Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise A, Question 1

Question:

A group of thirty college students was asked how many DVDs they had in their collection. The results are as follows.

12	25	34	17	12	18	29	34	45	6
15	9	25	23	29	22	20	32	15	15
19	12	26	27	27	32	35	42	26	25

Draw a stem and leaf diagram to represent these data.

a Find the median.

b Find the lower quartile.

c Find the upper quartile.

Solution:

Unordered

key 1|2 means 12

```
0 | 6  9
1 | 2  7  2  8  5  5  5  9  2
2 | 5  9  5  3  9  2  0  6  7  7  6  5
3 | 4  4  2  2  5
4 | 5  2
```

Ordered

key 1|2 means 12

```
0 | 6  9
1 | 2  2  2  5  5  5  7  8  9
2 | 0  2  3  5  5  5  6  6  7  7  9  9
3 | 2  2  4  4  5
4 | 2  5
```

(a) $\frac{30}{2} = 15$ therefore 15.5^{th} term $= 25$

(b) $\frac{30}{4} = 7.5$ therefore 8th term $= 15$

(c) $\frac{3(30)}{4} = 22.5$ therefore 23rd term $= 29$

© Pearson Education Ltd 2008

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise A, Question 2

Question:

The following stem and leaf diagram shows some information about the marks gained by a group of students in a statistics test.

stem	leaf									Key: 2|3 means 23 marks	
0	8	9									(2)
1	2	5	5	9							(4)
2	3	6	6	6	7						(5)
3	4	4	5	7	7	7	7	7	9		(9)
4	5	8	8	9							(4)

a Work out how many students there were in the group.

b Write down the highest mark.

c Write down the lowest mark.

d Write down how many students got 26 marks.

e Write down the modal mark.

f Find the median mark.

g Find the lower quartile.

h Find the upper quartile.

Solution:

(a) 24 (b) 49 (c) 8 (d) 3 (e) 37

(f) $\frac{24}{2} = 12$ therefore 12.5^{th} term $= 34$

(g) $\frac{24}{4} = 6$ therefore 6.5^{th} term $= \frac{19 + 23}{2} = 21$

(h) $\frac{3(24)}{4} = 18$ therefore 18.5^{th} term $= 37$

© Pearson Education Ltd 2008

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise A, Question 3

Question:

3 The number of laptops sold by a store was recorded each month for a period of 26 months. The results are shown in the stem and leaf diagram.

stem	leaf									Key: 1|8 means 18 laptops
1	8									(1)
2	3	6	7	9	9					(5)
3	2	6	6	6	7	8	8			(7)
4	4	5	5	5	7	7	7	7	9	(9)
5	2	7	7	9						(4)

a Find the median.

b Find the lower quartile.

c Find the upper quartile.

d Work out the interquartile range.

e Write down the modal number of laptops sold.

Solution:

(a) $\frac{26}{2} = 13$ therefore 13.5^{th} term $= \frac{38 + 44}{2} = 41$

(b) $\frac{26}{4} = 6.5 = 7^{\text{th}}$ term $= 32$

(c) $\frac{3(26)}{4} = 19.5 = 20^{\text{st}}$ term $= 47$

(d) IQR $= 47 - 32 = 15$

(e) 47

© Pearson Education Ltd 2008

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise A, Question 4

Question:

A class of 16 boys and 13 girls did a Physics test. The test was marked out of 60. Their marks are shown below.

Boys				Girls			
45	54	32	60	26	54	47	32
28	34	54	56	34	34	45	46
32	29	47	48	39	52	24	28
44	45	56	57	33			

a Draw a back-to-back stem and leaf diagram to represent these data.

b Comment on your results.

Solution:

(a) Unordered

```
                        Girls Key 2|6
      Boys                    means 26
_____
        9  8 | 2 | 6  4  8
      2  4  2 | 3 | 2  4  4  9  3
  5  4  8  7  5 | 4 | 7  5  6
  7  6  6  4  4 | 5 | 4  2
              0 | 6 |
```

Ordered

```
                        Girls Key 2|6
      Boys                    means 26
_____
        9  8 | 2 | 4  6  8
      4  2  2 | 3 | 2  3  4  4  9
  8  7  5  5  4 | 4 | 5  6  7
  7  6  6  4  4 | 5 | 2  4
              0 | 6 |
```

(b) Girls gain lower marks than boys

© Pearson Education Ltd 2008

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise A, Question 5

Question:

The following stem and leaf diagram shows the weekend earnings of a group of college students.

				Males		Females						Key: 5|1|0 means £15 for
			8	0	6							males and £10 for females
	7	6	5	1	0	5	5	5	8	8	
9	9	9	8	6	6	2	5	5	8	8	9
	8	8	5	5	5	3	5	5			
			8	5	4	0					

a Write down the number of male students and the number of female students.

b Write down the largest amount of money earned by the males.

c Comment on whether males or females earned the most in general.

Solution:

(a) 17 males and 15 females

(b) £48

(c) Males earned the most in general.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise B, Question 1

Question:

Some data are collected. The lower quartile is 46 and the upper quartile is 68.

An outlier is an observation that falls either $1.5 \times$ (interquartile range) above the upper quartile or $1.5 \times$ (interquartile range) below the lower quartile.

Work out whether the following are outliers using this rule.

a 7

b 88

c 105

Solution:

IQR = 68–46 = 22

$46 - 1\frac{1}{2} \times 22 = 13$

$68 + 1\frac{1}{2} \times 22 = 101$

(a) 7 is an outlier (b) 88 is not an outlier (c) 105 is an outlier.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise B, Question 2

Question:

Male and female turtles were weighed in grams. For males, the lower quartile was 400 g and the upper quartile was 580 g. For females, the lower quartile was 260 g and the upper quartile was 340 g.

An outlier is an observation that falls either $1 \times$ (interquartile range) above the upper quartile or $1 \times$ (interquartile range) below the lower quartile.

a Which of these male turtle weights would be outliers?

  400 g   260 g   550 g   640 g

b Which of these female turtle weights would be outliers?

  170 g   300 g   340 g   440 g

c What is the largest size a male turtle can be without being an outlier?

Solution:

a) no outliers
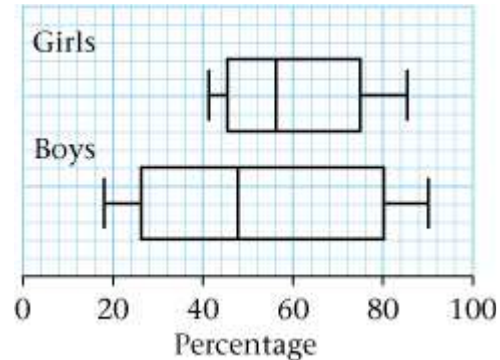
b) 170 g and 440 g are outliers

c) 760 g

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise C, Question 1

Question:

A group of students did a test. The summary data is shown in the table below.

Lowest value	Lower quartile	Median	Upper quartile	Highest value
5	21	28	36	58

Given that there were no outliers draw a box plot to illustrate these data.

