STATISTICS

Paper 4040/12
Paper 1

Key messages

It is always important to read a question carefully to understand fully the situation described and any specific constraints involved.

After obtaining the result of a calculation, it is worth pausing to consider whether or not the result seems possible and reasonable for the practical situation of the question.

Candidates need to regard the writing parts of questions, which may involve explanations and assumptions, as being as important as the routine calculations they have learned to carry out.

General comments

The standard of work involving calculations of a routine nature was again generally good. This was particularly true of finding the mean and standard deviation of a grouped frequency distribution (see **Question 7** below), finding a line of best fit (see **Question 9** below), and the use of pie and bar charts (see **Question 10** below). On some questions involving minimal amounts of working, especially that on a two-way table, and that on range as a measure of dispersion (see **Questions 4 and 6** below), performance varied a lot. Few candidates obtained full marks on all the probability elements on the paper (see **Questions 5 and 9(g)** below).

There were instances on the paper where some candidates had perhaps not read the question carefully enough in order to appreciate particular details of the situation or data provided (see **Questions 5 and 10(a)** below).

Answers to the writing parts of questions, involving explanations and assumptions, were generally very limited. These are the questions designed to test the candidates' proper understanding of what they have learned.

It has to be emphasised again, as it has been in past reports, that Statistics is a subject which is applied to real-life situations, and that candidates should try to be aware of whether or not their answers are reasonable for the situation of the question. It is not expected that candidates will be familiar with the statistics of the many practical situations which form the subject of examination questions. But they should have some appreciation of the size of numbers, and how, for example, 200 metres is of a very different order of magnitude from 5 metres. There were two particular instances in this paper where some highly unrealistic answers were presented, which should have given the candidate at least pause for thought about the working they had done leading to these answers (see **Questions 7(b) and 9(e)** below).

Comments on specific questions

Question 1

Almost all candidates were able to interpret the pictogram correctly.

Question 2

Parts **(b)** and **(c)** were answered well. Answers to **part (a)** varied more in quality. The best stated how the dependent variable changed, in the two cases, as the independent variable increased. Credit was also given to answers which explained clearly, in the two cases, the pattern of the points on scatter diagrams. However, candidates opting to answer in this way often produced more limited answers, sometimes with little meaning.

Question 3

Venn diagrams continue to create problems for some candidates. A common error in the interpretation in **part (a)** was omitting to say that that these teenagers, whilst liking two of the types of music, did not like the third. In **part (c)** a common error was to have 10 in the numerator rather than 17. Some candidates show confusion in interpreting a Venn diagram between what is represented by the full double intersection of such a diagram, and the part of the intersection which excludes the third category.

Question 4

There were many fully correct answers to **part (a)**. There were fewer such answers to **part (b)**. It seemed that here it was the 3–5 class that created the problem, some candidates treating the five people in this class visiting the dentist as having visited 3 times and 4 times and 5 times. Good answers to **part (c)** were rare. Most candidates did not seem aware of the fact that the data do indeed show that more visits to a doctor were made than to a dentist. But few were able to offer any one of many possible reasons why Thabisa's suggestion could be wrong. Too many wrote about the open '10 or more' class in visits to a doctor.

Question 5

It was essential to understand the rules of the game in this question, explained with an example, and it was apparent that not all candidates had done so. In **part (b)** those who showed they had understood the rules, by listing the possible ways a score of nine could be obtained, always earned some marks, even if they did not eventually obtain the correct answer. Full marks were awarded in **part (c)** to answers which used the answer to **part (b)** in the appropriate way, even if the answer to **part (b)** was incorrect. Too many candidates however simply multiplied their answer to **part (b)** by two.

Question 6

The first two parts of **part (a)** were generally answered correctly. But very little understanding was shown in **part (a)(iii)** of the limitations of range as a measure of dispersion, as a consequence of its use of only two values (the extreme values) in any set of data. **Part (b)** was answered moderately well, but there were fewer correct answers to **part (c).** Many candidates appeared not to realise that, to find the range requested in **part (c)**, the longest journey times had first to be found from the values in the second and third columns of the given data.

Question 7

In **part (a)** very good computational skills were shown, and many fully correct answers were seen. Occasionally, as has been noted in the past, there were small errors in the value of the standard deviation as a consequence of the use of only a three significant figure value for the mean in the formula. The table in such questions is given with plenty of blank space so that candidates can create additional columns of their own to order their working and keep it tidy. Many, however, do not do this.

Part (b) involved only a simple calculation using the total from **part (a)**, yet many candidates produced much confused arithmetic, sometimes resulting in an answer of totally unrealistic size. It is certainly not expected that candidates will be familiar with the details of cheese making; yet it ought to have been realised that, for a herd of animals producing milk for 18 weeks, an answer of only about 5 kg of cheese being made had to be incorrect.

