Displaying Data

KS2

KS4

1) Key Words
2) Tally Charts
3) Pictograms
4) Block Graph
5) Bar Graphs
6) Pie Charts
7) Grouped Tally Charts (KS2/3 analysis)
8) Grouped Frequency Diagrams
9) Frequency Polygons
10) Line Graphs
11) Scatter Diagrams
12) Cumulative Frequency Diagrams
13) Box Plots
14) Histograms
15) Grouped Tally Charts (KS4 analysis)
16) What Makes A Good Graph

* Analysing Data

Central South Consortium
Consortiwm Canolbarth y De
Joint Education Service
Gwasanaeth Addysg ar y Cyd

Key words

Axes

Continuous

Correlation

Data

Discrete

Frequency

Grouped

Interquartile

Labels

Linear

Median

Origin

Plot

Scale

x -axis

y -axis

Title

Tally

Types of data

Discrete data	can only take specific values, e.g. siblings, key stage 3 levels, numbers of objects
Continuous data	can take any value, e.g. height, weight, age, time, etc.

Tally Chart

A tally chart is used to organise data from a list into a table.

The data shows the number of children in each of 30 families.

~~2~~, ~~1~~, 5, 0, 2, 1, 3, 0, 2, 3, 2, 4, 3, 1, 2, 3, 2, 1, 4, 0, 1, 3, 1, 2, 2, 6, 3, 2, 2, 3

Number of children in a family	Tally	Frequency
0		
1	|	
2	|	
3		
4 or more		

Year 3/4/5/6:- represent data using: lists, tally charts, tables and diagrams

Tally Chart

This data can now be represented in a Pictogram or a Bar Graph

The data shows the number of children in each family. 30 families were studied.

Number of children in a family	Tally	Frequency
0	III	3
1	𝍸 I	6
2	𝍸 𝍸	10
3	𝍸 II	7
4 or more	IIII	4
Total		**30**

Add up the tally

Check the total is 30

𝍸 = 5

Year 3/4/5/6:- represent data using: lists, tally charts, tables and diagrams

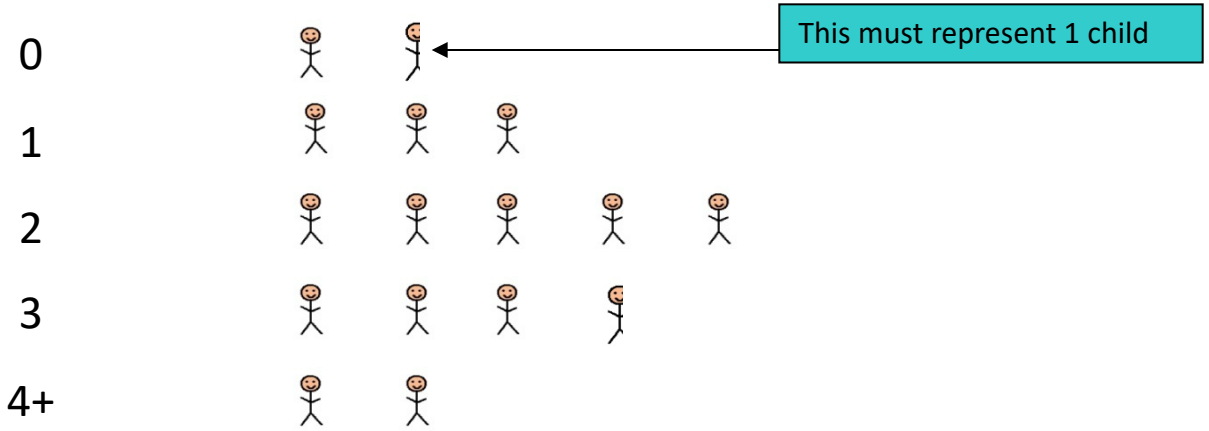
Pictogram

This data could be represented by a Pictogram:

Number of children in a family	Tally	Frequency
0	III	3
1	𝍸 I	6
2	𝍸 𝍸	10
3	𝍸 II	7
4 or more	IIII	4

Key:  **2 children**

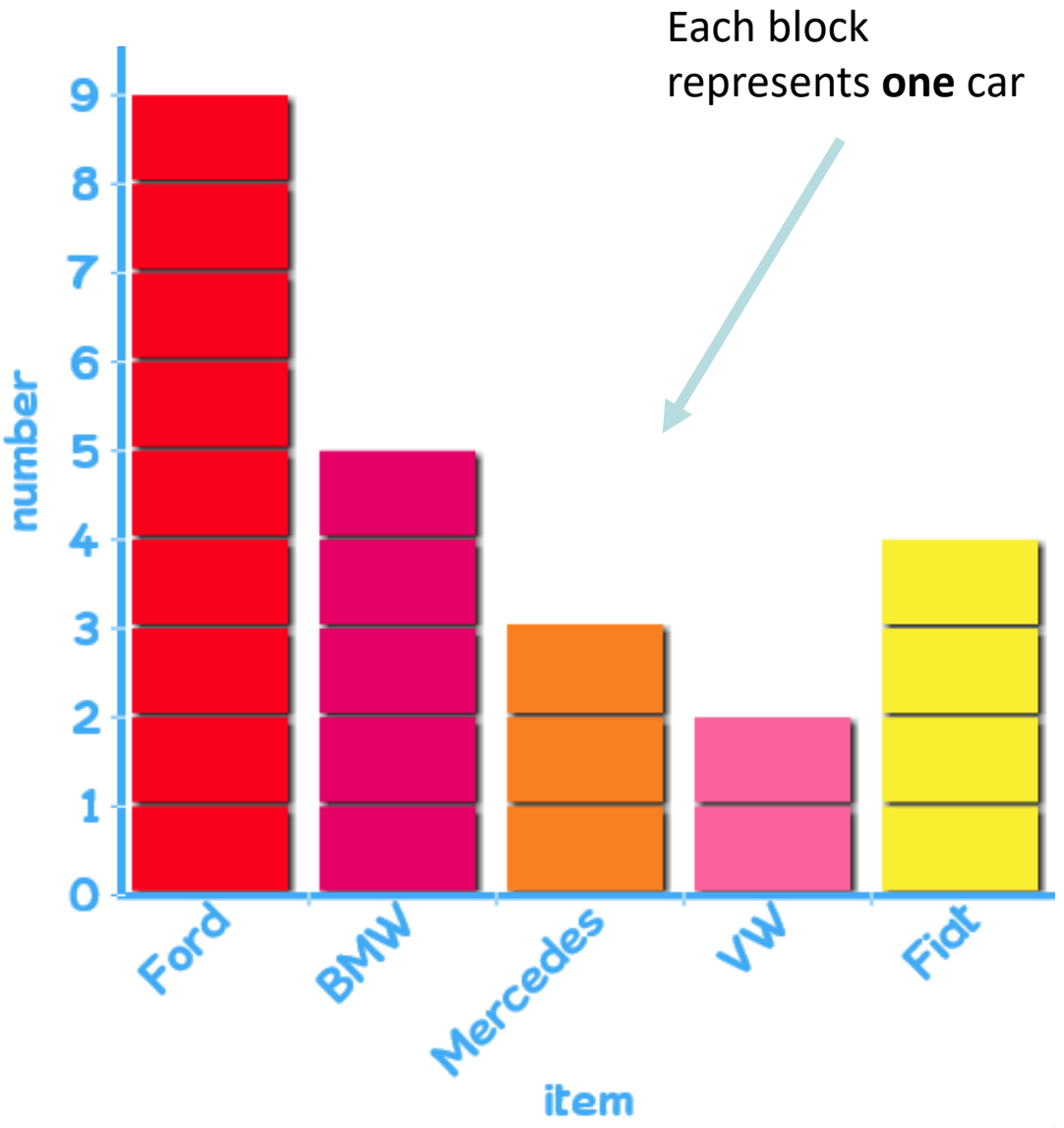
Number of Children



Year 3/4:- represent data using: pictograms where one symbol represents more than one unit using a key

Block Graph

item	number
Ford	9
BMW	5
Mercedes	3
VW	2
Fiat	4



Year 2: gather and record data from: lists and tables, diagrams, block graphs, pictograms where the symbol represents one unit

Bar Graph

The Bar Graph below represents the data in the tally chart as a graph

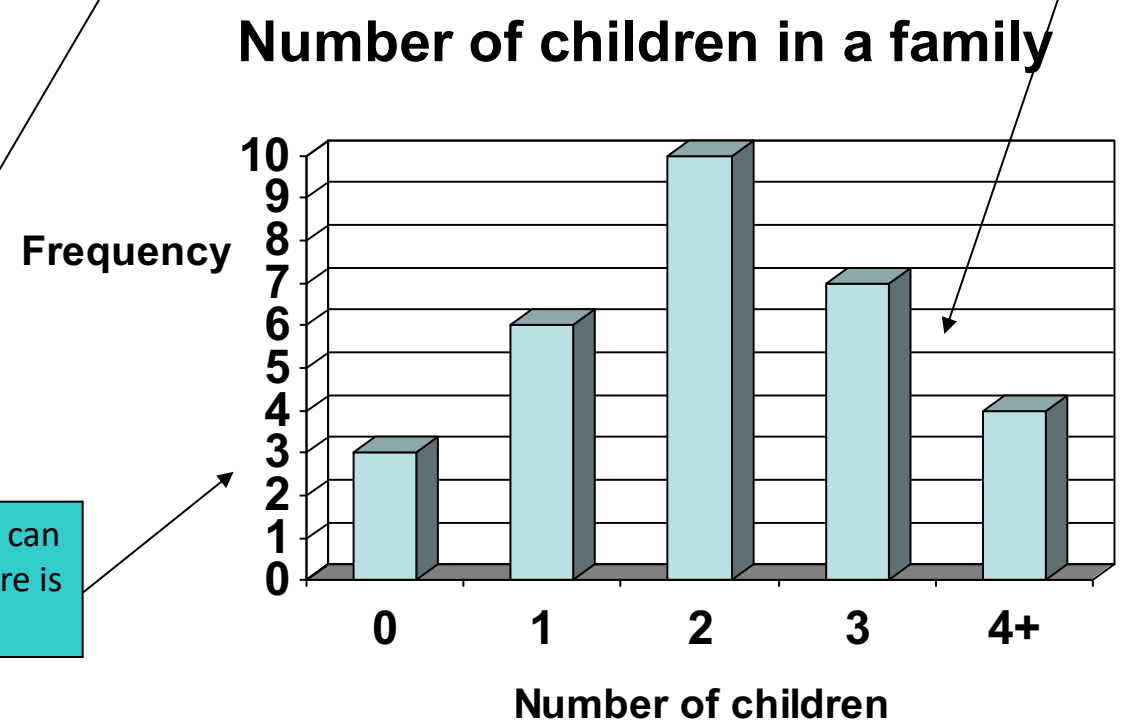
– remember to label each axes!

Number of children in a family	Tally	Frequency
0	III	3
1	IIII I	6
2	𝍸 IIII	10
3	𝍸 II	7
4 or more	𝍸	4

Graph scale must go up to 10

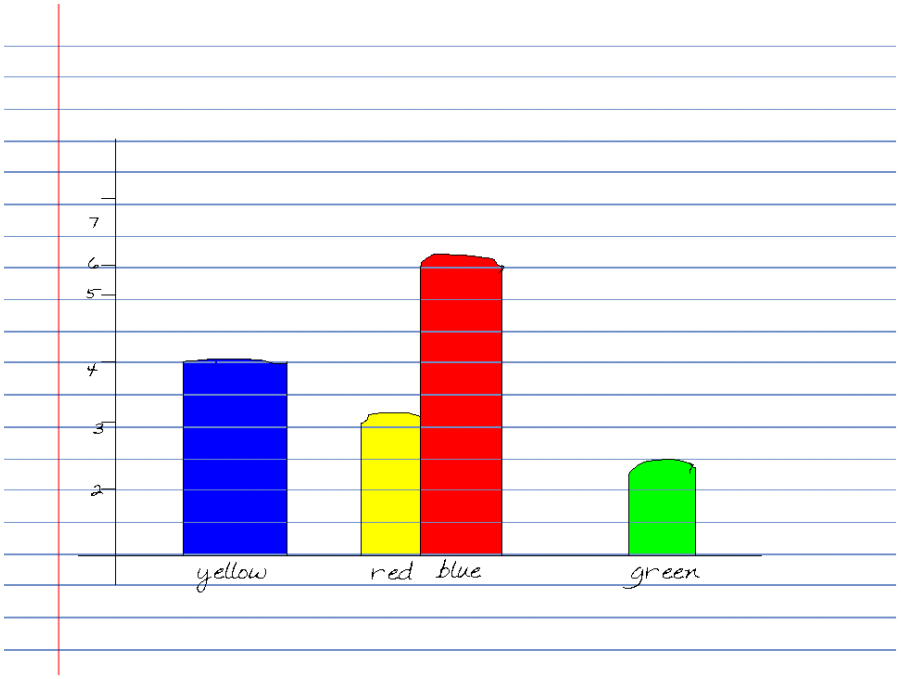
Gaps between bars for different subjects

