point 7

To write a small number in standard form:
- Place the decimal point after the first non-zero digit.
- How many places has this moved the digit? This is the negative power of 10.

Example 7

Write 0.003 52 in standard form.

$0.00352 = 3.52 \times 10^{-3}$

3.52 lies between 1 and 10.
Multiply by the power of 10 needed to give the original number.
0 . 0 0 3 5 2

Calculating with numbers in standard form

Multiplication and division

Example 3

Work out $(5 \times 10^3) \times (7 \times 10^6)$

$5 \times 7 \times 10^3 \times 10^6$ — Rewrite the multiplication grouping the numbers and the powers.

35×10^9 — Simplify using multiplication and the index law $x^m \times x^n = x^{m+n}$. This is not in standard form because 35 is not between 1 and 10.

$35 = 3.5 \times 10^1$ — Write 35 in standard form.

$35 \times 10^9 = 3.5 \times 10^1 \times 10^9 = 3.5 \times 10^{10}$ — Work out the final answer.

Work out $\dfrac{2.4 \times 10^5}{3 \times 10^2}$

Divide 2.4 by 3.
Use the index law $x^m \div x^n = x^{m-n}$ to divide 10^5 by 10^2.

$= 0.8 \times 10^3$
$= 8 \times 10^2$

Addition and subtraction

Write numbers in decimal form before adding and subtracting.

Write the answer in standard form.

Work out $3.6 \times 10^2 + 4.1 \times 10^{-2}$
$= 360 + 0.041$
$= 360.041$
$= 3.60041 \times 10^2$

Work out $2.5 \times 10^6 - 4 \times 10^4$

$$\begin{array}{r} 2\,500\,000 \\ -\ \underline{\quad 40\,000} \\ 2\,460\,000 \end{array}$$

$= 2.46 \times 10^6$

Terminology

- Any number can be raised to a power or index. The power or index tells you how many times the number is multiplied by itself. $3^4 = 3 \times 3 \times 3 \times 3$
- We read 3^4 as '3 to the power 4'.
- Some calculators have a power or index key. In maths we do not tell students which key presses to use, as calculators vary. Instead we would say 'Make sure you know how to input numbers raised to a power on your calculator'.
- Any number raised to the power zero = 1.
- The index laws:
    - To multiply powers of the same number, add the indices.
    - To divide powers of the same number, subtract the indices.
- Some of the powers of 10:

10^{-4}	10^{-3}	10^{-2}	10^{-1}	10^0	10	10^2	10^3	10^4
0.0001 or $\frac{1}{10\,000}$	0.001 or $\frac{1}{1000}$	0.01 or $\frac{1}{100}$	0.1 or $\frac{1}{10}$	1	10	100	1000	10 000

- Standard form is a way of writing very large or very small numbers as a number between 1 and 10 multiplied by a power of 10.
  $A \times 10^n$ where A is between 1 and 10 and n is the power of 10

- **Inconsistency:** When writing numbers in standard form, do not talk about 'moving the decimal point'. The position of the decimal point remains fixed. Multiplying by a power of 10 moves digits places to the left and dividing by a power of 10 moves digits places to the right.

- On some calculators you can enter numbers in standard form, or answers may be given in standard form. In maths we do not tell students which key presses to use, as calculators vary. Instead we would say 'Make sure you know how to enter and read numbers in standard form on your calculator'.

- Geographers may need to use the idea of standard form when dealing with big data sets, e.g. country populations. For some project work and independent study it would be useful to know how to manipulate data in a standard form.

9. Descriptive statistics – average, mean, quartiles/percentiles and standard deviation

Demand

Normal distribution is not covered in maths GCSE.

Foundation level students do not use quartiles and interquartile range.

Higher level students learn quartiles and interquartile range in year 11.

Percentiles are not on the GCSE maths specification.

Only Higher level students find means and medians from bar charts and histograms in GCSE maths.

Approach

Calculating the range

From a small data set

Range = largest value minus smallest value

For example, for this data: 2, 2, 5, 7, 2, 4, 6, 9

the range is 9 – 2 = 7

From a frequency table

Range = largest value minus smallest value

Number of cars per household	Frequency
1	2
2	15
3	6

Range = 3 – 1 = 2

NB it is the range of the data values, not of the frequencies.

From a grouped frequency table

An estimate of the range is the largest possible value minus the smallest possible value.

Worked example

In a survey, people were asked their age. The table shows the results.

Age, a (years)	Frequency
$0 \leqslant a < 10$	12
$10 \leqslant a < 20$	15
$20 \leqslant a < 30$	2
$30 \leqslant a < 40$	11

Work out an estimate for the range of ages.

From the frequency table, the smallest possible age is 0 years.

The largest possible age is 40 years.

So an estimate of the range is 40 − 0 = 40 years.

Finding the mode

From a data set

For the data set 2, 2, 5, 7, 2, 4, 6, 9, the mode is 2.

For the data set 1, 1, 3, 4, 2, 5, 3, 3, 2, 2, 1 the modes are 1 and 3.

From a frequency table

For this data, the mode is 2 cars per household.

Number of cars per household	Frequency
1	2
2	15
3	6

Common error

Students may give '15' as the mode (the highest frequency), rather than 2, which is the number of household cars with the greatest frequency.

From a grouped frequency table

For this data the modal class is $10 \leq a < 20$

Age, a (years)	Frequency
$0 \leqslant a < 10$	12
$10 \leqslant a < 20$	15
$20 \leqslant a < 30$	2
$30 \leqslant a < 40$	11

Finding the median and quartiles

First write the data in order.

0	2	4	5	7	8	10	12	12	12	13	14	14	15	16
			↑ The lowest quartile is the 4th value				↑ The median is the 8th value				↑ The upper quartile is the 12th value			

Common error

Students may not order the data before finding the median and quartiles.

The information below has been taken from a frequency table of fieldwork data from the coast (based on length of long axis).