Part (c) was generally answered well, with most candidates appreciating that not all class frequencies could be found simply by reading column heights. As the question did not specifically ask about class frequencies an approach was possible, working from the areas of the different parts of the histogram, without finding the frequencies. Some candidates did this successfully. Occasionally the central column was included, incorrectly, in the calculation.

Question 8

There was a larger proportion of correctly drawn graphs in **part (b)** than has often been the case in the past. Many candidates clearly knew that the cumulative frequencies had to be plotted against the upper boundaries of the classes. However a minority still made the serious error of plotting the points at either the mid-points or lower boundaries of the classes. Accuracy of later answers is always affected when this is done.

Part (c) was very well answered, but the remaining parts less so. In **part (d)** only the candidates who saw that the question involved the 18 tremors with a magnitude of 2.0 or more were able to gain credit. In **part (e)** few realised that all that was required was a magnitude reading from the graph corresponding to a cumulative frequency of 78. In **part (f)** there was limited appreciation of the fact that that the median is the better measure of central tendency here because of the particular nature of the distribution of data.

In their answers to all parts of the question, many candidates explained their answers extremely well with lines drawn on the graph. Candidates who did this are much to be commended. Examiners were then able to award marks for method even if there were numerical inaccuracies in the work.

Question 9

Parts **(a)**, **(c)** and **(d)** were answered well, with more variable answers to **part (b)**. Although the lower semi-average was often found correctly in **part (c)**, candidates are advised to use the two given averages when finding the line of best fit, in case the one not given has been calculated incorrectly. In **part (d)** a few candidates read points from their drawn line to find its equation. Whilst correct they should be aware that this is a less accurate way of proceeding than working from the averages.

The last three parts were less well answered. In **part (e)** good answers used $y = 1$, and found x either from the equation, or from reading the graph. Limited attempts, gaining no credit, set $y = 0$, or substituted some value for x. This was another question where occasionally completely unrealistic answers were offered. It is not expected that candidates should know a likely winning height for an athlete in the pole vault; but it should have been obvious that an answer of around 200 metres, or of half a metre, could not possibly be correct.

In **part (f)** there was less reference than expected to the problems of extrapolating beyond the limits of experimental data. And in **part (g)**, whilst the product of two 'without replacement' probabilities was often presented, at least one of numerator and denominator was frequently incorrect.

Question 10

Answers to **part (a)** were mixed. Candidates who saw that the populations of the countries were given in millions, so all that was required was a simple division of the values in the third column by those in the second, almost always earned at least two of the available marks. Others, possibly not having noticed this, needlessly introduced many zeros into their calculations, and ended with answers that were either excessively large or excessively small. In all cases errors were made frequently in the choice of country, many choosing the country with the smallest death rate instead of the one with the highest.

In **parts (b) and (c)** interpretation of the charts was very good, and many fully correct answers were seen. Almost all candidates recognised that in such cases only integer answers are possible. The combined use of pie chart and bar chart in **part (d)** was handled moderately well, with only a minority of candidates offering an answer derived from just one of the charts. The necessary assumption, that the proportions of deaths by road user type were the same in the South as in the country as a whole, was clearly expressed only in a few cases.

Responses to **part (e)** were very limited. Very few candidates were able to explain that a standard population is broken down into age group categories simply because death rates (for natural deaths) are strongly influenced by age; and few were able to offer the comment that age may not similarly be the main determining factor in road accident deaths, which are not natural deaths.

STATISTICS

Paper 4040/13
Paper 1

Key messages

It is always important to read a question carefully to understand fully the situation described and any specific constraints involved.

After obtaining the result of a calculation, it is worth pausing to consider whether or not the result seems possible and reasonable for the practical situation of the question.

Candidates need to regard the writing parts of questions, which may involve explanations and assumptions, as being as important as the routine calculations they have learned to carry out.

General comments

The standard of work involving calculations of a routine nature was again generally good. This was particularly true of finding the mean and standard deviation of a grouped frequency distribution (see **Question 7** below), finding a line of best fit (see **Question 9** below), and the use of pie and bar charts (see **Question 10** below). On some questions involving minimal amounts of working, especially that on a two-way table, and that on range as a measure of dispersion (see **Questions 4 and 6** below), performance varied a lot. Few candidates obtained full marks on all the probability elements on the paper (see **Questions 5 and 9(g)** below).

There were instances on the paper where some candidates had perhaps not read the question carefully enough in order to appreciate particular details of the situation or data provided (see **Questions 5 and 10(a)** below).

Answers to the writing parts of questions, involving explanations and assumptions, were generally very limited. These are the questions designed to test the candidates' proper understanding of what they have learned.

