AQA GCSE Maths: Higher

Your notes

Statistical Diagrams

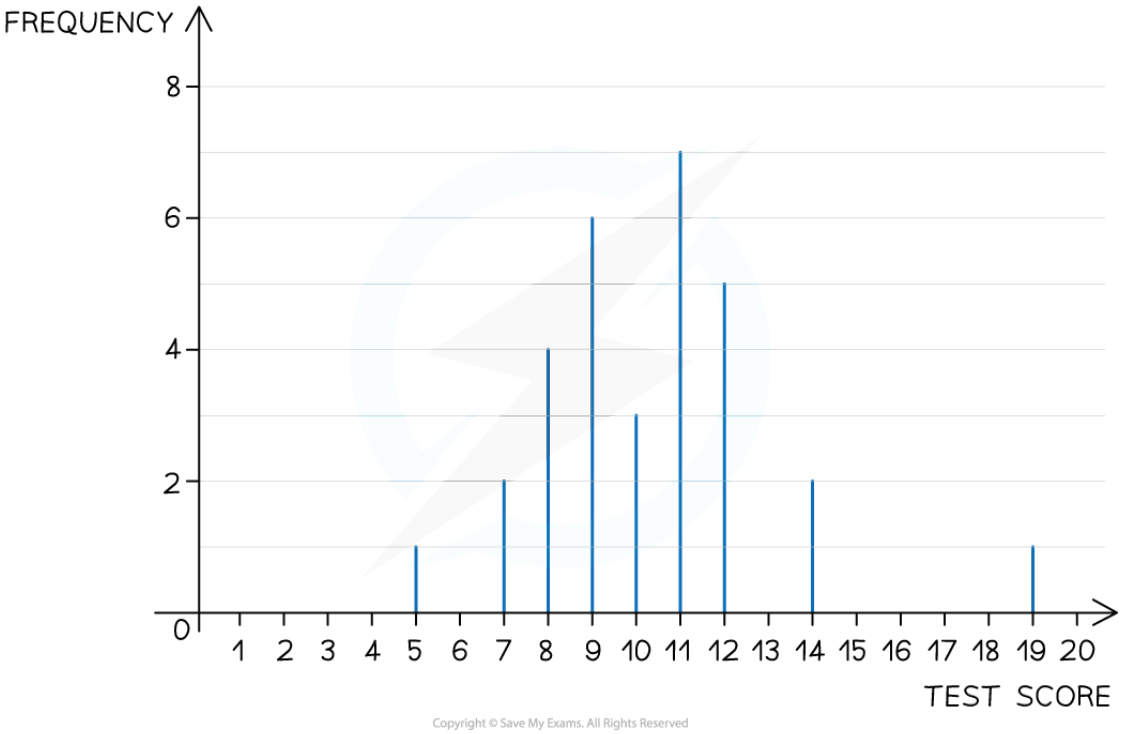
Contents

Your notes

Line Charts, Bar Charts & Pictograms

What is a line chart and what is it used for?

- Sometimes called a **vertical line chart**, this a visual way to represent **discrete** data
  - **Line charts** are used for **numerical** data (rather than categorical data)
    - They are particularly useful when there are lots of different options to show e.g. Results of a test where scores are given as percentages
- The **vertical axis** shows the **frequency**
  - The scale should start at zero and increase in equal amounts
- The **horizontal axis** shows the different **outcomes**
  - A vertical line is drawn for each outcome and its height is its frequency



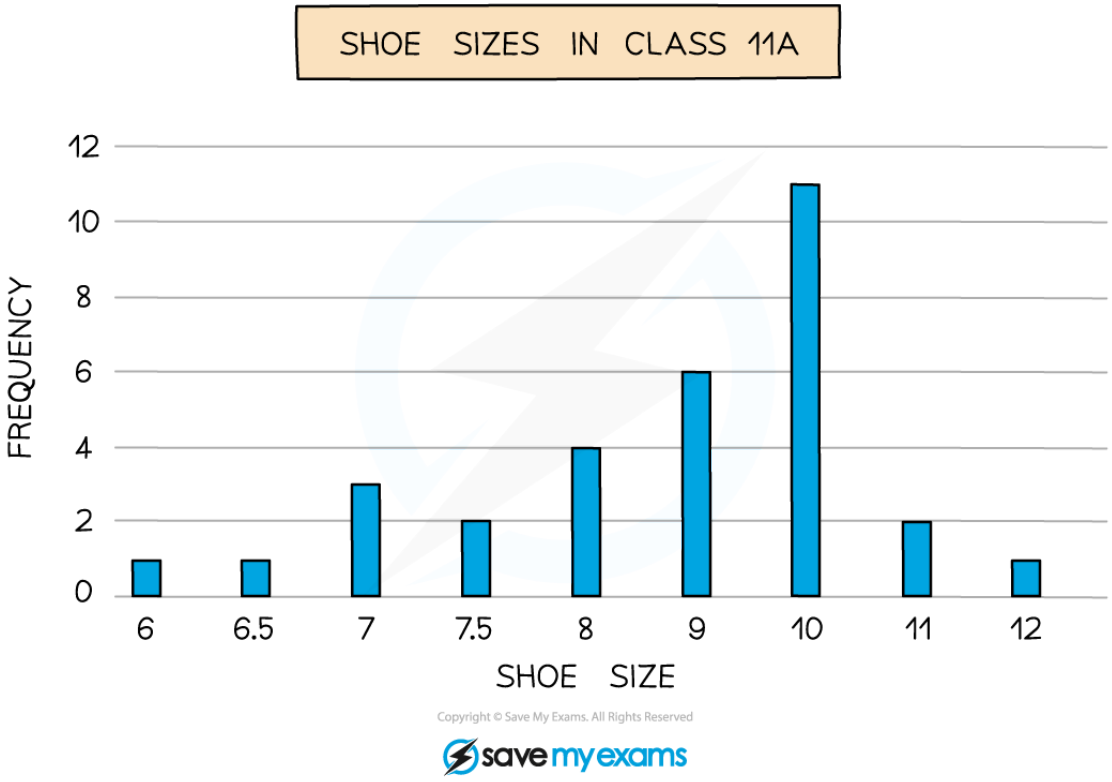
- You can easily identify the **mode** using a line chart
  - The **mode** is the most common outcome
  - This will be the **outcome** with the **highest** (tallest/longest) line
  - e.g. In the line chart above, 11 was the **modal** test score, with a frequency of 7
- You can quickly see how the data is **spread** using a line chart
  - Lines may be crowded around a particular group of options with only a few elsewhere
  - This may help identify anomalies or outliers in the data
  - e.g. In the line chart above we can see
    - the majority of the test scores, out of 20, were between 7 and 12
    - one pupil scored 19 out of 20, much higher than anyone else in the class


Your notes

What is a bar chart?

