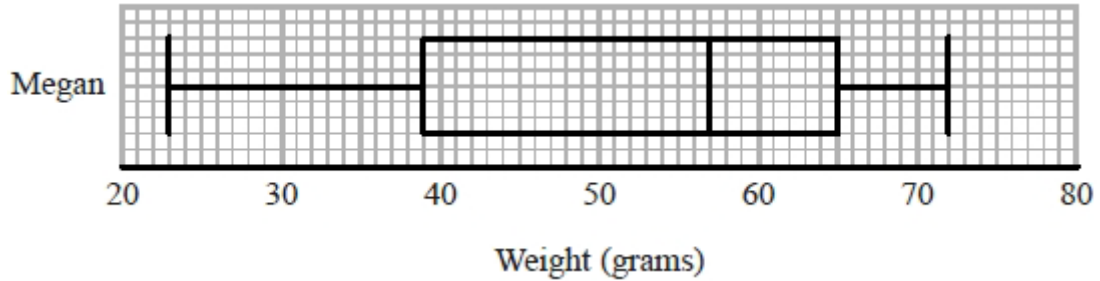## Questions

Q1.

Megan grows potatoes.

The box plot below shows information about the weights of Megan's potatoes.

Weight (grams)

Megan says that half of her potatoes weigh less than 50 grams each.

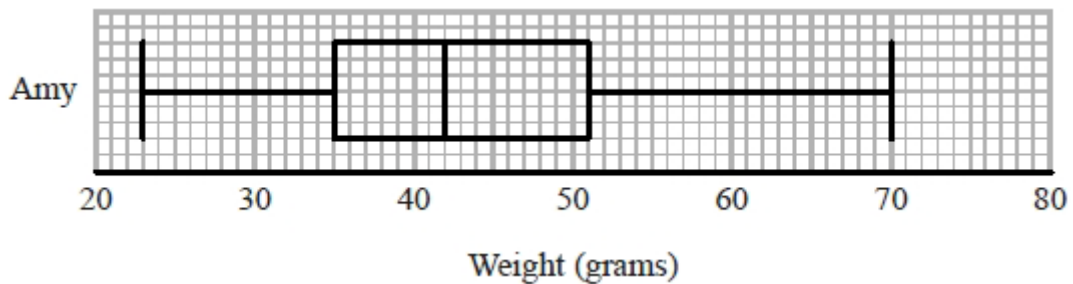(a) Is Megan correct?

Give a reason for your answer.

....

....

....

(1)

Amy also grows potatoes.

The box plot below shows information about the weights of Amy's potatoes.

Weight (grams)

(b) Compare the distribution of the weights of Megan's potatoes with the distribution of the weights of Amy's potatoes.
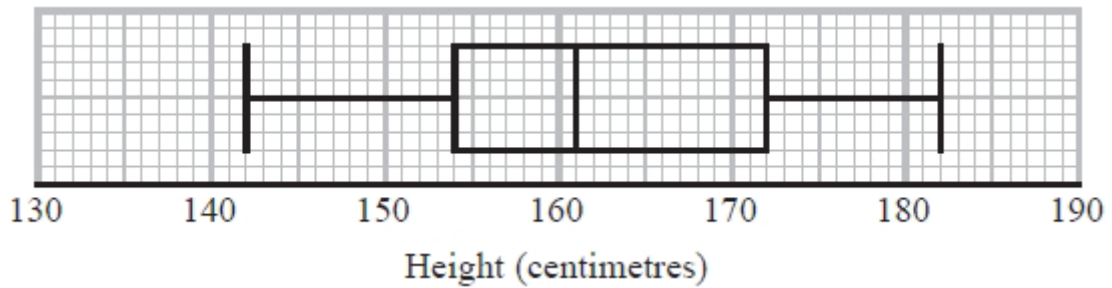
....

....

....

....

....

(2)

(Total for question = 3 marks)

Q2.

Aisha recorded the heights, in centimetres, of some girls.
She used her results to work out the information in this table.

Least height	142 cm
Lower quartile	154 cm
Interquartile range	17 cm
Median	162 cm
Range	40 cm

Aisha drew this box plot for the information in the table.
The box plot is **not** fully correct.

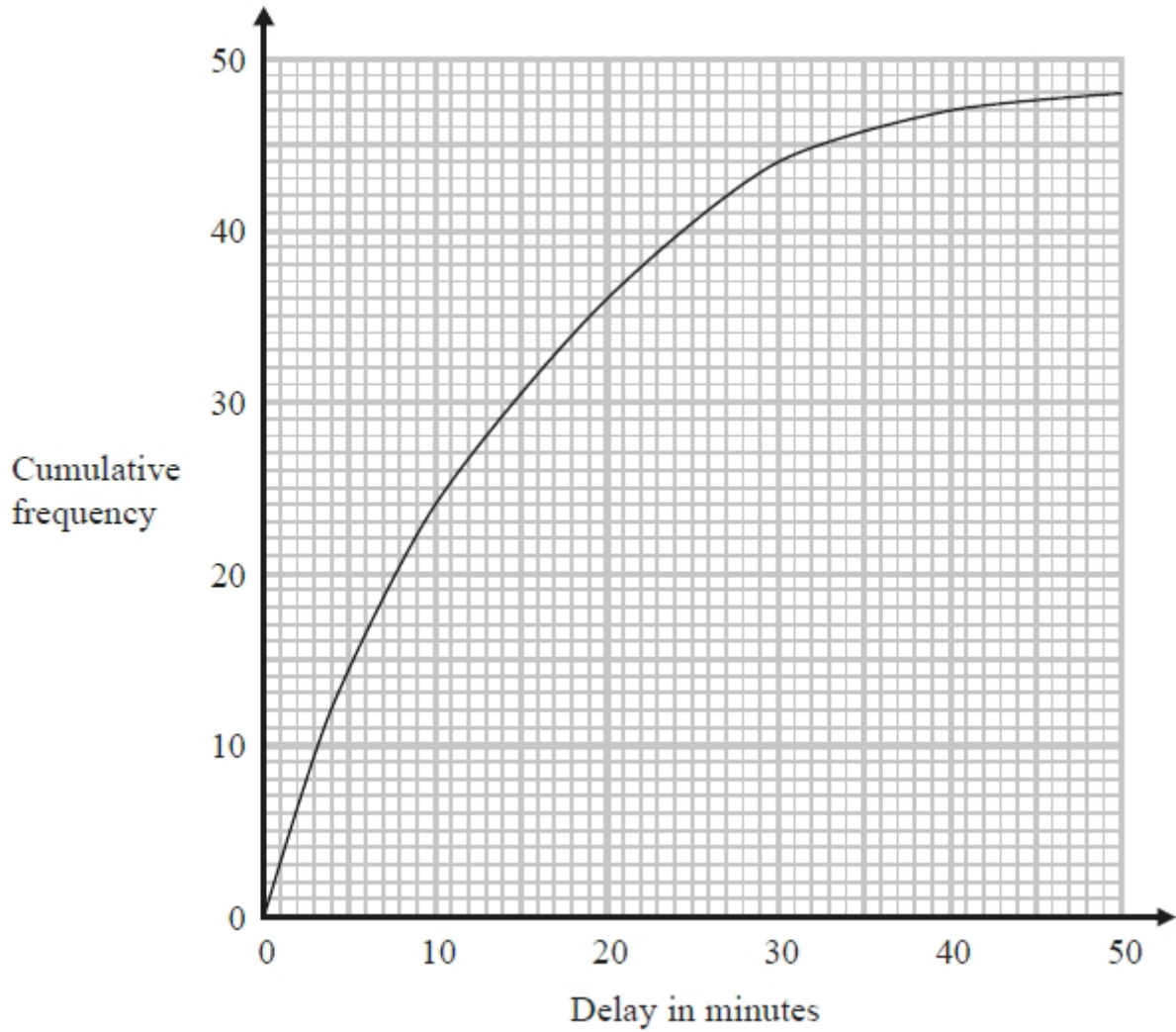Write down the two things Aisha should do to make the box plot fully correct.

