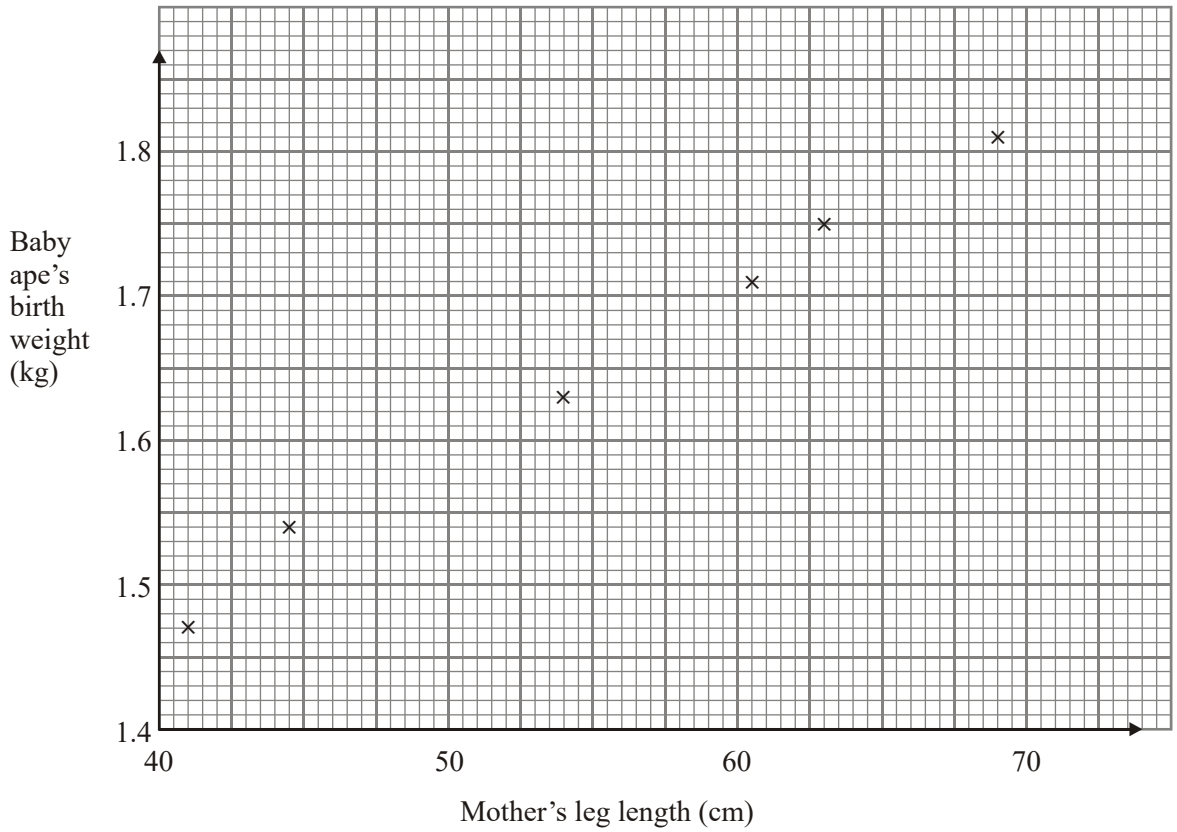
1. The scatter graph shows some information about six new-born baby apes.

For each baby ape, it shows the mother's leg length and the baby ape's birth weight.



The table shows the mother's leg length and the birth weight of two more baby apes.

Mother's leg length (cm)	50	65
Baby ape's birth weight (kg)	1.6	1.75

(a) On the scatter graph, plot the information from the table.

(1)

(b) Describe the **correlation** between a mother's leg length and her baby ape's birth weight.

……

(1)

(c) Draw a line of best fit on the diagram.

(1)

A mother's leg length is 55 cm.

