3b. Representing and interpreting data (S2, S3, S4, S5)	**Teaching time** 7–9 hours

OBJECTIVES

By the end of the sub-unit, students should be able to:

- Know which charts to use for different types of data sets;
- Produce and interpret composite bar charts;
- Produce and interpret comparative and dual bar charts;
- Produce and interpret pie charts:
  - find the mode and the frequency represented by each sector;
  - compare data from pie charts that represent different-sized samples;
- Produce and interpret frequency polygons for grouped data:
  - from frequency polygons, read off frequency values, compare distributions, calculate total population, mean, estimate greatest and least possible values (and range);
- Produce frequency diagrams for grouped discrete data:
  - read off frequency values, calculate total population, find greatest and least values;
- Produce histograms with equal class intervals:
  - estimate the median from a histogram with equal class width or any other information, such as the number of people in a given interval;
- Produce line graphs:
  - read off frequency values, calculate total population, find greatest and least values;
- Construct and interpret time–series graphs, comment on trends;
- Compare the mean and range of two distributions, or median or mode as appropriate;
- Recognise simple patterns, characteristics relationships in bar charts, line graphs and frequency polygons.

POSSIBLE SUCCESS CRITERIA/EXAM QUESTIONS

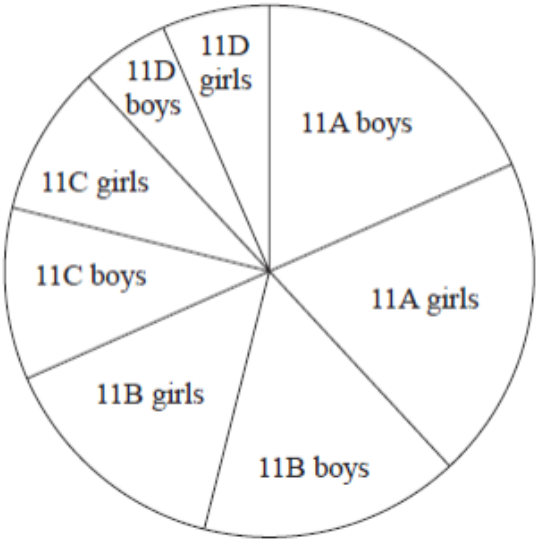
Use a time–series data graph to make a prediction about a future value.

Explain why same-size sectors on pie charts with different data sets do not represent the same number of items, but do represent the same proportion.

Make comparisons between two data sets.

The pie chart gives information about the 76 students in the same four Year 11 classes at Trowton School.

Number of students in Year 11 of Trowton School



Tolu says,

"It is more difficult to find out the numbers of students in each class from the pie chart than from the bar chart."

(*c*) Is Tolu correct?
You must give a reason for your answer.

(1)

(Total 5 marks)

New SAMs Paper 3F qu.6 (S2 – AO2)

OPPORTUNITIES FOR REASONING/PROBLEM SOLVING

Choose which type of graph or chart to use for a specific data set and justify its use.

Evaluate statements in relation to data displayed in a graph/chart.

NOTES

Interquartile range is covered in unit 16.

Misleading graphs are a useful activity for covering AO2 strand 5: Critically evaluate a given way of presenting information.

When doing time–series graphs, use examples from science, geography.

NB Moving averages are not explicitly mentioned in the programme of study but may be worth covering too.

14b. Cumulative frequency, box plots and histograms (S3, S4, S5)	**Teaching time** 6–8 hours

OBJECTIVES

By the end of the sub-unit, students should be able to:

- Use statistics found in all graphs/charts in this unit to describe a population;
- Know the appropriate uses of cumulative frequency diagrams;
- Construct and interpret cumulative frequency tables;
- Construct and interpret cumulative frequency graphs/diagrams and from the graph:
  - estimate frequency greater/less than a given value;
  - find the median and quartile values and interquartile range;
- Compare the mean and range of two distributions, or median and interquartile range, as appropriate;
- Interpret box plots to find median, quartiles, range and interquartile range and draw conclusions;
- Produce box plots from raw data and when given quartiles, median and identify any outliers;
- Know the appropriate uses of histograms;
- Construct and interpret histograms from class intervals with unequal width;
- Use and understand frequency density;
- From histograms:
  - complete a grouped frequency table;
  - understand and define frequency density;
- Estimate the mean from a histogram;
- Estimate the median from a histogram with unequal class widths or any other information from a histogram, such as the number of people in a given interval.

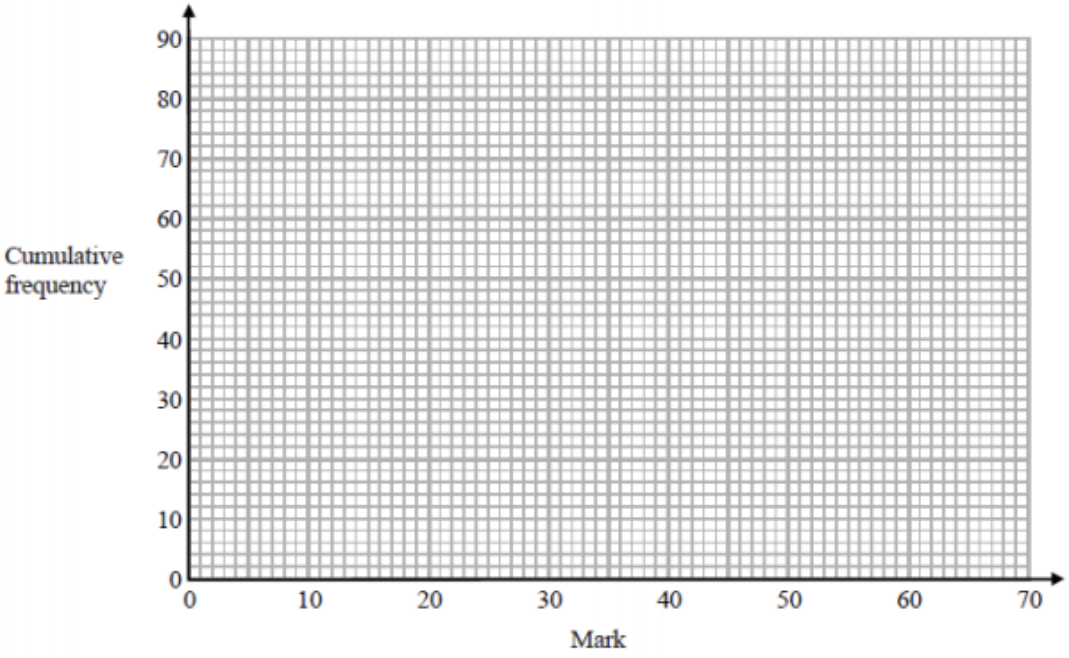
POSSIBLE SUCCESS CRITERIA/EXAM QUESTIONS

Construct cumulative frequency graphs, box plots and histograms from frequency tables.

Compare two data sets and justify their comparisons based on measures extracted from their diagrams where appropriate in terms of the context of the data.

Mark (m)	Cumulative frequency
$0 < m \leqslant 10$	8
$0 < m \leqslant 20$	23
$0 < m \leqslant 30$	48
$0 < m \leqslant 40$	65
$0 < m \leqslant 50$	74
$0 < m \leqslant 60$	80

(a) On the grid, plot a cumulative frequency graph for this information.



(2)

(b) Find the median mark.

(1)

Students either pass the test or fail the test.
The pass mark is set so that 3 times as many students fail the test as pass the test.