1 ....

....

....

2 ....

....

....

**(Total for question = 2 marks)**

Q3.
The times that 48 trains left a station on Monday were recorded.

The cumulative frequency graph gives information about the numbers of minutes the trains were delayed, correct to the nearest minute.

The shortest delay was 0 minutes.
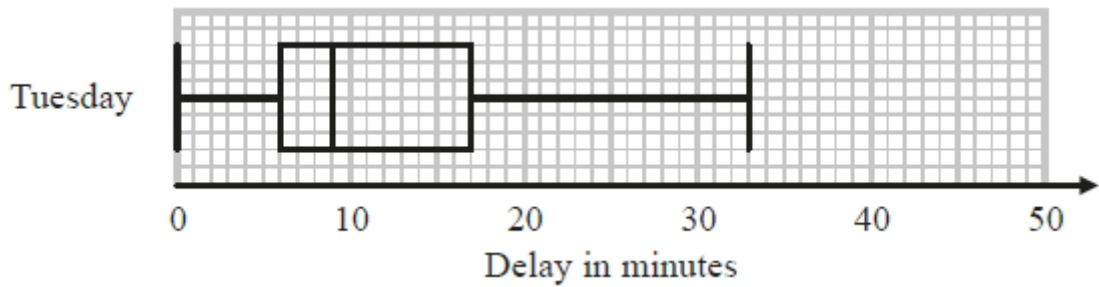The longest delay was 42 minutes.

(a) On the grid below, draw a box plot for the information about the delays on Monday.

Monday

0   10   20   30   40   50
Delay in minutes
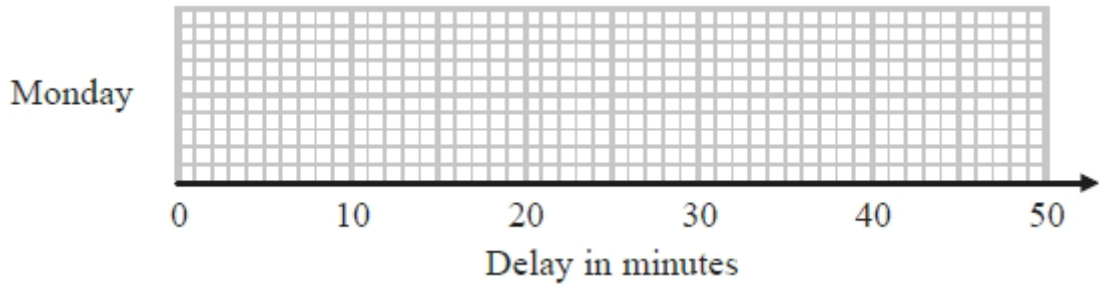
(3)

48 trains left the station on Tuesday.
The box plot below gives information about the delays on Tuesday.

Tuesday

0   10   20   30   40   50
Delay in minutes

(b) Compare the distribution of the delays on Monday with the distribution of the delays on Tuesday.

....

....

....

....

(2)

Mary says,

"The longest delay on Tuesday was 33 minutes.
This means that there must be some delays of between 25 minutes and 30 minutes."

(c) Is Mary right?
You must give a reason for your answer.

....

....

(1)

(Total for question = 6 marks)

Q4.
The box plot shows information about the length of time, in minutes, some people waited to see a doctor at a hospital on Monday.

Waiting time (minutes)
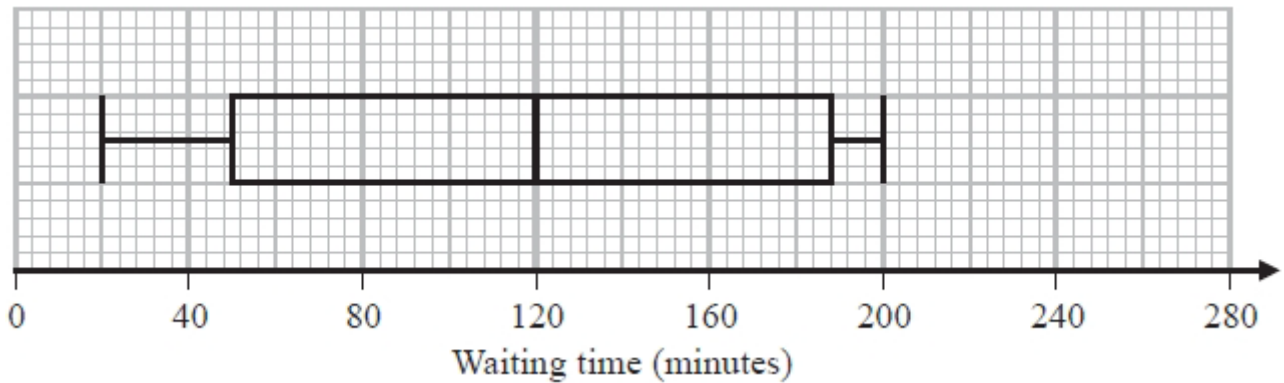
(a) Work out the interquartile range of the information in the box plot.