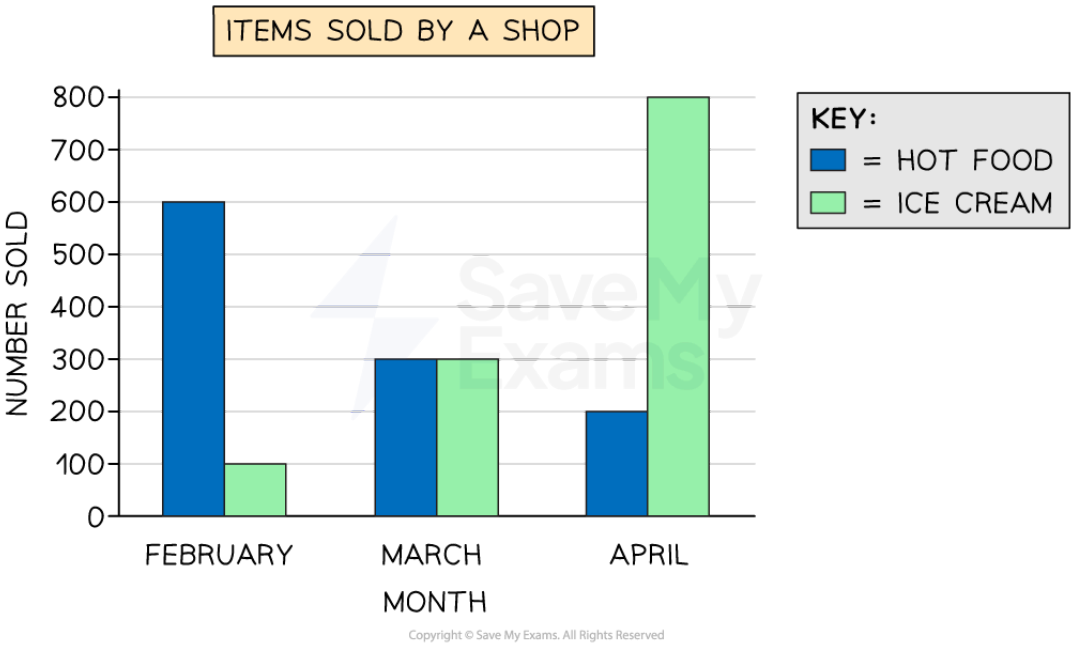
- A **bar chart** is a visual way to represent **discrete** data
  - Discrete data is data that can be **counted**
    - This can be **numerical** like shoe sizes in a class
    - Or **non-numerical** (categorical) like colours of cars down a road
- The **horizontal axis** shows the different **outcomes**
- The **vertical axis** shows the **frequency**
- The **heights** of the **bars** show the frequency
  - Bars should be **separated** by gaps
  - Bars should have **equal widths**

Your notes



- The **mode** is the **outcome** with the **highest** bar
- You can also get **dual bar charts** to compare two data sets
  - Bars are in **pairs** (side-by-side) for each outcome

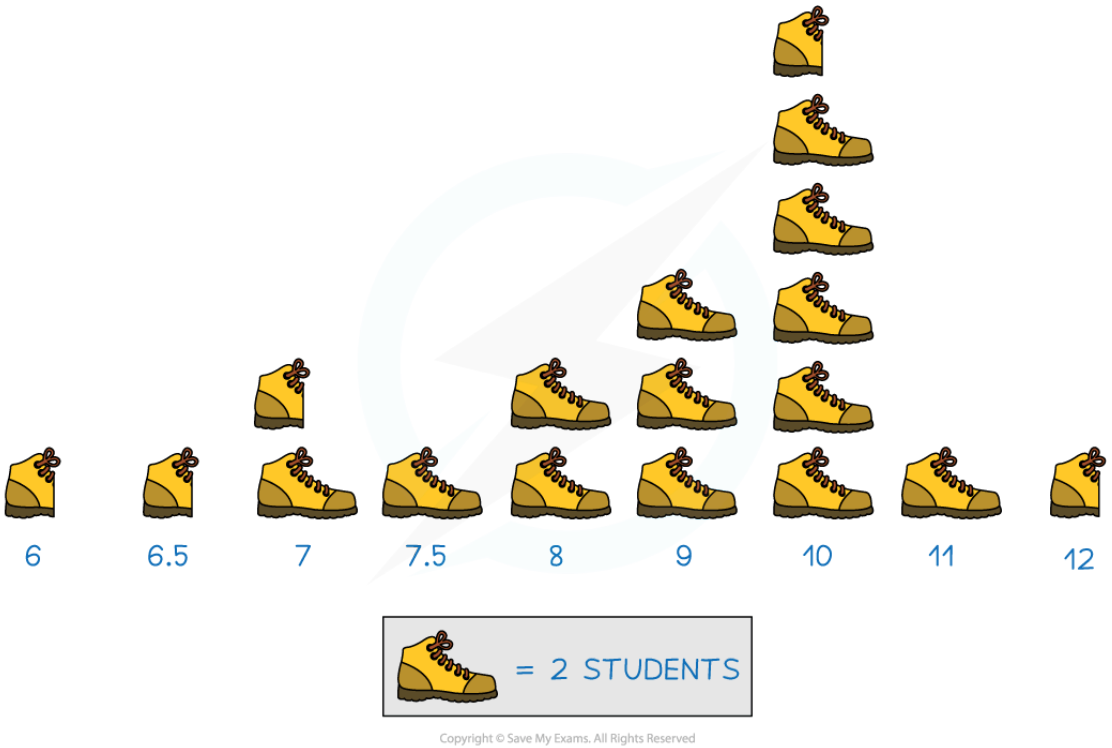
Your notes



What is a pictogram?