It has to be emphasised again, as it has been in past reports, that Statistics is a subject which is applied to real-life situations, and that candidates should try to be aware of whether or not their answers are reasonable for the situation of the question. It is not expected that candidates will be familiar with the statistics of the many practical situations which form the subject of examination questions. But they should have some appreciation of the size of numbers, and how, for example, 200 metres is of a very different order of magnitude from 5 metres. There were two particular instances in this paper where some highly unrealistic answers were presented, which should have given the candidate at least pause for thought about the working they had done leading to these answers (see **Questions 7(b) and 9(e)** below).

Comments on specific questions

Question 1

Almost all candidates were able to interpret the pictogram correctly.

Question 2

Parts **(b)** and **(c)** were answered well. Answers to **part (a)** varied more in quality. The best stated how the dependent variable changed, in the two cases, as the independent variable increased. Credit was also given to answers which explained clearly, in the two cases, the pattern of the points on scatter diagrams. However, candidates opting to answer in this way often produced more limited answers, sometimes with little meaning.

Question 3

Venn diagrams continue to create problems for some candidates. A common error in the interpretation in **part (a)** was omitting to say that that these teenagers, whilst liking two of the types of music, did not like the third. In **part (c)** a common error was to have 10 in the numerator rather than 17. Some candidates show confusion in interpreting a Venn diagram between what is represented by the full double intersection of such a diagram, and the part of the intersection which excludes the third category.

Question 4

There were many fully correct answers to **part (a)**. There were fewer such answers to **part (b)**. It seemed that here it was the 3–5 class that created the problem, some candidates treating the five people in this class visiting the dentist as having visited 3 times and 4 times and 5 times. Good answers to **part (c)** were rare. Most candidates did not seem aware of the fact that the data do indeed show that more visits to a doctor were made than to a dentist. But few were able to offer any one of many possible reasons why Thabisa's suggestion could be wrong. Too many wrote about the open '10 or more' class in visits to a doctor.

Question 5

It was essential to understand the rules of the game in this question, explained with an example, and it was apparent that not all candidates had done so. In **part (b)** those who showed they had understood the rules, by listing the possible ways a score of nine could be obtained, always earned some marks, even if they did not eventually obtain the correct answer. Full marks were awarded in **part (c)** to answers which used the answer to **part (b)** in the appropriate way, even if the answer to **part (b)** was incorrect. Too many candidates however simply multiplied their answer to **part (b)** by two.

Question 6

The first two parts of **part (a)** were generally answered correctly. But very little understanding was shown in **part (a)(iii)** of the limitations of range as a measure of dispersion, as a consequence of its use of only two values (the extreme values) in any set of data. **Part (b)** was answered moderately well, but there were fewer correct answers to **part (c).** Many candidates appeared not to realise that, to find the range requested in **part (c)**, the longest journey times had first to be found from the values in the second and third columns of the given data.

Question 7

In **part (a)** very good computational skills were shown, and many fully correct answers were seen. Occasionally, as has been noted in the past, there were small errors in the value of the standard deviation as a consequence of the use of only a three significant figure value for the mean in the formula. The table in such questions is given with plenty of blank space so that candidates can create additional columns of their own to order their working and keep it tidy. Many, however, do not do this.

Part (b) involved only a simple calculation using the total from **part (a)**, yet many candidates produced much confused arithmetic, sometimes resulting in an answer of totally unrealistic size. It is certainly not expected that candidates will be familiar with the details of cheese making; yet it ought to have been realised that, for a herd of animals producing milk for 18 weeks, an answer of only about 5 kg of cheese being made had to be incorrect.

Part (c) was generally answered well, with most candidates appreciating that not all class frequencies could be found simply by reading column heights. As the question did not specifically ask about class frequencies an approach was possible, working from the areas of the different parts of the histogram, without finding the frequencies. Some candidates did this successfully. Occasionally the central column was included, incorrectly, in the calculation.

Question 8

There was a larger proportion of correctly drawn graphs in **part (b)** than has often been the case in the past. Many candidates clearly knew that the cumulative frequencies had to be plotted against the upper boundaries of the classes. However a minority still made the serious error of plotting the points at either the mid-points or lower boundaries of the classes. Accuracy of later answers is always affected when this is done.

Part (c) was very well answered, but the remaining parts less so. In **part (d)** only the candidates who saw that the question involved the 18 tremors with a magnitude of 2.0 or more were able to gain credit. In **part (e)** few realised that all that was required was a magnitude reading from the graph corresponding to a cumulative frequency of 78. In **part (f)** there was limited appreciation of the fact that that the median is the better measure of central tendency here because of the particular nature of the distribution of data.

In their answers to all parts of the question, many candidates explained their answers extremely well with lines drawn on the graph. Candidates who did this are much to be commended. Examiners were then able to award marks for method even if there were numerical inaccuracies in the work.

Question 9

Parts **(a)**, **(c)** and **(d)** were answered well, with more variable answers to **part (b)**. Although the lower semi-average was often found correctly in **part (c)**, candidates are advised to use the two given averages when finding the line of best fit, in case the one not given has been calculated incorrectly. In **part (d)** a few candidates read points from their drawn line to find its equation. Whilst correct they should be aware that this is a less accurate way of proceeding than working from the averages.

