Qualification Accredited

OCR
Oxford Cambridge and RSA

AS and A LEVEL

Mathematical Skills Handbook

GEOLOGY

H014/H414
For first teaching in 2017

OCR Advanced Subsidiary and Advanced GCE in Geology

Version 1.4

Contents

OCR will update this document on a regular basis. Please check the OCR website (www.ocr.org.uk) at the start of the academic year to ensure that you are using the latest version.

Version 1.4 February 2020

Version 1.4
Changes of note made between Version 1.3 and Version 1.4:

1. Correction of the figures used in calculations for return periods in section M2.3

Version 1.3 June 2019

Version 1.3
Changes of note made between Version 1.2 and Version 1.3:

2. Correction of the definition of uncertainty in section M2.4

Version 1.2 April 2019

Version 1.2
Changes of note made between Version 1.1 and Version 1.2:

1. Change in M2.3 to the list of calculated return periods
2. Change in M2.12 to data column order in example Chi squared computation.
3. Change in M2.12 concluding paragraph of example Mann-Whitney U-test computation.

Version 1.1 March 2018

Version 1.1
Changes of note made between Version 1.0 and Version 1.1:

1. Change in M4.1 to the geometrical formulae that need to be recalled. Formulae for calculating surface area and volume of spheres will now be provided in assessments where needed.
2. Change to Appendix A to provide up to date list of formulae that must be recalled.
3. Change to Appendix B to provide up to date list of formulae that will be provided where needed in assessments.

Version 1.0 September 2017

Version 1.0
Original issue

DISCLAIMER

This resource was designed using the most up to date information from the specification at the time it was published. Specifications are updated over time, which means there may be contradictions between the resource and the specification, therefore please use the information on the latest specification at all times. If you do notice a discrepancy please contact us on the following email address: resources.feedback@ocr.org.uk

1 Introduction

In order to be able to develop their skills, knowledge and understanding in AS and A Level Geology, students need to have been taught, and to have acquired competence in, the appropriate areas of mathematics relevant to the subject as indicated in Appendix 5e of the specifications:

- **H014/H414 - OCR Geology A**

The assessment of all AS and A Level Geology qualifications will now include at least 10% Level 2 (or above) mathematical skills as agreed by Ofqual (see below for a definition of 'Level 2' mathematics). These skills will be applied in the context of the relevant geology.

This Handbook is intended as a resource for teachers, to clarify the nature of the mathematical skills required by the specifications, and indicate how each skill is relevant to the subject content of the specifications.

The content of this Handbook follows the structure of the Mathematical Requirements table in Appendix 5e of the specifications, with each mathematical skill, M1.1 – M4.3, discussed in turn. The discussion of each skill begins with a description and explanation of the mathematical concepts, followed by a demonstration of the key areas of the geological content in which the skill may be applied. Notes on common difficulties and misconceptions, as well as suggestions for teaching, may be included in each section as appropriate.

As this Handbook shows, all required mathematical skills for geology can be covered along with the subject content in an integrated fashion. However, as assessment of the mathematical skills makes up at least 10% of the overall assessment, OCR recommends that teachers aim to specifically assess students' understanding and application of the mathematical concepts as a matter of course, in order to discover and address any difficulties that they may have. This is particularly relevant for students who are not taking an AS or A Level Mathematics qualification alongside AS or A Level Geology.

Definition of Level 2 mathematics

Within AS or A Level Geology, 10% of the marks available within the written examinations will be for assessment of mathematics (in the context of geology) at a Level 2 standard, or higher. Lower level mathematical skills will still be assessed within examination papers, but will not count within the 10% weighting for geology.

The following will be counted as Level 2 (or higher) mathematics:

- application and understanding requiring choice of data or equation to be used
- problem solving involving use of mathematics from different areas of maths and decisions about direction to proceed
- questions involving use of A Level mathematical content (as of 2012) e.g. use of logarithmic equations.

The following will <u>not</u> be counted as Level 2 mathematics:

- simple substitution with little choice of equation or data and/or structured question formats using GCSE (9-1) Mathematics (based on 2012 GCSE mathematics subject content).

As lower level mathematical skills are assessed in addition to the 10% weighting for Level 2 and higher, the overall assessment of mathematical skills will form greater than 10% of the assessment.

M1 – Arithmetic and numerical computation

M1.1 Recognise and make use of appropriate units in calculations

Students should be able to demonstrate their ability to:

- convert between units e.g. ppb to gram per tonne
- use correct units as part of calculations for gold ore concentration factor
- work out the unit for a rate e.g. sedimentation rate.

Mathematical concepts

Units indicate what a given quantity is measured in. A measured quantity without units is meaningless, although note that there are some derived quantities in geology that do not have units, for example pH.

At GCSE, students will have used different units of measurement and would be required to recognise appropriate units for common quantities. For example, whilst cm is appropriate for a length or distance, students should be able to identify that cm^2 is used for area and cm^3 is used for volume. Students should be aware that in geology pipettes usually measure volume in ml or µl and should be able to readily convert between ml and cm^3, i.e. 1 ml = 1 cm^3. However, students should be aware that volume will usually be stated in cm^3 in assessments.

Students will be expected to be able to convert between different metric units, for example 0.5 m = 500 mm, without conversion 'facts' being given (i.e. 1 m = 1000 mm). Converting between different multiples is a matter of either multiplying or dividing by the appropriate factor, depending upon the direction of the conversion. For example, converting 7 µm to mm requires a *division* by 10^6, *not* a multiplication. It is a common misconception for students to believe that because millimetres are 'larger' than micrometres (in the sense that 1 mm is larger than 1 µm) that a multiplication is necessary to go to the larger unit. However, a simple check should reveal that 7 000 mm is not equal to 7 µm, and so a division is required.

Typical measures that would have been encountered at GCSE and would come up within AS/A Level Geology are distance, area, volume, density and mass.

Unit prefixes indicate particular multiples and fractions of units. A full list of SI unit prefixes is given in Table 1, with the prefixes that are most likely to be used within the AS/A Level Geology course highlighted.

Factor	Name	Symbol	Factor	Name	Symbol
10^{24}	yotta	Y	10^{-1}	deci	d
10^{21}	zeta	Z	10^{-2}	centi	c
10^{18}	exa	E	10^{-3}	milli	m
10^{15}	peta	P	10^{-6}	micro	µ
10^{12}	tera	T	10^{-9}	nano	n
10^{9}	giga	G	10^{-12}	pico	p
10^{6}	mega	M	10^{-15}	femto	f
10^{3}	kilo	k	10^{-18}	atto	a
10^{2}	hecto	h	10^{-21}	zepto	z
10^{1}	deca	da	10^{-24}	yocto	y

Table 1: SI unit prefixes

Rates of change would have also already been encountered at GCSE, for example speed in m s^{-1}. At AS and A Level, students should be able to work out the appropriate unit for a rate, e.g.

{quantity of 'stuff' they are measuring}/{Unit of Time}.

A common example is to work out the rate of change of temperature. For example the geothermal gradient in a sedimentary basin in which a temperature down a borehole (°C) is being measured over a distance (km), the rate of change would be in °C km^{-1} as the temperature change is measured against depth (by convention km).

In the AS and A Level Geology assessments, students will be expected to recognise and use compound units in the form °C km^{-1}, rather than °C/km (see Appendix C).

Other examples of common rates of change in geology include:

- river discharge ($m^3\ s^{-1}$)
- seafloor spreading rate (cm a^{-1})
- heat flow (mW m^{-2})
- a temperature change over time (°C s^{-1})

Within the OCR GCE Geology qualifications, students will in general be expected to use and recognise standard SI units. For example, dm^3 is used rather than l (litre). However, there are exceptions to this, e.g. degree (°) for angles, which is used in preference to the radian, annum (a) and hours (h) in addition to seconds (s), ml and μl in pipette use (as discussed above). In general, any other conversion to or from non-standard units that may be required in assessment would be provided in the question.

Contexts in geology

When making measurements during fieldwork or in an experiment, Geology students will have to decide on the appropriate units of measure to use, and the unit symbol, for example common quantities include,

- **distance: km, m, cm, mm, μm.**
- **area: km^2, m^2, cm^2, mm^2.**
- **volume: km^3, m^3, cm^3, mm^3.**
- **density: Mg/m^3, g/cm^3**
- **mass: t, kg, g**
- **time: Ga, Ma, a, d, h, min, s.**
- **temperature: °C, K.**
- **pressure: pa, Mpa**
- **angle: degrees°**
- **azimuth: degrees°**

When gradients are measured on graphs students will need to decide what the correct units of measurement are for the rate of change.

Students need to be aware at how crucial the choice of units are and the impact on the scale of results. For example, when considering measurements of sea floor spreading for millions of years and using these to calculate a yearly rate.

A sample of basalt from the Pacific Ocean is 2 million years old and is 100 km from the MOR, to calculate the spreading rate the distance must be converted into cm, 10 000 000 cm and years is 2 000 000. The spreading rate is calculated by dividing distance by time, therefore an average rate of 5 cm a^{-1}.

M1.2 Recognise and use expressions in decimal and standard form

Students should be able to demonstrate their ability to:

- use an appropriate number of decimal places in calculations e.g. for a mean
- carry out calculations using numbers in standard and ordinary form e.g. use of magnification
- convert between numbers in standard and ordinary form
- understand that significant figures need retaining when making conversions between standard and ordinary form e.g. 0.063 mm is equivalent to 6.3×10^{-2} mm

Mathematical concepts

Standard Form

Many numbers in geology will be written in standard form (scientific notation). Students are expected to be able to express results in standard form and be able to convert to and from decimal form. For example, the size of a crystal may be 100 μm or 0.0001 m in ordinary (decimal) form. This would be written in standard form as 1.0×10^{-4} m.

The mathematical notation for a number written in standard form is,

$$(-)a \times 10^{n}$$

Where n is an integer (whole number) and

$$1 \leq a < 10$$

When using standard form in calculations, a key mistake is for students to concentrate on the value of n, and forget the rule for the value of a. For example, they may write:

$$(2.0 \times 10^{4}) \times (8.0 \times 10^{3}) = 16 \times 10^{7}$$

Here the student has correctly dealt with the indices and has got the correct numerical answer, but the number is not in standard form. The 16 is incorrect and has to be 'scaled' down by dividing by 10. However if we divide by 10 we have to multiply the 10^{7} by 10 to keep the numerical value the same:

$$16 \times 10^{7} = 1.6 \times 10^{8}$$

Other problems occur if students are required to apply the power rule (see Appendix C), for example in,

$$(2 \times 10^{-3})^{2} = 4 \times 10^{-6}$$

Students will make a number of errors here. Common mistakes are,

- to forget to square the 2: $(2 \times 10^{-3})^{2} = 2 \times 10^{-6}$
- to ignore the negative power: $(2 \times 10^{-3})^{2} = 4 \times 10^{6}$
- to add the powers rather than multiplying them: $(2 \times 10^{-3})^{2} = 2 \times 10^{-1}$

Whatever form numbers are given in, students must use the appropriate number of decimal places or significant figures (as appropriate) in calculations. In the context of converting between standard and ordinary form, students must appreciate that significant figures need to be retained. For example:

$$0.0050\ \text{mol dm}^{-3} = 5.0 \times 10^{-3}\ \text{mol dm}^{-3}$$

Here the final zero in the expression on the left is a significant figure, and so must be retained in standard form.

Decimal Places and Significant Figures

Geology students are usually expected to record raw data to the same number of decimal places (rather than the same number of significant figures). For example, when recording the following volumes 9.0, 9.5, 10.0 and 10.5 ml, the measurements can be recorded to the same number of decimal places (but they do not have the same number of significant figures),

Measurement	Volume (ml) ✗	Volume (ml) ✓
1	9	9.0
2	9.5	9.5
3	10	10.0
4	10.5	10.5
		Mean = 9.75

Processed data can be recorded to up to one decimal place more than the raw data. For example, if the students were asked to calculate the mean for the above example, the answer could be recorded as 9.8 or 9.75 ml. (See the Practical Skills Handbook for more guidance on tables).

In the examinations, students may be asked to record their answer(s) to a particular number of decimal places or to a particular number of significant figures.

Decimal Places in Calculations

When adding and subtracting numbers that are quoted to the appropriate number of significant figures, the answer should be given using the lowest number of decimals used in the calculation. This can sometimes give different results than if the answer was given to the lowest number of significant figures (see *M1.3*).
For example:

> 25.5 – 8.3 = 17.2; answer given to the lowest number of decimal places, not lowest number of significant figures
>
> 105.5 – 93.75 = 11.8; answer given to the lowest number of decimal places, not the lowest number of significant figures.

Calculator Use

Students with access to a scientific calculator should be able to use it to convert between different decimal/standard form calculations, as well as enter numbers in standard form. Table 2 shows the required functions for common makes of calculator.

Calculator make	Convert standard / decimal	Enter standard form
Casio	S→D	$\times 10^x$
Texas Instruments	◄►	$10^{\square}$
Sharp	Change	EXP

Table 2: Calculator functions for standard form

For other models encourage students to investigate the appropriate functions for themselves.

It should be noted that calculators will not necessarily retain the correct number of decimal places required for the calculation. For example, 3.0×10^3 is correct to 2 significant figures, but once entered into a calculator the display could be 3×10^3, which loses 1 significant figure.

Contexts in geology

There are many areas where students will be required to recognise and use standard form. Any calculations which involve large or small numbers will require the use of standard form.

When making observations and descriptions of igneous rocks students may need to use expressions in decimal form, crystal grain size classifications are based on measurements and when assigning samples to different categories students may need to use decimal places.

<1 mm fine crystal size

1-5 mm medium crystal size

>5 mm coarse crystal size.

Measuring quantities by difference

This is a main area where students need to consider the role of decimal places in addition and subtraction.

The most common quantities measured by difference in practical work are mass, temperature and volume. The measurements made should be recorded to a specific number of decimal places, depending on the resolution of the instrument (see the Practical Skills Handbook for more on this topic). When calculating the difference between the measurements, this number of decimal places should be maintained.

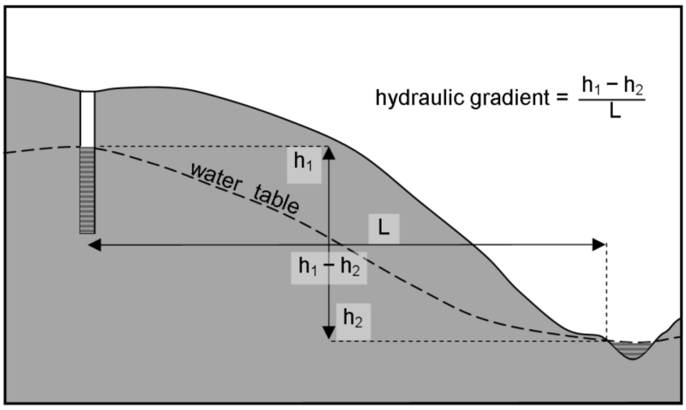
For example, a student conducting an investigation may record the following measurements:

Initial temperature	22.5 °C
Final temperature	29.5 °C
Temperature difference	7.0 °C

The temperature difference is given to 1 decimal place, to match the resolution of the measured values. The '0' is significant, so must be included.

There are many areas in geology where students may be required to calculate a rate of change or gradient, the method of which involves considering the difference between quantities. One such example is when calculating hydraulic gradient of the water table or piezometric surface between two points.

Hydraulic gradient = change in hydraulic pressure (h_1-h_2) / distance between 2 points (L)



M1.3 Use an appropriate number of significant figures

Students should be able to demonstrate their ability to:

- Report calculations to an appropriate number of significant figures given raw data quoted to varying numbers of significant figures
- understand that calculated results can only be reported to the limits of the least accurate measurement.

Mathematical concepts

The number of significant figures used to express particular values ultimately derives from the resolution of the measuring apparatus used to determine experimental values. See the Practical Skills Handbook for more on this topic, including the appropriate number of decimal places to use for certain apparatus.

Common rounding errors include:

- Forgetting to include zeroes where they are significant figures, rather than placeholders. For example, 4.99×10^5 rounded to 2 significant figures is 5.0×10^5, not 5×10^5
- Rounding sequentially; for example rounding 2.4478 first to 2.45 and then to 2.5. This is incorrect; the number should be rounded in a single step, giving 2.4 to 2 significant figures.

Students must understand that the lowest level of accuracy in the inputs of a calculation will determine the level of accuracy in the answer. If there are 3 inputs to a particular calculation and they are quoted as being correct to 2, 3 and 4 significant figures, respectively then the answer can only be quoted reliably correct to 2 significant figures. Note though that if the calculation only involves addition and subtraction, decimal places should be taken into account rather than significant figures – see Section *M1.2*.

In these calculations, students should be aware that certain numbers are 'exact', that is, there is no uncertainty in their value. These numbers can be treated as having an infinite number of significant figures; they do not affect the number of significant figures the result of a calculation should be reported to.

Contexts in geology

When using GPS to measure the absolute movement of tectonic plates data can be recorded which are both accurate and have a high precision, by using specialist geodetic GPS techniques. But when deriving average relative plate motion consideration must be given to the appropriate number of significant figures. Rates may vary, observations may not be representative of the whole plate, sources of error may exist and assumptions that rates have been constant may not be valid.

The number of significant figures used is important when analysing data collected in the field. For example when processing field data best practice is to express the calculated result to the same number of significant figures as the least precise measurement. However when multiple repeats are made, such as the length of clasts within a conglomerate bed, there may be some justification in expressing the mean length to an additional significant figure.

Thinking about significant figures is important in any calculation, and students should be particularly aware of the meaning of significant figures when performing calculations using experimentally determined values. Geology students could be asked to report an answer to a certain number of decimal places or to a certain number of significant figures. For example:

The diameter of a globular shaped foram was measured as 12 µm. Calculate the volume of the foram and report your answer to 3 significant figures.

$$Volume\ of\ a\ sphere\ = \frac{4}{3}\pi r^3 = \frac{4}{3}\pi 6^3$$

$$Answer = 904.78 = 905\ \mu m^3\ (to\ 3\ significant\ figures)$$

M1.4 Use ratios, fractions and percentages

Students should be able to demonstrate their ability to:

- calculate percentage yields
- calculate surface area to volume ratio
- use scales for measuring

Mathematical concepts

Ratios, fractions and percentages are related concepts. Many problems within geology will require students to have a good understanding of the relationships between these concepts, and to use them in calculations. The individual skills required will have been covered at GCSE (9-1) mathematics and science, but they are used in new contexts here.

Percentage change

There are many misconceptions when performing these calculations. This is often because students are 'over-taught' the method on how to find percentage increases and decreases, actually the calculations are relatively easy and shouldn't produce too much anxiety. The key is to understand that the multiplier 1 represents a change of 0%. A multiplier of 1.43 therefore represents an increase of 43% whilst a multiplier of 0.83 represents a decrease of 17% (note that it is the difference between the multiplier and 1 which is the change – it isn't a percentage decrease of 83%). Quantities and percentages can then be found using a simple formula:

$$Original\ Quantity\ \times Multiplier = New\ Quantity$$

This formula can be stated in a formula triangle and then applied to situations where the percentage change is required. For example, if the initial mass of a saturated rock sample is 56 g and after a night in a 75°C oven it is 47 g then to work out the percentage decrease we have:

$$56 \times Multiplier = 47$$

$$Multiplier = \frac{47}{56} = 0.84$$

This represents a percentage decrease of (1 - 0.84) = 16% decrease.

Or, say we expect hydrofracturing to cause a percentage increase of 10% in production what will the predicted daily production of an oil well that was producing 800 bopd (barrels oil per day):

$$800 \times Multiplier = New\ production\ rate$$

$$800 \times 1.1 = New = 880\ bopd$$

Contexts in geology

Surface area to volume ratio

Students will also need to calculate surface area to volume ratios and understand the implications that these ratios have. Students will have come across this at GCSE (for example in the context of insulation) and this concept will be covered in A Level geology (Hjulström curve, behaviour of clays, fluids in rocks etc.).

The formula to calculate the surface area to volume ratio is:

$$Ratio = \frac{Surface\ Area}{Volume}$$

See *M4.1* for a list of formulae to calculate the surface areas and volumes of regular shapes.

Scales

Students will need to use scales for measuring. The ability to use the Unitary Method (using the value of a common unit) to label diagrams and pictures is an essential skill. This can be denoted by a sentence, (i.e. 1 cm represents 1μm) or a scale bar or a ratio (1: 0.00001). A fiducial (or fiducial marks) in a photograph or scientific drawing is a scale bar or reference object (coin, compass-clinometer) which can be used to calculate the true size of features in the image.

Concentration factor

When considering metallic mineral resources, students calculate the average crustal abundance as a percentage and the percentage grade of that metal within a deposit. Students must use these values to calculate the concentration factor, how much the concentration of a metal must increase above its average to form an ore deposit. Students could be provided with any two pieces of data for a named metal and rearrange the formula to allow calculation of the third value.

$$concentration\ factor = \frac{\text{grade of metal in ore}}{\text{average crustal abundance}}$$

Metal in ore deposit	Average crustal abundance %	Grade %	Concentration factor
Silver	0.0000075	4	533333
Magnesium	2.3	3.0	1.3
Calcium	4.2	34.0	8.1

Reporting Ratios

When presenting information as a ratio of one quantity to another the ratio is reported in the form

x : 1

where x is found by dividing the first quantity by the second.

For example two species of brachiopod were counted in a sedimentary unit:

Fossil species	Number of individuals
Productid	78
Spiriferid	20

What is the ratio of productid to spiriferid brachiopods?

78 / 20 = 3.9

The ratio of spiriferids to productids is **3.9 : 1**

What is the ratio of spiriferid to productid brachiopods?

20 / 78 = 0.256 = 0.3

The ratio of productids to spiriferids is **0.3 : 1**

When more than two quantities are all being compared in this way the order in which the ratio is given once again follows the order in which the different quantities are named and the last one will always be given as 1, with the other numbers all relative to that.

Fossil species	**Productid**	**Spiriferid**	**Rhynchonellid**	**Orthida**
Number of individuals	127	50	32	8
Number of individuals / 8	15.9	6.3	4	1

The ratio of productid to spiriferid to rhynchonellid to orthida is **15.9 : 6.3 : 4.0 : 1**

M1.5 Make order of magnitude calculations

Students should be able to demonstrate their ability to:

- use and manipulate the magnification formula,

$$Magnification = \frac{size\ of\ \ image}{size\ of\ \ real\ \ object}$$

Mathematical concepts

Students will need to be able to deal with different orders of magnitude. For example, in microscopy the size of a structure on a micrograph may be measured in mm but the actual size is six orders of magnitude smaller (nm). Order of magnitude calculations can be used to approximate values without extensive calculations and are also a useful check to ensure calculated values are reasonable (linking back to *M1.6*).