- A **pictogram** is an **alternative** to a bar chart
  - It is used in the same situations
- There are **no axes**
  - **Frequency** is represented by **symbols**
  - A **key** shows the value of 1 symbol
    - For example, 1 symbol represents a frequency of 2
  - Half and quarter symbols are often used

Your notes



- The pictogram above shows the shoe sizes of students in a class
  - As 1 picture of a shoe represents 2 students
    - Half a shoe represents 1 student
  - The number of students with a shoe size of 7, is 3

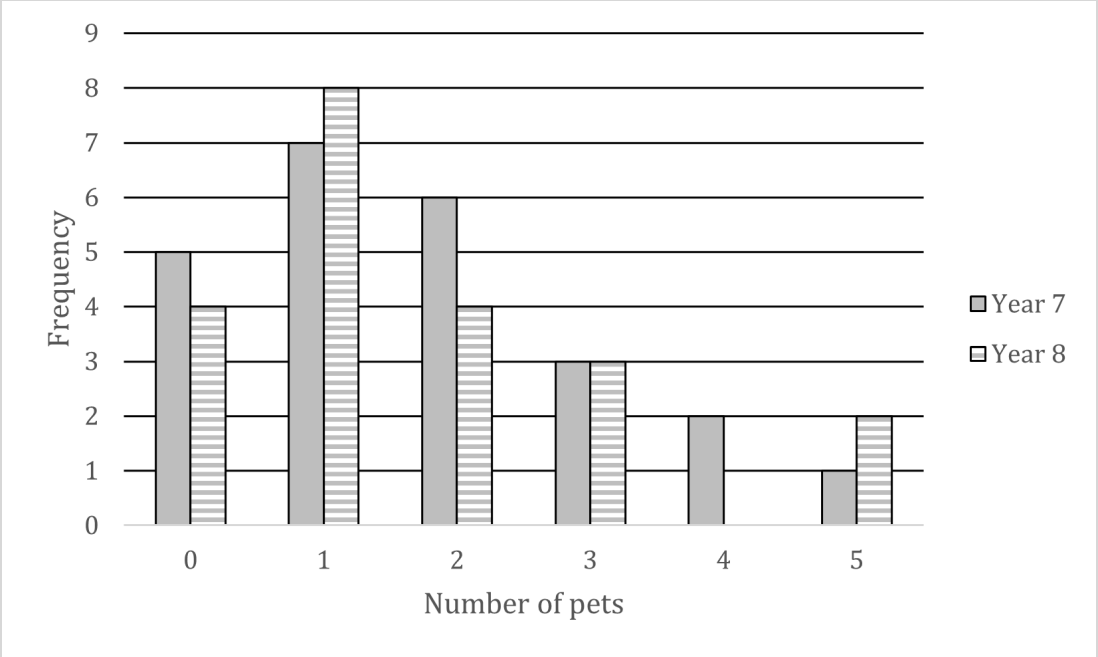
Examiner Tips and Tricks

- If asked to draw a bar chart, find the largest frequency and choose a scale which makes that fit in the space provided
- If asked to draw a pictogram, pick a symbol that is easy to duplicate and draw half (or quarter) of

Worked Example

Your notes

Mr Barr teaches students in Year 7 and Year 8.
He records the number of pets that students in each year have.
His results are shown below.



(a) Write down the modal number of pets for his Year 7 students.

The modal number (mode) is the number of pets that occurs the most
Visually, this will be the highest bar for Year 7s

The mode for Year 7 is 1 pet

(b) How many Year 8 students does he teach?

Add up all the heights (frequencies) of the Year 8 bars

$$4 + 8 + 4 + 3 + 0 + 2$$

He teaches 21 Year 8 students

Pie Charts

What is a pie chart?

- A **pie chart** is a circle divided into sectors which is used to **present data**
- The **sectors** represent different **categories**
  - They show the relative **proportions** of the categories
  - They do **not** show the actual **frequencies** of each category

How do I draw a pie chart?

- This is easiest explained through an example
- The following data shows the favourite colours of a class of students

Colour	Red	Purple	Blue	Green	Yellow	Orange
Students	11	4	9	3	2	1

- Write the frequencies as **fractions** of the **total number** of students, 30

Colour	Red	Purple	Blue	Green	Yellow	Orange
Students	11	4	9	3	2	1
Fractions	$\frac{11}{30}$	$\frac{4}{30}$	$\frac{9}{30}$	$\frac{3}{30}$	$\frac{2}{30}$	$\frac{1}{30}$

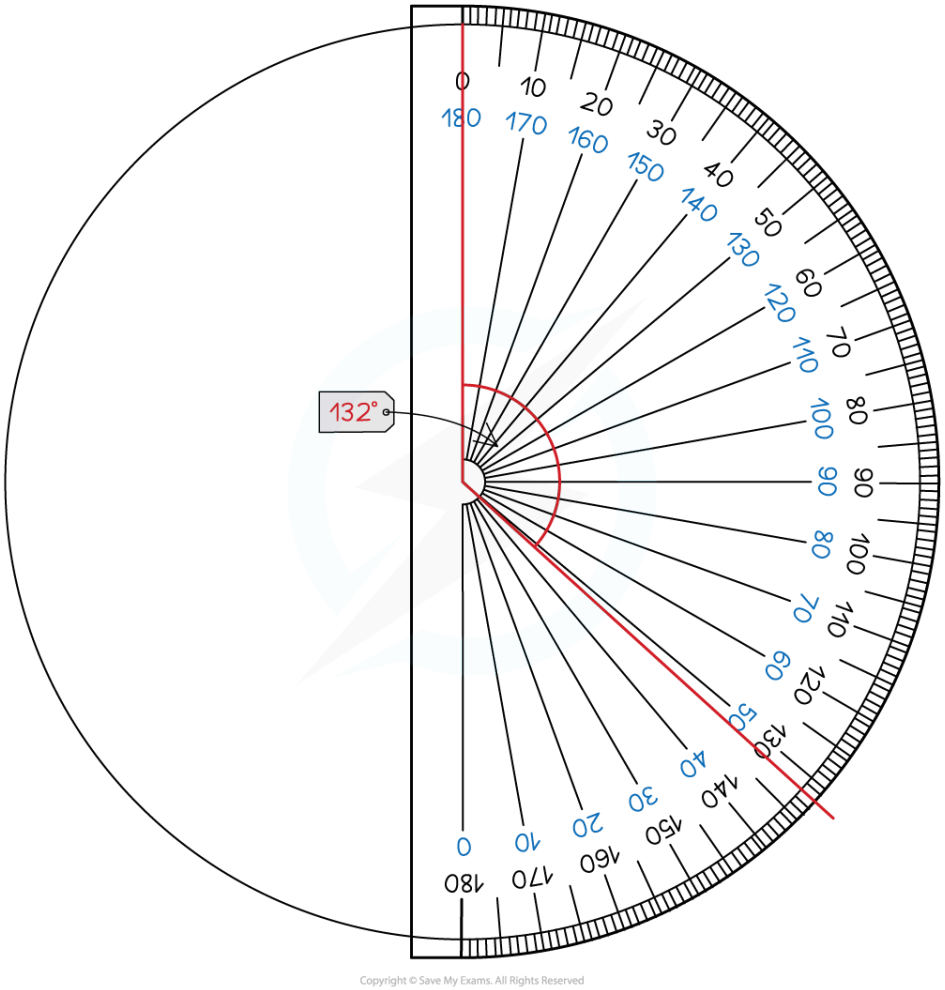
- Find the **angles** of the sectors by **multiplying each fraction** by **360°**
  - Then check all angles add up to 360°

