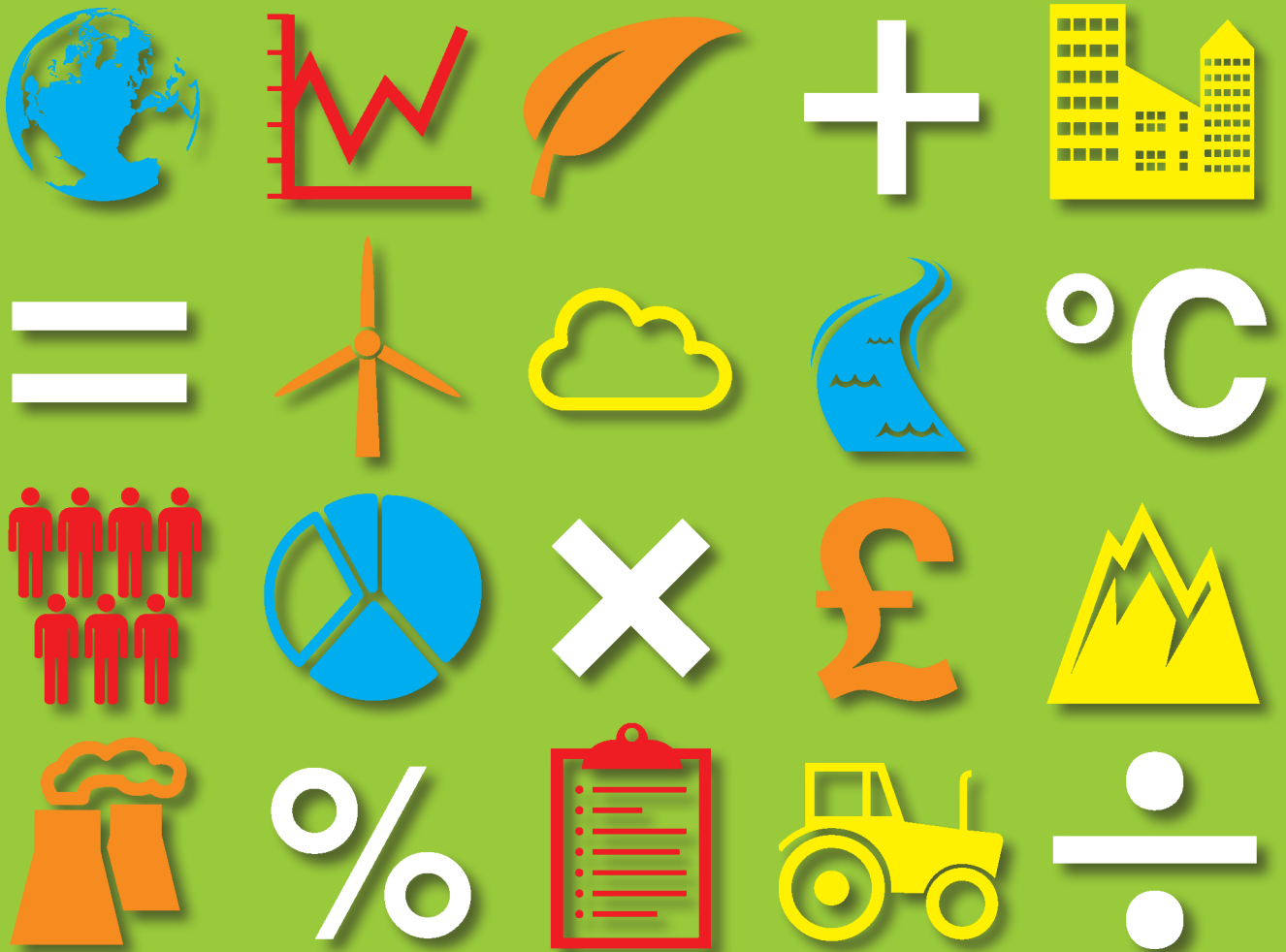


Essential Maths Skills for AS/A-level Geography

Helen Harris



Series Editor Heather Davis
Educational Consultant with Cornwall Learning

Contents

Introduction	5
1 Understanding data	
Nominal, ordinal and interval data	7
Ratios and fractions	9
Percentages	11
Densities	17
2 Measures of central tendency	
Mean, median and mode	20
Frequency distributions	25
3 Measures of dispersion	
Range, interquartile range and standard deviation	31
4 Measures of concentration	
Nearest neighbour index	39
Location quotient	44
Lorenz curves	47
5 Measures of correlation	
Scatter graphs and lines of best fit	56
Spearman rank correlation coefficient	60
6 Testing for differences between data sets	
Student's t test	68
Mann–Whitney U test	73
Chi-squared test	79