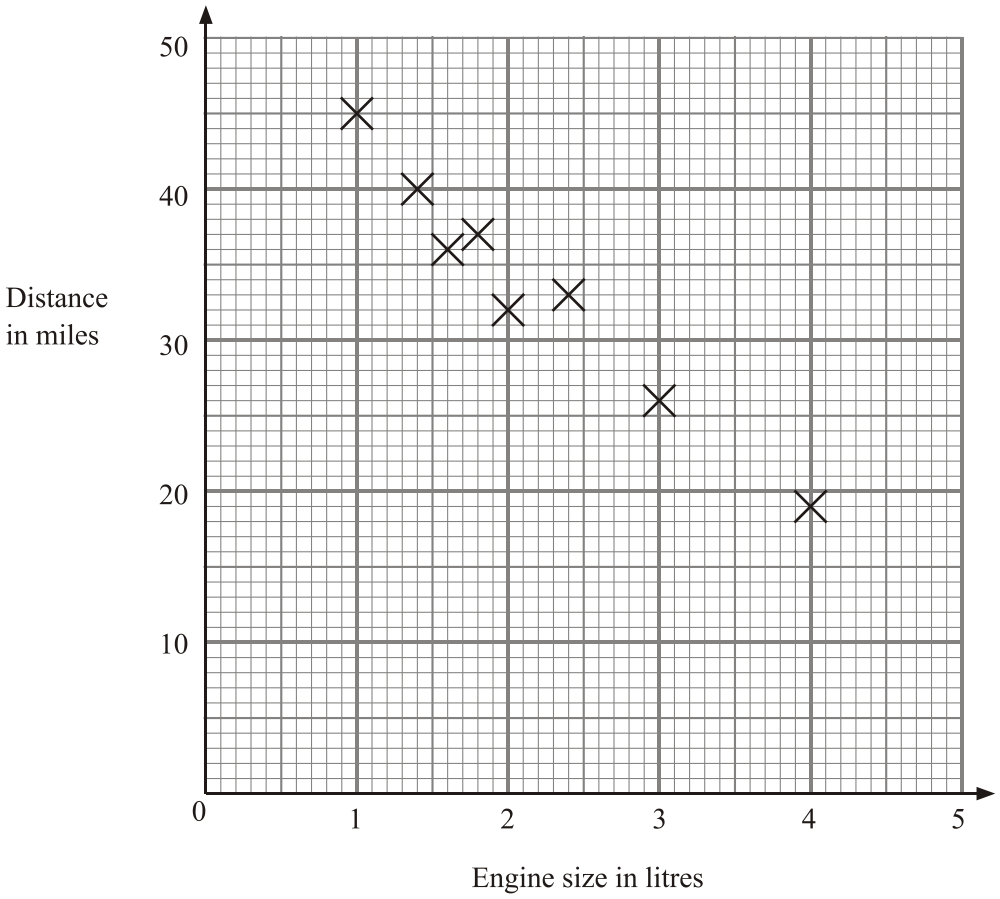
(d) Use your line of best fit to estimate the birth weight of her baby ape.

……kg

(1)

(Total 4 marks)

2. The scatter graph shows some information about 8 cars.
For each car it shows the engine size, in litres, and the distance, in miles, it travels on one gallon of petrol.



(a) What type of correlation does this scatter graph show?

.....

(1)

(b) Draw a line of best fit on the scatter graph.

(1)

(c) Use your line of best fit to estimate

(i) the distance travelled on one gallon of petrol by a car with an engine size of 3.5 litres,

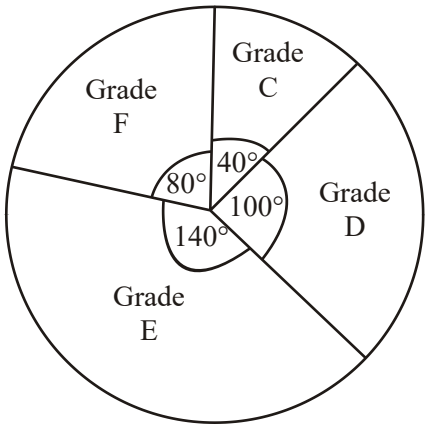
..... miles

(ii) the engine size of a car that travels a distance of 28 miles on one gallon of petrol.

..... litres

(2)
(Total 4 marks)

3. The pie chart gives information about the mathematics exam grades of some students.



Mathematics exam grades

Diagram **NOT** accurately drawn

(a) What grade was the mode?

.....

(1)

(b) What fraction of the students got grade D?

.....

(1)

8 of the students got grade C.

(c) (i) How many of the students got grade F?

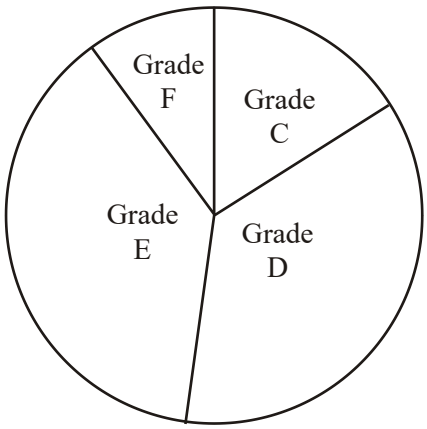
.....