Colour	Red	Purple	Blue	Green	Yellow	Orange
Students	11	4	9	3	2	1

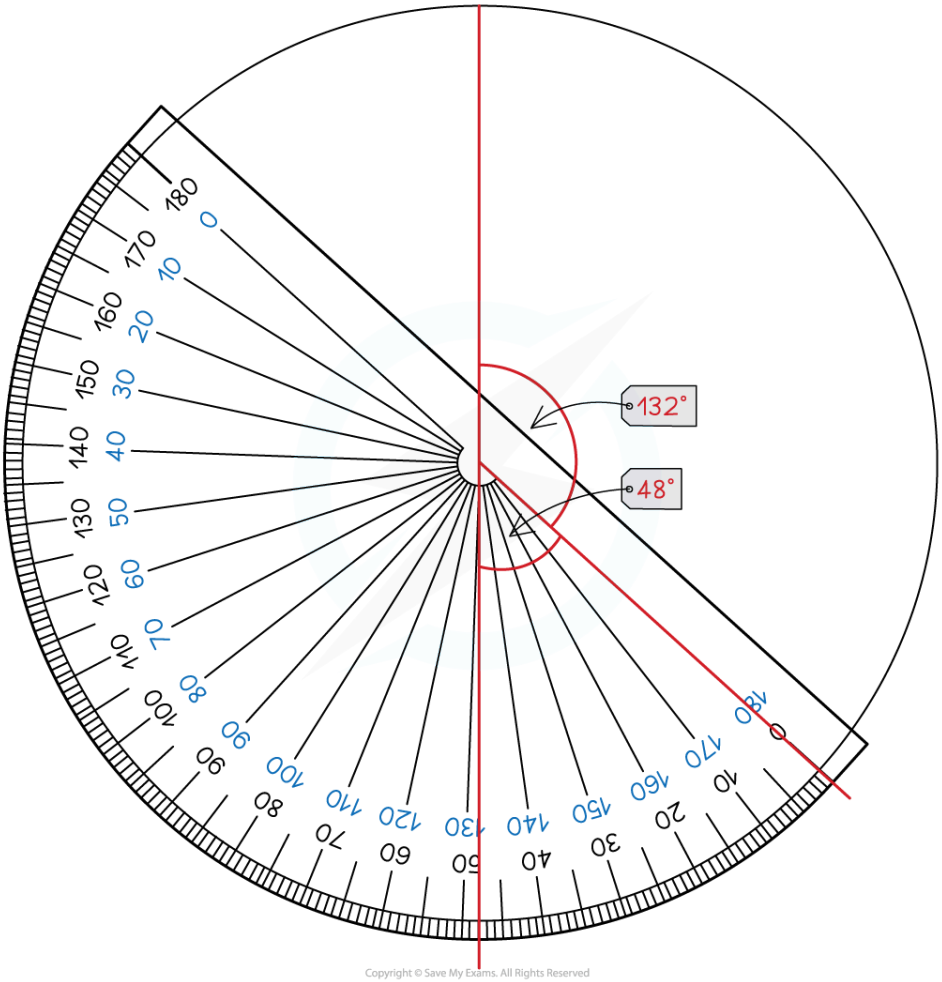
Fractions	$\frac{11}{30}$	$\frac{4}{30}$	$\frac{9}{30}$	$\frac{3}{30}$	$\frac{2}{30}$	$\frac{1}{30}$
Angles	132°	48°	108°	36°	24°	12°

Your notes

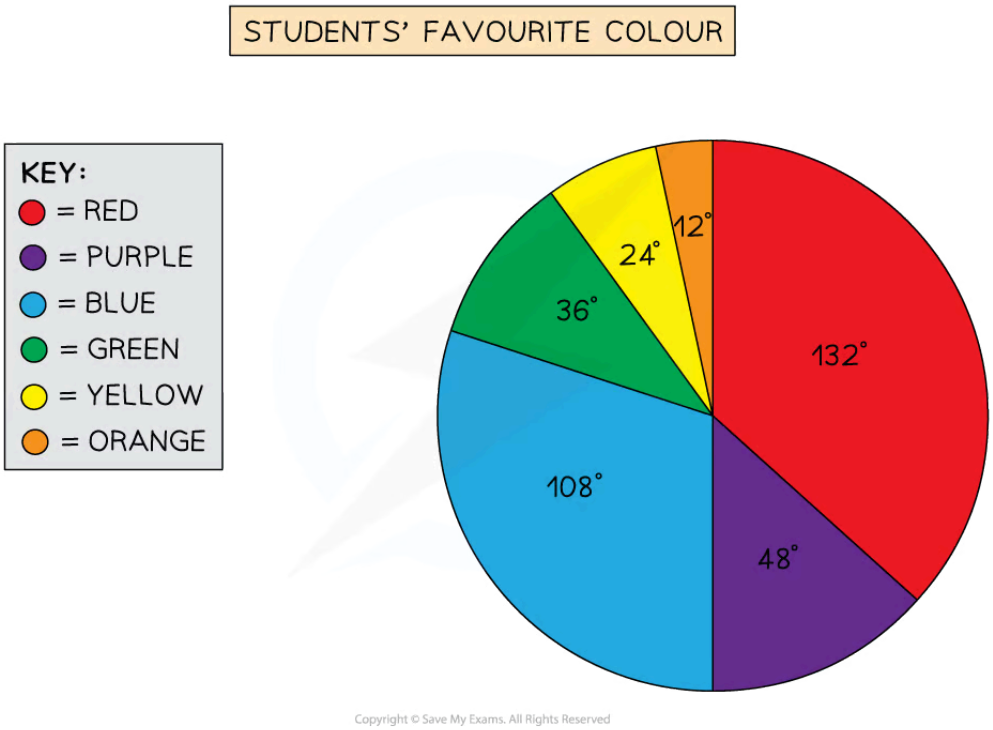
- Draw a vertical line from the circle's centre to the top, then use a **protractor** to mark off the first angle
  - Draw a line from the centre to this first mark
    - Then, from this line, mark off the next angle (and so on)



Your notes



Your notes



How do I solve problems with pie charts?

