



**General Certificate of Secondary Education  
June 2011**

**Statistics**

**43102**

**(Specification 4310)**

**Unit 2: Controlled Assessment**

***Report on the Examination***

---

Further copies of this Report on 43102 are available from: [aqa.org.uk](http://aqa.org.uk)

Copyright © 2011 AQA and its licensors. All rights reserved.

**Copyright**

AQA retains the copyright on all its publications. However, registered centres for AQA are permitted to copy material from this booklet for their own internal use, with the following important exception: AQA cannot give permission to centres to photocopy any material that is acknowledged to a third party even for internal use within the centre.

Set and published by the Assessment and Qualifications Alliance.

The Assessment and Qualifications Alliance (AQA) is a company limited by guarantee registered in England and Wales (company number 3644723) and a registered charity (registered charity number 1073334).  
Registered address: AQA, Devas Street, Manchester M15 6EX.

## General

This year there were fewer issues with documentation and the teething problems of last year did not recur. Almost all centres managed to find and complete the necessary documentation. The majority of centres were able to send marks in on time or early, although work did occasionally then follow after a delay rather than within the week expected.

It is noticeable that those centres who did attend one of the teacher standardising meetings were less likely to need mark adjustments. Again, those who annotate and provide clear evidence of how marks have been awarded do make it easier for a moderator to agree to the marks awarded.

The 'Houses' task proved to be far more popular than the 'Football' task, and was generally better done. This appeared not to be due to the task itself.

As last year, there was no overall difference between the marks gained by candidates in the two parts of the controlled assessment. Centres are to be congratulated on the variety of work produced by their candidates, and only a small number of centres had candidates who produced similar pieces of work.

## Houses Investigational Task

It was rare to see 'The small data set' used alone.

The websites provided more than sufficient data for the students to identify appropriate populations from which to sample. Those who chose to use their own websites, such as Zoopla, which gave the candidate mean values, often missed the fact that the mean values may well be for houses sold over different time periods, making comparisons difficult, if not impossible, in some cases.

Some candidates were able to list a large number of house prices from the websites and used this to select from. Those who obtained one value at a time from the websites clearly required more time to do this. A long period of data collection usually means less time available for analysis.

Some candidates used very small samples which not only impinges on the mark in strand 2, but can also affect the mark awarded in strand 3, as it is not appropriate to use particular techniques, such as cumulative frequency graphs, with small quantities of data. Some candidates failed to appreciate there were different types of house that could be chosen, or different sizes within these categories.

Some candidates appeared to choose two locations which they felt would give them large differences in house prices. These candidates did not gain by doing this, and usually failed to capitalise on the large differences which resulted when interpreting their results and made only weak comments. They also lost out on the opportunities for discussion of results which often arose from using better reasons for selection of location.

Knowledge of the assumptions being made when choosing variables was tested in the written assessment, and it was clear from the responses seen that many candidates appeared not be aware of the assumptions they were making at the start of their work. Consideration of these often allowed candidates to develop a strategy with which the task itself could be approached.

Interpretation was often seen to be a weak area, with interpretation of diagrams being generally slightly better than interpretation of calculations.

### Strand 1 Hypothesis and strategy (0-2)

The vast majority of candidates were able to state a hypothesis. Plans varied in detail and depth, but those just listing the methods they were going to use were restricted to 1 mark. Sometimes the depth of planning was linked to the selection and collection of data, showing some awareness of how strata were to be identified, and why, with a link to the reasons for their choice of variable(s). Another way candidates could develop their planning was to consider what it is they are attempting to analyse, indicating what variables were involved and then limiting the effects of selected extraneous variables. For example, a candidate who attempted to compare house prices in the north and south may have decided that the size or nature of the 'towns' where the houses were located may have affected the price of housing, and so the towns selected were chosen so that they were comparable. This type of decision does, however, need to be explicit.