Contexts in geology

Students need to use the orders of magnitude calculation to consider the evidence in thin sections and micrographs. When comparing the environments of deposition of a range of sandstones by measuring grains in thin section, together with knowledge of the magnification the size of the real grain size is calculated and could indicate the strength of the palaeocurrent.

For example, take the micrograph showing glacial sand, with a large quartz grain circled:

If the magnification is ×125 and the length of the circled grain is 34 mm, to calculate the actual size of the grain,

$$\frac{\text{size of image}}{\text{magnification}} = \text{size of real object}$$

$$\frac{34}{125} = 0.272\,\text{mm} = 272\,\mu\text{m}$$

However, if a student had accidently multiplied the two values together,

$$34 \times 125 \ = 4{,}250 \text{ mm} = 4{,}250{,}000 \text{ μm}$$

Then an approximate value for the size of the quartz grain would be useful to show the student that this calculated value is not reasonable.

M1.6 Estimate Results

Students should be able to demonstrate their ability to:

- Estimate results to sense check that the calculated values are appropriate.

Mathematical concepts

Being able to make an estimate for the size of a given measure is a notoriously difficult concept. The best advice for making reasonable estimates is to start at a 'known' quantity and then extrapolate from that fact to the object to be estimated. For example, let's say the mass of a boulder is to be estimated. If the mass of an average small family car is 1500 kg then perhaps a sensible first estimate is to say that the boulder is 2-3 times the mass of the car. Hence the mass of the boulder could be estimated as 1500 x 2.5 = 3750 kg.

In making estimations in calculations care should be taken with how the numbers are rounded as it could lead to confusion as to whether there is an under or over estimation.

Estimating is a valuable skill; if you are able to estimate an approximate answer to a calculation, it is easier to spot if you have made a mistake in carrying out the actual calculation. For example the calculation:

4.9/1.10

could be estimated as

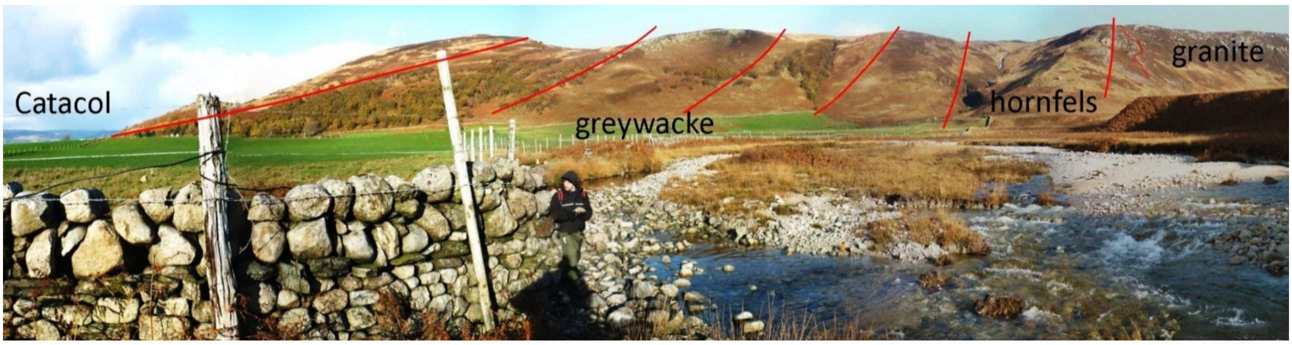
5/1 = 5

If the answer is then calculated as 4.45, the estimate gives reassurance that this is a reasonable answer. However, if the calculation gives an answer of 0.45, the estimate will help to recognise that a decimal point error has been made.

Estimating can become easier if students are familiar with the types of answer that are typical for a particular situation.

Contexts in geology

When working towards understanding the structure and relationships between rock units in the field students may begin with considering the topography. The shape of a hillside can be a useful way to estimate the dip of the strata. Once this measurement has been estimated students can focus on particular sites in the field area where they can use their compass clinometer to make accurate measurements to plot on a map and later use in constructing cross sections.



The view northeast across Glen Catacol, Arran, the topography clearly depicts the effect of the forceful intrusion of the North Arran granite by dipirism. The apparent dip of the beds in the Southern Highland Group greywackes increases from 21° at Catacol village to close to vertical close to the contact with the granite. By estimating the dip before starting a more detailed investigation would allow students to identify locations where the structure is different to regional trends when investigating individual localities.

M2 – Statistics and probability

M2.1 Find arithmetic means

Students should be able to demonstrate their ability to:

- find the mean of a range of data e.g. the mean clast size.

Mathematical concepts

Means and weighted means

The mean ($\bar{x}$) is calculated using a simple formula:

$$\bar{x} = \frac{\sum x}{n}$$

where Σx is the sum of the data values and n is the number of data values.

Most students will be familiar with this from GCSE and it is best taught as a rather ad-hoc message: ‘add them all up, divide by how many’. There are few misconceptions with this when dealing with raw, listed data as the calculations involved are quite simple. However students may attempt to use arithmetical mean with circular data, such as current directions, rather than calculating the vector mean (see *M2.6*).

Anomalies and selecting data

When analysing experimental data, for example calculating the mean value of a number of measurements, students must be able to identify anomalies (outliers) and exclude them from the calculation. An anomaly/outlier is a value in a set of results that is judged not to be part of the inherent variation (see *The Language of Measurement* 2010, ASE-Nuffield). It should be emphasised that there is no hard and fast rule on how to deal with anomalies; they should be treated case by case.

A simple checklist for fieldwork and experimental data is this:

- was the suspected anomaly recorded in error?
- was the suspected anomaly recorded in different conditions to the other values?

If the answer to any of these questions is yes, then the anomaly should be omitted from the data set and the mean should be calculated without this value. If a potential anomaly is spotted at the time of the experiment then students should question whether the experiment should be repeated (see the Practical Skills Handbook for more on this topic). It is possible to use mathematical techniques to identify anomalies in data however this goes beyond the scope of AS/A level Geology, but may be appropriate for student studying Level 3 Core Maths.

Calculator use

Many different scientific calculators have a Statistics mode where the mean can be calculated automatically. The latest scientific and graphical calculators have statistical functions to gain full summary statistics, such as quartiles, standard deviation, sum of squares, correlation coefficients etc.). It will be of benefit to students to familiarise themselves with using these statistical functions on their own calculator. Whilst there is no computational advantage of doing this for just finding the mean, it is easier to review their data and then correct any identified input errors.

Contexts in geology

Calculating the mean for sets of data in geology can be very useful, for instance when using replicates in experiments or for comparison purposes. For example, students might be asked to compare the size of clasts in conglomerate beds. The length of five clasts in each bed has been recorded:

Clast size in bed 1	Clast size in bed 3
45	6
48	9
47	11
50	7
46	7

To find the mean for each set of data,

$$\frac{(45+48+47+50+46)}{5}=47$$

$$\frac{6+9+11+7+7}{5}=8$$

Then the students can use the calculated means to infer the energy of the depositional system at the time that the clasts were deposited and compare the events that deposited the conglomerate in the two beds.

As stated in *M1.2*, processed data such as the mean can be recorded to up to one decimal place more than the raw data.

M2.2 Construct and interpret frequency tables and diagrams, bar charts and histograms

Students should be able to demonstrate their ability to:

- represent a range of data in a table with clear headings, units and consistent decimal places
- interpret data from a variety of tables, e.g. data relating to intrusive dykes
- plot a range of data in an appropriate format, e.g. grain size distribution as a cumulative frequency graph
- interpret data for a variety of graphs, e.g. explain seismograph traces.

Mathematical concepts

Students will be familiar with frequency tables, bar charts and histograms from GCSE (9-1) mathematics and science. There are a few misconceptions however for the construction of histograms and bar charts as there is a subtle difference.

Bar charts are used when the data is *categorical/discrete*. The data can only take specific values. Variables such as Mohs hardness and lithology can only take certain values, there are no in-between values; Mohs hardness could be 1,2, … 10, but could not be 1.5. Lithology can only be schist, granite, etc. With a bar chart there are gaps between the bars when it is plotted.

Histograms are used when the data is *numerical/continuous*. Length, mass, temperature, time are all examples of continuous data. With histograms there are *no* gaps between the bars. For more information on tables and graphs please see the Geology Practical Skills Handbook.

Contexts in geology

Plotting a histogram

Below is a table of data which can be plotted as a histogram. The lengths of the a axis of clasts in a conglomerate were measured to the nearest cm and displayed in this table.

Length of the a axis (cm) to nearest cm	Frequency
$2 \le x < 4$	8
$4 \le x < 6$	7
$6 \le x < 8$	9
$8 \le x < 12$	3
$12 \le x < 16$	2

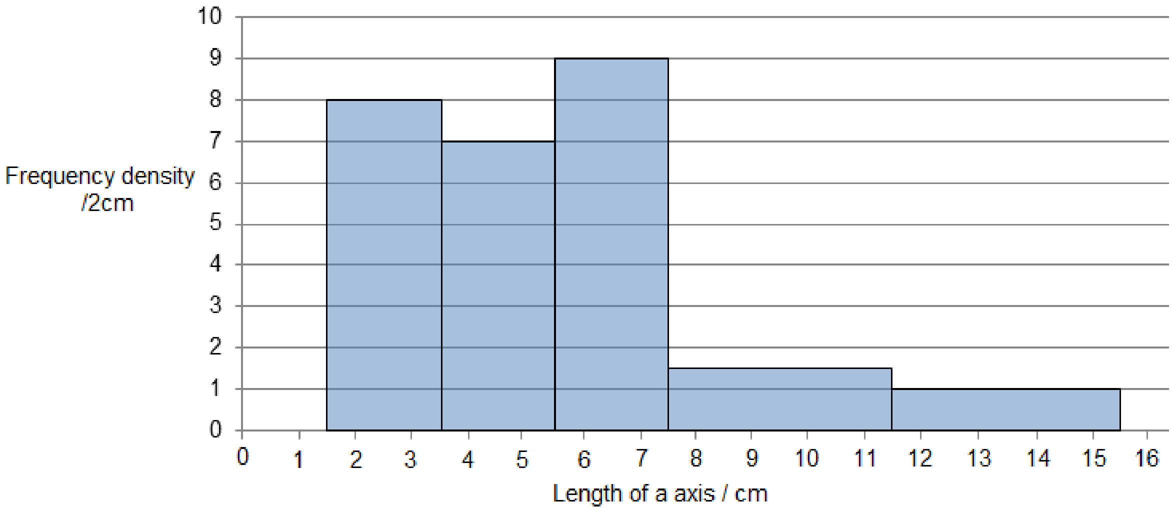
The data above is continuous data. The lengths are recorded to the nearest cm and therefore the groups should reflect this rounding.

Length of the a axis (cm) to nearest cm	Frequency
$1.5 \le x < 3.5$	8
$3.5 \le x < 5.5$	7
$5.5 \le x < 7.5$	9
$7.5 \le x < 11.5$	3
$11.5 \le x < 15.5$	2

This data can be plotted as a histogram, when plotting note that the last two groups are larger than the first three, this must be taken into account when plotting, the area of the bars must be proportional to the frequency. To account for this calculate frequency density, as the common width is 2cm the frequency density can be frequency/2cm.

Length of the a axis (cm) to nearest cm	Frequency	Frequency density/2cm
1.5≤x<3.5	8	8
3.5≤x<5.5	7	7
5.5≤x<7.5	9	9
7.5≤x<11.5	3	1.5
11.5≤x<15.5	2	1

A graph to show the length of a axis of clasts in conglomerate sample



Students are required to interpret the data and suggest what the evidence reveals about the environment of deposition. The presented sediment size data could indicate sorting or energy levels which support particular paleoenvironments.

Interpreting data

Students are often required to interpret data presented in tables or graphs in AS/A Level assessments. These skills are always linked to the information these representations contain. For example, it would be expected that students are familiar with seismogram traces and the information these contain.

M2.3 Understand simple probability

Students should be able to demonstrate their ability to:

- use the terms *probability* and *chance* appropriately
- understand the probability associated with return periods for geohazards.

Mathematical concepts

The fundamental misconception when dealing with probability is the idea of *randomness*. There are a number of different definitions for the word *random* but to keep it simple emphasise to the students that a random process doesn't produce rare results but it produces results that are impossible to *predict*. Randomness is based on our inability to predict what will happen at a *given* moment of time (see *M2.7*).

This can be modelled by rolling a die or, if there are just two equally likely outcomes, flipping a coin.

Contexts in geology

Probabilities are used in a number of contexts including hazard analysis, extinctions and extinction rates, resource forecasting, and design parameters for engineering geology. In geoscience the model can be run multiple times with different starting parameters to evaluate the probability of a particular outcome (for example the formation of the terrestrial planets).

Return periods of geohazards consider how likely and the magnitude of extreme events. Return periods are calculated based on past data and to provide a probability that a given magnitude event will occur in any given year.

$$T = \frac{n+1}{m}$$

Where T = Recurrence interval, n = number of years on record, m = magnitude ranking

The probability of an event with a recurrence interval T is:

$$P = \frac{1}{T}$$

Where P is probability and T is the recurrence interval.

For example Katla on the south coast of Iceland is one of Europe's most explosive volcanoes but the major hazard from the volcano are jökulhlaups (glacial outburst floods). During the 1918 eruption a flood wave with a discharge twice that of the Amazon River reached the sea 2 hours after the start of the major eruption and extended the coast of Iceland 5km southward in one flood event. On average there is a time interval of 70 years between eruptions, so are we overdue another large eruption and flood event? We can use the historical records and data to compile a record of Katla's eruptions and jökulhlaups over the period 871 to 2017 (1147 years).

Date	Ash Layer		Flood deposit	
	Size	**Rank**	**Size**	**Rank**
2011		—	small	12
1918	large	2	large	2
1860	small	21	small	10
1823	small	20	medium	8
1755	large	3	large	4
1721	medium	7	large	3
1660	medium	8	medium/large	6
1625	large	5	medium	7
1612	small	19	small	11
1580	small	18	small/medium	9
1500	large	4	medium/large	5
1460	small	17	—	—
1440	small	16	—	—
1416	medium	9	—	—
1357	medium	10	—	—
1262	large	6	—	—
1245	small	15	—	—
1179	small	14	—	—
1140	small	13	—	—
934	large	1	very large	1
920	small	12	—	—
871	small	11	—	—

Return period for eruptions:

Large eruptions = (1147 + 1) / 6 = 191 years, probability = 1 / 191 = 5.2×10-3

Medium eruptions = (1147 + 1) / 4 = 287 years, probability = 1 / 287 = 3.5×10-3

Small eruptions = (1147 + 1) / 11 = 104 years, probability = 1 / 104 = 9.6×10-3

Any eruption = (1147 + 1) / 21 = 55 years, probability = 1 / 55 = 1.8×10-2

Large jökulhlaups = (518 + 1) / 3 = 47 years, probability = 1 / 47 = 2.1×10-2

Probability of a large eruption with a large jökulhlaups = 3/6 = 0.50

Multiplying the probabilities gives the respective probability of a large eruption happening this year and that eruption resulting in a large jökulhlaup:

$$P = (5.2 \times 10^{-3}) \times 0.50 = 2.6 \times 10^{-3} = 0.26\%$$

$$T = \frac{1}{(2.6 \times 10^{-3})} = 385\ years$$

Therefore although the average time between eruptions is about 70 years the return period for a large eruption accompanied by a large flood is 385 years or a probability of 0.26% in any one year.

M2.4 Identify uncertainties in measurements and use simple techniques to determine uncertainty when data are combined Students should be able to demonstrate their ability to:

Students should be able to demonstrate their ability to:

- calculate percentage error where there are uncertainties in measurement.

Mathematical concepts

When a measurement is taken there will be an uncertainty due to the resolution of the measuring apparatus. When multiple measurements are combined, the uncertainty in the final result will be a combination from the individual uncertainties in each measurement.

When adding or subtracting measurements, the absolute uncertainties are simply added together

combined uncertainty = uncertainty in A + uncertainty in B

The *absolute* uncertainty is the amount a measurement could be 'out' by. The *relative* uncertainty is the ratio of the absolute uncertainty to the quantity measured.

$$\%\, uncertainty = \frac{absolute\ uncertainty}{quantity\ measured} \times 100\%$$

In GCE Geology, students do not need to be able to combine uncertainties in more complex operations, such as when multiplying or dividing.

Contexts in geology

Where equipment is used to make measurements is subject to uncertainty whether this is carried out in the laboratory or the field. For example, a compass clinometer is graduated every 2° and an individual measurement can be interpolated to 1°, however the absolute uncertainty is ±1° in dip angle, or the azimuth of dip or strike.

A digital calliper is graduated in divisions every 0.01 mm and the manufacturer data sheet gives the "instrumental error" as ±0.05 mm. The overall uncertainty in any distance measured always comes from two readings, so the overall uncertainty = 2 × 0.05 mm = 0.1 mm.

In a distance measurement covering the entire 50 mm length of the calliper, the uncertainty is small

percentage uncertainty = $\frac{2 \times 0.05}{50.0} \times 100\%$ = 0.2%

For shorter distances, the percentage uncertainty becomes more significant. For measuring a distance of 0.25 mm:

percentage uncertainty = $\frac{2 \times 0.50}{0.25} \times 100\%$ = 40%

M2.5 Understand the principles of sampling as applied to scientific data

Students should be able to demonstrate their ability to:

- estimate optimum sample size from a plot of number of clasts sampled vs running mean of mean b-axis length.

Mathematical concepts

AS/A Level geology students will cover random and non-random (systematic) sampling.

In random sampling the positions at which samples are taken (when doing fieldwork) or more generally the sampling strategy used (when sampling in the lab or elsewhere) are intended to take an unbiased sub-set of the whole population (or the whole potential dataset).

In non-random sampling the positions at which samples are taken or the strategy for sampling are chosen by the investigator according to the question being investigated and some assumptions about the system being investigated.

The key thing to bear in mind when choosing how to perform sampling is the *purpose* of the investigation.

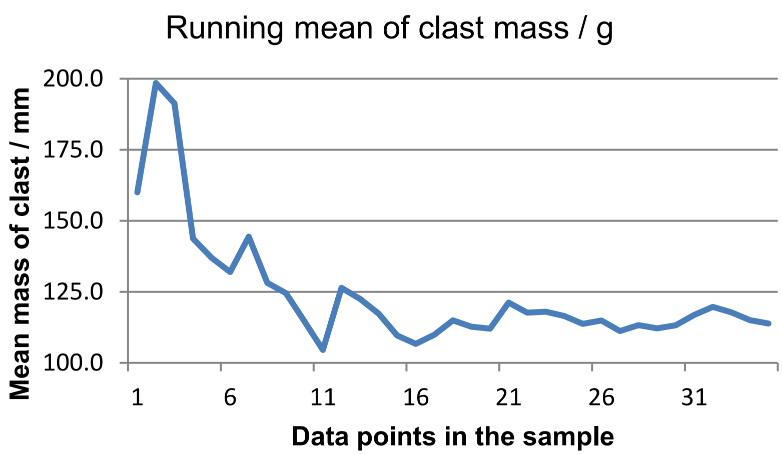
For example if the purpose of the work is to estimate the total number of a particular fossil on a large bedding plane a random sampling approach would be a natural choice, especially in an initial study where little was known about the distribution of that species. In this case positions for sampling should be chosen at random. A useful way to obtain this 'randomness' is to use as your coordinates for each sample, numbers from a random number generator function on a calculator or a list of random numbers.

On the other hand if the purpose of the work is to assess how distance from sediment source affects grain size a systematic sampling approach could be more appropriate with samples being taken at regular intervals moving further and further from the source of the sediment.

Contexts in geology

General scientific principles of sampling are not always easily applied to fieldwork. When planning fieldwork, it may cover a wide area where access is limited or difficult, due to terrain or land use. The exposure of outcrop needed to carry out the investigation is likely to be limited at times or variable. Therefore it may only be possible to sample by opportunity.

When estimating the optimum sample size students can use trial data to calculate the running mean and determine how many observations are required before the effect of extreme values on the sample mean is negligible. In the example below the optimum sample size is around 16.



If exposure of a bedding plane is good and a wide area can be covered then random number sampling can be used, a random number generator or tables can be used to direct where measurements should be made on that surface.

Imagine a 1 m^2 area of a bedding plane is to be sampled to count fossils. To sample without bias, the area needs to be split up into smaller squares of *equal* size (say for example 0.01 m^2 squares labelled from 00 to 99:

09	19	29	39	49	59	69	79	89	99
08	18	28	38	48	58	68	78	88	98
07	17	27	37	47	57	67	77	87	97
06	16	25	36	46	56	66	76	86	96
05	15	25	35	45	55	65	75	85	95
04	14	24	34	44	54	64	74	84	94
03	13	23	33	43	53	63	73	83	93
02	12	22	32	42	52	62	72	82	92
01	11	21	31	41	51	61	71	81	91
00	10	20	30	40	50	60	70	80	90

If we want to take 6 samples then we require 6 random numbers to eliminate the bias caused if we were to choose the squares to be sampled. Using the #RAN button on a calculator or using the '=randbetween(1,99)' function on a spreadsheet program these can be obtained. Below are six random numbers generated by Excel:

40, 58, 69, 91, 01, 15

Therefore samples should be taken from squares 40, 58, 69, 91, 01 and 15.

Random sampling is time consuming and exposures may not lend themselves to this approach therefore systematic sampling where samples are taken at regular intervals on a grid or on a transect along a bed are easier to apply. A common error when sampling coarse sediments is to introduce bias by picking a clast that fits in the hand or is in some other way attractive. It is important for students to measure the clast at the sampling point on the grid or transect determined by the random or systematic sampling procedure.

