



YEAR 9  
KNOWLEDGE ORGANISERS



BLOCK: REASONING WITH DATA

The data handling cycle

Measures of location

"MATHS OPENS DOORS"

# YEAR 9 - REASONING WITH DATA... The data handling cycle

## What do I need to be able to do?

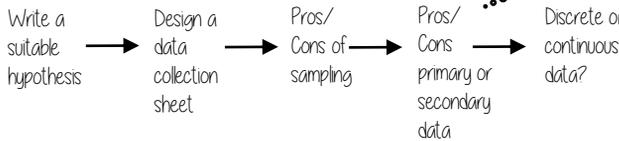
By the end of this unit you should be able to:

- Set up a statistical enquiry
- Design and criticise questionnaires
- Draw and interpret multiple bar charts
- Draw and interpret line graphs
- Represent and interpret grouped quantitative data
- Find and interpret the range
- Compare distributions

## Keywords

- Average:** a measure used to find the middle (central tendency) of a data set. The three commonly used averages are mean, median and mode.
- Continuous data:** quantitative (numerical) data which can take any value within a range, e.g. height, length, weight
- Discrete data:** quantitative (numerical) data which can only take certain values, usually whole numbers.
- Hypothesis:** a statement which might be true, that can then be tested
- Primary data:** data which you have collected yourself
- Proportion:** a part share or number considered by comparing it to a whole
- Sampling:** taking a selection from a larger group (the "population") that will let you find out things about the larger group
- Secondary data:** data which has been collected by someone else
- Spread:** the variation of values within a set of data

## Set up a statistical enquiry



Features of a data collection sheet

Data Title	Tally	Frequency
Grouped or ungrouped categories		Total number of that group observed

## Design and criticise a questionnaire

**The Question** - be clear with the question - don't be too leading/ judgemental

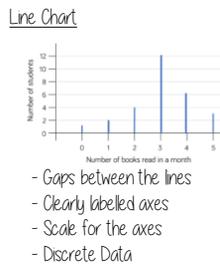
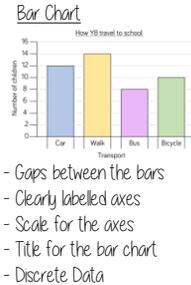
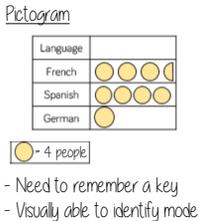
e.g. How much pocket money do you get a week?

**Responses** - do you want closed or open responses? - do any options overlap? - Have you an option for all responses?

Zero option →  £0    £0.01 - £2    £2.01 - £4    more than £4 ← More option

NOTE: For responses about continuous data include inequalities  $< x \leq$

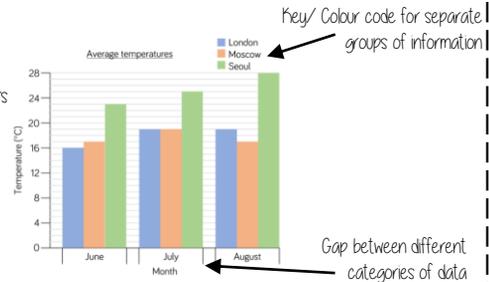
## Pictograms, bar and line charts



## Multiple Bar chart

Compares multiple groups of data

- Clearly labelled axes
- Scale for axes
- Comparable data bars drawn next to each other



## Draw and interpret Pie Charts

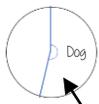
R

Remember a circle has 360°

Type of pet	Dog	Cat	Hamster
Frequency	32	25	3

There were 60 people asked in this survey (Total frequency)

**Multiple method**  
As 60 goes into 360 - 6 times  
Each frequency can be multiplied by 6 to find the degrees (proportion of 360)



Use a protractor to draw  
This is 192°

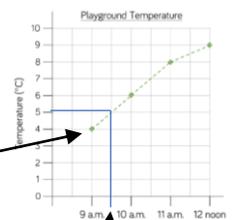
Represents quantitative, discrete data

## Draw and interpret line graphs

- Commonly used to show changing over time
- The points are the recorded information and the lines join the points

Line graphs do not need to start from 0

More than one piece of data can be plotted on the same graph to compare data



It is possible to make estimates from the line  
e.g. temperature at 9.30am is 5°C