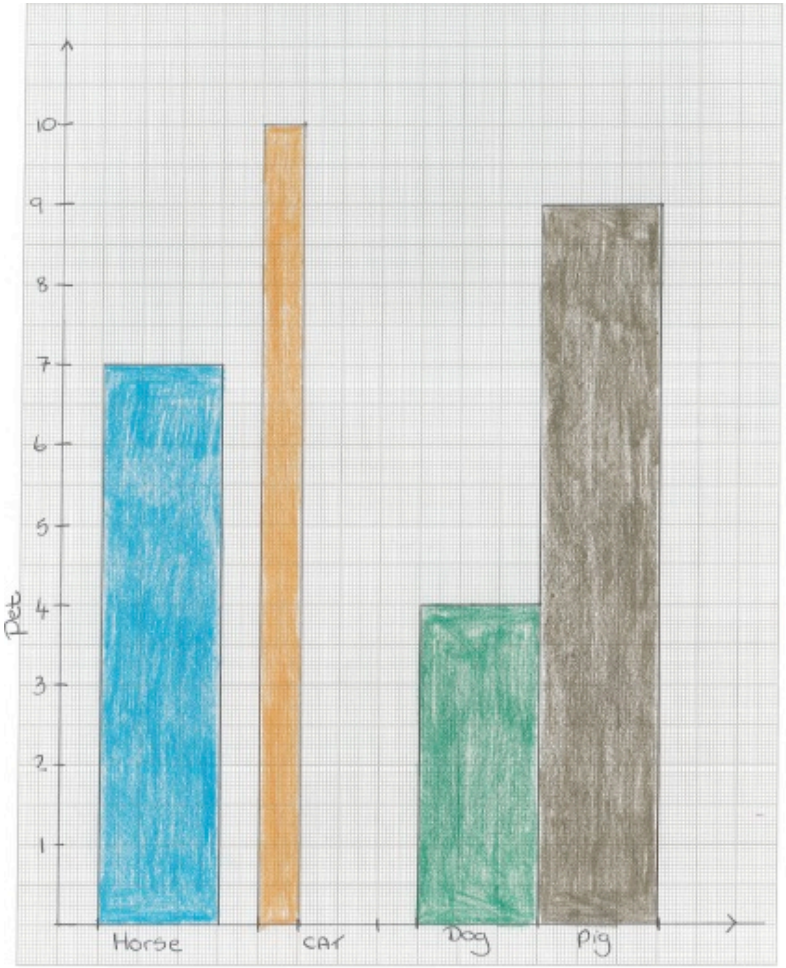
Scale is going up in 1. This can go up in 2, 5 or even 10 if there is more data.



Year 3/4: represent data using: tally charts, bar charts and bar line graphs labelled in 2s, 5s & 10s
Year 5/6: represent data using: tally charts, frequency tables, bar charts

Examples of Incorrect Bar Graphs





Q. Can you spot all that is wrong with these bar graphs?

Gaps v No Gaps

Discrete data	can only take specific values, e.g. siblings, key stage 3 levels, numbers of objects	Gaps between bars
Continuous data	can take any value, e.g. height, weight, age, time, etc.	No gaps between bars

Tally Chart (Grouped Data)

A tally chart can also be used for grouped data.

The data shows the pocket money 30 pupils receive each per week.

~~£1.50~~, ~~£2.00~~, £1.60, £4.00, £5.50, £3.20, £4.60, £1.00, £10.00, £8.00, £6.50, £5.50, £2.50, £3.80, £1.20, £8.00, £7.50, £1.00, £3.50, £7.50, £2.50, £4.60, £5.50, £1.50, £6.50, £2.50, £4.10, £1.00, £7.60, £4.20

As there is a large amount of data, it needs to be grouped.

Make sure the groups do not overlap

£2.00 is between £2.00 and £3.99 so we need a tally in this group.

Pocket Money	Tally	Frequency
£0 - £1.99	|	
£2.00 - £3.99	|	
£4.00 - £5.99		
£6.00 - £7.99		
£8.00 - £9.99		
£10.00 - £11.99		

£1.50 is between £0 and £1.99 so we need a tally in this group.

Year 5/6: represent data using: lists, tally charts, tables, frequency tables, grouped data charts

Tally Chart (Grouped Data)

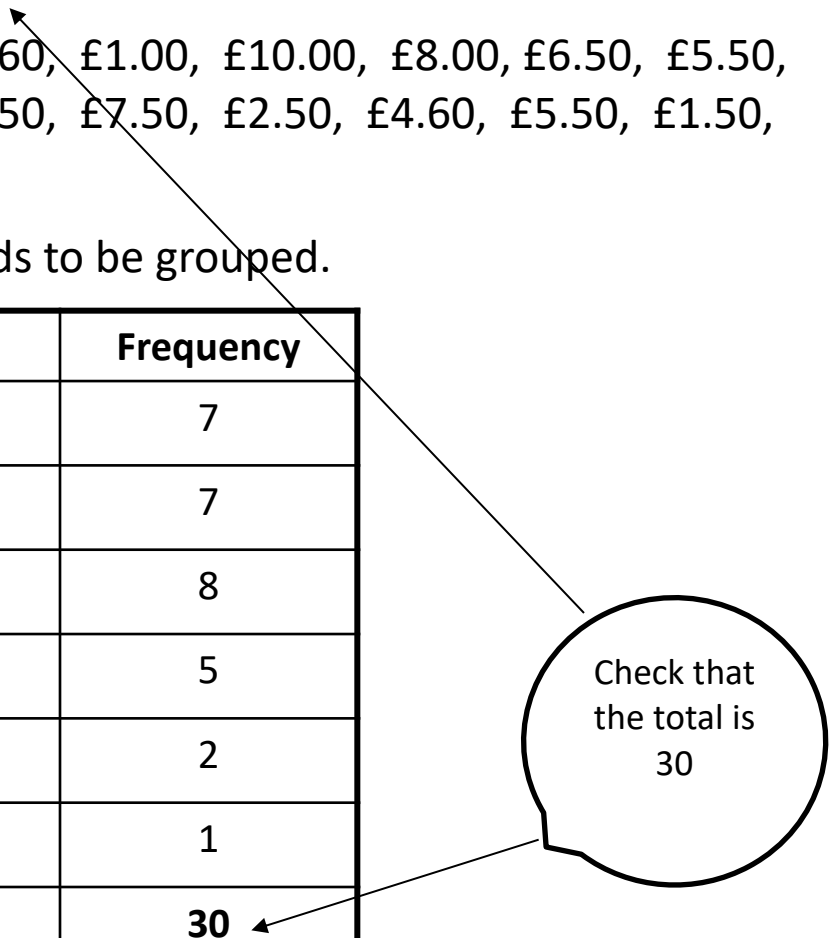
A tally chart can also be used for grouped data.

The data shows the pocket money 30 pupils receive each per week.

£1.50, £2.00, £1.60, £4.00, £5.50, £3.20, £4.60, £1.00, £10.00, £8.00, £6.50, £5.50, £2.50, £3.80, £1.20, £8.00, £7.50, £1.00, £3.50, £7.50, £2.50, £4.60, £5.50, £1.50, £6.50, £2.50, £4.10, £1.00, £7.60, £4.20

As there is a large amount of data, it needs to be grouped.

Pocket Money	Tally	Frequency
£0 - £1.99	𝍸 II	7
£2.00 - £3.99	𝍸 II	7
£4.00 - £5.99	𝍸 III	8
£6.00 - £7.99	𝍸	5
£8.00 - £9.99	II	2
£10.00 - £11.99	I	1
TOTAL		**30**



Grouped Frequency Diagram

From this data we can construct a grouped frequency diagram.

Pocket Money	Tally	Frequency
£0 - £1.99	𝍸 II	7
£2.00 - £3.99	𝍸 II	7
£4.00 - £5.99	𝍸 III	8
£6.00 - £7.99	𝍸	5
£8.00 - £9.99	II	2
£10.00 - £11.99	I	1
TOTAL		**30**

The biggest bar will be 8

The bars will touch

Both axis will need a scale

KS2/3 Analysis

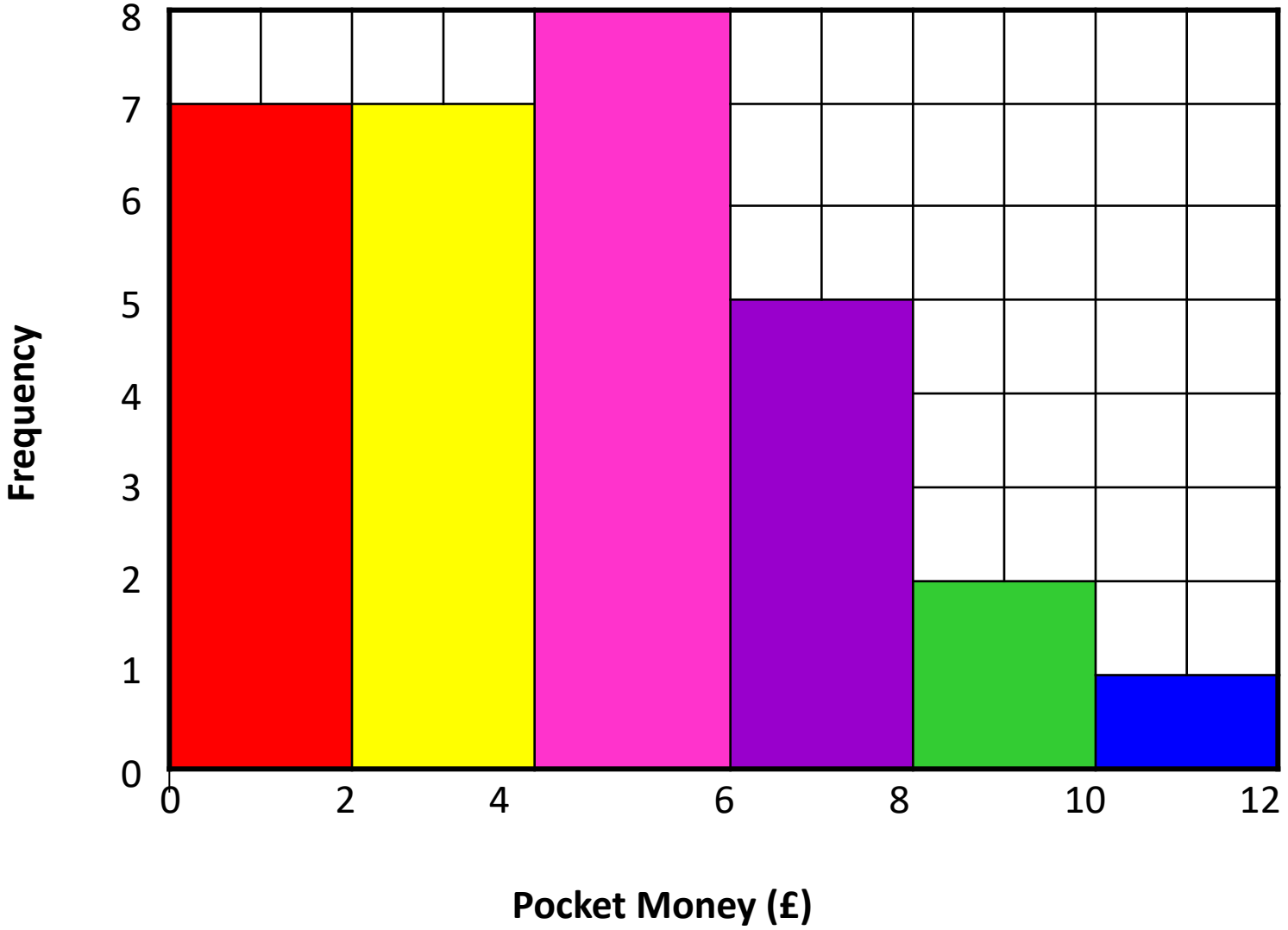
Modal Group = £4.00 - £5.99

Range = £11.99 - £0 = £11.99

Year 7: construct frequency tables for sets of data, grouped where appropriate, in equal class intervals (**groups given to learners**)

Year 8: construct frequency tables for sets of data in equal class intervals, **selecting groups as appropriate**

Weekly Pocket Money



Pie Chart

Non- calculator

A pie chart is a circular chart divided into sectors, illustrating numerical proportion!

Example:

30 people were asked which newspapers they read regularly.

The results were :

Newspaper	Number of people
The Guardian	8
Daily Mirror	7
The Times	3
The Sun	6
Daily Express	6

Pie Chart

Non- calculator

There are 30 people in the survey and **360º** in a full pie chart.