Exam-style questions

AS level questions	86
A-level questions	88

Appendix 1

Statistical tables	90
--------------------------	----

Appendix 2

Choosing statistical methods for geographical fieldwork	93
---	----

Appendix 3

Key terms.....	94
----------------	----

Appendix 4

Specification cross-reference.....	96
------------------------------------	----

2 Measures of central tendency

Mean, median and mode

Measures of **central tendency** are used in geography to identify simple features of data or differences between data sets. These measures give a single figure which provides a useful **summary** of the data set. When tables of data are generated from fieldwork or collected from secondary sources, a preliminary overview can be obtained by calculating measures of central tendency. For example, if you wanted to study differences in stream flow over a number of sites or variations in property values in different parts of an urban area, it would be convenient to have a single number representing all the data collected from each location.

Geographers mostly use the following three standard measures of central tendency:

- **Mean** — the most common measure, calculated by adding up all the individual data values and dividing by the total number of data items. The mean is what is usually meant by the term ‘average’.
- **Median** — the middle value in the data set. To find the median, arrange all the data items in order (a ‘rank order’) and identify the middle value in the sequence. If there is an even number of data items, the median is halfway between the two middle values; that is, add them together and divide by 2.
- **Mode** — the most straightforward measure, obtained by identifying the most frequently occurring value in a data set or, if the data has been put into categories, the most common category.

A Worked examples

Table 2.1 contains secondary data on the number of migrants entering the UK between 2005 and 2014. You want to use some basic measures of central tendency to obtain an overall summary of the data set.

Table 2.1 Immigration to the UK, 2005–2014

Year	Number of immigrants to the UK (thousands)
2005	567
2006	596
2007	574
2008	590
2009	567
2010	591
2011	566
2012	498
2013	526
2014	632

Source: Office for National Statistics (ONS)

a Calculate the mean value of this data set.

Step 1: add the numbers of immigrants in all the years to find the total number for the 10-year period. Remember that the data in the table is in thousands.

$$567 + 596 + 574 + 590 + 567 + 591 + 566 + 498 + 526 + 632 = 5707 \text{ (thousand)}$$

Step 2: divide the total by the number of years, in this case 10, to find the mean.

$$5707 \div 10 = 570.7 \text{ (thousand)}$$

Rounding to the nearest whole number, we find that the mean annual number of immigrants to the UK from 2005 to 2014 is 571 thousand.

b Calculate the median value of this data set.

Step 1: order the data values from lowest to highest.

$$498, 526, 566, 567, 567, 574, 590, 591, 596, 632$$

Step 2: the median is the middle value in this ordered list.

Because there is an even number of values (10), the median is halfway between the 5th and 6th values.

The 5th value is 567; the 6th value is 574. Add them and divide by 2:

$$567 + 574 = 1141$$

$$1141 \div 2 = 570.5$$

Rounding to the nearest whole number, we find that the median annual number of immigrants to the UK from 2005 to 2014 is 571 thousand.

c Find the mode of this data set.

From the ordered list in Worked example **b**, it can be seen that 567 is the only repeated value (occurring twice), so the mode is 567 thousand.

Comparing the mean, median and mode

Mean

- Calculation of this average makes use of all the data values and gives you a simple overview of the whole data set.
- It works best when the data values span a fairly narrow range. If there are several exceptionally high or low values (called ‘anomalies’ or ‘outliers’), the mean will not be very representative of the data set.

Median

- As the median is the middle value when the data items are ranked from lowest to highest, it is quite straightforward to calculate.
- It is not affected by extreme values in the data set.

Mode

- The mode is the easiest measure of central tendency to obtain and it can be used for categorical data, whereas the mean and median can only be used with ordinal (including interval and ratio) data.
- It is not affected by extreme values in the data set.
- However, a data set can have no modal value (if each data value occurs only once) or more than one modal value, which could be confusing.

As a geographer, after calculating a measure or measures of central tendency, you should then use this information to analyse the data further. For instance, in the Worked example, you could examine the data to see if certain years had well above (632 thousand in 2014) or well below (498 thousand in 2012) the average intake of migrants and the next step would be to investigate the possible reasons for this.

B Guided question

Copy out the workings and complete the answers on a separate piece of paper.

- 1 Table 2.2 shows rainfall totals recorded in Oban and Norwich for each month in 2014.

Table 2.2 Monthly rainfall in Oban and Norwich in 2014

Month	Rainfall in Oban (mm)	Rainfall in Norwich (mm)
January	150	42
February	75	39
March	98	44
April	68	39
May	76	36
June	89	52
July	55	45
August	80	51
September	98	63
October	130	59
November	128	65
December	110	61

- a Calculate the mean monthly rainfall for Oban and for Norwich.

