edexcel

Teacher's Guide

Edexcel GCSE in Statistics

ALWAYS LEARNING

PEARSON

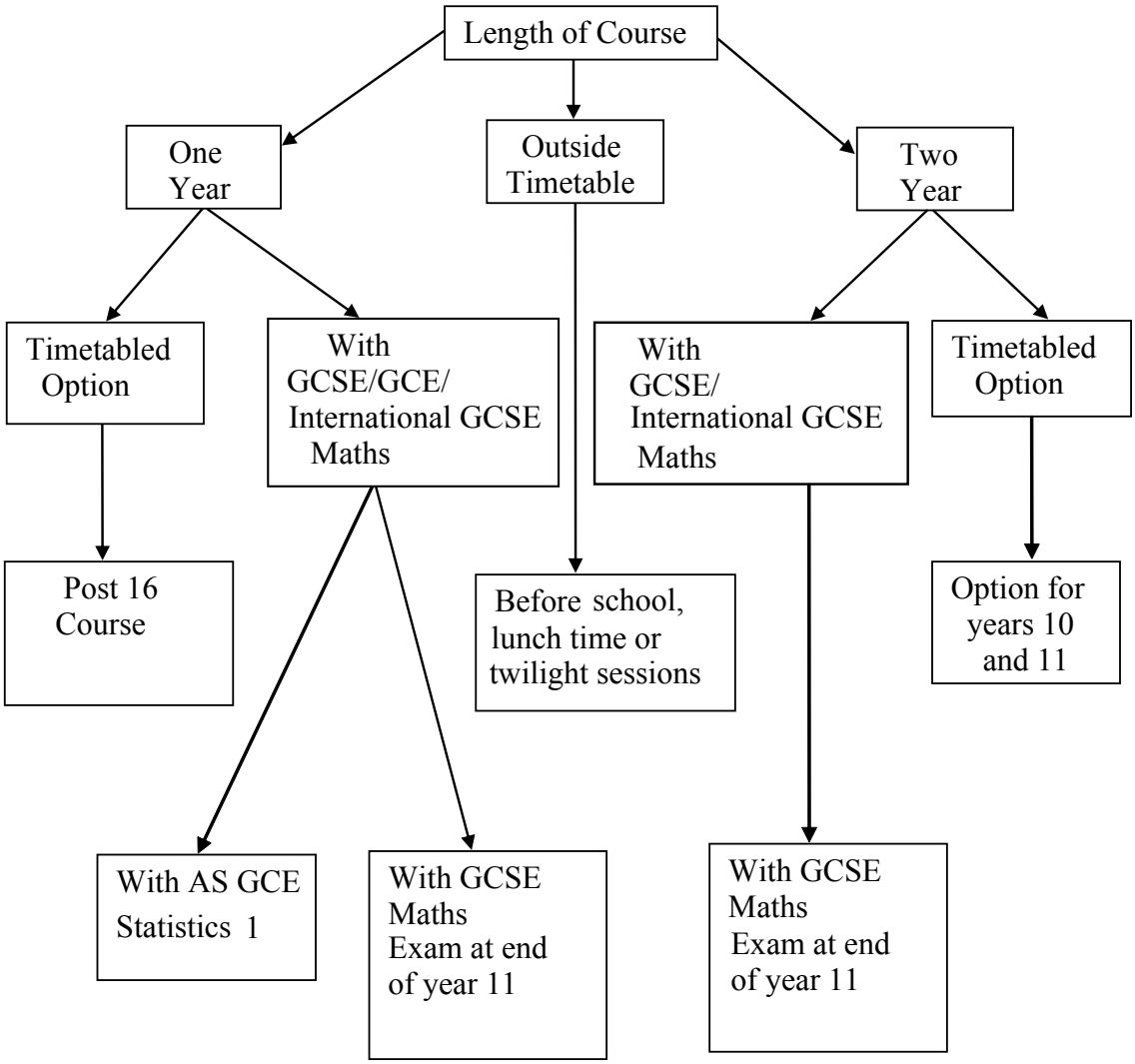
Contents

Delivery models

There are many ways of delivering the Edexcel GCSE in Statistics course.

The way you choose to deliver the course will depend on the amount of time you have available and the approach you wish to take.

The flowchart below gives an overview of most of the options available.



Here are some suggestions for planning for the delivery method you choose.

Teaching GCSE Statistics as a stand-alone course

When teaching GCSE Statistics, the Edexcel scheme of work will provide a good route through the specification. This is available for download on the Edexcel website.

Teaching alongside GCSE Mathematics courses

There are topics in the Statistics specification that are not covered within the Mathematics specification and time will have to be given to covering these separately.

A scheme of work outlining the detail of this content, and how to plan for it in your teaching, will be available before the start of the course.

Planning for the controlled assessment

Some timetabled time will be needed for the controlled assessment.

This needs to be carried out no later than early in the spring term in the examination year so that students have time to collect their data. If it is started too early students will not know enough about the statistical techniques they need for the work.

Students are encouraged to use ICT but this is not essential.

There are several topics that will be rarely used in the controlled assessment and it is a good idea to leave teaching of these until the end. This will enable the controlled assessment to be carried out in time to be sent to your moderator by the beginning of May.

Topics that students are unlikely to use in the controlled assessment are:

At Foundation Tier

- Most Probability
- Simple Index Numbers
- Chloropleth Maps

At Higher Tier

- Most Probability
- Binomial Distribution
- Index Numbers (simple, chain, weighted)
- Quality Control

Teaching ideas

Here are some ideas on the sort of approaches you can use to introduce some topics.

Stem and leaf diagrams, Foundation Tier, targeting grades E to G

Objective

To introduce stem and leaf diagrams.

Key vocabulary

Stem, leaf, key.

Introductory activity

Show the 'word': *ofhrsaromdrceo*

Discuss what it says. Emphasise that to make sense of this the letters have to be put in a particular order: *orderfromchaos*.

Still doesn't mean much! You need to arrange them into groups to make sense: *order from chaos*.

Main activity

Ask quick-fire questions about this list of numbers:

34 48 56 23 42 38 36 44 58 52 23 42 46 48 59 22

Which is the lowest number? Which is the highest?

Discuss how this could be made easier, emphasising ordering:

23 23 22 All the numbers beginning with a 2

34 38 36 All the numbers beginning with a 3

48 42 44 42 46 48 All the numbers beginning with a 4

56 58 52 59 All the numbers beginning with a 5

Discuss how to make it even simpler, leading to using the first digit once:

2	3 2 3
3	4 8 6
4	8 2 4 2 6 8
5	6 8 2 9

Now go the final step and order all the numbers, adding a key:

2	2 3 3
3	4 6 8
4	2 2 4 6 8 8
5	2 6 8 9

Key
2 | 3
Means 23

Discuss how you can think of this diagram as being like a tree where each row begins with the stem and the numbers that join it are leaves.

Follow up

Use practice exercises from any resources you are using to help you deliver the course.

Introduction to probability, Foundation Tier, targeting grades D to G

Objectives

To understand the definition of probability. Measures of probability.

Key vocabulary

Certainty, impossible, likely, unlikely, evens.

Preparation

You need red, green and blue counters or cubes and a bag or box.

Main activity

- Put only red counters in a bag.

Ask students: 'If I take one counter out, what is the likelihood that it will be red?' What is the likelihood that it will be green?

Encourage correct language: You are **certain** to get a red. It is **impossible** to get a green.

- Add 2 green counters to the bag.

Discuss: 'If I take one counter out, what is the likelihood that it will be red?' What is the likelihood that it will be green?

Encourage correct language: You are **likely** to get a red top. It is **unlikely** to be a green counter.

- Put an equal number of red and green counters in the bag.

Discuss: 'Which is most likely a red or a green?'

Emphasise that **the probability** of getting a red is the same as that of getting a green. The probabilities are **evens**. It is equally likely to be red as green.

- Put 6 green and 4 blue counters in the bag.

Discuss: 'Which colour am I most likely to pick out?'

Emphasise: green, because there are 6 green and only 4 blue.

Emphasise that the **probability** of getting a green is 6 out of 10. Discuss how to write 6/10 as a decimal: 0.6

Discuss: 'What is the probability of getting a blue?'

The probability of a blue is 4 out of 10 you write this as 0.4

- Introduce a probability line with 0 for impossible and 1 for certain. Discuss where to mark in probabilities of 0.6 and 0.4

Encourage students to suggest where to put the words certain, impossible, likely, unlikely and evens onto the line.

Follow up

Work out a variety of probabilities using 10 counters in a bag with different ratios of 2 colours. Then 15 in a bag with different ratios of 2 colours. Move on to 3 colours in the bag.

Petersen capture/recapture method, Higher Tier, targeting grades D to C

Objective

To introduce the capture/recapture method.