Each person is therefore represented by **360º ÷ 30 = 12º**

360º ÷ Total Frequency

We can now calculate the angle for each category:

Newspaper	No of people	Working	Angle
The Guardian	8	8 × 12º	96º
Daily Mirror	7	7 × 12º	84º
The Times	3	3 × 12º	36º
The Sun	6	6 × 12º	72º
Daily Express	6	6 × 12º	72º
Total	30		360º

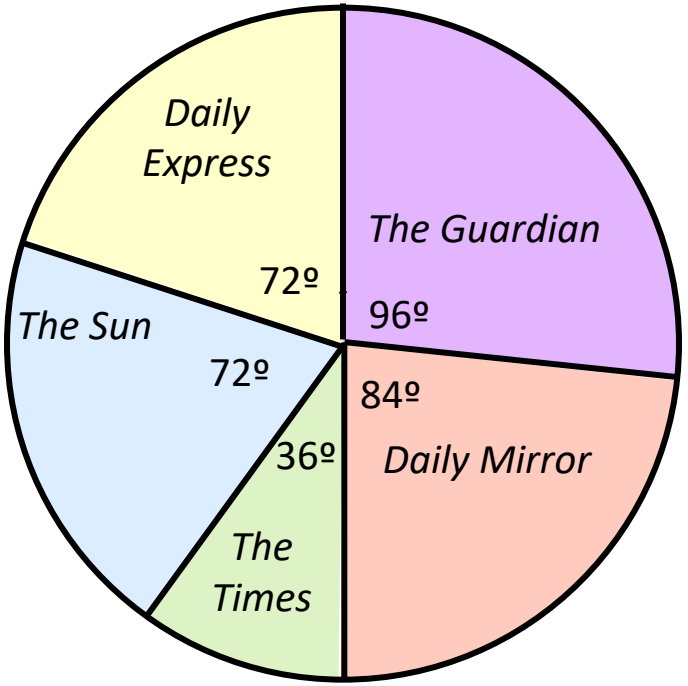
Year 9: **construct** and interpret graphs and diagrams (including **pie charts**) to represent discrete or continuous data, with the learner choosing an appropriate scale.

Pie Chart

Non- calculator

Once the angles have been calculated you can draw the pie chart:

- Start by drawing a circle using compasses.
- Draw a radius from the centre to the **top of the circle**.
- Measure an angle of 96º from the radius using a protractor and label the sector (go clockwise around)
- Measure an angle of 84º from **the last line** you drew and label the sector.
- Repeat for each sector until the pie chart is complete.



Year 9: **construct** and interpret graphs and diagrams (including **pie charts**) to represent discrete or continuous data, with the learner choosing an appropriate scale.

Pie Chart

Calculator

The number of seats won by the political parties in the May 2010 general election is shown in the table below.

Party	**Number of seats**
Conservatives	306
Labour	268
Liberal Democrats	57
Other	29

Draw a pie chart, as accurately as possible, to show this information.
Show how you calculate the angles of your pie chart.

Year 9: **construct** and interpret graphs and diagrams (including **pie charts**) to represent discrete or continuous data, with the learner choosing an appropriate scale.

Pie Chart

Calculator

There are 650 people in the survey and **360º** in a full pie chart.

Each person is therefore represented by **360º ÷ 650 = 0.5538461538º**

360º ÷ Total Frequency

We can now calculate the angle for each category using a **calculator**.

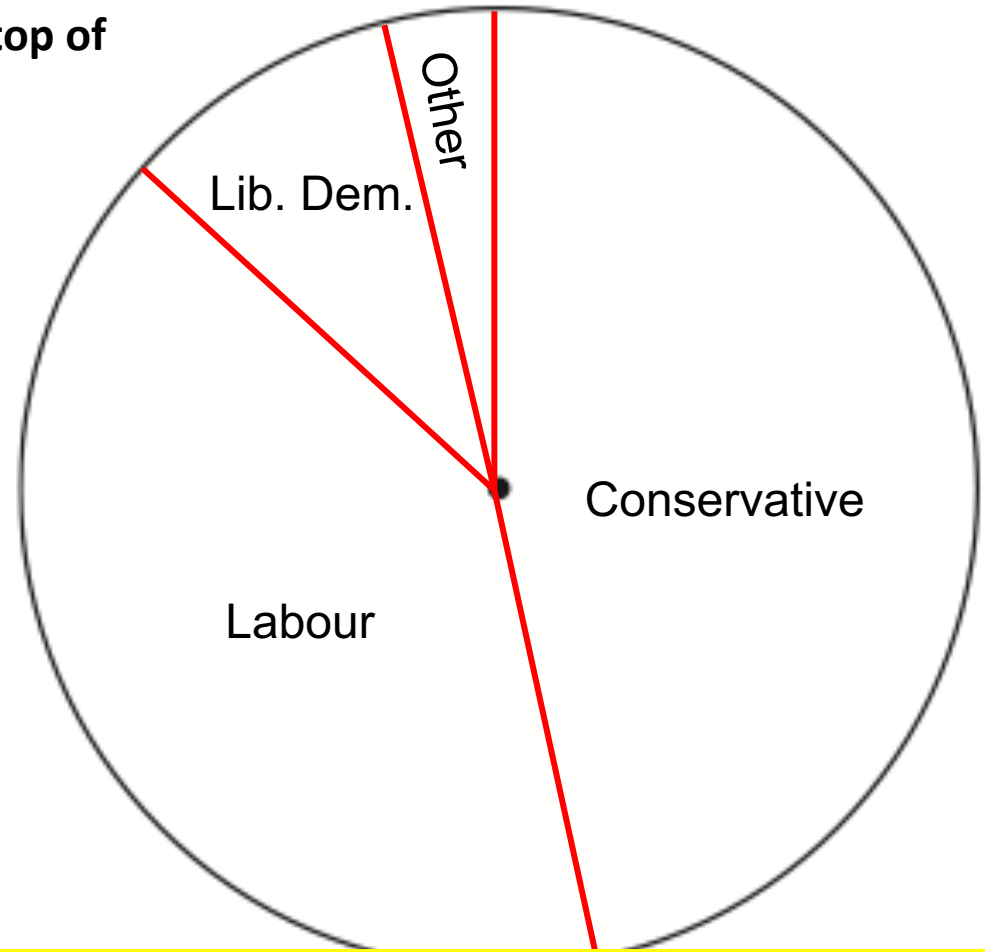
Remember to use the full calculator display for accuracy!

Round angles to the **nearest whole number**. Check they add to 360º

Party	Number of seats	Calculation	Angle
Conservatives	306	306 x 0.553	169.476 = **169º**
Labour	268	258 x 0.553	142.892 = **143º**
Liberal Democrats	57	57 x 0.553	31.569 = **32º**
Other	29	29 x 0.553	16.061 = **16º**
	total= 650		total = 360

Pie Chart

- Start by drawing a circle using compasses.
- Draw a radius from the centre to the **top of the circle**.
- Measure an angle of 169º from the radius using a protractor and label the sector (go clockwise around)
- Measure an angle of 143º from **<u>the last line</u>** you drew and label the sector.
- Repeat for each sector until the pie chart is complete.

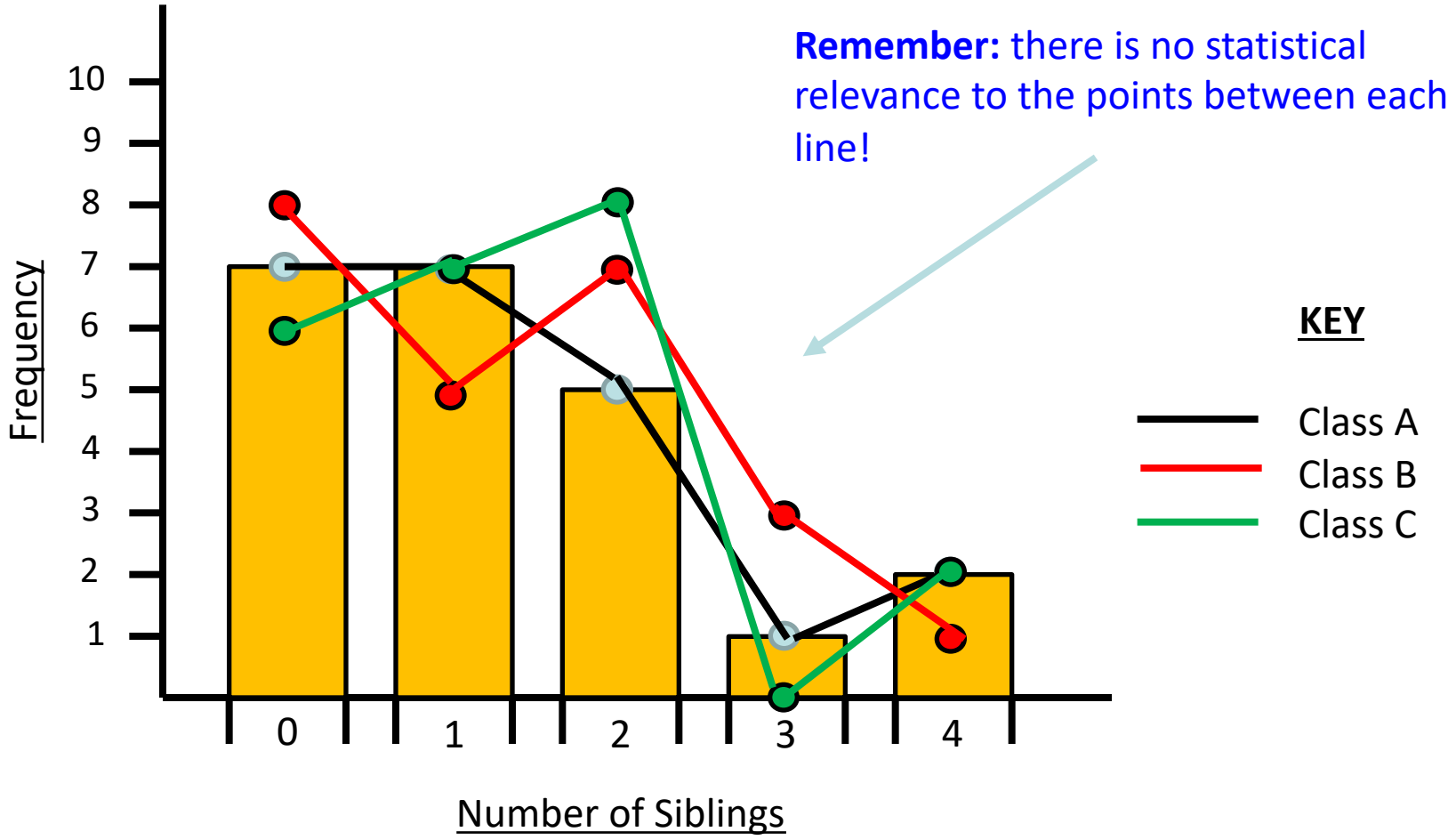


Year 9: construct and interpret graphs and diagrams (including **pie charts**) to represent discrete or continuous data, with the learner choosing an appropriate scale.

Frequency Polygons

- Often a frequency polygon is a more straightforward and better way of comparing different sets of data compared to a bar chart
- A frequency polygon shows the trend of the data
- For grouped data, you always plot the **midpoint** of the group against the frequency.

Example of converting a bar graph to a frequency polygon with discrete data

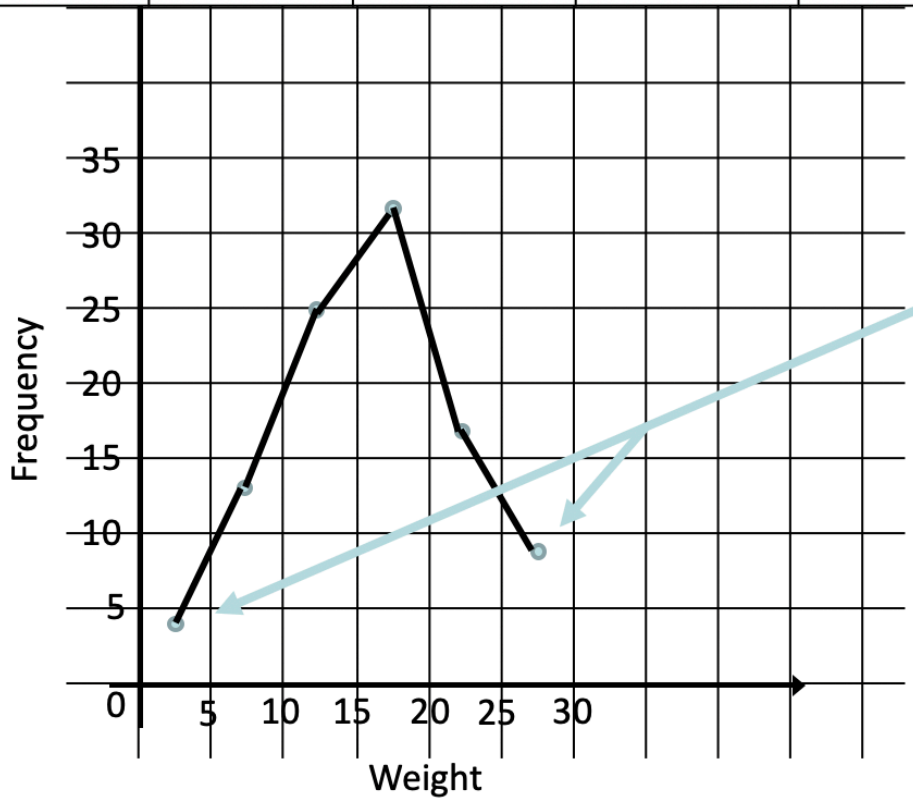


Year 5/6: extract and interpret information from an increasing range of diagrams, timetables and graphs (including pie charts)

The weight of 100 dogs at a dogs home are shown in the table below.

