



**General Certificate of Secondary Education
June 2011**

Statistics

43101H

(Specification 4310)

Unit 1: Written Paper (Higher)

Report on the Examination

Further copies of this Report on 43101H are available from: 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 was the first paper based on the new specification. There is a greater emphasis on data interpretation and inferential statistics rather than just the application of learned techniques. Candidates appeared to adapt well to this change with few failing to attempt all questions on the paper.

Questions on the new specification topics were not particularly well answered and there are examples of these in the specimen materials.

The presentation of answers was good as was the graphing of data.

Candidates still seem reluctant to make effective use of calculators and are unsure as to the levels of accuracy expected for an acceptable numerical answer.

Topics that were well attempted included:

- Histograms using frequency density
- Use of Venn diagrams and simple application of probability rules
- Regression analysis.

Topics that were not answered well included:

- Cumulative frequency step polygons
- Standardised scores
- Comparative pie charts
- Inter-observer bias
- Aspects of sample design.

Question 1

This question was generally well attempted although in part (a) a number thought that a census was necessarily the National Census and that a sample would be more representative and as such less prone to bias.

In (b)(i) many used the rental company or the number of complaints as suggested categories, whilst in (b)(ii) a number missed the point, writing instead about why stratification is used in sample design rather than why they would choose to use a particular stratification factor.

Part (c) was usually well answered although a number omitted the time frame in the question or suitable non overlapping response sections to record the answers.

Generally for part (d) answers were correct with appropriate explanations provided.

Question 2

Parts (a) and (b) were usually correct although there were odd numerical errors in arriving at the missing parts of the Venn diagram.

In part (c) a number gave 88 as the denominator and were consistent in using this in their answer to (c)(ii)

Many in answer to part (d) did not identify 14 as the correct % to apply instead, working with a figure of 32% (10 + 14 + 8)

Question 3

This question was not well attempted. Many misinterpreted inter-observer bias as meaning that an observer's opinion of a candidate was biased. This included comments on bias resulting from gender, race, age and other criteria.

As a result of confusion in part (a) many incorrect responses were seen in (b). Most suggested having one or an odd number of observers or averaging results.

A number in part (c) did not appear to know the meaning of the term "extraneous variable".

Question 4

Parts (a) and (b) were well answered although a few made comparisons based on a single data point, whilst a minority thought the data was over 8 consecutive days.

In part (c)(i) most recognised there were a higher number of 85+ year olds in Eastbourne compared to N. Lincs. Part (c)(ii) was not so well answered with a number stating it was because Eastbourne had a smaller population or simply that the average ages were not known.

Question 5

In part (a) most gained the method mark but then recorded their answer as 2006. Some missed the word "increase" and went for the year 2008 as having the highest number of births. Most realised in part (b) that there were more males than females but missed the constant ratio mark. A number incorrectly referenced the mode or stated "averages out to..."

There were many good and diverse answers to part (c)(i). For (c)(ii) a number suggested the mode or median would be better whilst others incorrectly commented on the range.

For part (d) a minority of candidates did not refer back to the table whilst others made vague reference to "births to foreigners".

In part (e) most scored the one mark for "cost/time" but the second mark was more elusive. Many wrote "easier"/ "more reliable" or attempted a definition of primary/secondary data.

Question 6

In part (a), many candidates approximated too early and thus lost the accuracy mark as 3.69 was an exact answer. In part (b) a number failed to appreciate the “squaring” part of the question and simply calculated 1.1×24.6 . A number were able to find 29.8 based on $(1.1)^2$ but were then unable to get any further. Most however, were restricted to one mark on this part by recognising that 126° should be used somewhere. Many were awarded full marks in answer to part (c).

Question 7

In part (a) most candidates were able to calculate the cumulative frequencies but then opted for a c.f. curve rather than a step polygon. Of those who did attempt the correct diagram most failed to complete the final step or did not relate the steps correctly in relation to their horizontal scale. Vertical plots were generally accurate however many candidates still remain unaware of the existence of a cf. step polygon.

For (b)(i) and (ii), a number lost marks as a result of an incorrect diagram in (a) or in part (ii) giving an answer based on “ $72 - 27 = 45$ ”

Part (c) (i) was generally correct but some used “23” as a numerator. Parts (ii) and (iii) were usually attempted with replacement and many in answer to part (iii) failed to use the “x 3”

Question 8

This question on Index Numbers was much better attempted than under the old specification. A significant number lost the accuracy mark for not rounding to 1dp in part (a). A few chose 2008 in answer to part (b) and only a minority were able to correctly show how the geometric mean had been estimated. Many in fact summed the indices and divided by 5 or calculated the mean of the 6 cost figures.

Question 9

There were many correct answers for part (a). Some did however make vague references to heights and area without meaningful explanation. In part (b) most were able to score 5 marks losing the one mark for incorrect or missing label. A number did, however, attempt to plot “ $ucl \times f$ ” or “ $midpoint \times f$ ”. A minority thought that the class widths were actually one unit bigger than given.

Part (c) was generally correct but few right answers were seen for part (d). Most found 63 but failed to realise they needed to take a proportion of the fourth class.

Some were able to score the first 3 marks in answer to part (e)(i) but then failed to give a convincing explanation as to whether it should be Tracy or Bill. A number based their calculations simply on mean/standard deviation. In answer to part (e)(ii) a number went down the route of $(8.3 - 13)/1.5$ but did not then understand the significance of 3 standard deviations. Similarly those who attempted to answer the question the other way around used 1 or 2 standard deviations rather than 3.

Question 10

Part (a), was generally well answered showing a good understanding of Spearman's formula. A small number did not know how to identify " n " and some missed a mark by failing to clearly evaluate the numerator and denominator separately. Very few made the error of $(1 - 6 \times 41.5)/\dots$. Part (b) was poorly attempted as almost all candidates referenced expenditure rather than the relative ranks.

Question 11

Part (a) was generally correct although some answered "2" for part (i) and stated lowest takings for (ii).

Part (b) was well attempted. There were, however, some errors in reading the scale when estimating " m " and when trying to find " c " by substitution. Some missed the " x " out of the final equation. Few candidates scored full marks in part(c). Many noted the car park being closed or a small sample but references to other factors were often very vague.

Question 12

In part (a) answers to **Y** were often too informal such as "don't care". Sometimes the headings given were the wrong way around. For part (b) most recognised the correct pairing but failed to realise they were contradictory because most people agreed with them. A number chose 1 and 3 which may have appeared contradictory but the responses showed they were not.

Few understood the concept of multistage sampling in part (c). Some scored one mark for cheaper/quicker but many thought it more representative.

Mark Ranges and Award of Grades

Grade boundaries and cumulative percentage grades are available on the [Results statistics](#) page of the AQA Website.

UMS conversion calculator www.aqa.org.uk/umsconversion