Solution:
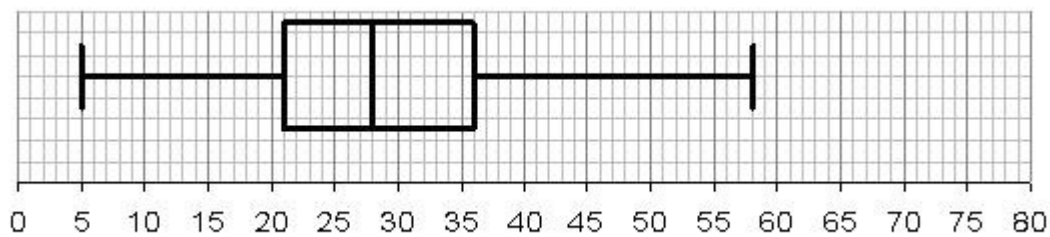
© Pearson Education Ltd 2008

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise C, Question 2

Question:

Here is a box plot of marks in an examination.
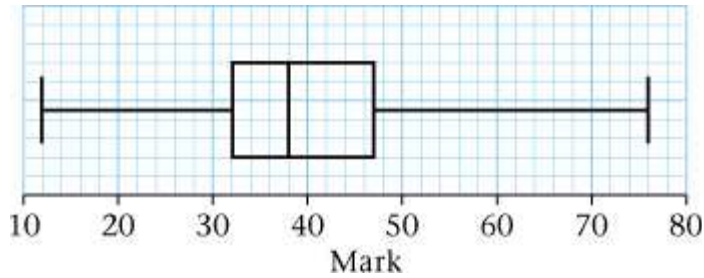
a Write down the upper and lower quartiles.

b Write down the median.

c Work out the interquartile range.

d Work out the range.

Solution:

(a) 47 and 32 (b) 38 (c) IQR = 47 − 32 = 15 (d) Range = 76 − 12 = 64

© Pearson Education Ltd 2008

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise D, Question 1

Question:

A group of students took a statistics test. The summary data for the percentage mark gained by boys and by girls is shown in the box plots opposite.

a Write down the percentage mark which 75% of the girls scored more than.

b State the name given to this value.

c Compare and contrast the results of the boys and the girls.

Solution:

(a) 45

(b) Lower quartile

(c) Boys have a lower median and bigger IQR/range (or girls have a higher median and lower IQR/range)

The person with the highest mark was a boy. The person with the lowest mark was a boy.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise D, Question 2

Question:

Male and female turtles were weighed in grams. Their weights are summarised in the box plots opposite.
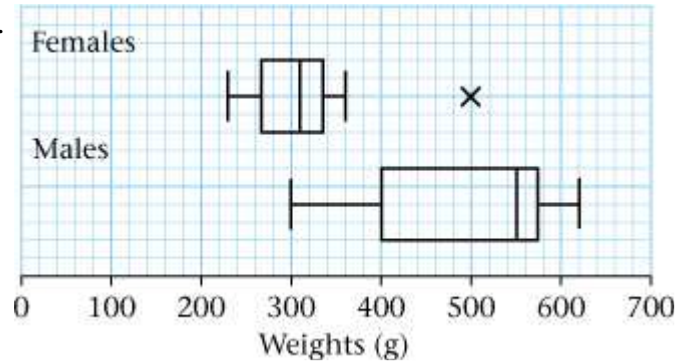
a Compare and contrast the weights of the male and female turtles.

b A turtle was found to be 330 grams in weight. State whether it is likely to be a male or a female. Give a reason for your answer.

c Write down the size of the largest female turtle.

Solution:

(a) Male turtles have a higher median weight (or Females have a lower).

Males have a bigger range (or IQR) (females have a lower range IQR).

(b) It is more likely to be a female. Hardly any male turtles would weigh this little, but more than a quarter of female turtles would weigh more than this.

(c) 500 g

© Pearson Education Ltd 2008

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise E, Question 1

Question:

The heights of a year group of children were measured. The data are summarised in the group frequency table.

Height (cm)	Frequency	Class width	Frequency density
135–144	40	10	
145–149	40	5	
150–154	75		
155–159	65		
160–174	60		

a Copy and complete the table.

b Draw a histogram for these data.

Solution:

a 5, 5, 15; 4, 8, 15, 13, 4

b
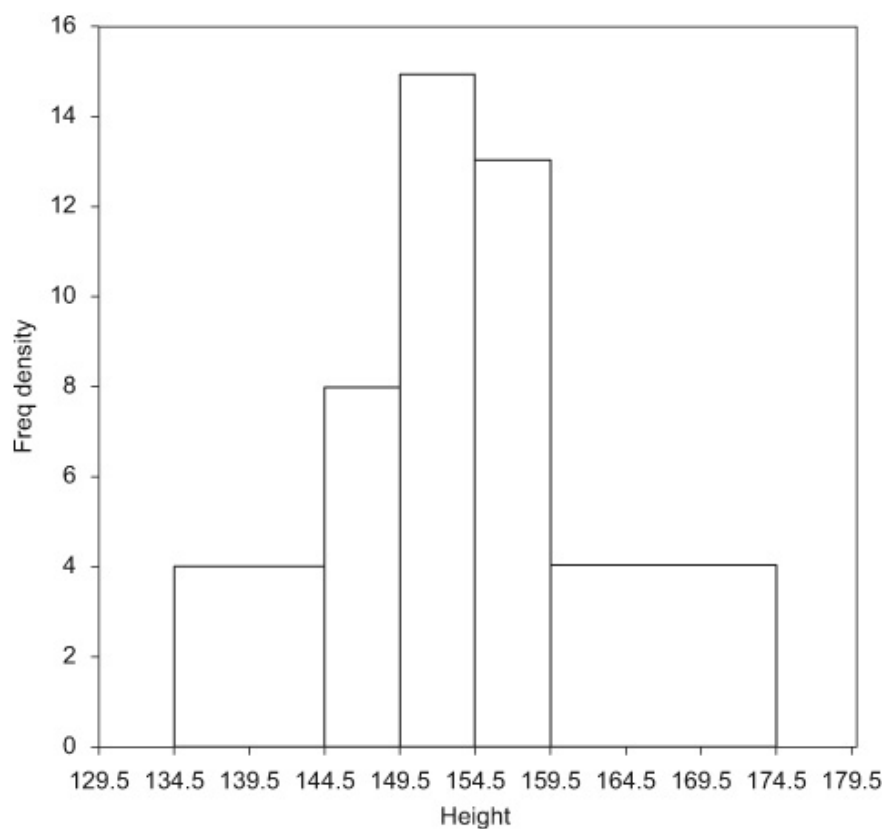
Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise E, Question 2

Question:

Some students take part in an obstacle race. The time it took each student to complete the race was noted. The results are shown in the histogram.
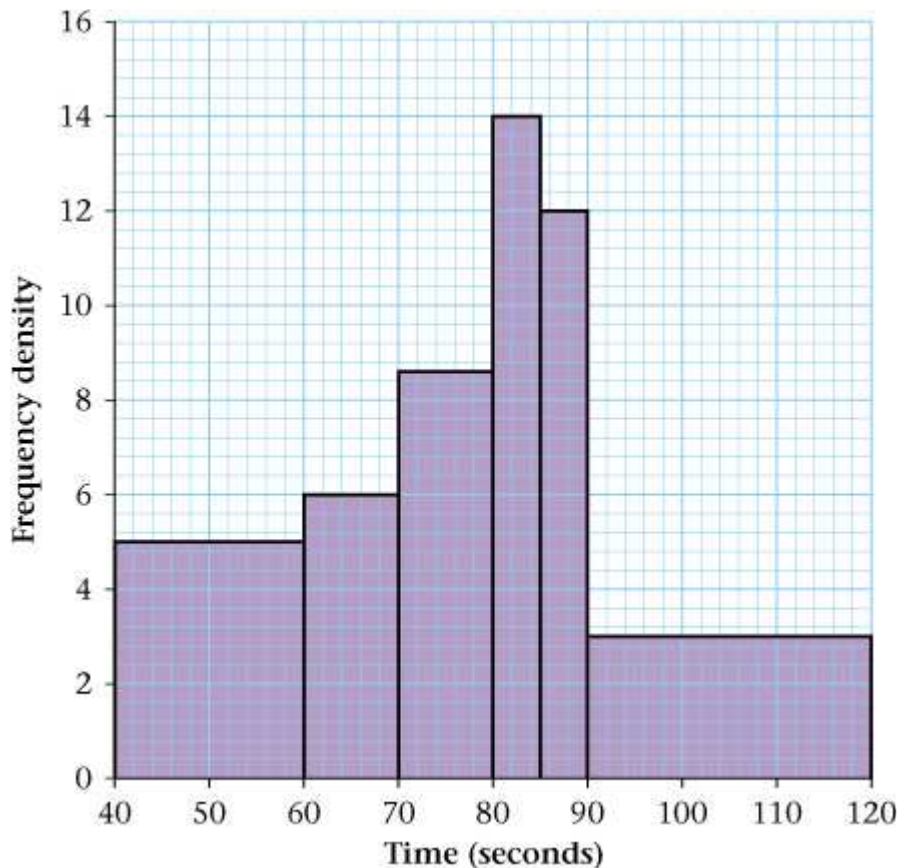
a Give a reason to justify the use of a histogram to represent these data.

The number of students who took between 60 and 70 seconds is 90.

b Find the number of students who took between 40 and 60 seconds.

c Find the number of students who took 80 seconds or less.

d Calculate the total number of students who took part in the race.

Solution:

a Time is continuous data

b Area of $60 - 70$ seconds bar is $10 \times 6 = 60$ units squared

1 unit squared $= 90/60 = 1.5$ students

Area of $40 - 60$ seconds bar is $20 \times 5 = 100$ units squared

Number of students = $100 \times 1.5 = \mathbf{150}$

c Area for 80 seconds or less = $20 \times 5 + 10 \times 6 + 10 \times 8.6 = 246$ units squared.

Number of students = $246 \times 1.5 = \mathbf{369}$

d Total Area = $246 + 5 \times 14 + 5 \times 12 + 30 \times 3 = 466$ units squared.

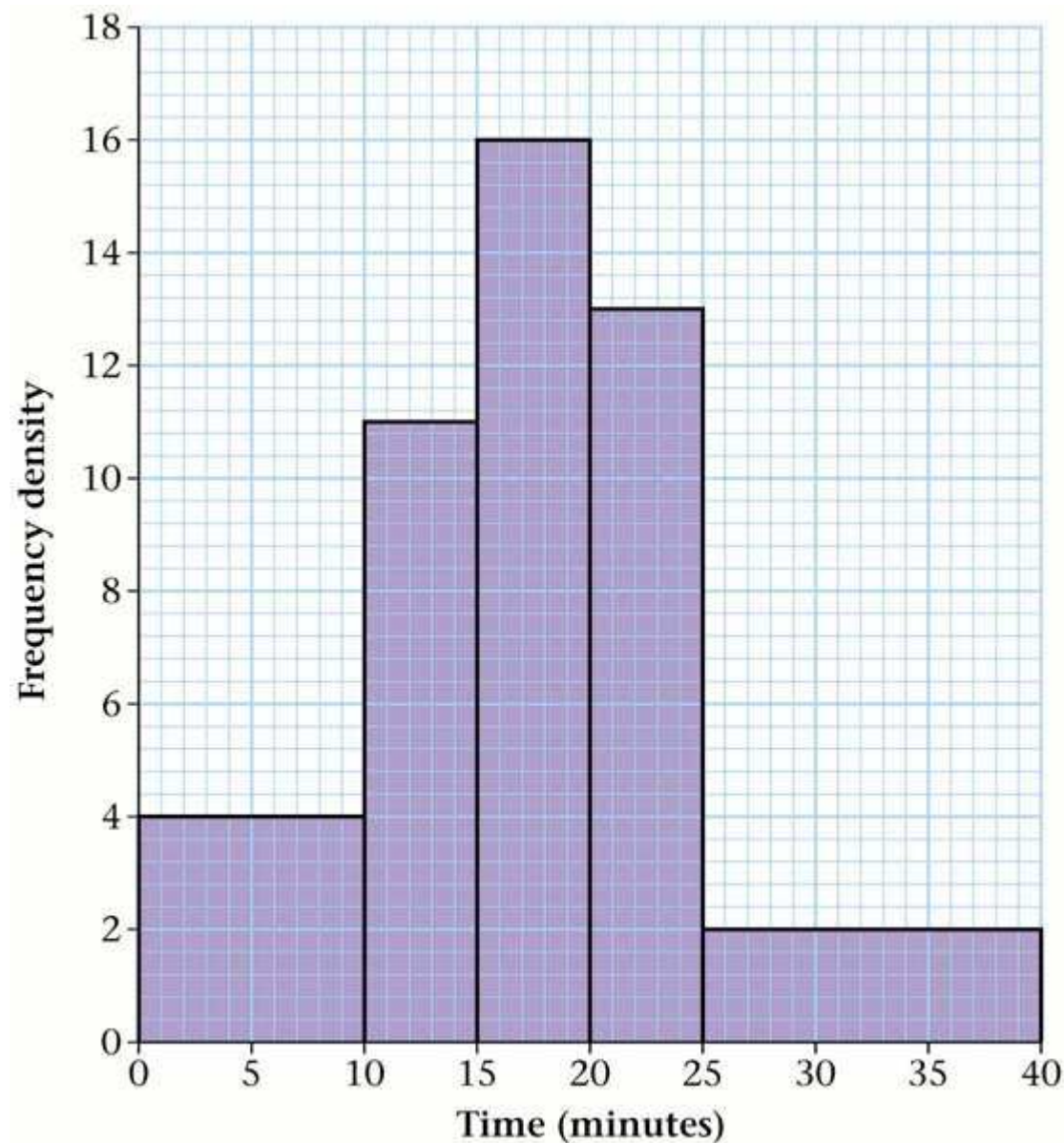Number of employees = $466 \times 1.5 = \mathbf{699}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise E, Question 3

Question:

The time taken for each employee in a company to travel to work was recorded. The results are shown in the histogram.