Key vocabulary

Estimate, capture/recapture.

Introductory activity

Tell students they are going to find out how many fish there are in a pond. Show them a bag and say it represents a pond. You will need a pile of stones (minimum 50 so students can't count them). They should all be about the same size, so pebbles are good. Show students the pile of stones and put them in the bag. Tell students that each stone represents one fish, and that some fish will be bigger than others.

Main activity

Discuss how many fish students think are in the pond. Students record their guesses.

Discuss what makes a good guess.

Emphasise that guesses based on logical reasoning are called **estimates**. An estimate is better than a guess.

Discuss how to get a better estimate of n, the number of fish in the pond.

Take out 10 fish and mark them in some way. Return them to the pond and mix again. Emphasise that there are now 10 fish out of n in the pond.

Prompt students to suggest taking out another sample – of, say 20 fish.

Count the number of marked fish in the sample. (We'll say there's x.)

Emphasise that if the sample proportion of marked fish is the same as the proportion in the pond, then $\frac{x}{20} = \frac{10}{n}$

Discuss how to work out n.

$$n = \frac{20 \times 10}{x}$$

Replace the 20 fish, shake up and try another sample of 20. Work out n again.

Why are the two values calculated different? Discuss estimation.

Follow up

Count the number of fish and see which was better-guessing or estimating.

Follow up with exercises from any resources you are using to help you deliver the course.

Standardised scores, Higher Tier, targeting grades B to A*

Objective

To introduce standardised scores.

Key vocabulary

Standardised score.

Introductory activity

Jack and Olivia were the most popular names on the list published by the Office of National Statistics in 2006. Students can look up popular names at www.statistics.gov.uk/CCI/nugget.asp?ID=184, so they can see if they have a popular name.

Tell students that Jack and Olivia are twins. Their father said that the one that did best in the exam for their favourite subject would get extra pocket money.

Show the exam results in the table:

Favourite exam	John's mark	Olivia's mark	Mean mark	Standard deviation
English		62	50	8
Mathematics	60		42	12

Main activity

Discuss who students think should get extra pocket money.

Record how many think Jack and how many think Olivia.

Expect these arguments:

- Olivia's mark was higher than Jack's mark. 62 > 60. She should get the extra pocket money. Discuss what's wrong with this argument.
- Jack said it should be the number of marks above average that counts and (60 – 42) > (62 – 50). Discuss whether this is fair.

Explain that Jack's standard deviation, which is a measure of spread about the mean, is bigger than Olivia's. Discuss what this means and whether it makes it easier to be more above average.

Follow up

Students decide that they would take the number of standard deviations above the mean as a measure of success.

Olivia had $\frac{(62-50)}{8} = 1\frac{1}{2}$ and Jack had $\frac{(60-42)}{12} = 1\frac{1}{2}$

They both received a rise in pocket money. So they were both wrong.

The measure used is called the standardised score.

$$\text{standardised score} = \frac{\text{score} - \text{mean}}{\text{standard deviation}}$$

Students can practise using exercises in any resources you are using to help you deliver the course.

Student guide

Why should I study this course?

The age of statistics is upon us.

Statistics are being used at an increasing rate in business, politics and the sciences.

In the real world we are constantly bombarded, through the media, with graphs, economical data such as retail price index (RPI) and various other statistical statements.

You will find that understanding statistics is important in all aspects of everyday life.

When you have completed this course you should have a far greater understanding of how to interpret and use statistical statements.

If you are continuing in education, you will have acquired important skills required for studying other subjects. Biology, psychology, economics, geography, engineering and archaeology are among the many subjects that rely at times on an understanding of statistical data.

If you decide to continue with a mathematics subject to AS Level you will find that you already know a lot of the content of the statistics options.

What will I learn on the course?

Statistics give us information in one form or another.

In this course, you will learn how to collect unbiased data and how to summarise and represent these data.

An important part of statistics is that you should use the available data to predict what will happen in the future. You will learn how to reason, interpret and discuss results, and how to make forecasts based on these results.

This course will give you the skills to read and analyse data. You will be able to interpret various diagrams and statistical statements accurately. You will also be able to appreciate when statistical statements are exaggerated.

We cannot always be certain that a forecast will be correct, so you will also consider the probability of something happening.

How will I be assessed?

The course is assessed in two parts.

Task (teacher assessed)	Written paper
Approximately 8 to 10 hours	Foundation Tier 1 hour 30 mins Higher Tier 2 hours
This will involve you choosing, with the help of your teacher, a task to investigate. You will make a plan, collect data, analyse it and then write up a report on your work. Some of this will be done during lesson time.	Each paper will have some short and some long questions. Many of them will involve the use of real-world data. You could be asked for example to design questionnaires, draw box plots and scatter diagrams, interpret diagrams, carry out calculations and interpret statements.

What can I do after I've completed the course?

You can use the skills you have acquired when studying further subjects at all levels. You might also use these skills to make decisions about things in your own life such as investing savings, choosing electrical equipment and buying cars etc.

You might consider studying statistics at a higher level. It forms part of the GCE Mathematics course.

Statistics can also be studied at university where it can lead on to a variety of fascinating jobs. These include working in most areas of manufacturing, insurance and the civil service. Statisticians are needed in most walks of life.

What do I need to know, or be able to do, before taking this course?

You will need very little previous statistical knowledge to be able to take this course. In Key Stage 3 you will have done some data-handling work which will be repeated at the beginning of this course.

Foundation students need to have a basic understanding of accuracy, fractions, decimals and percentages.

Higher-level students need, in addition, to be familiar with the equation of a straight line and have a basic understanding of the shapes of exponential curves.

Next steps!

Go online and have a look at some of the resources available for GCSE students, such as www.revision-notes.co.uk, and then talk to your teacher about how suitable this course is for you.

Assessment overview

The grid below gives an overview of the assessment for this course.

Most of this information can be given to your students if you think it would be helpful.

Unit	Level/type of assessment	% of total	Mark	Time	Availability
1	Foundation Written paper Grades G to C available	75	80	1 ½ hours	June
1	Higher Written paper Grades D to A* available	75	100	2 hours	June
2	Foundation/Higher controlled assessment	25	40	Teacher assessed Recommended time approximately 8–10 hours	June

Type of assessment	Description	Knowledge and skills
Written paper	Each written paper will consist of: • some short and some long questions • questions involving real-life data • questions ramped in difficulty with easy questions at the beginning and getting progressively more demanding • questions on any aspect of the specification content • compulsory questions.	The Assessment Objectives covered in this assessment, as a percentage of the total assessment, are: AO1: 4–14% AO2: 5–15% AO3: 32–42% AO4: 19–29%
Controlled	Students produce a statistical task in formally and informally supervised stages. Students choose one from three tasks. Tasks with a teacher booklet, a student booklet and task-specific assessment schemes will be given to the centre in advance of the assessment. See the controlled assessment guide on page 31 for more information.	The Assessment Objectives covered in this assessment, as a percentage of the total assessment, are: AO1: 6.25% AO2: 5% AO3: 7.5% AO4: 6.25% The work will be done in three stages: 1 planning 2 collecting, processing and representing data 3 interpreting and analysing data.

Examination questions

This section contains practice examination questions with model answers so you can see the sorts of questions that will be asked and the standards required.

Examiner comments are provided to help you understand what students need to do in the written paper, and the common mistakes you can help them overcome.

There is a mixture of short questions which test basic ideas and longer questions which apply statistical lines of enquiry.

The Assessment Objectives for which marks are given, AO1, AO2, AO3 or AO4, are shown next to each mark.

Question 1 Foundation Tier, targeting grades E, F and G

Choice Engineering Ltd has an equal number of male and female employees.

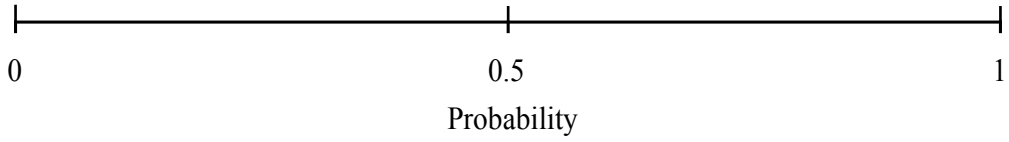
An employee is to be chosen at random.

On the probability scale below:

(i) use *A* to mark the probability that the employee chosen is male,

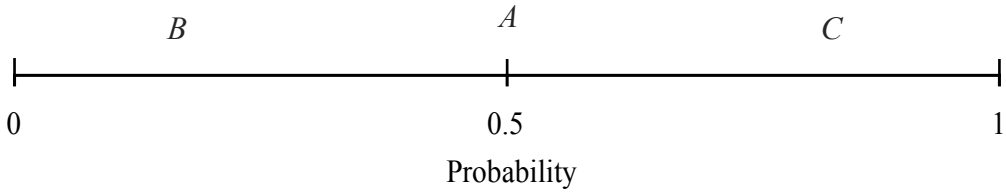
(ii) use *B* to mark the probability that the employee chosen was born on a Tuesday

(iii) use *C* to mark the probability that the employee chosen was **not** born on a Sunday.