Weight	$0 < w \leq 5$	$5 < w \leq 10$	$10 < w \leq 15$	$15 < w \leq 20$	$20 < w \leq 25$	$25 < w \leq 30$
Frequency	4	13	25	32	17	9
Midpoints	2.5	7.5	12.5	17.5	22.5	27.5

These must be calculated



Do not join these two points up!

Year 5/6: extract and interpret information from an increasing range of diagrams, timetables and graphs (including pie charts)

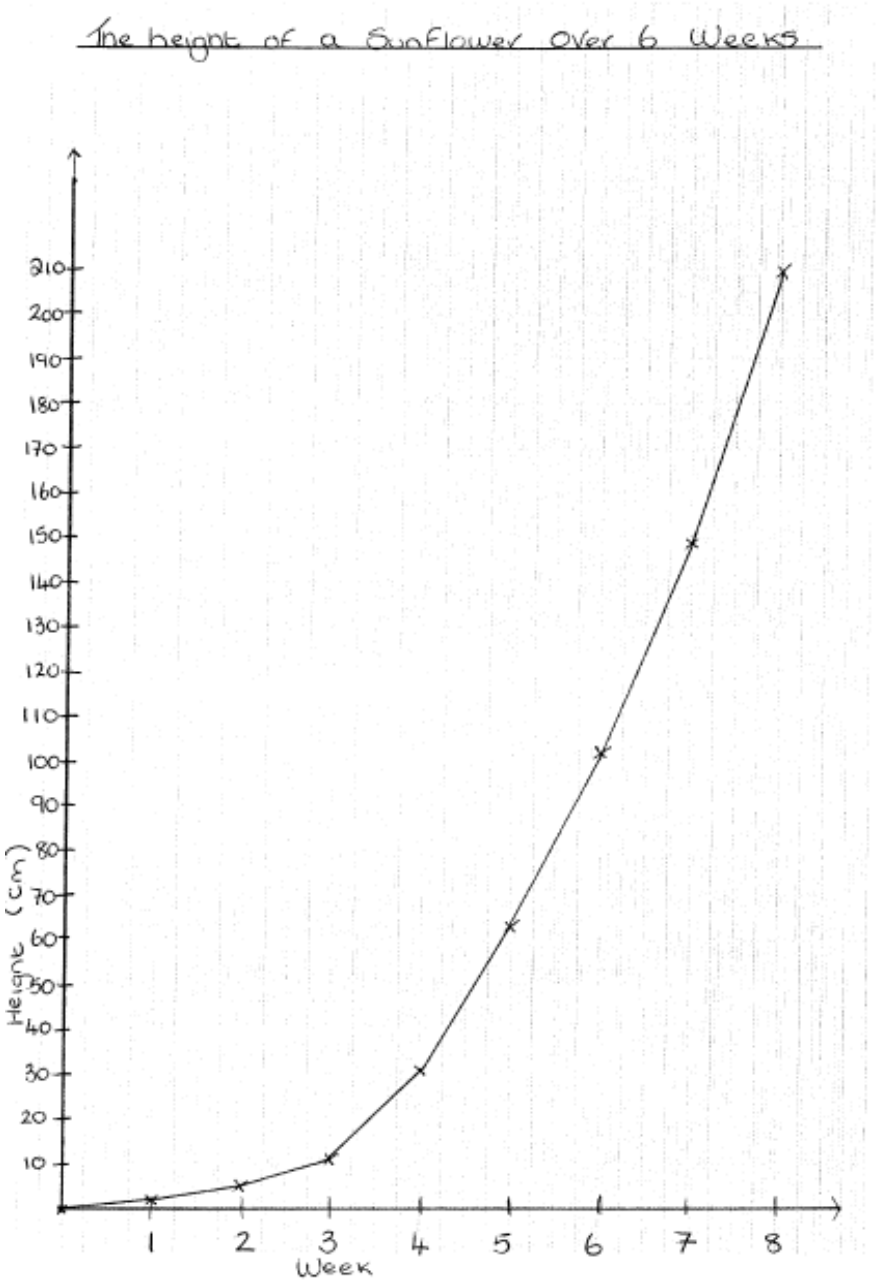
Year 7: construct a wide range of graphs and diagrams to represent the data and reflect the importance of scale

Year 8: construct a wide range of graphs and diagrams to represent discrete and continuous data

Line Graph

- A line graph is a way of representing data. A diagram which shows how two sets of information are related, in the form of a line
- Suitable for continuous data
- Useful when showing a pattern over time.

Line Graph


The height of a Sunflower over 6 Weeks
Height (cm)
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
1
2
3
4
5
6
7
8
Week

Everyday examples of line graphs

- A health visitor plots and displays the weight of a baby over 12 months
- Hours of sunshine in a travel brochure
- Changes in stocks and shares.

Scatter Diagrams

Scatter diagrams show the relationship between two sets of data. Points are plotted very much like co-ordinates.

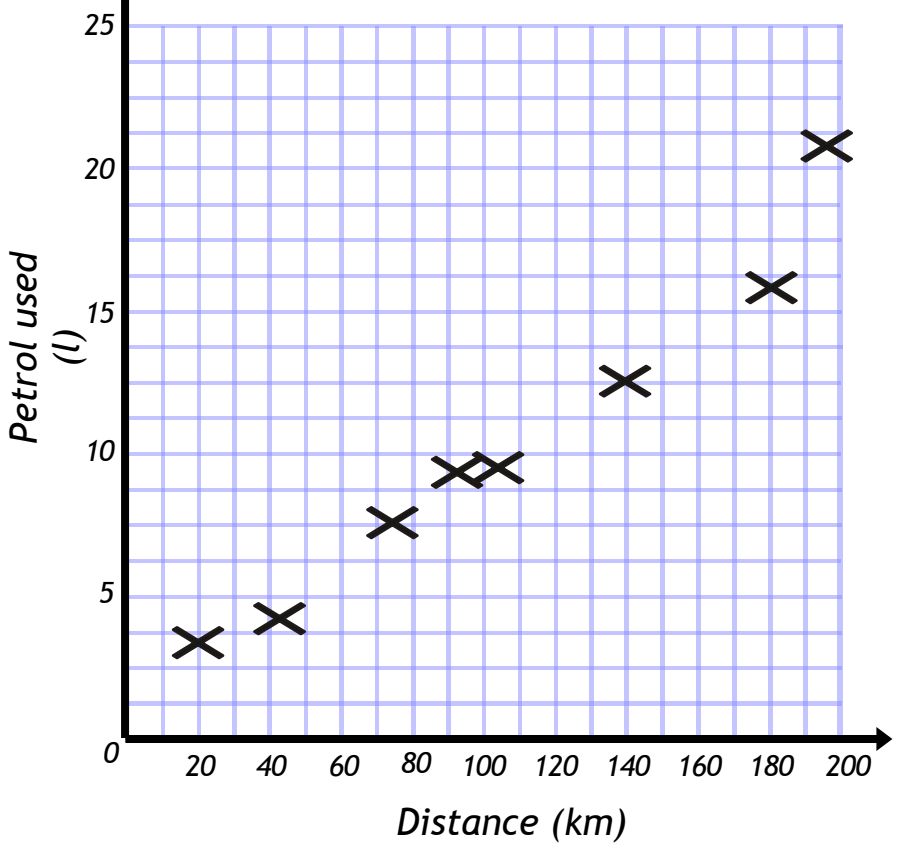
Below is in the information taken from 8 different car journeys.

Journey	1	2	3	4	5	6	7	8
Distance (km)	75	140	197	180	20	93	104	42
Petrol used (l)	7	12.5	21.3	16	3.5	8.9	9.5	4.1

Year 5/6: extract and **interpret information** from an increasing range of diagrams, timetables and graphs (including pie charts)
Year 7: **interpret** diagrams and graphs (including pie charts)
Year 8: **construct** graphs to represent data including **scatter diagrams** to investigate correlation

How a scatter graph should look

Journey	1	2	3	4	5	6	7	8
Distance (km)	75	140	197	180	20	93	104	42
Petrol used (l)	7	12.5	21.3	16	3.5	8.9	9.5	4.1



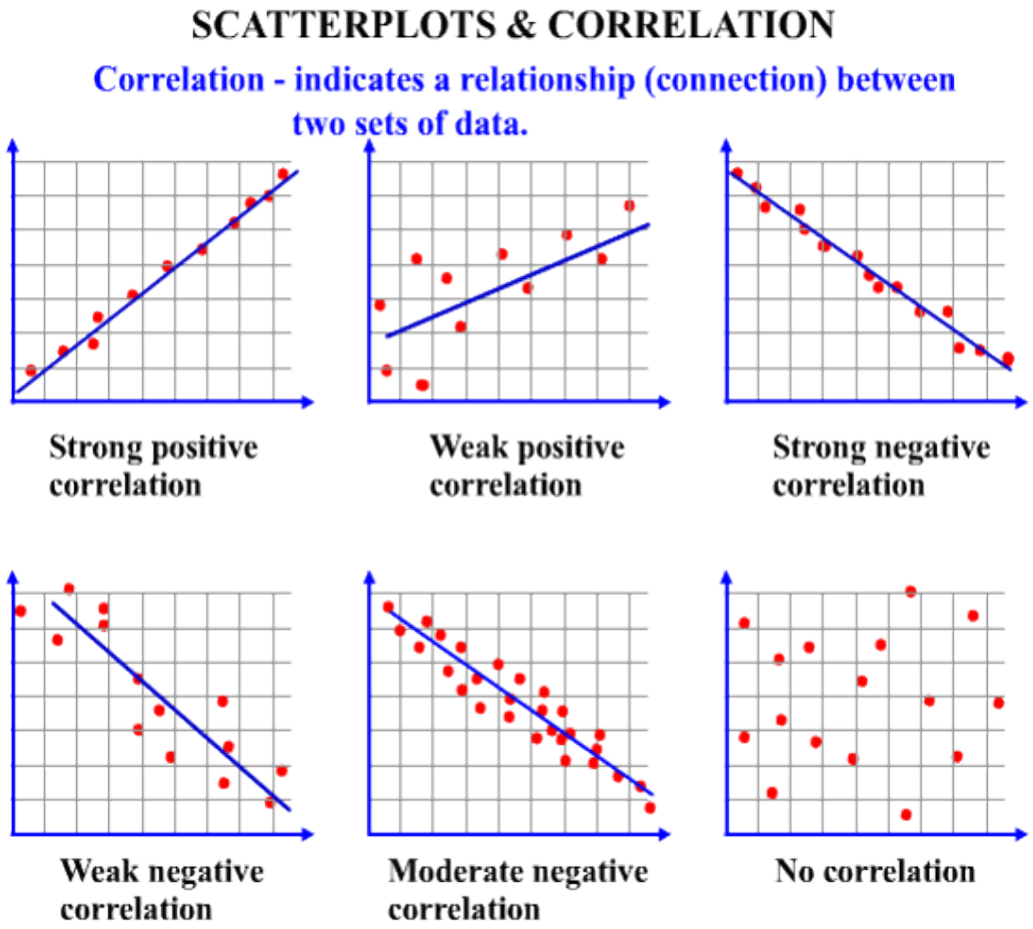
Year 5/6: extract and **interpret information** from an increasing range of diagrams, timetables and graphs (including pie charts)

Year 7: **interpret** diagrams and graphs (including pie charts)

Year 8: construct graphs to represent data including **scatter diagrams** to investigate correlation

What do scatter graphs tell us?

Scatter graphs show the relationship between our two sets of data.
We describe this relationship using correlation.
There are basically 3 types of correlation:
Positive, Negative and No Correlation



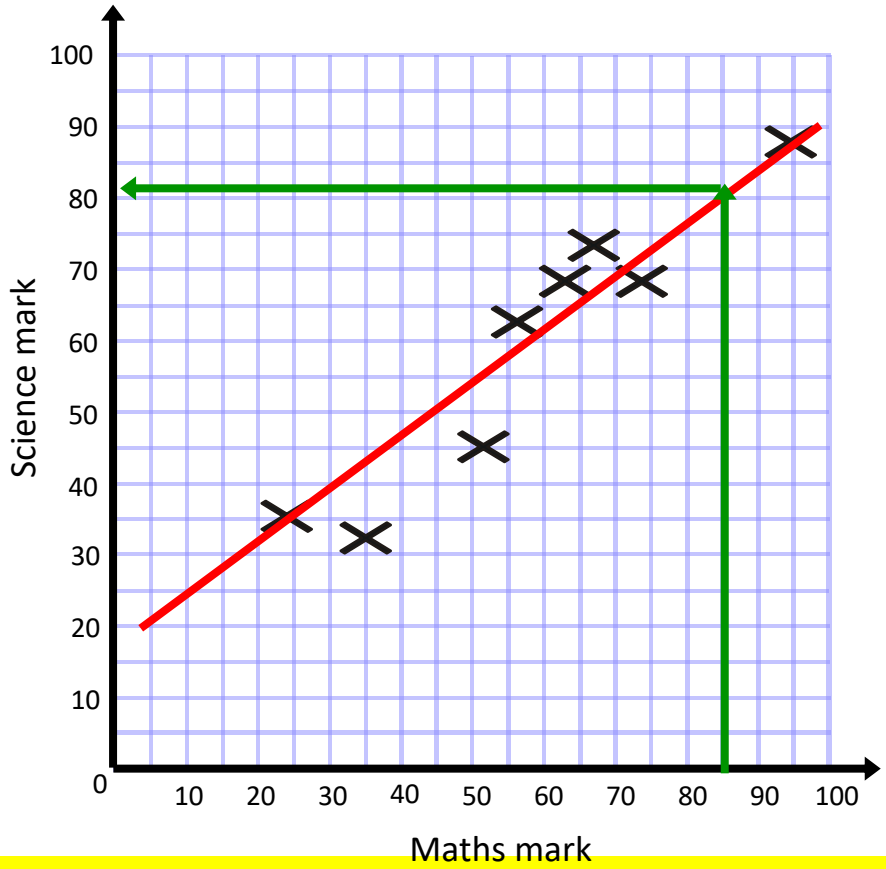
The more in line the points, the stronger the correlation.