**Strand 2** Selection and Collection of Data (0-4)

Many candidates had no problems in gaining 2 marks in this strand, having collected some data, identified the source and indicated that sampling had taken place. In order to gain 3 or 4 marks, considerably more work in the way of sampling is required, and it was noted that a number of centres were slightly generous in their interpretation of the assessment criteria here. The sampling method was not always named and was sometimes incorrectly named. Details of sampling were often weak, using software to select a random sample is not appropriate nor is using a website that selects random samples as neither provides sufficient detail. Samples were sometimes far too small, sometimes smaller than the small data set offered for weaker candidates. Justification also was weak and candidates would be well advised to consider the population when deciding on an appropriate sampling method to use, as this will help provide the necessary, appropriate reason for their choice. To get a 'fair' sample, or to get 'representation' are not sufficient on their own to justify their choice of sampling method.

There were some good, clear attempts at sampling made by some candidates, including multistage sampling, and stratification of some kind. It was disappointing to see some candidates merely select the highest prices from two areas rather than attempting to find a sample of prices representing the populations.

Some centres appeared not to have noticed the bullet points on the Assessment Criteria and elaboration document which are there to help in the decision making process when awarding marks.

**Strand 3** Use of Graphs and Diagrams to test hypotheses (0-7)

A good variety of graphs and diagrams were drawn by many candidates. There are three keywords in the assessment criterion for this strand. They are 'appropriate', 'accurate', and 'use'. The diagrams need to be appropriate for the type of variable/data being represented. They need to be accurate. This will usually require the use of graph paper, an obvious exception being pie charts. Diagrams such as histograms will usually be better for being drawn using a ruler. Diagrams also need to be of a size that can be used. 'Use' means interpretation, and here it should be within the context of the situation. Merely commenting on skewness shown in a boxplot is not an interpretation and on its own could not gain more than 2 marks.

Some candidates confused bar charts and histograms. Bar charts should be drawn to show frequencies for separate values with each rectangle labelled, and not for showing mean values as was sometimes seen. Histograms should be drawn for grouped, continuous data, and should have a linear horizontal scale. Histograms with equal classes are a foundation tier diagram, and histograms with unequal classes are a higher tier diagram. Histograms, as last year, were not interpreted well, and generally contributed to a mark of 5 at most.

This year saw more candidates drawing cumulative frequency graphs and boxplots than last year. Centres are reminded that provided these are accurate, appropriate and interpreted well, they can be awarded a maximum of 4 marks.

There are various ways candidates may move into the 5-7 mark box. Having already gained 3-4 marks for appropriate diagrams being accurately drawn and interpreted, the candidate may have included higher tier diagrams as one of two different types, in which case a mark of 5 may be awarded. Alternatively, a candidate may have already gained 3-4 marks for appropriate diagrams being accurately drawn and interpreted; and should the diagrams be of different types and the interpretation sophisticated, then conclusions about the validity of the hypothesis may be drawn, and again a mark of 5 may be awarded. Without interpretation, diagrams cannot contribute to a mark above 2, regardless of whether they are higher tier or foundation tier diagrams.

The best candidates often used a series of boxplots to compare the different strata from their sample.

**Strand 4** Use of calculations to test hypotheses (0-7)

Simple calculations such as percentage calculations can often be useful for making basic comparisons, and so can be awarded up to 2 marks. This also includes pie chart calculations. Calculations of simple statistics with brief comments such as ‘the mean house price for area X is higher than the mean house price for area Y’, may also be awarded up to 2 marks. As with strand 3 the key words are ‘appropriate’, ‘accurate’ and ‘use’. Many candidates were able to score reasonably well due to accurate and appropriate calculations, although poor interpretation often prevented higher marks being awarded. Not all candidates appreciated the difference between a measure of average and a measure of spread, sometimes commenting that a higher range/standard deviation/IQR meant that the values are generally higher or that ‘my mean and standard deviation show house prices are higher for area X’ without any explanation as to how they do this.