## Grouped quantitative data

Time (minutes)	Frequency
$0 \leq t < 5$	4
$5 \leq t < 10$	6
$10 \leq t < 15$	5
$15 \leq t < 20$	8
$20 \leq t < 25$	10
$25 \leq t < 30$	1



Grouping the data is useful if there is a large spread of data to begin with

"More than or equal to 25 and less than 30 minutes"

The use of inequalities shows that this will be a frequency diagram

## Find and interpret the range

The range is a measure of **spread**

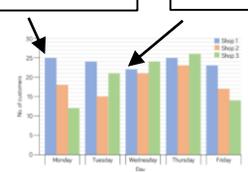
A smaller range means there is less variation in the results - it is more consistent data

A range of 0 means all the data is the same value

Shop 1 has the smallest range - this indicates it has a more consistent flow of customers each week.

Difference between the biggest and smallest values

Shop 1 highest value      Shop 1 lowest value



Range of customers =  $25 - 22 = 3$  (Shop 1)

# YEAR 9 - REASONING WITH DATA...

## Measures of location

### What do I need to be able to do?

By the end of this unit you should be able to:

- Understand and use mean, median and mode
- Choose the most appropriate average
- Identify outliers
- Compare distributions using averages and range

### Keywords

**Average:** a measure used to find the middle (central tendency) of a data set. The three commonly used averages are mean, median, and mode.

**Consistent:** free from contradiction

**Frequency:** the number of times a particular item appears in a set of data

**Outlier:** a value that "lies outside" (is much smaller or larger than) most of the other values in a set of data

**Represent:** to serve as an example of something

**Spread:** the variation of values within a set of data

**Total:** the sum or whole amount, the result of an addition

### Mean, Median, Mode

#### The Mean

A measure of average to find the central tendency... a typical value that represents the data

24, 8, 4, 11, 8

Find the sum of the data (add the values) 55

Divide the overall total by how many pieces of data you have  $55 \div 5$

Mean = 11

#### The Median

The value in the center (in the middle) of the data

24, 8, 4, 11, 8

Put the data in order

4, 8, 8, 11, 24

Find the value in the middle

4, 8, 8, 11, 24

Median = 8

NOTE: If there is no single middle value, find the mean of the two numbers left

#### The Mode (The modal value)

This is the number OR the item that occurs the most (it does not have to be numerical)

24, 8, 4, 11, 8

This can still be easier if the data is ordered first

4, 8, 8, 11, 24

Mode = 8

### Choosing the appropriate average

The average should be a representative of the data set – so it should be compared to the set as a whole - to check if it is an appropriate average

Here are the weekly wages of a small firm

£240 £240 £240 £240 £240  
£260 £260 £300 £350 £700

Which average best represents the weekly wage?

The Mean = £307

The Median = £250

The Mode = £240

Put the data back into context

Mean/Median – too high (most of this company earn £240)

Mode is the best average that represents this wage

It is likely that the salaries above £240 are more senior staff members – their salary doesn't represent the average weekly wage of the majority of employees

### Identify outliers

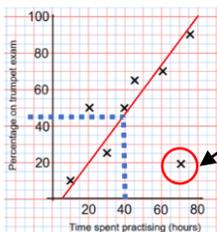
Outliers are values that stand well apart from the rest of the data

Outliers can have a big impact on range and mean. They have less impact on the median and the mode

Sometimes it is best to not use an outlier in calculations

Height in cm  
152 150 142 158 182 151 153 149 156 160 151 144

Where an outlier is identified try to give it some context. This is likely to be a taller member of the group. Could it be an older student or a teacher?



Outliers can also be identified graphically e.g. on scatter graphs

### Comparing distributions

Comparisons should include a statement of average and central tendency, as well as a statement about spread and consistency

Here are the number of runs scored last month by Lucy and James in cricket matches

Lucy: 45, 32, 37, 41, 48, 35

James: 60, 90, 41, 23, 14, 23

Lucy

Mean: 39.6 (1dp), Median: 38, Mode: no mode, Range: 16

James

Mean: 41.8 (1dp), Median: 32, Mode: 23, Range: 76

James has two extreme values that have a big impact on the range

"James is less consistent than Lucy because his scores have a greater range. Lucy performed better on average because her scores have a similar mean and a higher median"