How do we use scatter graphs?

- We can use scatter graphs to estimate results based upon other results. We do this by drawing a "line of best fit".
- A "line of best fit" is a straight line (drawn with a ruler) that goes through as many of the plotted points as possible. This is an estimate, but try to get half of the points on either side of it **and** go through at least two points exactly.
- The "line of best fit" **does not** have to go through the origin!
- If you have calculated a ‘mean point’ then the line of best fit must go through this point.
- When doing an estimate from a scatter graph you must draw the "line of best fit" and read from the line onto the axis.
- If there is no correlation, you cannot draw a "line of best fit".

Example of a scattergraph

Imogen missed the science test because she was ill. She sat the maths test and got 85.
Use your scatter graph to estimate what Imogen would have achieved on the science test.



We *expect* Imogen would have got a score of 80 in the science test.

Year 10: construct and interpret graphs and diagrams (including pie charts) to represent discrete or continuous data, with the learner choosing the most appropriate representation, including frequency polygons and lines of best fit on scatter diagrams

Year 10: use a scatter diagram to make predictions about the data from a line of best fit drawn by eye v understand the effects of extrapolation and interpolation on reliability

Year 11: use a scatter diagram to make predictions about the data from a line of best fit that passes through the mean

Cumulative Frequency Diagrams

The weights of dogs coming into a vets is shown below. Calculate an estimate for the median and quartiles.

Weight, w (kg)	Frequency	Cumulative Frequency
$0 \leq w < 5$	8	8
$5 \leq w < 10$	23	31
$10 \leq w < 15$	20	51
$15 \leq w < 20$	10	61
$20 \leq w < 25$	5	66
$25 \leq w < 30$	6	72

Add up each frequency in turn

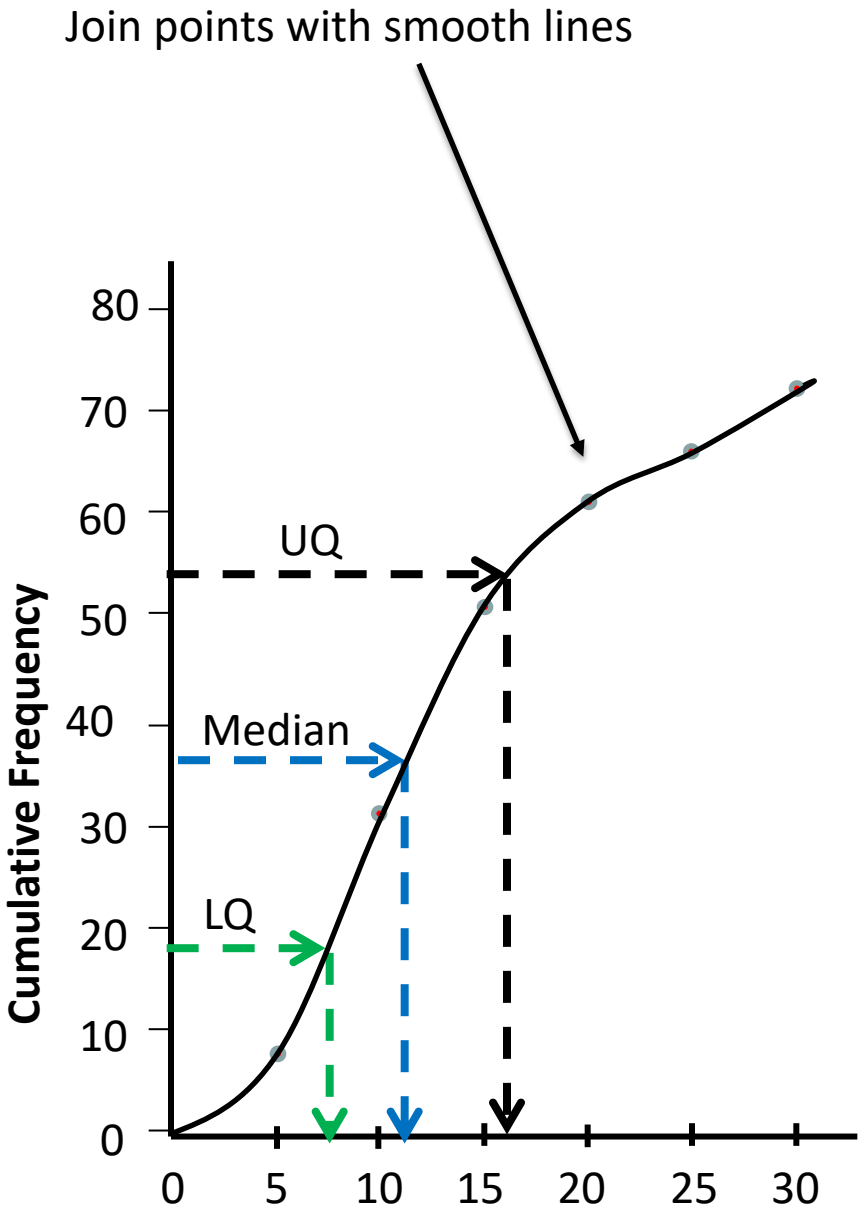
Median Position = Total Frequency x 0.5

Lower Quartile (LQ) Position = Total Frequency x 0.25

Upper Quartile (UQ) Position = Total Frequency x 0.75

Inter Quartile Range (IQR) = UQ – LQ

*Read *across* from these positions to the cumulative frequency curve, then read down to the x-axis for the values.



Weight, w (kg)	Frequency	Cumulative Frequency
$0 \leq w < 5$	8	**8**
$5 \leq w < 10$	23	**31**
$10 \leq w < 15$	20	**51**
$15 \leq w < 20$	10	**61**
$20 \leq w < 25$	5	**66**
$25 \leq w < 30$	6	**72**

Median Position = $72 \div 2 = 36^{th}$ value

Median = 11kg

LQ Position = $\frac{1}{4} \times 72 = 18^{th}$ value

LQ = 8kg

UQ Position = $\frac{3}{4} \times 72 = 54^{th}$ value

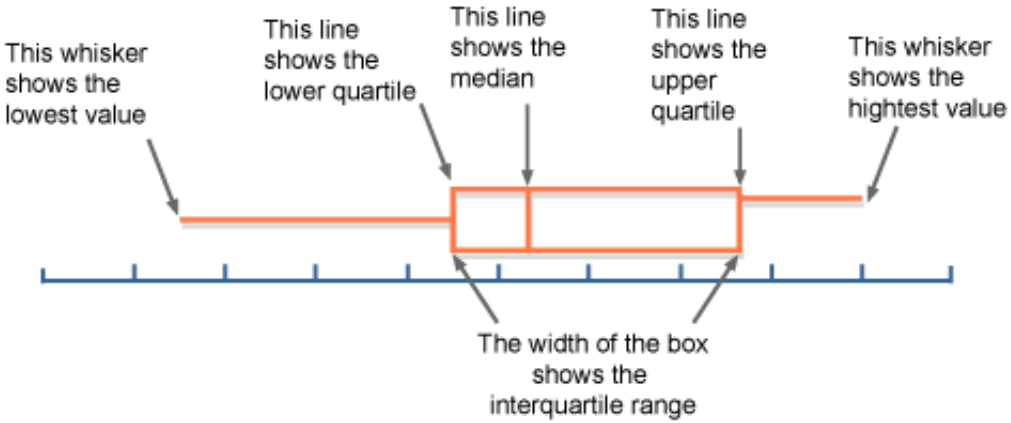
UQ = 16kg

IQR = 16 – 8 **= 8kg**

Year 11: use a cumulative frequency curve to estimate the median, quartiles and interquartile range

Box Plots

A **box and whisker diagram** (also know as a **box plot**) is used to display information about the range, the median and the quartiles. It is usually drawn alongside a number line, as shown;

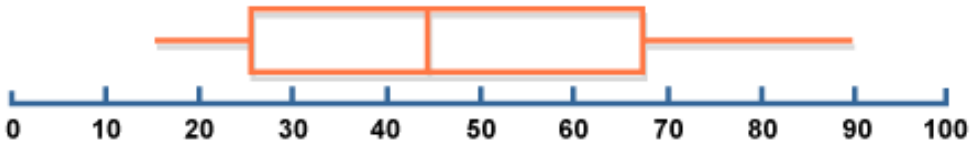


Example

The oldest person in Mathsminster is 90. The youngest person is 15.

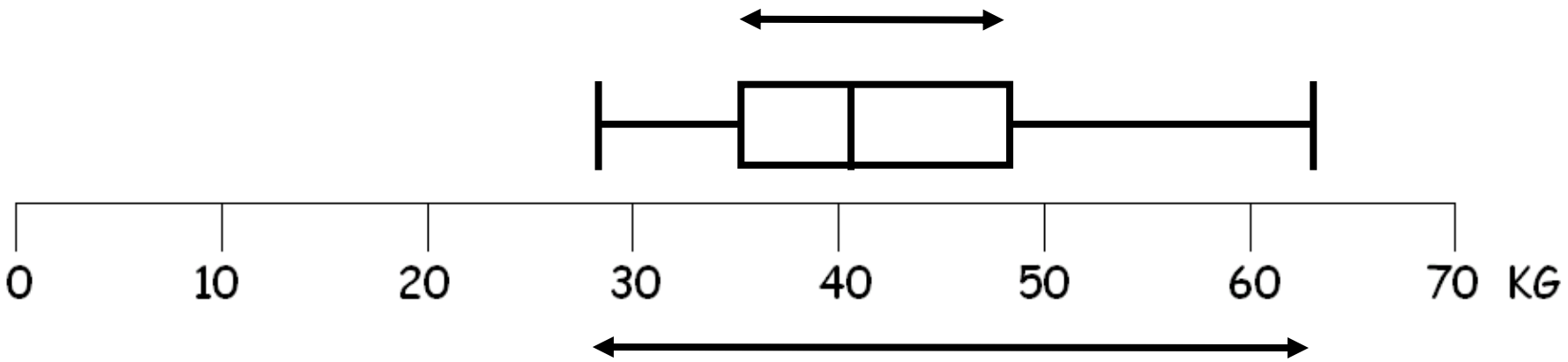
The median age of the residents is 44, the lower quartile is 25, and the upper quartile is 67.

Represent this information with a box-and-whisker plot.