The last three parts were less well answered. In **part (e)** good answers used $y = 1$, and found x either from the equation, or from reading the graph. Limited attempts, gaining no credit, set $y = 0$, or substituted some value for x. This was another question where occasionally completely unrealistic answers were offered. It is not expected that candidates should know a likely winning height for an athlete in the pole vault; but it should have been obvious that an answer of around 200 metres, or of half a metre, could not possibly be correct.

In **part (f)** there was less reference than expected to the problems of extrapolating beyond the limits of experimental data. And in **part (g)**, whilst the product of two 'without replacement' probabilities was often presented, at least one of numerator and denominator was frequently incorrect.

Question 10

Answers to **part (a)** were mixed. Candidates who saw that the populations of the countries were given in millions, so all that was required was a simple division of the values in the third column by those in the second, almost always earned at least two of the available marks. Others, possibly not having noticed this, needlessly introduced many zeros into their calculations, and ended with answers that were either excessively large or excessively small. In all cases errors were made frequently in the choice of country, many choosing the country with the smallest death rate instead of the one with the highest.

In **parts (b) and (c)** interpretation of the charts was very good, and many fully correct answers were seen. Almost all candidates recognised that in such cases only integer answers are possible. The combined use of pie chart and bar chart in **part (d)** was handled moderately well, with only a minority of candidates offering an answer derived from just one of the charts. The necessary assumption, that the proportions of deaths by road user type were the same in the South as in the country as a whole, was clearly expressed only in a few cases.

Responses to **part (e)** were very limited. Very few candidates were able to explain that a standard population is broken down into age group categories simply because death rates (for natural deaths) are strongly influenced by age; and few were able to offer the comment that age may not similarly be the main determining factor in road accident deaths, which are not natural deaths.

STATISTICS

Paper 4040/22
Paper 2

Key message

It is important that candidates read instructions in the questions carefully. For example, **Questions 7a, 7c** and **7e** contained the instructions 'include a key', 'use values from your table to justify your answer', and 'include all necessary labels', respectively. In each of these cases there was evidence that some candidates had not followed these instructions carefully.

Candidates should be advised that it is good practice to check that any calculated values seem sensible in the context of the question. For example, some values greater than 1 were seen as answers to questions requiring a probability. In **Question 6a**, any working leading to a median outside of the range 600 g to 1000 g should have been checked by candidates. In **Question 8c**, candidates should have checked working resulting in a positive seasonal component for quarter 2, noting that the number of swans was always below the centred moving average value in this quarter.

Where a question requires interpretation, this should always be done in the context of the problem. **Question 2a** required candidates to justify the use of a specific technique, and this should be done in the context of the aims study.

General comments

Most candidates made a good start to the paper and appeared well-prepared. Presentation of work was mostly good, with accurately plotted graphs. Axis labels and keys on graphs and diagrams are essential if the data presented is to be understood, and these were not always present. Most candidates appeared to have sufficient time to complete the paper.

A small number of candidates offered more than one solution to a question, without making it clear which solution they wanted to be marked.

Comments on specific questions

Question 1

The first two parts of **Question 1** were answered well. In **part (a)**, a small number of candidates gave an incomplete description, omitting 'continuous' and giving only 'quantitative' as their description for the distance travelled, but most candidates scored full marks.

In **part (b)**, a small number of candidates found the range, rather than the interquartile range, or misread the scale when finding the lower quartile, but the majority found the interquartile range correctly.

Most candidates found **part (c)** more difficult, with the difference between the highest value and the upper quartile, namely 28 – 15 = 13, being a common wrong answer. Of those that got the answer correct, most found ¾ of 144 first, and then subtracted from 144, rather than simply finding ¼ of 144.

Question 2

The best answers to **part (a)** usually came from candidates giving a sectional bar chart as their alternative chart and providing a reason in the context of finding the most popular sport from observing the tallest bar. Reasons such as 'it will be easier to draw' or 'it will show the results clearly' were quite common and did not score the mark. Some candidates gave a percentage sectional bar chart as their chosen alternative, but this was only acceptable if they made it clear that it was to be a single bar combining the data from the boys and the girls, divided into the percentage of the children who preferred each of the three sports.

Most candidates found the correct probability in **part (b)**, but the probabilities in **parts (c)** and **(d)** were sometimes switched, or one or both given as 3/25.

Question 3

Fully correct solutions were often seen to **part (a)**, but some candidates omitted to conclude that the sample was not representative, having provided all the correct initial calculations. A common incorrect answer involved finding that there were 2 of each battery size in the given sample and concluding, without further calculations, that therefore the sample was representative or in some cases crossing out correct calculations leading to the representative strata sizes of 1, 3 and 2.

