

REASONING WITH DATA...

The data handling cycle

@whisto_maths

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

Hypothesis: an idea or question you want to test

Sampling: the group of things you want to use to check your hypothesis

Primary Data: data you collect yourself

Secondary Data: data you source from elsewhere e.g. the internet/ newspapers/ local statistics

Discrete Data: numerical data that can only take set values

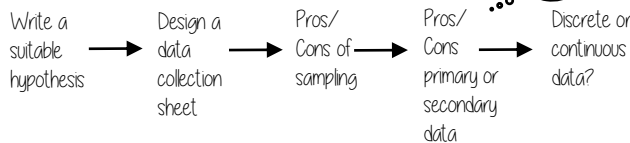
Continuous Data: numerical data that has an infinite number of values (often seen with height, distance, time)

Spread: the distance/ how spread out/ variation of data

Average: a measure of central tendency – or the typical value of all the data together

Proportion: numerical relationship that compares two things

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

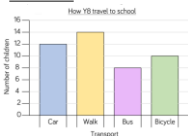
Pictogram

Language	
French	4 circles
Spanish	3 circles
German	1 circle

1 circle = 4 people

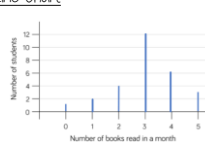
- Need to remember a key
- Visually able to identify mode

Bar Chart



- Gaps between the bars
- Clearly labelled axes
- Scale for the axes
- Title for the bar chart
- Discrete Data

Line Chart



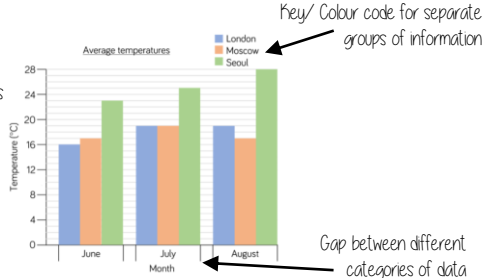
- Gaps between the lines
- Clearly labelled axes
- Scale for the axes
- Discrete Data

Represents quantitative data

Multiple Bar chart

Compares multiple groups of data

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



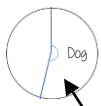
Draw and interpret Pie Charts

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

This fraction of the 360 degrees represents dogs

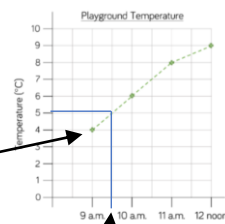
$$\frac{32}{60} \times 360 = 192^\circ$$

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

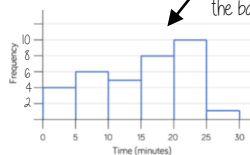
More than one piece of data can be plotted on the same graph to compare data



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



This is a frequency diagram
There are no gaps between the bars

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

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

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

Find and interpret the range

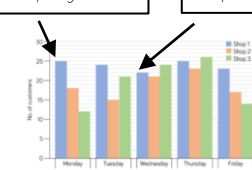
The range is a measure of spread

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

Difference between the biggest and smallest values

Shop 1 highest value Shop 1 lowest value



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

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