Year 11: construct and interpret graphs and diagrams (including pie charts) to represent discrete or continuous data, with the learner choosing the most appropriate representation, including cumulative frequency curves and boxplots

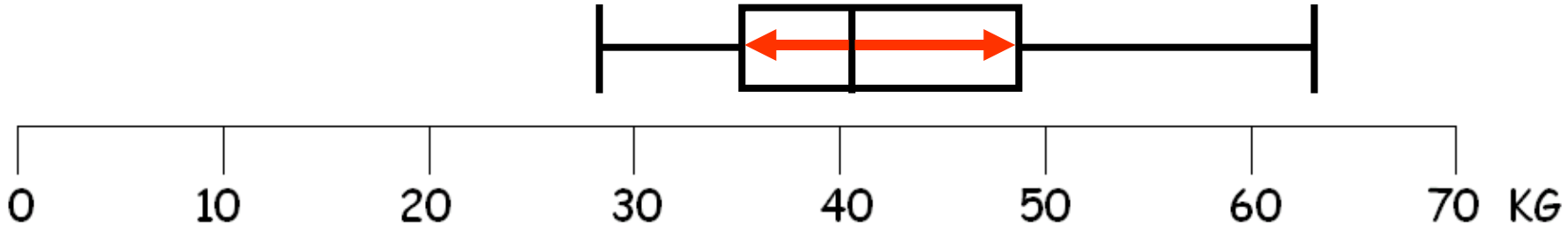
UQ – LQ = Interquartile Range (IQR)



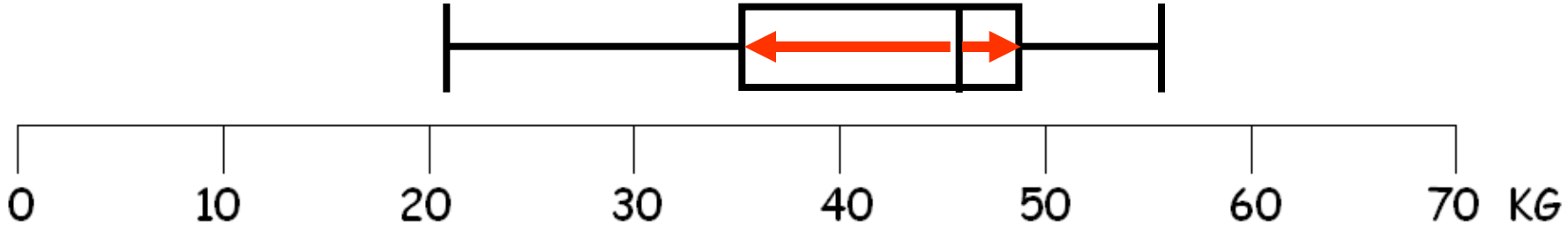
Max – Min = Range

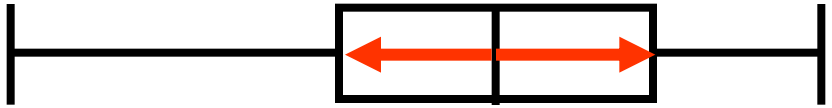
Terminology

Positive skew: median closer to LQ than UQ



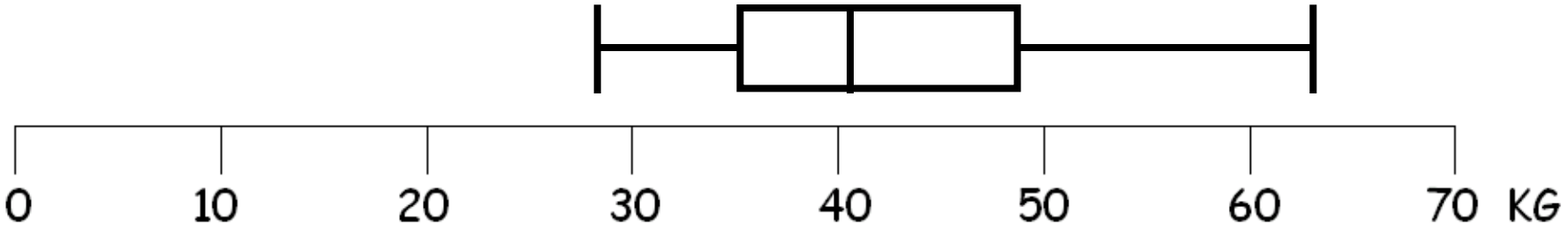
Negative skew: median closer to UQ than LQ





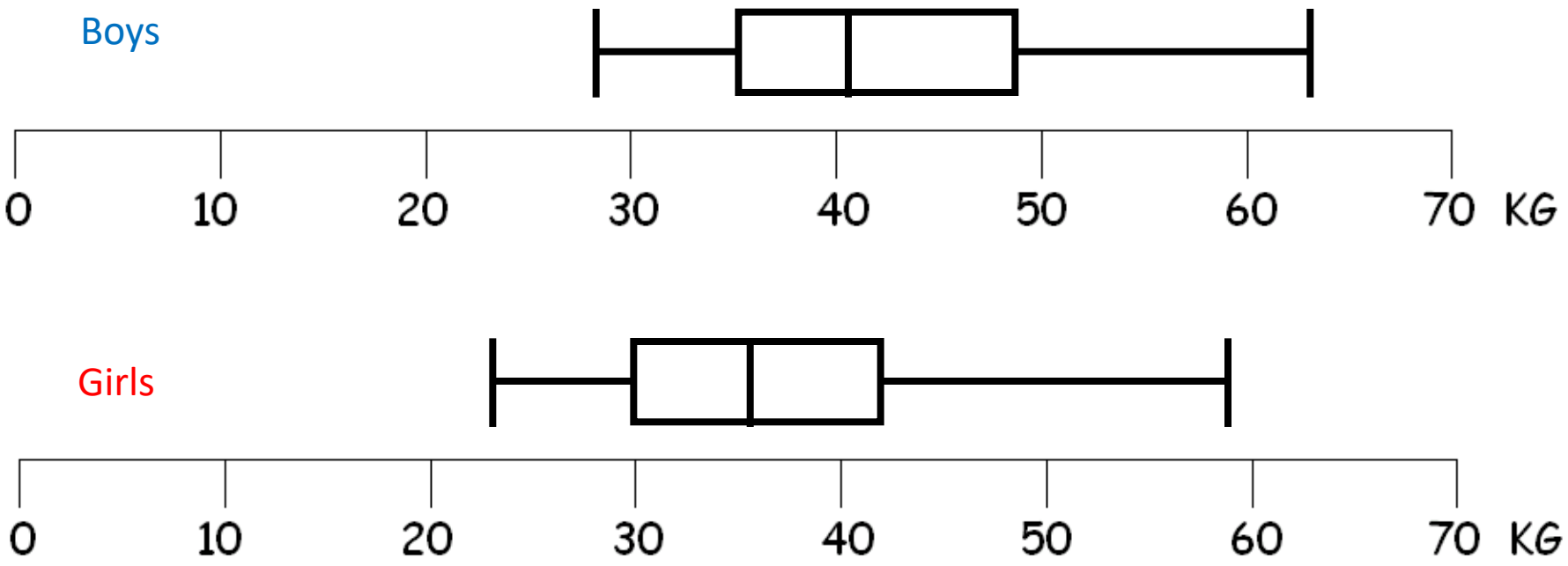
Symmetrical distribution

Interpreting the box plot



- Easily see lightest / heaviest and range
- The ‘box’ contains the middle 50% of people (the most ‘representative half’)
- The ‘whiskers’ show the lightest 25% and heaviest 25% of people (extremes)

Comparing groups



"Lightest girl lighter than lightest boy"

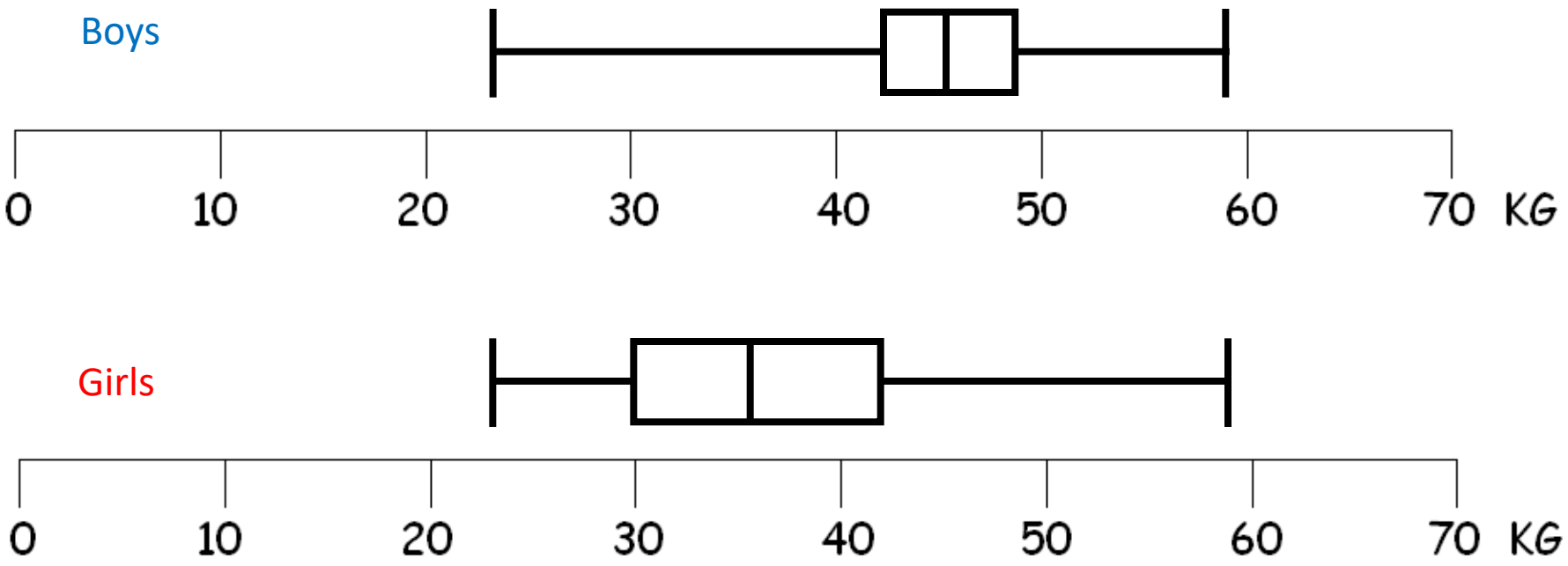
"Heaviest boy heavier than heaviest girl"

"Most representative half of girls generally lighter than most representative half of boys"

Year 11: use the interquartile range to compare distributions

Year 11: compare sets of data and their distributions, using appropriate methods, including those that involve describing central tendency, dispersion, correlation

Comparing groups



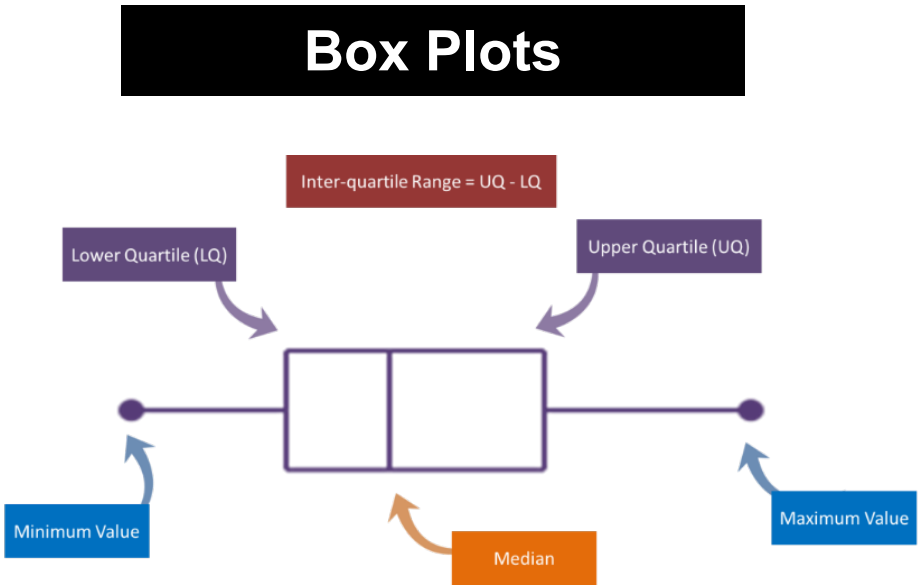
"Lightest girl same as lightest boy"

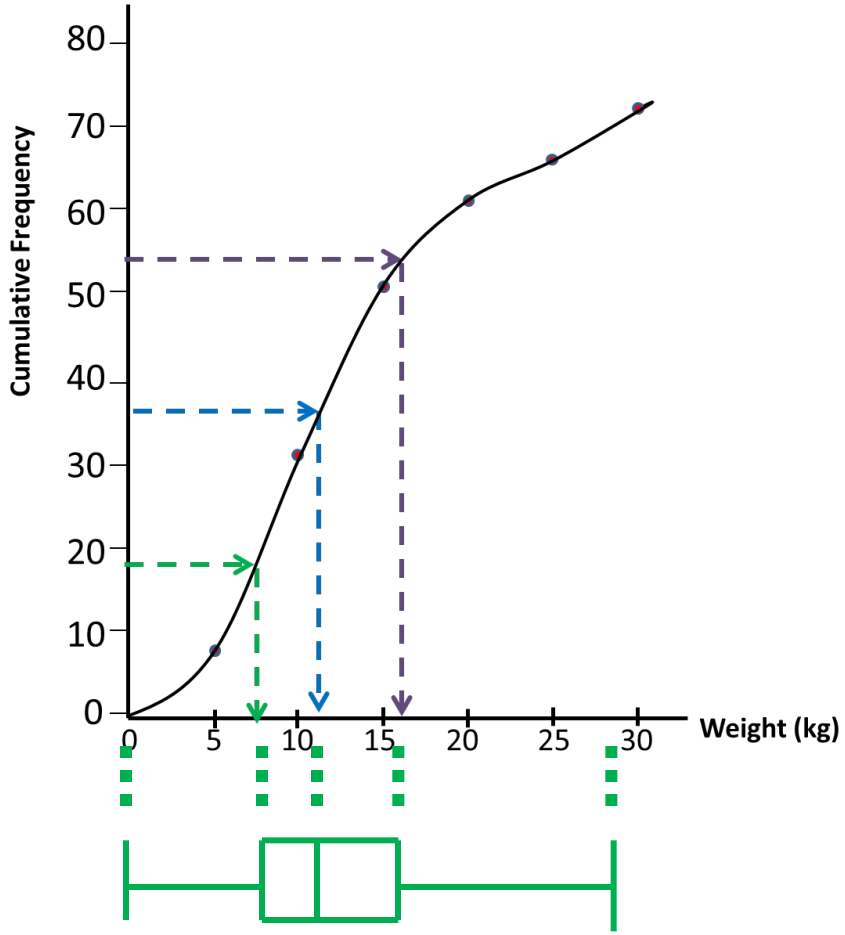
"Heaviest boy same as heaviest girl"

"All of the most representative half of girls
lighter than most representative half of boys"

"Three quarters of girls lighter than
three quarters of boys"

The links between Box Plots and Cumulative Frequency Diagrams





Histograms

- Histograms are used to represent data that is grouped into unequal intervals.

- Remember that in a bar chart the height of the bar represents the frequency. It is therefore correct to label the vertical axis 'frequency'.

- However, as in a **histogram**, it is the **area** which represents the frequency.

- It would therefore be incorrect to label the vertical axis 'frequency' and the label should be '**frequency density**'.

Frequency density = frequency ÷ class width

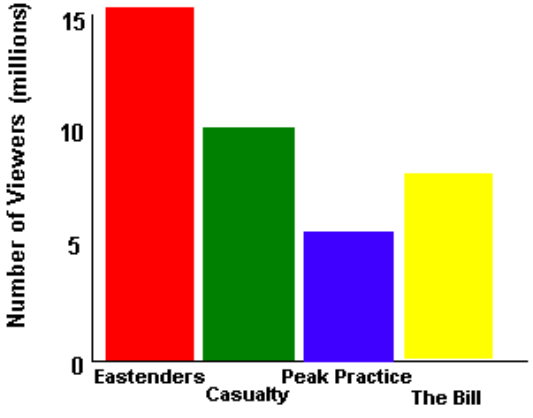
The Differences

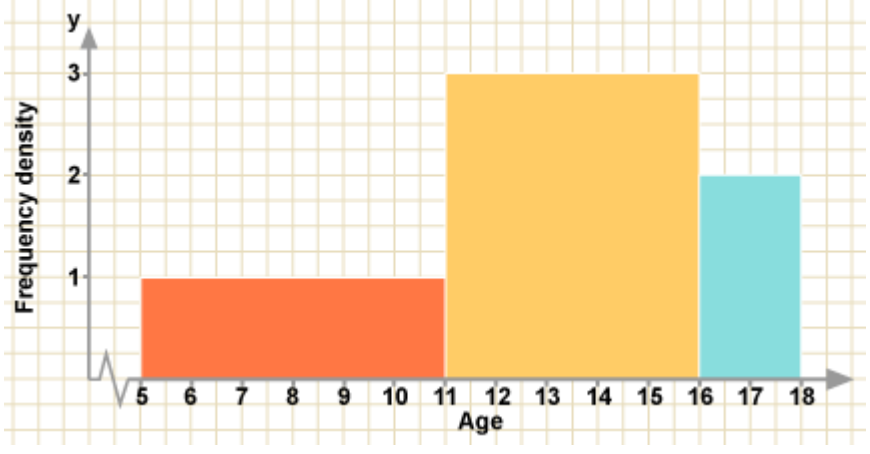
Bar Chart

Category	Property
Bars	There are gaps between the bars
X-Axis	Words or categories (Discrete)
Y-Axis	Number of people or frequency

Histogram

Category	Property
Bars	There are no gaps between the bars and they are different widths
X-Axis	Numbers (Continuous)
Y-Axis	Frequency density – what's that?





Histogram Example

A survey has been conducted on how many hours of TV some children watched last week. Draw a histogram for this data.

Hours (h) spent watching TV last week	Frequency
$0 \leq h < 2$	3
$2 \leq h < 5$	6
$5 \leq h < 10$	10
$10 \leq h < 20$	25
$20 \leq h < 40$	10

We do not plot a bar chart for this data as the groups are of <u>different widths.</u>

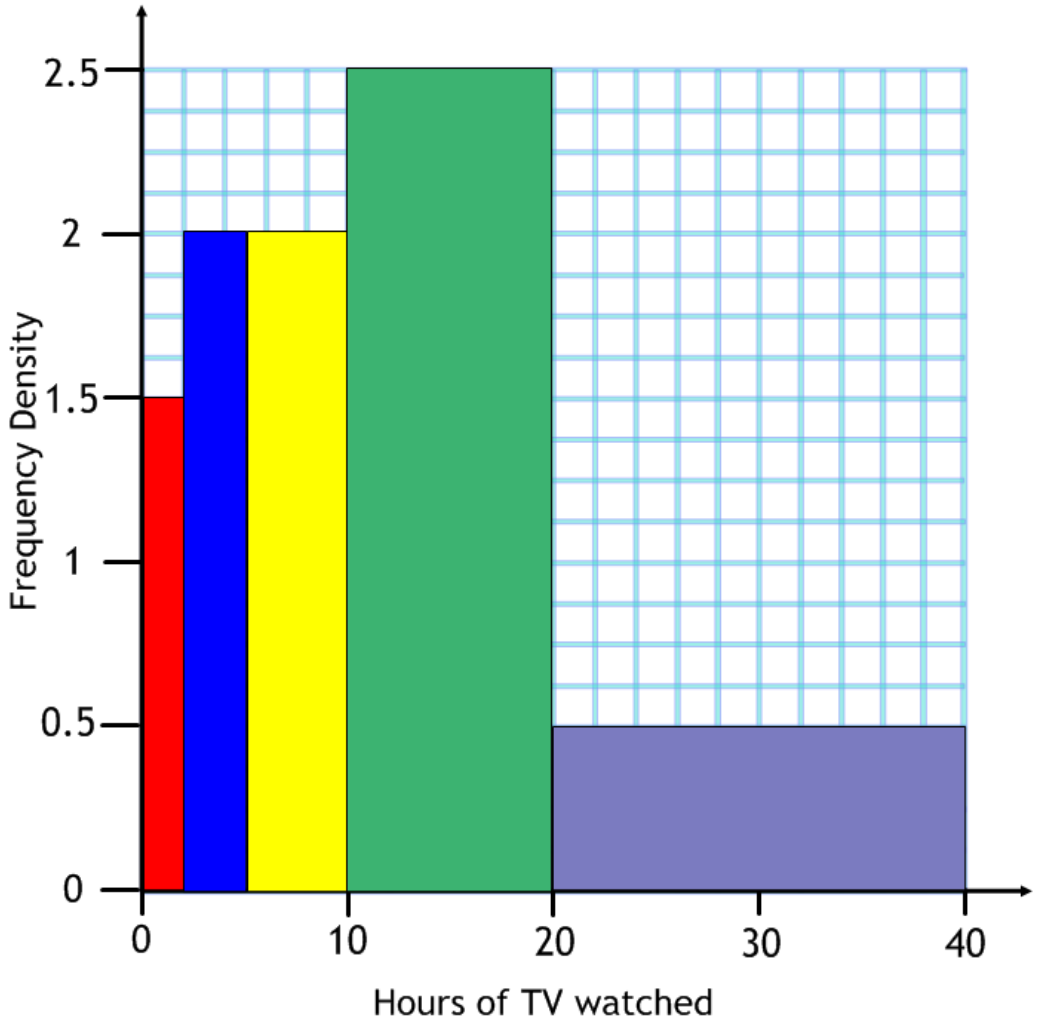
How to draw a histogram

A survey has been conducted on how many hours of TV some children watched last week. Draw a histogram for this data.

Hours (h) spent watching TV last week	**Frequency**	**Frequency Density (Frequency ÷ Group Width)**
$0 \leq h < 2$	3	3 ÷ 2 = **1.5**
$2 \leq h < 5$	6	6 ÷ 3 = **2**
$5 \leq h < 10$	10	10 ÷ 5 = **2**
$10 \leq h < 20$	25	25 ÷ 10 = **2.5**
$20 \leq h < 40$	10	10 ÷ 20 = **0.5**

Since the groups are all different widths we need to calculate the frequency density by dividing the frequency by the group width.

Drawing A Histogram



Things to notice:

- The widths of the bars are the group widths
- We plot the frequency density not the frequency
- The **area** of the bars represent the frequency

Grouped Frequency Diagram

From this data we can construct a grouped frequency diagram.

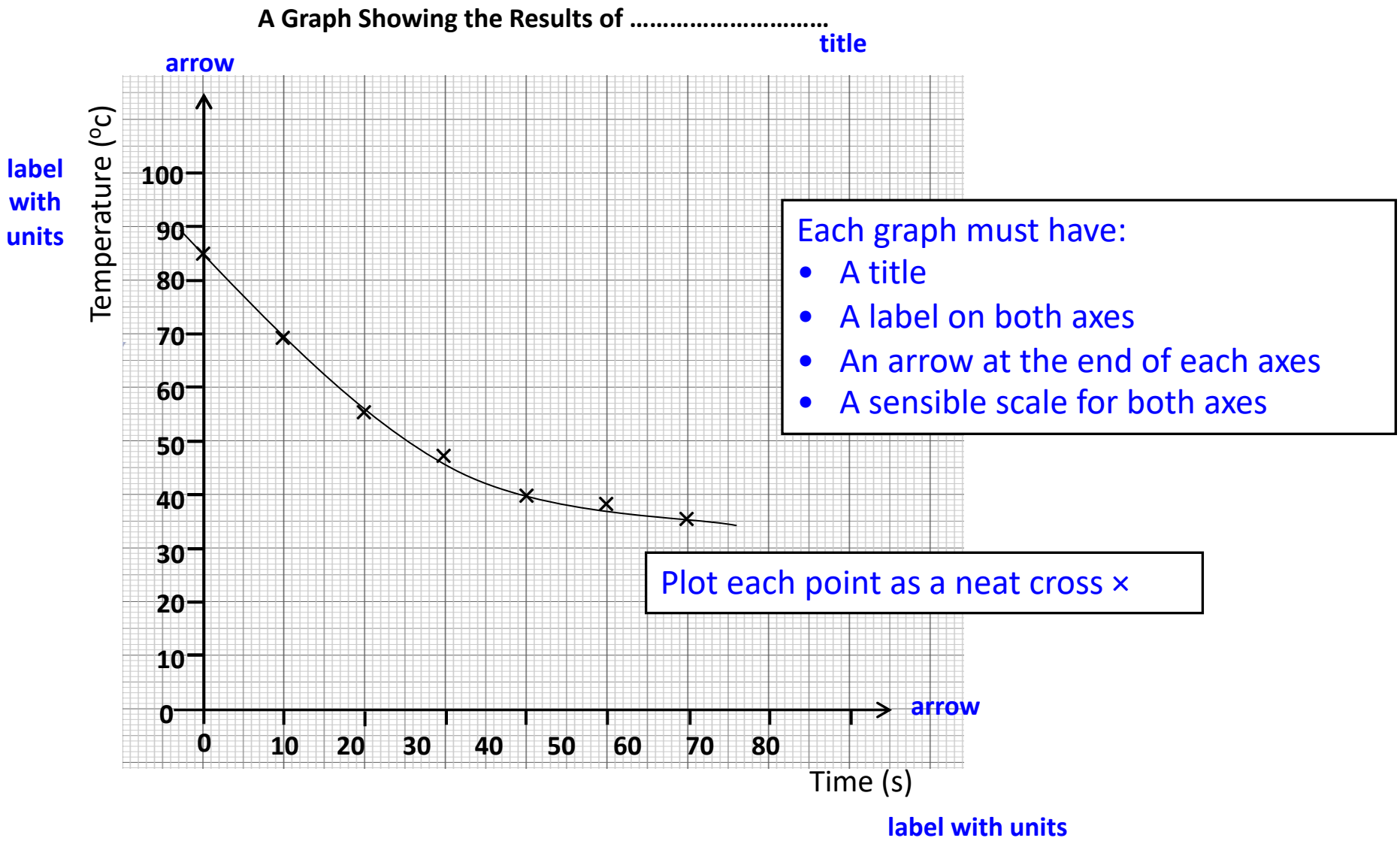
Pocket Money	Tally	Frequency(f)	Mid point (MP)	Mid point x f
£0 - £1.99	𝍸	7	£1	7
£2.00 - £3.99	𝍸	7	£3	21
£4.00 - £5.99	𝍸 I	8	£5	40
£6.00 - £7.99	𝍸	5	£7	35
£8.00 - £9.99	II	2	£9	18
£10.00 - £11.99	I	1	£11	11
TOTAL		**30**		**132**

KS4 Analysis

Median Person = (30 + 1) / 2 = 15.5 = 16th person.
Median Group = £4.00 - £5.99

Estimated Mean = 132 / 30 = £4.40

What Makes A Good Graph?