A clear sampling strategy is also important for laboratory based investigations. Microscope stages will sometimes have a mechanism to allow it to be moved in the x or y direction by set increments to allow for systematic sampling. For sediment samples care is needed to avoid bias both in the collection in the field and the analysis in the classroom. Bagged loose sediment is affected by a number of processes which tend to segregate it into layers; the Brazil nut effect, settling out of finer grains and shape sorting all modify the bagged sample. Cone and quartering can be used to recover a well sorted representative sample:

- weigh the bagged sample and determine what fraction is required for sieving
- pour the sample out onto a paper sheet and flatten the top of the cone
- using a glossy card or metal sheet cut the cone into four quarters
- select and mix together two opposite quarters then repeat the process until the required sample size is achieved. Return the other two quarters to the sample container.

https://www.911metallurgist.com/blog/laboratory-methods-of-sample-preparation

M2.6 Understand the terms mean, median and mode

Students should be able to demonstrate their ability to:

- calculate or compare the mean, median and mode of a set of linear data e.g. Folk and Ward graphic statistics from sieve analysis of sand samples from different sedimentary environments
- calculate (graphically) or compare vector mean, median and mode of a set of circular data e.g. palaeocurrent directions in an aeolian sandstone.

Mathematical concepts

The mean, median and mode are all measures of *central tendency* of a data set. They act as a representative value for the whole data set. These quantities are easy to calculate and should be well known to the students but there are some subtleties. The statistical functionality of the new calculators will generally find mean, median and mode together.

The list below represents some data:

$$25, 24, 27, 28, 19, 31, 25, 31$$

The arithmetic mean has been mentioned in *M2.1* and is the sum of the data values divided by the number of data values:

$$\bar{x} = \frac{210}{8} = 26.25$$

To find the median the data have to be reordered:

$$19, 24, 25, 25, 27, 28, 31, 31$$

The median is the *middle* value or in a formal manner the $\frac{n+1}{2}$th piece of data where *n* is the number of pieces of data. In this example there are 8 pieces of data and hence the median lies on the (8+1)/2 th piece of data which is the 4.5th piece of data. This doesn't really make sense until you realise that the 4.5 th data is halfway between the 4th and 5th items; 25 and 27 and hence the median (usually denoted by Q_2) is:

$$Q_2 = 26$$

The mode is the easiest to spot and is the most 'popular' item of data, the most frequent item. In the above example there is no single most frequent item of data with 25 and 31 both occurring twice. The data set is therefore *bi-modal* with modes 25 and 31.

As a general rule of thumb the mean is the most commonly quoted statistical measure because it uses all the items of data. If there are outliers however, then the median or mode is more representative because they are less sensitive to outliers. If for instance a data value of 100 was added to the example above, the mean would change to 34.4 but the median would move to 27, a far more representative measure, and the mode would remain the same.

Students of geology frequently collect data on the direction of geological structures, unlike the linear data discussed above, circular data cannot be treated in the same way. For example the mean of 340°, 355°, 005°, 015°, 015° is not the arithmetic mean (340+355+5+15+30)/5 = 149 but the vector mean 005°. Students are familiar with vector arithmetic from GCSE (9-1) Mathematic. Raw data plots (see *M2.11*) can be used to in a similar way to box and whisker plots or dispersion diagrams to identify mode, median and interquartile range.

Contexts in geology

Students could be asked to calculate and/or compare these quantities for any context where measurements are taken, such as a set of data collected on an outcrop of glacial till could be plotted as a cumulative frequency curve which can be used to quote median value and mean value. The mode is the most frequent outcome, therefore in this data set the group with the greatest percentage of results. The median is the middle value, therefore in the data set the group at 50% of results. Mean average is calculated using mid points of groups and frequency.

Grain size analysis of a conglomerate – glacial moraine

Grain size / mm	Mass / g	Mass / %	Cumulative frequency / %
0-1	0.1	0.03	0.03
1-5	5.6	1.77	1.80
5-10	10.8	3.41	5.20
10-20	12.3	3.88	9.08
20-30	14.6	4.60	13.69
30-40	12.8	4.04	17.72
40-50	30.2	9.52	27.25
50-60	50.4	15.89	43.14
60-70	110.0	**34.69**	77.83
70-80	70.3	22.17	100.00
Total	**317.1**	**100**	

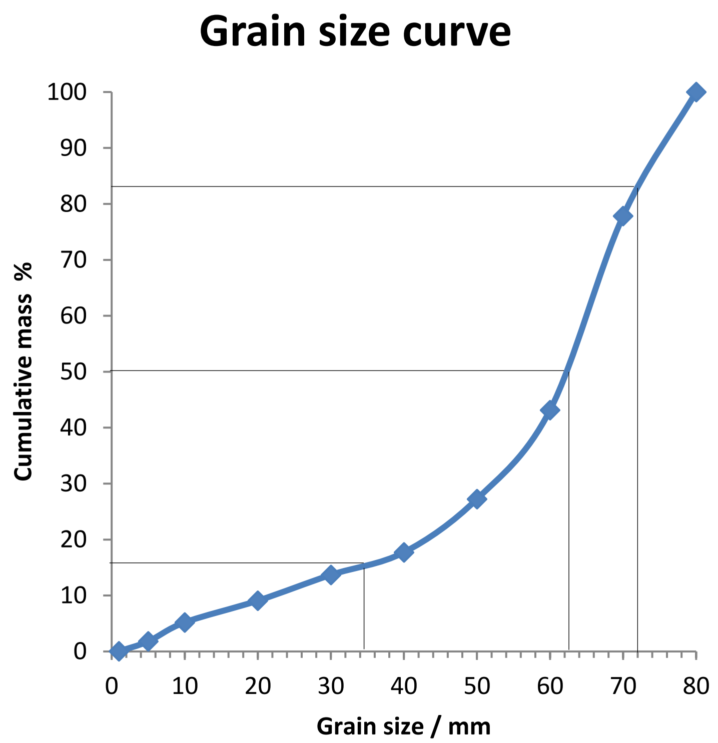
modal class

To calculate the grain size statistics it is necessary to first plot a cumulative frequency curve and then abstract the percentile values from the curve to calculate the Folk–Ward graphic statistics:

$$mean\ grain\ size = \frac{16\% + 50\% + 84\%}{3} = \frac{35 + 63 + 72}{3} = 57\ mm$$

$$median\ grain\ size = 50th\ percentile = 63\ mm$$

$$modal\ grain\ size\ class = 60\ to\ 70\ mm$$



M2.7 Know the characteristics of normal and skewed distributions

Students should be able to demonstrate their ability to:

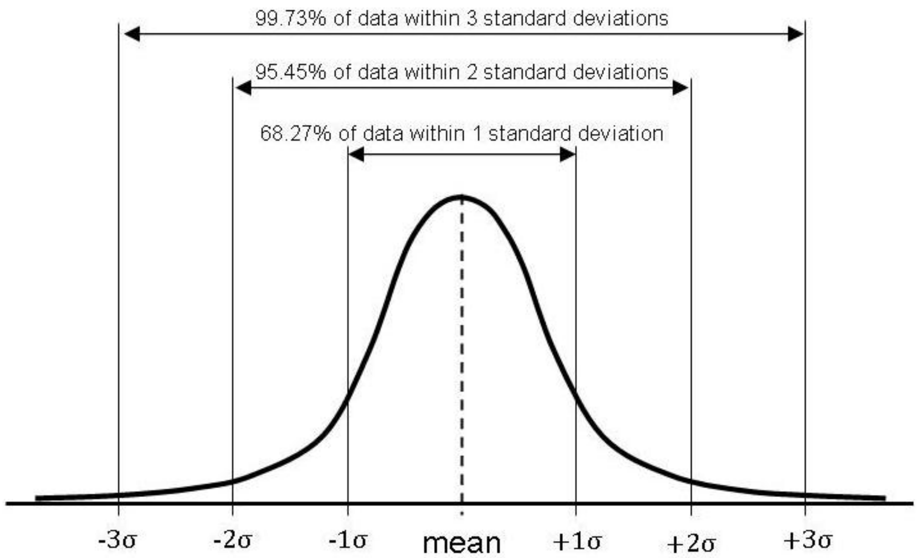
- being presented with a set of data for crystal size in an igneous intrusion and being asked to indicate the position of the mean (or median, or mode)
- interpret size analysis data from sieving of different sands.

Mathematical concepts

Most observations of natural systems are affected by multiple factors and contain random errors, however they can be described by a normal distribution. Even where the underlying distribution is not normal the means of multiple samples will form a normal distribution. For the shoe sizes of male or female students, class test scores, daily temperature readings are all normal distributions, particularly as the sample size increases.

The normal distribution is sometimes called a Gaussian distribution or a bell curve. When the mean and standard deviation of a sample is known the normal distribution can be described mathematically, so for example 68.27% of observations will be within one standard deviation of the mean, and 95.45% within two standard deviations of the mean.

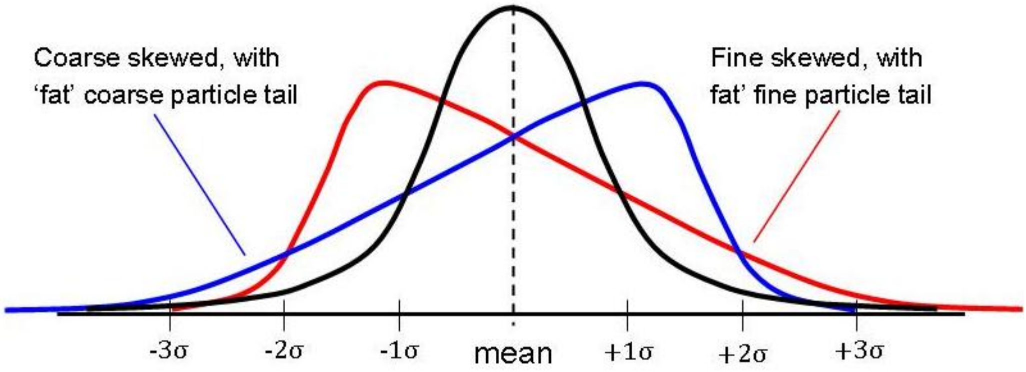
The compound effects of the minor variations in the multiple factors that affect our observations cause this randomness (see *M2.3*). Although we cannot predict what will happen at a given moment of time we can describe the probability that an individual observation forms part of the inherent variability in the data set.



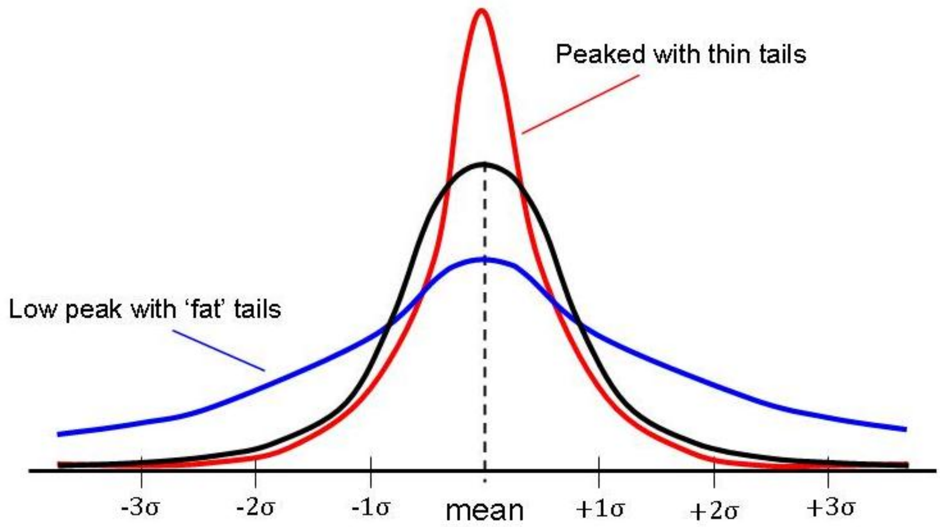
Contexts in geology

Students could be asked to compare grain size distribution curves for sediments deposited in different environments and sedimentary regimes. The effect of sorting and preferential removal of different sediments can be interpreted from the shape of the distribution curve. Only where sediment has been transported and deposited by a single process would normal distribution be expected, for example windblown dune sands.

Processes in glacial environments may deposit sediments with a coarsely skewed distribution while deep ocean sediments often display a finely skewed grain size distribution curve.



Some processes can result in very well sorted sediment with a peaked distribution curve while others such as subglacial lodgement till (boulder clay) are very poorly sorted with a low peak and broad tails.



M2.8 Understand measures of dispersion, including standard deviation and range

Students should be able to demonstrate their ability to:

- calculate the standard deviation
- understand why interquartile range might be a more useful measure of dispersion for a given set of data than standard deviation e.g. where there is an extreme observation which is part of the inherent variation

Mathematical concepts

Standard deviation is a quantitative measure of the spread of data about the mean. Students should be aware that there are different formulas for standard deviation. If using the statistical functionality on their calculators then they should recognise the notation used on their model in order to quote the correct value.

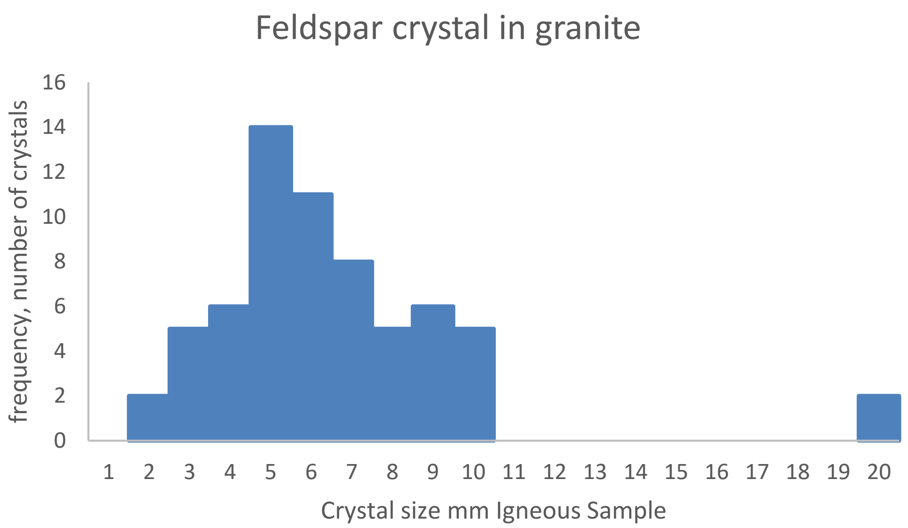
For a set of sample data, the formula for the standard deviation is:

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$

The $n - 1$ occurs because in a sample there are only $n - 1$ independent data values. Because in this calculation the mean has already been found, we are free to choose $n - 1$ items of data with complete freedom. The last item of data will be set because the mean is set. Take for example a set of data with mean 5 and $n = 4$ and we wish to choose 4 numbers with a mean of 5. We can choose the first one with freedom, say 6, the second with freedom say 2, and the third with freedom say 7. The final item of data cannot be chosen with freedom however as it is constrained by the fact the mean is 5. Therefore for the mean to be five the data have to add to 20; 2+7+6=15 and hence the final item of data has to be 5. Therefore we divide by $n - 1$ as this is the number of *degrees of freedom*.

Contexts in geology

When comparing dispersion of data students can compare the range (highest-lowest value) and the interquartile range The example below is based on a classroom exercise measuring feldspar crystals in a granite sample. The students were able to identify and measure 64 euhedral feldspar crystals between 1mm and 20mm long. Their results are displayed below in a histogram.



When describing the dispersion of data the students can calculate the range (highest-lowest values), the interquartile range or the sample standard deviation.

For the feldspar crystal data set below the range is 20-1 = 19, the extreme value of 20 has a significant effect. The interquartile range looks at the difference between the upper and lower quartile quartiles by calculating the difference between the result at 25% and 75%, for this data it is 8-5 = 3. This seems to reflect the data more clearly.

Students may need to calculate standard deviation when analysing field or experimental data. For the feldspar crystal data the following table shows the students' calculations.

Crystal size/ mm	Frequency / f	$x \times f$	$x - \bar{x}$	$(x - \bar{x})^2$	$f(x - \bar{x})^2$
1	0	0			
2	2	4	-4.55	20.67	41.35
3	5	15	-3.55	12.58	62.90
4	6	24	-2.55	6.49	38.92
5	14	70	-1.55	2.39	33.50
6	11	66	-0.55	0.30	3.29
7	8	56	0.45	0.21	1.64
8	5	40	1.45	2.11	10.56
9	6	54	2.45	6.02	36.11
10	5	50	3.45	11.92	59.62
11	0	0			
12	0	0			
13	0	0			
14	0	0			
15	0	0			
16	0	0			
17	0	0			
18	0	0			
19	0	0			
20	2	40	13.45	180.99	361.97
Total	**64**	**419**			**649.86**
	Mean	**6.5**		Standard Deviation	**3.19**

The mean crystal size is 6.5 mm and the standard deviation is ±3.2 for this set of results on crystal size.

- mean crystal size = 6.5 mm
- median crystal size = 6 mm
- range = 19 mm
- interquartile range = 3 mm
- standard deviation = 3.2 mm

For a normal distribution 68% of values are predicted to be within one standard deviation of the mean (3.3 to 9.7 mm) and 95% of the values within 2 standard deviations of the mean (0.1 to 16.1 mm). For this data set 78% of values are within one standard deviation of the mean and 97% within two standard deviations of the mean.

M2.9 Plot two variables from experimental or other linear data

Students should be able to demonstrate their ability to:

- select an appropriate format for presenting data, e.g. bar charts, histograms, line graphs and scattergrams.

Mathematical concepts

There are a number of different graphs that can be plotted for different types of data. If there is a single variable being collected then a bar chart or histogram is used (e.g. clast size, % silica). If the data that has been collected has two variables (e.g. time vs concentration, anterior–posterior length vs dorsal–ventral length) then a scatter graph/scatter-plot is used. A line of best-fit can be constructed to identify trends (usually with the independent variable as the x axis).

Plotting a graph should be a straightforward concept but the following guidelines are useful:

- Points plotted must be within 1 square of the correct value
- Appropriate linear scale used on axes
- Graph should make good use of available space
- Scales should be 'sensible', i.e. using decimal or otherwise straightforward scale
- Scales must be chosen so that all points fall within the graph area – points must not be plotted outside the graph area
- Axes must be labelled, with units included
- There should be an informative title.

If there is a trend in the data, a line of best fit (which may be a curve) should be drawn. A line of best fit must be drawn so that it achieves a balance of points above and below the line and minimises the distance of all data points from the line. Common errors are to draw a line which connects all the data points or to force the line through the origin when this is not supported by the data. In certain cases, such as rate–concentration, the lines of best fit will pass through the origin and students must assess when this is appropriate from the data.

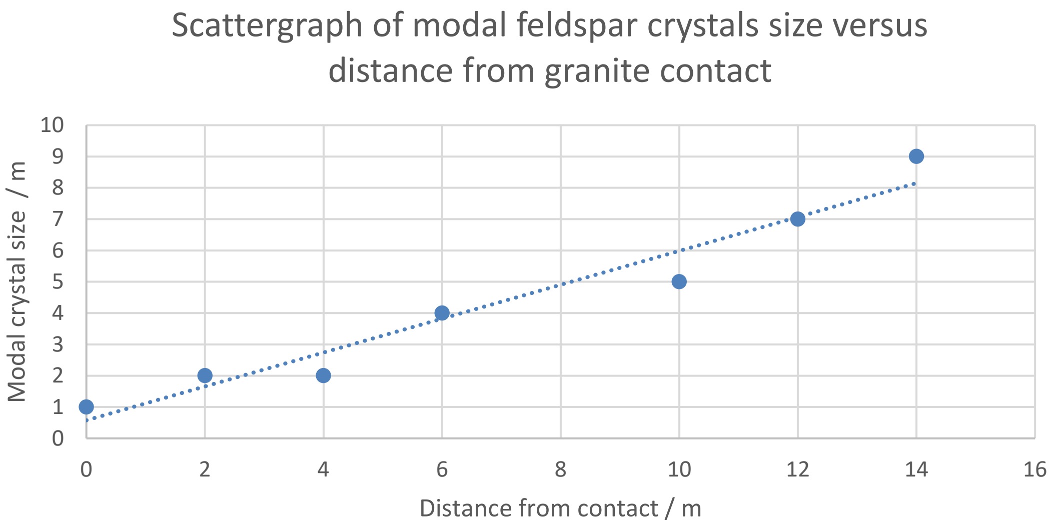
Extrapolation of a line of best fit is required in some instances, for example to determine the intercept with the *y*-axis. Students will only be asked to extrapolate linear graphs; extrapolation should be achieved by extending the line of best fit to the appropriate point. If broken axis are used the interpolation must not extend across the break in the axis.

Please see the Practical Skills Handbook for more on the topic of graphs.

Contexts in geology

There are many opportunities when students can plot graphs within this course in order to consider the relationship between two variables. For example the contact between an igneous intrusion and the country rock could be studied in the field, students could consider the relationship between distance from the contact and modal size of feldspar crystals.

Distance form contact / m	0	2	4	6	10	12	14
Modal feldspar crystal size / mm	1	2	2	4	5	7	9



When drawing the graph a suitable scale should be chosen to make full use of the page. Common errors are to squeeze a small graph onto the lefthand corner of sheet of A4 graph paper or draw the graph so large that axes, titles and labels have to be squeezed onto the page. As a rule of thumb a box enclosing all the plotted data points should contain at least 50% of the available plotting area of the graph, and the graph should be balanced within the available space on the page.

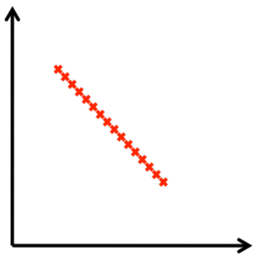
M2.10 Use a scatter diagram to identify a correlation between two variables

Students should be able to demonstrate their ability to:

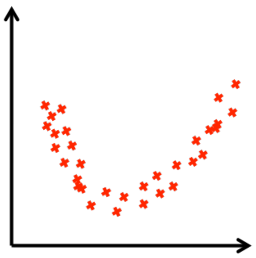
- interpret a scattergram, e.g. total length of subducting plate margin.

