1. Finding the midpoints of the given groups was predominantly carried out correctly with very few errors seen. In contrast, attempts at finding the width and height of the 26 –30 group were extremely varied, with most candidates finding this particularly challenging, especially in finding the height. In the majority of cases, candidates obtained the wrong width and height, mostly with no clear strategy, although these did multiply together to make 20.8 in some cases. Calculation of the mean was carried out successfully on the whole, although there were some apparent misconceptions, with quite a few candidates merely summing the midpoints (without multiplying by the frequency) and dividing this by 56.

The standard deviation proved to be more problematic, with frequent mistakes in both the formula and in their calculations. Some candidates used the sum of the f^2x's and others the sum of the $(fx)^2$ or the sum of the fx's all squared in their formula. Very few candidates calculated s. Most candidates were able to use the correct interpolation technique to obtain the median, although many lost the accuracy mark through their use of 21 as the lower class boundary (which was relatively common) and /or 4 as the class width. Quite a few candidates worked with 28.5. A few candidates tried to apply the correct formula to the wrong class interval, however. Some candidates appeared to have a limited understanding of the class boundaries and failed to recognise the continuous nature of the data.

The majority of candidates were able to carry out a suitable test to determine the skewness of the data correctly. This mostly involved comparing $Q_3 - Q_2$ to $Q_2 - Q_1$ (with or without explicit substitutions), although the wrong conclusion was often drawn, either following on from a previous error in evaluating the median or from a lack of understanding of what their result was showing. A few students evaluated 3(mean – median)/standard deviation. Quite often the result of their test was described in words not figures, for example Q_2 is closer to Q_3 than Q_1. Some candidates merely attempted to describe the skewness without carrying out any test.

2. Parts (a) and (b) were answered very well although a few candidates gave the upper quartile as 39 or 39.5 (usually as a result of incorrectly rounding $\frac{3n}{4}$) however the follow through marks meant that no further penalty need occur. A few found the upper and lower quartiles but failed to give the interquartile range. Most found the limit for an outlier using the given definition, although a few used 1.5 × IQR, and went on to make a suitable comment about the one employee who needed retraining. There were some excellent box plots seen with all the correct features clearly present but a number failed to plot the outlier appropriately and simply drew their lower whisker to 7. A not insignificant minority were confused by the absence of an upper whisker and felt the need to add one usually at Q_3 + IQR.

3. Part (a) gave most candidates two easy marks but the rest of the question proved more demanding. The calculation of the mean in part (b) was usually answered well but there were still some dividing by 8 and a few using $\frac{\sum fx^2}{\sum fx}$. The calculation of the standard deviation was better than on previous occasions with many reaching 0.421 but there is still some confusion over the formula with $\sqrt{(15889.5 - \bar{x}^2)}$, a hybrid of the correct formula and S_{xx} , being quite common. Candidates should be aware that the formula for standard deviation is very sensitive to rounding errors and an accurate value for the mean (stored on their calculator) should be used rather than a rounded answer. A number of candidates failed to use the given values for Σfx and Σfx^2 and lost marks because of numerical slips. The attempts at interpolation in part (d) were much improved with the correct fraction often being added to a lower class boundary. Unfortunately many used 2.95 or 2.5 as their class boundary and lost the marks. In part (e) the better candidates used their values for the mean and median and made an appropriate comment. Some spent the next page calculating Q_1 and Q_3 , often correctly, in order to use the quartiles to justify their description of the skewness.

4. There were many good answers to this question. The Venn diagram was often totally correct although a number failed to subtract for the intersections and obtained value of 35, 40 and 28 instead of 31, 36 and 24 for the numbers taking two options. Parts (b) and (c) were answered very well with only a minority of candidates failing to give probabilities. Part (d) proved straightforward for those who knew what was required but some attempted complicated calculations, often involving a product of probabilities, whilst others simply gave their answer as 4/180.

5. Although there were more correct solutions than in previous papers for this type of question the process required to answer this question was not applied successfully by a large number of candidates. The most common error in part (a) was to give an answer of 0.8. In tandem with this was an answer to part (b) of 7.5 where candidates recognised that the answer to part (a) times the answer to part (b) should be 6. Many candidates divided 9 by 3 in part (b) but failed to multiply by 2. Other candidates however produced two correct answers but nothing else. The variety of approaches may suggest some logical thinking rather than a taught approach to this type of problem.