For Oban:

Step 1: add up the rainfall amounts in each month.

$$150 + 75 + 98 + 68 + 76 + 89 + 55 + 80 + 98 + 130 + 128 + 110 = \text{_____ mm}$$

Step 2: divide the total by the number of months.

$$\text{_____} \div 12 = \text{_____}$$

The mean monthly rainfall for Oban is _____ mm.

For Norwich:

Step 1: add up the rainfall amounts in each month.

$$42 + \text{__} + \text{__} + \text{__} + \text{__} + \text{__} + \text{__} + \text{__} + \text{__} + \text{__} + \text{__} + 61 = \text{_____ mm}$$

Step 2: divide the total by the number of months.

$$\text{_____} \div 12 = \text{_____}$$

The mean monthly rainfall for Norwich is _____ mm.

b Calculate the median monthly rainfall for Oban and for Norwich.**For Oban:**

Step 1: order the values from lowest to highest.

55, 68, 75, 76, 80, 89, 98, 98, 110, 128, 130, 150

Step 2: because there is an even number of values (12), the median is halfway between the 6th and 7th values (so that there are five values above and five values below).

The 6th value is ____; the 7th value is _____. Add them and divide by 2:

$$\underline{\quad} + \underline{\quad} = \underline{\quad}$$

$$\underline{\quad} \div 2 = \underline{\quad}$$

The median monthly rainfall for Oban is _____ mm.

For Norwich:

Step 1: order the values from lowest to highest.

36, ____, ____, ____, ____, ____, ____, ____, ____, ____, ____, ____, 65

Step 2: because there is an even number of values (12), the median is halfway between the 6th and 7th values.

The 6th value is ____; the 7th value is _____. Add them and divide by 2:

$$\underline{\quad} + \underline{\quad} = \underline{\quad}$$

$$\underline{\quad} \div 2 = \underline{\quad}$$

The median monthly rainfall for Norwich is _____ mm.

c Identify the modal monthly rainfall for Oban and for Norwich.**For Oban:**

Step 1: order the values from lowest to highest.

55, 68, 75, 76, 80, 89, 98, 98, 110, 128, 130, 150

Step 2: identify the most common value in the data set.

The value ____ occurs twice; all the other values occur only once.

The modal monthly rainfall for Oban is _____ mm.

For Norwich:

Step 1: order the values from lowest to highest.

36, ____, ____, ____, ____, ____, ____, ____, ____, ____, ____, ____, 65

Step 2: identify the most common value in the data set.

The modal monthly rainfall for Norwich is _____ mm.

d What is the difference in average annual rainfall between Oban and Norwich?

If it is not specified which kind of ‘average’ is meant, usually we take it to be the mean. In this case, let us calculate the difference between the means, the medians and the modes for comparison.

Difference in mean rainfall of Oban and Norwich = ____ - ____ = ____ mm

Difference in median rainfall of Oban and Norwich = ____ - ____ = ____ mm

Difference in modal rainfall of Oban and Norwich = ____ - ____ = ____ mm

C Practice questions

2 Table 2.3 shows data for the top ten countries in imports and exports in 2013.

Table 2.3 Top ten exporters and importers in 2013 (values are in billions of US dollars)

Rank	Exporters	Value	Rank	Importers	Value
1	China	2 209	1	United States	2 329
2	United States	1 580	2	China	1 950
3	Germany	1 453	3	Germany	1 189
4	Japan	715	4	Japan	833
5	Netherlands	672	5	France	681
6	France	580	6	UK	655
7	Republic of Korea	560	7	Hong Kong	622
8	UK	542	8	Netherlands	590
9	Hong Kong	536	9	Republic of Korea	576
10	Russian Federation	523	10	Italy	477

Source: World Trade Organization (WTO)

- a What is the **mean** value of the
- exports?
 - imports?
- b What is the **median** value of the
- exports?
 - imports?
- c Do the median values for imports and exports accurately reflect the range of data in the table?
- 3 Table 2.4 shows fieldwork data collected from three different sites on a beach in North Wales. The values in the second, third and fourth columns are the numbers of pebbles of each size found at that particular site.

Table 2.4 Data from a survey of pebble sizes on a beach in North Wales

Long axis (mm)	Site 1	Site 2	Site 3
<10	67	17	6
10–19	54	24	12
20–39	20	23	47
>40	14	23	53

Sea → Inland

What is the modal category for each of the three sites?

TIP

In mathematical notation the mean is written as

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

where n is the number of data items and Σ (Sigma, the Greek capital letter 'S') stands for 'sum', so that Σx represents the sum of all the data values 'x'.