Mathematical concepts

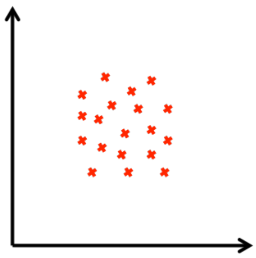
There are many types of correlation and in the absence of a mathematical procedure to calculate the correlation coefficient a lot of the interpretation comes down to judgement. The following graphs illustrate different types of correlation for two variables plotted against one another:



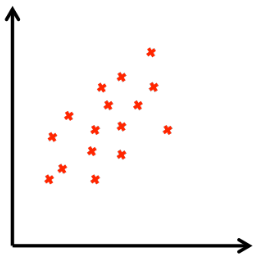
Perfect negative linear correlation



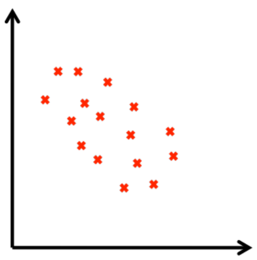
Quadratic correlation



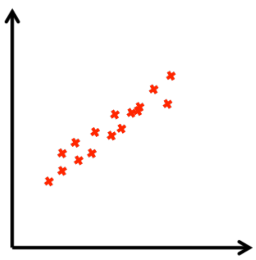
No correlation



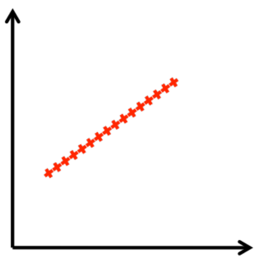
Weak positive linear correlation



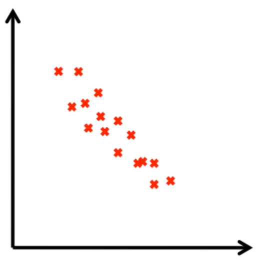
Weak negative linear correlation



Strong positive linear correlation



Perfect positive linear correlation



Strong negative linear correlation

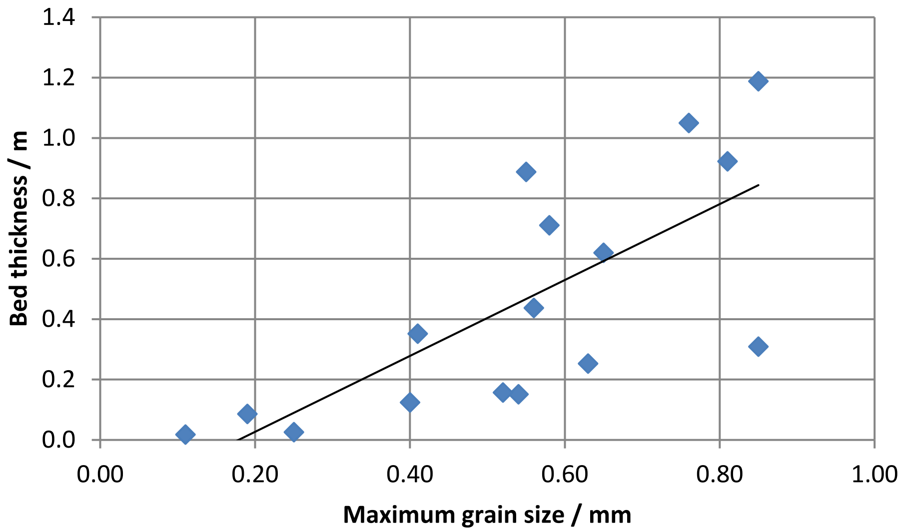
A number of points need to be made. From GCSE (9-1) Mathematics students will only be familiar with ***line of best fit*** when it is a straight line. However, just because the data is not in a straight-line does not mean that there is no correlation or ***line of best fit***. From the second diagram above it is obvious that there is a strong quadratic correlation but extremely weak *linear* correlation. For the most part, linear correlation is what we are looking for most of the time. If both sets of data are *normally distributed* then the shape of scatter will be approximately elliptical in shape and for most geological variables this will be the case. (Actually this is a condition for the correlation coefficient (see *M2.12*) to be used; that the data be normally distributed and in the absence of a formal mathematical test one can look at the data and see it is approximately an elliptical shape.)

It is very important that students appreciate that ***correlation does NOT necessarily imply causation***. Even if two variables display a high level of correlation it does not mean that there is a link between them. For example in the Western Canada Sedimentary Basin, if seismic events and the subcrop of Devonian fossil coral reefs are plotted on a map there is a strong spatial correlation. In fact the correlation is stronger than between hydrofracturing wells and seismic events for which there is a plausible mechanism. Mapping of the Swan Hills Formation reefs show that they grew on the edges of uplifted horsts, over deep basement faults (the likely cause of the seismicity), the point to make is that a strong correlation does not necessarily mean the variables are linked or dependent on one another. Coral reefs do not cause movement on faults, but both reefs and seismicity may be related to faulting but for different reasons.

Contexts in geology

Students could be asked to plot data and/or interpret a scattergram. The example used in M2.9 above shows the distance from the contact with the country rock and the modal size of feldspar crystals in the groundmass of a granite intrusion: This *appears* to show a positive linear correlation therefore one could make a conjecture that the greater the distance from the contact the larger the modal crystal size. However, this graph does not prove that there is a causal link between these two variables and so any probable causation would need to be justified

For coarse clastic rocks (such as conglomerates and turbidites) the casual observation is often made in the field that the largest clasts are found in the thickest beds. A student decided to test the validity of this hypothesis by measuring the largest grain in a sequence of turbidite beds exposed on the coast near Aberystwyth.



Having drawn a line of best fit the student could extrapolate the trend to predict the thickness of beds based on the maximum grain size in the bed or points chosen between existing points to interpolated intermediate data values. Students should be aware that extrapolation is only possible based on the assumption that the identified trend continue beyond the range of results recorded, and there are risks in making predictions if this is not the case.

The student has identified a linear trend and calculated the Spearman's "r_s" statistic to be 0.785 which indicates a strong positive correlation (M2.12). The student knows that turbidite beds are deposited by density currents in the deep sea so that a stronger density current would be likely both to carry more sediment and also be competent to carry larger clasts. However the student has not proved causation.

M2.11 Plot variables from experimental or other circular data

Students should be able to demonstrate their ability to:

- select an appropriate format for presenting data, raw data plot, circular bar graph, rose diagram and polar equal area stereonet (polar plots only not projections or great circles).

Mathematical concepts

Students of geology frequently collect data on direction of features. As discussed in *M2.6* above circular data cannot be treated in the same way as linear data as this is likely to lead to errors and spurious associations. Students are familiar with linear data and its treatment (e.g. arithmetical treatment, suitable graph types, summary statistics, and normal distributions). However while they are aware of circular data (e.g. azimuth, dip, 24 hour day, 365 day year) it is a common misconception that circular data can be treated in the same way as linear data. However circular data requires treatment using graphical or vector methods to calculate summary statistics and because the distribution about the mean may wrap around (i.e. 000°/360°) it is not always possible to apply the same statistical tests as for linear data.

Students are expected to be able to select and use appropriate graphs depending on the type of data available and the purpose of the display. A common misconception is to use a rose diagram for the display of any circular data, however because it does not display frequency density , rose diagrams are more suitable to qualitative data display, similar to the use of pie charts or bar graphs for displaying linear data.

There are a number of different graphs that can be plotted for different types of data (Appendix D). The following table compares the suitability and equivalence of different data display methods:

Numerical / continuous data	Categorical / discrete data	Linear data	Circular data
✓		Dispersion diagram	Raw data plot
✓	✓	Pie chart	Rose diagram
	✓	Bar graph	Rose diagram
✓		Histogram	Circular bar graph
✓		Scattergraph or scatter-plot	Stereonet (polar projection)

Plotting a graph should be a straightforward concept but the following guidelines are useful:

- Points plotted must be within 1 square of the correct value
- Appropriate linear scale used on axes
- Graph should make good use of available space
- Scales should be ‘sensible’, i.e. using decimal or otherwise straightforward scale
- Scales must be chosen so that all points fall within the graph and axes must be labelled, with units included
- There should be an informative title.

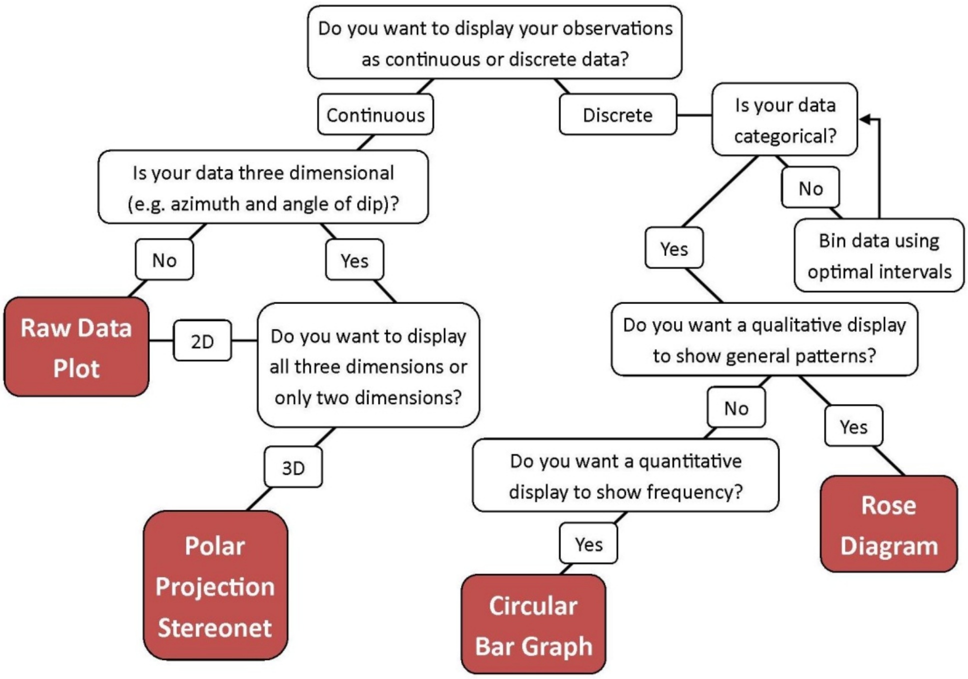
The stereonet used in GCE Geology is intended as a basic introduction to this graphical technique. Only the use of the equal area polar projection stereonet is required but it is important to appreciate that it is being used as a “circular scatterplot” and that the plotting of poles to structural measurements is not required and is beyond the scope of A level Geology. Suitable data would include azimuth and dip angle of foresets, bedding on folds, imbricated clasts, or azimuth and a linear quantity (e.g. height of cosets in herringbone ripples, depth of weathering on slopes).

Please see the Practical Skills Handbook for more on the topic of graphs.

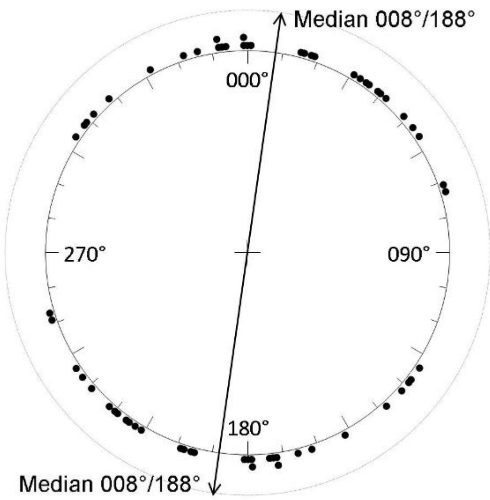
Contexts in geology

A large proportion of the data collected in fieldwork is directional and is therefore most appropriately plotted in circular graphs. It is important for students to be aware of the difference between ***orientation*** data (such as dip direction or imbricate structures) and ***trend*** data (such as the direction of strike, fault traces, dykes, clast axis), a useful teaching analogy is to compare the direction of the A1 trunkroad (***trend*** - north **and** south) with the River Severn (***orientation*** - north **to** south). By convention ***trend*** data is tabulated between 000° and 179° and ***orientation*** data between 000° and 359°. A common misconception when plotting ***trend*** data is to only plot the tabulated data but not the mirror image in the 180° to 359° sector.

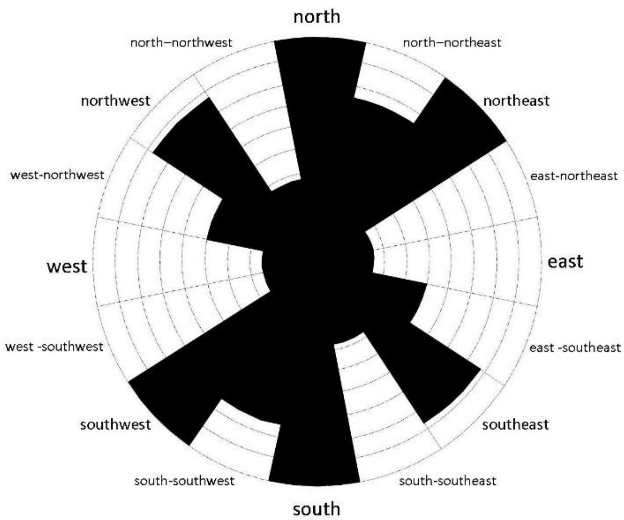
The following dichotomous key can be used to select an appropriate data display



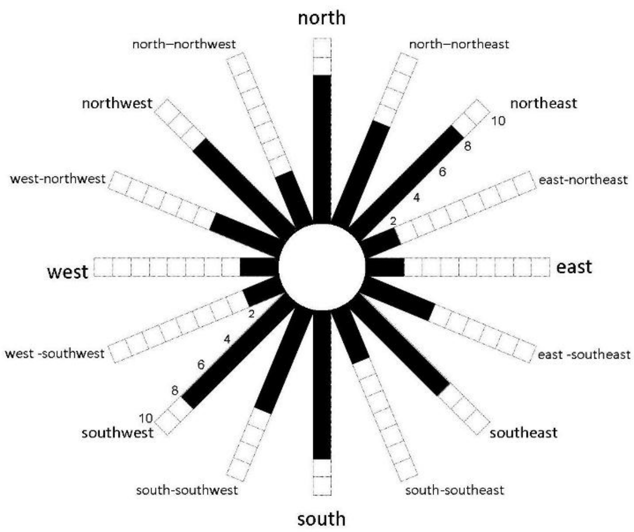
Examples of each data display type are shown below:



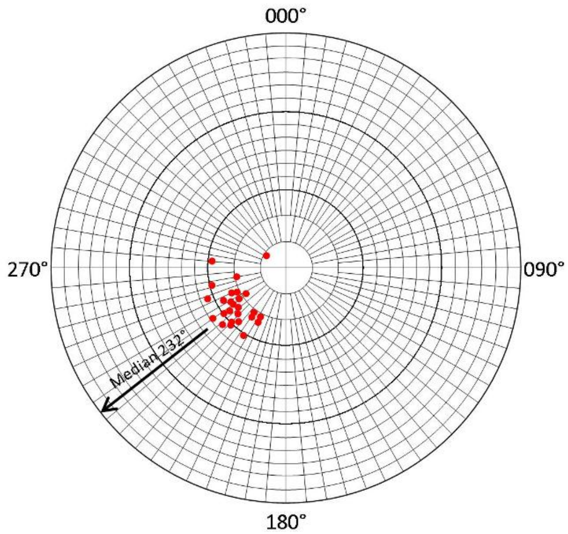
Raw data plot of phenocrysts long axis trends in a granite exposure. Median calculated graphically



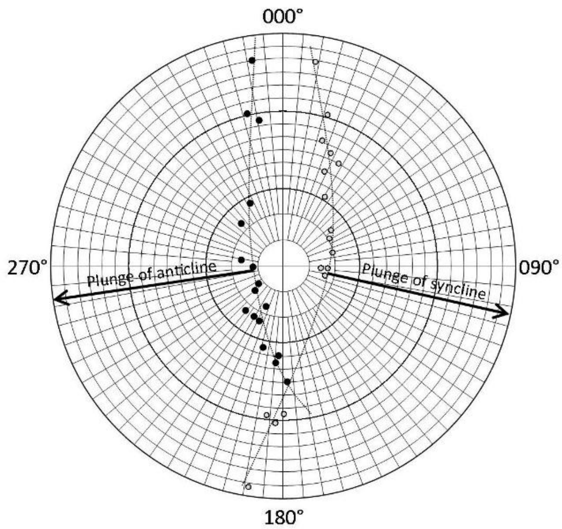
Rose diagram of phenocrysts long axis trends in a granite exposure from another student's data.



Circular bar graph displaying the same data as the rose diagram of phenocrysts long axis trends in a granite exposure.



Stereonet (*circular scatter graph*) of dip azimuths and dip angle of large aeolian dune foresets, Permian sandstone, Morayshire.



Stereonet (*circular scatter graph*) of dip azimuths and dip angle on two plunging folds in Pembrokeshire (Broadhaven and Stackpole Quay).

M2.12 Select and use a statistical test

Students should be able to demonstrate their ability to select and use:

- the chi squared test (χ^2) to test the significance of the difference between observed and expected results
- the Mann–Witney U test e.g. clast sizes in two conglomerate beds
- the Spearman's rank correlation coefficient.

Mathematical concepts

In the context of OCR AS/A level Geology the statistical tests are relatively few and it should be straightforward to select the appropriate test. (Critical values tables are available in Appendix D - Statistical Tables.)

A hypothesis test is generally used when comparing new data with research previously carried out or when comparing a sample with a theoretical distribution. The hypothesis is stated in terms of a parameter, and it is the way in which this parameter differs from the hypothesis that determines whether the test is one or two tailed. A one tailed hypothesis test is used if the data show a change in a particular direction, either as an increase or a decrease. A two tailed hypothesis test is used when we predict that the parameter is different, but we do not know whether it is larger or smaller.

In general students of geology use two tailed tests as in geology we are typically looking for a difference between data sets rather than predicting the direction (e.g. larger or smaller) that this difference may go. While single tailed tests are more powerful in discriminating between data sets they should only be used by students who have a clear understanding of the conditions where they are valid (i.e. students studying A level Further Maths, A Level Statistics or Level 3 Core Maths); the use of single tailed tests is not a requirement of the OCR GCE Geology specifications.

There are two possible outcomes for a hypothesis test, either it can be inferred that the evidence supports accepting the proposed alternative hypothesis (H_1), or there is insufficient evidence to reject the null hypothesis (H_0).

The chi squared (χ^2) test is used on data which has been observed given we know the expected values. In some cases our expected values are based on previous experimental or sample data. In the absence of such data, our expectation might simply be equal proportions in each category.

The Mann–Whitney *U*-test is used when comparing two smaller sets of the same type of data to see if there is a difference in the median and distribution of each of them. This is statistically equivalent to the Wilcoxon *rank sum* test but not with the similarly named Wilcoxon *signed rank sum* test. This test can only be carried out if the data has similar distributions but they are not required to be normal distributions. Alternatively the Student's *t*-test could be used when comparing larger sets of data to see if there is a difference in the mean of each of them. The OCR GCE Geology specifications will only cover Mann–Whitney *U*-test. Student's *t*-test is not a requirement.

The correlation test is used to see if two different variables are correlated in a linear fashion in the context of a scatter-graph via Spearman's rank correlation coefficient alternatively Pearson's Product Moment Correlation coefficient could be used. The OCR GCE Geology specifications will only cover Spearman's rank correlation coefficient. Pearson's Product Moment Correlation coefficient is not a requirement.

Contexts in geology

Chi squared test

Suppose that the orientation for phenocrysts in a granite (*M2.11*) are expected to be completely random, it would be expected on average that the phenocrysts would be aligned equally in all directions. If the orientations of a-axis of crystals were then measured this would be recorded as observed and the chi squared statistical test could be applied:

Phenocryst a-axis trend	Expected Frequency	Observed Frequency	$(E-O)$	$(E-O)^2$	$\frac{(E-O)^2}{E}$
North	5	8	-3	9	1.800
North Northeast	5	6	-1	1	0.200
Northeast	5	8	-3	9	1.800
East northeast	5	2	3	9	1.800
East	5	2	3	9	1.800
East southeast	5	4	1	1	0.200
Southeast	5	7	-2	4	0.800
South southeast	5	3	2	4	0.800
Total	**40**	**40**		**46**	**9.200**

The question to answer is 'is there a significant difference between the expected and the observed frequencies?' To answer this we calculate the χ^2 statistic which is given by the formula:

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

Simply, we find the difference between the observed frequencies and the expected, square them, divide by the expected frequency and then sum them for each trend direction:

$$\chi^2 = \frac{(8-5)^2}{5} + \frac{(6-5)^2}{5} + \frac{(8-5)^2}{5} + \frac{(2-5)^2}{5} + \frac{(2-5)^2}{5} + \frac{(4-5)^2}{5} + \frac{(7-5)^2}{5} + \frac{(3-5)^2}{5} = 9.20$$

To perform the test we need a significance level (what percentage probability are we happy to accept for there to be a mistake in our conclusion) and the number of degrees of freedom.

There are 8 trends (categories). The number of degrees of freedom is one less than the number of categories (see note on standard deviation M2.8) so in this case the degrees of freedom is 7. Choosing a significance level of 5% and looking up the critical value in the probability tables gives 14.07.

Our value of 9.20 is smaller than the critical value of 14.07 so we can say that there is no significant difference between the observed and the expected values. If the χ^2 statistic were larger than the critical value then we would have concluded that there is evidence to suggest that there is a significant difference between the observed and expected frequencies.

Man-Whitney *U*-test

The b-axis length of granite clasts at either end of a shingle beach on the Isle of Arran were measured. The two sites were 3 km apart and the samples were taken at the high tide mark. Twelve clasts were measured at Catacol (data set ***1***) and nine at Lochranza (data set ***2***). Longshore drift moves sediment towards Lochranza so it was predicted that the median and distribution of the samples would be different.

Catacol clast b-axis / cm		Lochranza clast b-axis / cm	
60	0	38	7
60	0	36	7
58	0	36	7
48	0	34	7.5
47	0	32	8
44	0	31	8
39	0	30	8
34	3.5	29	8
23	9	24	8
14	9	***U₂=***	**68.5**
9	9		
8	9		
U₁=	**39.5**		

The *U* statistic for Catacol (U_1) is calculated by counting the number of values in data set ***2*** (Lochranza) which exceed each of the values in data set ***1***. The *U* statistic for Lochranza (U_2) is calculated by repeat for all the values in data set **2**. If there is a tie between an observation in data set ***1*** and data set ***2*** then this is treated as a half value. An alternative method is to rank the values in the combined sample as described on the MEI (Maths in Education and Industry) website.

As $U_1 + U_2 = (n_1 \times n_2)$ there is a cross check that the *U* values are correct (i.e. 39.5+68.5 = 12×9).

The null hypothesis (H_0) is that the two samples are drawn from the same underlying population. The alternative hypothesis (H_1) is that the two samples come from populations with different medians.

This is a two-tailed test as the alternative hypothesis (H_1) is 'not equal' and it will be tested at the 5% probability level. Using the table below find the critical values for a significance level of 5% for a two-tailed test:

Size of the smaller sample	Size of the larger sample										
		5	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
	5	2	3	5	6	7	8	9	11	12	13
	6	3	5	6	8	10	11	13	14	16	17
	7	5	6	8	10	12	14	16	18	20	22
	8	5	8	10	13	15	17	19	22	24	26
	9	6	10	12	15	17	20	23	26	28	31
	10	7	11	14	17	20	23	26	29	33	36
	11	8	13	16	19	23	26	30	33	37	40
	12	9	14	18	22	26	29	33	37	41	45
	13	11	16	20	24	28	33	37	41	45	50
	14	13	17	22	26	31	36	40	45	50	55

For n_1 = 12 and n_2 = 9 the critical value for U is 26. As the smaller of *U* = 39.5 ≥ 26 hence the null hypothesis cannot be rejected; there is insufficient evidence to reject H_0. The samples come from the same population and the difference in median values is not significant at the 5% level.

Spearman's rank correlation coefficient test

For a sequence of seven conglomerate beds a student calculated the mean clast size. In order to use Spearman's rank correlation coefficient, first the data have to be *ranked*:

Bed thickness / cm	Rank x	Mean clast size / cm	Rank y
20	3	3	6
5	6.5	20	2
40	2	5	4
15	4	4	4
5	6.5	25	1
10	5	18	3
50	1	1	7

They are ranked from 1 to 7 with 1 being the highest and 7 being the lowest. Notice that for the bed thicknesses there are two 5 cm items of data occupying the 6th and 7th highest ranks. Therefore they are given the average rank of (6+7)/2=6.5.

Next find the difference between the ranks d and then square d^2:

Bed thickness / cm	Rank x	Mean clast size / cm	Rank y	d	d^2
20	3	3	6	3	9
5	6.5	20	2	4.5	20.25
40	2	5	4	2	4
15	4	4	4	0	0
5	6.5	25	1	5.5	30.25
10	5	18	3	2	4
50	1	1	7	6	36
				Total	**103.5**

The sum of $\sum d^2$ is 103.5. Finally use the formula:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6 \times 103.5}{7(7^2-1)} = 1 - \frac{621}{336} = -0.848$$

To test to see if this is significant the critical value has to be found.

The critical value for the two-tailed test (H_0: There is no correlation between x and y) at the 5% level for 7 pairs of data is 0.786.

As the test statistic $r_s = 0.848 > 0.786$ (we ignore the sign) then we have evidence to reject H_0 and accept H_1; therefore we can say, with 95% confidence, there is a negative correlation between bed thickness and the mean clast size.

M3 – Algebra and graphs

M3.1 Understand and use the symbols: =, <, <<, >>, >, $\propto$, ~

Students should be able to demonstrate their ability to:

- use these symbols appropriately and correctly in their given contexts
- understand these symbols in the contexts of formulae given.

Mathematical concepts

Students should have had exposure to the symbols = (equals), < (less than), ≤ (less than or equal to), << (much less than), >> (much greater than), ≥ (greater than or equal to), and > (greater than) from an early age, and should understand how and why they are used.

The symbol $\propto$ means 'is proportional to'. If two quantities *A* and *B* are *directly proportional* then the appropriate mathematical statement is,

$$A \propto B$$

If the two quantities are *inversely proportional* then the appropriate relationship is,

$$A \propto \frac{1}{B}$$

The symbol ~ means 'is roughly equal to' or 'of the same order'. This symbol may be used in the context of approximations made in calculations of quantities.

Students will be required to understand and use these symbols as they arise in various contexts. The more important aspect here is that students *understand* the symbols when they are used. When describing mathematical relationships, students would often be able to use descriptions in place of the symbols, for example stating that one value is directly proportional to another rather than giving the formal mathematical statement. Conversely, students might prefer to use symbols rather than descriptions for reasons of brevity. This is fine, but students must be sure to use the correct symbol.

Contexts in geology

Students will be expected to recognise these symbols in a range of geological contexts.

M3.2 Change the subject of an equation

Students should be able to demonstrate their ability to:

Use and manipulate equations, e.g. magnification.

Mathematical concepts

The most common equations to be rearranged can often be posed in the form of a formula triangle. This is where three quantities a, b and c are linked by the simple relationship:

$a = bc$

This is quite an easy equation to arrange for the other variables. For example dividing by b yields the formula for c:

$$c = \frac{a}{b}$$

whilst dividing by c will give the formula for b:

$$b = \frac{a}{c}$$

This mathematical principle should be familiar to students. However, even students who are comfortable with this area of mathematics may struggle with its application in the science classroom, because the equations are not presented in a familiar way. Students may need help to see that e.g.

mass = amount of substance × molar mass ($m = nM$)

is equivalent to

$a = bc$

and

$N = N_0 e^{-\lambda t}$ rearranges to $InN = -\lambda t + InN_0$ – see *M3.5*

is equivalent to

$y = ax + b$

You may wish to discuss this application of algebra with maths teachers in your school, to ensure you can approach this skill in a way that help students to make links between use of equations in science and what they have previously learnt in maths.

Contexts in geology

Students may need to recall and rearrange a geological formula within an assessment. For example in microscopy, to calculate magnification the equation is given by:

$$magnification = \frac{size\ of\ image}{size\ of\ real\ object}$$

This equation can be used to find the magnification given the size of the image and the size of the real object are known.

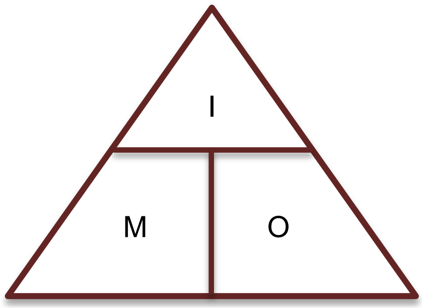
This can be rearranged to make the size of the real object the subject:

$$size\ of\ real\ object = \frac{size\ of\ image}{magnification}$$

or for the size of the image:

$$size\ of\ image = magnification \times size\ of\ real\ object$$

As mentioned above, students often find the easiest way to visualise the relationship between three quantities is to use a triangle, e.g.



It is often necessary for students to rearrange the spreading rate equation in a similar way depending on the data they are provided with.

$$seafloor\ spreading\ rate = \frac{distance\ between\ dated\ samples}{time\ difference\ between\ dated\ samples}$$

M3.3 Substitute numerical values into algebraic equations using appropriate units for physical quantities

Students should be able to demonstrate their ability to:

Use a given equation e.g. Darcy's law,

$$Q = -\kappa A\left(\frac{h_2 - h_1}{L}\right)$$

Mathematical concepts

Students should be aware of the principles of algebraic equations from GCSE Maths but a few misconceptions may remain. The most common problem is when dealing with powers and negative quantities in formulae.

The expression x^2 for example, whilst innocuous enough can cause issues when a negative number is substituted. Substituting $x = -2$ for example should be calculated as $(-2)^2 = 4$ not $-2^2 = -4$.

There can be confusion when substituting numbers of different signs. The expression 'two negatives make a positive' is often over-used. It is *always* true that two negatives multiplied/divided equal a positive, so

–3 × –5 = +15

whilst if only one of them is negative then the answer is negative

2 × –3 = –6

The over-use of this rule arises when addition is involved. For example:

–3 + –5

Students could think that because there are two negatives being added then they become positive and the answer is

3 + 5 = 8

Actually for addition the changes occur only when the signs are the same in the 'middle' of the sum. In the above sum one of the signs in the 'middle' is negative so it becomes negative. So the above example should actually be read as '–3 minus 5' which is –8. However, in the formula:

–3 – –5

Here there are two negatives in the 'middle' and there it becomes a plus. Hence the sum is '–3 plus 5' which is +2.

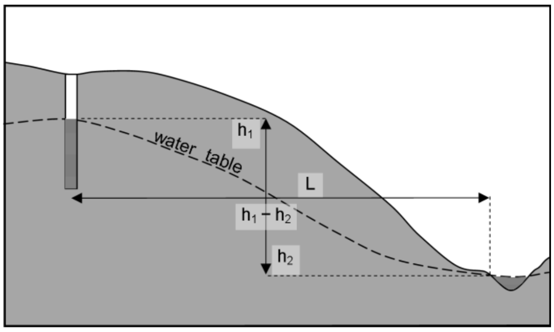
Additionally in the context of a formula such as $A = \pi r^2$ (for the area of a circle) students should be aware that the radius squared is being *multiplied* by pi despite the absence of a multiplication sign.

In general the laws of BIDMAS should be adhered to where the operations should be completed in the order of Brackets, Indices, Division, Multiplication, Addition and Subtraction.

Contexts in geology

Darcy's Law

$$Q = -\kappa A\left(\frac{h_2 - h_1}{L}\right)$$



Where L is the horizontal distance between points. h_1 and h_2 show the difference in height between two points, κ is the hydraulic conductivity and A is the area of the aquifer where the water is flowing.

To calculate Volume of water discharged in a period of time (Q), we substitute values into the Darcy's Law equation, so if:

L = 100 m

h_1 and h_2 = 160 m and 165 m

κ = 0.1 cm s^{-1} = 0.001 m s^{-1}

A = 200 m^2

$Q = -\kappa A\left(\frac{h_2-h_1}{L}\right) = -0.001 \times 200\left(\frac{165-160}{100}\right) = -0.200 \times 0.05 = 0.010$ m^3 s^{-1}

M3.4 Solve algebraic equations

Students should be able to demonstrate their ability to:

solve equations in a geological context, e.g.

$$\varphi = -\log_2\left(\frac{D}{D_0}\right)$$

Mathematical concepts

Solving an equation usually involves substituting values into a formula and realising that there is one unknown unaccounted for. Finding the value of this unknown is the same as solving the equation.

In order to calculate the unknown, it may be necessary to first rearrange the equation. Skills *M3.2* and *M3.3* are therefore often also needed in solving equations; indeed, the three skills are rarely encountered in isolation.

Take for example the formula:

- $E = U + pV$

If we were to substitute some values in for E, U and p, the formula becomes an equation for V; the only variable that remains unknown:

- $7 = 2 + 3V$

To find V we have to 'unlock' what is happening to V. By this we mean we have to 'undo' the operations that link V to the other numbers. First we subtract the '2' from both sides to get the 3V by 'itself':

- $5 = 3V$

To 'undo' the multiplication by 3 we divide by 3 and solve the equation:

$$V = \frac{5}{3} = 1.667$$

Contexts in geology

Sediment grain sizes are often given as phi (φ) sizes which is a logarithmic scale (*M3.5*). The phi size can be calculated using:

$$\varphi = -log_2 \frac{D}{D_0}$$

φ is the Krumbein phi scale

D is the diameter of the particle in mm.

D_0 is a reference diameter, equal to 1 mm (to make the equation dimensionally consistent.)

For example the mean diameter of a sieved sample was determined to be 2.2φ, what is the equivalent value in mm?

$$D = D_0 \times 2^{-\varphi} = 1 \times 2^{-2.2} = 0.2176$$

So $2.2\varphi = 0.22$ mm

M3.5 Use calculators to find and use power, exponential and logarithm functions

Students should be able to:

- Use a calculator to perform calculations involving powers of numbers, exponentials and logarithms (with different bases)
- solve for unknowns in radionuclide decay problems e.g.

$$N = N_0 e^{-\lambda t}$$

Mathematical concepts

The chief difficulty with calculating powers on a calculator is that different models have different ways of entering powers and using them. The students in your class will potentially own a wide range of calculator models. Different calculator symbols used include 'x^y', '10^x', '^' and 'exp'. It is worth taking the time to become familiar with the different models used in your class, and ensuring that students understand how to use these functions correctly. Also make sure they understand the different operations for e.g.:

3.6^3

and 3.6×10^3

Multiplications such as 4 × 4 × 4 can be written in power form as 4^3. The chief difficulty with calculating powers on a calculator is the potential wide range of models that students will own which all have different ways of entering powers and using them. For example different calculators will have " x^y ", "$\blacksquare^{\blacksquare}$" or "^" as the symbols for the power function. Students should be encouraged to find the appropriate button on their own model and learn how to use it effectively.

The exponential function has the common form e^x, where e is a mathematical constant (approximately 2.718). It is sometimes written as exp(x) and models relationships in which the rate of change of a quantity is dependent on the instantaneous amount of the quantity. On a calculator it appears as e^x or $e^{\blacksquare}$.

All students will have done the ground work for logarithms when they studied power and exponential in GCSE (9-1) Mathematics, students in A level geology only need to consolidate this prior learning. A logarithm is the reverse operation of raising a number to a power:

if $2^4 = 16$

then $log_2 16 = 4$

The chief difficulty for students using logarithms is that in different contexts different bases need to be used. A Level geology students will need to use logarithm operations for base 10, base 2 and base e. Students need to be aware that a log to base e is denoted by "ln" for some calculators and that in general the "log" button on a calculator will mean log to base 10.

Where students need to use log for any other base, and there is no "$\log_n$" button they can use this sequence on their calculator:

- $log_2(x) = log_{10}(x) \div log_{10}(2)$ or
- $log_2(x) = log_e(x) \div log_e(2)$ or
- $log_2(x) = In(x) \div In(2)$ or

Contexts in geology

Numerical age

Students can apply this skill when solving for unknowns in radionuclide decay problems.

A sample of bone from a mammoth in a glacial deposit contains 12% of the original carbon nuclides as an equivalent from a modern bone. Half life of $_{14}C$ is 5730 years, how old is the bone?

$$N = N_0 e^{-\lambda t}$$

represents an exponential decay with initial 'amount' N_0.

N = Number of atoms of parent isotope in the sample at time t (present)

λ = Decay constant

- $\frac{N}{N_0} = e^{-\lambda t}$

However as $\frac{N}{N_0}$ =0.12 (this is because $_{14}C$ remaining/ $_{14}C$ original x 100 = 12%) we can substitute

- $0.12 = e^{-\lambda t}$
- $ln 0.12 = -\lambda t$
- $t = -\frac{ln 0.12}{\lambda}$
- $\lambda = \frac{0.693}{t_{½}} = \frac{0.693}{5730}$ = 121 × 10^4 years
- $t = -\frac{ln 0.12}{1.21 \times 10^4} = 17527.5\ a$

Phi (φ) size

Students can apply this skill when converting between phi and mm or μm.

The mean grain size of an aquifer sandstone was calculated to be 175 μm based on measurements of a thin section using a microscope. To compare with modern sediment the grain size needs to be converted to φ units:

- $\varphi = -log_2 \frac{D}{D_0}$

D = 175 μm = 0.175 mm

D_0 = 1 (by convention)

- $\varphi = -log_2 \frac{0.175}{1}$
- $\varphi = -log_2 0.175$
- $\varphi = 2.514573 = 2.514$ (as the original measurement was to three s.f. see *M3.6*)

M3.6 Use logarithms in context with quantities that range over several orders of magnitude

Students should be able to:

- use a logarithmic scale in the context of geology e.g. decay law of radioactivity / Udden–Wentworth grain size scale

Mathematical concepts

Logarithms are basically powers. If we take the following calculation:

$10^2 = 100$

this can be expressed as

the power of 10 that gives 100 is 2

or in formal notation

$\log_{10} 100 = 2$

and we usually drop the 10 as it is assumed to be base 10 unless stated otherwise:

$\log 100 = 2$

Logarithms provide a better scale when dealing with quantities that vary exponentially (get big/small very quickly). For example, imagine sketching a graph where the scale goes from 10, 100, 1000, 10 000, 100 000 and so on. This would be impossible to do on a standard graph. Taking the logarithms of these quantities gives 1, 2, 3, 4, 5, which is far more manageable to handle and to spot trends.

The natural logarithm is denoted by $\ln x$, which is shorthand for $\log_e x$. Here e is the mathematical constant approximately equal to 2.7182818. This number is of central importance in mathematics, and often occurs in situations where quantities change exponentially over time. Like π, which students should be aware of, it is an irrational number, meaning it cannot be represented as a repeating decimal.

A note on significant figures

In numbers expressed as logarithms, the whole number represents the power of 10, and the decimal represents the value. So e.g. in the logarithmic number

2.86

the '2' represents the power of 10, and '.86' is the actual value. The whole number is thus *not* significant; the number above is given to 2 significant figures, not 3.

Given the result of a pH calculation

2.44977…

where the lowest number of significant figures in the data provided was 3, the final answer should be given as

2.450 (3 significant figures)

Contexts in geology

Logarithmic scales allow students to easily compare quantities that vary over a wide range of magnitude. The Moment Magnitude scale is one example of a logarithmic scale where for each whole step increase in magnitude corresponds to a $10^{1.5}$ increase in the energy released by the earthquake. Power's roundness scale (i.e. very angular/1 to well-rounded/6), although it is usually thought of as a categorical scale, is also a logarithmic scale where:

$$Powers\ roundness = 6^{\left(\frac{mean\ radius\ of\ corners}{radius\ of\ inscribed\ circle}\right)}$$

(Note: not OCR GCE Geology specification content)

The most common context in which students used logarithms is the Udden–Wentworth grain size scale particularly when used with φ (phi) units. Without using the φ scale it would be very difficult to compare and analyse graphically sediment grains which can range over seven orders of magnitude from microscopic clay mineral grains (<0.00004 mm) to boulders (>256mm).

<table>
<tr><th colspan="4">Clast diameter (b-axis)</th><th rowspan="2">Udden-Wentworth grade</th><th rowspan="2">Udden-Wentworth class</th><th rowspan="2" colspan="2">Rock type</th></tr>
<tr><th>φ</th><th>m</th><th>mm</th><th>µm</th></tr>
<tr><td>-12</td><td>4.1</td><td>4096</td><td></td><td colspan="2">megaclast</td><td colspan="2">megaclastic rock</td></tr>
<tr><td>-8</td><td>2.5</td><td>256</td><td></td><td rowspan="4">gravel</td><td>boulder</td><td rowspan="4" colspan="2">conglomerate
or
breccia</td></tr>
<tr><td>-6</td><td>0.6</td><td>64</td><td></td><td>cobble</td></tr>
<tr><td>-2</td><td>0.04</td><td>4</td><td></td><td>pebble</td></tr>
<tr><td>-1</td><td>0.02</td><td>2</td><td>2000</td><td>granule</td></tr>
<tr><td>0</td><td>0.01</td><td>1.0</td><td>1000</td><td rowspan="5">sand</td><td>very course sand</td><td rowspan="5" colspan="2">sandstone</td></tr>
<tr><td>1</td><td></td><td>0.5</td><td>500</td><td>coarse sand</td></tr>
<tr><td>2</td><td></td><td>0.25</td><td>250</td><td>medium sand</td></tr>
<tr><td>3</td><td></td><td>0.125</td><td>125</td><td>fine sand</td></tr>
<tr><td>4</td><td></td><td>0.063</td><td>62.5</td><td>very fine sand</td></tr>
<tr><td>5</td><td></td><td>0.032</td><td>31.3</td><td rowspan="4">silt</td><td>coarse silt</td><td rowspan="4">siltstone
or
shale</td><td rowspan="5">mudstone or shale</td></tr>
<tr><td>6</td><td></td><td></td><td>15.6</td><td>medium silt</td></tr>
<tr><td>7</td><td></td><td></td><td>7.8</td><td>fine silt</td></tr>
<tr><td>8</td><td></td><td></td><td>3.9</td><td>very fine silt</td></tr>
<tr><td>14</td><td></td><td></td><td>0.06</td><td>clay</td><td>clay</td><td>claystone</td></tr>
</table>

M3.7 Translate information between graphical, numerical and algebraic forms

Students should be able to demonstrate their ability to:

- understand that data may be presented in a number of formats and be able to use these data, e.g. time distance curves for earthquakes.

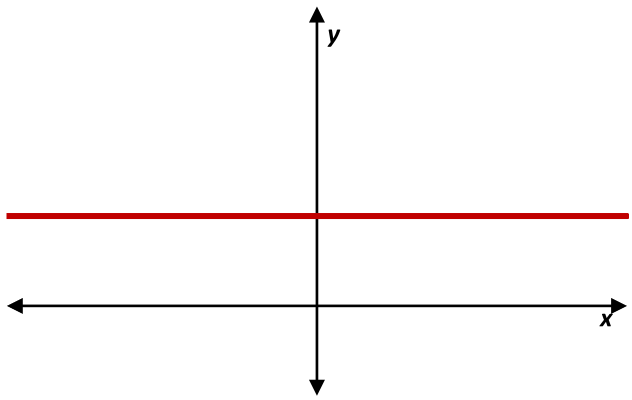
Mathematical concepts

There are several situations in GCE Geology where data may be presented graphically. Students should be familiar with the types of graphical representations used, the conventions for variables used on the graph axes, and how to interpret the information provided in the graph.

Students will need to be able to read co-ordinates for points on graphs (both *x*- and *y*-co-ordinates).

In geology, students do not need an elaborate understanding of how to convert graphs into algebraic equations, but there are a few instances where it is useful if students can judge the relationship between the plotted variables from the shape of the graph.

The following general graphs are useful for students to know and recognise.



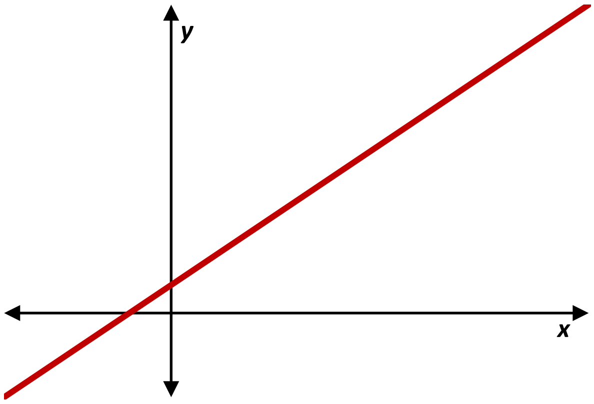
A graph showing a horizontal line parallel to the *x*-axis shows that the variable plotted on the *y*-axis is independent of the variable plotted on the *x*-axis. In mathematical terms, this relationship can be expressed as

$$y \propto x^0$$

or

$$y = \text{constant}$$

Where the constant in question is given by the *y*-axis value of any point on the line.



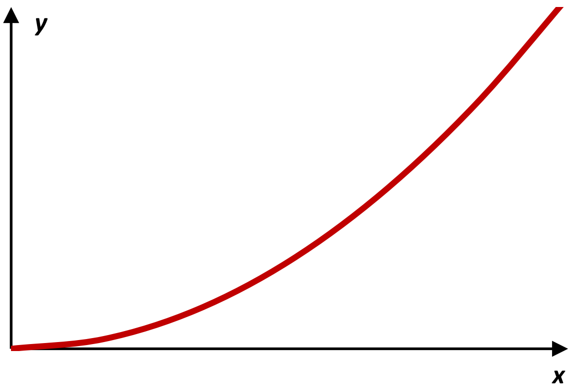
A non-horizontal straight line shows that the variable plotted on the *y*-axis is proportional to the variable plotted on the *x*-axis, or

$y \propto x$

In mathematical terms, this graph can be expressed as:

$y = mx + c$

m is the gradient of the graph, and *c* is the value of the intercept on the *y*-axis (link to *M3.3*).



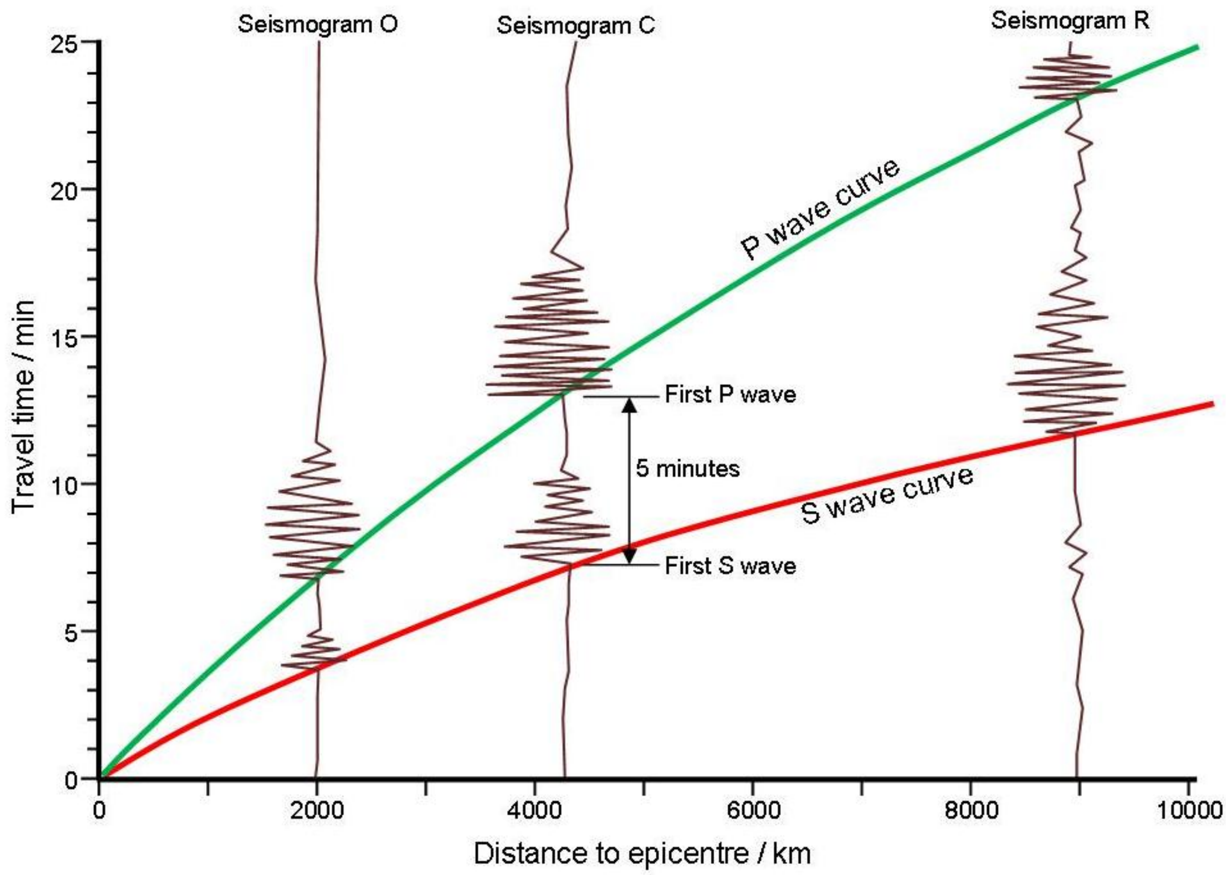
A curved graph that passes through the origin indicates that the variable plotted on the *y*-axis is proportional to a power of the variable plotted on the *x*-axis that is greater than 1, or

$y \propto x^n$ where $n > 1$

Students will not be required to determine the exact mathematical relationship of such graphs.

Contexts in geology

Students must understand that data can be presented in a number of formats and be able to use these to gain information and evidence.



The example above shows how seismograph traces are used to construct a graph displaying distance to the epicentre of an earthquake based on the difference in the travel time for the first P wave and S wave observed on the seismograms recorded at three different locations.

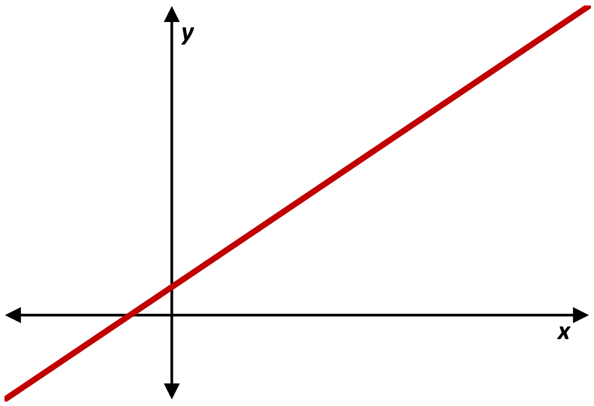
M3.8 Understand that $y = mx + c$ represents a linear relationship

Students should be able to demonstrate their ability to:

- Predict/sketch the shape of a graph with a linear relationship, e.g. burial curves in a sedimentary basin or the effect of intrusion size on the width of the baked margin.

Mathematical concepts

This concept is a two-way process. As discussed in *M3.1*, students should be able to work out a linear relationship given a graph and be able to sketch a graph given a linear relationship.



A non-horizontal straight line shows that the variable plotted on the y-axis is proportional to the variable plotted on the x-axis, or

$$y \propto x$$

In mathematical terms, this graph can be expressed as:

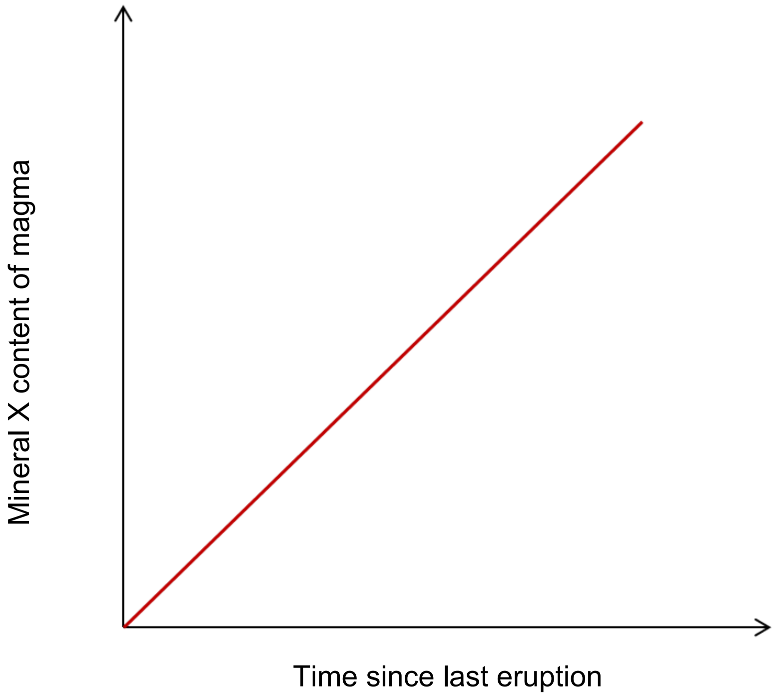
$$y = mx + c$$

m is the gradient of the graph, and c is the value of the intercept on the y-axis (see *M3.9*).

Students should understand that a positive m represents a line going ‘up’ from left to right and a negative m represents a line going ‘down’ from left to right. When sketching the graph students should always start from the y-intercept and then use the gradient to determine another point on the graph. Once this extra point has been determined the line can be drawn as it is only necessary for two points to be known to sketch a straight line.

Contexts in geology

Following many opportunities in fieldwork or experimental investigations students will be able to plot scatter graphs and may find linear relationships between variables. Using the equation above students can work out the equation of the line, m represents the gradient which can be measured on the graph and c is where the graph intercepts the y axis. Once the equation is constructed students can use it to predict y values for any x and vice versa.



This is an example of a straight-line graph and hence will have the form $y = mx + c$. In fact because it goes through the origin, then the y intercept is 0 and will be of the form:

$$Mineral\ X\ content\ of\ magma = m \times time$$

where *m* is the gradient of the line (see M3.9 and *M3.10*).

Magnitude Moment scale

The equation to calculate the Magnitude Moment of an earthquake takes the form of the equation for a straight line, although the exact equation used will depend on how the magnitude of the earthquake was measured. For example:

$$M_w = {}^{2}/_{3}\, logE - 6.1$$

M_w is the Moment Magnitude which is dimensionless

E is the energy released by the earthquake in Joules

An earthquake of magnitude 4.7 occurred in the North Sea on 30 June 2017. How much energy was released by the earthquake? – see *M3.5* for log rules.

$$log_{10}E = \frac{3(6.1 + M_w)}{2} = \frac{3(6.1 + 4.7)}{2} = \frac{32.4}{2} = 16.2$$

$$E = 10^{16.2} = 1.59 \times 10^{16} = 1.59 \times 10^{7}\ gigajoules$$

Note: see also *M3.4, M3.5* and *M3.6* for treatment of logarithm and power.

M3.9 Determine the slope and intercept of a linear graph

Students should be able to:

- read off an intercept point from a graph e.g. the initial velocity of a velocity time graph for a density current
- find the gradient (slope) of a linear graph.

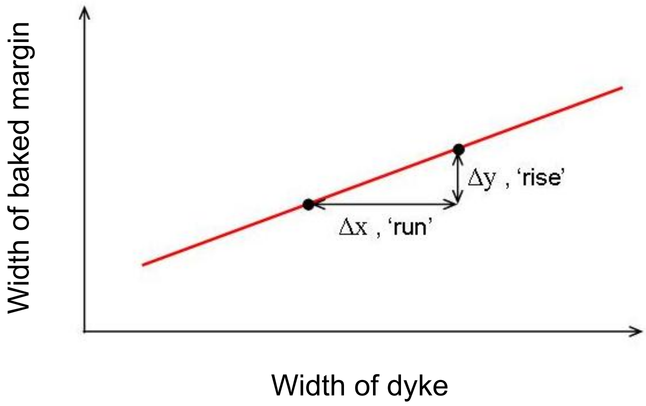
Mathematical concepts

The straightforward way to find the *y*-intercept is to examine where the line of the graph crosses the *y*-axis. However, this must be done at the point where $x = 0$ – this point is not always visible on an appropriately drawn graph (see Section M3.7 for comments on axis scales). In such cases, the *y*-intercept can be determined mathematically.

To find the gradient the following formula is useful:

$$\text{gradient} = \frac{\text{rise}}{\text{run}} \left(= \frac{\Delta y}{\Delta x} \right)$$

The 'rise' represents the vertical step between two points, and the 'run' represents the horizontal step between the same two points. Both of these quantities could be negative and care has to be taken in these cases.



The principle is that two points on the line of best fit are selected. Measuring the horizontal distance between the points gives the run, and the vertical distance gives the rise. The division according to the formula above gives the gradient.

When determining the gradient of a graph plotted from fieldwork/experimental data, the points used must be on the line of best fit. Students should **not** use the plotted raw data to determine the gradient (unless the plotted data falls exactly on the line of best fit).

Contexts in geology

See *M3.8* above.

M3.10 Calculate rate of change from a graph showing a linear relationship

Students should be able to demonstrate their ability to:

- calculate a rate from a graph e.g. geothermal gradient through the lithosphere.

Mathematical concepts

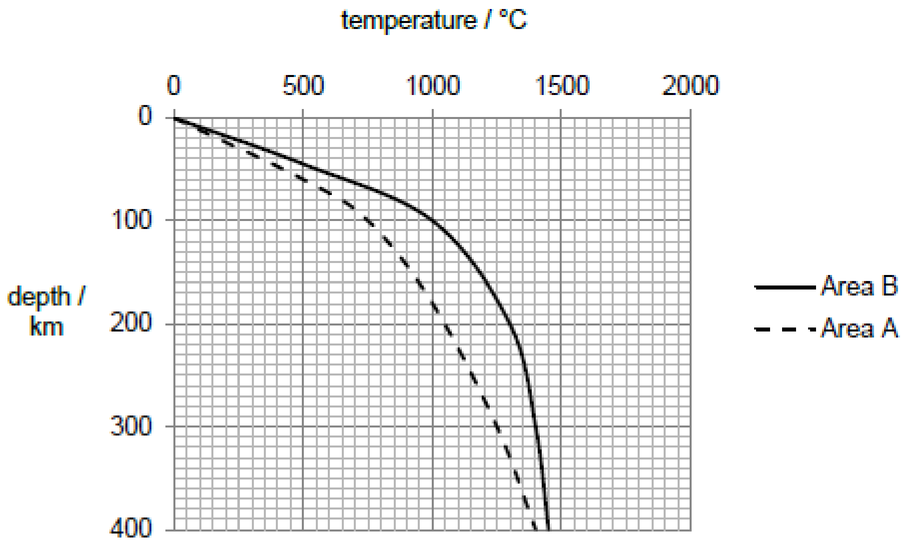
The gradient of a linear graph is always a measure of the rate of change between the two variables. The gradient of a linear graph with formula $y = mx + c$ basically measures the rate of change of y with respect to x. In words, the gradient expresses how quickly y changes as x changes.

A positive gradient means a quantity that increases as x increases whilst a negative gradient is a decreasing quantity as x increases.

The procedure for calculating a rate of change is therefore the same as for calculating the slope of a graph, as described in Section M3.9.

Contexts in geology

A student could be asked to calculate the geothermal gradient (a rate) from a linear portion of a geotherm curve, for example:



The geothermal gradient can be estimated in Area B between 230km and 400km depth:

$$Geothermal\ gradient = \frac{change\ in\ temperature}{change\ in\ depth} = \frac{200}{170} = 1.2°C km^{-1}$$

M3.11 Interpret logarithmic plots

Mathematical Concepts

Students should be able to demonstrate their ability to:

- use logarithmic plots with decay law of radioactivity

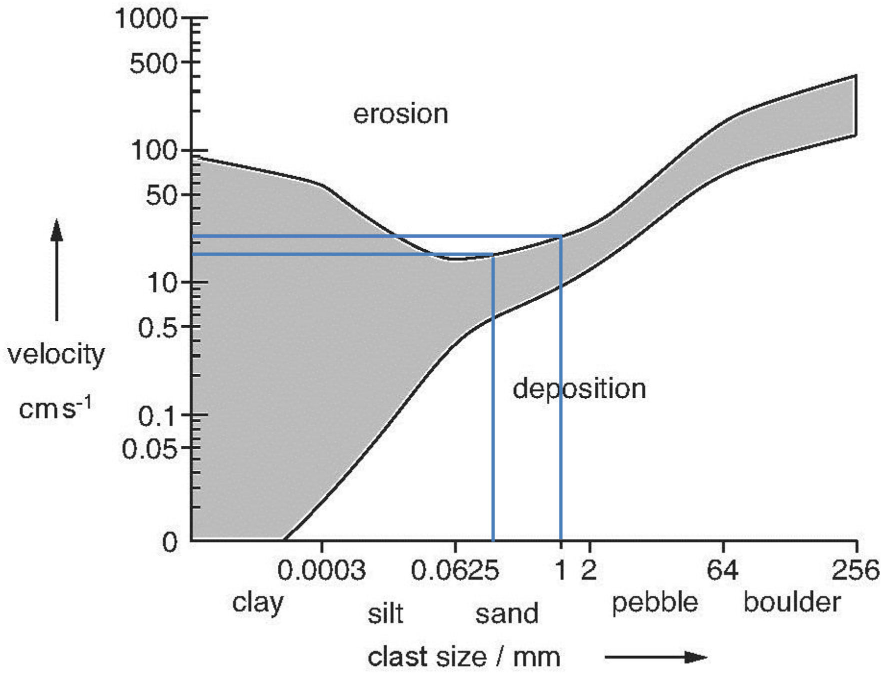
From a set of data a mathematical relationship that links the two variables is to be found. The two models that are tested in this context are:

- $y = ka^x$
- $y = kx^n$

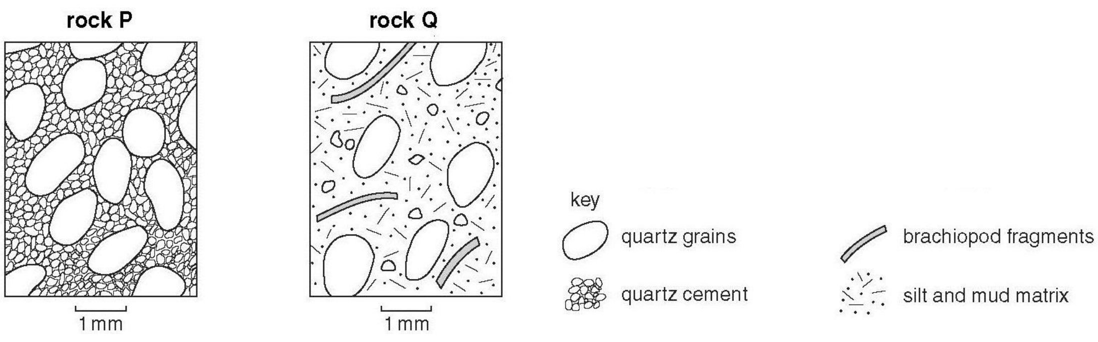
Notice in the first model the *x* variable is in the power and this is an example of an exponential function. The second model has the *x* variable as the base and this is a Power function.

Contexts in Geology

Students could be asked to interpret a Hjulström diagram in which both axis are drawn using base 10 logarithmic scales. The graph below shows the current velocity needed erode sediment particles and maintain them in transport.



What would the range of velocities at which the sand grains in rock Q were first transported? The sand grains would be eroded and start moving when the current velocity exceeded 17 cm s^{-1} (0.2 mm) to 22 cm s^{-1} (1 mm).



Alternatively students could be asked to interpret a radioactivity decay curve.

M4 – Geometry and measures

M4.1 Calculate the circumferences, surface areas and volumes of regular shapes

Students should be able to demonstrate their ability to:

- calculate the circumference and area of a circle
- calculate the surface area and volume of rectangular prisms, of cylindrical prisms and of spheres e.g. calculate the surface area or volume of a longwall panel.

Mathematical concepts

These calculations will have been covered during GCSE maths, and students need to be aware that they could come up in the context of an AS/A Level Geology assessment. The formulae will not be provided for these calculations.

The following list of formulae will therefore need to be recalled: students are expected to be able to use the following list of formulae:

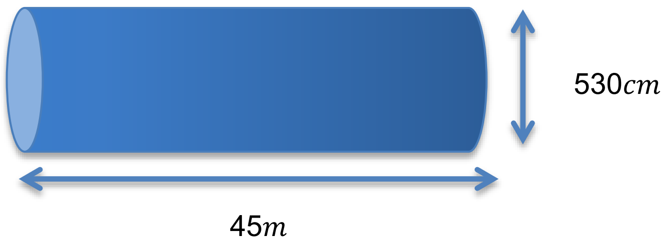
- Area of a right-angles triangle $\frac{1}{2} base \times height = \frac{1}{2} bh$
- Area of a triangle $\frac{1}{2} ab \sin C$, a and b are sides containing the interior angle C
- Circumference of a circle $2\pi r$ with r the radius
- Area of a circle πr^2
- Surface area of a cuboid $2(bh + bl + hl)$, b is base, l is length and h is height
- Surface area of a right prism (including cylinder) $\Sigma\,(area\ of\ each\ face)$
- Volume of a cuboid hbl
- Volume of a right prism (including cylinder) $area\ of\ cross\ section\ \times height$

The following list of formulae will need to used but not recalled: students will be given the formulae in the exam, or in a list from which they select and apply as appropriate:

- Surface area of a sphere $4\pi r^2$,
- Volume of a sphere $\frac{4}{3}\pi r^3$

Contexts in geology

Students could be asked to consider a scenario where a corroded pipe has leaked an organic solvent contaminating a volume of soil at a brown field site which is being redeveloped. The zone of contamination is approximately a cylinder with length 45 m and diameter 530 cm



The volume and surface area of the zone of contamination need to be calculated in order to decide whether to mediate the hazard by excavating the contaminated soil and treating it off site, or aerating the soil to encourage microbial oxidation of the pollutant in place.

The volume is calculated as:

$$V = \pi r^2 l = \pi \times 2.65^2 \times 45 = 993\ m^3$$

The surface area is calculated as:

$$S.A = 2\pi r(r + l) = 2\pi \times 2.65 \times (2.65 + 45) = 793\ m^2$$

Based on these figures and applying them to information on the proposed mediation techniques, a student could calculate that:

- a minimum of 34 truckloads of contaminated soil would have to be removed
- bioremediation could be completed in 16 weeks if aeration lances are used to encourage optimum microbial action in the soil.

M4.2 Visualise and represent 2-D and 3-D forms including 2-D representations of 3-D objects

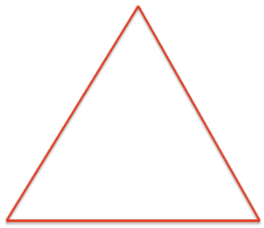
Students should be able to demonstrate their ability to:

- draw geological cross-sections interpreted from geological maps
- interpret block diagrams to show geological structures in 3D
- interpret field exposures and record 3D geological structures using field sketches.

Mathematical concepts

Symmetry is a notoriously difficult topic to teach. It relies entirely on a student's spatial awareness and reasoning. There are a few tricks that can help students to improve. Symmetry is a measure of the ability of a shape to be 'messed around' with but still keep its essential structure the same.

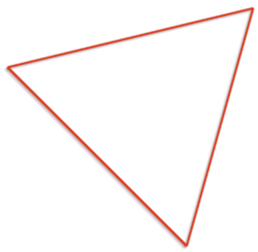
Take an equilateral triangle:



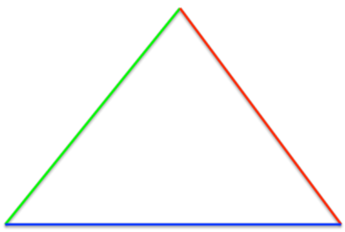
It displays a number of forms of symmetry.

- It has *rotational* symmetry: if you rotate it by 120° or 240° it looks the same.
- It has *mirror* symmetry: if you flip it about the vertical axis (or either of the other planes of symmetry) it looks the same.

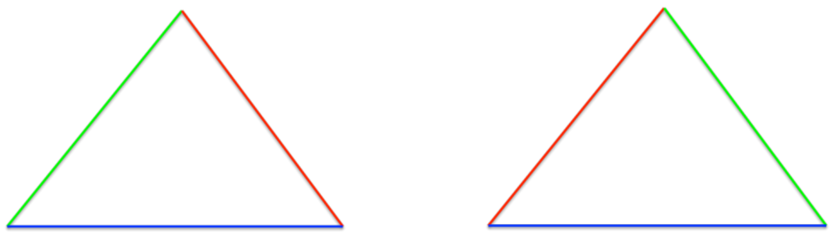
In the following diagram the triangle has been rotated to a different position, but its structure is still the same. It is identical to the first triangle, only its position is different. It can be rotated to make it look the same as the triangle above.



The following triangle is not symmetrical.



It cannot be rotated (unless by 360°, which takes it back to the starting point) or reflected to make it look the same. This has an effect on the nature of the mirror image of this triangle:



These triangles have the same position and structure. The colours are the same and in the same positions relative to each other. But the triangle on the right cannot be rotated in such a way as to produce the triangle on the left. The mirror images are *non-superimposable*.

In A Level Geology, students need to apply these principles in 3 dimensions. If the 3-D shape of a molecule lacks symmetry, then its mirror image cannot be rotated so that it will look the same as the original (see below).

Looking at 3-D models can help to grasp the principle that molecules can be mirror images, but cannot be rotated so that they look the same. From there, students must apply the principle to 2-D representations of molecules, bearing in mind that the 2-D representation may 'mask' any lack of symmetry.

Contexts in Geology

There are many situations where students must consider how features appear in 2D and 3D and be able to construct and transfer information between different forms. Basic problem maps are likely to be the starting point here before considering geological maps of a field location. Horizontal strata will appear as a simple outcrop pattern when the topography is relatively flat, undulating topography on the other hand will show a complex geological pattern in outcrop for horizontal bedding. Vertical structures are not affected by the land contours and will appear as a straight line across the map. All other rock layers must be dipping.

For dipping beds the dip direction is at 90°to the strike. Any angle measured at less the 90° to the strike is an apparent dip. Structure contours can be drawn parallel to the strike line. If a bed crosses the same topographic contour twice the two points can be drawn to make a structure contour. Steepest dipping beds have the narrowest outcrops and thicker beds will have the widest outcrops. When drawing maps and cross sections beds are considered to have a constant thickness.

Folding is identified by rocks dipping in different directions on a map along with repetition of strata. An antiform is identified by strata dipping away from the axis whereas a synform is identified by rocks dipping towards the axis. If the axis of a fold is horizontal the limbs will show parallel outcrop on a map.

Faulting can be identified by a repetition of bedding which show the same dip amount or there may be some strata missing. The downthrow side of the fault is identified by finding where younger rocks are in contact with older.

M4.3 Use sin, cos and tan in physical problems

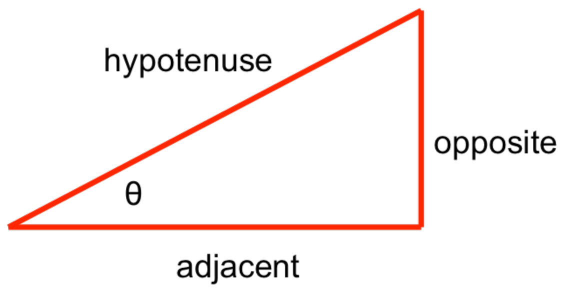
Students should be able to demonstrate their ability to:

- determine true thickness of rock units
- interpret block diagrams to show geological structures in 3D
- crustal extension or shortening.

Mathematical concepts

- Know and apply the trigonometrical ratios in relation to right-angled triangles
- Know and apply the sine and cosine rule for non-right angled triangles

For a right-angled triangle



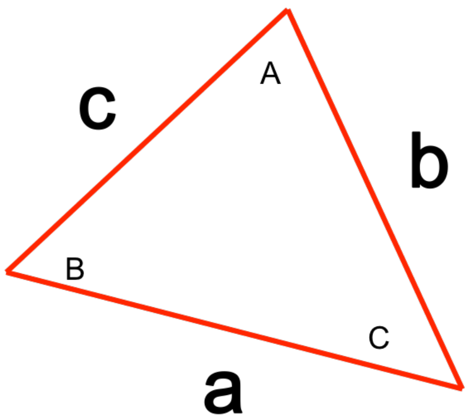
the following ratio's apply

$$\sin\theta = \frac{o}{h}$$

$$\cos\theta = \frac{a}{h}$$

$$\tan\theta = \frac{o}{a}$$

whilst for a non right-angled triangle



the rules are

$$\frac{\sin A}{a} = \frac{\sin B}{b} = \frac{\sin C}{c}$$

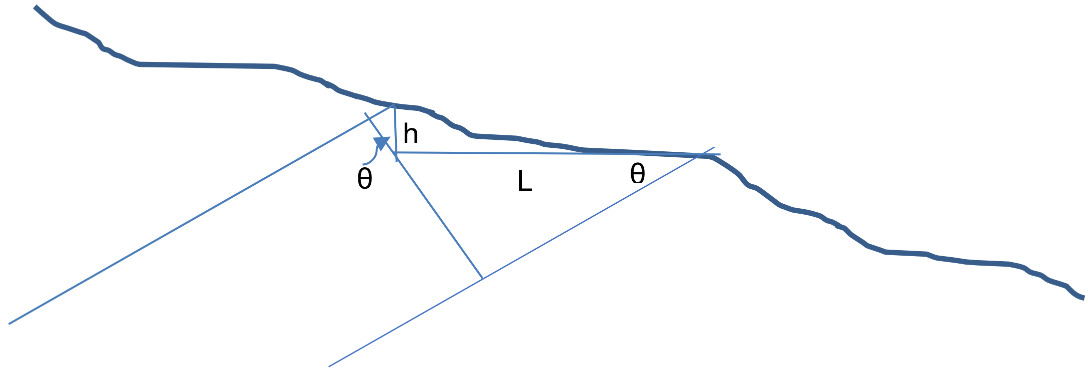
$$c^2 = a^2 + b^2 - 2bc\cos C$$

Notice that in the non-right angled triangles the sides and angles are labelled according to the side they are opposite to, with angles in capitals and sides in lower-case. So for example Angle A is opposite length *a*.

For the right-angled triangles the opposite and adjacent are labelled corresponding to whatever angle is given in the triangle (or the angle that is to be found).

Contexts in geology

Students may need to use trigonometry during mapwork to calculate the true thickness of a rock unit from a given dip amount and width of the outcrop on the map and the elevation differences between the contacts.



True thickness of bedding plane = (h cos θ) + (L sin θ)

Using the geological map a student obtained the following measurements which they used to calculate the true thickness of the bed:

h = difference in height between the base and top of the unit = 120 m – 108 m = 12 m

L = horizontal distance, perpendicular to strike, between the base and top of the unit = 67m

- $True\ thickness\ of\ unit = (h \times cos\theta) + (L \times sin\theta) = (12 \times 0.914) + (67 \times 0.407)$
- $True\ thickness\ of\ unit = 10.968 \times 27.269 = 27.269$

The true thickness of the rock unit is 27 m

Appendix A – Useful formulae for geology

As indicated in the specification content and the mathematical appendix, students should recall the following,

Magnification (2.1.1, 2.1.2, 2.1.3, 3.3.1):

$$Magnification = \frac{size\ of\ image}{size\ of\ real\ object}$$

Concentration factor (3.1.2, 5.5.1):

$$Concentration\ factor = \frac{\text{grade of metal in ore}}{\text{average crustal abundance}}$$

Return period (6.1.2):

$$Return\ period = \frac{\text{number of years on record} + 1}{\text{number of events of that magnitude}}$$

Lithostatic pressure (6.2.1):

$$p = \rho gh = density\ of\ rock \times g \times thickness\ of\ overburden$$

Phi scale (2.1.3, 5.5.1), not required at AS*:*

$$Phi = -log_2\left(\frac{\text{diameter of grain in mm}}{1}\right)$$

How to calculate rates - including,

- *Geothermal gradient (3.2.1, 5.3.2)*
- *Radioactivity (2.2.2, 5.3.2)*
- *Relative plate motion (3.2.1, 5.3.2)*
- *Sedimentation rate (2.2.2, 5.1.1, 7.2.2)*

$$rate = \frac{change\ in\ quantity}{time\ taken}$$

Mean (not content specific):

$$\bar{x} = \frac{\sum x}{n}$$

Percentage Change (not content specific):

$$percentage\ change = \frac{new\ quantity}{original\ quantity} \times 100$$

Percentage Yield (not content specific):

$$\%\ yield\ = \frac{Actual\ Amount}{Theoretical\ Amount} \times 100$$

Relative uncertainty in a value calculated by difference (not content specific):

$$\%\ uncertainty = \frac{2 \times absolute\ uncertainty}{quantity\ measured} \times 100\%$$

Surface Area to Volume Ratio (not content specific):

$$Ratio = \frac{Surface\ Area}{Volume}$$

Note: See *M4.1* for the formulae for circumferences, surface areas and volumes of regular shapes.

Appendix B – Formulae that will be provided in the assessments

Students should be aware of the following formulae. They should know how and when to use them.

Surface area of a sphere

$$4\pi r^2$$

Volume of a sphere

$$\frac{4}{3}\pi r^3$$

Chi Squared

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

Mann–Whitney U

U_1 is found by counting the number of values in sample 2 which exceed each of the values in sample 1. U_2 is found by counting the number of values in sample 1 which exceed each of the values in sample 2.

Spearman's Rank Correlation Coefficient

$$r_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

Standard Deviation

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$

Darcy's Law

$$Q = -\kappa A\left(\frac{h_2 - h_1}{L}\right)$$

Moment Magnitude scale e.g.

$$M_w = {}^2/_3\, log E - 6.1$$

(This will always take the form of the equation for a straight line i.e. $\mathrm{y = mx + c}$)

Radioactive decay

$$N = N_0 \mathrm{e}^{-\lambda t}$$

Stokes Law

$$v = \frac{gd^2(\rho_p - \rho_w)}{18\eta t}$$

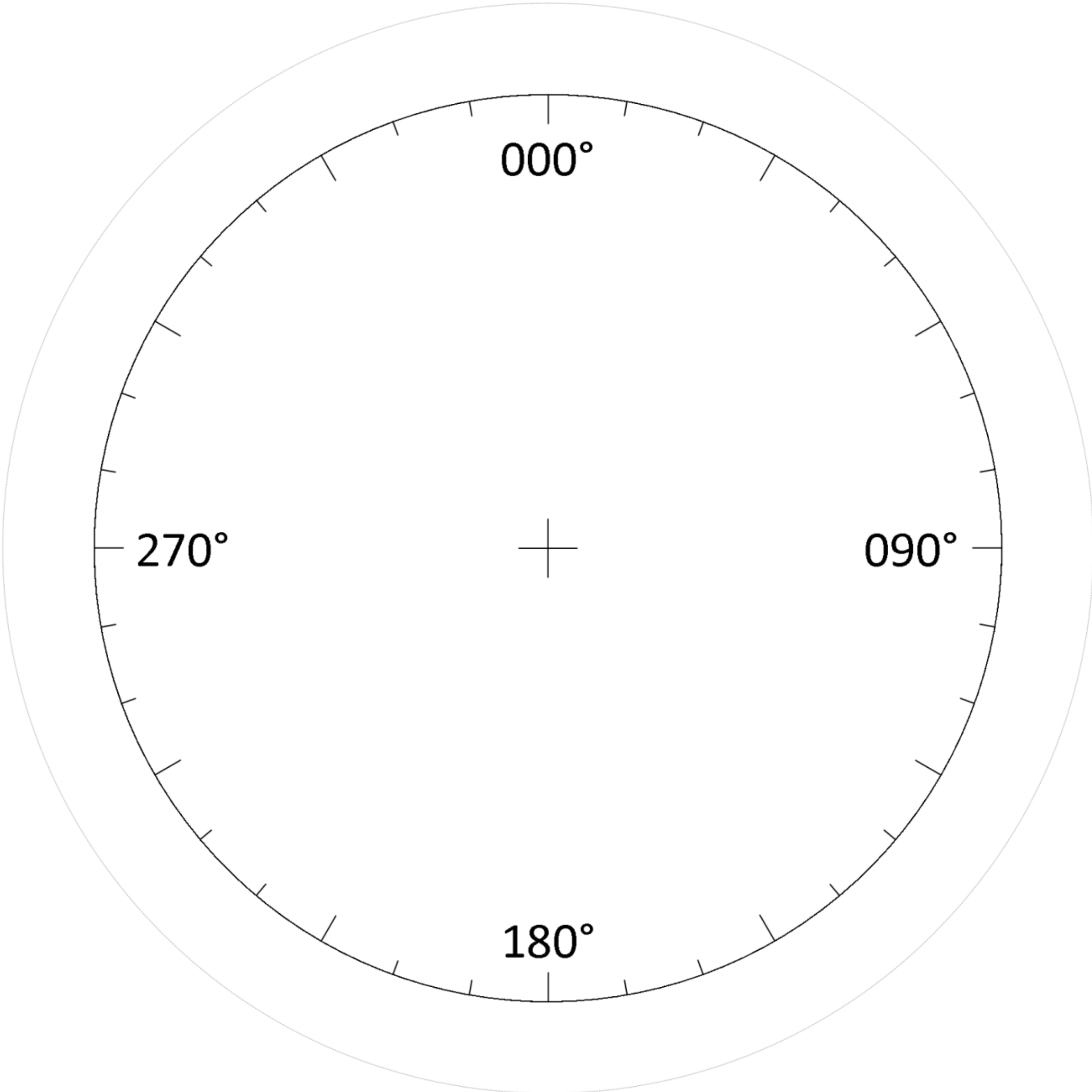
Appendix C – Key power laws

It is useful for students to be aware of the following power laws to help in certain mathematical skills, as referenced in the text.

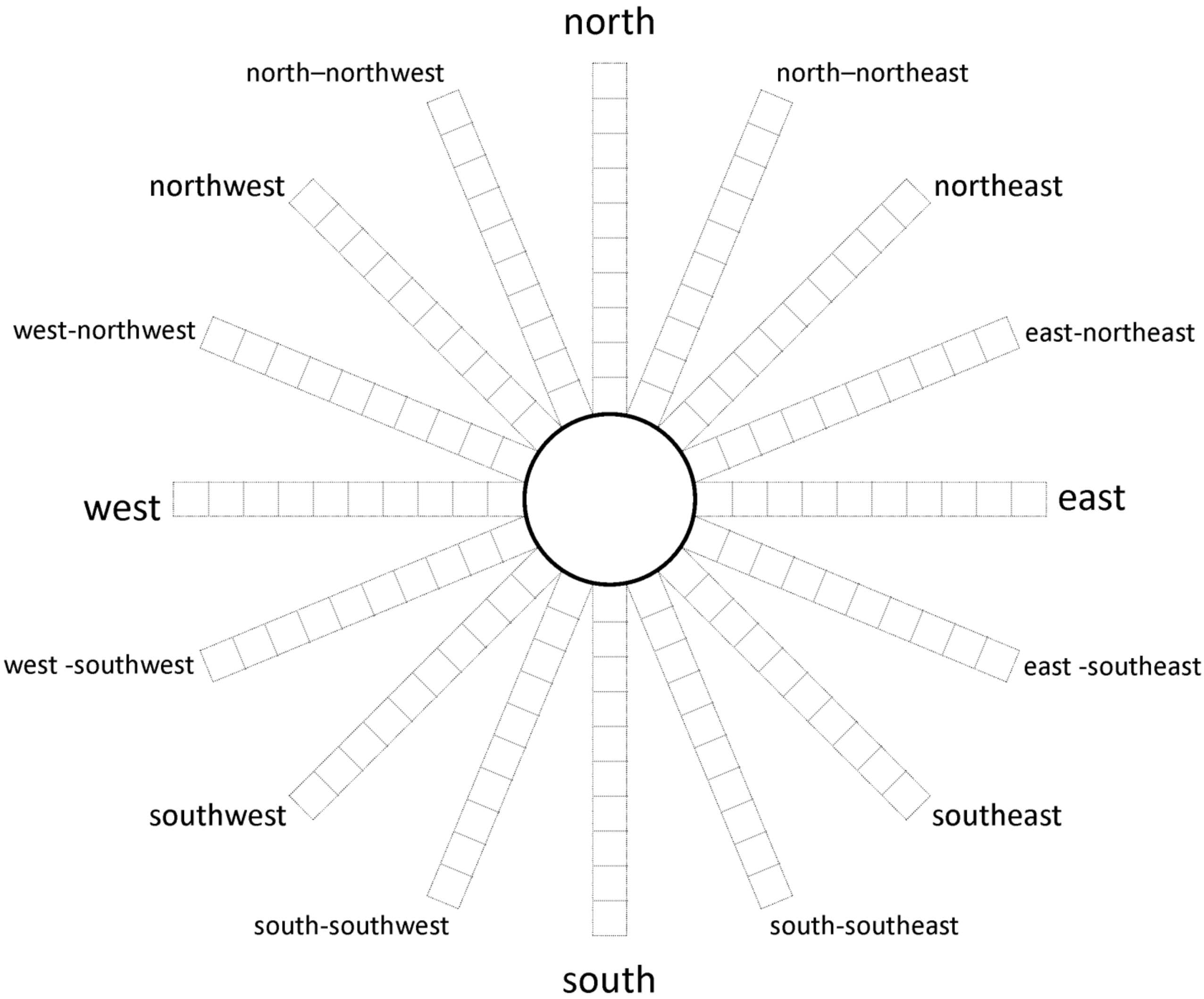
$x^n \times x^m = x^{n+m}$	multiplicative rule
$\dfrac{x^n}{x^m} = x^{n-m}$	division rule
$(x^n)^m = x^{nm}$	power rule
$x^{-1} = \dfrac{1}{x^n}$	reciprocal rule
$x^{n/m} = \sqrt[m]{x^n}$	root rule

