

2. The time taken to complete a puzzle, in minutes, is recorded for each person in a club. The times are summarised in a grouped frequency distribution and represented by a histogram.

One of the class intervals has a frequency of 20 and is shown by a bar of width 1.5 cm and height 12 cm on the histogram. The total area under the histogram is 94.5 cm^2

Find the number of people in the club.

(3)

2. A rugby club coach uses club records to take a random sample of 15 players from 1990 and an independent random sample of 15 players from 2010. The body weight of each player was recorded to the nearest kg and the results from 2010 are summarised in the table below.

Body weight (kg)	75–79	80–84	85–89	90–94	95–99	100–104	105–109
Number of Players (2010)	1	2	2	4	3	2	1

- (a) Find the estimated values in kg of the summary statistics a , b and c in the table below.

	Estimate in 1990	Estimate in 2010
Mean	83.0	a
Median	82.0	b
Variance	44.0	c

Give your answers to 3 significant figures.

(6)

The rugby coach claims that players' body weight increased between 1990 and 2010.

- (b) Using the table in part (a), comment on the rugby coach's claim.

(2)

2. The weights, to the nearest kilogram, of a sample of 33 female spotted hyenas living in the Serengeti are summarised in the stem and leaf diagram below.

Weight (kg)	Totals
3 2 3 7	(3)
4 1 3 3 4 5 5 6 9	(8)
5 1 2 2 3 4 4 5 5 5 7 8 8 9 9 9	(15)
6 2 3 3	(3)
7 1 4 7	(3)
8 4	(1)

Key: 3|2 means 32

- (a) Find the median and quartiles for the weights of the female spotted hyenas.

(3)

An outlier is defined as any value greater than c or any value less than d where

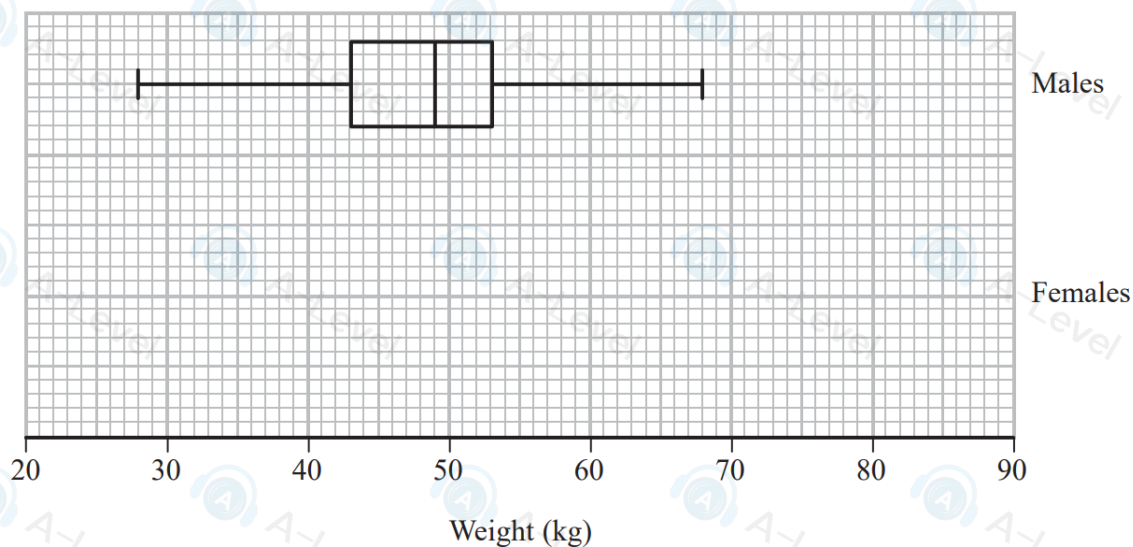
$$c = Q_3 + 1.5(Q_3 - Q_1)$$

$$d = Q_1 - 1.5(Q_3 - Q_1)$$

- (b) Showing your working clearly, identify any outliers for these data.

(3)

The weights, to the nearest kilogram, of a sample of male spotted hyenas living in the Serengeti are summarised below.



- (c) In the space provided in the grid above, draw a box and whisker plot to represent the weights of female spotted hyenas living in the Serengeti. Indicate clearly any outliers. (A copy of this grid is on page 9 if you need to redraw your box and whisker plot.)

(3)

- (d) Compare the weights of male and female spotted hyenas living in the Serengeti.

(2)

1. A researcher is investigating the growth of two types of tree, Birch and Maple. The height, to the nearest cm, a seedling grows in one year is recorded for 35 Birch trees and 32 Maple trees. The results are summarised in the back-to-back stem and leaf diagram below.

Totals	Birch		Maple	Totals
(2)	9 8	2	5 7 7 8 9	(5)
(8)	9 9 9 6 5 3 1 1	3	0 2 6 6 8 9 9	(7)
(9)	9 8 8 7 6 3 1 1 1	4	1 1 1 <i>k</i> 7 8	(6)
(9)	7 7 7 5 4 3 2 1 0	5	0 1 2 3 4 4 4	(7)
(3)	7 6 5	6	3 4 6	(3)
(3)	6 5 4	7	0 7	(2)
(1)	5	8	0 0	(2)

Key: 5 | 6 | 3 means 65 cm for a Birch tree and 63 cm for a Maple tree

The median height that these **Maple** trees grow in one year is 45 cm.