.... minutes

(2)

Becky says,

"50% of the people waited for at least 2 hours."

(b) Is Becky correct?

Explain why.

....

....

....

(1)

The table gives information about the length of time, in minutes, some people waited to see a doctor at the same hospital on Tuesday.

	Waiting time (minutes)
Shortest time	20
Lower quartile	50
Median	100
Upper quartile	140
Longest time	210

Becky was asked to compare the distribution of the lengths of times people waited on Monday with the distribution of the lengths of times people waited on Tuesday.

She wrote,

"People had to wait longer on Tuesday than on Monday."

(c) Give **one** reason why Becky may be wrong.

....

....

....

(1)

(Total for question = 4 marks)

Q5.

*Some students were asked how many times they each used their mobile phones last week.

The box plots give information about the male students' answers and about the female students' answers.
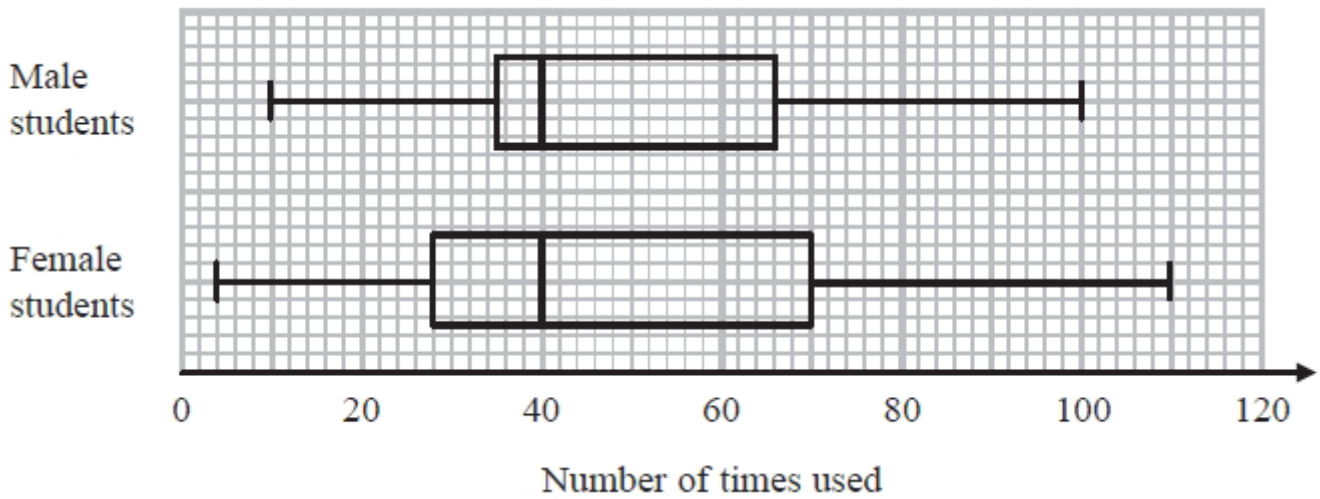
Compare the two distributions represented by the box plots.

....

....

....

....

(Total for question = 3 marks)

Q6.

Presta recorded the number of lemons on each of 60 lemon trees.

The incomplete table and box plot give information about her results.

	Number of lemons
Smallest number	
Lower quartile	8
Median	11
Upper quartile	
Greatest number	22
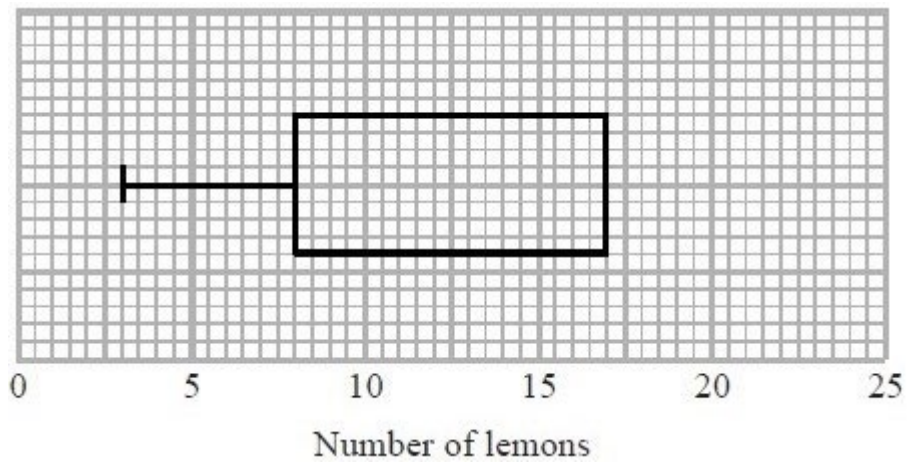
Number of lemons

(a) (i) Use the information in the table to complete the box plot.

(ii) Use the information in the box plot to complete the table.

(3)

Some of these 60 lemon trees have 8 or more lemons on them.

(b) Find an estimate for the number of lemon trees with 8 or more lemons on them.

....

(2)

(Total for Question is 5 marks)

Q7.

The table below shows information about the times, in minutes, a group of students took to answer 10 maths questions.

	Least	Lower quartile	Median	Upper quartile	Greatest
Time in minutes	14	18	20	25	30

On the grid below, draw a box plot to show the information in the table.