The number of employees who took less than 10 minutes to travel to work is 48.

a Find how many employees took less than 15 minutes to travel to work.

b Estimate how many employees took between 20 and 30 minutes to travel to work.

c Estimate how many employees took more than 30 minutes to travel to work.

Solution:

a Area less than 10 minutes is $10 \times 4 = 40$ units squared

1 unit squared $= 48 \div 40 = 1.2$ students

Area less than 15 minutes is $40 + 5 \times 11 = 95$ units squared

Number of employees $= 95 \times 1.2 = \mathbf{114}$

b Area of $20 - 30$ minutes bars is $5 \times 13 + 5 \times 2 = 75$ units squared

Number of employees $= 75 \times 1.2 = \mathbf{90}$

c Area for more than 30 minutes seconds bars is $10 \times 2 = 20$ units squared

Number of employees $= 20 \times 1.2 = \mathbf{24}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise E, Question 4

Question:

A Fun Day committee at a local sports centre organised a throwing the cricket ball competition. The distance thrown by every competitor was recorded. The data were collected and are shown in the histogram. The number of competitors who threw less than 10 m was 40.
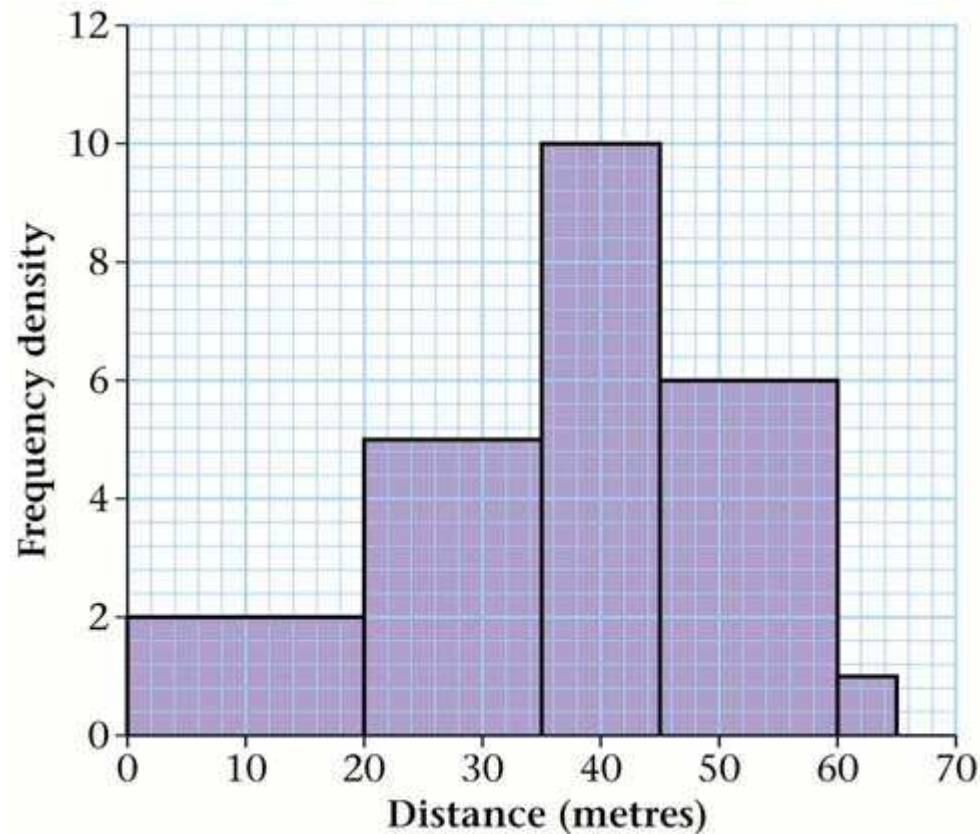
a Why is a histogram a suitable diagram to represent these data?

b How many people entered the competition?

c Estimate how many people threw between 30 and 40 metres.

d How many people threw between 45 and 65 metres?

e Estimate how many people threw less than 25 metres.

Solution:

a Distance is continuous data.

b Area for less than 10m is $10 \times 2 = 20$ units squared

1 unit squared $= 40 \div 20 = 2$ people

Total Area $10 \times 2 + 10 \times 2 + 15 \times 5 + 10 \times 10 + 15 \times 6 + 5 \times 1 = 310$ units squared

Number of people $= 310 \times 2 = $ **620**

c Area for 30 − 40 m is $5 \times 5 + 5 \times 10 = 75$ units squared.

Number of people $= 75 \times 2 = $ **150**

d Area for 45 − 65 m is $15 \times 6 + 5 \times 1 = 95$ units squared.

Number of people $= 95 \times 2 = $ **190**

e Area for less than 25 m is $10 \times 2 + 10 \times 2 + 5 \times 5 = 65$ units squared.

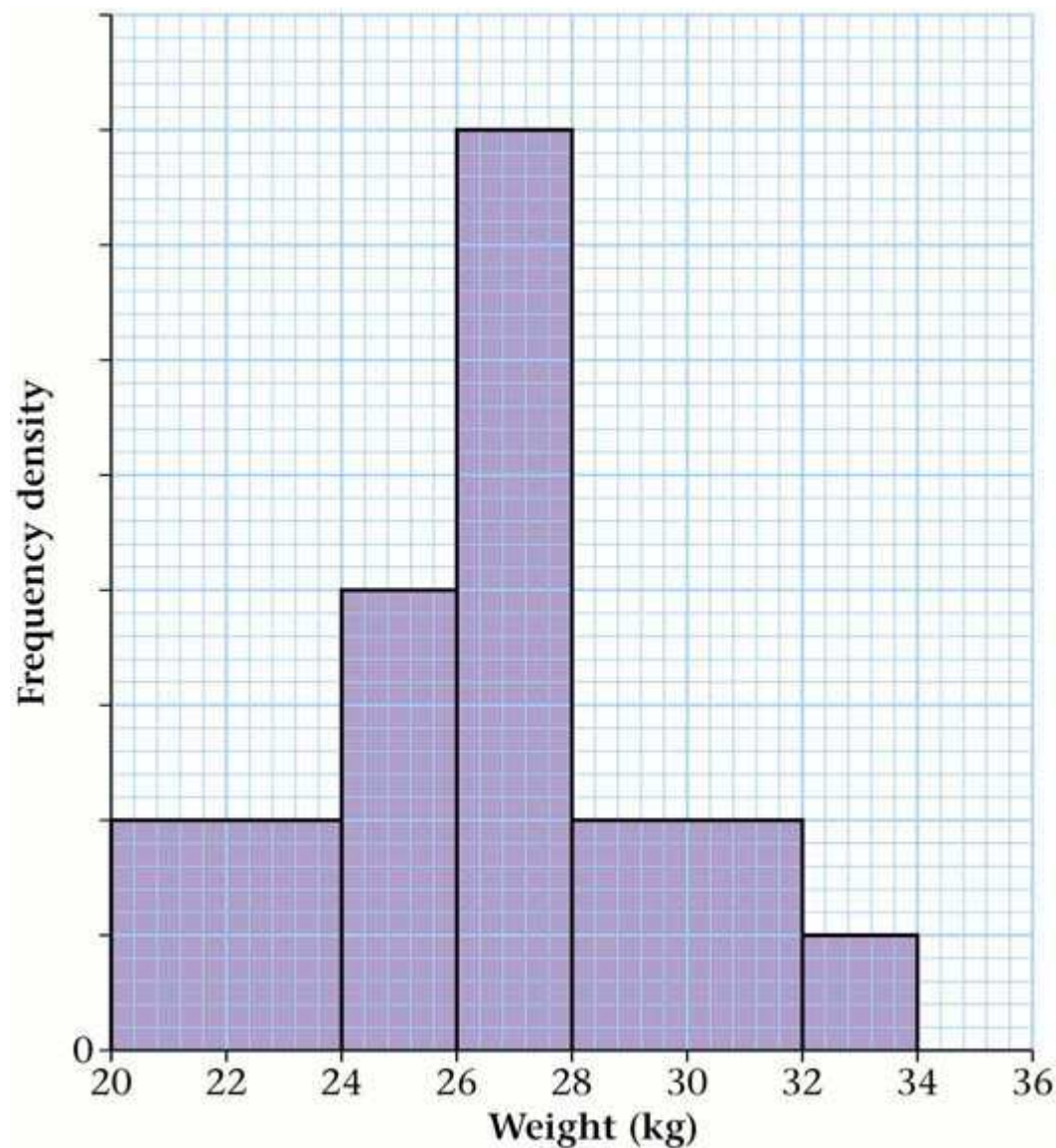Number of people $= 65 \times 2 = $ **130**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise E, Question 5

Question:

A farmer weighed a random sample of pigs. The weights were summarised in a grouped frequency table and represented by a histogram.

One of the classes in the grouped frequency distribution was 28–32 and its associated frequency was 32. On the histogram the height of the rectangle representing that class was 2 cm and the width was 2 cm.

a Give a reason to justify the use of a histogram to represent these data.

b Write down the underlying feature associated with each of the bars in a histogram.

c Show that on this histogram each pig was represented by 0.125 cm^2.

d How many pigs did the farmer weigh altogether?

e Estimate the number of pigs that weighed between 25 and 29 kg.

Solution:

a Weight is continuous data.

b The area of the bar is proportional to the frequency.

c Area for 28 – 32 kg is $2 \times 2 = 4$ units squared

Area for 1 pig is $4 \div 32 = 0.125$

d Total area in cm^2 is $2 \times 2 + 1 \times 4 + 1 \times 8 + 2 \times 2 + 1 \times 1 = 21$ units squared.

Number of pigs = $21 \div 0.125 = $ **168**

e Area for $25 - 29$ m is $\frac{1}{2} \times 4 + 1 \times 8 + \frac{1}{2} \times 2 = 11$ units squared.

Number of pigs = $11 \div 0.125 = $ **88**

© Pearson Education Ltd 2008

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise F, Question 1

Question:

In a survey of the earnings of some sixth form students who did Saturday jobs the median wage was £36.50. The 75th percentile was £45.75 and the interquartile range was £30.50.

Use the quartiles to describe the skewness of the distribution.

Solution:

1. $Q_1 = 45.75 - 30.5 = 15.25$

$Q_3 - Q_2 = 45.75 - 36.5 = 9.25$

$Q_2 - Q_1 = 36.5 - 15.25 = 21.25$

$Q_2 - Q_1 > Q_3 - Q_2$ therefore it is negatively skewed

© Pearson Education Ltd 2008

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise F, Question 2

Question:

A group of estate agents recorded the time spent on the first meeting with a random sample of 120 of their clients. The times, to the nearest minute, are summarised in the table.

Time	Number of clients
10–15	2
15–20	5
20–25	17
25–30	38
30–35	29
35–45	25
45–80	4
Total	120

a Calculate estimates of the mean and variance of the times.

b By interpolation obtain estimates of the median and quartiles of the times spent with customers.

One measure of skewness is found using $\frac{3(\text{mean} - \text{median})}{\text{standard deviation}}$.

c Evaluate this measure and describe the skewness of these data.

The estate agents are undecided whether to use the median and quartiles, or the mean and standard deviation to summarise these data.

d State, giving a reason, which you would recommend them to use.

Solution:

(a) Mean $= \frac{\sum fx}{n} = \frac{12.5 \times 2 + 17.5 \times 5 + 22.5 \times 17 + 27.5 \times 38 + 32.5 \times 29 + 40 \times 25 + 62.5 \times 4}{120}$

$\quad = 31.104166 = 31.1$ minutes

Variance $= \frac{\sum fx^2}{\sum f} - \mu^2 = \frac{12.5^2 \times 2 + 17.5^2 \times 5 + 22.5^2 \times 17 + 27.5^2 \times 38 + 32.5^2 \times 29 + 40^2 \times 25 + 62.5^2 \times 4}{120} - 31.1^2$
$\quad = 1045.36 - 967.31 = 78.05$

Standard deviation $= \sqrt{78.05} = 8.835$

(b) Median is the $\frac{120}{2} = 60$ th value so Q_2 is in class $25 - 30$

$\frac{Q_2 - 25}{30 - 25} = \frac{60 - 24}{62 - 24}$
$\qquad Q_2 = \textbf{29.7 minutes}$

$Q_1 : \dfrac{120}{4} = 30^{\text{th}}$ value so Q_1 is in class $25 - 30$

$\dfrac{Q_1 - 25}{30 - 25} = \dfrac{30 - 24}{62 - 24}$

$\qquad Q_1 = \mathbf{25.8 \ minutes}$

$Q_3 : 3 \times \dfrac{120}{4} = 90^{\text{th}}$ value so Q_3 is in class 30–35

$\dfrac{Q_3 - 30}{35 - 30} = \dfrac{90 - 62}{91 - 62}$

$\qquad Q_3 = \mathbf{34.8 \ minutes}$

(c) Skew $= \dfrac{3(31.10 - 29.74)}{8.835} = 0.46$

Positive skew.

(d) Median and quartiles because of the skew.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise F, Question 3

Question:

The following stem and leaf diagram summarises the wing length, to the nearest mm, of a random sample of 67 owl moths.

	Wing length											Key:5|0 means 50	
5	0 0 0 1 1 2 2 3 3 3 4 4												(12)
5	5 5 6 6 6 7 8 8 9 9												(10)
6	0 1 1 1 3 3 4 4 4 4												(10)
6	5 5 6 7 8 9 9												(7)
7	1 1 2 2 3 3												(6)
7	5 7 9 9												(4)
8	1 1 1 2 2 3 3 4												(8)
8	7 8 9												(3)
9	0 1 1 2												(4)
9	5 7 9												(3)

a Write down the mode of these data.

b Find the median and quartiles of these data.

c On graph paper, construct a box plot to represent these data.

d Comment on the skewness of the distribution.

e Calculate the mean and standard deviation of these data.

f Use a further method to show that these data are skewed.

g State, giving a reason, which of **b** or **e** you would recommend using to summarise the data in the table.
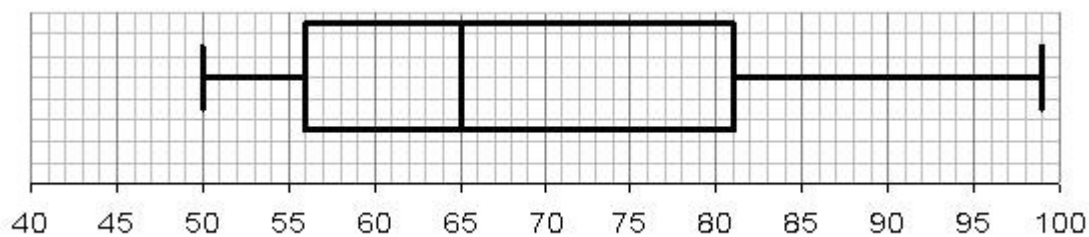
Solution:

(a) 64

(b) $Q_2 = \frac{67}{2} = 33.5$ therefore 34th term = 65

$Q_1 = \frac{67}{4} = 16.75 = 17^{th}$ term = 56

$Q_3 = 16.75 \times 3 = 50.25 = 51^{st}$ term = 81

(c)

(d) Positive skew

(e) mean 68.72 sd 13.73

(f) $Q_2 - Q_1 = 65 - 56 = 9$ and $Q_3 - Q_2 = 81 - 65 = 16$

$Q_2 - Q_1 < Q_3 - Q_2 \Rightarrow$ Positive skew

$\dfrac{3(Q_3 - Q_1)}{Q_2} = \dfrac{3(81 - 56)}{65} = 1.15 \Rightarrow$ Positive skew

(g) (b) because of the skew.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise F, Question 4

Question:

A TV company wishes to appeal to a wider range of viewers. They decide to purchase a programme from another channel. They have the option of buying one of two programmes. The company collects information from a sample of viewers for each programme. The results are summarised in the table. State which programme the company should buy to increase the range of their viewers. Give a reason for your answer.

	Mean age	Standard deviation of age
Programme 1	50	5
Programme 2	50	10

Solution:

Program 2 because it has a bigger standard deviation and hence bigger range in the age groups watching it.

© Pearson Education Ltd 2008

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise G, Question 1

Question:

Jason and Perdita decided to go for a touring holiday on the continent for the whole of July. They recorded the number of kilometres they travelled each day. The data are summarised in the stem and leaf diagram below.

stem	leaf								Key: 15|5 means 155 kilometres
15	5								
16	4	8	9						
17	3	5	7	8	8	8	9	9	9
18	4	4	5	5	8				
19	2	3	4	5	5	6			
20	4	7	8	9					
21	1	2							
22	6								

a Find Q_1, Q_2, and Q_3

Outliers are values that lie outside $Q_1 - 1.5(Q_3 - Q_1)$ and $Q_3 + 1.5(Q_3 - Q_1)$.

b Find any outliers.

c Draw a box plot of these data.

d Comment on the skewness of the distribution.

Solution:

(a) $Q_2 = \frac{32}{2} = 16$ therefore 16.5^{th} term $= 185$
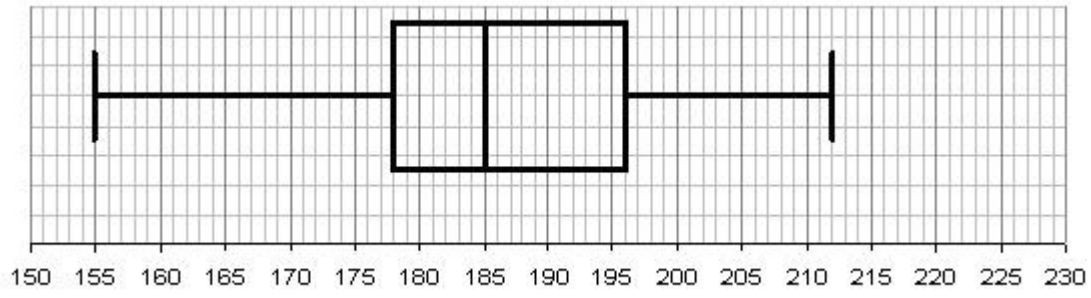
$Q_1 = \frac{31}{4} = 7\frac{3}{4} \Rightarrow 8^{th}$ term $= 178$

$Q_3 = \frac{3(31)}{4} = 23\frac{1}{4} \Rightarrow 24^{th}$ term $= 196$

(b) $Q_1 - 1.5(Q_3 - Q_1) = 178 - 1.5(196 - 178) = 151$

$Q_3 + 1.5(Q_3 - Q_1) = 196 + 1.5(196 - 178) = 223$ Outliers are 226.

(c)

(d) Positive skew

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise G, Question 2

Question:

Sophie and Jack do a survey every day for three weeks. Sophie counts the number of pedal cycles using Market Street. Jack counts the number of pedal cycles using Strand Road. The data they collected are summarised in the back-to-back stem and leaf diagram.

	Sophie	Stem	Jack								
	9 9 7 5	0	6 6								
7 6 5 3 3 2 2 2 1 1		1	1 1 5								
	5 3 3 2 2	2	1 2 2 2 3 7 7 8 9								
	2 1	3	2 3 4 7 7 8								
		4	2								

Key: 5|0|6 means Sophie counts 5 cycles and Jack counts 6 cycles

a Write down the modal number of pedal cycles using Strand Road.

The quartiles for these data are summarised in the table below.

	Sophie	Jack
Lower quartile	X	21
Median	13	Y
Upper quartile	Z	33

b Find the values for X, Y and Z.

c Write down the road you think has the most pedal cycles travelling along it overall. Give a reason for your answer.

Solution:

(a) 22

(b) $X = \frac{21}{4} = 5\frac{1}{4} \Rightarrow 6^{\text{th}}$ item = 11

$Z = \frac{3(21)}{4} = 15\frac{3}{4} \Rightarrow 16^{\text{th}}$ item = 22

$Y = \frac{21}{2} = 10.5$ therefore 11th item = 27

(c) Strand road has the highest median.

© Pearson Education Ltd 2008

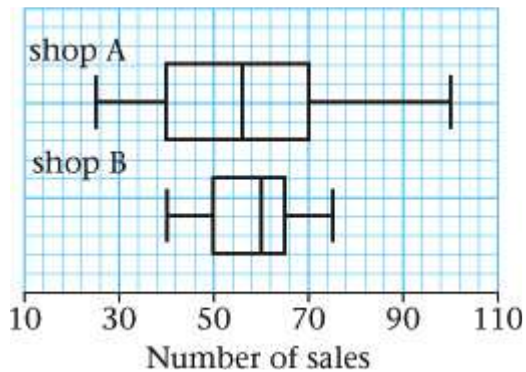# Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise G, Question 3

Question:

Shop A and Shop B both sell mobile phones.

They recorded how many they sold each day over a long period of time. The data they collected are represented in the box plots.

a Shop B says that for 50% of the days they sold 60 or more phones a day. State whether or not this is a true statement. Give a reason for your answer.

b Shop A says that for 75% of the days they sold 40 or more phones a day. State whether or not this is a true statement. Give a reason for your answer.

c Compare and contrast the two box plots.

d Write down the shop you think had the most consistent sales per day. Explain the reason for your choice.

Solution:

(a) This is a true statement. The median is 60 phones a day so for half the days they sold 60 or more.

(b) True. The lower quartile is 40.

(c) Shop A has a lower median and a bigger IQR/range. Overall shop A sells less phones but the daily quantity sold is more variable. Shop A had both the highest and the lowest number of phones sold in a day.

(d) Shop B as the interquartile range smaller.

© Pearson Education Ltd 2008

# Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise G, Question 4

Question:

4 Fell runners from the Esk Club and the Irt Club were keen to see which club had the fastest runners overall. They decided that all the members from both clubs would take part in a fell run. The time each runner took to complete the run was recorded.

The results are summarised in the box plot.
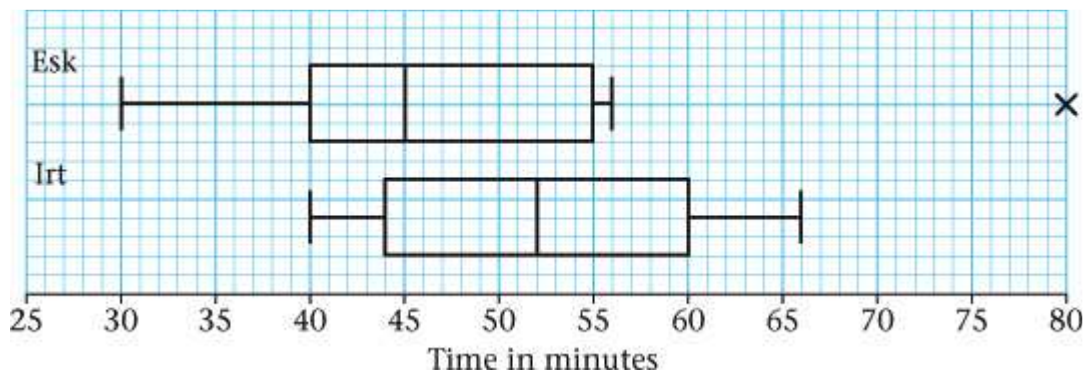
a Write down the time by which 50% of the Esk Club runners had completed the run.

b Write down the time by which 75% of the Irt Club runners had completed the run.

c Explain what is meant by the cross (×) on the Esk Club box plot.

d Compare and contrast these two box plots.

e Comment on the skewness of the two box plots.

f What conclusions can you draw from this information about which club has the fastest runners?

Solution:

(a) 45 minutes

(b) 60 minutes

(c) This is an outlier that does not fit the pattern.

(d) The Irt club had the highest median so overall they had the slowest runners.

The IQR ranges were about the same.

(e) Esk club times are positively skewed. Irt club times are symmetric.

(f) Esk had the fastest runners because they had the lower times.

© Pearson Education Ltd 2008

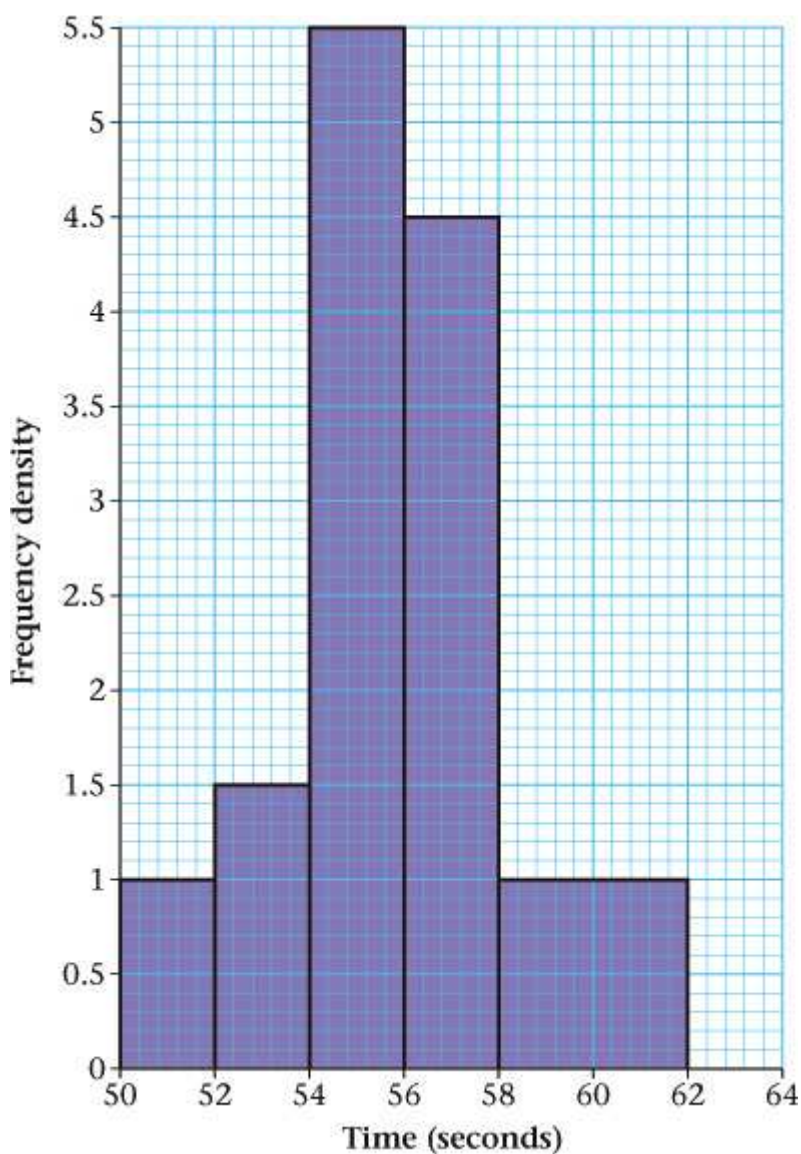# Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise G, Question 5

Question:

The histogram shows the time taken by a group of 58 girls to run a measured distance.

a Work out the number of girls who took longer than 56 seconds.

b Estimate the number of girls who took between 52 and 55 seconds.