(ii) How many students took the exam?

.....

(3)

This accurate pie chart gives information about the English exam grades for a different set of students.



English exam grades

Sean says "More students got a grade D in English than in mathematics."

(d) Sean could be **wrong**.
Explain why.

.....

.....

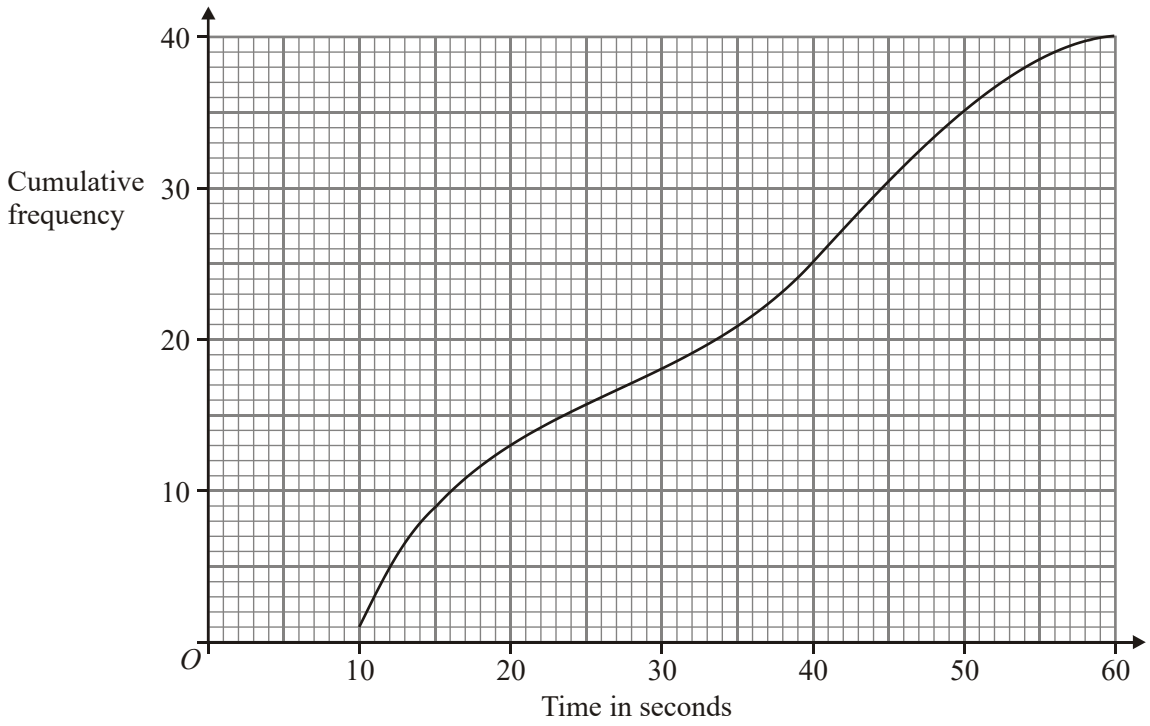
(1)
(Total 6 marks)

4. 40 boys each completed a puzzle.
The cumulative frequency graph below gives information about the times it took them to complete the puzzle.

(a) Use the graph to find an estimate for the median time

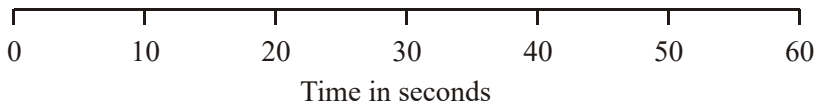
..... seconds

(1)



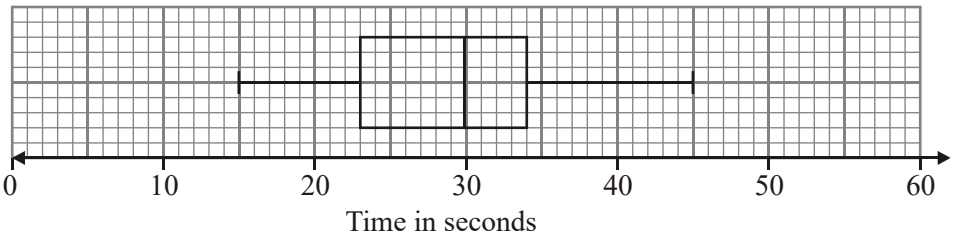
For the boys
the minimum time to complete the puzzle was 9 seconds
and the maximum time to complete the puzzle was 57 seconds.