0   5   10   15   20   25   30   35
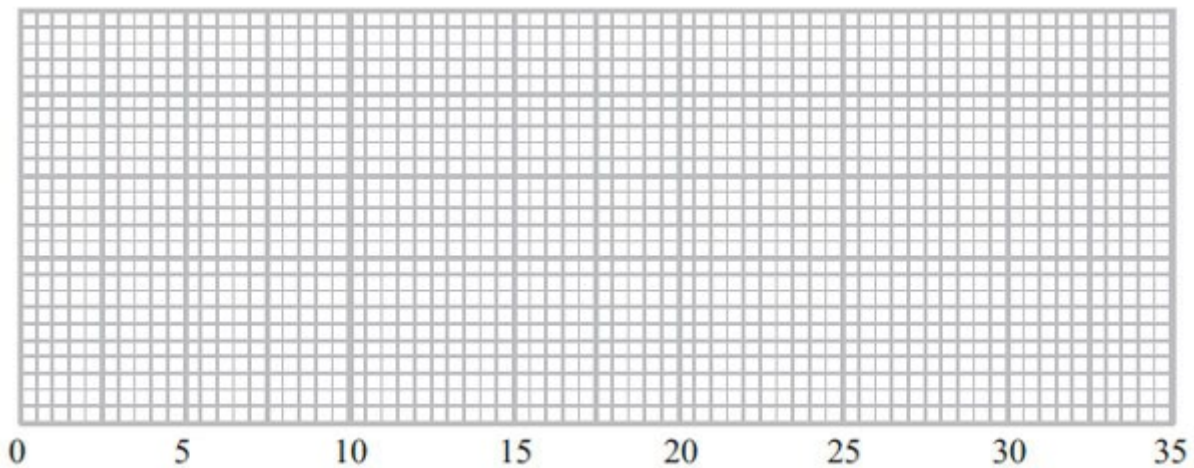
(Total for Question is 3 marks)

Q8.

The table gives some information about the heights of 80 girls.

Least height	133 cm
Greatest height	170 cm
Lower quartile	145 cm
Upper quartile	157 cm
Median	151 cm

(a) Draw a box plot to represent this information.

120   130   140   150   160   170   180

height (cm)

(b)

(3)
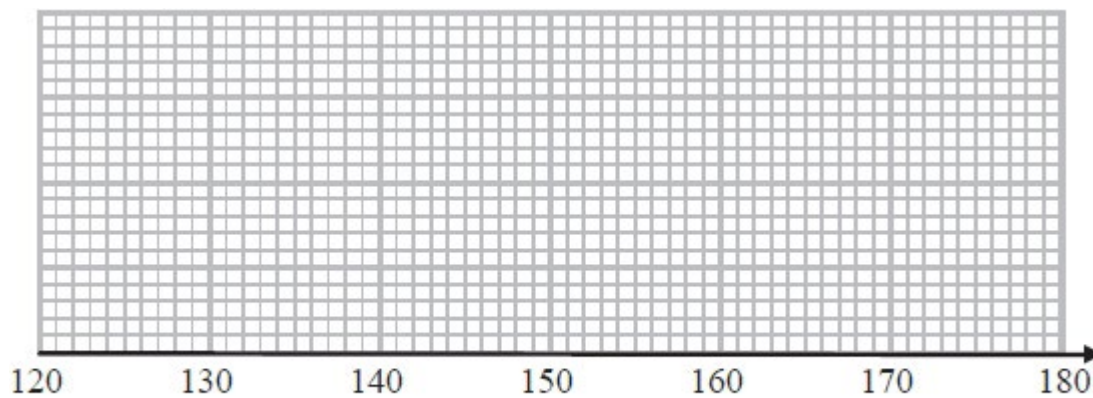
(b) Work out an estimate for the number of these girls with a height between 133 cm and 157 cm.

....

(Total for question = 5 marks)

Q9.
The table gives information about the weights of 60 parcels.

Lowest	1.5 kg
Highest	6.3 kg
Lower quartile	2.7 kg
Interquartile range	1.7 kg
Median	3.2 kg

Draw a box plot for this information.

Weight (kg)
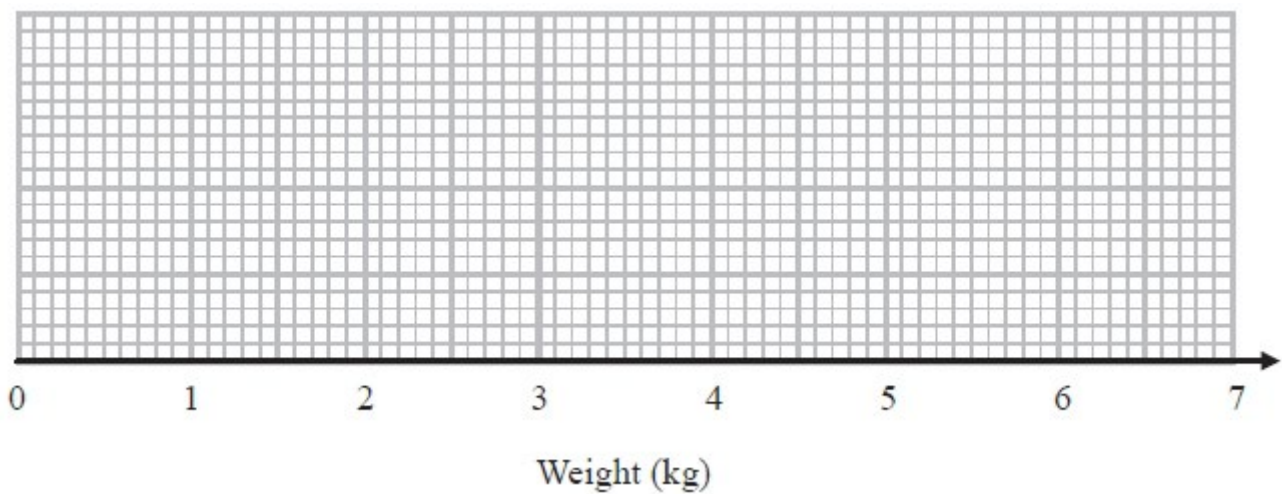
(Total for question = 3 marks)

Q10.
Here are the ages, in years, of 15 women at West Ribble Tennis Club.