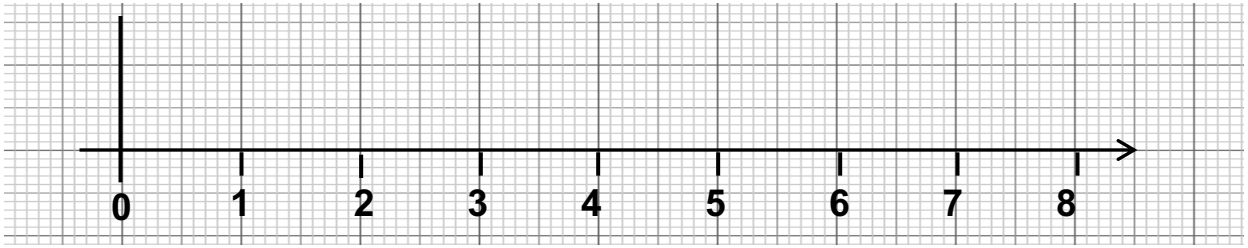


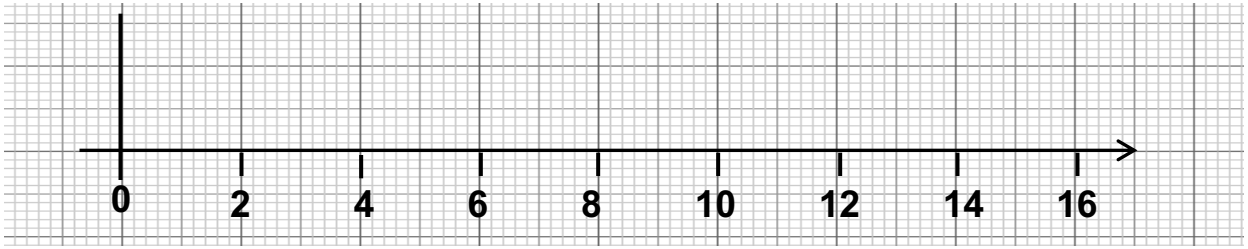
P-TASK Graph Checklist

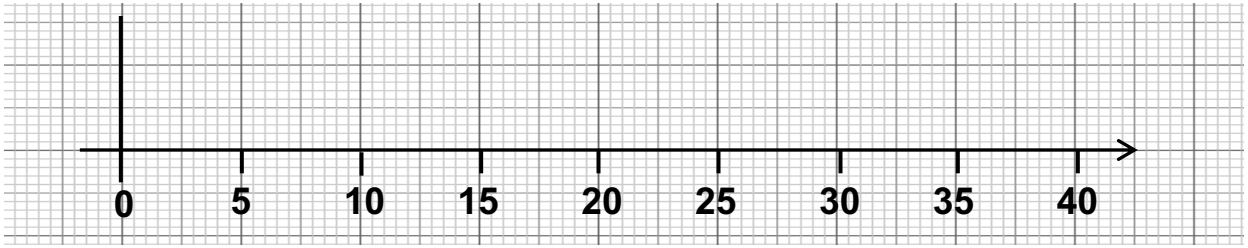
- PAPER – have I used graph paper?
- TITLE – have I used a title?
- AXES – have I labelled my axes and used units?
- SCALE – is it suitable and have I labelled the divisions?
- KEY – Do I need to include a key to explain what things mean.

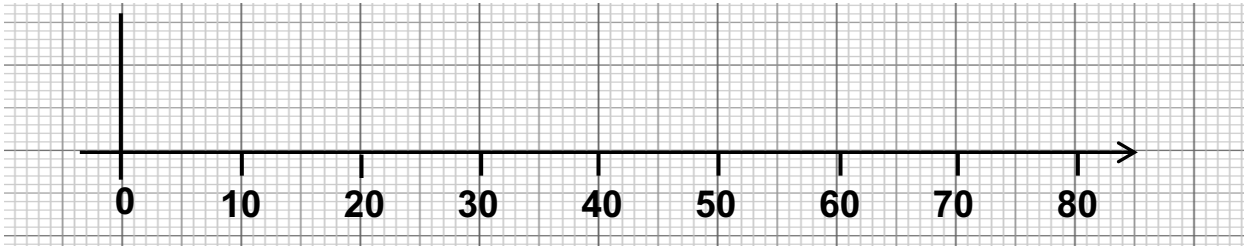
Choose a scale for the horizontal (x) axis

(don't forget the label)





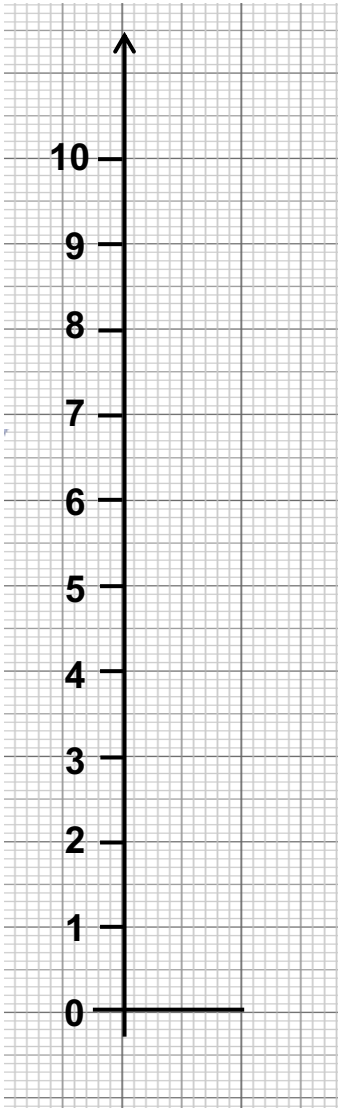


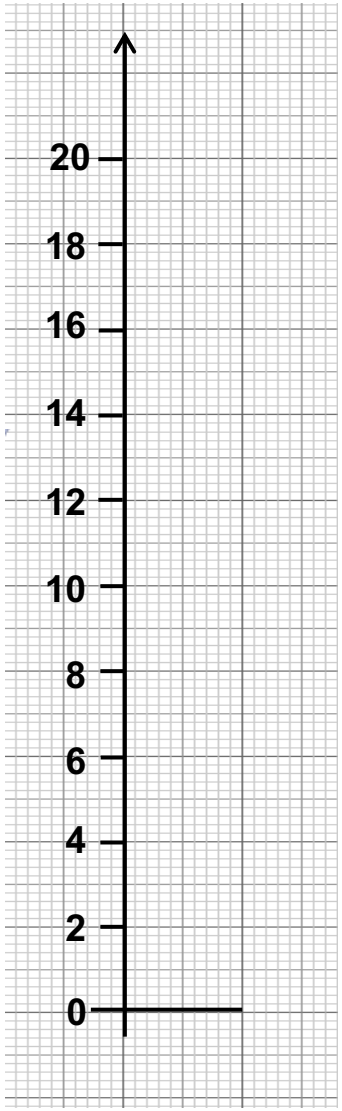


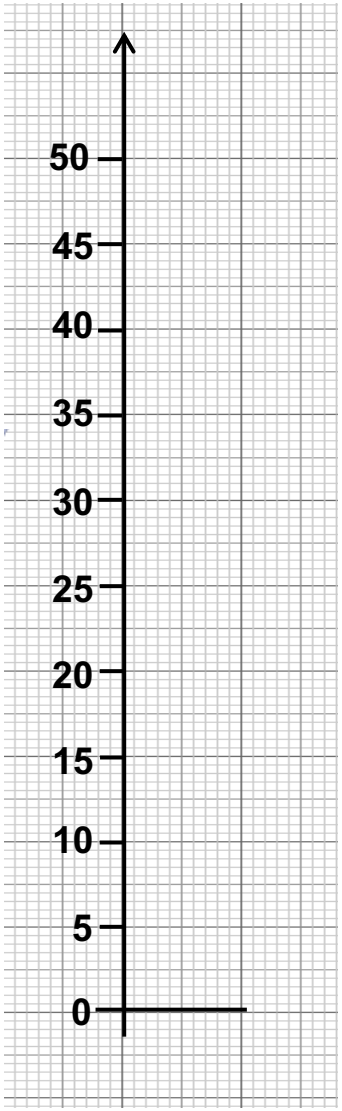
- Consider what your maximum and minimum values are.
- Use equal spacing on the axis.
- Make sure the axis fills as much of the space as possible – not too squashed, not too big that is doesn't fit. Planning is key!

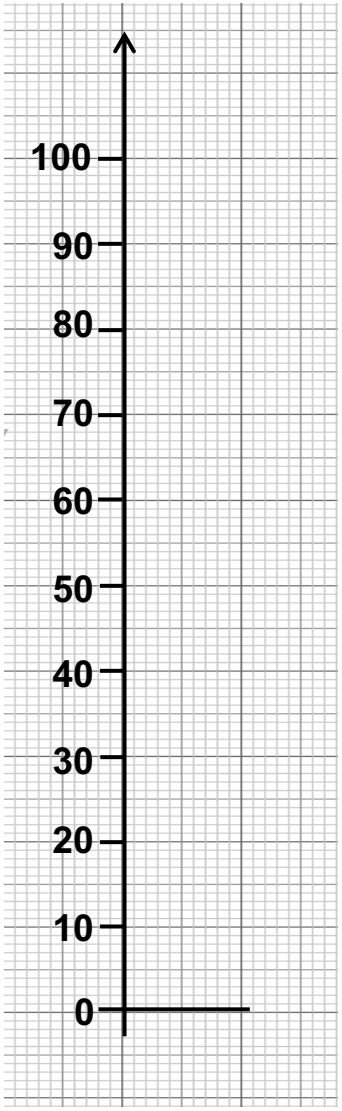
Choose a scale for the vertical (y) axis

(don’t forget the label)

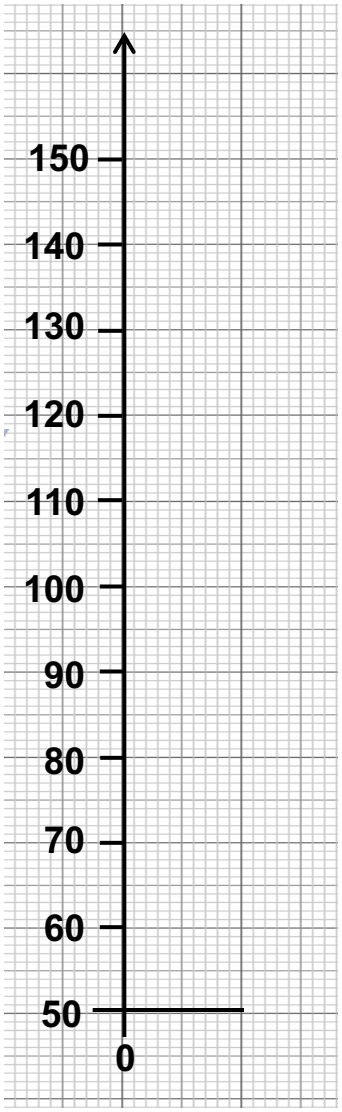


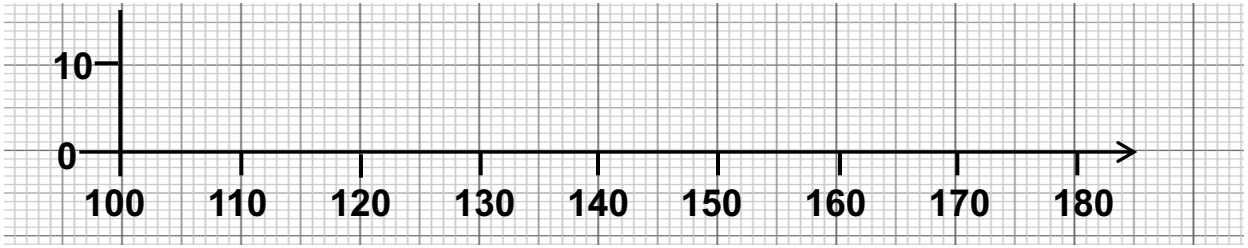


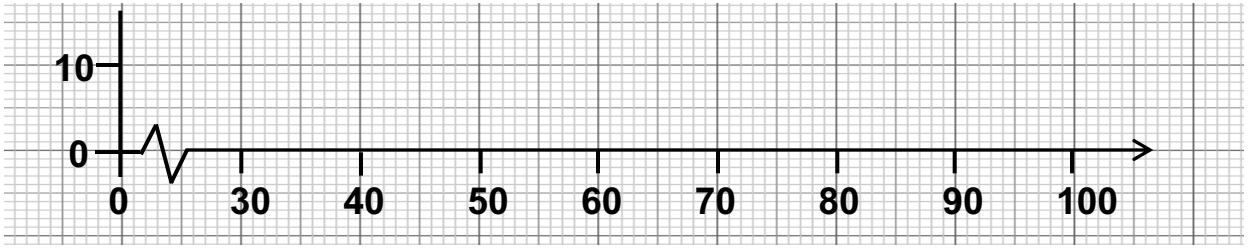




Your axes might not start from zero







Analysing Data

Averages

There are three different types of average that can be calculated:

MODE

The **mode** is the most common or most popular data value. It is sometimes called the modal value.

MEDIAN

To find the **median** of a set of data, put the values in order of size, the median is the middle value.

For **n** data values,

$$\frac{n+1}{2}$$

gives the position of the median.

MEAN

To find the **mean**, find the total of all the data values and divide the total by the number of data values.

Type of Average	Advantages	Disadvantages
Mean	Uses all the data	Distorted by extreme values
	Most accurate value	Mean is not always a data value
Median	Unaffected by extremes	Not always a data value
	Easy to calculate if data is ordered	Not easy to use for further analysis
Mode	Very easy to find	There is not always a mode
	Can be used with non-numerical data	Not easy to use for further analysis
	Mode is always a data value	