It was often noticeable that candidates who interpreted what they produced immediately after producing it, often experienced more success than those who did their calculations and diagrams one after another, then writing their interpretation all together at the end. Mainly this was due to the ease with which they could refer to what was being interpreted. Those who did attempt interpretation only at the end often referred to the ‘results showing ...’ or ‘this shows ...’ or ‘it shows ...’ giving no clues as to what the ‘it’ actually was. On its own this cannot count as interpretation, but only as ‘little attempt to interpret’ contributing towards a mark of 1 or 2 in strands 3 and/or 4.

There was a strong feeling, from some candidates, that a list of calculations should be given regardless of whether they were appropriate or to be used. These often appear to be following use of a software package. This type of work will only contribute to a mark of 1 or 2 unless accurate, appropriate, and interpreted in context.

**Houses Written Assessment**

It is important that the investigational task is available for candidates when they do the written assessment. Some referred to work they had not done and stated assumptions they clearly had not made.

Likewise, it is important that the assessor at the centre has access to the investigational task when marking the written assessment, to ensure that work referred to has actually been done, and to check that assumptions stated were actually made and that variables identified as extending the work have not already been pursued within the work.

This did not always appear to be the case.

**Question 1**

The vast majority managed to correctly name a diagram they had used. Many were also able to describe what it had been used for. Many candidates then failed to state a valid reason why the diagram was appropriate or inappropriate for their data. Though some did name histograms as being appropriate for their continuous data, or bar charts for their categorical data, the majority repeated their answer to 1(b). Unfortunately, many of these had been marked as correct by centres.

**Question 2**

This question was quite well answered. The first calculation offered was usually a measure of average, generally the mean, with a valid reason gaining 1 mark. The second calculation offered was often a different measure of location, or a measure of spread without a valid reason. Reasons were sometimes unclear.

This question was generally marked well by centres.

### **Question 3**

Candidates need to understand the assumptions they make when they carry out a statistical study. It was evident from many of the responses that this was not the case as many candidates repeated their hypothesis. The question asked specifically for assumptions that had been made when choosing the variables, so assumptions about the website being up to date or having correct data were not assumptions made when choosing the variable. Having selected their variables at the outset of the work, candidates may have then decided not to consider whether properties had large or small gardens, or some other variable that may have had an effect on house prices, assuming that their effects on prices would be small, or they may have assumed that one particular type of house best represented house prices for each area.

This question was generally marked well by centres.

### **Question 4**

Whilst many candidates did gain credit here, some incorrectly thought that more data or more calculations or diagrams would extend or develop the work. Whilst more data may improve their work and the validity of their conclusions, possibly due to too small a sample in the first place, it would not actually move the task on. To do this would usually require a new variable to be introduced. Some candidates successfully made a link between questions 3 and 4, and used their assumptions to select new variables, for example, having assumed that size of garden did not affect the price of a house in the task and stated this in question 3, they could then select this as a possible new variable to investigate to see if it did affect house prices. Others identified the number of bedrooms, or whether the property had a garage.

### **Question 5**

Those who read the question carefully or who had gained from working on a practice piece of work generally scored well here.

This question was also very accurately assessed by centres.

Some candidates turned the question around and treated it as an investigation involving secondary data with data coming from the internet. Some decided to investigate the type of house people actually live in. Both of these errors limited the number of marks available to candidates.

Most candidates had a hypothesis. Identifying a variable was sometimes done within the hypothesis, although many candidates failed to make a full statement recognising different types of housing.

Many gained 1 mark for identifying a sampling method they would use, quite a few then went on to contradict themselves and stated a different method. Some identified the problem as being to do with the sampling method, so changed their minds.

Many valid problems were identified, the most common of which was poor response rate for the survey/questionnaire they were going to use.

### **Question 6**

This question was answered poorly by some candidates. Scatter diagram was identified more frequently than a correlation coefficient.

## Football Investigational Task

It was rare to see 'The small data set' used alone.