(b) Use this information and the cumulative frequency graph to draw a box plot showing information about the boy's times.



(3)

The box plot below shows information about the times taken by 40 girls to complete the same puzzle.



(c) Make **two** comparisons between the boys' times and the girls' times.

.....

.....

(2)
(Total 6 marks)

5. Daniel took a sample of 100 pebbles from Tawny Beach.
He weighed each pebble and recorded its weight.
He used the information to draw the cumulative frequency graph shown on the grid.

(a) Use the cumulative frequency graph to find an estimate for

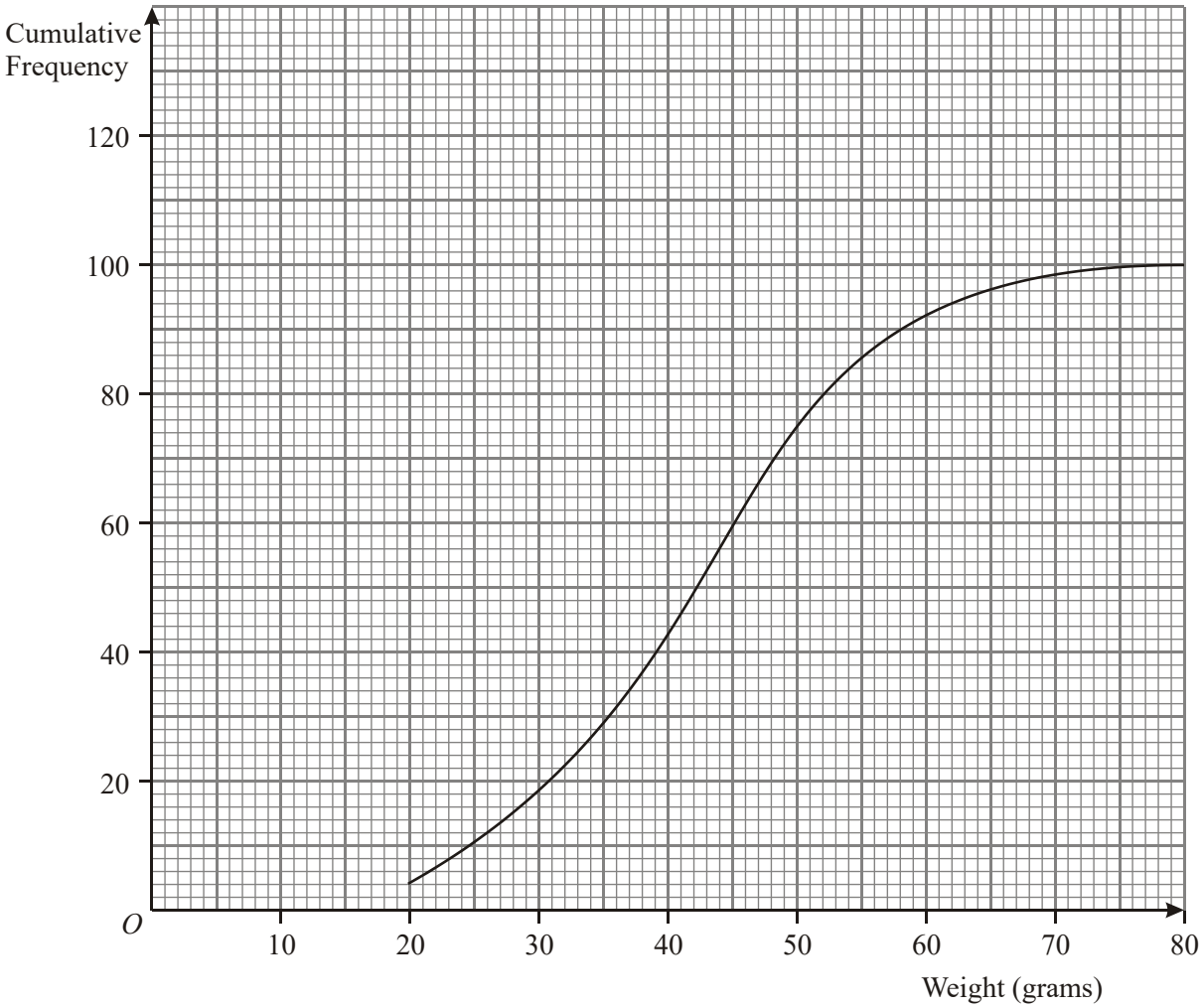
(i) the median weight of these pebbles,

..... grams

(ii) the number of pebbles with a weight more than 60 grams.