Classification of sediment shape	Frequency
1 (Very rounded)	4
2 (Rounded)	8
3 (Poorly rounded)	6
4 (Subangular)	2
5 (Angular)	0
6 (Very Angular)	0
Total	20

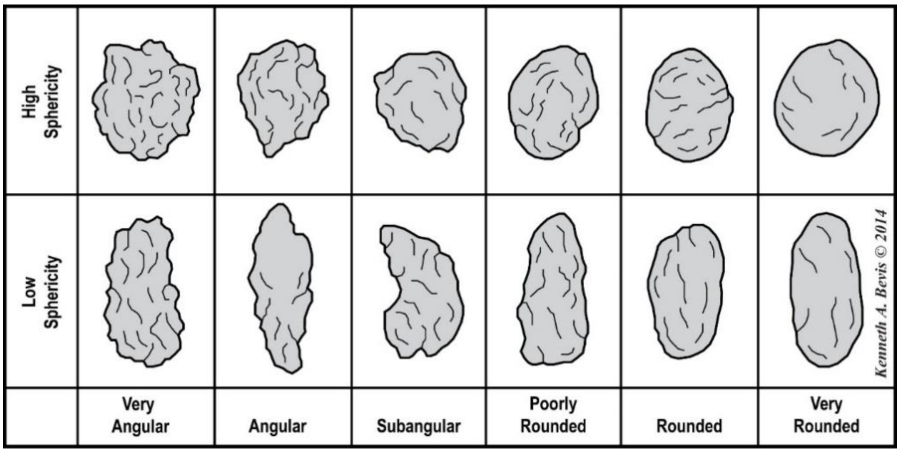


Figure 41: A table showing grain shape; sedimentary particles become more rounded and spherical as they are transported

Source: http://intheplaygroundofgiants.com/geology-of-the-grand-canyon-region/the-geology-of-sedimentary-rocks/

There are 20 pieces of data in the table.

The median is the $\frac{20+1}{2}$ = 10.5th data item (i.e. between the 10th and 11th items).

Classification of sediment shape	Frequency	
1	4	4
2	8	4 + 8 = 12
3	6	
4	2	
Total	20	

Add up the frequencies to find the 10th and 11th data items.

The 10th and 11th items are both shape 2 examples, so the median is shape classification 2.

Calculating the quartiles from a frequency table

The lower quartile is the $\frac{20+1}{4}$ = 5.25th data item (i.e. between the 5th and 6th items). The upper quartile is the $\frac{3(20+1)}{4}$ = 15.75th data item (i.e. between the 15th and 16th items).

Classification of sediment shape	Frequency	
1	4	4
2	8	4 + 8 = 12
3	6	12 + 6 = 18
4	2	
Total	20	

Add up the frequencies to find the 5th and 6th data items and the 15th and 16th data items.

The 5th and 6th items are both classification 2, so the lower quartile is 2.

The 15th and 16th items are both classification 3, so the upper quartile is 3.

Finding the interval containing the median from a grouped frequency table

Age, a (years)	Frequency
$0 \leqslant a < 10$	12
$10 \leqslant a < 20$	15
$20 \leqslant a < 30$	2
$30 \leqslant a < 40$	11

12
12 + 15 = 27

Add up the frequencies to find the 20th and 21st data items.

Total frequency = 40

Median = $\frac{40+1}{2}$ = 20.5th data item.

The 20th and 21st data items are in the interval $10 \leq a < 20$.

Calculating the mean of a small data set

Example 1

Work out the mean of 3, 6, 7, 7 and 8.

$3 + 6 + 7 + 7 + 8 = 31$

Add the values first to find the total.

$\frac{31}{5} = 6.2$

There are 5 values, so divide the total by 5.

The mean is 6.2

Source: *Edexcel Mathematics (9-1) Foundation*

When using a calculator to calculate a mean, students may add the numbers and not press = before dividing, which will give an incorrect value.

Calculating the mean from a frequency table

Worked example

Jack asked students in his class how many pets they had.
Here are his results. Work out the mean.

Number of pets	Frequency	Total number of pets
0	7	$0 \times 7 = 0$
1	8	$1 \times 8 = 8$
2	6	$2 \times 6 = 12$
3	3	$3 \times 3 = 9$
4	1	$4 \times 1 = 4$
Total	25	33

Add a column to the table to work out the total numbers of pets.

Work out the total frequency (number of people in the survey) and the total number of pets.

mean = $33 \div 25 = 1.32$

mean = total number of pets ÷ number of people

Source: *KS3 Maths Progress*

Calculating an estimate of the mean from a grouped frequency table

Worked example

In a survey, people were asked their age. The table shows the results.

Age, a (years)	Frequency
$0 \leqslant a < 10$	12
$10 \leqslant a < 20$	15
$20 \leqslant a < 30$	2
$30 \leqslant a < 40$	11

Calculate an estimate for the mean age.

Age, a (years)	Frequency	Midpoint of class	Midpoint × Frequency
$0 \leqslant a < 10$	12	$\frac{0 + 10}{2} = 5$	$5 \times 12 = 60$
$10 \leqslant a < 20$	15	$\frac{10 + 20}{2} = 15$	$15 \times 15 = 225$
$20 \leqslant a < 30$	2	25	$25 \times 2 = 50$
$30 \leqslant a < 40$	11	35	$35 \times 11 = 385$
Total	40		720

Add a column to calculate an estimate of the total age for each class.

Calculate the total number of people in the survey and the sum of their ages.

mean = sum of ages ÷ total number of people

$= \frac{720}{40}$

$= 18$

Source: *KS3 Maths Progress*

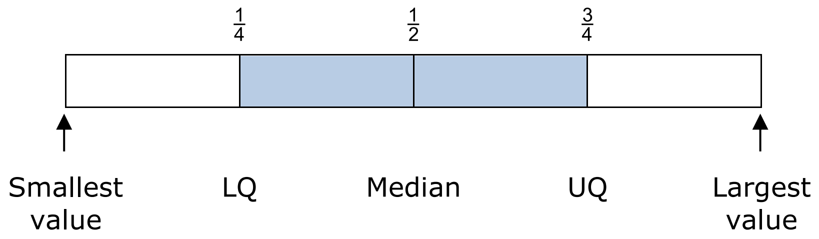
Calculating means and medians from bar charts and histograms

Make a frequency table for the bar chart or histogram, and use the appropriate method shown above.

Terminology

- Mean, median and mode are all averages. Usually when someone says 'average' they are talking about the mean. In geography, students need to be aware of the importance of this term to describe data, as well as its advantages and disadvantages. See the end of this section.
- The range is a measure of spread. It is calculated as largest value – smallest value. Note that the range is a single number; not, for example, 3-12, i.e. two numbers separated by a hyphen. In a maths question we would say 'Work out' the range – you need to do a calculation to find it.
- A larger range means the data is less consistent. A smaller range means the data is more consistent.
- You can estimate the range from a grouped frequency table, as largest possible value minus smallest possible value.
- Mean = $\frac{\text{total of all the values}}{\text{number of values}}$
- You can calculate the mean from an ungrouped frequency table. For a grouped frequency table, you can calculate an estimate of the mean (because you use the midpoint of each group as an estimate of the data values in that group). Please word such questions as 'Calculate an estimate of the mean'.
- The mode is the most common value. In a frequency table, this is the value with the highest frequency. The mode is one of the data values. A set of data can have more than one mode. For grouped data, the modal class is the class interval with the highest frequency.
- The median is the middle value when the data is written in order. It may not be one of the data values (e.g. it could be halfway between two values).
- For an ordered set of data with an even number of values, the median is the mean of the two middle values (which is the same as the value midway between them).
- For a set of n items of data, the median is the $\frac{n+1}{2}$th data item. When n is very large, you can use the $\frac{n}{2}$th data item.
- If you have an anomalous value (sometimes called an outlier in maths – see page 24), i.e. one that is likely to have been a recording error, you can ignore this when calculating the mean. In geography we sometimes think anomalies are incorrect, especially when linked to fieldwork. They may not be – they could be true values that are simply unexpected, and should still be included in the calculation of the mean.
- For a set of ordered data, the median is the value halfway through the data. The lower quartile is the value one quarter of the way into the data set. The upper quartile is the value three quarters of the way into the data set.
- For a set of n items of data, the lower quartile is the $\frac{n+1}{4}$th data item and the upper quartile is the $\frac{3(n+1)}{4}$th data item. When n is very large, you can use the $\frac{3n}{4}$th data item.

- The interquartile range is value calculated by: upper quartile minus lower quartile. It shows how spread out the middle 50% of the data is.



- The 50th percentile is the median. The 10th percentile is the value 10% of the way into the data set, when the data is in order.
- NB Students do not learn about percentiles in GCSE maths.
- To describe a set of data, give at least one average and a measure of spread.
- To compare two sets of data, compare one average and one measure of spread.

Complexity of 'averages' in geography

Example question

Explain one reason why sometimes it is better to use median rather than mean when describing centrality in a data set. (2 marks)

In the case of this type of question in geography, we can use some household income data from 2016 to help students visually understand the idea.

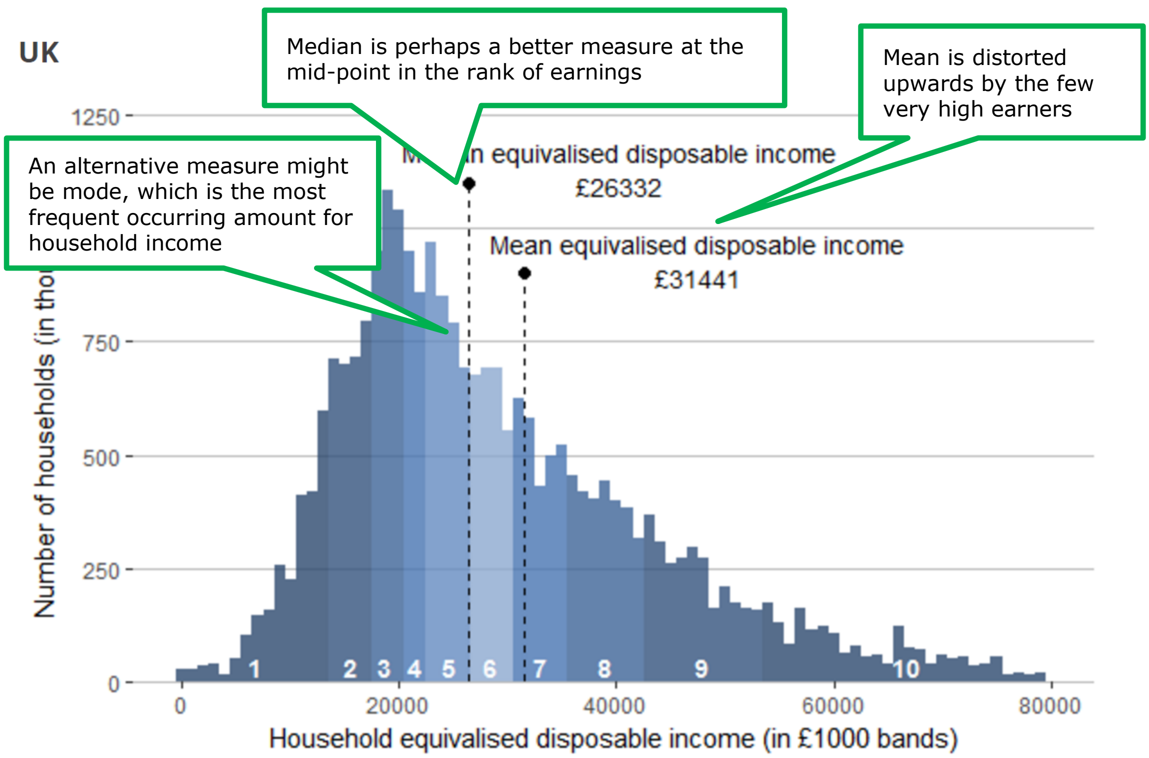


Figure 42: A dispersion diagram showing household disposable income

Source:https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/bulletins/theeffectsoftaxesandbenefitsonhouseholdincome/financialyearending2016

So one explanation for this question might be:

Arithmetic averages (means) are often influenced by outliers, especially in data sets that are not normally distributed and have many values at the extremes. Using median, rather than mean, would minimise the impacts of the extreme values and give a truer indication of the average.

Standard deviation

Standard deviation measures how far data is from the mean. It is based around a calculation of variance, or spread; the average of the squared differences from the mean.

The procedure requires the tabulation of results, and then each individual result (x) to be squared (x^2). Then calculate the $\sum x$ and $\sum x^2$. There are different formulas, but Edexcel will use the one below:

$$\textbf{SD} = \sqrt{\frac{\sum (x-\bar{x})^2}{n}}$$

n = number of observations / data points

$\sum$ = sum of

x = each value in the data set

$\bar{x}$ = mean of the values in the data set

Note this formula will be given in an AS / A level exam, so students would not be expected to memorise the formula.

For the purposes of the NEA, students should use a spreadsheet to calculate standard deviation rather than working it out by hand.

Below are the stages for the SD calculation:

Step 1: Find the mean.

Step 2: For each data point, find the square of its distance to the mean.

Step 3: Sum the values from Step 2.

Step 4: Divide by the number of data points.

Step 5: Take the square root.

Below is a worked example of a question, showing a sequence and then illustrative mark tariff.

The table shows the total footfall data for a small independent high street retailer over the course of a year. Figures are totals per month, rounded to the nearest thousand people.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Footfall	2	4	5	4	4	7	8	6	4	9	9	10

Calculate the mean monthly footfall. (1 mark)

Calculate the standard deviation for the data in the table. Show your working. (4 marks)

Here are the answers.

*The mean is: 6 (This is **Step 1**)*

The working for the standard deviation question could be set out as follows.

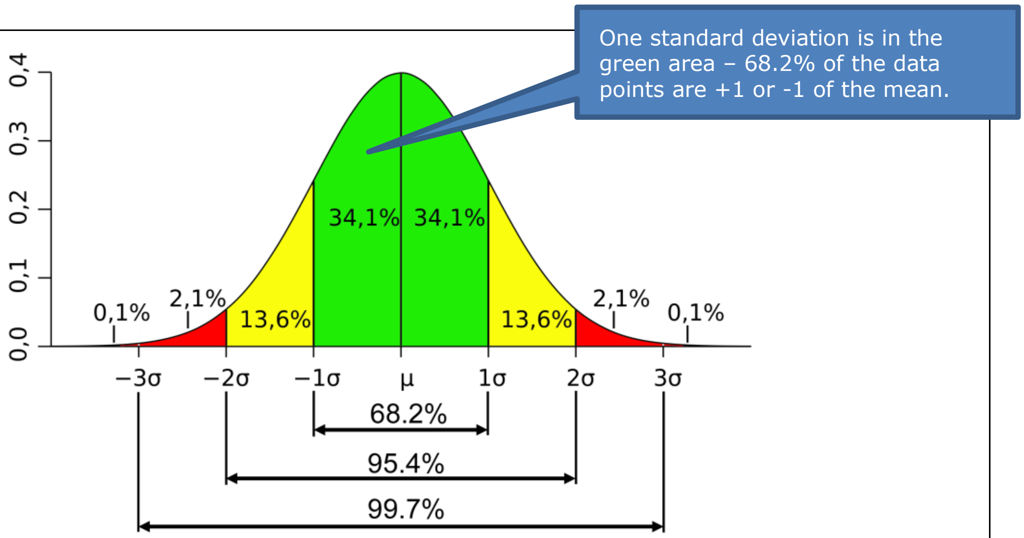
Month	**Data (x)**	**Difference (d) (to mean) ($x - \bar{x}$)**	**d²** **($x - \bar{x}$)²**
Jan	2	2-6 = -4	16
Feb	4	4-6 = -2	4
Mar	5	5-6 = -1	1
Apr	4	4-6 = -2	4
May	4	4-6 = -2	4
Jun	7	7-6 = 1	1
Jul	8	8-6 = 2	4
Aug	6	6-6 = 0	0
Sep	4	4-6 = -2	4
Oct	9	9-6 = 3	9
Nov	9	9-6 = 3	9
Dec	10	10-6 = 4	16
	72		

*Sum the squares of the distances (**Step 3**)= $\sum d^2$ = 72 (1 mark)*

*Divide by the number of data points (**Step 4**) √(72 / 12) (2 marks)*

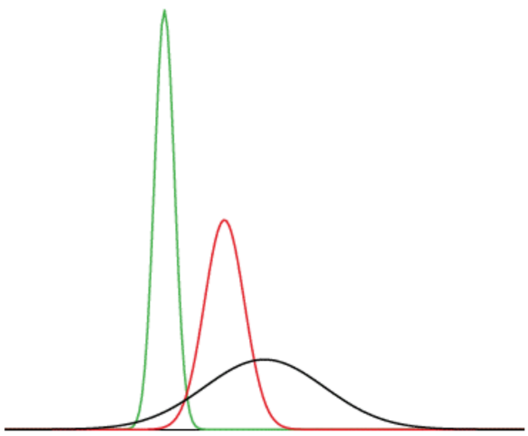
*Take the square root (**Step 5**) The standard deviation is: 2.45 (1 mark)*

It might be useful to visualise the standard deviation by reference to normal distributions.



Normal distributions, or bell curves, as they are sometimes known, are important in geography. Many systems both natural and physical have normalised distribution, e.g. the household earning data that we have already seen in this section. The mean is in the middle, with an equal number of smaller and larger values either side of it.

Normal distributions are often shown as in the example illustrated above, or the three below made up of green, red and black lines. In these examples the standard deviations and frequency data are different, hence the different shapes of the profiles.



Examples of different normal distributions, with different shapes.

10. Inferential statistical tests: t-test, Spearman's Rank, Chi-squared

Demand

Inferential statistical tests are not covered in GCSE maths.

Students should be able to recognise correlation from scatter plots and to extrapolate trends.

Classifying types of statistics

In AS and AS/A level geography, the DfE criteria refers to descriptive statistics, inferential statistics and relational statistics. The table below provides summary classifications.

	Inferential – uses hypothesis testing and sample populations			Descriptive
	Difference	**Relational**	**Categorical data**	
Overview	Looking for differences between populations	Correlations and regression are used to measure the strength of an observed relationship	Using data that belongs to one of a number of categories Distribution of data might be unknown	Mainly to do with summarising data to identify patterns Includes measures of central tendency as well as measures of spread (see Section 9)
Quantitative technique specified (p.91 AS/A level specification)	t-test (assuming the data are normally distributed)	Spearman's Rank	Chi-squared	Modes, means, medians, inter-quartiles, standard deviation, ranges
Alternative and other tests that may be relevant	Mann-Whitney U test (can be used when the data are not normally distributed). Also useful when using counts for data collection.	Calculate line of best fit Pearson product-moment	n/a	Can also include measures of skewness and kurtosis

Inferential statistics are normally used in fieldwork since they allow the data to be generalised to the larger population. This is because fieldwork relies on population samples that infer what the entire population is assumed to be like. This sampling approach is discussed in the Edexcel / Pearson AS and A Level Fieldwork Planner and Guide (2017).

Independent vs dependent variables

- Independent variables are causal variables. They are variables the investigator manipulates (i.e. changes) – assumed to have a direct effect on the dependent variable. An example would be the frequency of wind (number of days) from a particular direction.
- Dependent variables are effect variables. They are variables the investigator measures, after making changes to the independent variable that are assumed to affect the dependent variable. An example might be number of days (frequency) of driving rain. In this instance this variable is dependent on the independent variable.

Note that in any investigative work there might also be extraneous variables, which can affect the results. These may not be directly linked to the dependent variables.

Exploring meaning: null hypotheses, confidence limits and significance levels

Null hypothesis	• This is opposite to a hypothesis (which is an idea or tentative theory). Inferential statistical tests aim to disprove the null hypothesis, thereby accepting the (alternative) hypothesis. • The null hypothesis is usually the 'boring case'… there is no difference/no association/no correlation. • The alternative hypothesis is usually what the evidence might be indicating… there is a difference / association / correlation.
Degrees of freedom	• In carrying out the various hypothesis tests, you will come across the term 'degrees of freedom'. This is something that affects the sample size in your test. Usually it makes the sample size slightly smaller. You need to know the number of degrees of freedom in order to use the statistical tables.
Significance levels and confidence limits	• The term significant has a precise mathematical definition – it concerns the reliability of the data expressed at a percentage value. • For example, if we say that the information/results are significant at 95% level, then this means that only 5 times out of 100 would this outcome occur by chance. 99% means only a 1 in 100 chance. • Significance levels are indicated in published tables, usually 0.05 and 0.01 respectively (5% and 1%). A

	result significant at 0.01 means that there is only a 1% chance the null hypothesis is correct, so the alternative hypothesis is accepted. Confidence levels are closely related to significance levels.

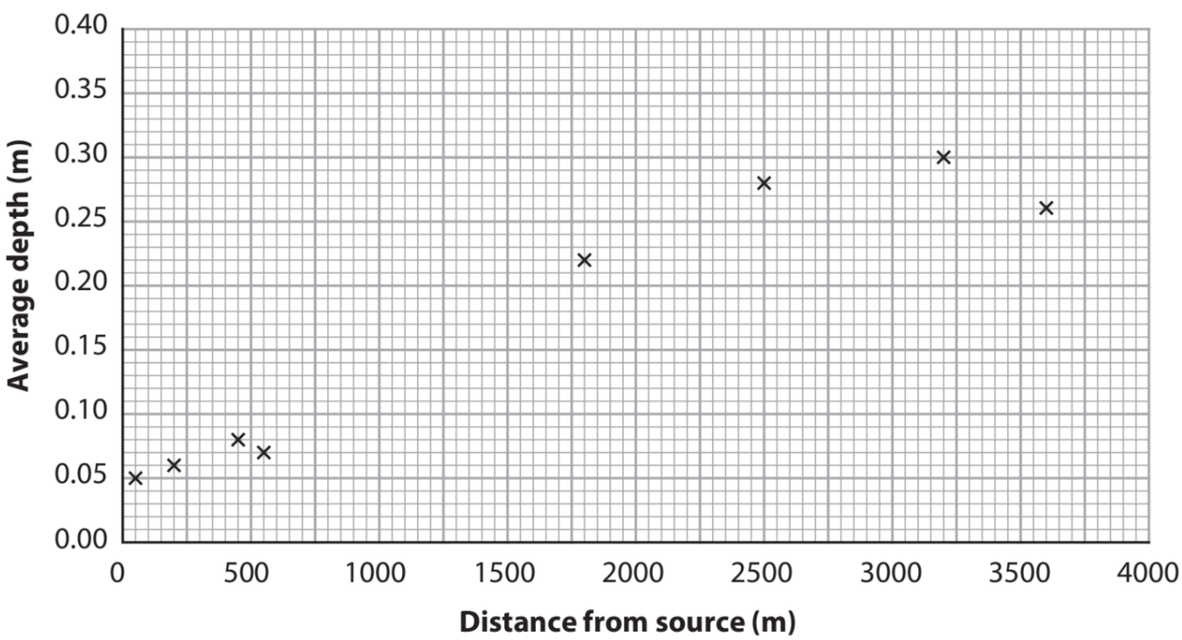
Test for relationships – Spearman's Rank

Spearman's Rank is used to test for a relationship between two variables. A coefficient is calculated (R_s) which determines the strength of the observed relationship.

The simplest way to illustrate a relationship between two variables is first to construct a scatter graph (see Section 2). This may indicate possible relationships, but will not give an indication of the strength – hence the need to use a relationship test to determine a precise measure. The R_s value lies between -1 and +1.

- A positive relationship means that as variable X increases, so does variable Y.
- A negative relationship indicates that as one variable increases, the other decreases.

The scatter graph below, for example, shows a positive correlation between depth (m) and distance from a river's source (m). There are 8 data points, which is generally regarded as the minimum for a Spearman's Rank test to be valid. As it *looks like* there might be a reasonable correlation, it makes sense to further test the strength of the possible relationship using a statistical test such as Spearman's Rank.



Spearman's makes use of ranked ordinal data and is especially useful as a good general assessment of a relationship. In order to perform the procedure, the two sets of data are put into a table and then ranked from highest to lowest.

The formula is given,

$$R = 1 - \frac{6\sum d^2}{n^3 - n}$$

where:

n = the number of observed data pairs

$\sum d^2$ = sum of (d) which is the differences in the rank

As with all the inferential tests, a null and alternative hypothesis are required. The test results are compared to a critical value in the statistical table, which then determines whether or not the null hypothesis can be accepted or rejected.

Note this formula will be given in an AS / A level exam, so students would not be expected to memorise the formula.

Below is an example of how Spearman's Rank is used within an Edexcel/Pearson AS/A Level exam (Paper 1 – Sample Assessment).

Coastal districts in Japan	Deaths	Rank	Tsunami wave height (m)	Rank	d	d^2
Ishinomaki	3735	1	7.6	4	3	9
Rikuzentakata	1846	2	8	3	1	1
Kesennuma	1356	3	7.2	6	3	9
Otsuchi	1286	4	8.1	2	2	4
Higashimat	1105	5	7.3	5	0	0
Kamaishi	1047	6	4.1	9	3	9
Natori	966	7	6.3	7	0	0
Onagawa	915	8	3.4	10	2	4
Minamis	845	9	5.1	8	1	1
Soma	458	10	9.3	1	9	81
					Σ	

Figure 1

Deaths and tsunami wave height resulting from the 2011 Tohoku tsunami in 10 coastal districts in Japan

In this instance, the candidate is required to calculate the Σd^2 and then apply the formula, which is provided as part of the question. In this question, there are 3 marks available for completing the calculation.

The final part of the question requires candidates to link their R_s value (the test statistic) to a critical value/confidence limit and level. Then they are required to comment on whether they should accept the null and (alternative) hypothesis. Note that candidates are provided with all the information to do this within the exam.

Confidence level	**0.10 (90% significance)**	**0.05 (95% significance)**	**0.01 (99% significance)**
Critical value of Spearman's rank R value	0.48	0.6	0.78

Null Hypothesis: There is no significant relationship between the tsunami wave height and the number of deaths in coastal districts.
Hypothesis: There is a significant relationship between the tsunami wave height and the number of deaths in coastal districts.

Using the Spearman's rank correlation R value calculated in part (i), state which hypothesis can be accepted.

(1)

....

In the example above, the mark scheme shows:

Award 1 mark for the sum of d^2 column (Σ) = 118

Award 1 mark for the correct working of equation:

$$1 - \frac{6 \times 118}{10^3 - 10} \quad \text{or} \quad 1 - \frac{708}{990} = R$$

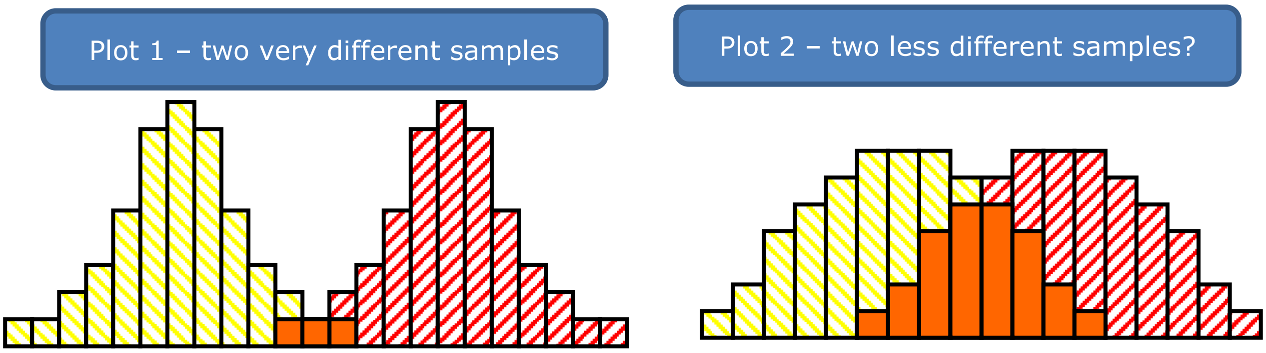
Award 1 mark for answers that round to R = 0.28
OR
Award 1 mark for the correct value of R alone (0.2848).

Award 1 mark for accept null hypothesis as R value is less than critical value at 0.1 confidence level

Remember that correlation does not mean causation (a difficult concept for many students to grasp). Here is one example of a discussion about correlation and causation: www.khanacademy.org/math/probability/scatterplots-a1/creating-interpreting-scatterplots/v/correlation-and-causality

Comparing two sets of data – the t-test

In the two plots below, which show two sets of data on each plot, there are obvious differences in frequency distributions of the two sets of data. In Plot 2 there is much more overlap, so it is harder to establish, at least from a statistical point of view, whether the two populations are really different or distinct.



The (student's) t-test is a test of the difference between two samples. It measures the degree of overlap between the two data sets. *The formula for a t-test is provided in questions (and see below)*.

Calculation of a t-test is arduous by hand. It's much easier to use a calculator, or even better to use a spreadsheet. There are several online calculators, e.g. http://ncalculators.com/statistics/t-test-calculator.htm

This would be good practice in the context of using this test in relation to the NEA (Unit 4).

This is the formula for a t-test:

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{{S_1}^2}{N_1} + \frac{{S_2}^2}{N_2}}}$$

There are other variations of this equation as well,

where

$\bar{x}_1$ is the mean of first data set

$\bar{x}_2$ is the mean of first data set

S_1^2 is the standard deviation of first data set

S_2^2 is the standard deviation of first data set

N_1 is the number of elements in the first data set

N_2 is the number of elements in the first data set

- The top part of the equation compares the means of the two sets of data. This is the first factor that contributes to the overlap between the two sets of data.
- The bottom part of the formula includes measurements of the variability of the data (S_1^2 and S_2^2). This is the other factor that contributes to the amount of overlap.

Like the other inferential tests, a table needs to be set up to perform the calculation.

Part of an example table is provided below, including some data.

Measurement / observation	Location / Site 1		Location / Site 2	
n	X_1	X_1^2	X_2	X_2^2
1	4.5	20.25	4.5	20.25
2	5.0	25.00	3.1	9.61
3	4.5	20.25	4.1	16.81
4	4.4	19.36	3.1	9.61

The basic steps to calculate *t* are the same as those to calculate standard deviation (page 66).

There are several stages in the calculation:

1. Square the individual values of x_1 and x_2 (as in the table).
2. Calculate the sums (∑) of all the x values in the 4 columns in the table.
3. Calculate the means for x_1 and x_2. Do this by dividing ∑ by the number of measurements/observations.
4. Calculate the variances s_1^2 and s_2^2. See the standard deviation calculation on page 66.
5. Substitute the mean and variances from steps (3) and (4) into the formula for *t*.

- If the two sets of data have widely separated means and small variances (i.e. tightly clustered data), they will have little overlap and a **big value of *t***. This would be typical of Plot 1.

- If the two sets of data have means that are close together and large variances (widely spread), they will have a lot of overlap and a **small value of *t*** – like Plot 2.

As in the other inferential tests, critical values are needed to see the 'cut-off' points, so that the hypothesis can be correctly interpreted. For this test there will also need to be consideration of the **degrees of freedom**, which takes into account the reliability of the data by relating the critical value to the number of measurements collected.

Tally chart for clast size		
Size Class (cm)	East Shore	West Shore
1.5-1.9		
2.0-2.4	I	IIII
2.5-2.9	III	III
3.0-3.4	IIII	IIIIII
3.5-3.9	IIIIII	IIIIIIIIIII
4.0-4.4	IIIIIIIIIII	IIIIIIIIIII
4.5-4.9	IIIIIIIIIII	IIIIIIIIIII
5.0-5.4	IIIIII	IIIII
5.5-5.9	IIIII	III
6.0-6.4	II	I
6.4-6.9	I	
Mean	4.46	3.46

Students could be required to calculate data based around tallies such as this

Alternatively, these means may not be provided and students might be required to work them out.

Chi-squared χ^2

This test is used to look for differences in data which is categorical and where the expected distribution of the data is not known. In other words, data must be collected so that it is grouped into classes. As in other tests, a null hypothesis must also be developed. This usually takes the form:

'X' has a random distribution

In this worked example, data has been collected on the orientation of corries. This data was obtained using a GIS program.

The null hypothesis would be 'The orientation of corries is random', or 'There is no preferred alignment in corrie orientation'.

Angle (°)	**0–89**	**90–179**	**180–269**	**270–359**
No. of corries	14	6	4	5

The calculation of Chi-squared (χ^2) means that the data must be put into a new table to determine the expected frequencies. Expected frequencies are just averages for each orientation. As there are 29 corries in total, we would 'expect' 7.25 in each orientation (29/4).

Angle (°)	**0–89**	**90–179**	**180–269**	**270–359**
Observed (O)	14	6	4	5
Expected (E)	7.25	7.25	7.25	7.25
$(O-E)^2$	45.56	1.56	10.56	5.06
$(O-E)^2/E$	6.28	0.22	1.46	0.70

$\chi^2 = 6.28 + 0.22 + 1.46 + 0.70 = 8.66$

In this calculation, degrees of freedom = 4 – 1 = 3

The critical value at 5% or 0.05 = 7.82

As the χ^2 is above the critical value (8.66>7.82), then we can reject the null hypothesis and confirm there is some preferred orientation. The higher the χ^2 value then the less likely is the fact that the data occurred by chance.

This is an example of a one-column / one-row test. If you were using two or more rows and columns, you would then change the expected values so that you have proportional expected values. This might be relevant for the NEA, for instance.

There is an excellent introduction to the Chi-squared test here: www.khanacademy.org/math/statistics-probability/inference-categorical-data-chi-square-tests/chi-square-goodness-of-fit-tests/v/chi-square-distribution-introduction

Statistical tables can be downloaded from a range of online websites. The Royal Statistical Society versions can be accessed through this link:

https://www.rss.org.uk/Images/PDF/pro-dev/RSS-Tables-2012-watermarked.pdf

Some final statistical considerations

This support material is intended to provide an introduction to inferential statistics and their application in geography. There is a range of other online support that can be used to supplement and provide additional worked examples.

The FSC has material here:
https://www.geography-fieldwork.org/a-level/before-starting/analysis/

Also the RGS:
http://www.rgs.org/OurWork/Schools/Teaching+resources/Key+Stage+5+resources/Data+skills+and+thinking+geographically/FSC+statistical+methods.htm

There is also useful material here from the Data Skills section of the RGS website:
http://www.rgs.org/OurWork/Schools/Teaching+resources/Key+Stage+5+resources/Data+skills+and+thinking+geographically/Data+skills+and+thinking+geographically.htm

Finally, there is the useful RGS Individual Investigation Guide, available online, which has a dedicated statistics section:
http://www.rgs.org/OurWork/Schools/Teaching+resources/Key+Stage+5+resources/Data+skills+and+thinking+geographically/A+guide+to+the+A+Level+Independent+Investigation

Additional statistical information can be found in this new book from the Geographical Association ***Fieldwork at A level: your guide to the independent investigation***

http://www.geography.org.uk/shop/shop_detail.asp?ID=847

When discussing statistics with students, please reinforce the following:

- Inferential statistics need to use null and alternative hypotheses in order to make sense of the outcome.
- Outcomes should always be stated in the context of a significance level, i.e. the results at 99% are 'stronger' than the ones at the 95% level.
- Smaller data sets may be heavily affected by anomalies and distorted patterns.
- Causal relationships may be shown, but what is the underlying theory?
- Statistics give us a lot more support in making conclusions and evaluating. They provide statistical evidence, which may or may not be useful in the context of what you are trying to find out.

The Pearson Edexcel Fieldwork Guide has additional information about statistics in the context of the NEA.

11. Sampling and data

Demand

Students learn that interpolation is more accurate than extrapolation, though may not use these terms to describe it.

Students learn to write and criticise questionnaire questions in KS3, but this is not on the GCSE specification.

All students learn about choosing a random sample to avoid bias in GCSE maths. They should understand the concept of random numbers and using these to generate a random sample, but they may not know how to generate random numbers on a calculator.

Only Higher level students learn stratified sampling and the capture-recapture method in GCSE maths.

Approach

Estimating from samples

Use arrow diagrams to scale up

For example, a student is planning some fieldwork to investigate differences in dandelion plants between two fields. The dandelions are being used as one indicator of tourism pressure in the area, so this sampling method is used to estimate how many dandelions there might be in total in each of the fields.

In the first field, a quadrat $0.5\,\text{m} \times 0.5\,\text{m}$ is randomly thrown 5 times.

The number of dandelions in each quadrat is counted. The mean is 10 plants.

The field has area $2000\,\text{m}^2$.

Estimate the number of dandelions in the field.

		Area	Dandelions	
Sample		$0.25\,\text{m}^2$	10	
	$\times 4$			$\times 4$
In $1\,\text{m}^2$		$1\,\text{m}^2$	40	
	$\times 2000$			$\times 2000$
Field		$2000\,\text{m}^2$	80 000	

Terminology

- **Inconsistency:** In maths, students learn that data is either continuous, discrete (rather than discontinuous) or categorical. Data that is measured is continuous.
- Inconsistency: In maths, students do not manipulate data (which has negative connotations) – they process it.

- In geography, good questionnaire questions have tick boxes (closed) to reduce answer options. Make sure ranges for the tick boxes do not overlap, all options are covered, questions are set in a clear timeframe (e.g. how many times do you go shopping per month, rather than just 'how many times do you go shopping?'), and are not leading or biased. A survey should be anonymous in order to get more reliable outcomes. There may also be ethical considerations to consider.
- In a random sample, every member/item within a population has an equal chance of being selected.
- To select a random sample of size 10, you can:
  - give every population member a number
  - generate 10 random numbers and select the members with those numbers for the sample.
- For a small sample size, you can put 'numbers into a hat' and draw them out at random. This is not practical for a large sample size.
- A sample that is too small can give biased results. This is particularly true for geographical fieldwork. Surveying two sites along a river, for example, would yield insufficient data and therefore would count as an unreliable sample. Similarly, measurement of a river's depth should take into account variability of the river bed. An uneven bed should have more samples than a regular cross-section.
- For data that is grouped, for example by age bands, you can take a stratified sample, so the sample reflects the proportion of each age group in the population. This would be important in a fieldwork survey of a town, for example, to get a fair and reliable survey or sample of people's attitudes and opinions.
- You can use lines of best fit or follow the trend of a graph to estimate 'missing' data values. Estimating values that lie within the range of given values is called interpolation, though students may not learn this term. Estimating values that lie outside the range of given values is called extrapolation, though students may not learn this term. Interpolation is more likely to be accurate than extrapolation.

Data reliability and validity

Data reliability and validity are terms frequently used by people and organisations and have taken on a ‘blurred’ meaning. Care needs to be taken to use these terms with a degree of accuracy and technical correctness – see the table below for definitions.

Data evaluative term	**Description**
Reliability	The extent to which measurements are consistent if the investigation/experiment were repeated. A reliable design and / or investigation would lead to small variations between results.
Validity	The suitability of the method to answer the question that it intended to answer. In other words, the confidence in a set of results and outcomes.
Accuracy	How close the value or result is to the true or actual value. In many instances (especially when conduction fieldwork), the true value can never be determined with any degree of certainty.
Precision	Precise measurements have very little spread around the mean value, so the term refers to the closeness of two or more readings or measurements. Precision is independent of accuracy.

Some investigators also talk about a ‘true value’, which is the result of the investigation obtained using an ideal measurement.

Using this terminology is an essential part of the process of evaluation.

Acknowledgements

***AS and A Level Geography: Guide to Maths for Geographers* by David Holmes. Acknowledgements prepared on 4th October 2017. Amended 6th October 2017.**

We are grateful to the following for permission to reproduce copyright material:

Figures 19, 20, and 35 from Edexcel AS Paper 1 June 2014, Edexcel 6GE01 June 2016 Paper and Pearson Edexcel A level Geography Unit 3 June 2016, copyright © Pearson Edexcel; Figures 24 and 25 from *The Spirit Level* by Richard Wilkinson and Kate Pickett, 2009 and www.equalitytrust.org. Reproduced with permission of The Equality Trust; Figure 29 from http://landregistry.data.gov.uk/app/ukhpi/explore, © Crown copyright 2017. Open Government Licence. All content is available under the Open Government Licence v3.0; Figure 33 from, www.gapminder.org/world/. Reproduced for free from gapminder.org; Figure 36 from ArcGIS Online. Copyright © 2017 Esri, ArcGIS Online, and the GIS User Community. All rights reserved. Used with permission; Figure 37, www.wunderground.com/wundermap/, copyright © Mapbox © OpenStreetMap. Reproduced with permission; Figure 38 from *An Introduction to Geographical Landscape Drawing* by G.E. Hutchings, 1955, copyright © Field Studies Council; Figure 39 by Chris Warn. Reproduced by kind permission of Sue Warn; Figure 40 from https://fehweb.ceh.ac.uk/GB/map, copyright © NERC (CEH) 2017, Natural Environment Research Council Centre for Ecology & Hydrology; Figure 41 'Lizard Wood', Ordnance Survey Maps, © Crown copyright 2016, OS 100030901; Figure 'Grain shape; sedimentary particles become more rounded and spherical as they are transported' Figure 7 in *The Geology of Sedimentary Rocks* http://intheplaygroundofgiants.com/geology-of-the-grand-canyon-region/the-geology-of-sedimentary-rocks/ copyright © Dr Kenneth A. Bevis, 2014. Reproduced by permission; and Figure "Distribution of UK household disposable income, financial year ending 2016" from *Effects of taxes and benefits on UK household income: financial year ending 2016*, Figure 1, 25 April 2017, Crown copyright 2017. All content is available under the Open Government Licence v3.0.

In some instances we have been unable to trace the owners of copyright material, and we would appreciate any information that would enable us to do so.