Solution:

(a) Area = frequency

$2k(1 + 1.5 + 5.5 + 4.5) + 4k(1) = 58$

$29k = 58$ so $k = 2$

Number of girls who took longer than 56 seconds = 2 {(4.5 × 2) + (1 × 4)} = 26 girls

(b) Number of girls between 52 and 55 seconds = 2{(1.5 × 2) + (1 × 5.5)} = 17 girls

© Pearson Education Ltd 2008

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise G, Question 6

Question:

The table gives the distances travelled to school, in km, of the population of children in a particular region of the United Kingdom.

Distance, km	0–1	1–2	2–3	3–5	5–10	10 and over
Number	2565	1784	1170	756	630	135

A histogram of this data was drawn with distance along the horizontal axis. A bar of horizontal width 1.5 cm and height 5.7 cm represented the 0–1 km group.

Find the widths and heights, in cm to one decimal place, of the bars representing the following groups:

a 2–3,

b 5–10.

Solution:

$1.5 \times 5.7 \times k = 2565$, so $k = 300$

a width = 1.5 cm height = $\dfrac{1170}{2565} \times 5.7 = 2.6$ cm

b width = $5 \times 1.5 = 7.5$ cm height = $\dfrac{630}{300 \times 7.5} = 0.28$ cm

© Pearson Education Ltd 2008

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise G, Question 7

Question:

The labelling on bags of garden compost indicates that the bags weigh 20 kg.

The weights of a random sample of 50 bags are summarised in the table opposite.

Weight in kg	Frequency
14.6–14.8	1
14.8–18.0	0
18.0–18.5	5
18.5–20.0	6
20.0–20.2	22
20.2–20.4	15
20.4–21.0	1

a On graph paper, draw a histogram of these data.

b Estimate the mean and standard deviation of the weight of a bag of compost.

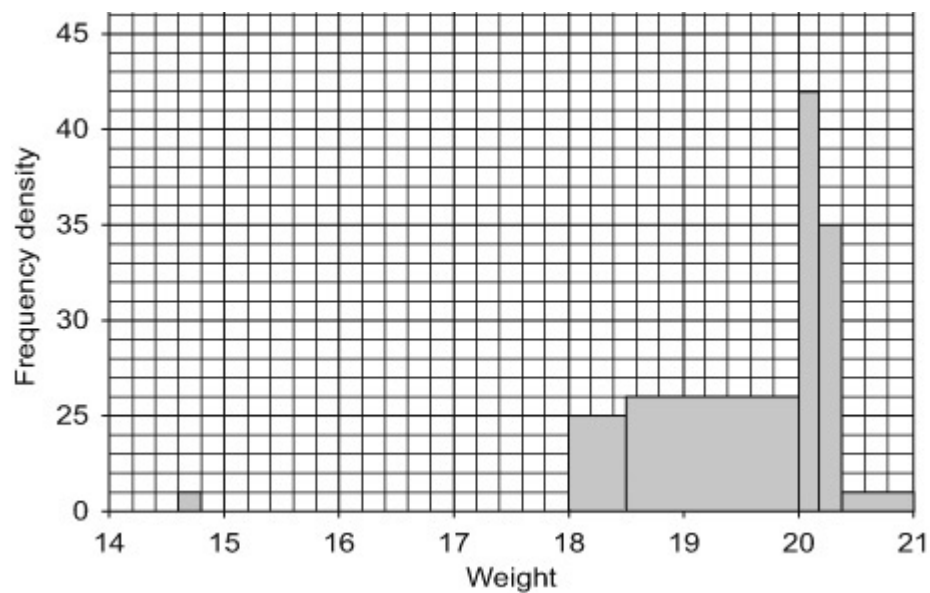[You may use $\Sigma fy = 988.85$, $\Sigma fy^2 = 19\ 602.84$]

c Using linear interpolation, estimate the median.

One coefficient of skewness is given by

$\frac{3(\text{mean} - \text{median})}{\text{standard deviation}}$.

d Evaluate this coefficient for the above data.

e Comment on the skewness of the distribution of the weights of bags of compost.

Solution:

(b) Mean $= \dfrac{\Sigma fy}{n} = \dfrac{988.85}{50} = 19.777$

sd $= \sqrt{\dfrac{\Sigma fy^2}{n} - \mu^2} = \sqrt{\dfrac{19602.84}{50} - 19.777^2} = \sqrt{0.927} = \sqrt{0.927} = 0.963$

(c) median $= 20 + \dfrac{13}{22} \times 0.2 = 20.118181818 = 20.118$

(d) $\dfrac{3(19.777 - 20.118)}{0.963} = -1.0623 = -1.06$

(e) The distribution of the weights of bags of compost is negatively skewed.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise G, Question 8

Question:

The number of bags of potato crisps sold per day in a bar was recorded over a two-week period. The results are shown below.

20   15   10   30   33   40   5   11   13   20   25   42   31   17

a Calculate the mean of these data.

b Draw a stem and leaf diagram to represent these data.

c Find the median and the quartiles of these data.

An outlier is an observation that falls either $1.5 \times$ (interquartile range) above the upper quartile or $1.5 \times$ (interquartile range) below the lower quartile.

d Determine whether or not any items of data are outliers.

e On graph paper draw a box plot to represent these data. Show your scale clearly.

f Comment on the skewness of the distribution of bags of crisps sold per day. Justify your answer.

Solution:

(a) 22.285714 bags

(b)
Number of bags

0	5				
1	0	1	3	5	7
2	0	0	5		
3	0	1	3		
4	0	2			

key 1|1 = 11 bags

(c) Median = $\frac{15}{2}$ = 7.5 therefore 8th item = 20

$Q_1 = \frac{14}{4} = 3.5 \Rightarrow 4^{\text{th}}$ item = 13

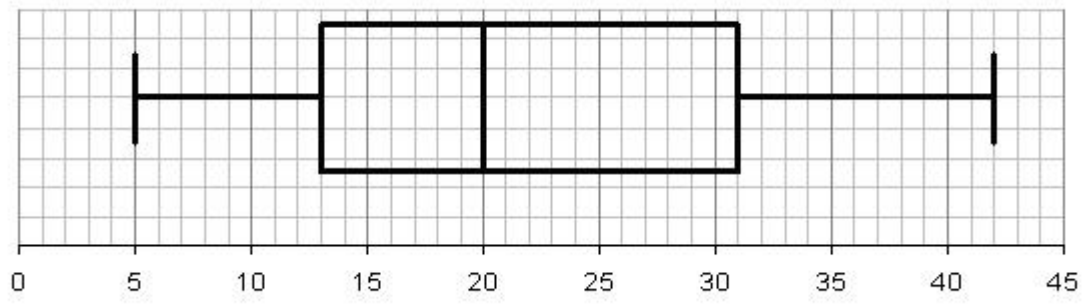$Q_3 = \frac{3(14)}{4} = 10.5 \Rightarrow 11^{\text{th}}$ item = 31

(d) IQR = 31 − 13 = 18 so $1.5 \times$ IQR = 27

13 − 27 = −14

31 + 27 = 58

There are no outliers

(e)

(f) Positive skew. $Q_2 - Q_1 < Q_3 - Q_2$