.....

(3)



Daniel also took a sample of 100 pebbles from Golden Beach.
The table shows the distribution of the weights of the pebbles in the sample from Golden Beach.

Weight (w grams)	Cumulative frequency
$0 < w \leq 20$	1
$0 < w \leq 30$	15
$0 < w \leq 40$	36
$0 < w \leq 50$	65
$0 < w \leq 60$	84
$0 < w \leq 70$	94
$0 < w \leq 80$	100

(b) On the same grid, draw the cumulative frequency graph for the information shown in the table.

(2)

Daniel takes one pebble, at random, from his sample from Tawny Beach and one pebble, at random, from his sample from Golden Beach.

(c) Work out the probability that the weight of the pebble from Tawny Beach is more than 60 grams **and** the weight of the pebble from Golden Beach is more than 60 grams.

.....

(4)
(Total 9 marks)

01. (a) Points plotted 1
B1 ± 1 full (2 mm) square

(b) positive 1
B1 cao

(c) Line of best fit 1
B1 Must pass through (42.5, 1.45), (42.5, 1.55) AND (67.5, 1.75), (67.5, 1.85)

(d) ~1.65 1
B1 ft from single line segment with positive gradient ± 1 full (2 mm) square

[4]

02. (a) negative 1
B1 cao

(b) line of best fit 1
B1 straight line passing between (4, 15) and (4, 20) and between (1, 40) and (1, 45)

(c) (i) ~22 2

B1 ft from single line segment with negative gradient ± 1 full (2mm) square

(ii) ~2.8

B1 ft from single line segment with negative gradient ± 1 full (2mm) square

[4]

03. (a) Grade E 1

B1 for E, e Grade E, e, or 140°

(b) 100/360 1

B1 5/18 oe

(c) (i) 8 × 2 = 16 3

B1 cao

(ii) 360/40 × 8 = 72

M1 360/40 × 8 oe, or 360/80 × "16" oe, or "16" × 4.5 or attempts to find an association eg 8 + 16 + 20 + 28
A1 cao or ft from (i)

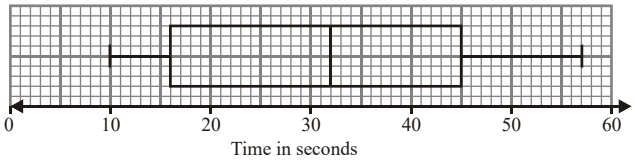
(d) Reason 1

B1 reason (eg %, not actual numbers; do not know how many students, etc)

[6]

04. (a) 32 1

B1 for 32 (accept 31.5 to 33.5 inclusive)

(b) 

3

B1 for ends of whiskers at 9 and 57 (with a box)
B1 for ends of box at 16 and 45/46 (± 0.5)
B1for median marked at "32"or complete box and whisker diagram drawn with a median inside the box

(c) Median(B)>Median(G); on average boys take longer
IQR(B)>IQR(G); times for boys have a greater spread 2

B1 eg for comparison of medians (ft on diagrams)
B1 eg for comparison of (interquartile) ranges (ft on diagram)

[6]

05. (a) 42g
8 3

Median at 50.5 (50)
100 – 92

B1 for 42g to 43g

M1 for reading correctly from graph ± $\frac{1}{2}$ sq and subtracting from 100
A1 for 7, 8 or 9

(b) cf 2

B1 for plots (condone one error) ± $\frac{1}{2}$ sq

B1 (dep) for joining points to give cf graph
SC: B1 if points plotted consistenly within intervals (condone one error) and joined

(c) 0.0128 4

100 – 84 = 16
0.16
0.08 × 0.16

B1 for $\frac{"8"}{100}$ oe (Tawny Beach)

B1 for $\frac{15}{100}$ or $\frac{16}{100}$ or $\frac{17}{100}$ oe (Golden Beach)

M1 for multiplying two probabilities
A1 ft (dep on B2)

[9]

01. Specification A

Foundation Tier

A large number of candidates plotted the points accurately in part (a). If one error were made, it was usually with (65, 1.75). In part (b), candidates were more likely to describe the relationship than the correlation; this received no credit. The term "positive" was required. In part (c), there were many lines of best fit which were within the permitted tolerance but there were also many curves and zig-zags joining all the points. Although, in this case, an acceptable line of best fit could pass through the origin, candidates should not assume that a line of best must do this.