16, 18, 18, 20, 25, 25, 27, 28, 30, 35, 38, 42, 45, 46, 50

(a) On the grid, draw a box plot for this information.

age (years)

(3)
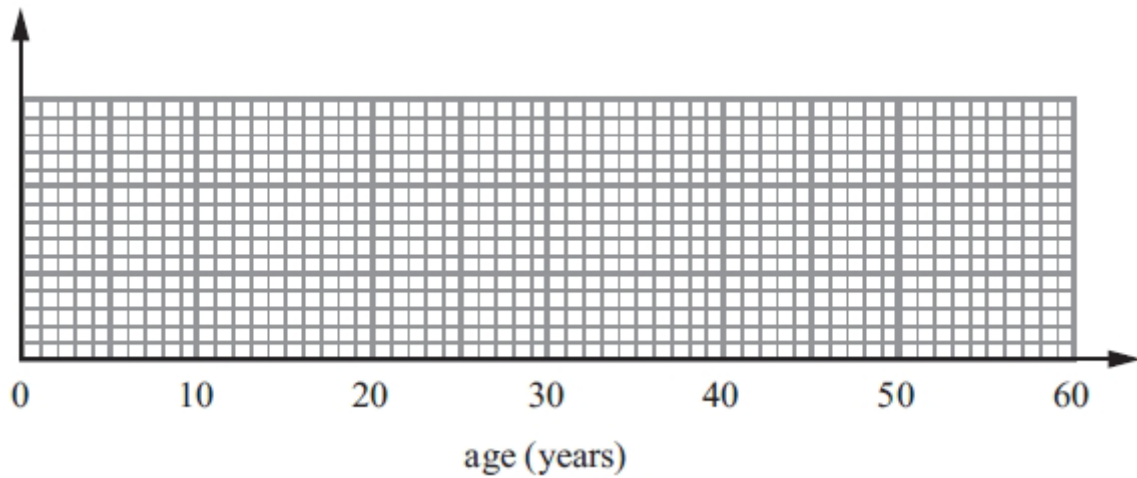
The box plot below shows the distribution of the ages of the men at West Ribble Tennis Club.

age (years)
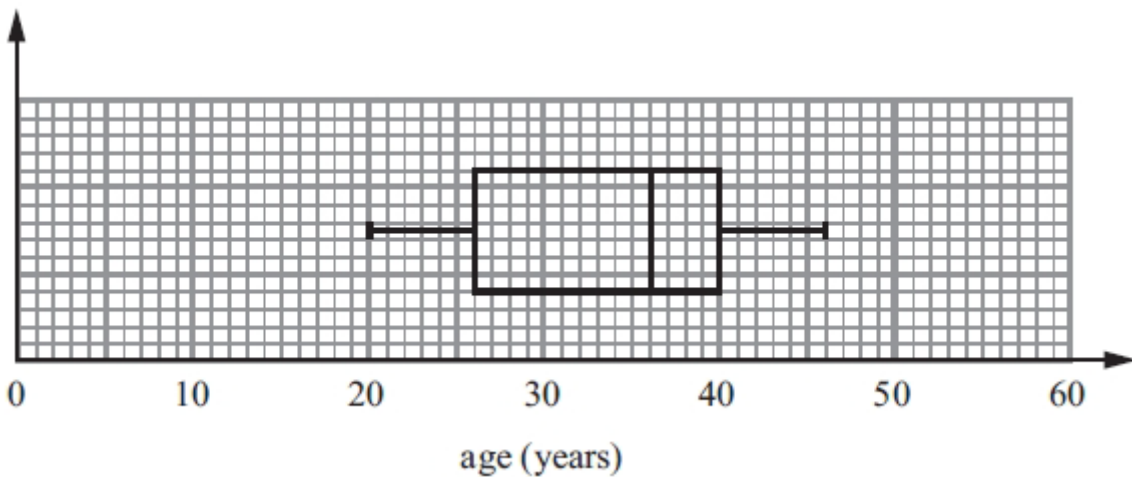
*(b) Use the box plots to compare the distributions of the ages of these women and the distributions of the ages of these men.

....

....

....

....

....

(2)

(Total for Question is 5 marks)

Q11.

Some girls did a sponsored swim to raise money for charity.

The table shows information about the amounts of money (£) the girls raised.

Least amount of money (£)	10
Greatest amount of money (£)	45
Median	25
Lower quartile	16
Upper quartile	42

(a) On the grid, draw a box plot for the information in the table.
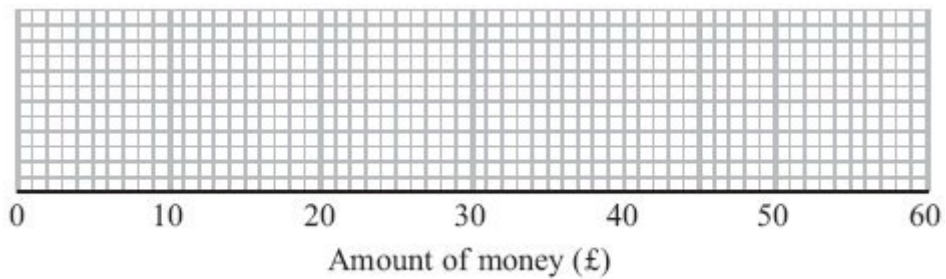
Amount of money (£)

(2)

Some boys also did the sponsored swim.

The box plot shows information about the amounts of money (£) the boys raised.
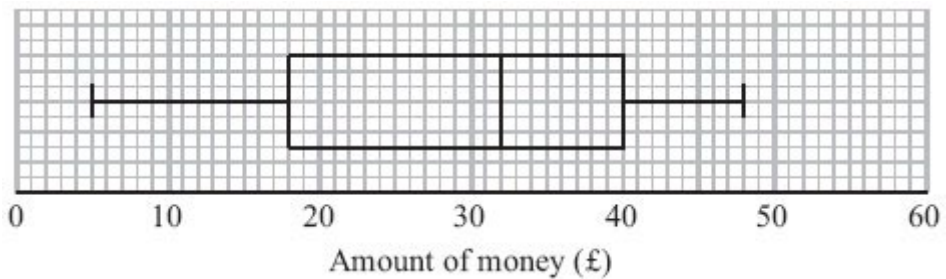
Amount of money (£)

(b) Compare the amounts of money the girls raised with the amounts of money the boys raised.

....

....

....

....

(2)

(Total for Question is 4 marks)

Q12.

Here are the times, in seconds, that 15 people waited to be served at Rose's garden centre.

5   9   11   14   15   20   22   25   27   27   20   22   25   27

(a) On the grid, draw a box plot for this information.

Time (seconds)

(3)
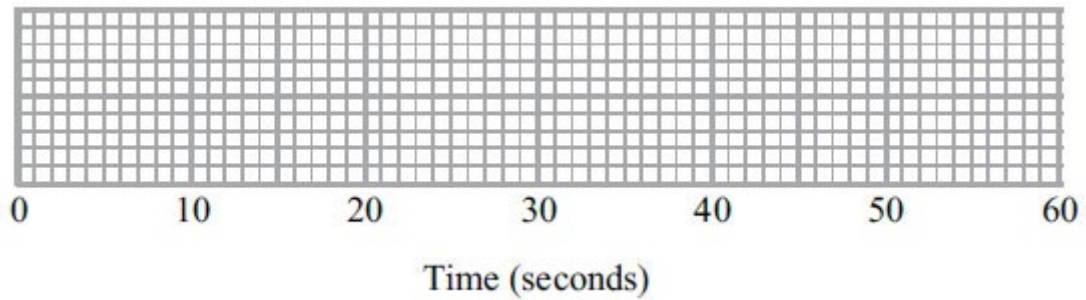
The box plot below shows the distribution of the times that people waited to be served at Green's garden centre.

Time (seconds)
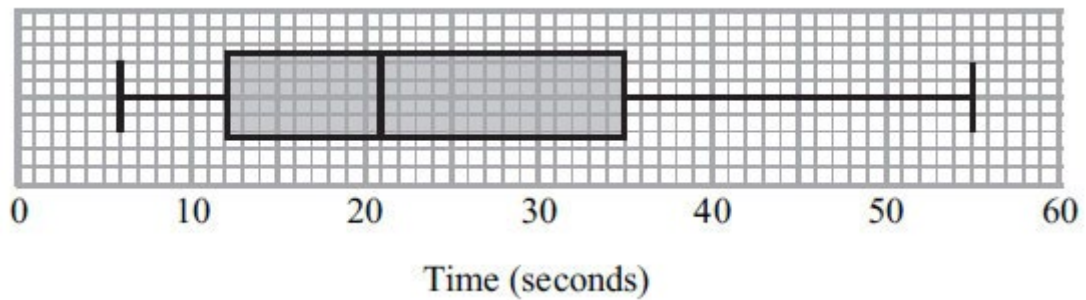
(b) Compare the distribution of the times that people waited at Rose's garden centre and the distribution of the times that people   waited at Green's garden centre.

....
....
....
....

(2)

(Total for Question is 5 marks)

Q13.
The table gives some information about the weights of 60 babies.

Lowest	2.0 kg
Highest	6.5 kg
Lower quartile	2.8 kg
Upper quartile	4.2 kg
Median	3.0 kg

(a) Draw a box plot to show this information.
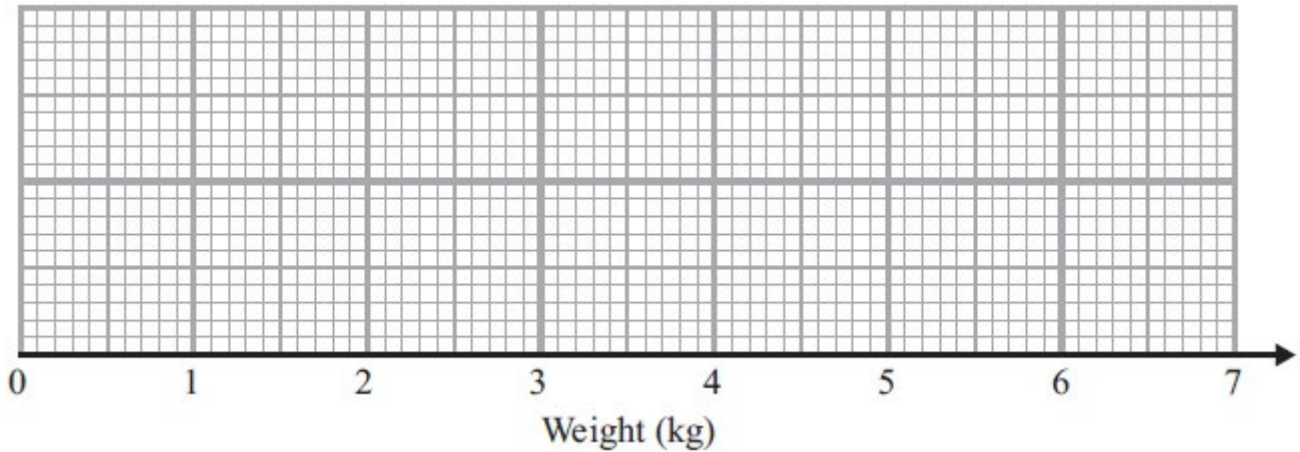
Weight (kg)

(2)

There are 60 babies.

(b) Work out an estimate for the number of these babies with a weight greater than 2.8 kg.

....

(2)

(Total for Question is 4 marks)

Examiner's Report

Q1.

Students need to be made aware that just using the term "average" within the context of statistics is not enough; throughout correct statistical terms were needed such as "median" or "IQR". Equally if figures were used, they had to be correct.

In part (a) those that recognised "half of" was associated with the median generally got the mark. Many responses quoted incorrect values or identified the incorrect parts of the box plot and a common misconception was that the box plot represented frequencies, so they would identify a "total" for the number of potatoes and then use it to try and find where the middle value was.

In part (b) a number of students did not understand what was being asked for or had not realised they needed to compare the median and (interquartile) range. Common was listing the values of each without a comparison or making an arithmetic error in calculation. Students often selected the wrong vocabulary, using mean instead of median and distribution or spread instead of the range. The IQR and range were often transposed. Students need to be reminded that they need to put their responses like this in context.

Q2.
It was really pleasing to see so many good responses to this question. Most were able to recognise both errors, but almost all finding at least one. Here the two errors were in the plotting of the median and the upper quartile. To gain the marks it was not sufficient to simply state they were plotted wrong, but students had to explicitly quote figures relating to the error.

Q3.
In part (a), the most common error from those students who understood anything about cumulative frequency graphs was to consider $\frac{1}{4}, \frac{1}{2}$ and $\frac{3}{4}$ of 50 instead of 48 resulting in an inaccurate box; the whiskers were usually correct gaining one mark. Many responses were seen where the whiskers were correct, but no box was drawn.