(3 marks at AO3)

Model answer



Examiner's comment

The line is usually drawn to a sensible length (eg 10 cm). Students rarely measure the line to work out positions. A is often correct but B and C are often wrong.

Question 2 Foundation Tier, targeting grades E, F and G

A nature reserve is divided into 16 regions of equal size.

The number of orchids in each region is counted.

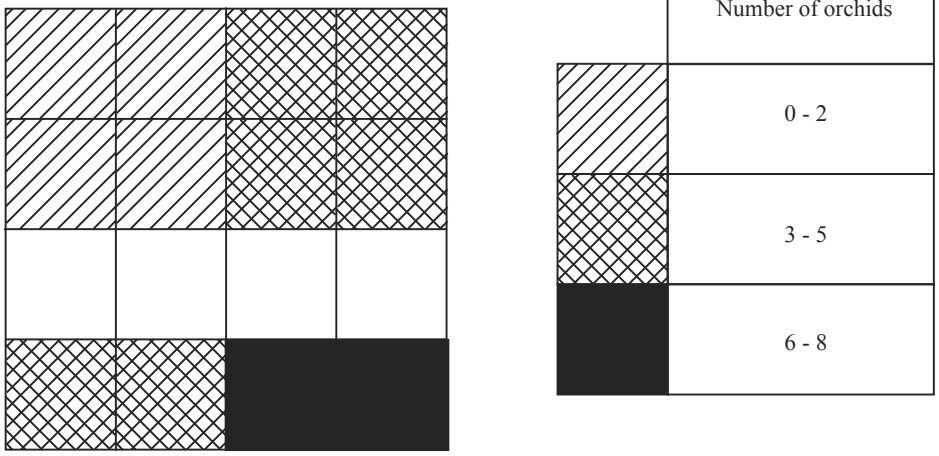
The results are shown in the following diagram.

0	0	3	4
1	1	4	3
2	1	4	6
4	5	7	8

5 means 5 orchids grow in this region

(a) Use the information in the diagram to complete the choropleth map.

12 regions have already been shaded.



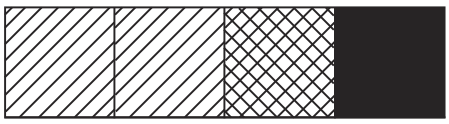
(1 mark at AO3)

(b) Describe how the orchids are spread across the reserve.

(1 mark at AO4)

Model answer

(a)



(b) Most orchids are at the bottom right of the reserve, the least at the top left.

Examiner's comments

With choropleth maps, students will be expected to either fill in with texture, as in (a), and answer questions as in (b)

OR

answer questions on a map that is already drawn.

Question 3 Foundation Tier, targeting grades C to G

The monthly rainfall figures, in mm, for Bretown in England in 2006.

95 60 58 51 68 55 59 73 73 88 86 91

(a) Write down the mode of the data.

(1 mark at AO3)

(b) Work out the median monthly rainfall.

(2 marks at AO3)

(c) Work out the mean monthly rainfall.

(2 marks at AO3)

The Bretown tourist centre wants to let visitors know how much rainfall they can expect.

(d) Which of the above figures should they use if they wish to give a true picture of the rainfall? Give a reason for your answer.

(2 marks at AO4)

(e) What other measure should be used in order to give a complete picture of the rainfall?

(1 mark at AO4)

Model answer

(a) 73

(b) 51 55 58 59 60 68 73 73 86 88 91 95

Median is $\frac{(60 + 68)}{2} = 64$

(c) Mean $= \frac{795}{12} = 66.25$

(d) The mean, because it uses all the data.

(e) Range or IQR or standard deviation.

Examiner's comments

(b) Students often get mean and median mixed up. 60 or 68 are common incorrect answers.

(d) Understanding when to use each measure is usually poor and students lose marks here.

(e) Common incorrect answers are mean and mode. The importance of spread is not well understood at Foundation Tier.

Question 4 Foundation Tier, targeting grades C to E

The local council is planning to build a new community centre.

The councillors want to get the views of the local people.

Councillor Ali wants to take a census of the town's population.

(a) (i) Give **one** advantage of taking a census.

(ii) Give **one** disadvantage of taking a census.

(2 marks AO1)

Councillor Jones suggests taking a sample of the people who attend the council offices.

(b) Explain why this would not be a representative sample.

(1 mark AO1)

Councillor Khan suggests taking a simple random sample of 100 people.

(c) Describe, in detail, how the council could take a simple random sample.

(2 marks AO2)

The council decide to use a questionnaire.

Councillor Turner suggests they ask the question,

'Do you agree that a new community centre would be a good idea?'

This is a leading question.

(d) Rewrite this so that it is not a leading question.

(Remember to include answer boxes.)

(2 marks AO1)

Model answer

(a) (i) Unbiased or accurate or takes in whole population.

(ii) Time consuming or expensive or difficult to ensure whole population is surveyed.

(b) The sample would be biased since the majority of people do not visit the council offices.

(c) Number each person. Generate 100 random numbers. Select people corresponding to the random numbers.

(d) 'Do you or do you not think that a new community centre would be a good idea?'

Yes ☐ *No* ☐

Examiner's comments

With the increased weighting on Assessment Objective AO1 more questions of the above type are to be expected.

(c) Students tend to forget to number the people.

(d) 'Do you think' is a common answer and useful in this sort of question. Boxes are often left out leading to lost marks.

Question 5 Foundation Tier/Higher Tier, targeting grades C and D

A council wants to find out people's views on recycling.

It wants to know if there is a difference in the support for recycling between males and females and between those aged 25 and under and those aged over 25.

(a) Write down **two** suitable hypotheses for investigation.

(2 marks at AO1)

(b) Use the best word from the list to complete the sentences below.

qualitative primary quantitative secondary ratio

(i) Age is..... data.

(ii) Gender is..... data.

(2 marks at AO2)

(c) Which of the following sampling frames would be most suitable for the council to use? Give a reason for your answer.

A A list of names in a telephone directory.

B A list of council tax payers.

C A list of residents in one street.

(2 marks at AO1)

Model answer

(a) There is a difference between males and females in their support for recycling. There is a difference between those aged 25 and under and those over 25 in their support for recycling.

(b) (i) Quantitative (ii) Qualitative.

(c) B. Taxpayers have to pay for the recycling and they elect the council.

Examiner's comments

More emphasis will be on AO1 in the new specification and this is an example of the type of question that may be met in the future.

Answers to (a) and (c) might vary but must always include sensible reasons.

In part (c), spellings have to be recognisable.

Question 6 Foundation Tier/Higher Tier, targeting grades C and D

A researcher wants to choose a random sample of 300 people from her town.

She considers choosing the people that live near her house.

(a) Write down one advantage and one disadvantage of this method of choosing her sample.

(2 marks at AO1)

There are 30 000 people on the electoral register in her town.

(b) Describe how the researcher should choose a systematic sample of 300 people from the electoral register.

(2 marks at AO2)

Model answer

(a) Advantage — it is convenient. Disadvantage — it is likely to be biased.

(b) Get a random number between 1 and 100. Starting at this number select every 100th person on the register.

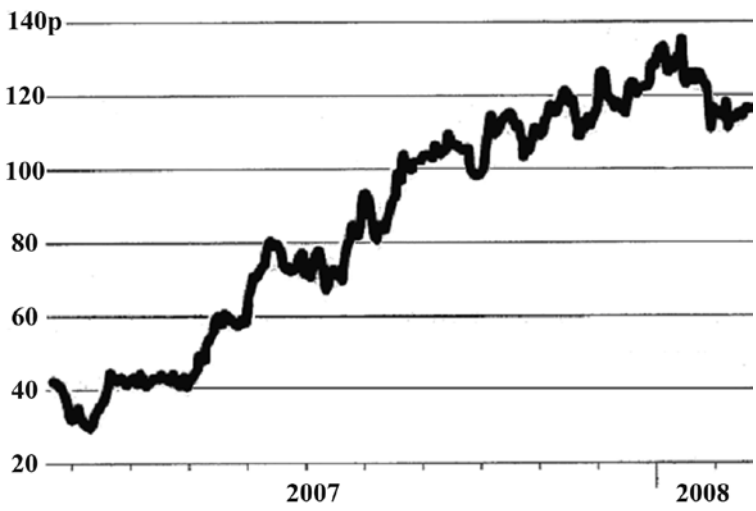
Examiner's comments

(a) Other wording may be acceptable.