Appendix D – Circular data display templates

Raw data plot

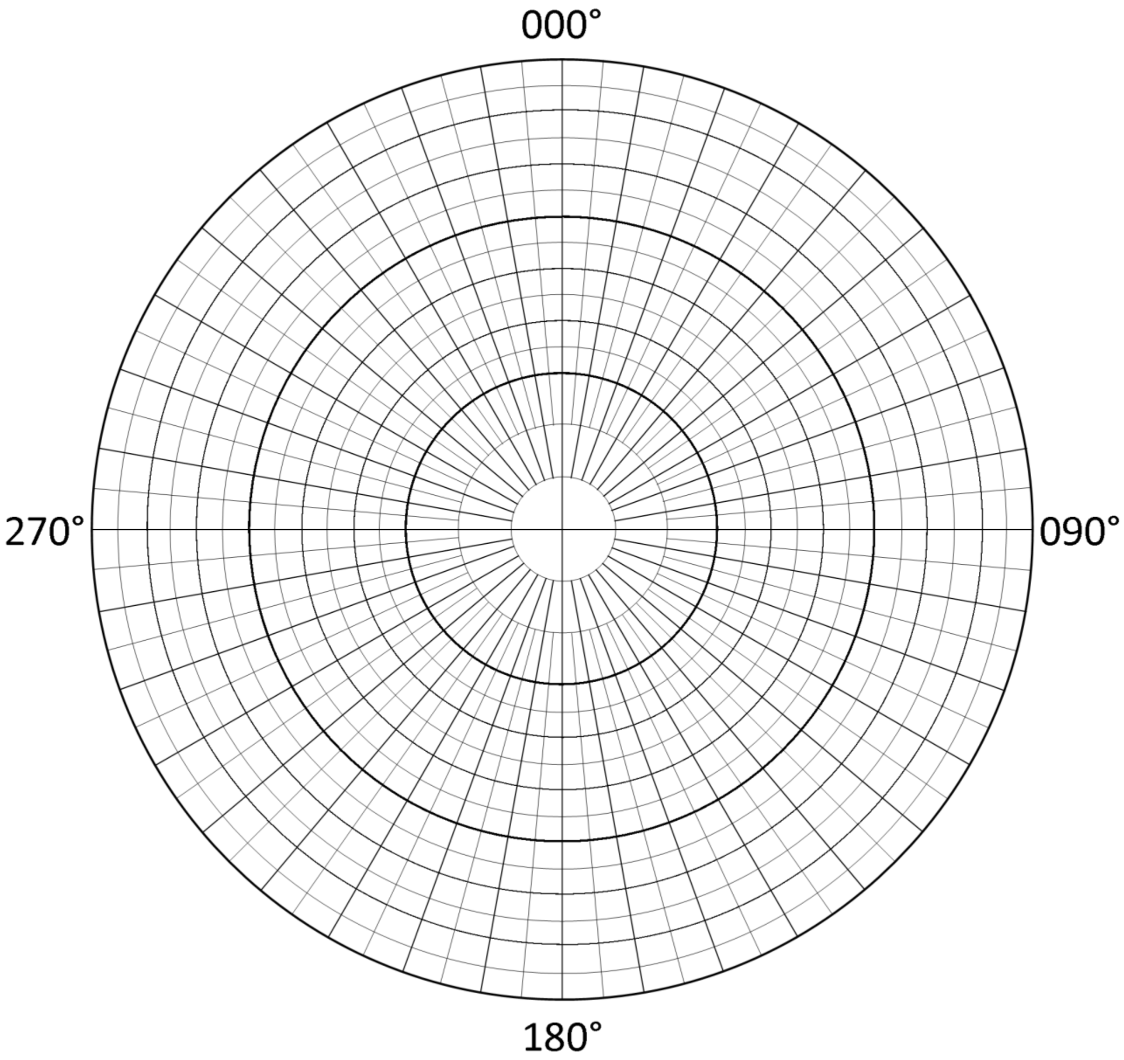


Circular bar graph

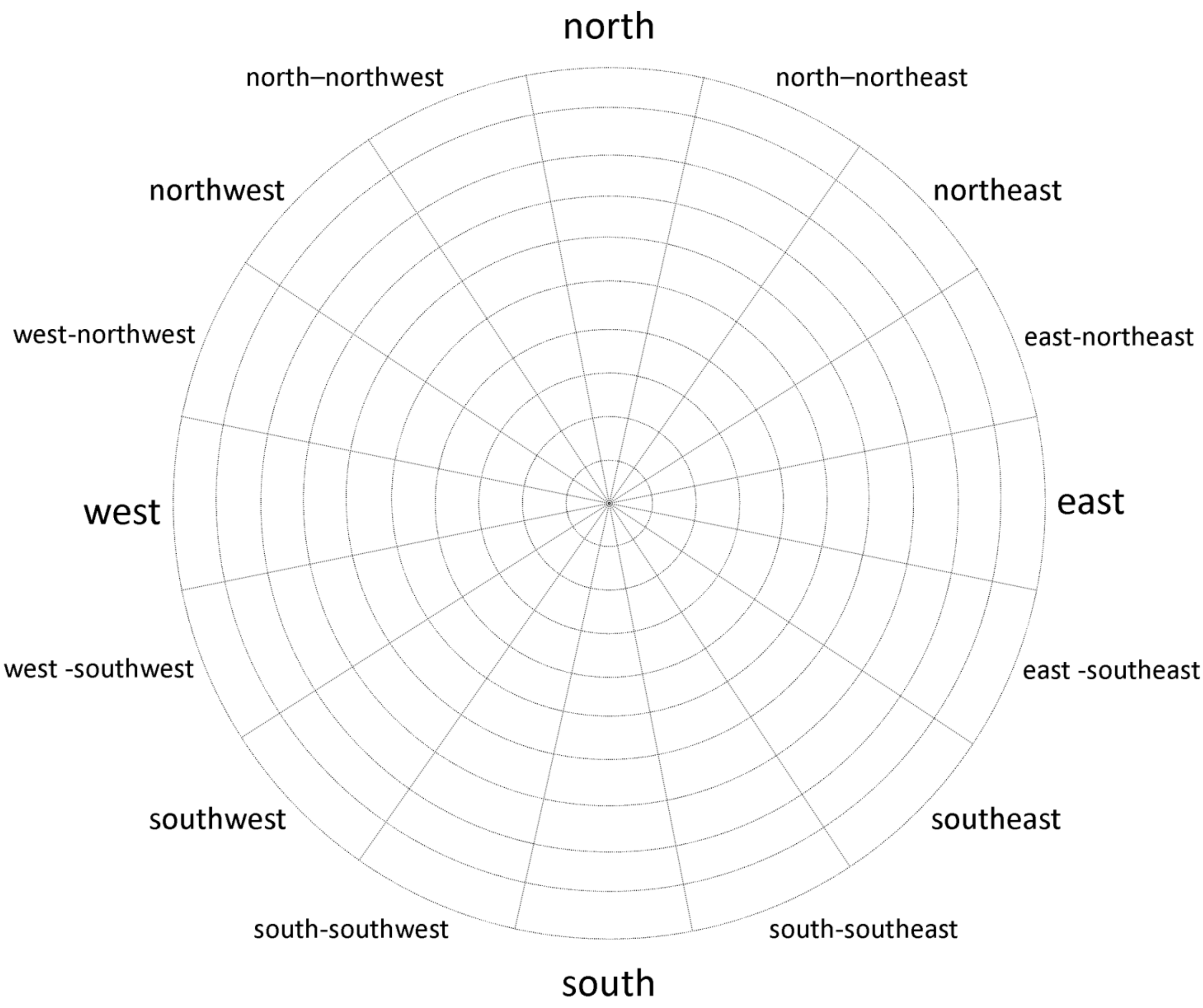


Direction	Observations	Direction	Observations
349° to 011°		169° to 191°	
012° to 033°		192° to 213°	
034° to 056°		214° to 236°	
057° to 078°		237° to 258°	
079° to 101°		259° to 281°	
102° to 123°		282° to 303°	
124° to 146°		304° to 326°	
147° to 168°		327° to 348°	

Polar equal area stereonet

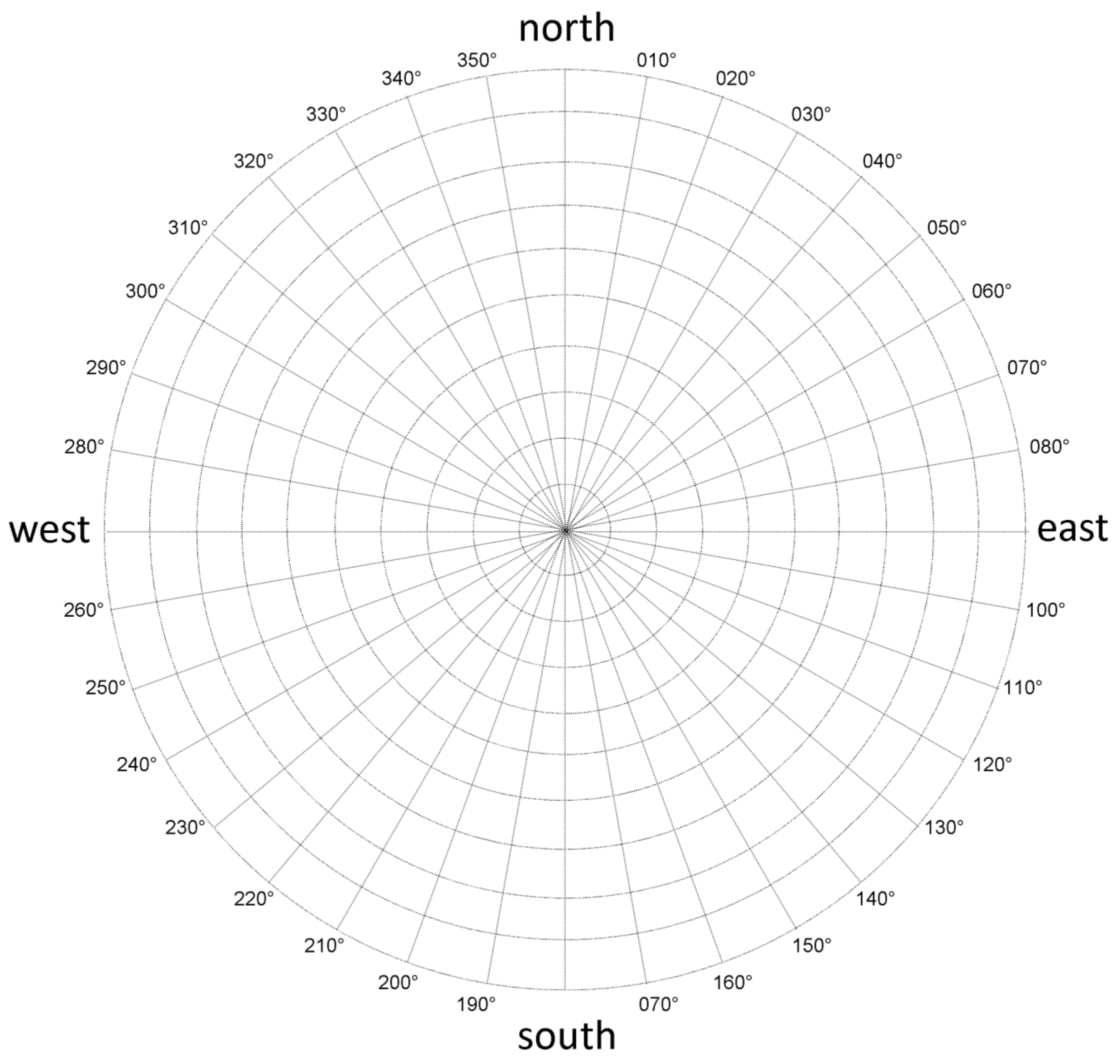


Rose diagram – cardinal and intercardinal directions – 22.5° intervals



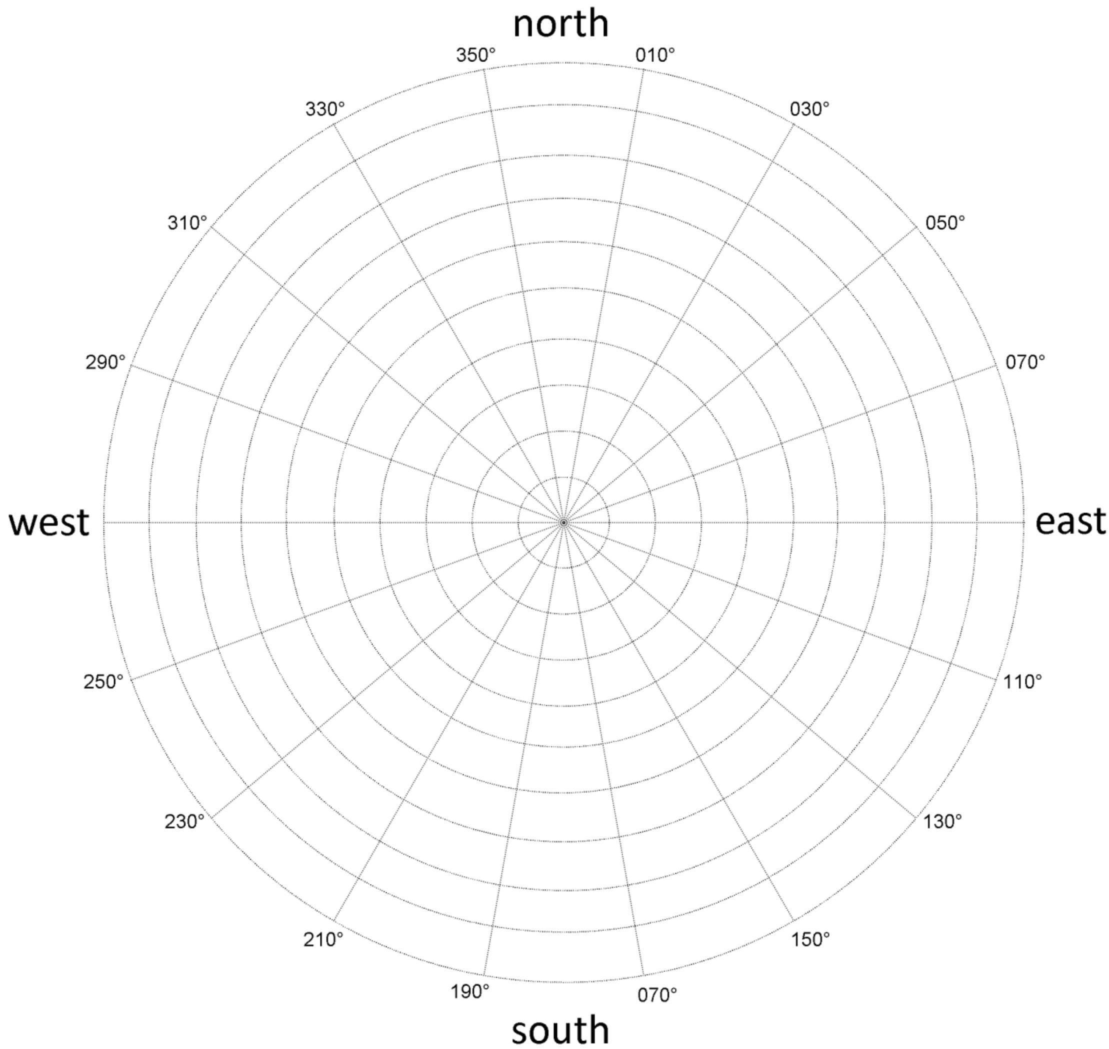
Direction	Observations	Direction	Observations	Tally
349° to 011°		169° to 191°		
012° to 033°		192° to 213°		
034° to 056°		214° to 236°		
057° to 078°		237° to 258°		
079° to 101°		259° to 281°		
102° to 123°		282° to 303°		
124° to 146°		304° to 326°		
147° to 168°		327° to 348°		

Rose diagram – 10° intervals



Direction	Observations	Direction	Observations	Direction	Observations
000° to 009°		120° to 129°		240° to 249°	
010° to 019°		130° to 139°		250° to 259°	
020° to 029°		140° to 149°		260° to 269°	
030° to 039°		150° to 159°		270° to 279°	
040° to 049°		160° to 169°		280° to 289°	
050° to 059°		170° to 179°		290° to 299°	
060° to 069°		180° to 189°		300° to 309°	
070° to 079°		190° to 199°		310° to 319°	
080° to 089°		200° to 209°		320° to 329°	
090° to 099°		210° to 219°		330° to 339°	
100° to 109°		220° to 229°		340° to 349°	
110° to 119°		230° to 239°		340° to 359°	

Rose diagram – 20° intervals



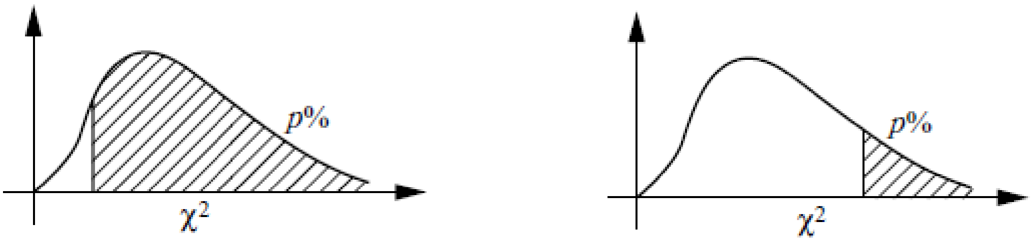
Direction	Observations	Direction	Observations	Tally
351° to 010°		171° to 190°		
011° to 030°		191° to 210°		
031° to 050°		211° to 230°		
051° to 070°		231° to 250°		
071° to 090°		251° to 270°		
091° to 110°		271° to 290°		
111° to 130°		291° to 310°		
131° to 150°		311° to 330°		
151° to 170°		331° to 350°		

Appendix E – Statistical tables

It is useful for students to be aware of the following statistical tables and how to use them (M2.12).

Chi Squared (χ^2)

Percentage points of the χ^2 (chi-squared) distribution



p%	99	97.5	95	90		10	5	2.5	1	0.5
v = 1	.0001	.0010	.0039	.0158		2.706	3.841	5.024	6.635	7.879
2	.0201	.0506	0.103	0.211		4.605	5.991	7.378	9.210	10.60
3	0.115	0.216	0.352	0.584		6.251	7.815	9.348	11.34	12.84
4	0.297	0.484	0.711	1.064		7.779	9.488	11.14	13.28	14.86
5	0.554	0.831	1.145	1.610		9.236	11.07	12.83	15.09	16.75
6	0.872	1.237	1.635	2.204		10.64	12.59	14.45	16.81	18.55
7	1.239	1.690	2.167	2.833		12.02	14.07	16.01	18.48	20.28
8	1.646	2.180	2.733	3.490		13.36	15.51	17.53	20.09	21.95
9	2.088	2.700	3.325	4.168		14.68	16.92	19.02	21.67	23.59
10	2.558	3.247	3.940	4.865		15.99	18.31	20.48	23.21	25.19
11	3.053	3.816	4.575	5.578		17.28	19.68	21.92	24.72	26.76
12	3.571	4.404	5.226	6.304		18.55	21.03	23.34	26.22	28.30
13	4.107	5.009	5.892	7.042		19.81	22.36	24.74	27.69	29.82
14	4.660	5.629	6.571	7.790		21.06	23.68	26.12	29.14	31.32
15	5.229	6.262	7.261	8.547		22.31	25.00	27.49	30.58	32.80
16	5.812	6.908	7.962	9.312		23.54	26.30	28.85	32.00	34.27
17	6.408	7.564	8.672	10.09		24.77	27.59	30.19	33.41	35.72
18	7.015	8.231	9.390	10.86		25.99	28.87	31.53	34.81	37.16
19	7.633	8.907	10.12	11.65		27.20	30.14	32.85	36.19	38.58
20	8.260	9.591	10.85	12.44		28.41	31.41	34.17	37.57	40.00
21	8.897	10.28	11.59	13.24		29.62	32.67	35.48	38.93	41.40
22	9.542	10.98	12.34	14.04		30.81	33.92	36.78	40.29	42.80
23	10.20	11.69	13.09	14.85		32.01	35.17	38.08	41.64	44.18
24	10.86	12.40	13.85	15.66		33.20	36.42	39.36	42.98	45.56
25	11.52	13.12	14.61	16.47		34.38	37.65	40.65	44.31	46.93
26	12.20	13.84	15.38	17.29		35.56	38.89	41.92	45.64	48.29
27	12.88	14.57	16.15	18.11		36.74	40.11	43.19	46.96	49.64
28	13.56	15.31	16.93	18.94		37.92	41.34	44.46	48.28	50.99
29	14.26	16.05	17.71	19.77		39.09	42.56	45.72	49.59	52.34
30	14.95	16.79	18.49	20.60		40.26	43.77	46.98	50.89	53.67
35	18.51	20.57	22.47	24.80		46.06	49.80	53.20	57.34	60.27
40	22.16	24.43	26.51	29.05		51.81	55.76	59.34	63.69	66.77
50	29.71	32.36	34.76	37.69		63.17	67.50	71.42	76.15	79.49
100	70.06	74.22	77.93	82.36		118.5	124.3	129.6	135.8	140.2

Critical values of Mann–Whitney U coefficient at the 5% level (2-tail test) and 2.5% level (1-tail test)

		Size of the larger sample																				
		5	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**
Size of the smaller sample	**5**	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20	22	23	24	25	27
	6	–	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27	29	30	32	33	35
	7	–	–	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44
	8	–	–	–	13	15	17	19	22	24	26	29	31	34	36	38	41	43	45	48	50	53
	9	–	–	–	–	17	20	23	26	28	31	34	37	39	42	45	48	50	53	56	59	62
	10	–	–	–	–	–	23	26	29	33	36	39	42	45	48	52	55	58	61	64	67	71
	11	–	–	–	–	–	–	30	33	37	40	44	47	51	55	58	62	65	69	73	76	80
	12	–	–	–	–	–	–	–	37	41	45	49	53	57	61	65	69	73	77	81	85	89
	13	–	–	–	–	–	–	–	–	45	50	54	59	63	67	72	76	80	85	89	94	98
	14	–	–	–	–	–	–	–	–		55	59	64	67	74	78	83	88	93	98	102	107
	15	–	–	–	–	–	–	–	–	–	–	64	70	75	80	85	90	96	101	106	111	117
	16	–	–	–	–	–	–	–	–	–	–	–	75	81	86	92	98	103	109	115	120	126
	17	–	–	–	–	–	–	–	–	–	–	–	–	87	93	99	105	111	117	123	129	135
	18	–	–	–	–	–	–	–	–	–	–	–	–	–	99	106	112	119	125	132	138	145
	19	–	–	–	–	–	–	–	–	–	–	–	–	–	–	113	119	126	133	140	147	154
	20	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	127	134	141	149	156	163
	21	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	142	150	157	165	173
	22	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	158	166	174	182
	23	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	175	183	192
	24	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	192	201
	25	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	211

Critical values for Spearman's rank correlation coefficient, r_s

1-Tail Test	5%	2½%	1%	½%	1-Tail Test	5%	2½%	1%	½%
2-Tail Test	10%	5%	2%	1%	2-Tail Test	10%	5%	2%	1%
n					*n*				
1	–	–	–	–	31	0.3012	0.3560	0.4185	0.4593
2	–	–	–	–	32	0.2962	0.3504	0.4117	0.4523
3	–	–	–	–	33	0.2914	0.3449	0.4054	0.4455
4	1.0000	–	–	–	34	0.2871	0.3396	0.3995	0.4390
5	0.9000	1.0000	1.0000	–	35	0.2829	0.3347	0.3936	0.4328
6	0.8286	0.8857	0.9429	1.0000	36	0.2788	0.3300	0.3882	0.4268
7	0.7143	0.7857	0.8929	0.9286	37	0.2748	0.3253	0.3829	0.4211
8	0.6429	0.7381	0.8333	0.8810	38	0.2710	0.3209	0.3778	0.4155
9	0.6000	0.7000	0.7833	0.8333	39	0.2674	0.3168	0.3729	0.4103
10	0.5636	0.6485	0.7455	0.7939	40	0.2640	0.3128	0.3681	0.4051
11	0.5364	0.6182	0.7091	0.7545	41	0.2606	0.3087	0.3636	0.4002
12	0.5035	0.5874	0.6783	0.7273	42	0.2574	0.3051	0.3594	0.3955
13	0.4835	0.5604	0.6484	0.7033	43	0.2543	0.3014	0.3550	0.3908
14	0.4637	0.5385	0.6264	0.6791	44	0.2513	0.2978	0.3511	0.3865
15	0.4464	0.5214	0.6036	0.6536	45	0.2484	0.2974	0.3470	0.3822
16	0.4294	0.5029	0.5824	0.6353	46	0.2456	0.2913	0.3433	0.3781
17	0.4142	0.4877	0.5662	0.6176	47	0.2429	0.2880	0.3396	0.3741
18	0.4014	0.4716	0.5501	0.5996	48	0.2403	0.2850	0.3361	0.3702
19	0.3912	0.4596	0.5351	0.5842	49	0.2378	0.2820	0.3326	0.3664
20	0.3805	0.4466	0.5218	0.5699	50	0.2353	0.2791	0.3293	0.3628
21	0.3701	0.4364	0.5091	0.5558	51	0.2329	0.2764	0.3260	0.3592
22	0.3608	0.4252	0.4975	0.5438	52	0.2307	0.2736	0.3228	0.3558
23	0.3528	0.4160	0.4862	0.5316	53	0.2284	0.2710	0.3198	0.3524
24	0.3443	0.4070	0.4757	0.5209	54	0.2262	0.2685	0.3168	0.3492
25	0.3369	0.3977	0.4662	0.5108	55	0.2242	0.2659	0.3139	0.3460
26	0.3306	0.3901	0.4571	0.5009	56	0.2221	0.2636	0.3111	0.3429
27	0.3242	0.3828	0.4487	0.4915	57	0.2201	0.2612	0.3083	0.3400
28	0.3180	0.3755	0.4401	0.4828	58	0.2181	0.2589	0.3057	0.3370
29	0.3118	0.3685	0.4325	0.4749	59	0.2162	0.2567	0.3030	0.3342
30	0.3063	0.3624	0.4251	0.4670	60	0.2144	0.2545	0.3005	0.3314

The small print

www.ocr.org.uk

OCR Customer Support Centre

General qualifications

Telephone 01223 553998

Facsimile 01223 552627

Email general.qualifications@ocr.org.uk

5400719668