- (a) Find the value of *k*, used in the stem and leaf diagram.

(1)

- (b) Find the lower quartile and the upper quartile of the height grown in one year for these **Birch** trees.

(2)

The researcher defines an outlier as an observation that is

greater than $Q_3 + 1.5 \times (Q_3 - Q_1)$ or less than $Q_1 - 1.5 \times (Q_3 - Q_1)$

- (c) Show that there is only one outlier amongst the Birch trees.

(2)

The grid on page 3 shows a box plot for the heights that the Maple trees grow in one year.

- (d) On the same grid draw a box plot for the heights that the Birch trees grow in one year.

(4)

- (e) Comment on any difference in the distributions of the growth of these Birch trees and the growth of these Maple trees.

State the values of any statistics you have used to support your comment.

(1)

The researcher realises he has missed out 4 pieces of data for the **Maple** trees. The heights each seedling grows in one year, to the nearest cm, in ascending order, for these 4 Maple trees are 27 cm, *a* cm, 48 cm, 2*a* cm.

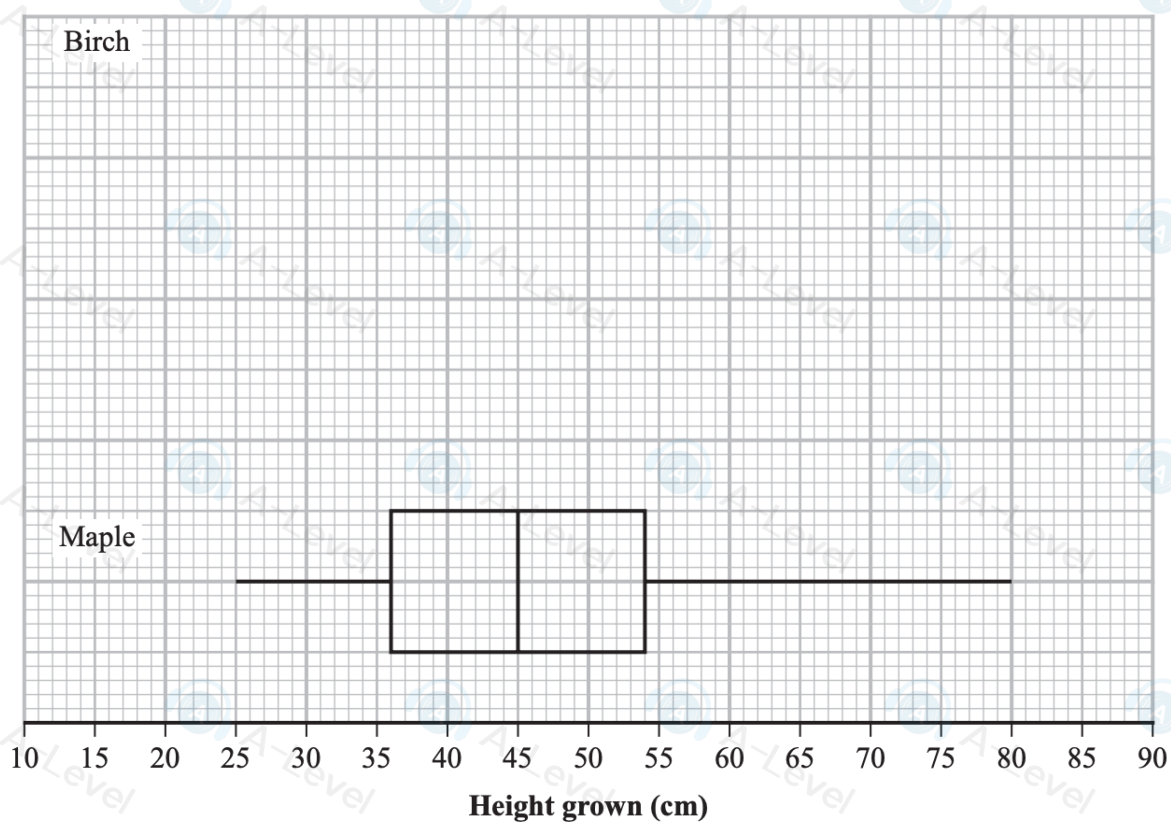
Given that there is no change to the box plot for the **Maple** trees given on page 3

- (f) find the range of possible values for *a*
Show your working clearly.