(b) Most students forget to establish a random starting point.

Question 7 Foundation Tier/Higher Tier, targeting grades C and D

The graph gives information about the price, in pence, of buying a share in Bodgers Unlimited.



The newspaper said 'Shares in Bodgers Unlimited have increased a lot over the past year'.

(a) Write down **one** misleading feature of the graph that makes the increase look greater than it was.

(1 mark at AO4)

(b) Write down **one** other misleading feature of this graph.

(1 mark at AO4)

Model answer

(a) Scale does not start at zero.

(b) Line too thick or shadow on line makes it difficult to read.

Examiner's comments

Students may be asked to comment on pie charts and other diagrams too.

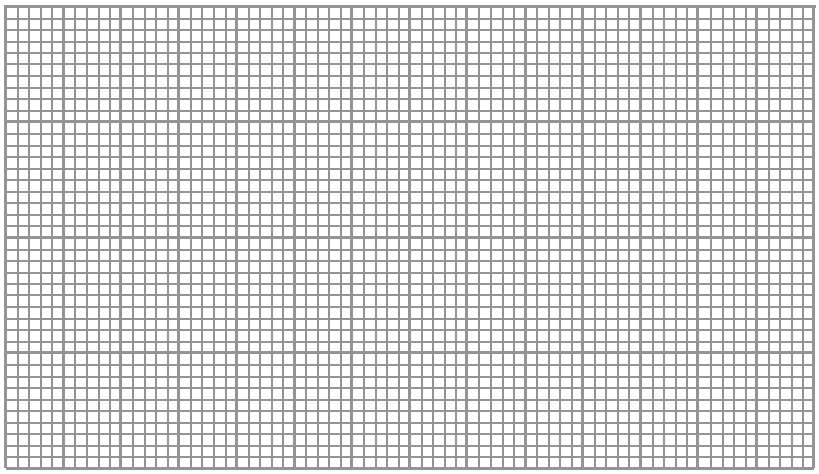
Note that the zigzag of the line and the position of the title are not misleading.

Question 8 Higher Tier, targeting grades A to C

300 female and 220 male turtles are weighed. The summary statistics of these data are shown in the table below.

Turtle weight (grams)	Minimum	Lower quartile	Median	Upper quartile	Maximum
Female	346	513	604	704	952
Male	621	987	1110	1211	1383

(a) Draw two box plots to represent these data on the grid.



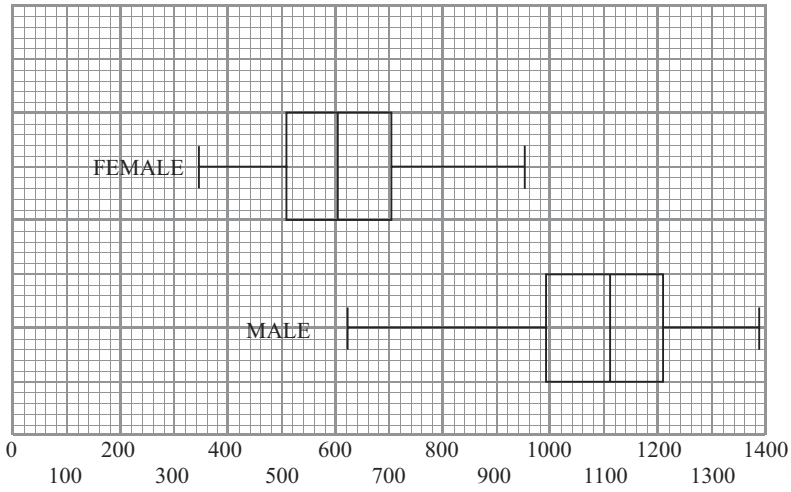
(4 marks AO3)

(b) Compare and contrast the distributions of the weights of adult turtles.

(2 marks AO4)

Model answer

(a) Box with whiskers. Medians correct. Quartiles correct. End points correct. (Some latitude given.)



(b) Male has greater median weight than female. Male has greater IQR (or range). So, males weigh more on average than females but their weight is more variable.

Examiner's comments

A measure of central tendency and a measure of dispersion is the minimum needed to compare distributions.

Students tend to compare end points.

In part (b), a correct reference to skew would get a mark.

Question 9 Higher Tier, targeting grades A* to B

The table below shows how the price of a house bought by first-time buyers has changed over a period of five years. It also shows some of the chain base index numbers for the prices.

The prices to the nearest £ 1000 are taken from the Nationwide Database of House Prices.

Year	2001	2002	2003	2004	2005
Price (£ 1,000s)	£70	£87.5	£ 105	£ 136.5	£ 163.8
Chain base index number	123	125	120		

(a) Calculate, to the nearest whole number, the chain base index numbers for years 2004 and 2005, and enter them in the table.

(2 marks at AO3)

(b) What is the trend of the chain base index numbers?

(1 mark at AO4)

Model answer

(a) $136.5 - 105 = 120$ $163.8 - 136.5 = 120$

(b) Fairly level or even.

Examiner's comments

Students find chain base index numbers very difficult, and these types of questions are usually answered badly.

Students often explain each movement rather than the trend. This does not receive any credit.

Question 10 Higher Tier, targeting grades A* to D

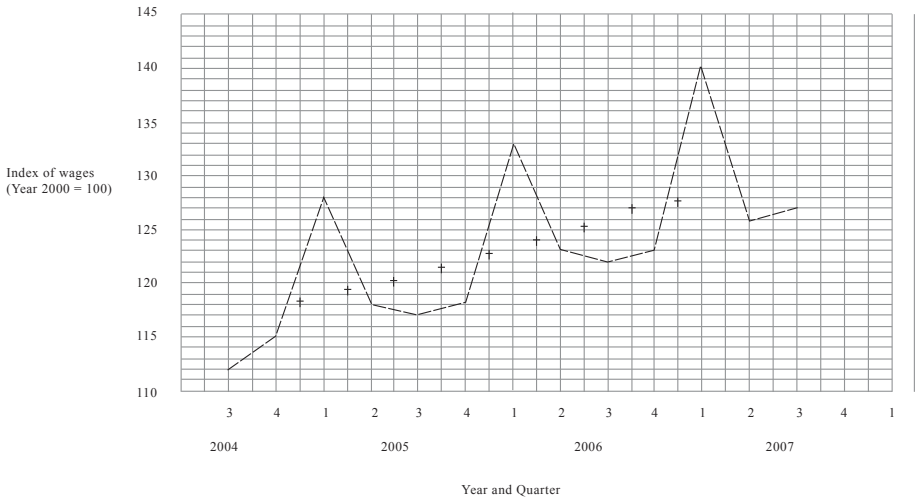
The table shows the quarterly index of wages from the third quarter of 2004 to the third quarter of 2007. The year 2000 is the base year for the index.

	Quarter			
Year	1	2	3	4
2004			112	115
2005	128	118	117	118
2006	133	123	122	123
2007	140	126	127	

Data source: statistics.gov.uk

These data have been plotted as a time series on the graph below.

All the 4-point moving averages, except the last one, are also plotted.



(a) Calculate the last 4-point moving average and plot it on the graph.

(2 marks at AO3)

(b) Draw a trend line on the graph.

(1 mark at AO3)

(c) Write down the quarter in which wages were at their highest each year.

(1 mark at AO4)

The average wage was £18 000 in 2000.

(d) Taking 2000 as the base year, use the time series information to work out the average wage in quarter 1 of 2005.

(2 marks at AO3)

(e) To the nearest whole number, work out the average seasonal variation for quarter 1. Show your working.

(2 marks at AO3)

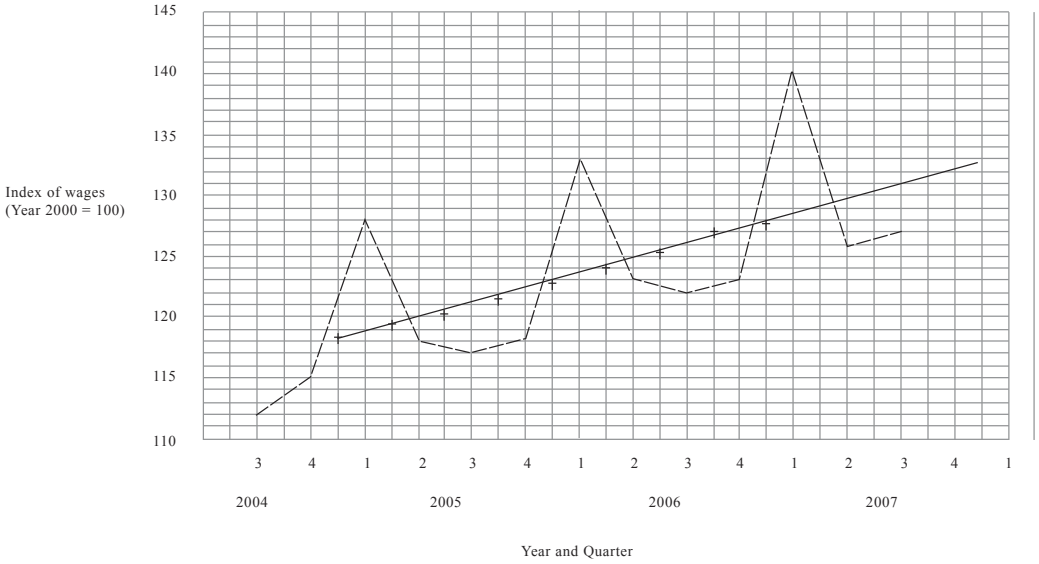
(f) Predict the value of the index of wages for quarter 1 of 2008.