- Use the following facts
  - **angles** are **proportional** to the **frequencies** of each category
  - **360°** represents the **total frequency**
- For harder problems, it helps to work out
  - **what frequency** is represented **by 1°**
  - **what angle** is represented by **1 unit** of **frequency**
- For example, if a sector of 30° represents 15 people, then
  - 1° = 0.5 people (dividing by 30)
  - 2° = 1 person (multiplying by 2)
- These relationships can then be scaled **up** or **down** accordingly
  - If 1° = 0.5 people

- then 360° = 180 people (multiplying by 360)

Your notes

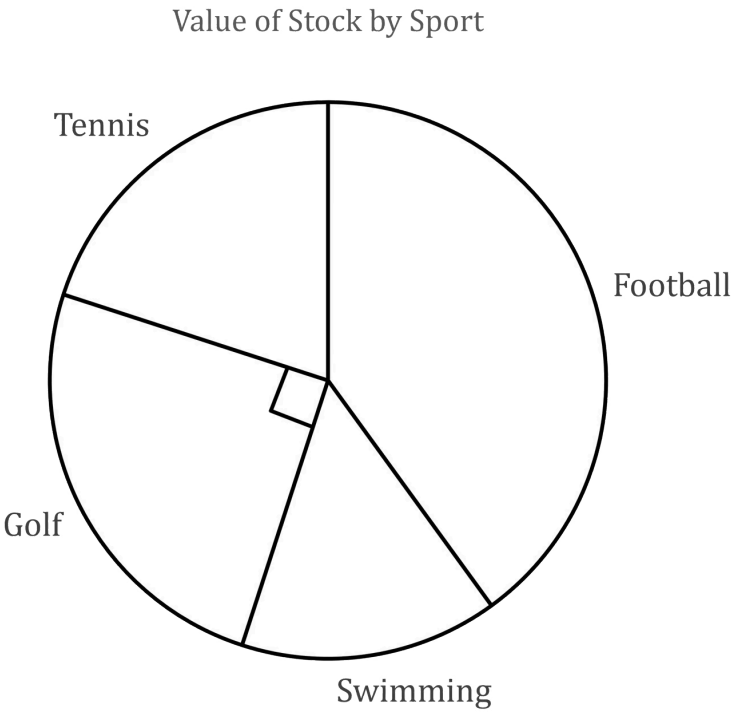
Examiner Tips and Tricks

- If the pie chart says 'not to scale', then examiners want you to use ratio and proportion methods to answer the questions
  - Don't measure angles using a protractor!

Worked Example

The following pie chart represents the values of items stocked in a sports shop.

[not to scale]



(a) Given that the shop stocks $12 000 of golf items, find the total value of the shop's stock.

Find a relationship between an angle and a value

$$90° = \$12\,000$$

The total value is represented by 360°
Multiply by 4 to get from 90° to 360°

$$360° = 4 \times 12\,000$$

Total value is $48 000

You can also do this question by finding 1° first

(b) If the angle on the pie chart for tennis is 72°, find the value of tennis items that are stocked by the shop.

It is quickest to find the fraction $\frac{72}{360}$ of the total value, found in part (a)

$$\frac{72}{360} \times 48\,000 = 9\,600$$

The value of tennis items is $9 600

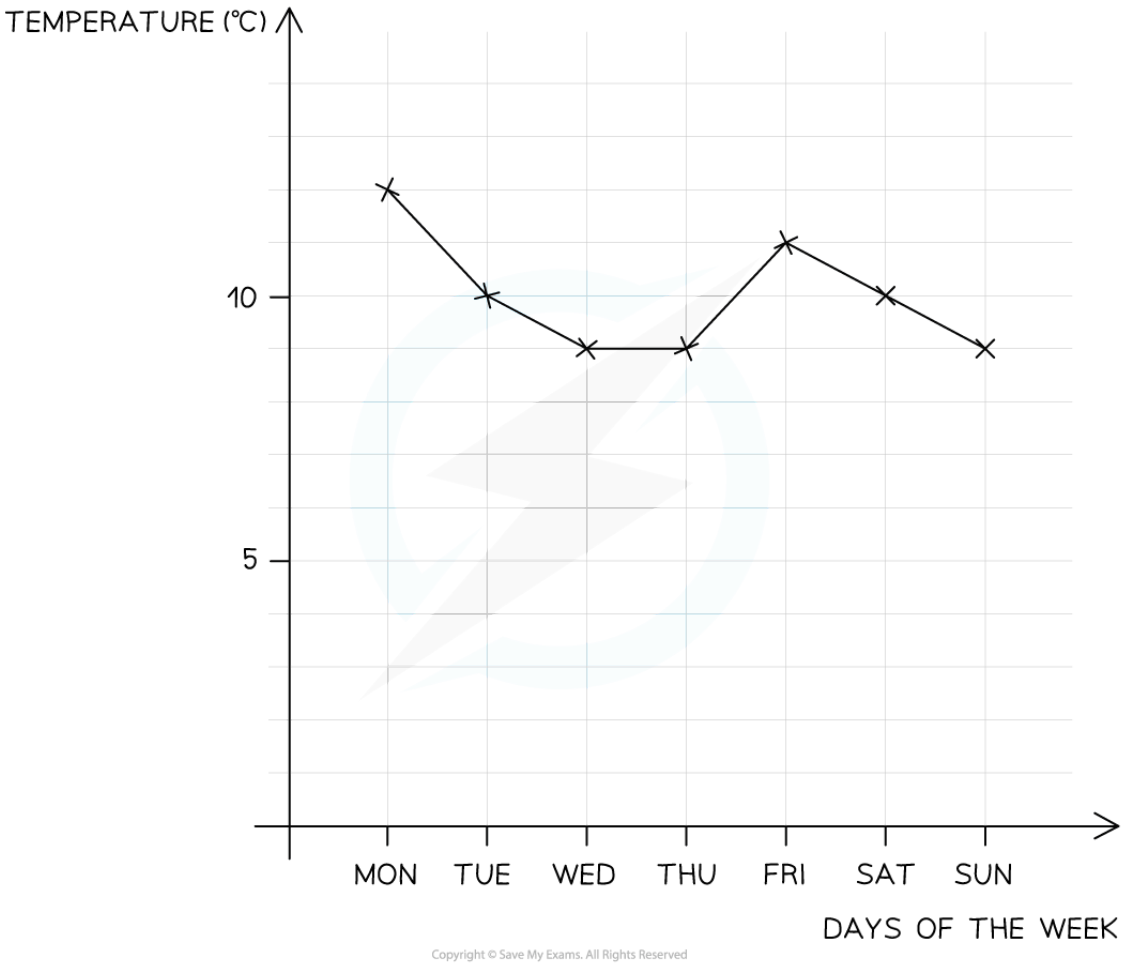
Your notes

Your notes

Time Series Graphs

What is a time series graph?

- A **time series graph** is sometimes called a **line graph**
- A time series graph shows how a quantity (**continuous** data) changes over time
  - e.g. How the outside temperature changes during a week (shown below)
- **Measurements** of the quantity are taken at particular times
  - **Measurements** should be taken at **regular** time intervals
  - These are then **plotted** as points on a time series graph and **joined** together with **straight lines**
  - The straight lines help us to **identify patterns** and **features** in the data
- Time series graphs can show changes over **short** or **long periods** of **time**
  - e.g. Changes to the temperature of two chemicals for the first few **minutes** after they've been mixed
  - Or changes to the temperature of the earth over several **years**



Your notes

How do I draw a time series graph?

- The **horizontal** axis (*x*-axis) will be the **time** axis
- The **vertical** axis (*y*-axis) will be the **quantity** being measured/recorded
- **Plot** the data as a series of **points**
- **Join** one point to the next, in order, with **straight lines**
  - Use a ruler
- Sometimes a time series graph may have **more than one data set**
  - e.g. one line for car emissions and one line for motorbike emissions
  - Plot one data set and join the points up before moving on to the second data set

- This will ensure you do not muddle the points up
- You could use crosses (×) for one set of points, and dots (•) for the other
- You could use different colours or dotted/dashed lines when joining the points up
- Always include a **key** in such cases to make it clear which line is which data set


Your notes

How do I use and interpret a time series graph?

- This involves looking at **patterns** in the data as well as **specific** points
- If a question asks you to interpret or describe a time series graph look for:
  - A **general trend**
    - e.g. The rate of inflation may fluctuate (go up and down) but is generally going down over a decade
  - **Seasonal** differences
    - e.g. Sales are higher in the summer months than winter months
- Other things to look for
  - a **horizontal** line between points - **no change** (constant)
  - the **steepest** line (**gradient**) would indicate the **greatest change**
    - this could be an **increase** ('uphill' left to right, like /)
    - or a **decrease** ('downhill' left to right, like \)
  - **Unusual** 'one-off' readings - **'spikes'** or **'dips'**
    - e.g. A spike in mobile phone network activity at midnight on New Year's eve
- For line graphs with two (or more) data sets, be clear about **which line** you are describing
  - Use the **key**
  - Double check which data set is 'higher' or 'lower' (or they may be equal) at a particular time

Examiner Tips and Tricks

- If you are asked to **describe** or **interpret** a line graph then it may help to:
  - **Draw vertical or horizontal lines** on the graph from the respective **time** or **measurement** to ensure an accurate reading

- Use a ruler to do this!
- **Highlight** any particular **points** that you mention in your description

Your notes

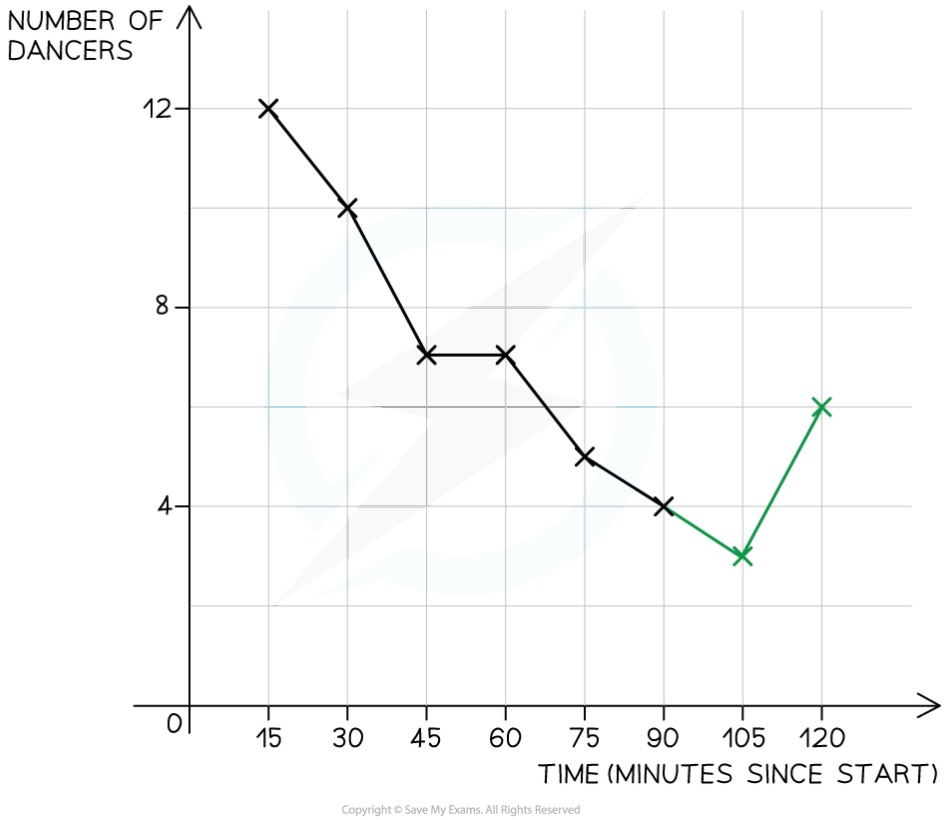
Worked Example

In a two-hour charity 'dance-a-thon' dancers can join and leave the dance floor as they choose.
The number of dancers on the dance floor is recorded every 15 minutes.
The times-series graph below shows the data for the first two hours of the 'dance-a-thon'.

At 1 h 45 m and 2 h the number of dancers were 3 and 6 respectively.

(a) Add these two recordings to the time-series graph.

Plot the points 105 minutes (1 h 45 m) on the time axis against 3 on the number of dancers axis and 120 minutes against 8 dancers.



(b) Use the time-series graph to find

(i) the time at which there were the fewest number of dancers,

(ii) the two times at which the number of dancers remained the same.

(i) Look for the lowest point on the graph, including those added in part (a)

The lowest point is (105, 3)

The final answer is the **time**

105 minutes (1 h 45 m)

(ii) The keyword here is **remain** - a horizontal line would indicate the number of dancers remaining the same

There is a horizontal line between 45 and 60 minutes

The number of dancers remained the same at 45 m and 60 m (1 h)

(c)

Comment on the general pattern of the number of dancers shown by the time-series graph.

Apart from the last point, and where it remained the same, the number of dancers decreased over the two hours.

In general, the number of dancers decreased during the first two hours of the 'dance-a-thon'


Your notes

Your notes

Reading & Interpreting Statistical Diagrams

How do I interpret statistical diagrams?

- Read and **understand** the initial sentences describing the **situation** (context)
  - **Underline** important words if necessary
- Look for any **keys** that may help you to understand the diagram
  - For example
    - 1 unit represents 20 people
    - Year 10 is the **solid** line, Year 11 is the **dotted** line
    - Class A is **shaded**, class B is **striped**
- **Read** the **titles** of diagrams and all **axes labels**
  - A graph for '**new** students at a school' is different to '**all** students at the school'
  - A **frequency** axis that **starts at 50** is different to one starting at 0
- Understand which **units** are being used
  - Individual lengths may be in **centimetres** but total length may be in **metres**
  - Populations may be measured in **thousands**
- Look out for any **extreme values** (outliers / anomalies)
  - One month's temperature might be unusually high
    - Was it a heat wave or a recording error?

How do I draw conclusions from diagrams?

- Look for **overall trends** in the diagram
  - Prices **increase** year on year
  - The temperature **peaks** in June
- Use **numbers** from the graphs
- Refer to any **changes**

- The **steepness** (gradient) of graph may change

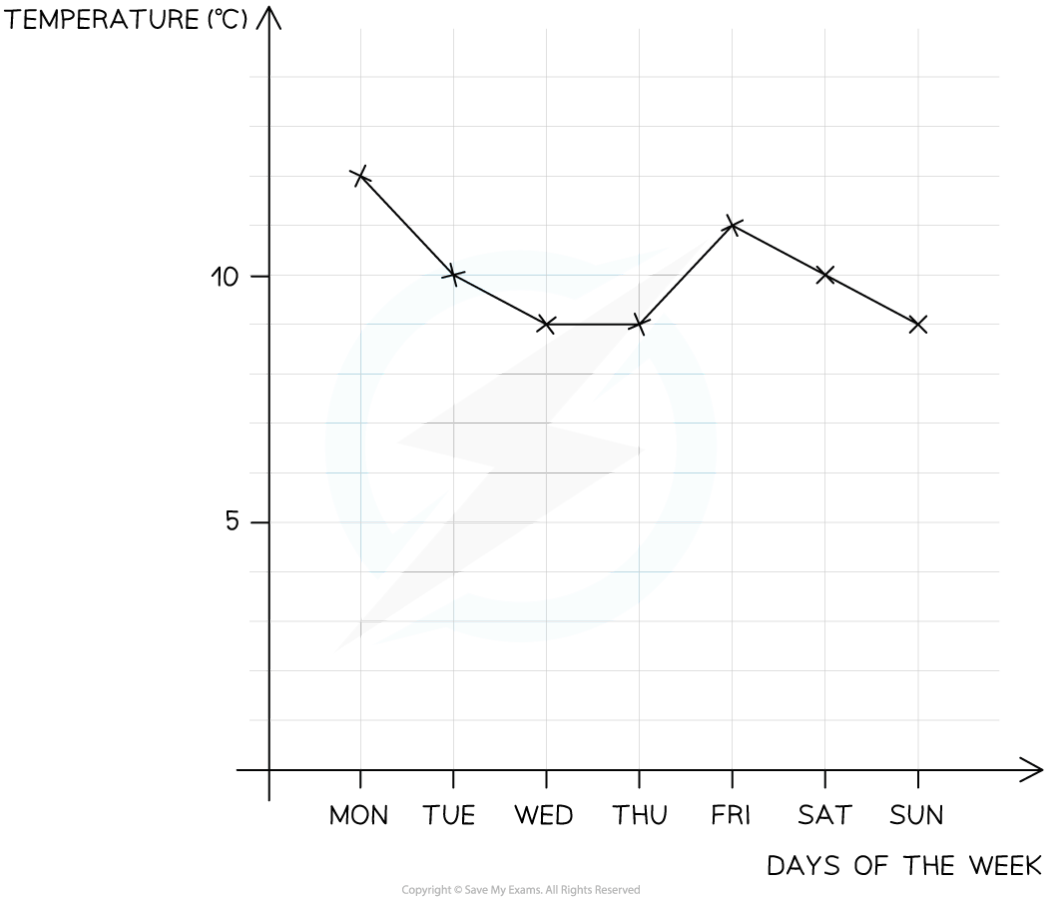
- Write in **full sentences** that copy the **exact wording** from the question
  - 'The number of goats in farm A has decreased by 12 over the 8 month period'
  - **Not** 'There are fewer of them now'
- You may need to calculate the **mode**, **median**, **mean** or **range** to support any explanations
- Understand why drawing conclusions may **not** be **suitable**
  - The data set may be **too small** to be representative
  - The data set may be **biased**
  - Consider the **scope** of the data
    - e.g. Data for January to March cannot be used to predict August

Worked Example

The diagram below shows the temperature of a garden in the UK, recorded at 7am on each day of a particular week in March.

Your notes

Your notes



(a) Describe the change in temperature over the first four days.