In part (b), far too many students simply compared specific values. Students need to know that, at this level, comparisons must be made of spread and central tendencies within the context of the question. Students found it difficult to interpret the data often confusing it with the number of trains delayed as opposed to delay time in minutes. Referring to delay time was necessary to put their comparisons into context.

Answers to part (c) clearly showed that very few students fully understand the dynamics of cumulative frequency diagrams. Rarely was any reference made to the parts of the diagram between 17 and 25 and 30 and 33. This is an area where centres could concentrate on when delivering this topic.

Q4.
Part (a) was not well answered with students showing they either didn't understand a very standard higher tier topic or that they couldn't accurately read the scale on an axis. It was the former that was most common, but many didn't know the difference between range and interquartile range, or simply had no idea. Of those who did know, many were unable to read the upper quartile from the diagram.

Part (b) proved a challenge for many. The mark was for understanding the proportional representation of a box plot and that 50% of waiting times will be up to the median. Part (c) was answered better with the most common response relating the median wait time being higher on Monday than Tuesday. Reference to measures of spread were not able to gain credit in part (c), and this was costly for some.

Q5.
It was encouraging to see that the majority of students were able to score at least one mark on this question, generally for stating that the male and female students had the same median.

However many used the word "medium" instead of "median" which scored no marks. Some of these students then went on to say that the females had a higher range or interquartile range than the males and a few students were able to put either of their statements into context. Unfortunately many quoted incorrect values for the range or IQR which meant they could not access the mark for this comparison. Students struggled to write a comparison in context which meant three marks on this question was

seldom scored. Many students just quoted facts without any comparison, e.g. "The IQR for females is 42 and the IQR for males is 31". This sort of statement could not score any marks without a comparison being made. When trying to compare range and IQR in context the majority of students who attempted this interpreted a wider range/IQR to mean that females used their phones more rather than identifying less consistent results.

Q6.

The scale of the box plot was simple, yet too many candidates mis-read values when either drawing the box plot, or completing the table. Part (b) was not well understood, as evidenced by estimates, guesses, and proportional calculations involving 8 and 60. Candidates need to be reminded that box plots effectively divide the distribution of the data into four parts.

Q7.

Many candidates knew what a box plot (box and whisker diagram) should look like and indeed knew how the plot related to the five given values. However, very many candidates lost marks through inaccurate drawing. The three upper values caused few problems as they were multiples of 5 and so were plotted above the corresponding values on the time axes. But very many candidates plotted the 18 at 19, not taking note of the scale. Similar poor plotting happened with the lowest value of 14, which was often plotted at 12.

Q8.

Part (a) was a straightforward question for the vast majority of students who knew how to draw a box plot. Some students lost one mark for an error in plotting one of the five values.

In part (b), fewer students than might have been expected realised that 3/4 of the girls have a height between 133 cm and 157 cm. Those who did were almost always able to work out 3/4 of 80 and score 2 marks. A few students just wrote 75% or 3/4 which did not gain any credit. A surprising number of students did not connect the stated values with their positions within the data set so were unable to recognise that they needed to work with 3/4 of the 80 girls. Common incorrect methods included 157 – 133 = 24, 4/5 × 80 = 64 and 24/37 × 80.

Q9.

Nearly all students attempted to draw a box plot but a sizeable minority of them plotted all the values given in the table rather than realise that they needed to calculate the value of the upper quartile by adding the interquartile range to the lower quartile. As a result many students restricted the number of marks gained to one.

Q10.

Part (a) required the candidates to draw an accurate box plot from the given data. There were very few non attempts and nearly all candidates, managed to score at least a mark for two correct values drawn with a box or whiskers. The majority managed to have the maximum and minimum values in the correct places, although the UQ and LQ values proved challenging to many.

In part (b), many candidates stated values without comparative statements thus losing marks. Those that used the word mean instead of median also lost that mark although 'average' was acceptable. Quite a few commented on the oldest and youngest members of women/men without any mention of range. Others stated that there were 'more older men than women' attending the tennis club, losing the mark as there is nothing to suggest the quantities of men at the club

Q11.

In part (a) most scored full marks.

In part (b) there were some trivial comparisons, but most candidates were able to gain a single mark from comparing the median or IQR. To gain full marks at least one of these needed to be expressed in terms of the context of the question, making reference to the money. Simply listing the values for the measures is not comparative and should be discouraged.

Q12.

In part (a), it was common for candidates to draw a box plot with the correct minimum and maximum values. The median was correct more often than the lower and upper quartiles.

In part (b), it was common to see very general comments rather than comments comparing the medians and the ranges or interquartile ranges. Specific mathematical language should be used. Many responses

compared numbers from within the datasets rather than comparing the distributions, eg 'the median at Rose's was 25 and at Green's it was 21' should be presented as 'the median at Rose's was higher'. Some candidates used incorrect terminology, eg the word 'average' instead of 'median' for comparisons, so did not gain the mark.

Q13.
Very few candidates failed to construct an accurate box plot showing the given information in part (a). Some failed to draw a box whilst a minority showed all 5 pieces of information in a box.

Part (b) was very much less well done. Many candidates clearly appear not to understand the meaning of quartiles in the context of a problem and failed to grasp that they had to find ¾ of the data. The most common incorrect approach was to try to use the Weight values as frequency values.
  For example 6.5 – 2.8 = 3.7

Mark Scheme

Q1.

Question	Answer	Mark	Mark scheme	Additional guidance
(a)	Explanation	C1	eg 'No' the median is 57	
(b)	Comparison	C1	(ft) a correct comparison of medians eg the median weight for Megan was greater than the median weight for Amy	Simply quoting values for median, range and IQR is insufficient, they must be compared
		C1	a correct comparison of a measure of spread eg the interquartile range of weights for Megan was greater than the interquartile range of weights for Amy For the award of both marks at least one of the comparisons must be in the context of the question	Median Range IQR Megan 57 49 26 Amy 42 47 16 Figures given must be correct. Comparisons can relate to the range or the IQR

Q2.

Question	Answer	Mark	Mark scheme	Additional guidance
	Two changes	C1	plot the median at 162, not 161 oe	
		C1	plot the upper quartile at 171, not 172 oe **Acceptable examples** the median has been plotted at 161 / upper quartile at 172 the upper quartile should be 171 (not 172) UQ is wrong as IQR is 17 not 18 **Not acceptable examples** the median / upper quartile have been plotted / drawn wrong the upper quartile has been worked out incorrectly She needs to work out the UQ	

Q3.