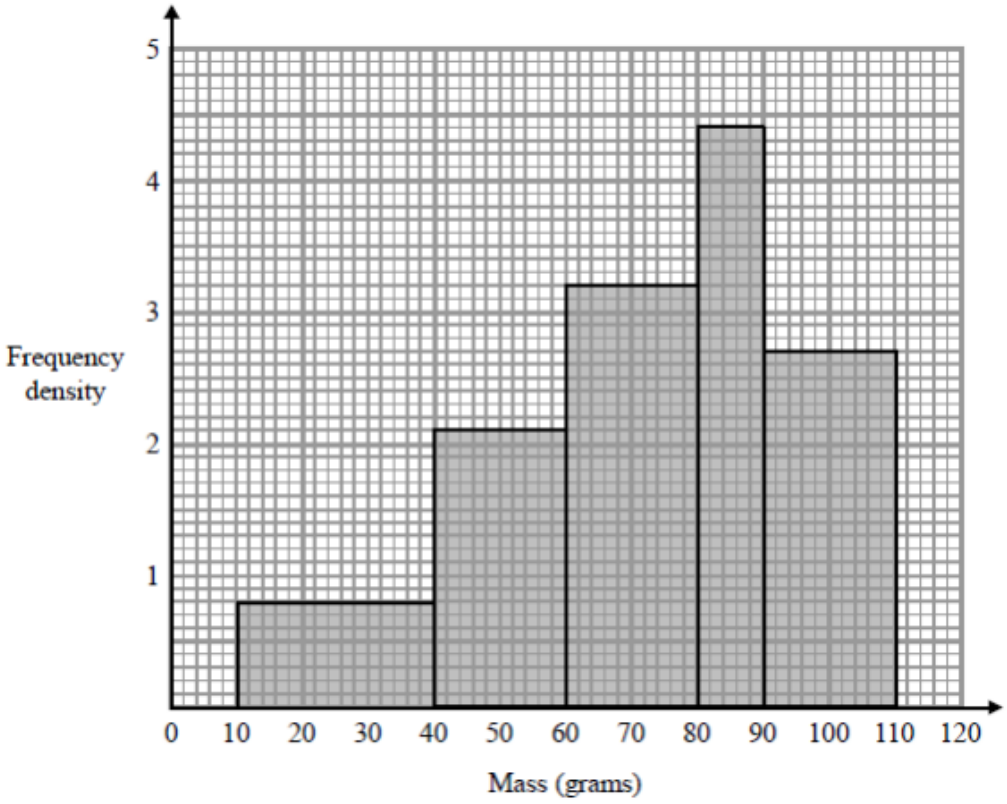
(c) Find an estimate for the lowest possible pass mark.

(3)

(Total 6 marks)

New SAMs Paper 3H qu.7 (S3, R9 – AO1/AO2/AO3)

A biologist is studying the effects of global warming on animal size.
The histogram gives information about the masses of a species of snail in a sample he took in 2013 from a large lake.



The mean mass of the same species of snail taken from the lake in 2003 was 75 grams.

(*a*) Is there any evidence to support the hypothesis that the mass of this species of snail has decreased?

(5)

(*b*) Explain whether it is possible to state what the mode is from this histogram.

(1)

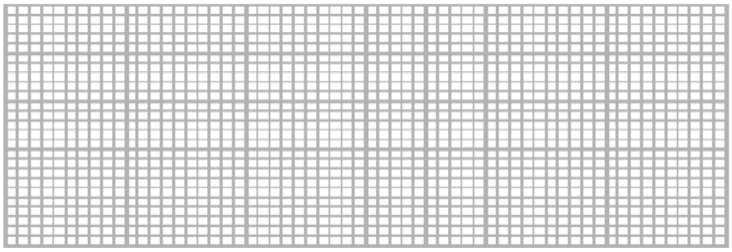
(Total 6 marks)

Original SAMs Paper 2H qu.14 (S3, S4, S5 – AO1/AO2/AO3)

Ben played 15 games of basketball.
Here are the points he scored in each game.

17 18 18 18 19 20 20 22 23 23 23 26 27 28 28

(*a*) Draw a box plot for this information.



(3)

Sam plays in the same 15 games of basketball.

The median number of points Sam scored is 23
The interquartile range of these points is 12
The range of these points is 20

(*b*) Who is more consistent at scoring points, Sam or Ben?
You must give a reason for your answer.

(2)

(Total 5 marks)
New SAMs Paper 1H qu.14 (S4 – AO2)

OPPORTUNITIES FOR REASONING/PROBLEM SOLVING

Interpret two or more data sets from box plots and relate the key measures in the context of the data.

Given the size of a sample and its box plot calculate the proportion above/below a specified value.

COMMON MISCONCEPTIONS

Labelling axes incorrectly in terms of the scales, and also using 'Frequency' instead of 'Frequency Density' or 'Cumulative Frequency'.

Students often confuse the methods involved with cumulative frequency, estimating the mean and histograms when dealing with data tables.

NOTES

Ensure that axes are clearly labelled.

As a way to introduce measures of spread, it may be useful to find mode, median, range and interquartile range from stem and leaf diagrams (including back-to-back) to compare two data sets.

As an extension, use the formula for identifying an outlier, (i.e. if data point is below

LQ – 1.5 × IQR or above UQ + 1.5 × IQR, it is an outlier). Get them to identify outliers in the data, and give bounds for data.