6. Very few candidates got full marks for this question, being unable to perform the calculations for grouped data, although the mean caused the least problems. Those candidates with good presentation particularly those who tabulated their workings tended to fare better. In spite of the well defined groups many candidates subtracted or added 0.5 to the endpoints or adjusted the midpoints to be 0.5 less than the true value with the majority getting part (a) incorrect as a result. As usual all possible errors were seen for the calculation of Σfx^2 i.e. $(\Sigma fx)^2$, $\Sigma (fx)^2$, Σf^2x and Σx^2. Use of 17.1 for the mean in the calculation of the standard deviation led to the loss the accuracy mark. Candidates are once again reminded not to use rounded answers in subsequent calculations even though they usually gain full marks for the early answer. The comment in part (c) was often forgotten perhaps indicating that candidates are able to work out the figures but do not know what they mean, although many did appreciate in part (d) that there is no skew in a normal distribution. As opposed to question 1, correlation was often mentioned instead of skewness although again this is becoming less common.

7. This question was usually answered well. In part (b) some did not realise that they needed to check the lower limit as well in order to be sure that 110 was the only outlier. Part (c) was answered very well although some lost the last mark because there was no gap between the end of their whisker and the outlier. Part (d) was answered very well and most gave the correct values for $\sum y$ and $\sum y^2$ in the appropriate formula. A few tried to use the $\sum (y - \bar{y})^2$ approach but this requires all 10 terms to be seen for a complete "show that" and this was rare.

Part (e) was answered well although some gave the answer as –5.7 having forgotten the 10^{-3}, or failed to interpret their calculator correctly. Many candidates gave comments about the correlation being small or negative in part (f) but they did not give a clear reason for rejecting the parent's belief. Once again the interpretation of a calculated statistic caused difficulties.

8. Part (a) was not answered well. Many candidates attempted to calculate frequency densities but they often forgot to deal with the scale factor and the widths of the classes were frequently incorrect. There are a variety of different routes to a successful answer here but few candidates gave any explanation to accompany their working and it was therefore difficult for the examiners to give them much credit. The linear interpolation in part (b) was tackled with more success but a number missed the request for the Inter Quartile Range. Whilst the examiners did allow the use of (n + 1) here, candidates should remember that the data is being treated as continuous and it is therefore not appropriate to "round" up or down their point on the cumulative frequency axis. Although the mean was often found correctly the usual problems arose in part (c) with the standard deviation. Apart from those who rounded prematurely, some forgot the square root and others used $\sum f^2 x$ or $\left(\sum fx\right)^2$ or $\frac{\sum fx^2}{\sum fx}$ instead of the correct first term in their expression and there was the usual crop of candidates who used n = 6 instead of 104. The majority were able to propose and utilise a correct test for the skewness in part (d) with most preferring the quartiles rather than the mean and median. Few scored both marks in part (e) as, even if they chose the median, they missed mentioning the Inter Quartile Range. A number of candidates gave the mean and standard deviation without considering the implications of their previous result.

9. Most candidates were able to draw a six-branched tree diagram correctly, although a number of candidates had incorrect or missing labels. From a correct diagram most gained full marks in part (b). The conditional probability in part (c) once again caused difficulty for many of the candidates. Many of the responses in part (d) were, incorrectly, referring to the importance of testing people for a disease rather than referring to the probability in part (c).

10. Parts (a) to (d) represented a chance for all students who had an average grasp of statistics to score highly. The median in part (b) was incorrectly identified by a significant number of candidates, but the standard deviation was often correct.

Part (e) was done surprisingly well with students appearing to have a much greater understanding of what is required for a comparison than in previous years. Often numbers were stated without an actual comparison. Confusion was evident in some responses as skewness was often referred to as correlation. A small minority of candidates had failed to take note of the 'For the Balmoral Hotel' and had done some correct statistics for all 55 students.