Question	Answer	Mark	Mark scheme	Additional guidance
(a)	box plot drawn	B1	ends of whiskers at 0 and 42 with a box	The box can be of any height. Accept ends that are marked (eg line, cross, dot) or defined by the end of the whiskers if clear.
		B1	median at 10 inside a box	Has to be inside a box; whiskers not required
		B1	for ends of box at 4 and 20	An independent mark that can be awarded for just a box; do not need whiskers for this mark.
(b)	Comparison	C1	for a correct comparison of medians, eg. the median delay time on Mon was greater than the median delay time on Tues. or ft (a)	Simply quoting values for median, range and IQR is insufficient, they must be compared
		C1	for a correct comparison of a measure of spread, eg. the interquartile range (range) of delay times on Mon was greater than the interquartile range (range) of delay times on Tues. or ft (a) For the award of both marks at least one of the comparisons must be in context	Comparisons can relate to the median, and then either the range or the IQR.
(c)	statement	C1	'No' with statement explaining that there might not be any delays between 25 minutes and 30 minutes as in the upper 25% (12 trains) the delays may all be between 17 and 25 or 30 and 33	The 'No' may be implied from their wording, and could be written next to the "?" The statement must mention (or imply) values above the UQ of 17

Q4.

Question	Answer	Mark	Mark scheme	Additional guidance
(a)	138	M1	for upper quartile = 188 or lower quartile = 50 or an indication that they are trying UQ – LQ	Could be written on the grid
		A1	cao	
(b)	Yes, with reason	C1	Yes, with reason **Acceptable examples** Yes, because the median is at 2 hour (120 min) Yes, since 50% is at the 2 hour mark Yes, because the middle is at 2 hours **Not acceptable examples** No ….. The median is at the 2 hour mark Yes, because 50% is exactly half way between "188" and "50"	
(c)	statement	C1	**Acceptable examples** The median is lower on Tuesday (higher on Monday) The upper quartile is lower on Tuesday (higher on Monday) There may just have been one person waiting for 210 mins We don't know how many people were waiting for each time **Not acceptable examples** The range is bigger for Tuesday (smaller for Monday) The IQR is smaller for Tuesday (bigger for Monday)	M T Shortest time 20 20 Lower quartile 50 50 Median 120 100 Upper quartile 188 140 Longest time 200 210 Range 180 190 IQR 138 90

Q5.

Question	Working			Answer	Mark	Notes
*		M	F	Compare: medians and spread	3	C1 for any correct comparison of the medians C1 for any correct comparison of the IQRs or range C1 for a comparison of medians, IQRs or ranges written in context
	Median	40	40			
	Range	90	106			
	IQR	31	42			

Q6.

5MB1H_01				
Question	Working	Answer	Mark	Notes
(a)(i)		11, 22	3	B3 for all of: median 11 drawn, greatest number 22 drawn, smallest number 3 in table, upper quartile 17 in table
(ii)		3, 17		(B2 for 3 correct, B1 for 2 correct)
(b)		45	2	M1 for 0.75×60 or 60 − 0.25×60 oe or 15 seen (may be seen in diagram) A1 cao

Q7.

		Working	Answer	Mark	Notes
			Box plot drawn	3	B3 for fully correct box plot (B2 for 3 correctly plotted values including box and tails) (B1 for 2 correctly plotted values including box or tails)

Q8.

Question	Answer	Mark	Mark scheme	Additional guidance
(a)	Box plot drawn	B3	for a fully correct box plot	Condone the lack of a vertical marker at the end of the tails
		(B2	for at least 3 correctly plotted values including box and whiskers/tails)	Note that a box must be present, as must "tails"
		(B1	for at least 2 correctly plotted values including box or whiskers/tails or 5 correct values plotted or clearly identified and no box or whiskers/tails)	
(b)	60	M1	for a method to find ¾ of 80 eg 20 + 20 + 20 or ¾ × 80	
		A1	cao	

Q9.

5MB1H_01 November 2015				
Question	Working	Answer	Mark	Notes
		Box plot	3	M1 for UQ = 1.7 + 2.7 (= 4.4) or 4.4 shown on box plot B1 for box plot with at least 3 pieces of information correctly plotted B1 for fully correct box plot NB: there must be a box

Q10.

5MB1H_01					
Question	Working	Answer	Mark	Notes	
(a)	Lowest = 16 LQ = 20 Median = 28 UQ = 42 Highest = 50	Box plot drawn	3	B3 for fully correct box plot (B2 for at least 3 correctly plotted values including box and whiskers/ tails) (B1 for at least 2 correctly plotted values including box or whiskers/tails or 5 correct values plotted and no box or whiskers/tails)	
(b)			2	C2 ft for two relevant comparisons one for median/average and one for IQR or range (C1ft for one relevant comparison) ie IQR or range or use of median for comparison	

Q11.

		Working	Answer	Mark	Notes
	(a)		Box plot overlay	2	M1 for a box drawn with at least 2 correct points from LQ, Med and UQ A1 for a fully correct box plot
	(b)		Comparison of a measure of spread plus a comparison of medians (in context)	2	B1 for a correct comparison of a measure of spread (using either range or iqr) B1 for a correct comparison of medians For the award of both marks at least one of the comparisons made must be in the context of the question.

Q12.

		Working	Answer	Mark	Notes
	(a)	minimum = 5 lower quartile = 14 median = 25 upper quartile = 30 maximum = 44	box plot	3	B3 for fully correct box plot (B2 for at least 3 correct values plotted including box **and** tails **or** 5 correct values indicated) (B1 for at least 2 correct values plotted including box **or** tails **or** 3 or 4 correct values indicated)
	(b)		comparisons	2	B1 for a correct comparison (ft) of medians B1 for a correct comparison (ft) of ranges or IQRs

Q13.

Question		Working	Answer	Mark	Notes
	(a)		Box plot	2	B2 for correct box plot (B1 for box plot with at least 3 pieces of information correctly plotted) Note: There must be a box
	(b)	75% × 60	45	2	M1 for 0.75 × 60 oe A1 cao [SC: B1 for an answer of 15 if M0 scored]