(3 marks at AO4)

Model answer

(a) $\frac{(123 + 140 + 126 + 127)}{4} = 129$ *(plotted below)*

(b)



(c) *Quarter 1*

(d) $\frac{128 \times £18\,000}{100} = £23\,040$

(e) $\frac{(9 + 9\frac{1}{2} + 11)}{3} = 9\frac{5}{6} = 9.8$

(f) *9.8 + 133 = 142.8*

Examiner's comments

(a) Plotting at the correct place on the ***x*** axis is often incorrect.

(b) Some latitude is allowed here.

(c) The question says **'each year'** — Q1 in 2007 is a common incorrect answer.

(d) This is 128% of £18 000.

(e) Again some latitude is allowed on this answer.

(f) Their answer to (e) + 133 is acceptable.

Question 11 Higher Tier, targeting grades A and B

On a production line in a factory, lentil soup is put into tins.

The label on each tin says that the contents weigh 400 g.

(a) Give two reasons why it is not practical to check the weight of the contents of each tin.

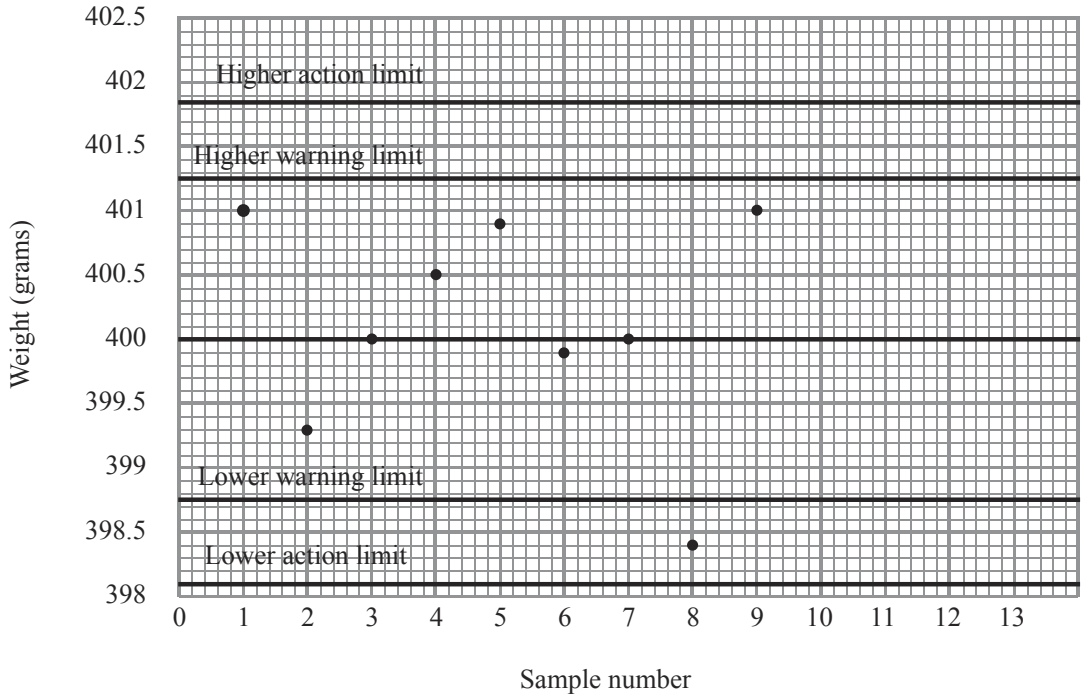
(2 marks at AO4)

Samples are taken at intervals of half an hour and the weights of the tins' contents are found.

The mean weight of each sample is plotted on a quality control chart.

The mean weights of the 10th and 11th samples were 401 g and 402 g.

(b) Plot the 10th and 11th samples on the chart below.



(2 marks at AO3)

(c) Comment on any action that would have been taken during the 6-hour shift.

(2 marks at AO4)

Model answer

(a) It would take too long.

(b) Correctly plotted points.

(c) A second sample would have been taken after sample 8. The machine would be stopped and reset after sample 11.

Examiner's comments

This process is used in industry so it is a practical application of statistics.

In part (c) students often mix up the new sample/reset limits.

Question 12 Higher Tier, targeting grades A* and A

A baker finds that the total quantity of flour used by his company each week is normally distributed with a mean of 600 kg and a standard deviation of 8 kg.

(a) Calculate an estimate of the maximum quantity of flour the company is likely to use per week.

(1 mark at AO4)

The manager decides to make a regular order for the same amount of flour to be delivered each week on a Monday.

(b) Discuss which of the mean and maximum is a suitable figure for the quantity to be ordered.

(2 marks at AO4)

Model answer

(a) $600 + 3 \times 8 = 624$ kg

(b) The mean of 600 kg would be correct in the long run. 624 kg would result in the amount of stock growing over time.

Examiner's comments

(a) 616 is a common incorrect answer.

(b) A variety of answers could be acceptable.

Answers must refer to what would happen if more than the mean is ordered each time.

Controlled assessment

Controlled assessment requires students to carry out a statistical task. It is similar to the coursework projects except that controlled stages have been added to help you ensure the work is the student's own.

There are three activities where controls apply: task setting, task taking and task marking. The level of control for each activity was specified by QCDA.

This section explains the level required for each activity and what it means for you and your students.

Task setting

This is very similar to the current arrangements, so will be familiar.

What is the level of control?

High.

What does this mean?

Edexcel will produce three tasks each year.

Students choose one of the three tasks to investigate.

The variety of tasks each year will help you choose the task that will engage your students. There will be at least one task which requires the collection of primary data, and at least one which allows students to use secondary data, providing links to websites, databases and other information sources.

How often will the tasks change?

The tasks will be replaced every year. The window for the tasks will run from late spring term to late spring term.The tasks will be pre-released in early spring each year for teachers to plan teaching.

Any students wanting to retake the controlled assessment unit will need to use the one available for the session in which they are retaking, regardless of what task they did originally.

If students are taking the same task again in the same year, they must start from scratch and do the whole task again.

Task marking

This is similar to the current arrangements, so will be familiar. You mark the task. You then fill in the student record form (which you will find at the back of the specification) to show all the marks achieved. Edexcel will ask for a sample of the work to moderate, including student work with high and low scores. Edexcel will moderate the work and you will receive a report on the moderation.

You should mark the planning part first before students embark on the next stage.

Training from Edexcel courses on marking tasks will be available to help you mark work effectively. Our specification experts can also provide support, just email gcsestatistics@edexcelexperts.co.uk.

Task taking

Task taking is split into three stages. The levels of control and the effect are different for each stage. You should allow between 8 and 120 hours in total.

Breaking the task into stages provides a structure and allows students to focus on one stage at a time.

A student booklet and a teacher booklet will be available for each task to support you through each stage. You can support and question students throughout the investigation.

Students should be encouraged to use ICT throughout, and especially during the collecting and representing stage.

1 *Planning (1–2 hours, formal supervision)*

The first part of the investigation is the planning stage.

You introduce the task to students.

Students plan their investigation under formal supervision in the classroom.

This means that the teacher must be present at all times to authenticate the work. The teacher monitors students' work and can check that it is their own by benchmarking against previous work and expectations. You can advise the student, provide guidance as the work progresses, and make suggestions if a student is attempting a project which is not going to allow them to demonstrate their full statistical ability.

You should not suggest a specific direction for students to take.

It is recommended that students undertake their planning over more than one session so that you can give feedback before they finalise the plan.

At the end of the planning stage, students will give you their plans to mark, keeping a copy. Write any feedback on the student record form.

2 *Collecting and representing data (4–7 hours, limited supervision)*

In this stage, students collect data, perform calculations and draw diagrams. Limited supervision means that students are able to do this independently without a teacher being present, as for current coursework arrangements.

The project should follow the student's plan but may be adapted as required.

3 *Interpreting and evaluating data (up to 2 hours, formal supervision)*

This stage is carried out under formal supervision in the classroom so that you are sure the student's interpretation is their own.

This stage is not an exam and requires supervision rather than invigilation. There is no need to set up the room like an exam or for the room to be silent. The key requirement is that students are supervised at all times.

Teachers now mark stages 2 and 3 and complete the student record form.

Controlled assessment support materials

For each of the tasks produced each year, Edexcel will publish a support booklet for teachers and a support booklet for students to ensure you are clear about the requirements and students perform as well as they can in the assessment. The support booklets for the sample controlled assessment materials are available in the sample assessment document. Here is a general booklet which you can adapt to use for practice tasks if you have certain ones you want to use.

Teacher's booklet

Students will need guidance in choosing the questions they are going to pose. Questions will need to enable students to use statistical techniques based on the specification content of the GCSE tier at which they are entered. They will also need advice about data collection.

The controlled assessment will be done in three stages.

Stage 1: Planning

Formal supervision (classroom or ICT suite) 10 marks, up to 2 hours

This should take place under formal supervision.

Students must be supervised at all times but this does not take place under examination conditions and students should be encouraged to discuss their planning with the teacher.

Whilst word processors may be useful, internet access is not allowed.

You should discuss the task and how it can be approached with students.

Discuss the following with students.

- What sort of questions can be asked?
- Draw out of students the questions they might consider.
- This will determine the sort of data they might collect.
- What data will be needed? This should include discussion of the population and how much data will be needed. (Generally, at least 30 pieces required but at least 60 are needed for more complex work.)
- How could the data be collected?
- What data collection sheets will be needed? Students may collect data in groups but they will need to give an explicit account in their plan of how they will do this.
- What statistical techniques can be used?

Whilst the use of techniques must be an integral part of the project, students will need some guidance as to what is expected at each level.

It should be pointed out that choosing techniques with reasons is essential if students are to achieve marks for planning. Students should be encouraged to be selective.

Some statistical techniques are more sophisticated than others. Students should choose techniques that suit their ability.

More able students should be advised that the complexity of their investigation will be limited if their hypotheses and choice of techniques do not allow them to produce work of a reasonable degree of complexity.

Students should produce a written plan at the end of this highly controlled stage.

You should review this and have given feedback to the student before any further work is done.

Feedback should be recorded on the student record sheet.

At this point students can amend their plans in the light of your comments.

You should advise regarding the suitability of student choices in your feedback at the end of Stage 1.

Choosing too trivial a project will not allow a student to reach their potential.

Here are some suggested questions and related analyses.

- Students targeting lower grades (G–E) might consider collecting qualitative data. Questions such as 'What is the most common X?' or 'What is the mean number of x?' or 'Are there more x than y?'. This would allow sampling from a single class (G–F) or sampling from several classes (F–E). Totals could be worked out and bar charts or pie charts drawn. There is scope for averages and simple comparisons.
- Students targeting middle grades (E–B) might look at associations between measurements or variables. Such questions as 'Is there a relationship between x and y?' or 'Does your ability to x depend on y?'
- This level of work could involve two measurements. This should enable stratified sampling to be done.
- Any of the following could be used — scatter diagrams, cumulative frequencies, box plots, averages, quartiles, comparisons, Spearman's rank correlation etc.
- Students targeting higher grades (B–A*) need to produce more individual work. They need to develop further some of the ideas for the middle ability students to include such questions as 'Can x be modelled by a normal distribution?' or 'How does the spread of x for girls compare with the spread of x for boys?
- These students need to pose questions that give them the scope to cover the techniques suggested for middle ability students plus the use of statistical techniques such as histograms, standard deviation or normality.

Stage 2: Collecting, processing and representing data

Limited supervision

Students in their own time, under informal supervision, need to complete the following.

Data collection (8 marks)

This must have been planned in Stage 1 and data collection sheets prepared.

Students may want to collect data in groups, but their individual contributions should be clear in their plan. A group could visit a class and collect all the data needed for the various hypotheses to be tackled. A spreadsheet could be created.

Process and present data (12 marks)

This means that students need to draw relevant diagrams and carry out relevant calculations. The use of ICT is encouraged but not essential. Evidence of calculations or hand-drawn diagrams are not needed as these skills are assessed in the external assessments. It is important that ICT is used sensibly and diagrams have sensible scales and labels.

Stage 3: Interpret and evaluate

Formal supervision (classroom or IT suite) (10 marks), up to 2 hours

Students may use the internet and bring in work from outside the classroom but you will need to monitor the work in the classroom to ensure it relates to the initial plan. The initial plan may have been developed and adapted, and this is to be encouraged, but only following discussion with the teacher, with feedback and amendments noted. This is to ensure that the project is the student's work.

Students bring all the work they have compiled to the classroom and put together their whole report. They will need to produce the following in final form.

- The written-up hypothesis/es with their planned strategy.
- The data collection discussion describing exactly what they did. Raw data should be in an appendix with summary tables in the main body of the report. Problems and limitations of the data should be discussed.
- Reasons for choice should accompany the analysis. Diagrams and calculations should be interpreted.
- There should be an interpretation in the context of the whole investigation relating back to the original questions and hypotheses.
- Conclusions linking together the strands of their enquiry.
- Evaluation of the work discussing any limitations.
- A completed, student record form should accompany the work alongside the authentication sheet signed by you and the student.

Work can be handwritten or word processed.

When completed the work is handed in. This must be at the end of the highly controlled time – there must be no extension for finishing at home.

You will need to complete the student record form, which should be attached to the front of the project with your marks and any other information.

Quality of Written Communication

Controlled assessments in GCSE Statistics provide opportunities across ability ranges to assess Quality of Written Communication (QWC).

Each stage of the controlled assessment offers opportunities where:

(i) ***text is legible***
spelling, punctuation and grammar are accurate
meaning is clear

(ii) ***students select and use a form and style of writing appropriate to purpose and to complex subject matter***

(iii) ***students organise information clearly and coherently, using specialist vocabulary when appropriate.***

Throughout the controlled assessment there are opportunities to assess strand (i) of QWC, and to ensure students are using clear and legible writing and checking their punctuation, grammar and spelling throughout their work.

Strand (ii) of QWC can be assessed differently through the stages. In Stages 2 and 3, for example, students may find it is more appropriate to use diagrammatic or tabular representations of information, with short sentences linking the different considerations within that stage. However, in Stages 1 and 4 an extended writing structure would be appropriate.

Strand (iii) of QWC can be assessed throughout the stages of the controlled assessment, where students can be expected to express themselves logically and clearly, using appropriate technical language and notation, and the overall investigation should be organised clearly and coherently.

Where students are required to provide clear aims, strategies, lines of enquiry, explanations, justifications or reasons for their work are all opportunities to assess QWC (ii) and QWC (iii).

In the assessment criteria, specific indicators where QWC (ii) and QWC (iii) can be assessed have been included. QWC (i) can be assessed throughout the controlled assessment.

Student booklet

The project must be carried out in three stages.

The **first stage** must be carried out under formal supervision with your teacher present and you cannot take this work out of the classroom until it has been completed. You will plan your investigation in this stage.

In the **second stage** you can work under informal supervision outside the classroom. In this stage you will collect your data, draw diagrams and carry out calculations.

In the **third stage** you must work under formal supervision again with your teacher present. You will write up your investigation. This will involve analysing and interpreting your data, drawing conclusions and evaluating your work.

Quality of Written Communication will be taken into account throughout the marking of your investigation and includes clarity of expression, the structure and presentation of ideas and grammar, punctuation and spelling.

Stage 1: Planning

You need to plan how you will carry out your investigation.

What are you going to investigate? Discuss this with your teacher.

Your teacher will give you some ideas but the highest marks are for originality – that means your own ideas.

Develop a question or hypothesis to investigate.

You can have just one question or hypothesis, but it would be better if you considered related questions or hypotheses.

Remember

- The questions you investigate will influence the sample you collect.
- You need to think about what you expect the answers to your questions or hypotheses to be and why you expect these answers.

Data collection

Think about the data you want to collect.

- Where will your data come from?
- How reliable is your source?
- How big a sample will you need? Discuss a suitable sample size with your teacher. The sample size will depend on the questions you are trying to answer but you should usually collect at least 30 pieces of data for each variable you are investigating.
- How are you going to collect the data? What sampling method will you use?
- You can work with others to share the data collection but you must write down exactly what you did and what they were asked to do.
- How are you going to record the data you collect?
- What will you do about outliers or anomalies? (Pieces of data that appear to be out of place or incorrect.)

Diagrams

- Which diagrams would give the most information about your questions and what are the reasons for your choice?
- You should always choose the best diagram – there is no need to draw several diagrams to represent the same data.