11. A lot of fully correct Venn Diagrams were seen in part (a) although it was surprising the number who resorted to decimals rather than just using straightforward fractions; this often led to loss of many accuracy marks. A significant minority had negative numbers in their Venn diagram and saw nothing wrong in this when converting them to probabilities later in the question. Fewer candidates forgot the box this time. Part (c) proved to be the only difficult part, as many candidates struggled with the concept of conditional probability, and many denominators of 300 were seen.

12. The mean was calculated accurately by the majority of the candidates but the standard deviation calculation still caused problems for many. There were few summation errors but missing square roots or failing to square the mean were some of the more common errors.

Part (b) was poorly answered. The examiners were disappointed that such a sizeable minority failed to order the list and worked quite merrily with a median larger than their upper quartile. Some worked from the total of 2757 to get quartiles of 689.5, 1378.5 and 2067.5 whilst others used cumulative totals and obtained quartiles in the thousands but still failed to see the nonsensical nature of these figures. Those who did order the list used a variety of methods to try and establish the quartiles. Whilst the examiners showed some tolerance here any acceptable method should give a median of 198 but many candidates used 186. Those who knew the rules usually scored full marks here and in parts (c) and (d).

The examiners followed through a wide range of answers in part (c) and most candidates were able to secure some marks for correctly identifying patients *B* and *F* and in part (d) for describing their skewness correctly.

13. The common error here was to assume that frequency equals the area under a bar, rather than using the relationship that the frequency is <u>proportional</u> to the area under the bar. Many candidates therefore ignored the statement in line 1 of the question about the histogram representing 140 runners and simply gave an answer of 12 × 0.5 = 6. A few candidates calculated the areas of the first 7 bars and subtracted this from 140, sadly they didn't think to look at the histogram and see if their answer seemed reasonable. Those who did find that the total area was 70 usually went on to score full marks. A small number of candidates had difficulty reading the scales on the graph and the examiners will endeavor to ensure that in any future questions of this type such difficulties are avoided. A small number of candidates had difficulty reading the scales on the graph and the examiners will endeavor to ensure that in any future questions of this type such difficulties are avoided.

14. Parts (a), (b) and (c) were generally well done, although in part (c) there were many with strange ideas of heavy instruments. In part (d) the majority of candidates were able to make a credible attempt at this with most giving one of the two possible solutions with a reason. The majority used the median and quartiles to find that the distribution was symmetrical. The use of the words 'symmetrical skew', similar to 'fair bias', is all too often seen but was accepted. Equal, even or normal skew were also often seen and were given no credit.
Part (e) was attempted successfully by a minority of candidates. A large number of candidates did not understand the distinction between z–values and probabilities. A lot gave 0.68 as z-value leading to the loss of the accuracy mark. Others tried to put various values into standard deviation formulae.

15. This was generally well answered. The majority of the errors occurred in part (c) by rounding too early and getting 18.4 for a. The regression line was often inaccurately plotted. In part (e) many used chocolate content to justify answer and often did not use the regression line to get a suitable price. Some misunderstood the question and attempted to find the best value in the second part of part (e)

16. Many candidates started well with this question, but a large number of inaccurate answers were seen for the latter parts. Part (a) was usually correct and part (b) was generally done well. In part (c) there were a lot of mistakes in finding midpoints and also $\sum f$. Most knew the correct method for finding the mean, but rather fewer knew how to find the standard deviation in part (d) although most remembered to take the square root. Part (e) was very badly answered, with the majority unable to interpolate correctly which was often due to wrong class boundaries and / or class widths. In part (f), although the majority got an incorrect numerical value, most picked up the mark for interpreting their value correctly.

17. This question caused problems for many candidates. Part (a) did not always generate a comment about the skewness of the data and many who did eventually mention skewness thought it was negative. The calculation of the median in part (b) often caused difficulties. An endpoint of 19.5 was often used, but some thought the width was 9 not 10 and many simply opted for the midpoint of 24.5. The calculation of the mean in part (c) was sometimes the only mark scored by the weakest candidates and the examiners were disappointed at how many candidates were unable to find the standard deviation. Aside from the usual error of missing the square root or failing to square the mean, a number were using formulae such as $\sqrt{\frac{\sum fx^2}{\sum fx}}$. Most scored some marks in part (d) for attempting to use their values in the given formula, but the final mark required an answer accurate to 3 sf and this was rarely seen. In part (e) many failed to comment on the sign of their coefficient and there was often a discussion of correlation here rather than skewness. Of those who attempted the last two parts, part (g) was often successful, but in part (f), candidates often chose the mean because it used all the data rather than the median, which wouldn't be affected by the extreme values.