In **part (b)** candidates needed to make sensible rounding choices to find the strata sizes for a sample of size 5. Most candidates showed correct initial calculations, but many were unable to round appropriately, with common wrong answers being 0, 3, 2 or 1, 3, 1. Some candidates simply left the strata sizes as 1, 3, 2, without attempting to find the best possible sample of size 5.

Question 4

In this probability question, it appeared that the use of the words 'or' and 'and' rather than union and intersection symbols caused confusion for some candidates. Others did not use the fact that the events were independent. Common incorrect answers in **part (a)** came from simply doing 0.3 × 0.4 or from 0.3 + 0.4.

The correct answer for **part (a)** was often seen as the answer to **part (b)**.

Successful attempts at **part (c)** came either from using **part (a)**, and doing 1 – 0.58, or, more often, from doing (1 – 0.3) × (1 – 0.4). A common incorrect method seen was 1 – 0.3 × 0.4.

Question 5

In **part (a)**, most candidates were able to find the total amount spent on rent and on electricity last year before simplifying the ratio of the three categories. The most common error was for the rent to be left as the monthly figure rather than the annual amount.

Part (b) of this question was less structured than similar questions in the past and it was pleasing to see many candidates providing well-explained solutions, with some providing tables showing the price relatives and corresponding weights. A small number used a price relative of 102 rather than 98 for the other costs. Those that were unable to calculate the price relatives correctly were usually able to use their values in an appropriate formula for the weighted aggregate index and so most candidates achieved some, if not all, of the available marks.

Candidates found **part (c)** difficult, with many incorrectly suggesting reasons that are already accounted for in the price relatives, such as an increase in the cost of electricity or rent. Those that were successful usually gave a change in the number of units of electricity used as the explanation. This question was left blank by some candidates.

Question 6

This was a challenging question on linear interpolation. In **part (a)** it was quite common to see final answers that were outside of the range 600 g to 1000 g; candidates should have questioned such answers. Fully correct solutions were however seen, and these answers were usually rounded as instructed in the question.

There were only a small number of correct answers to **part (b)** with many candidates leaving this answer blank or stating that they had assumed that the median was the 50^{th} percentile or that 62 per cent of the parcels had masses between 600 g and 1000 g. Correct answers explaining that the assumption was that the masses of the parcels increase linearly were quite rare.

Question 7

In **part (a)** it was pleasing to see many completed stem-and-leaf diagrams with correctly ordered leaves. A few candidates did not attempt to line up the leaves appropriately, giving a distorted visual impression of the data. The most common cause of a loss of marks was for a missing or incomplete key. Incomplete keys sometimes only explained a female value on the right of the diagram, for example. Others, who attempted a key to explain values on the right and the left of the diagram, described the values of 33 and 30, say, as representing numbers of females and males rather than ages of a female and a male.

Many fully correct answers were seen in **part (b)**. The most common errors came from locating the median, say, by finding the 7.5th value rather than the 8th value.

In **part (c)**, the instruction in the question required the use of values from the table to justify the answer. Thus, in assessing Azeeb's claim, it was necessary to compare the median values and, in assessing Tebogo's claim, a comparison of the interquartile ranges was required. It is important that candidates read instructions such as this carefully.

Many fully correct answers to **part (d)** were seen, with the most common errors coming from putting the value 30, for example, into the $20 \leqslant x < 30$ class rather than the $30 \leqslant x < 40$ class.

Frequency polygons in **part (e)** usually had an appropriate key, but labels on the axes were often missing and upper class boundaries, rather than midpoints, were sometimes used.

Many correct advantages of a stem-and-leaf diagram over a frequency polygon were seen in **part (f)**, with reference made to the preservation of the original values. A few comments were too vague, such as 'it is easily interpreted', without an explanation as what it is that makes the stem-and-leaf easily interpreted.

Part (g) of this question proved very challenging for most candidates. A common incorrect answer seen in **part (i)** was 6, coming from 54 – 48. Of those that scored the mark in **part (i)**, very few were able to correctly answer **part (ii)**.

Question 8

Many candidates wrote down two correct reasons for finding moving average values in **part (a)**, with their use in finding the trend being the most common of the correct reasons given. A few candidates appeared to have misread the question and described reasons for centring the moving average values.

Most candidates correctly completed the table in **part (b)**, with the small number of errors that did occur appearing to be the result of an arithmetic error rather than a misunderstanding of the statistics.

Many fully correct calculations of the seasonal component were seen in **part (c)**. A few candidates performed the required subtractions the wrong way around and ended up with a positive seasonal component for quarter 2. Checking the original data should have alerted these candidates to their error. Some candidates tried to use the 2019 quarter 2 value in their calculation, but this value did not have a corresponding centred moving average value. Other errors seen were the use of the moving average values rather than the centred moving average values, so, for example, 332 – 1307.25 or 320 – 1294 seen as part of the calculation. Weaker candidates gave an answer of 326.3 from finding the mean of the three quarter 2 numbers of swans.

In **part (d)** most candidates plotted the centred moving average values accurately and drew an appropriate trend line. The most common plotting error came from the plot for 1297.75 being seen at approximately 1299.75, suggesting a misread of the vertical scale. It was acceptable for candidates to either include or ignore the first of the centred moving average values when it came to drawing the trend line, as this first value did not closely follow the pattern of the remaining ones. Joining the first plot and the last plot did not, however, form an appropriate trend line.

Most candidates did not comment on the fact, in **part (e)**, that the ranger's assistant had referred to 'each quarter'. Thus, the fact that the ranger's assistant was not correct, because the trend line shows a fall over time or each year, but not a fall each quarter, was missed by most candidates.

Accurate readings from the trend line in **part (f)** were usually seen, but some candidates gave this as the final answer without introducing the seasonal component. As the question requires an estimate of the number of swans, the final answer should have been given as a whole number.

Question 9

Many candidates correctly answered the first three questions in **part (a)**, but it was common to see errors in **part (iv)**. Most candidates, rather than use the previous answers, recalculated the answers to **parts (i)** and **(ii)** and added these, without subtracting the answer to **part (iii)**. Weaker candidates seemed to misunderstand the context of the problem, treating the question as if it involved one selection from six cards, rather than two independent selections, each from three cards.

Although few fully correct solutions were seen in **part (b)**, many candidates demonstrated some understanding of the requirements of the question. It was quite common to see the complement of their first probability used in an equation of the correct form. Some solutions would have benefitted from clearer explanations, such as labelling of the probabilities to indicate which represented the probability of identical cards and which represented the probability of non-identical cards.

The set up in **part (c)** seemed to be more easily understood than that in **part (a)**. Many candidates dealt correctly with the fact that the cards were not replaced between selections. Most also realised that in **part (i)**, for example, they needed to consider both square then triangle and triangle then square. Many candidates did not notice that the three options presented in **parts (i)** to **(iii)** covered all the options and therefore the answer to **part (iii)** could be simply obtained from the previous parts.

Most candidates understood the requirement for a fair game in **part (d)**, using their probabilities from **part (c)** correctly. Some weaker candidates simply added the prize values together.

Question 10

Most candidates took the correct approach to compare the scores in **part (a)** by attempting to scale or standardise them. Often this was done correctly with a correct conclusion, but sometimes the conclusion was incorrect and sometimes the subtraction had been done the wrong way around.

In **part (b)** many correct calculations of the mean were seen, but it was much less common to see a correct calculation of the standard deviation. The most common error when calculating the standard deviation was to use $9858^2 + 9588^2$ as the value for the sum of the squares. Some weaker candidates ignored the different number of candidates in each year group and simply found the mean of 62 and 68 and the mean of 8 and 10.

Candidates tended to be more successful with **part (c)**, and clearly set out solutions allowed for the award of part-marks when solutions were incomplete. Many, for example, got as far as finding the combined total score for the two year groups, even if they could progress no further.

In **part (d)**, most candidates correctly increased the mean by 10 per cent, but many did not realise that increasing each candidate's score by 10 per cent would also increase the standard deviation by 10 per cent.

In **part (e)**, many candidates realised that the mean would increase, but it was quite rare to see the effect on the standard deviation correct, with many thinking that it would stay the same.

STATISTICS

Paper 4040/23
Paper 2

Key message

It is important that candidates read instructions in the questions carefully. For example, **Questions 7a, 7c** and **7e** contained the instructions 'include a key', 'use values from your table to justify your answer', and 'include all necessary labels', respectively. In each of these cases there was evidence that some candidates had not followed these instructions carefully.

Candidates should be advised that it is good practice to check that any calculated values seem sensible in the context of the question. For example, some values greater than 1 were seen as answers to questions requiring a probability. In **Question 6a**, any working leading to a median outside of the range 600 g to 1000 g should have been checked by candidates. In **Question 8c**, candidates should have checked working resulting in a positive seasonal component for quarter 2, noting that the number of swans was always below the centred moving average value in this quarter.

Where a question requires interpretation, this should always be done in the context of the problem. **Question 2a** required candidates to justify the use of a specific technique, and this should be done in the context of the aims study.

General comments

Most candidates made a good start to the paper and appeared well-prepared. Presentation of work was mostly good, with accurately plotted graphs. Axis labels and keys on graphs and diagrams are essential if the data presented is to be understood, and these were not always present. Most candidates appeared to have sufficient time to complete the paper.

A small number of candidates offered more than one solution to a question, without making it clear which solution they wanted to be marked.

Comments on specific questions

Question 1

The first two parts of **Question 1** were answered well. In **part (a)**, a small number of candidates gave an incomplete description, omitting 'continuous' and giving only 'quantitative' as their description for the distance travelled, but most candidates scored full marks.

In **part (b)**, a small number of candidates found the range, rather than the interquartile range, or misread the scale when finding the lower quartile, but the majority found the interquartile range correctly.

Most candidates found **part (c)** more difficult, with the difference between the highest value and the upper quartile, namely 28 – 15 = 13, being a common wrong answer. Of those that got the answer correct, most found ¾ of 144 first, and then subtracted from 144, rather than simply finding ¼ of 144.

Question 2

The best answers to **part (a)** usually came from candidates giving a sectional bar chart as their alternative chart and providing a reason in the context of finding the most popular sport from observing the tallest bar. Reasons such as 'it will be easier to draw' or 'it will show the results clearly' were quite common and did not score the mark. Some candidates gave a percentage sectional bar chart as their chosen alternative, but this was only acceptable if they made it clear that it was to be a single bar combining the data from the boys and the girls, divided into the percentage of the children who preferred each of the three sports.

Most candidates found the correct probability in **part (b)**, but the probabilities in **parts (c)** and **(d)** were sometimes switched, or one or both given as 3/25.

Question 3

Fully correct solutions were often seen to **part (a)**, but some candidates omitted to conclude that the sample was not representative, having provided all the correct initial calculations. A common incorrect answer involved finding that there were 2 of each battery size in the given sample and concluding, without further calculations, that therefore the sample was representative or in some cases crossing out correct calculations leading to the representative strata sizes of 1, 3 and 2.

In **part (b)** candidates needed to make sensible rounding choices to find the strata sizes for a sample of size 5. Most candidates showed correct initial calculations, but many were unable to round appropriately, with common wrong answers being 0, 3, 2 or 1, 3, 1. Some candidates simply left the strata sizes as 1, 3, 2, without attempting to find the best possible sample of size 5.

Question 4

In this probability question, it appeared that the use of the words 'or' and 'and' rather than union and intersection symbols caused confusion for some candidates. Others did not use the fact that the events were independent. Common incorrect answers in **part (a)** came from simply doing 0.3 × 0.4 or from 0.3 + 0.4.

The correct answer for **part (a)** was often seen as the answer to **part (b)**.

Successful attempts at **part (c)** came either from using **part (a)**, and doing 1 – 0.58, or, more often, from doing (1 – 0.3) × (1 – 0.4). A common incorrect method seen was 1 – 0.3 × 0.4.

Question 5

In **part (a)**, most candidates were able to find the total amount spent on rent and on electricity last year before simplifying the ratio of the three categories. The most common error was for the rent to be left as the monthly figure rather than the annual amount.

Part (b) of this question was less structured than similar questions in the past and it was pleasing to see many candidates providing well-explained solutions, with some providing tables showing the price relatives and corresponding weights. A small number used a price relative of 102 rather than 98 for the other costs. Those that were unable to calculate the price relatives correctly were usually able to use their values in an appropriate formula for the weighted aggregate index and so most candidates achieved some, if not all, of the available marks.

Candidates found **part (c)** difficult, with many incorrectly suggesting reasons that are already accounted for in the price relatives, such as an increase in the cost of electricity or rent. Those that were successful usually gave a change in the number of units of electricity used as the explanation. This question was left blank by some candidates.

Question 6

This was a challenging question on linear interpolation. In **part (a)** it was quite common to see final answers that were outside of the range 600 g to 1000 g; candidates should have questioned such answers. Fully correct solutions were however seen, and these answers were usually rounded as instructed in the question.

There were only a small number of correct answers to **part (b)** with many candidates leaving this answer blank or stating that they had assumed that the median was the 50th percentile or that 62 per cent of the parcels had masses between 600 g and 1000 g. Correct answers explaining that the assumption was that the masses of the parcels increase linearly were quite rare.

Question 7

In **part (a)** it was pleasing to see many completed stem-and-leaf diagrams with correctly ordered leaves. A few candidates did not attempt to line up the leaves appropriately, giving a distorted visual impression of the data. The most common cause of a loss of marks was for a missing or incomplete key. Incomplete keys sometimes only explained a female value on the right of the diagram, for example. Others, who attempted a key to explain values on the right and the left of the diagram, described the values of 33 and 30, say, as representing numbers of females and males rather than ages of a female and a male.

Many fully correct answers were seen in **part (b)**. The most common errors came from locating the median, say, by finding the 7.5th value rather than the 8th value.

In **part (c)**, the instruction in the question required the use of values from the table to justify the answer. Thus, in assessing Azeeb's claim, it was necessary to compare the median values and, in assessing Tebogo's claim, a comparison of the interquartile ranges was required. It is important that candidates read instructions such as this carefully.

Many fully correct answers to **part (d)** were seen, with the most common errors coming from putting the value 30, for example, into the $20 \leqslant x < 30$ class rather than the $30 \leqslant x < 40$ class.

Frequency polygons in **part (e)** usually had an appropriate key, but labels on the axes were often missing and upper class boundaries, rather than midpoints, were sometimes used.

Many correct advantages of a stem-and-leaf diagram over a frequency polygon were seen in **part (f)**, with reference made to the preservation of the original values. A few comments were too vague, such as 'it is easily interpreted', without an explanation as what it is that makes the stem-and-leaf easily interpreted.

Part (g) of this question proved very challenging for most candidates. A common incorrect answer seen in **part (i)** was 6, coming from 54 – 48. Of those that scored the mark in **part (i)**, very few were able to correctly answer **part (ii)**.

Question 8

Many candidates wrote down two correct reasons for finding moving average values in **part (a)**, with their use in finding the trend being the most common of the correct reasons given. A few candidates appeared to have misread the question and described reasons for centring the moving average values.

Most candidates correctly completed the table in **part (b)**, with the small number of errors that did occur appearing to be the result of an arithmetic error rather than a misunderstanding of the statistics.

Many fully correct calculations of the seasonal component were seen in **part (c)**. A few candidates performed the required subtractions the wrong way around and ended up with a positive seasonal component for quarter 2. Checking the original data should have alerted these candidates to their error. Some candidates tried to use the 2019 quarter 2 value in their calculation, but this value did not have a corresponding centred moving average value. Other errors seen were the use of the moving average values rather than the centred moving average values, so, for example, 332 – 1307.25 or 320 – 1294 seen as part of the calculation. Weaker candidates gave an answer of 326.3 from finding the mean of the three quarter 2 numbers of swans.

In **part (d)** most candidates plotted the centred moving average values accurately and drew an appropriate trend line. The most common plotting error came from the plot for 1297.75 being seen at approximately 1299.75, suggesting a misread of the vertical scale. It was acceptable for candidates to either include or ignore the first of the centred moving average values when it came to drawing the trend line, as this first value did not closely follow the pattern of the remaining ones. Joining the first plot and the last plot did not, however, form an appropriate trend line.

Most candidates did not comment on the fact, in **part (e)**, that the ranger's assistant had referred to 'each quarter'. Thus, the fact that the ranger's assistant was not correct, because the trend line shows a fall over time or each year, but not a fall each quarter, was missed by most candidates.

Accurate readings from the trend line in **part (f)** were usually seen, but some candidates gave this as the final answer without introducing the seasonal component. As the question requires an estimate of the number of swans, the final answer should have been given as a whole number.

Question 9

Many candidates correctly answered the first three questions in **part (a)**, but it was common to see errors in **part (iv)**. Most candidates, rather than use the previous answers, recalculated the answers to **parts (i)** and **(ii)** and added these, without subtracting the answer to **part (iii)**. Weaker candidates seemed to misunderstand the context of the problem, treating the question as if it involved one selection from six cards, rather than two independent selections, each from three cards.

Although few fully correct solutions were seen in **part (b)**, many candidates demonstrated some understanding of the requirements of the question. It was quite common to see the complement of their first probability used in an equation of the correct form. Some solutions would have benefitted from clearer explanations, such as labelling of the probabilities to indicate which represented the probability of identical cards and which represented the probability of non-identical cards.

The set up in **part (c)** seemed to be more easily understood than that in **part (a)**. Many candidates dealt correctly with the fact that the cards were not replaced between selections. Most also realised that in **part (i)**, for example, they needed to consider both square then triangle and triangle then square. Many candidates did not notice that the three options presented in **parts (i)** to **(iii)** covered all the options and therefore the answer to **part (iii)** could be simply obtained from the previous parts.

Most candidates understood the requirement for a fair game in **part (d)**, using their probabilities from **part (c)** correctly. Some weaker candidates simply added the prize values together.

Question 10

Most candidates took the correct approach to compare the scores in **part (a)** by attempting to scale or standardise them. Often this was done correctly with a correct conclusion, but sometimes the conclusion was incorrect and sometimes the subtraction had been done the wrong way around.

In **part (b)** many correct calculations of the mean were seen, but it was much less common to see a correct calculation of the standard deviation. The most common error when calculating the standard deviation was to use $9858^2 + 9588^2$ as the value for the sum of the squares. Some weaker candidates ignored the different number of candidates in each year group and simply found the mean of 62 and 68 and the mean of 8 and 10.

Candidates tended to be more successful with **part (c)**, and clearly set out solutions allowed for the award of part-marks when solutions were incomplete. Many, for example, got as far as finding the combined total score for the two year groups, even if they could progress no further.

In **part (d)**, most candidates correctly increased the mean by 10 per cent, but many did not realise that increasing each candidate's score by 10 per cent would also increase the standard deviation by 10 per cent.

In **part (e)**, many candidates realised that the mean would increase, but it was quite rare to see the effect on the standard deviation correct, with many thinking that it would stay the same.