Calculations

- Which calculations will best summarise your data and answer the questions you have asked?
- Remember to consider outliers.
- You should **not** carry out calculations that do not answer your initial questions. You will get no extra credit and will waste your time.

Now write up your plan and give it in to your teacher.

Your plan should state the questions you are going to investigate and information on your plans for data collection, diagrams and calculations. You should write down reasons for your choices of all these.

Your teacher will discuss your plan with you. If you wish, at this point, you will be able to adapt your plan, taking into account your teacher's comments.

Stage 2: Collecting, processing and representing data

Now you have completed the first stage of your work you need to do some work by yourself.

You need to carry out the following.

- Collect your data as you have planned in Stage 1, using any collection sheets you have designed, and keep a record of the raw data.
- Write down any problems that arose.
- Put your data into the appropriate form.
- Carry out your calculations and produce your diagrams neatly, by hand or using IT.
- Collect together everything you need to complete your report.

Stage 3: Interpreting and evaluating data

The final writing of the rest of your report will be in the classroom.

You will already have completed most of your diagrams and calculations, but now everything needs to be put together so that it makes a complete report.

You need to carry out the following.

- Place your **Plan** at the beginning of your report.
- Under the title **Data Collection**, present the data you have collected in tables.
- Write down any extra information about the data that was not in your plan. This might include problems that you found and how you resolved them.
- Under the title **Analysis,** present your diagrams and calculations in a form which helps you to interpret them.
- Draw everything together in a **Conclusion** that relates to the questions and hypotheses you posed in your plan.
- **Evaluate** your work. Did you get the results you expected? Suggest improvements and discuss limitations.
- Create an **Appendix** for your raw data.
- Complete the student record sheet, sign the authentication form and put them at the front of your report.
- Number the pages and fix your report together with a treasury tag or piece of string.
- Hand in your report to your teacher.

Controlled assessment exemplars

There are two exemplars, both based on the estimation task. There is:

- an average piece of work from a student, followed by the moderator's comments as to how it could be improved
- an outline of an excellent piece of work which would achieve the higher grades.

Using these exemplars with students will help them understand the requirements and give them an insight into how to be successful in their own investigation.

You will find the task and support booklets in the sample assessment materials.

Exemplar 1: Sarah's work

Stage 1: Planning

In this coursework I am looking at people's skill in estimating. I am going to get a number of students to estimate the length of a line on A4 paper. I want to find out if girls are better at estimating than boys.

I will draw a line on a large piece of paper and ask the teacher to ask the students in year 10 classes to estimate its length. I will ask the teacher to give each student a slip of paper blue for boys and pink for girls. They will be asked to estimate the length of the line working on their own and not to compare their result with anyone else. It would be better if I could watch them estimate to ensure there was no cheating. The data will be continuous quantitative data.

I think boys will be better at estimating than girls.

I will take a random sample of 25 boys and 25 girls from the estimates which are collected.

I am going to use Box plots to compare the estimates as this will allow me to compare the Median and the spread of the data.

Moderator's comments

Expectations are given.

The student has chosen to use box plots to make a simple comparison between boys and girls and given an outline of how they will collect their data. The student is planning to collect an adequate amount of data.

She has foreseen that students may cheat but has not considered what they will do with anomalies.

This is an example of planning work produced by an average student and would achieve around 6 marks.

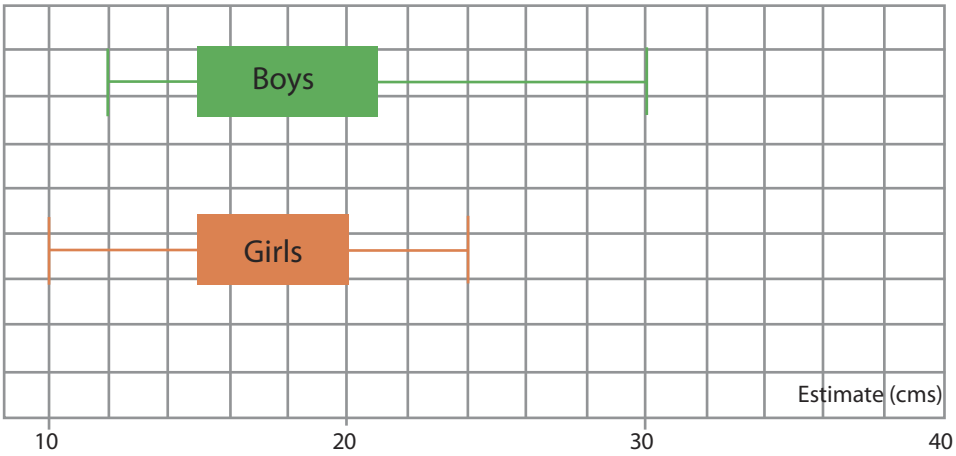
Stage 2: Collecting

I took a random sample of 25 boys and 25 girls by mixing up all of the slips and drawing them out of a box.

Moderator's comments

The student uses a recognised sampling method and explains where the data has come from. There are two data sets with sufficient data to make comparisons. This is worth around 4 marks.

Stage 2: Processing and representing data



	Boys	Girls
Number in sample, n	27	27
Mean, x	18.9259	17.4444
Standard deviation, x	4.33697	3.35916
Range, x	18	14
Lower quartile	15	15
Median	20	17
Upper quartile	21	20
Semi I.Q. Range	3	2.5

Moderator's comments

The student has entered their data into autograph which has calculated the median, quartiles and maximum and minimum points. Two box plots are drawn to compare the boys' and girls' estimates. They have units. The statistics quoted below the box plot have been taken from a computer program.

Outliers have not been considered and the actual length of the line is not shown.

This is worth 6 marks.

Stage 3: Analysing and evaluating

I said that boys would have a more accurate result than girls in estimating the line.

The length of the line was 17 cm so the median of 17 cm for the girls was exactly right. The girls also had a smaller inter-quartile range and range showing that their results were more consistent.

The median for the boys was 20 cm which was too high showing that more than half of the boys overestimated the length. 'Both the IQR and the range are bigger than the girls showing that more of them estimated accurately though all estimated better than the worst estimator from the girls.

Moderator comments

The student has summarised their results and related their statistical diagrams to the context of the investigation.

It could be improved by consideration of outliers and why the maximum value for boys is so large. They have compared a measure of central tendency (the median) and the spread. This is worth 6 marks.

The statistics box they have copied from the computer program has unused values (the mean, standard deviation and semi-interquartile range) which can be ignored at this level but would detract from the work of an A/A* student.

Exemplar 2: Outline for a top grade controlled assessment investigation

Stage 1: Planning

The student will carry out the following.

- Give an introduction planning to estimate people's ability to estimate the length of a line.
- Set up related hypotheses:
  - ❑ girls will estimate the length of a line better than boys
  - ❑ students who can estimate a drawn line accurately will be more accurate in drawing a line of a given length
  - ❑ students will find it easier to estimate the length of a curved line rather than a straight line.
- Create a questionnaire with a straight line, curved line and a line to be drawn with space for the subject to write their estimate along with their name and gender.
- Plan to pilot their questionnaire in their own class.
- Refine the questionnaire following the pilot.
- Use the pilot to set limits for outliers.
- Plan to take a sample stratified by year group of 50 boys and 50 girls across the school. They justify their choice and explain explicitly how students will be selected randomly.
- Explain that the data is continuous but that it is unlikely students will give more than 1 decimal place in their estimate.
- Explain how the data (which may be collected by a group) will be collated and shared.
- Plan to compare using box plots but also showing the actual length of the line.
- Plan to formally calculate outliers and to mark these on their box plots.
- Plan to draw scatter graphs comparing the estimates of lengths of given lines with those drawn by the student.
- Plan to use Spearman's rank correlation if the scatter graph shows correlation.
- Explain what they expect Spearman's rank correlation to show.
- Plan to draw histograms to compare with a normal curve.

Moderator's comments

A clearly written plan which contains the above elements is typical of a top grade student.

The student has planned to consider a number of related variables using a variety of statistical techniques some of which are above A grade.

The data collection and questionnaire may be planned collectively within a small group or with the class. A database of the estimates for classes across the school may be collated and made available for all students to select from. (The database may include other estimations for example weights, angles, capacity.)

The hypotheses, sampling decisions and choice of statistical techniques must come from the student.

This would be worth 9 marks.

To achieve 10 the student would need to:

- highlight problems of students not taking the activity seriously or cheating and explain how this will be overcome
- plan a formal comparison with the normal distribution
- plan to use percentage errors rather than raw estimates with reasons
- plan to use the box plot to help decide the class intervals for the histogram.

Stage 2: Collecting

The student sets the limits for outliers and collects the data as planned.

Problems met during data collection for example missing or completely erroneous data are discussed and acted on.