18. Parts (a) and (b) were not answered well. Few mentioned the type of variable in part (a) and in part (b) many simple stated that the frequency equals the area rather than stating that it was proportional to the area.

Many were able to give a correct calculation in part (c) but they sometimes failed to state that the 0.8 related to each individual child; the question was a "show that" and a final comment was required. The calculation in part (d) was usually correct.

19. Part (a) often scored full marks although some still mention 'mean' instead of 'median'. Part (d) was very straightforward for the vast majority of candidates. Those candidates who used a scale of 4cm to 10 units were sometimes prone to placing the median inaccurately. Part (e) was also quite well done but some only listed the 5 important values with little or no mention of IQR, range, outliers or skewness. There was occasional confusion thinking the bigger numbers meant school B had done better.

20. In part (a) there were very few correct solutions. It was rare for a candidate to appreciate that the selection was without replacement. The rest of this question was well answered by many, although a surprising number averaged the two means in part (c).

21. This question was a good source of marks with most candidates able to find the correct values for the mode and median, but too many getting the upper quartile wrong. A surprising number of candidates had problems with finding the standard deviation. In quite a few cases the square root was omitted but more often marks were lost due to the misinterpretation and misuse of standard formulae. This was not helped by some candidates ignoring given totals and calculating their own. In part (e) a large number of responses gave one reason rather than two.

22. It was very disappointing that so few candidates could carry out a simple analysis of a set of data. Few scored well.

(a) Relatively few candidates were able to state, "distance is a continuous variable." The most common wrong answer in this part referred to the unequal class widths.

(b) In general frequency densities were well done. The most common mistake was to calculate the incorrect class width, taking the first class width as 4. Other mistakes were class width divided by frequency or frequency multiplied by class width although these were less prevalent than in previous examinations.

(c) Interpolation was not familiar to many candidates. Those pupils who did attempt to interpolate to find the median and quartiles were on the whole successful, common errors being the use of 50 instead of 50.5 or the wrong class interval. Many used the mid-point of the class for the quartiles or more frequently used 134/2 or (134 + 1)/2 as their responses for an estimate of the median.

(d) The mean was calculated successfully by the vast majority of candidates with only occasional error through using the sum of fx^2 as opposed to the sum of fx. The standard deviation proved more difficult – where students used the wrong formula, omitted the square root or lost accuracy marks through using the rounded value of the mean. Some candidates wasted time by recalculating the values given.

(e) Those candidates with sensible values for their quartiles managed to substitute successfully to calculate the coefficient although it was surprising how many could not get 0.14 from a correct expression. On the whole they drew the correct conclusion about the data being positively skewed, although a small number of candidates managed a correct calculation and then concluded negative skewness.

(f) Although a fair number of students could give a reason to confirm that the skewness was positive, most lost this mark by not justifying their comment using numerical values.

23. (a) The vast majority of candidates were able to make an attempt at drawing a box plot though labels were not always added and the upper whisker often extended to 63. For many candidates this was the only mark they obtained for the question. Few candidates bothered, or were able, to use the information regarding 1.5 IQR in order to identify the limits of acceptable data. Of those candidates who did show some working more often than not, they did not do so in enough detail. The number 24 was usually implied but candidates often ignored showing working for the lower end. The numbers of 52 and -12 were visible on a number of papers, but the conclusion about which numbers were outliers was often omitted.

(b) The majority of candidates recognized positive skewness, but many did not justify their answer numerically. Some candidates did not understand the request about the "distribution of delays" and gave an interpretation more suited to part (c).

(c) Most managed a comment on the distribution that was relevant, but few wrote in terms of whether passengers would be bothered by the delays – the majority of students used technical statistical terms, referring to quartiles and percentages of the data, rather than simply interpreting the data in non-technical language.