The websites provided more than sufficient data for the students to identify appropriate populations from which to sample.

Some candidates used very small samples which not only impinges on the mark in strand 2, but can also affect the mark awarded in strand 3, as it is not appropriate to use particular techniques, such as cumulative frequency graphs, with small quantities of data. Interpretation was often seen to be a weak area, with interpretation of diagrams being generally slightly better than interpretation of calculations. It was clear that many candidates did not identify the population from which they were going to sample before they started the work. This is an important part of any statistical study and should be done in the first open discussion lesson. It will help inform the hypothesis and also help candidates decide on the sampling method to be used, or whether sampling is appropriate at all. In this case candidates were also expected to be able to identify the population in the written assessment.

### **Strand 1** Hypothesis and strategy (0-2)

The vast majority of candidates were able to state a hypothesis. Plans varied in detail and depth, and those who just listed the methods they were going to use were restricted to 1 mark. Sometimes candidates developed their planning by considering what it was they were attempting to analyse, indicating what variables were involved and then limiting the effects of selected extraneous variables. This does need to be explicit in the work.

Most candidates chose to pursue number of goals scored, number of matches won or the number of times countries reached the final stages of the competition over a number of years.

### **Strand 2** Selection and Collection of Data (0-4)

Many candidates had no problems in gaining 2 marks in this strand, having collected some data, identified the source and indicated that sampling had taken place. In order to gain 3 or 4 marks considerably more work in the way of sampling is required, and it was noted that a number of centres were slightly generous in their interpretation of the assessment criteria. The sampling method was not always named and was sometimes incorrectly named. Details of sampling were often weak, using software to select a random sample is not appropriate nor is using a website that selects random samples as neither provides sufficient detail. Samples were sometimes far too small, sometimes smaller than the small data set offered for weaker candidates. Justification also was weak.

Candidates would be advised to consider the population when deciding on an appropriate sampling method to use. This will help provide the necessary, appropriate reason for their choice. To get a 'fair' sample, or to get 'representation' are not sufficient on their own to justify their choice of sampling method.

The majority of candidates decided that their population was the countries with teams which qualified for the World Cup. This immediately reduced the amount of data available to them, so that there was little to sample from. The small data set included data from the qualifying matches, which offered a reasonable amount of data from which to sample.

Some centres appeared not to have noticed the bullet points on the Assessment Criteria and elaboration document which are there to help in the decision making process when awarding marks.

### **Strand 3** Use of Graphs and Diagrams to test hypotheses (0-7)

There are three keywords in the assessment criterion for this strand. They are 'appropriate', 'accurate', and 'use'. The diagrams need to be appropriate for the type of variable/data being represented. They need to be accurate. This will usually require the use of graph paper, an obvious exception being pie charts. Diagrams such as histograms will usually be better for being drawn using a ruler. Diagrams also need to be of a size that can be used. 'Use' means interpretation, and here it should be within the context of the situation. Merely commenting on skewness shown in a boxplot or the correlation shown in a scatter diagram, is not an interpretation and on its own could not gain more than 2 marks.

There are various ways candidates may move into the 5-7 mark box. Having already gained 3-4 marks for appropriate diagrams being accurately drawn and interpreted, the candidate may have included higher tier diagrams as one of two different types, in which case, a mark of 5 may be awarded. Alternatively, a candidate may have already gained 3-4 for appropriate diagrams being accurately drawn and interpreted; and should the diagrams be of different types and the interpretation sophisticated, then conclusions about the validity of the hypothesis may be drawn, and again a mark of 5 may be awarded. Without interpretation, diagrams cannot contribute to a mark above 2, regardless of whether they are higher tier or foundation tier diagrams.

Some candidates attempted to put several sets of data on to one set of axes for a scatter diagram, and consequently found it difficult to interpret.