Intermediate Tier

This as a well answered question. Most were able to plot the points correctly, the most common error caused by plotting (65, 1.75) incorrectly. In part (b) some candidates described a relationship rather than the correlation. The quality of the lines of best fit was very good this year, with far fewer attempting to connect the points together, though a significant minority continue to want to draw the line from the origin. Most candidates could also read off from their line quite accurately in part (d), the only common error was reading off from 52.5 rather than 55.

Specification B

The two points were usually plotted accurately although it was not uncommon to see the point (65, 1.75) plotted at the point (62.5, 1.75) in part (a).
In part (b) many candidates described the relationship between the two variables instead of commenting on the correlation.
Lines of best fit were usually within guidelines. Lines starting at the point (40, 1.4) was the major reason for loss of the mark in part (c). There was evidence in this part of the question that some candidates did not have access to a ruler.
Use of the line of best fit was generally good, however it was not uncommon for a leg length of 52.5 cm to be used instead of 55 cm, showing a lack of care with the scale on the horizontal axis.

02. This question was answered well. The majority of candidates showed awareness that 'positive' or 'negative' is required when describing correlation and that a straight line is required when drawing the line of best fit. The most common error in part (a) was to identify the correlation as positive. In part (b), the lines of best fit were generally well drawn with a ruler and within the bounds required. Some were outside the bounds because candidates drew the line to go through one of the corners of the grid. Most candidates were able to read accurately from their line of best fit in part (c). Some misread the vertical scale, giving, for example, 20.4 instead of 24

03. Part (a) was well answered, but few candidates gained the mark in part (b). Many attempted to estimate the fraction of the diagram, hence many gave $\frac{1}{4}$ or $\frac{1}{3}$ as the answer. Of those who used the 100°, the error for many was in giving it out of a number other than 360°.

In part (c) most candidates gained some credit, sometimes by showing evidence of using inventive methods. Some found and used a scaling factor such as 4.5. Others found an association using the relationship of the angles, showing 8 + 16 + 20 + 28, or equivalent methods.

Part (d) was a discriminator, and it was encouraging to find half the candidates were able to distinguish between proportion and actual values, giving an acceptable explanation why Sean was wrong.

04. **Mathematics A**

Paper 3

It was disappointing that less than half of the candidates estimated the median time correctly in part (a). Most candidates were able to draw a recognisable box plot in part (b) and many drew correct whiskers with the ends at 9 and 57. Candidates were much less successful, though, at finding the quartiles from the cumulative frequency graph and positioning the ends of the box at 16 and 45. Unfortunately, many of the comparisons given in part (c) only referred to the times of individuals (e.g. the fastest boy and the fastest girl) and did not make use of the median and spread to compare the distributions. Those who used the word ‘median’ often gave a correct interpretation.

Paper 5

This question on box plots was answered much better than the corresponding question last summer. Most grade B and above candidates correctly obtained an estimate for the median although some wrong reading of scales (30.2) was seen. Again in part (b) many gained at least partial credit (normally at least 2 out of 3) with a common wrong answer of plots at 10 and 60 for the ends of the whiskers. In part (c) the examiners were looking for two different comparisons between the times of boys and the times of girls rather than anything which related to single individuals. Most grade B and above candidates gained at least one of the two marks. Those who gained just one mark frequently gave ‘two’ comparisons which effectively were the same as illustrated by “the interquartile range of the boys’ times is greater”, “the girls’ times are closer together”.

Mathematics B Paper 16

Part (a) was answered correctly by 42% of the candidates.
The concept of box and whisker diagrams does now seem to be well known and understood by many, 78% gaining at least one mark, usually for the median. Many candidates confused the maximum and minimum values with the quartiles.
In part (c) very many candidates quoted exact figures from the box plots without any attempt at comparison between the boys and girls. Medians in “similar positions” or reference to means gained no marks.

05. Parts (a) and (b) were well done by the majority of candidates. In part (c) most candidates were able to write down the relevant probabilities correctly but these were then frequently added rather than multiplied. A common arithmetic error in this question was to give the answer to 100×100 as 1000.