24. A lack of detailed labelling in the box plot was common. Candidates should realise that 3 marks for parts where they are comparing etc. requires them to find three relevant points. Many only had one or 2 points and seemed to think that if they wrote enough about one point they could get the 3 marks. The last part was not well interpreted by many. They were likely to just say that the 2 values for Q_3 were the same. Most candidates can find quartiles and know how to display the information in box plots. There are still some candidates who do not draw a clearly labelled axis for their scale. Candidates need to remember that the purpose is to compare data so the scale needs to be the same for both sets of data. Some candidates can give good comparisons referring to range, IQR, median and quartiles, but many give vague descriptions concerning 'spread' and 'average' which gain no marks. They should be encouraged to be specific in their descriptions. Very few can interpret the upper quartile in context.

25. Many candidates were able to calculate the median and both quartiles accurately; the most common error was to give Keith's median as 201. Some candidates ignored or failed to show the calculation of outliers. However, the vast majority of candidates were able to draw a reasonable box plot although the scale was often unlabelled. Other common errors were to extend the left hand whisker to 146 and the right hand whisker to 269 with 266 marked on as a bound for outliers.

In part (c), a failure to show working was all too common. Asif's distribution was often said to be "negatively skewed", with only a minority qualifying it as weak or almost symmetrical. Some candidates confused negative skew with positive skew.

26. Part (a) produced a poor response, with very few candidates realising when a histogram should be used. Part (b) was answered correctly by a very large majority of candidates. Most candidates appreciated the need for frequency densities and they were usually calculated accurately. The chosen scale was often good, but unsuitable scales are still seen too frequently at this level. Candidates generally labelled their axes. The heights if the rectangles were usually correctly plotted, but unsuitable scales sometimes proved a hindrance to candidates.

27. Candidates knew how to tackle this question but too many of them did not pay sufficient attention to detail. They often calculated the variance and not the standard deviation and their arithmetic was not always as accurate as it should have been. Some poor computational methods were seen when calculating the standard deviation. The stem and leaf diagram caused few problems for the candidates but few of them presented it in the most appropriate form. The mode and quartiles were often correct and it was pleasing to see many of them making a good attempt to compare and contrast the attendance data.

28. This was a long question that needed an understanding of several different but frequently linked concepts and many of the candidates did not have the stamina for such a question. Common errors included the drawing of a bar chart; fd = class width/frequency; poorly drawn histograms; quartiles calculated without using a lower class boundary; IQR omitted; a variety of incorrect mid-points; poor arithmetic and no appreciation of the skewness of the data. For a routine type of question the overall response was very disappointing.

29. This question posed few conceptual problems for the candidates but few of them gained full marks. The majority of lost marks were the result of candidates not paying sufficient attention to detail. The standard deviation was rarely given to an appropriate degree of accuracy and some candidates did not look for outliers at both ends of the data set. Although choosing a scale that would fit on the grid supplied for the candidates was not easy, most managed to do so but then forgot to label the axis or the two box plots. Others ignored the outliers even though they had identified them. Whilst candidates tried to find two sensible comparisons very few expressed them clearly.

30. There was still some uncertainty about how to calculate frequency densities. There were examples of candidates using (class width)/frequency or mid-point × frequency. Wrongly labelled axes, poor scales and wrong bases for the histogram bars lost marks for many candidates. A completely correct histogram was not very common.

31. Most candidates could make a reasonable attempt at this question. The stem and leaf diagram rarely had a label and although almost all candidates gave the correct value for the median, far too many did not give correct values for the other quartiles. Showing that there were no outliers was not always well attempted since many could not calculate the IQR correctly. The box plot was often spoiled by having no label, a poor or no proper scale and inaccurate plotting. Positive skewness was usually recognised and a correct justification was often given.

32. This question was generally well answered but some candidates could not work out frequency densities correctly. Many histograms were poorly labelled and many candidates gained marks on this question only because examiners followed through their frequency densities. There were still too many candidates who drew bar charts instead of histograms.

33. The upper quartile caused problems for many candidates but they should have seen an example with 20 observations and been able to deal with the quartiles. The values obtained by the candidates were followed through and this allowed many of them to score most of the marks. Working for the outliers was often omitted, with a loss of marks as stated in the rubric, and the label on the box plot was often missing. The mean was usually correct and some attempt to comment on the skewness was nearly always made.

34. No Report available for this question.

35. No Report available for this question.