### **Strand 4** Use of calculations to test hypotheses (0-7)

Simple calculations such as percentage calculations can often be useful for making basic comparisons, and so can be awarded up to 2 marks. This also includes pie chart calculations. Calculations of simple statistics with brief comments such as 'the mean number of goals for team X is higher than the mean number of goals for team Y', may also be awarded up to 2 marks. As with strand 3 the key words are 'appropriate', 'accurate' and 'use'. Many candidates were able to score reasonably well due to accurate and appropriate calculations, although poor interpretation often prevented higher marks being awarded. Not all candidates appreciated the difference between a measure of average and a measure of spread, sometimes commenting that a higher range/standard deviation/IQR meant that the values are generally higher.

It was often noticeable that candidates who interpreted what they produced immediately after producing it, often experienced more success than those who did their calculations and diagrams one after another, then writing their interpretation all together at the end. Mainly this was due to the ease with which they could refer to what was being interpreted. Those who did attempt interpretation only at the end often referred to the 'results showing ....' or 'this shows ...' or 'it shows ...' giving no clues as to what the 'it' actually was. On its own this cannot count as interpretation, but only as 'little attempt to interpret' contributing towards a mark of 1 or 2 in strands 3 and/or 4.

There was a strong feeling, from some candidates, that a list of calculations should be given regardless of whether they were appropriate or to be used. These often appear to be following use of a software package. This type of work will only contribute to a mark of 1 or 2 unless accurate, appropriate, and interpreted in context.

### **Football Written Assessment**

It is important that the investigational task is available for candidates when they do the written assessment. Some referred to work they had not done and stated assumptions they clearly had not made.

Likewise it is important that the assessor at the centre has access to the investigational task when marking the written assessment, to ensure that work referred to has actually been done, and to check that assumptions stated were actually made and that variables identified as extending the work have not already been pursued within the work.

This did not always appear to be the case.



### **Question 1**

The vast majority managed to correctly name a diagram they had used. Many were also able to describe what it had been used for. Many candidates then failed to state a valid reason why the diagram was appropriate or inappropriate for their data. Though some did name histograms as being appropriate for their continuous data, or bar charts for their categorical data, the majority repeated their answer to 1(b) and, unfortunately, many of these had been marked as correct by centres.

### **Question 2**

Quite well answered by many. The first calculation offered was usually a measure of average, generally the mean, with a valid reason gaining 1 mark. The second calculation offered was often a different measure of location, or a measure of spread without a valid reason. Reasons were sometimes unclear.

This question was generally marked well by centres.

### **Question 3**

Candidates need to appreciate the importance of identifying the population from which they sample before sampling takes place. It was clear from the responses that this was not the case for all students.

Those who stated that their sample was not representative of the population usually gave a valid reason and identified the population correctly so gained 2 marks. Those who thought their sample was representative often gave a reason suggesting that this was evident as their hypothesis had been shown to be correct.

### **Question 4**

Whilst many candidates did gain credit here some incorrectly thought that more data or more calculations or diagrams would extend or develop the work. Whilst more data may improve their work and the validity of their conclusions, possibly due to too small a sample in the first place, it would not actually move the task on. To do this would usually require a new variable to be introduced.

### **Question 5**

Those who read the question carefully or who had gained from working on a practice piece of work generally scored well here.

Some misunderstood what was being asked and neglected to notice that the age and gender were both required.

Most candidates had a hypothesis. Many gained 1 mark for identifying a sampling method they would use, quite a few then went on to contradict themselves and stated a different method. Some identified the problem as being to do with the sampling method, so changed their minds.

Some valid problems were identified, the most common of which was people may not wish to give their age. Some thought that not asking enough people was a problem they may have and their solution then was to ask more. This is not really a problem which is likely to be encountered, rather a problem of ones own making in the selection of sample size.

This question was very accurately assessed by centres.

### **Question 6**

This question was answered poorly by some candidates. Scatter diagram was identified more frequently than a correlation coefficient.

This question was very accurately assessed by centres.