The trend shows a decrease in the first three days, then a constant temperature
Find numbers from the graph to use in your answer
Refer to the steepness of the change

The temperature decreases from 12°C on Sunday to 9°C on Wednesday
The decrease is steeper over the first two days
There is then a constant temperature of 9°C on both Wednesday and Thursday

(b) A gardener claims that, based on the graph, Monday must have experienced the highest temperature that week.

Give a reason as to why this might not be true.

Reread the information at the start
These temperatures were recorded at 7am in the morning (we don't know how hot the rest of the day was)

The temperatures on the graph are at 7am each day
The maximum temperature may have been after 7am, on a day that was not Monday

(c) A journalist wants to use the data shown to claim that the average temperature that week was below 10°C.
The mean of the temperatures shown is 10°C.

Which type of average would you suggest they use? Explain your answer.

The mean of 10°C does not support the claim that the average temperature is below 10°C
Try calculating the mode instead

12, 10, 9, 9, 11, 10, 9

The most frequent number is 9

The modal temperature is 9°C
9°C < 10°C so using the mode would help the journalist's claim

You could try the median, but it is also 10°C


Your notes

Your notes

Comparing Statistical Diagrams

How do I compare statistical diagrams?

- You may be given **two graphs** for two different **data sets** with the same **context**
- Compare **trends** in the graphs
  - **Increases**, **decreases**, **maximum points**
  - **Steepness** of the change
- Comment on **differences** and **similarities**
- Explain **clearly** which **part** of the graph you are talking about
- Use **numbers** from each data set in your comparisons
- Comment on any **extreme values** (outliers)
- Calculate **means, medians** or **modes** to compare **averages**
- Calculate **ranges** to compare the **spread** of values

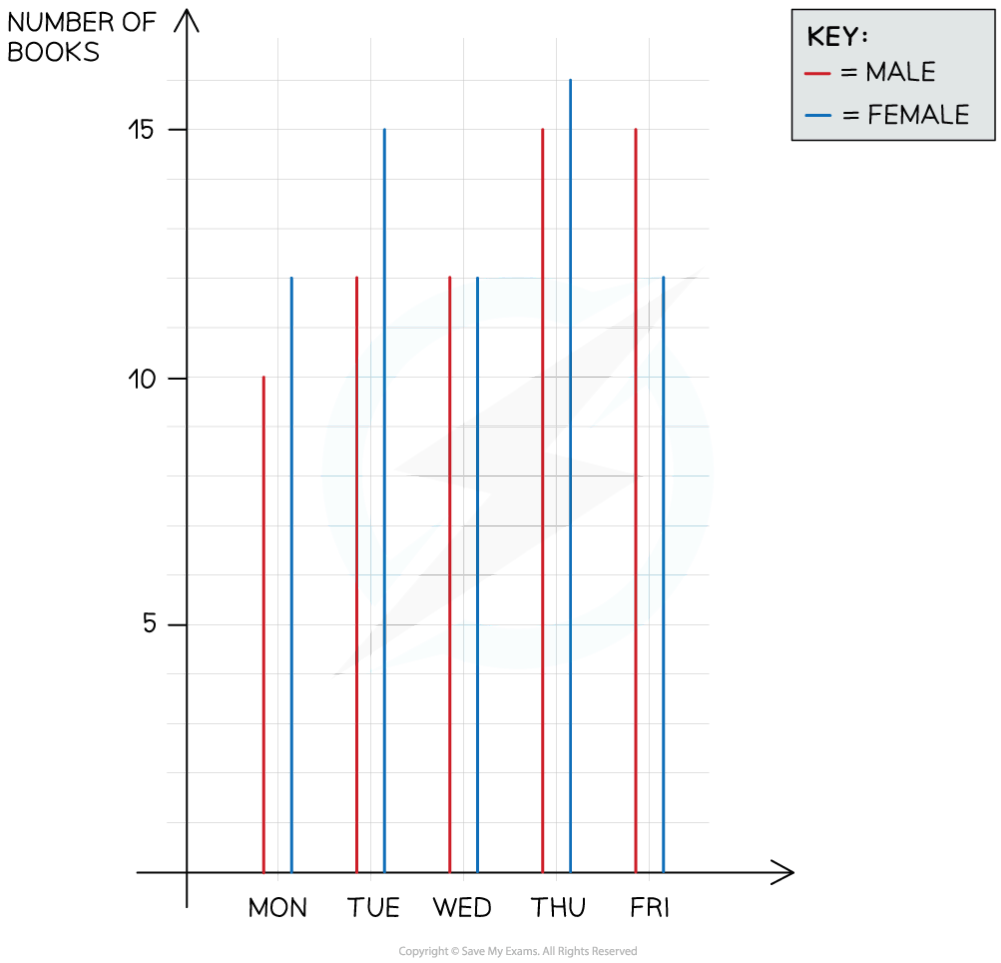
Examiner Tips and Tricks

- Always relate numbers and calculations back to the context of the question
  - If you don't, you won't score all the marks!

Worked Example

The number of books bought during the opening week of a new bookshop is shown below.
The shopkeeper wants to investigate shopping patterns between male and female customers.

Your notes



(a) Briefly summarise the overall trend in the number of books bought by male customers.

Look at the heights of the red lines
Comment on how they increase ('briefly' means you do not have to talk about every point)
Use numbers from the graph and exact phrases from the question

The number of books bought by male customers increases from 10 on Monday up to 15 on both Thursday and Friday

(b) The shopkeeper says that the changes in the blue lines suggest that the amount of books bought by females varies more than males.

By comparing a suitable statistical measure, show that this is not necessarily true.

Variations refers to the spread of data, so compare their ranges
For male and female, subtract the smallest value from the biggest value

Male range is 15 - 10 = 5

Female range is 16 - 12 = 4

Compare the numbers
Explain how this relates to the spread of the data
State how your findings do not agree with the shopkeeper

The range of male customers is greater than that of female customers, 5 > 4
This suggests male customers have a greater spread (variability) in books bought
This is the opposite of what the shopkeeper said

(c) Give one reason as to why the shopkeeper should not use the data shown to draw conclusions about shopping patterns.

Reread the sentences at the beginning
This data is for the opening week of the bookshop only
State that this is unrepresentative of a normal week
Give a specific real life example

The data shown is for the opening week of the bookshop
It is unlikely to be representative of a normal week
The numbers may increase as the bookshop becomes more popular

Or decrease if the customers lose interest
(Any reasonable example would work)

Your notes