All outliers are discussed.

Samples are included in the work.

Moderator's comments

The student uses a recognised sampling method and explains where the data has come from. There is sufficient data to make comparisons.

They have used the boundaries they set at the planning stage to identify outliers in their data but have chosen to include them.

They have discussed any problems which have arisen (not ignored them).

Adequate data has been collected. This is worth 7 marks.

To achieve 8 marks:

- the student could discuss ways in which bias could have arisen within their data collection and how they avoided this
- they should consider the proportion of outliers found and whether these were more prevalent in one stratum. They may use this as evidence in considering the reliability of their data collection
- they act on the plans they made to overcome problems, commenting on their effectiveness.

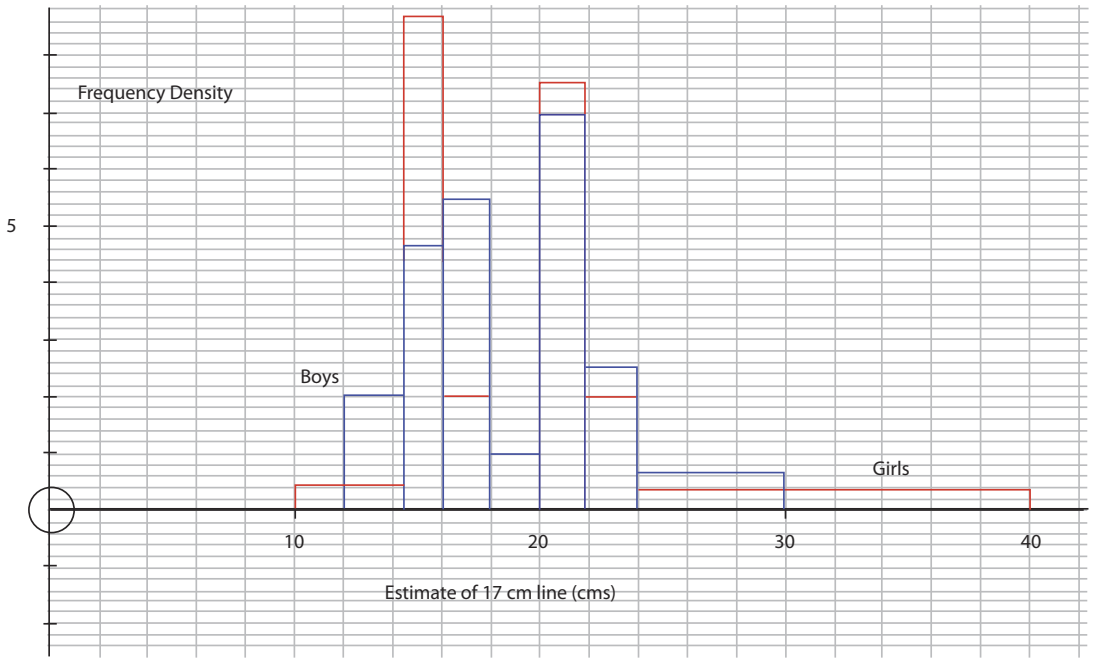
Stage 2: Processing and representing data

The student:

- produces box and plots for boys' and girls' estimates
- formally calculates outliers using Q + 1.5 x IQR
- marks outliers.

All diagrams are drawn for sensible-sized data sets.

All diagrams are drawn for ease of comparison.



Notices that box plot and histogram are not regular so comments that it is not sensible to draw a normal curve.

Draws scatter graphs and decides to calculate Spearman's rank correlation.

Uses a spreadsheet to help in Spearman's rank correlation calculation.

Moderator's comments

The student has worked intelligently and efficiently, producing only the diagrams and calculations mentioned in the plan.

They have used sensible scales

The student considered the effects of gender and the relationship between drawing a line of a given length and estimating a line which has already been drawn, using a number of different techniques to explore possible connections or effects. They have drawn box plots, histograms and scatter graphs and calculated averages, spread and Spearman's rank correlation to investigate their hypotheses.

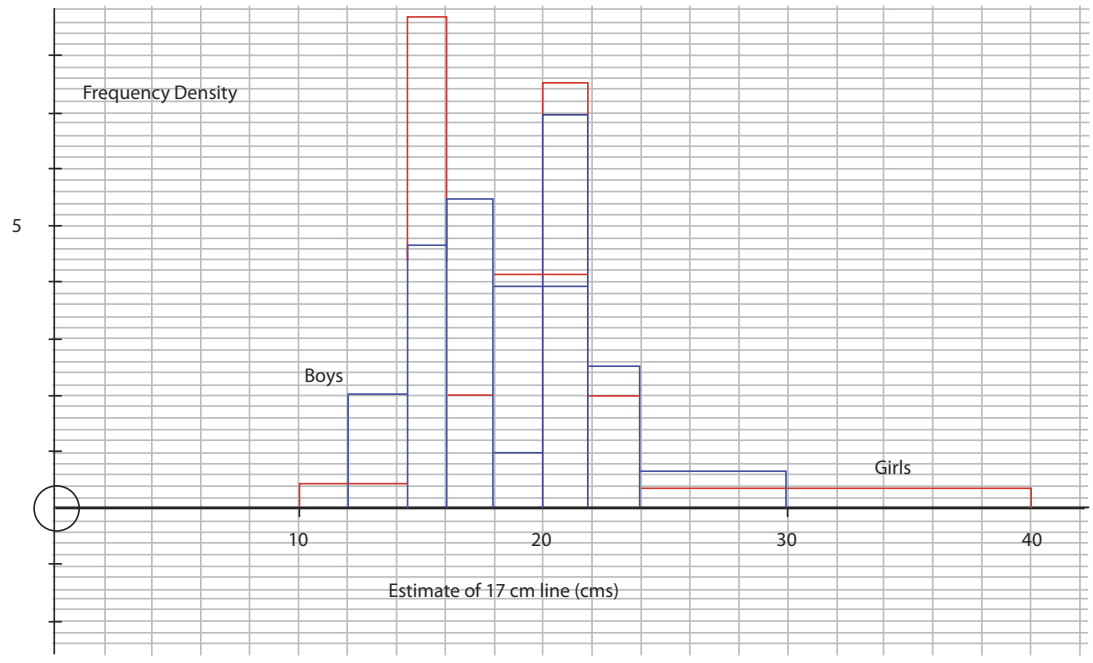
They have calculated outliers and marked these on the box plots and scatter diagrams.

There are no misconceptions and no serious errors in the work.

This is worth 11 of the 12 marks.

To achieve 12 the student could:

- use the box plot to decide that grouping the data in larger class intervals would be sensible at the extremes (upper and lower quartiles) but smaller ones would be needed in the middle where the data is more close together
- adjust the class intervals on an irregular histogram and redraw to consider normality



- on finding a more symmetric distribution, make a formal comparison with the normal distribution using the mean and standard deviation.

Stage 3: Analysing and evaluating data

The student gives a complete summary of their investigation quantifying and interpreting results in the real world.

They discuss each hypothesis in turn but ensure links occur between them.

Differences between the accuracy in estimation of boys and girls and the different lines are discussed. Statistical evidence is produced from their diagrams or calculations. This is interpreted in depth.

The effect of outliers is considered.

Every diagram and calculation is used and commented on.

The student evaluates their work and discusses any problems in data collection, choice of diagrams, scale or calculations.

Moderator's comments

The student summarises their strategy.

Sampling is discussed and the effect/number of anomalies commented on.

The student has used a range of diagrams: box plots, histograms and scatter diagrams in a sophisticated manner, and discussed all of them in their interpretation.

They consider both averages and spread.

They have used Spearman's rank correlation to measure the correlation and relate different factors.

This is worth 9 marks.

To achieve 10 requires a well-written, concise discussion with added depth which may include the following.

- Justification of unequal class intervals for example *using their box plots.*
- *Consideration of possible reasons why the data is bimodal for example when asked to estimate some people will round off to a whole number whilst those who are less confident may round to the nearest ten number giving a modal value at 20.*
- Grouping the class intervals differently to see if estimating overall showed a normal curve as expected.
- A formal exploration of the normal distribution using mean and standard deviation.

2019db130412S:LT\PD\GCSE 2012\TSMs\Edexcel GCSE 2012 Statistics TSM Issue 2.indd-48/2

For further copies of this publication, please send an email to the following addresses:

UK customers: publication.orders@edexcel.com
International customers: intpublication.orders@edexcel.com

Also, you can download copies at: www.edexcel.com

For more information on Edexcel and BTEC qualifications please visit our website: www.edexcel.com

Edexcel Limited. Registered in England and Wales No. 4496750
Registered Office: Edinburgh Gate, Harlow, Essex CM20 2JE
VAT Reg No GB 278 537121

Publication code UG031457