(3)

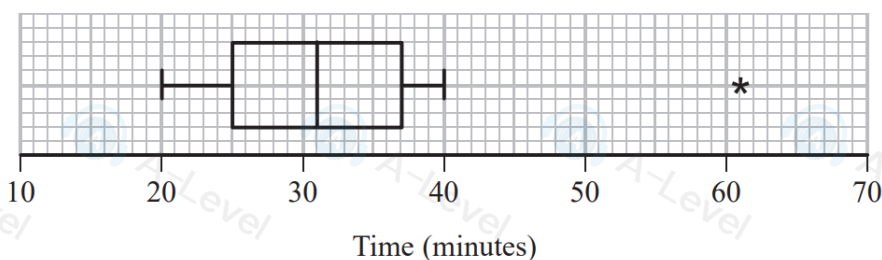
DO NOT WRITE

DO NOT WRITE IN THIS AREA



Turn over for spare grid if you need to redraw your answer for part (d)

2. The box plot shows the times, t minutes, it takes a group of office workers to travel to work.



- (a) Find the range of the times. (1)
- (b) Find the interquartile range of the times. (1)
- (c) Using the quartiles, describe the skewness of these data. Give a reason for your answer. (2)

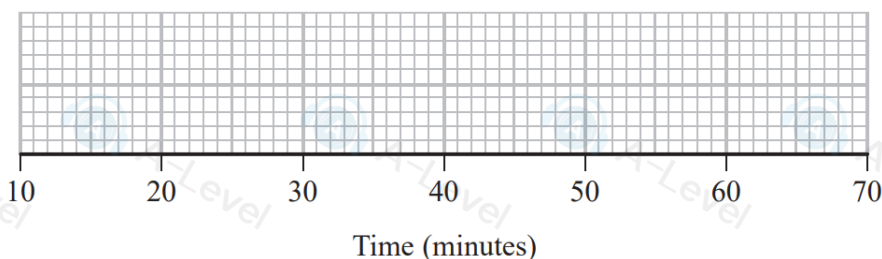
Chetna believes that house prices will be higher if the time to travel to work is shorter. She asks a random sample of these office workers for their house prices $\pounds x$, where x is measured in thousands, and obtains the following statistics

$$S_{xx} = 5514 \quad S_{xt} = 10 \quad S_{tt} = 1145.6$$

- (d) Calculate the product moment correlation coefficient between x and t . (2)
- (e) State, giving a reason, whether or not your correlation coefficient supports Chetna's belief. (2)

Adam and Betty are part of the group of office workers and they have both moved house. Adam's time to travel to work changes from 32 minutes to 36 minutes. Betty's time to travel to work changes from 38 minutes to 58 minutes. Outliers are defined as values that are more than 1.5 times the interquartile range above the upper quartile.

- (f) Showing all necessary calculations, determine how the box plot of times to travel to work will change and draw a new box plot on the grid on page 5. (3)



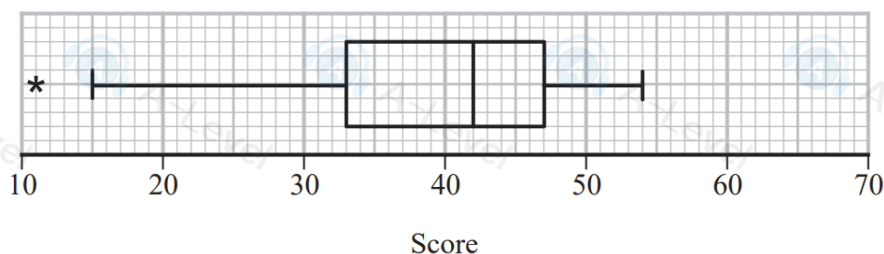
2. Chi wanted to summarise the scores of the 39 competitors in a village quiz. He started to produce the following stem and leaf diagram

Score

1	1 5 8 9
2	0 2 5 8 9
3	3 5 5 7 8 9 ...

Key: 2 | 5 is a score of 25

He did not complete the stem and leaf diagram but instead produced the following box plot.



Chi defined an outlier as a value that is

$$\text{greater than } Q_3 + 1.5 \times (Q_3 - Q_1)$$

or

$$\text{less than } Q_1 - 1.5 \times (Q_3 - Q_1)$$

(a) Find

- the interquartile range
- the range.

(2)

Albert and Beth asked for their scores to be checked.

Albert's score was changed from 25 to 37

Beth's score was changed from 54 to 60

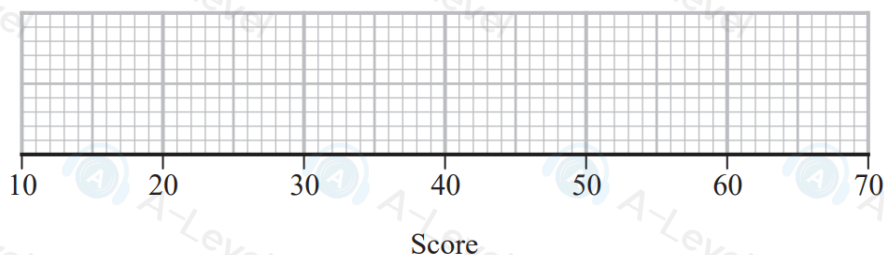
- (c) On the grid on page 5, draw an updated box plot.
Show clearly any calculations that you used.

(7)

Some of the competitors complained that the questions were biased towards the younger generation. The product moment correlation coefficient between the age of the competitors and their score in the quiz is -0.187

- (d) State, giving a reason, whether or not the complaint is supported by this statistic.

(2)



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN