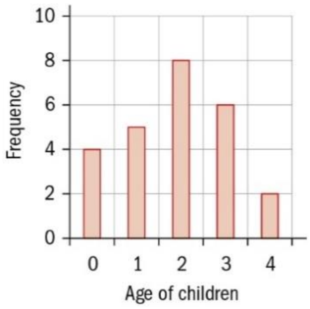
3 Representing and describing data: descriptive statistics

Skills check

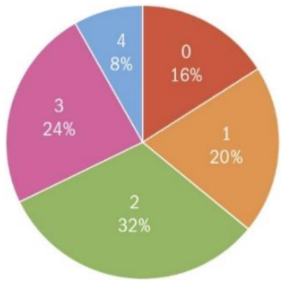
1 **a** ☺ = 1 child

0	☺☺☺☺
1	☺☺☺☺☺
2	☺☺☺☺☺☺☺☺
3	☺☺☺☺☺☺
4	☺☺

b

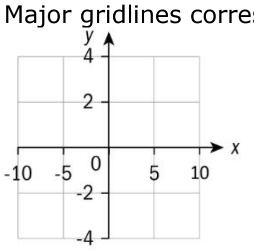


c



2 Mean = 7, Median = 8, Mode = 9, Range = 6.

3 Major gridlines correspond to 1cm.



Exercise 3A

1 **a** Discrete **b** Continuous **c** Discrete **d** Discrete
e Discrete **f** Continuous **g** Continuous **h** Discrete
i Continuous **j** Continuous

2

Number of sweets	Frequency
21	4
22	6
23	5
24	5
25	4
26	1

3

Height, in metres, *h*	Frequency
$2 \leq h < 3$	6
$3 \leq h < 4$	6
$4 \leq h < 5$	5
$5 \leq h < 6$	3

4

Weight, in kilograms, w	Frequency
$0 \le w < 10$	11
$10 \le w < 20$	8
$20 \le w < 30$	3
$30 \le w < 40$	3

Exercise 3B

1 **a** Mean = 10.1
Median = 9.0
Mode = 8.6

The most appropriate measures for this case are the mean and the median, because the data is continuous.

b Mean = 8.64
Median = 8.5
Mode does not exist

The most appropriate measures for this case are both mean and median. This is an example of a continuous data set where mode does not exist.

c Mean =32.62
Median = 30
Mode = 15

The most appropriate measure for this case is both mean and median as mode is clearly too small.

2 Find the modal class by determining which of the modal classes has the highest frequency. To calculate the mean and median, use mid values of the class intervals.

a The modal class: $150 \le n < 180$
an approximation for the mean: 112
an approximation for the median: 105

As it can be seen from the frequency table, the data in this case is *skewed,* which is also indicated by the modal class being at the highest end of the range. In this situation, an approximation for the median is the most appropriate measure of the central tendency as it is less affected by the skewed values.

b The modal class: $50 \le s < 55$
an approximation for the mean: 54.4
an approximation for the median: 52.5

The data set is well centred with all three measures agreeing well. Hence, the best measure of the central tendency in this case is the approximate mean which minimises the error for the guess of the next value.

c The modal class: $7 \le t < 8$
an approximation for the mean: 5.86
an approximation for the median: 5.5

The data does not have clear tendency. Best measure to use is the approximation for the median because the modal class is very high.

Exercise 3C

1 **a** Mean = 6.1, median = 5.2, mode = 7.5.

Possible outliers of the data set are 17.8 and 25. Excluding them from the calculation of mean and median gives the following values: mean = 4.7, median = 4.8, which are closer together than when the outliers are included.

b Mean = 3.5, median = 3.6, mode = 2.5.

Possible outlier is 6.1 as it is further away from the rest of the data points. Excluding 6.1 from the data set changes the mean to 3.4 but leaves median unchanged. The mean is now further from the median. Hence, 6.1 is not an outlier.

c Mean = 65, median = 62, mode = 62.

Possible outlier is 22 as it is further away from the rest of the data points. Excluding 22 from the data set changes the mean to 67 but leaves the median unchanged. Hence, 22 is not an outlier.

Exercise 3D

1 a Discrete

b It is useful to produce a frequency table of the data.

Number of daisies	Frequency
2	1
3	1
4	1
5	1
6	2
8	2
9	1
11	1
12	3
13	2
15	5
16	2
17	1
18	1
21	1
22	1
24	1
25	1
26	1
34	1

Mean = 13.9

Mode = 15

Median = 14

Since this is a discrete data set, **mode** is more appropriate measure of the central tendency, although all three measures agree well in this case.

c Find standard deviation using the formula $\sigma_n = \frac{1}{n}\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}$: $\sigma_n = 7.3$.

Since the standard deviation shows how the data values are related to the mean, the result indicates that the data points are quite spread out.

d Range = 34 – 2 = 32.

Interquartile range:

Q1 is the $\frac{30+1}{4} \approx 8^{th}$ number (a quarter of the data points are below this number): Q1 = 8

Q3 is the $\frac{3(30+1)}{4} \approx 23^{th}$ number (three quarters of the data points are below this number):

Q3 = 17

IQR = 17-8 = 9.

2 a Modal class: $30 \le c < 40$.

b Take middle values of each interval to obtain the estimates:

an estimate for the mean: 34

an estimate for the median: 35.

c 10.12

Since the standard deviation shows how the data values are related to the mean, the result indicates that the data points are very spread out.

d Variance = 102.3, Q_1 is the $\frac{60+1}{4} \approx 15^{th}$ number, Q_1 = 25, Q_3 is the $\frac{3(60+1)}{4} \approx 46^{th}$ number, Q_3 = 35. IQR = 35 – 25 = 10. They are all estimates because they use the mid value of the class intervals.

3 The new mean is US $3600, the new standard deviation is US $250. When every data point is shifted by an equal amount, the mean shifts by the same amount while standard deviation does not change (data is spread out by the same amount).

4 **a** Mean = 8.9, median = 10, mode = 12. Since mode is quite a bit higher than the other two measures, the mean and the median are the most appropriate to use.

b Standard deviation = 4.10 which shows that the data points have medium spread in relation to the mean.

c Range = 14 – 0 = 14. Q_1 is the $\frac{36+1}{4} \approx 9^{th}$ number, Q_1 = 6, Q_3 is the $\frac{3(36+1)}{4} \approx 28^{th}$ number, Q_3 = 12. IQR = 12 – 6 = 6. Range and IQR are relatively big implying that the data is reasonably spread out.

5 The new mean is 60, the new standard deviation is 6. When each data point is modified by a multiplier, both the mean and the standard deviation is modified by the same amount, i.e. both the location and the spread of the data is changed.

6 **a** Modal class = $180 \le x < 190$.

b Use mid interval values to estimate both measures: mean $\approx$ 180.2, standard deviation $\approx$ 11.0, so the data are strongly spread out.

7 **a** Mrs Ginger's new mean = 84, Mrs Ginger's new standard deviation = 16;

Mr Ginger's new mean = 80, Mr Ginger's new standard deviation = 20;

Miss Ginger's new mean = 76, Miss Ginger's new standard deviation = 24.

See questions 3 & 5 for explanations.

b New grades:

	Mrs Ginger	**Mr Ginger**	**Miss Ginger**
y	44	30	16
Zoe	70	62.5	55
Ans	92	90	88

8 **a** Basketball players: mean = 200.8, standard deviation = 10.5; males: mean = 172.3, standard deviation = 10.3.

b Basketball players are much taller on average, however both samples have nearly the same standard deviation, so the spread of the data is nearly the same.

9 **a** Males: mean = $2546.30, standard deviation = $729.78; Females: mean = $2114.58, standard deviation = $635.25.

b Male salaries have higher mean and higher spread than female salaries.

Exercise 3E

1 **a** Mean height of the whole school = 155.4.

b There are several ways to sample this data. For example, use stratified sampling technique taking an appropriate number of students from each grade: there are 160 students in total, so if a sample of 50 students should be chosen, take 9,9,8,8,8,8 students from grades 7,8,9,10,11,12 respectively. The final answer will differ depending on the random numbers. This method will be unbiased.

2 **a** Mean age of the 100 people = 25.92 which is well below 60 and the manager should not lose much revenue.

b Use a random number generator to obtain 35 numbers between 1 and 100 and average the ages represented by those numbers. The answers will differ.

c The result depends on the starting point. Starting with the first number the mean = 26.21.

d Usually the systematic sample mean will be closer to the population mean than the random sample mean. However, that depends heavily on the set of the random numbers used.

3 **a** Mean number of goals scored in all 50 matches = 3.58.

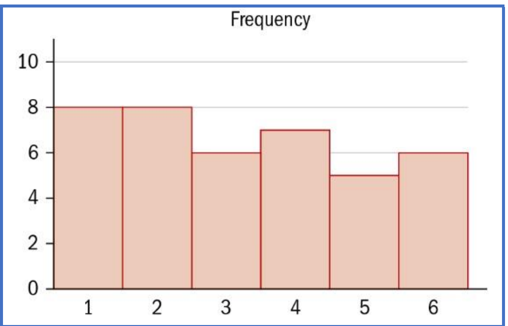
b Generate a set of 24 random numbers between 1 and 50. Calculate the mean.

c The estimated mean should be reasonably close to the actual mean but the answers can differ.

Exercise 3F

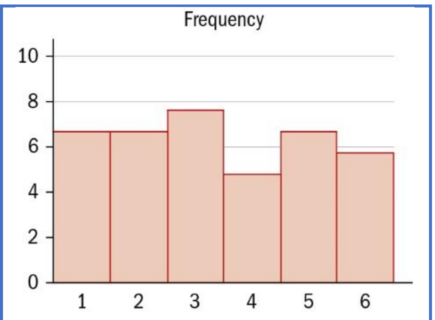
1 Theo's data:

Number	Frequency
1	8
2	8
3	6
4	7
5	5
6	6



Millie's data:

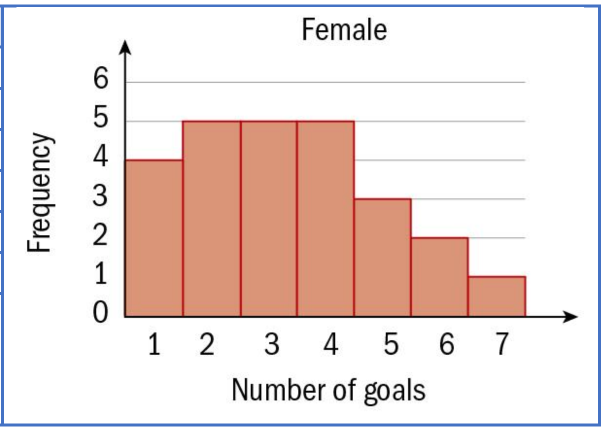
Number	Frequency
1	7
2	7
3	8
4	5
5	7
6	6



Comment: the distribution of the numbers are similar for both Theo and Millie and the frequencies for each number thrown are very similar.

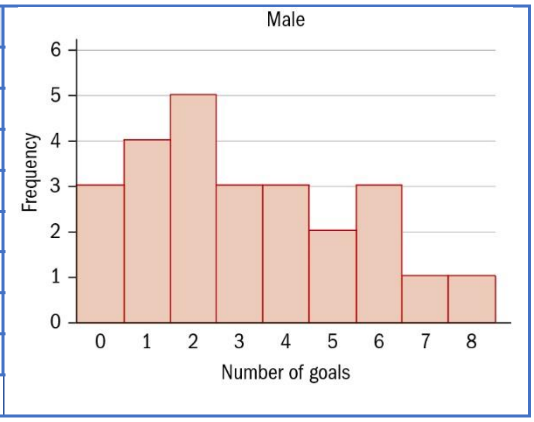
2 Female data:

Number of goals	Frequency
0	4
1	5
2	5
3	5
4	3
5	2
6	1



Male data:

Number of goals	Frequency
0	3
1	4
2	5
3	3
4	3
5	2
6	3
7	1
8	1

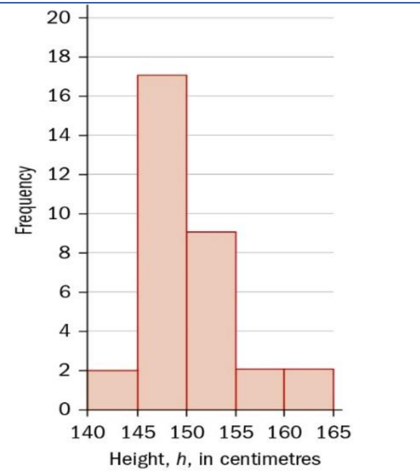


The range of the number of goals is bigger for males. The female data is more uniform than male data and more females scored no goals.

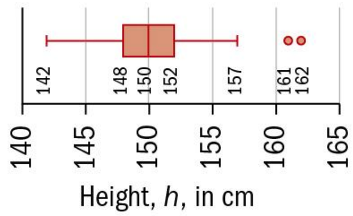
3 Frequency table and histogram:

a , b

Height, in centimetres, *h*	**Frequency**
$140<h\le145$	2
$145<h\le150$	17
$150<h\le155$	9
$155<h\le160$	2
$160<h\le165$	2



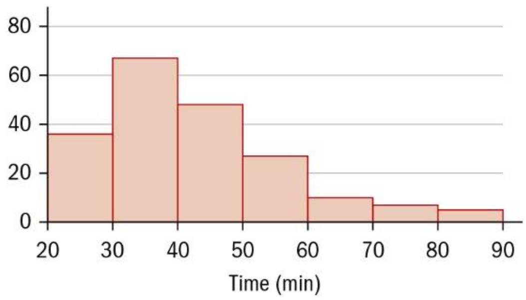
c



d

Use the box-and-whisker plot to determine whether the data is symmetric or not. From the box and whisker diagram, the data is not symmetric. When outliers are excluded, the data is symmetric.

4 a Histogram:



Time in minutes:

Mean = 42

Median = 35

LQ = 35

UQ = 45

Range = 70

Outliers: data points in $60\le x\le90$.

To find approximate values of the mean, median, LQ and UQ, use the midpoints of the given groups. These are approximations only because the original data have not been given.

To find median, identify the midpoint of the interval which contains the data point above the 100th point.

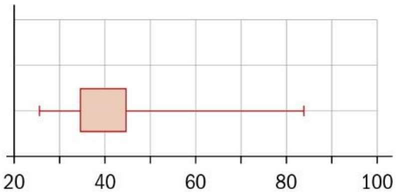
To find LQ, identify the midpoint of the interval which contains the data point above the 50th point.

To find UQ, identify the midpoint of the interval which contains the data point above the 150th point.

Data points are spread out between 20 and 90 minutes, hence the approximate range is 90-20=70 mins.

To determine the outliers, calculate IQR = UQ-LQ = 45-35 = 10. Then, outliers are the data points below LQ − 1.5 ×. IQR =20 and above UQ + 1.5 ×. IQR =60.

c Outliers are omitted from the box-and-whiskers graph. Note that median and LQ are equal.



d Marcus did worse than 75% of participants, so he may not be satisfied.

5 Boys and girls

a Boys:

Mean = 6

Median = 6

LQ = 5

UQ = 7

Range = 7

Girls:

Mean = 6

Median = 6

LQ = 5

UQ = 7

Range = 8

There are no outliers.

To find the means, multiply the score by its frequency, add up and divide by the number of girls or boys.

To find the median, find what score the 26th data point in each of the sets represents.

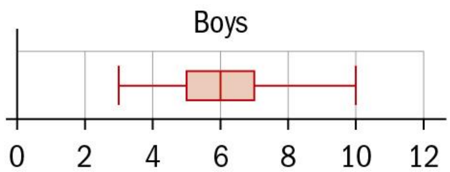
To find the LQ, find what score the 13th data point in each of the sets represents.

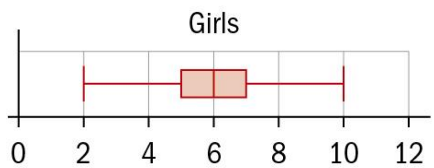
To find the UQ, find what score the 38th data point in each of the sets represents.

Note, range is the only different parameter between the girls and boys so far.

To find the outliers, calculate IQR = UQ − LQ = 7 – 5 = 2. Then, outliers are the data points below LQ – 1.5 × . IQR = 2 and above UQ + 1.5 × . IQR = 10. Hence, there are no outliers.

b Plots are almost identical apart from the minimum value being different (boys have higher minimum value).





c Whilst the boys data is not as symmetrical as that of the girls, both are clearly quite symmetrical.

6 a Time in minutes:

Mean = 17

Median = 16

LQ = 12

UQ = 21

Range = 37

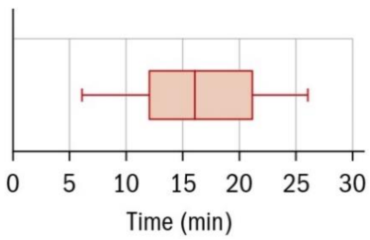
Outliers: 35, 43.

Order the data points from the smallest to the largest to find the quartiles. Median is the 18th data point, LQ is the 10th data point and UQ is the 28th data point.

Range = 43-6 = 37

To determine the outliers, calculate IQR = UQ-LQ = 21-12=9. Then, outliers are the data points below LQ – 1.5 × . IQR =-1.5 and above UQ + 1.5 × . IQR =34.5. Hence, the outliers are 35 and 43.

b .



c Since 16 is the median, there are 17 students who took longer to complete the puzzle. (Note, 16 students is also a valid answer, as there are two data points with the value of 16).

7 a Median = 3.

b IQR = 4 − 2 = 2.

c The data is almost symmetrical: the data between UQ and the maximum value are slightly more spread out than between the LQ and the minimum value.

8 **Boys**

a Median = 55

b IQR = 75-45 = 30

c 25% of boys scored between 45 and 55

e 15 boys scored below 45 (25% of 60).

Girls

a Median = 65

b IQR = 80-50 = 30

d 50% of the girls scored between 65 and 95 because 100% of the girls scored below 95

f 45 girls scored above 50 (75% of 60).

g Neither of the data sets are perfectly symmetrical. Girls data are symmetrical in the interquartile range, but more spread out between the UQ and the maximum value than between the LQ and the minimum value. Boys data are more spread out between the median and the maximum value than between the median and the minimum value.

9 **a** Median = 120.

b Range = 150 − 80 = 70.

c 10 pandas weigh less than 90kg (LQ = 90kg, there are 25% of pandas that weigh less than 90 kg and 25% of 40 is 10).

d 50% of pandas weigh between 120 and 160 kg (all the pandas above the median).

e 20 pandas weigh between 90 and 130 kg (LQ = 90, UQ = 130, there are 50% of pandas between these values and 50% of 40 is 20).

f Since the average weight corresponds with the median of the sample, we can deduce that the distribution of the weight of the pandas is skew with respect to the average weight.

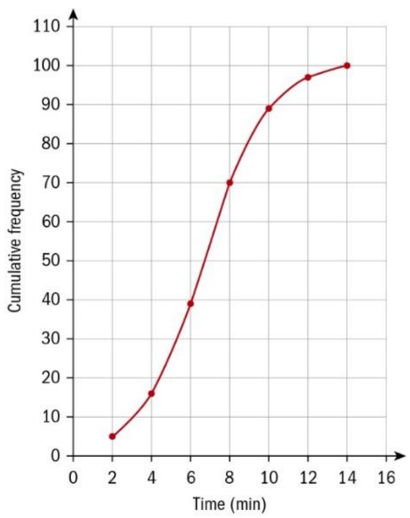
g Pandas in the sample must be mostly males because the box plot overlaps more with the range of male pandas weights.

Exercise 3G

1 **a**

Upper boundary	Cumulative frequency
2	5
4	16
6	39
8	70
10	89
12	97
14	100

b



c Median ≈ 6.75

IQR ≈ 3.5

To determine median, LQ and UQ, read off the time for cumulative frequency 50, 25 and 75 respectively. IQR = UQ − LQ = 8.5 − 5 = 3.5 mins.

(In minutes)

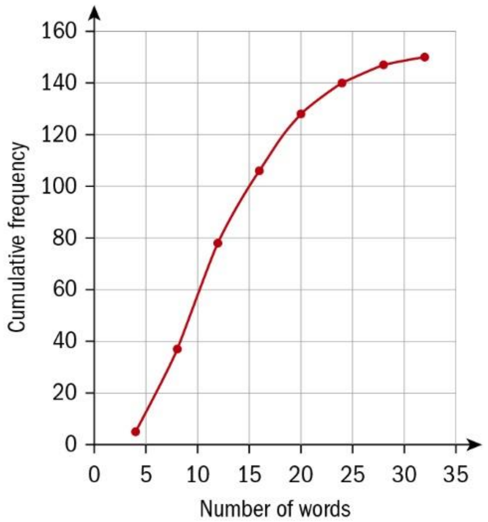
d 10^{th} percentile = 3 minutes — It is possible to use either the frequency table and the midpoints of the intervals or the cumulative frequencies graph to estimate the value of 10^{th} percentile.

e 5 people — Use the graph to read off that 11 minutes corresponds to 95^{th} percentile.

2 a

Upper boundary	Cumulative frequency
4	5
8	37
12	78
16	106
20	128
24	140
28	147
32	150

b



c Median ≈ 12
IQR ≈ 10

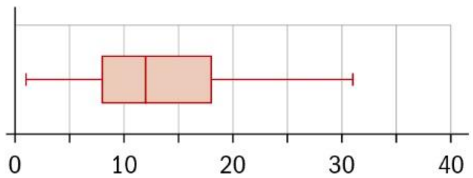
To determine median, LQ and UQ, read off the number of words for cumulative frequencies 80, 40 and 120 respectively.

IQR = UQ – LQ = 18 – 8 = 10 words.

d There are no outliers. — Outliers would be found below LQ – 1.5 × IQR = 8 – 15 = −7 and above UQ + 1.5 × IQR = 18 + 15 = 33.

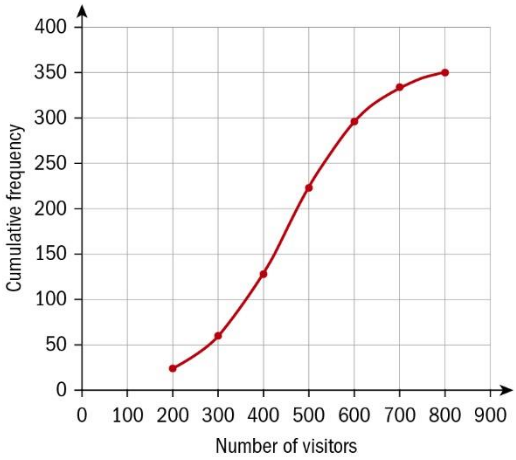
e 22 — 90^{th} percentile corresponds to 160 × 0.9 = 135 cumulative frequency data point. Read off the number of pages from the graph.

f



g Adult books because the mean 15 falls inside the interquartile range while the mean 8 does not.

3 a Cumulative frequency graph.



b Median ≈ 450

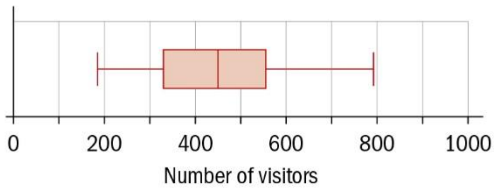
IQR ≈ 205

To determine median, LQ and UQ, read off the number of visitors for cumulative frequencies 175, 88 and 263 respectively. IQR = UQ – LQ = 555 – 350 = 205 visitors.

c There are no outliers.

Outliers would be found below LQ – 1.5 × IQR = 43 and above UQ + 1.5 × IQR = 863.

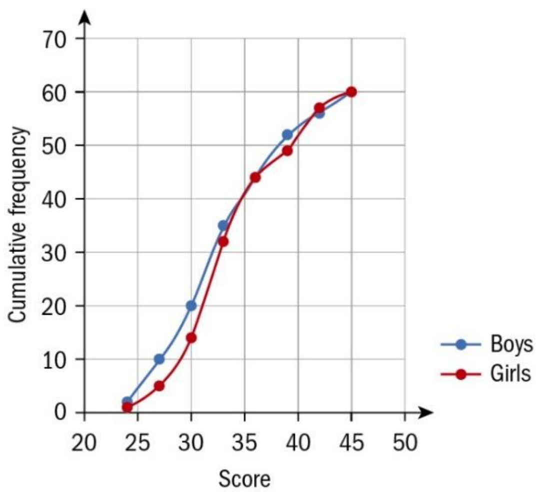
d



e 90 days.

Read off the cumulative frequency for 350 visitors.

4 a



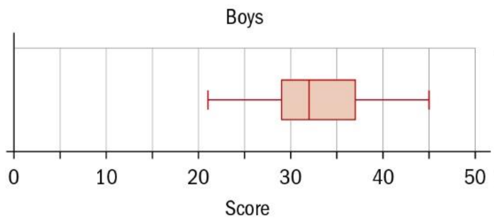
Both, the cumulative frequency and box-and-whiskers graphs are useful for analysing this situation.

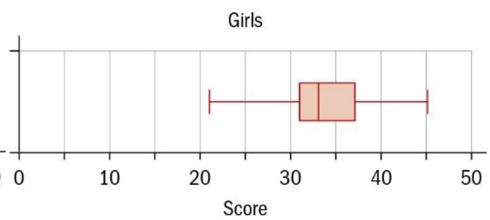
Use cumulative frequency graph to determine the median, LQ and UQ for both boys and girls.

Boys:
Median=32, LQ=29, UQ= 37.

Girls:

Median=33, LQ=31, UQ=37.





Assume 21 as the lowest and 45 as the highest scores for both girls and boys.

b Both graphs suggest fairly similar results. From the cumulative frequency graph it is seen that more boys than girls acquired the lower range scores. From the box-and-whiskers graph it is seen that both data sets are not perfectly symmetric. It is seen that the interquartile range of the boys' scores is wider than the girls' scores. The median score of the girls is just above the median score of the boys.

c Martin's score is the lower quartile for the boys, so Martin did better than 25% of the boys. however, Mary's score is below the lower quartile for the girls, so more than 75% of girls did better than Mary.

5 a Median ≈ 95

IQR ≈ 60

90th percentile ≈ 150.

To determine median, LQ and UQ, read off data points for cumulative frequencies 170, 85 and 255 respectively.

IQR = UQ − LQ = 130 − 70 = 60.

0.9*340 = 306.

b There are no outliers.

Outliers would be found below LQ − 1.5 × IQR = −20 and above UQ + 1.5 × IQR = 220.

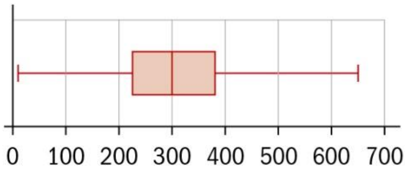
6 a Median ≈ 300

LQ ≈ 225

UQ ≈ 380

To determine median, LQ and UQ, read off data points for cumulative frequencies 50, 25 and 75 respectively.

b



c

Length, *l*, in cm	Cumulative frequency
50 ≤ *l* < 100	5
100 ≤ *l* < 150	5
150 ≤ *l* < 200	10
200 ≤ *l* < 250	12
250 ≤ *l* < 300	18
300 ≤ *l* < 350	15
350 ≤ *l* < 400	13
400 ≤ *l* < 450	8
450 ≤ *l* < 500	4
500 ≤ *l* < 550	4
550 ≤ *l* < 600	3
600 ≤ *l* < 650	2
650 ≤ *l* ≤ 700	1

d Mean ≈ 312, Standard deviation ≈ 132. Estimates found by using the midpoint of each interval.

Exercise 3H

1 a Negative, moderate.

b Positive, strong.

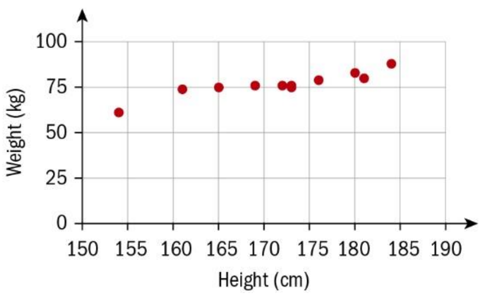
c Negative, strong.

d Positive, weak.

e No positive or negative correlation.

f No positive or negative correlation (but could split into intervals of strong positive and strong negative correlations).

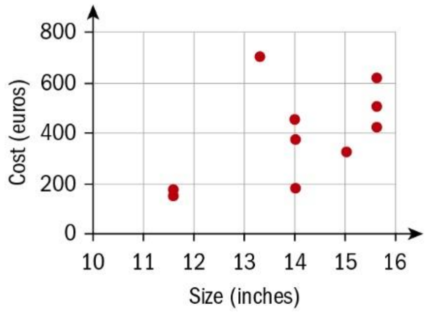
2 a



b There is a strong positive correlation between the heights and the weights of the football players. Hence, the taller the football player, the heavier he or she is expected to be.

c The correlation can indicate causation in this case because taller people have more body mass.

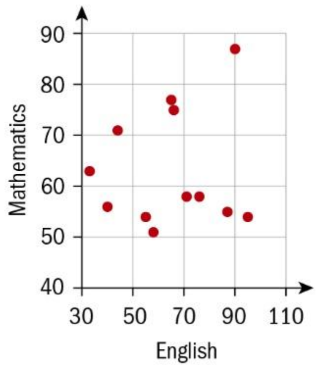
3 a



b There is a weak positive correlation between the size of the laptop screen and the cost of the laptop.

c Since the correlation is weak, the size of the screen has little influence on the cost of the laptop. This can be expected as there are many other factors that contribute towards the price, for example brand, type of the processor, whether or not the screen is touch sensitive.

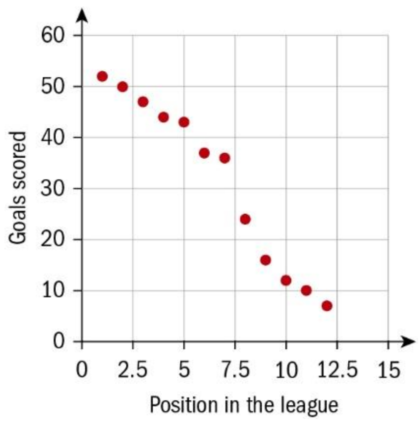
4 a



b There is no visible positive or negative correlation.

c The grade for the English test does not influence the grade for the Mathematics test and vice versa.

5 a



b There is strong negative correlation between the team's position in the league and the goals scored.

c Position in the league can be influenced by the goals scored.

Chapter review

1 **a** Discrete

Number of apples	**Frequency**
7	4
8	4
9	4
10	3
11	1
12	1

b Continuous. Group the data into convenient intervals to produce a frequency table.

Lengths, in cm, l	**Frequency**
$7 \le l \le 9$	5
$9 \le l \le 11$	2
$11 \le l \le 13$	2
$13 \le l \le 15$	4

c Discrete

Size	**Frequency**
33	1
34	2
35	4
36	3
37	4
38	3

2 **a** Mean = 54.8, median = 56, mode = 32, median and mean are best because this is continuous data.

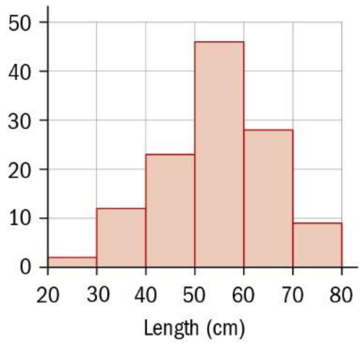
b Mean = 122.25, median = 87.5, mode = 62. Data is very skewed hence it is best to use median in this case.

c Mean = 6.42, median = 6, mode = 6. All measures agree well hence it is fine to use any of them.

3 **a** $50 \le l < 60$ since frequency is highest for this class.

b Use mid values of the intervals to estimate median ≈ 55, mean ≈ 54.42, standard deviation ≈ 11.28.

c Histogram:



4 The new mean is 76, the new standard deviation is 14. When each data point is modified by a multiplier, both the mean and the standard deviation is modified by the same amount, i.e. both the location and the spread of the data is changed.

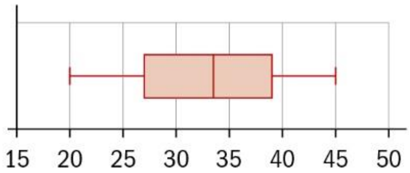
5 The new mean is 13, the new standard deviation is 1. When every data point is shifted by an equal amount, the mean shifts by the same amount while standard deviation does not change (data is spread out by the same amount).

6 **a** Mean = 32.8, standard deviation = 7.51 which indicates that the data points are reasonably spread out.

b Range = 45 – 20 = 25, IQR = Q3 − Q1 = 39 – 27 = 12.

c Min = 20, LQ = 27, Median = 34, UQ = 39, max = 45. Then, outliers are the data points below LQ – 1.5 × IQR = 9 and above UQ + 1.5 × IQR = 57. Min and max are within this range, so there are no outliers.

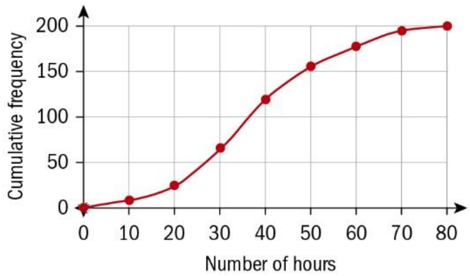
d



7 **a** Cumulative frequency table:

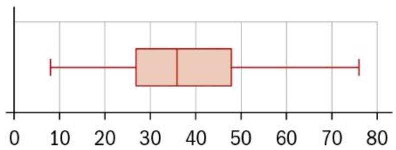
Upper boundary	**Cumulative frequency**
10	8
20	24
30	65
40	119
50	155
60	177
70	194
80	200

b

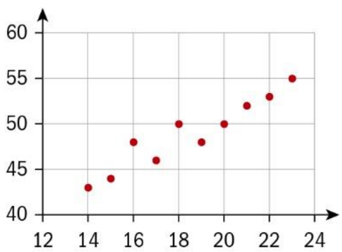


c Read off cumulative frequency 50, 100 and 150 for LQ = 27, Median = 36 and UQ = 48 respectively to give IQR = 21.

d



8 **a**



b Strong and positive correlation

c It is unlikely that weather is the cause of the increase/decrease of the number of eggs.

9 **a** $40 \le x < 60$ A1

b Use of mid-point M1

correct method M1

48.3% A1

c 22.6 M1A1

d The second class had slightly lower marks on average. R1

Their standard deviation was also lower, meaning their marks were more consistent, or less spread out than the first class. R1

10 a $\frac{398}{25} = 15.92$ g M1A1

b 13^{th} mouse has weight of 16 g M1

So median is 16 g A1

c $Q_1 = 14$ A1

$Q_3 = 19$ A1

Interquartile range is $Q_3 - Q_1 = 19 - 14 = 5$ M1A1

d $Q_3 + 1.5 \times \text{IQ range} = 19 + 1.5 \times 5 = 26.5$ g M1A1

11 a Using GDC, mean $= 23.58^\circ C$ M1A1

b Using GDC, SD $= 3.38^\circ C$ M1A1

c Using GDC, mean $= 22.83^\circ C$ M1A1

d Using GDC, SD $= 5.52^\circ C$ M1A1

e Tenerife has a higher mean temperature $(23.58^\circ C)$, so on average, temperatures could be said to be higher in Tenerife. R1R1

The standard deviation of Tenerife temperatures $(3.38^\circ C)$ is lower than that in Malta, therefore the temperatures can also be said to be more consistent. R1R1

12 a $5.25 + 3 = 8.25$ years M1A1

b 1.2 years M1A1

13 Total population is 65.12 million. A1

Number required from England: $\frac{54.8}{65.12} \times 5000 = 4208$ M1A1

Number required from Wales: $\frac{3.10}{65.12} \times 5000 = 238$ A1

Number required from Scotland: $\frac{5.37}{65.12} \times 5000 = 412$ A1

Number required from Northern Ireland: $\frac{1.85}{65.12} \times 5000 = 142$ A1

14 a $45 - 15 = 30$ M1A1

b $37 - 22 = 15$ M1A1

c 75% A1

d Interquartile range is 15

$Q_3 + 1.5 \times \text{IQ range} = 37 + 1.5 \times 15 = 59.5$ minutes M1A1

Yes, it would count as an outlier A1

15 a 69 M1A1

b 50 M1A1

c 83 M1A1

d 170 M1A1

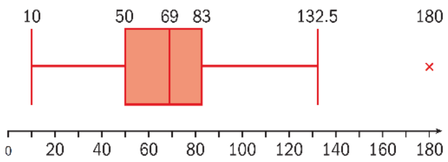
e Interquartile range is $83 - 50 = 33$ A1

$Q_3 + 1.5 \times \text{IQ range} = 83 + 1.5 \times 33 = 132.5$ M1

$Q_1 - 1.5 \times \text{IQ range} = 50 - 1.5 \times 33 = 0.5$ M1

Therefore 180 is an outlier A1

f



M1A1A1

16 a $45 \le x < 60$ A1

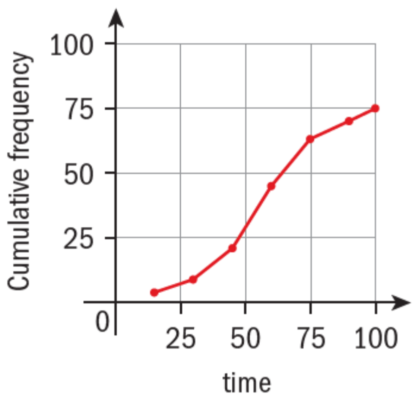
b $45 \le x < 60$ M1A1

c

Time, (t minutes)	c.f.
$0 \le x < 15$	4
$15 \le x < 30$	9
$30 \le x < 45$	21
$45 \le x < 60$	45
$60 \le x < 75$	63
$75 \le x < 90$	70
$90 \le x < 100$	75

M1A1A1

d



M1A1

e From graph, 37.5 on y-axis gives median of 55 M1A1

Interquartile range is approximately $70 - 40 = 30$ M1A1

17 a Using GDC, mean = $ 43 600 M1A1

b Median = $\frac{1}{2}(11+1)^{\text{th}}$ value = 6^{th} value, which is $ 25 000 M1A1

c LQ = $\frac{1}{4}(11+1) = 3^{\text{rd}}$ value, which is $ 25 000

UQ = $\frac{3}{4}(11+1)^{\text{th}} = 9^{\text{th}}$ value, which is $ 80 000 M1

IQR = $80\,000 - 25\,000 = 55\,000$ A1

d Analyst 8 would use the mean average R1

in order to suggest that their salary of $ 25 000 was significantly below the 'average' of $ 43 600. R1

e The managing director would use the median (or mode) R1

and say that Analyst 8 was already earning the 'average' salary of $ 25 000. R1

f Either: In this case, the median would be fairest R1

as there is no one earning a salary close to the mean value of $ 43 600, and the majority of workers earn the median salary. R1

18 a A random sample is a subset of a population where each member of the subset has an equal probability of being chosen R1

A stratified sample is where the population is divided into strata, based on shared characteristics. R1
Random samples of each strata are chosen in the same proportion as the strata found in the population. R1
A systematic sample is where members are chosen from a random starting point and a fixed periodic interval. R1
The interval is the population size divided by the sample size. R1

b **i** Population size is unknowable, so specific periodic choice of interval is difficult to determine. R1R1

ii Random sampling. R1
It is difficult to split rats into 'strata' as they all tend to look (and behave) the same. R1

19 a

Height, (x cm)	Frequency
$20 \le x < 25$	3
$25 \le x < 30$	3
$30 \le x < 35$	4
$35 \le x < 40$	7
$40 \le x < 45$	4
$45 \le x < 50$	2
$50 \le x < 55$	1

M1A1A1

b Mean height $= 35.8$ cm M1A1

c Variance $= 63.9$ cm^2 M1A1

d SD $= 7.99$ cm M1A1

e On average, the neighbour's garden's flowers had a slightly lower height compared to Eve's. R1
However, their standard deviation was smaller, indicating they were grown to a more consistent length. R1R1

20 a $\frac{611}{1200} \times 718824 = 366001$ M1A1

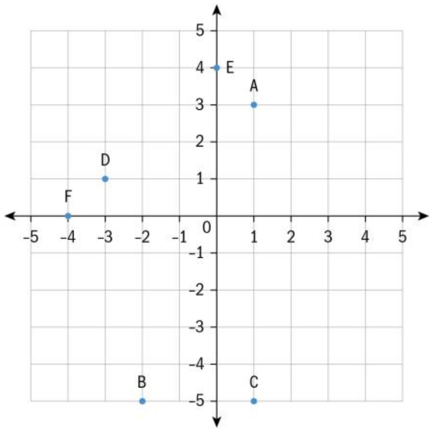
b Geographical location of residents; leisure / work travellers; age (anything appropriate or connected to travel) R1R1

4 Dividing up sapce: coordinate geometry, lines, Voronoi diagrams

Skills check

1 **a** $x=\frac{7}{2}$ **b** $x=\frac{5}{3}$

2



3 **a** 8 **b** $-\frac{1}{2}$

4 8cm

Exercise 4A

1 **a** $M\left(\frac{2+\sqrt{2}}{2},\frac{\frac{1}{3}+3.5}{2}\right)=M(1.71,1.92)$ (3 s.f.)

b $M\left(\frac{-1+7}{2},\frac{5+8.5}{2},\frac{-4.2-11}{2}\right)=M(3,6.75,-7.6)$

2 Midpoint of $A(-2,1)$ and $B(2,-4)$ is $M\left(\frac{-2+2}{2},\frac{1-4}{2}\right)=M\left(0,-\frac{3}{2}\right)$

3 Midpoint of $A(1,4)$ and $B(-3,-4)$ is $M_1\left(\frac{-3+1}{2},\frac{4-4}{2}\right)=M_1(-1,0)$. C is the midpoint of A and M_1,

i.e. $C\left(\frac{1-1}{2},\frac{4+0}{2}\right)=C(0,2)$

4 **a** $B(6,7,0)$ **b** $A(6,7,4)$

c $M\left(\frac{6+0}{2},\frac{7+0}{2},\frac{4+0}{2}\right)=M\left(3,\frac{7}{2},2\right)$

Exercise 4B

1 **a** $AB=\sqrt{(4+1.5)^2+(-3-5)^2}=\sqrt{5.5^2+8^2}=9.71$ (3 s.f.)

b $AB=\sqrt{(1+4)^2+(-2-0)^2+(10-3)^2}=\sqrt{5^2+2^2+7^2}=8.83$ (3 s.f.)

2 By Pythagoras' theorem, $PQ=2=\sqrt{(x-3)^2+(7-5)^2}=\sqrt{(x-3)^2+4}$. By squaring both sides of this expression, we find $(x-3)^2+4=4$, so $(x-3)^2=0$, i.e. $x=3$.

3 We calculate the lengths (and leave as roots for now): $AB=\sqrt{(-4-6)^2+(6-2)^2}=\sqrt{116}$.

$BC=\sqrt{(6+8)^2+(2+4)^2}=\sqrt{232}$, $AC=\sqrt{(-4+8)^2+(6+4)^2}=\sqrt{116}$. Therefore, as

$AB^2+AC^2=BC^2$, this triangle is a right angles triangle.

4 The diameter of the circle is $AB=\sqrt{(1+2)^2+(7-1)^2}=\sqrt{9+36}=\sqrt{45}$. The circumference is therefore $C=\pi\times AB=21.1(3\text{ s.f.})$

5 x = 4 + (4 − 1) = 7 and y = $\frac{6+3}{2}=4.5$

6 **a** Distance between the aircraft is

$\sqrt{(20-26)^2+(25-31)^2+(11-12)^2}=\sqrt{6^2+6^2+1^2}=\sqrt{73}=8.54$ km (3 s.f.)

b The radar will be able to detect both aircraft if they are both a distance less than 40km from the station at $O=(0,0,0)$. Using Pythagoras' theorem, we calculate the aircraft to be distances

$d_1=\sqrt{20^2+25^2+11^2}=33.9$, and $d_2=\sqrt{26^2+31^2+12^2}=42.2$. Therefore, the radar will be able to detect one, but not both, of the aircraft.

Exercise 4C

1 **a** A has the form $A(x,-8)$, where x is any number.

b A has the form $A(-3,y)$, where y is any number.

2 **a** $M=\frac{3-4}{2-(-1)}=\frac{-1}{3}$ **b** $M=\frac{2-7}{-11-2}=\frac{5}{13}$ **c** $M=\frac{2.2-(-0.3)}{-1.3-1\frac{1}{5}}=\frac{2.5}{-2.5}=-1$

3 Stair 1 gradient = $\frac{12}{25}=0.48$, stair 2 gradient = $\frac{10}{20}=0.5$, stair 3 gradient = $\frac{15}{18}=0.833$

a Stair 3 has the greatest gradient **b** Stair 1 has the least gradient.

4 See solutions

5 **a** $M=\frac{-1-2}{-1-3}=\frac{-3}{-4}=0.75$ **b** $M=\frac{2.5-5}{0-\left(-\frac{1}{3}\right)}=\frac{-2.5}{\frac{1}{3}}=-7.5$

6 The slope with gradient −3 is steeper, as the skier changes height more (3 units vs $\frac{1}{3}$ units) over a single unit distance travelled horizontally, compared to the slope with gradient $-\frac{1}{3}$.

7 Using the gradient formula, $M=0.35=\frac{h-10}{490-90}=\frac{h-10}{400}$. Re-arranging, we find that $h=10+400\times0.35=150$ m.

Exercise 4D

1 The point A has co-ordinates $A(30,2.5)$, so the gradient of the ramp is $M=\frac{2.5-0}{30-0}=0.0833$ (3 s.f.). Therefore, the ramp does not conform to safety regulations.

2 **a,b** See solutions

c Students should make a decision on the basis of the number of days they plan to work. If they plan to work fewer days than 6 then they should take the city guide job. If they plan to work longer than 6 days, the flower shop will pay more and they should take that job.

3 **a** The roof run is half the width, i.e 3.5 m.

b The roof gradient is $M = \frac{\text{rise}}{\text{run}} = \frac{1.6}{3.5} = 0.457$ (3 s.f.)

c The roof does not satisfy the requirements, as the gradient is steeper (larger) than the maximum specified by the regulations (0.457>0.17).

4 a The road has gradient $M = \frac{5}{20} = 0.25$, so "25%" would be written on the sign.

b The road with a sign indicating 15% is the steeper road.

c A 15% sign corresponds to a gradient of $M = 0.15$. In the road has horizontal change (run) of 5km, then the rise is $M \times 5 = 0.75$ km.

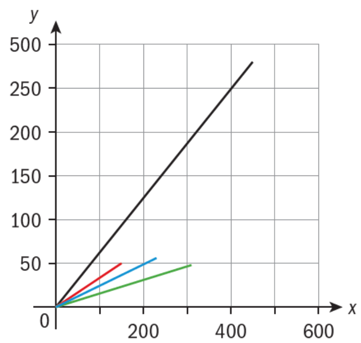
5 a i $M = \frac{280}{450} = 0.622$ (3 s.f.), so P_1 is classified as black (62.2% incline).

ii $M = \frac{50}{150} = 0.333$ (3 s.f.), so P_2 is classified as red (33.3% incline).

iii $M = \frac{48-0}{310-0} = 0.155$ (3 s.f.), so P_3 is classified as green (15.5% incline).

iv $M = \frac{56-0}{230-0} = 0.244$ (3 s.f.), so P_4 is classified as blue (24.4% incline).

b



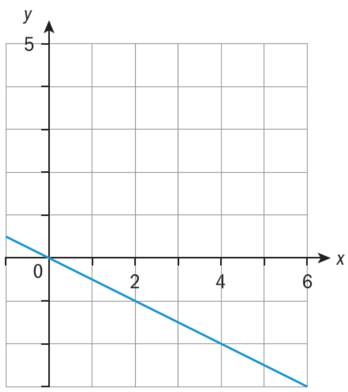
c Use the trigonometric ratio: $\tan(\alpha) = \frac{\text{opposite}}{\text{adjacent}} = \frac{50}{150} = \frac{1}{3}$. So $\alpha = \tan^{-1}\left(\frac{1}{3}\right) = 18.4°$ (3 s.f.).

Exercise 4E

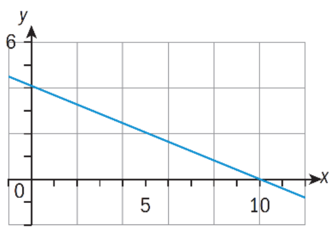
1 a $y - 4 = 3(x-1)$ **b** $y + 4 = -5(x-7)$ **c** $y - 3 = -\frac{1}{2}(x+1)$

d This line has gradient $M = \frac{4-8}{-1-1} = 2$. Therefore, the line has equation $y - 4 = 2(x+1)$.

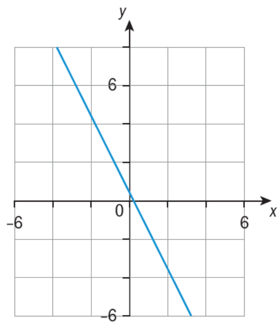
2 a



b



c



3 **a** $4 = \frac{y-5}{x+2} \Rightarrow y = 4x + 13$

b $\frac{2}{7} = \frac{y+2}{x-7} \Rightarrow y = \frac{2x}{7} - 4$

c The gradient is $M = \frac{1-0}{-3-2} = \frac{-1}{5}$, so the line has equation $\frac{-1}{5} = \frac{y}{x-2} \Rightarrow y = -\frac{1}{5}x + \frac{2}{5}$

d The gradient is $M = \frac{-2-0}{-1-2} = \frac{2}{3}$, so the line has equation $\frac{2}{3} = \frac{y}{x+2} \Rightarrow y = \frac{2}{3}x - \frac{4}{3}$

4 $A(-4,-5)$ has $x = -4, y = -5$. Therefore, as $-x - 9 = 4 - 9 = -5 = y, A$ does sit on the line $y = -x - 9$. However, $5x + 4y = 5\times-4 + 4\times-5 \neq 9$, so A does not sit on the line $5x + 4y = 9$.

5 **a** $y = \frac{1}{2}\times 6 + 4 = 3 + 4 = 7$

b From the equation of the line, $y = 9 = \frac{1}{2}x + 4$. Hence, $x = 2\times(9-4) = 10$

Exercise 4F

1 **a** $y = 3x - 2$

b This line has equation $M = -4 = \frac{y-5}{x+1}$, which is re-arranged to gradient-intercept form $y = -4x + 1$.

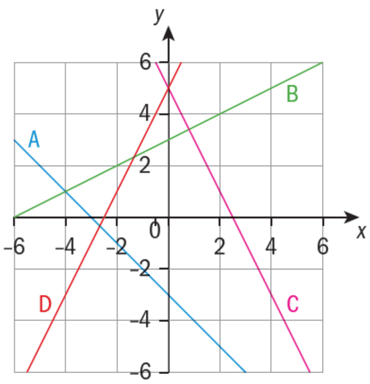
c The gradient of the line is $M = \frac{-3+1}{2+1} = \frac{-2}{3}$. Therefore, the equation of the line is $\frac{-2}{3} = \frac{y+1}{x+1}$, which is re-arranged to gradient intercept form: $y = \frac{-2}{3}x - \frac{5}{3}$.

2 **a** x-intercept $= (0,0)$, y-intercept $= (0,0)$.

b y-intercept $= (0,3)$. Re-arranging the equation of the line gives $x = -\frac{5}{2}(y-3)$, so the x-intercept is $(\frac{15}{2}, 0)$.

c When $x = 0$, then $y = 1$; the y-intercept is $(0,1)$. When $y = 0$, then $-0.2x = 1$, so the x-intercept is $(-5,0)$.

3 **a**

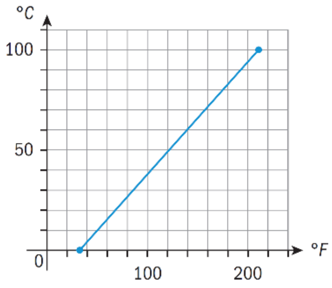


b **i** Setting $x = 0$in the line equation gives $y - 1 = -1\times 4 = -4$, so the y-intercept is $(0,-3)$. Setting $y = 0$ gives $-1 = -1(x+4)$, so the x-intercept is $(-3,0)$.

ii We write (i) as $y = -1(x+4) + 1$ and equate with (ii): $0.5x + 3 = -1(x+4) + 1 = -x - 3$. This re-arranges to give $1.5x = -6$, with solution $x = -4$. The corresponding y co-ordinate is $y = 0.5\times-4 + 3 = 1$; the intersection point is $(-4,1)$.

c The y-intercept of (iii) is 5 (by solving $y-5=0$). The line (iv) has equation $y-3=2(x+1)$, which has y-intercept at $y=3+2=5$, i.e. the y-intercept is $(0,5)$ therefore they both have the same intercept.

4 **a**



b The gradient of the line between the points is $M=\frac{212-32}{100-0}=\frac{9}{5}$. Therefore, the line has equation $\frac{9}{5}=\frac{F-32}{C}$, which may be re-arranged to give $C=\frac{5}{9}F-\frac{160}{9}$.

c The gradient of the equation in **b** is $\frac{5}{9}$

d The gradient is the amount that the temperature changes, measured in C when the temperature measured in F changes by 1.

e The y-intercept is at $F=\frac{-160}{9}$

f The temperature in °C at 0°F.

g Using the expression in 4ii) with $F=83$, we find $C=\frac{5}{9}\times 83-\frac{160}{9}=\frac{85}{3}=28.3°\text{C}$ (3 s.f.)

h Using the expression in 4ii) with $C=-10$, we find -10 $=\frac{5}{9}F-\frac{160}{9}$. This re-arranges to give

$$F=\frac{9}{5}\times\left(-10+\frac{160}{9}\right)=14.$$

i Adding $\frac{160}{9}$ to both sides of the expression in 4ii) gives $C+\frac{160}{9}=\frac{5}{9}F$. Then multiplying both sides by $\frac{9}{5}$ gives the expression $F=\frac{9}{5}C+32$.

5 We see that L_1 has a gradient $M=\frac{\text{rise}}{\text{run}}=\frac{1}{2}$. This line goes through $O(0,0)$ and therefore has equation $y=\frac{1}{2}x$. L_2: $x=-1$ L_3: $y=5$.

L_4 This line passes through

$A(0,4)$ and $B(4,0)$, so has gradient $M=\frac{4}{-4}=-1$. The equation of the line is

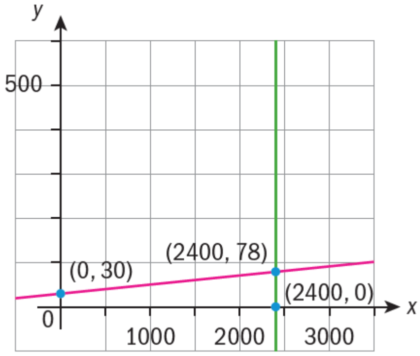
$y-0=-1(x-4)$, i.e. $y=-x+4$.

6 **a** $A=30+0.02x$

b The gradient is 0.02, which represents the amount Maria ears (in USD) per dollar spent in the restaurant.

c The y-intercept is 30. It represents the amount Maria would earn in a day if no food was sold.

d



e Using the formula in 6i) with $x = 2400,$ we find $A = 30 + 0.02\times 2400 = 78$ USD.

Exercise 4G

1 a -1

b Gradient is undefined (line is parallel to the y axis)

c This line can be re-arranged to $y = \frac{1}{3}x - \frac{5}{3}$. The gradient of any line parallel to it is therefore $\frac{1}{3}$.

d $\frac{-2}{5}$

2 a The line has a gradient of -3, so any line perpendicular to it has gradient $M = \frac{-1}{-3} = \frac{1}{3}$.

b By re-arranging as $y = \frac{x}{8}$, the gradient is seen to be $\frac{1}{8}$. Therefore, any line perpendicular to it has gradient

$M = \frac{-1}{\left(\frac{1}{8}\right)} = -8.$

c The line y = −3 has a gradient of 0, so any line perpendicular to it has an undefined gradient.

d The line has gradient $M = \frac{2}{3}$, so any line perpendicular to it has gradient $M_{\perp} = \frac{-1}{M} = \frac{-3}{2}$.

3 a L_1 has gradient $M_1 = \frac{-4+7}{0+1} = 3, L_2$ has gradient $M_2 = \frac{2-0}{-3-3} = \frac{-1}{3}$. Since $M_1 \times M_2 = -1, L_1$ and L_2 are perpendicular.

b The equation of L_1 can be re-arranged to $y = \frac{1}{4}x - \frac{3}{2}$: L_1 has gradient $M_1 = \frac{1}{4}$. As L_2 has gradient $0.25 = \frac{1}{4}$, these lines are parallel.

c L_1 has gradient $M_1 = \frac{-2}{5}$. L_2 can be re-arranged to $y = 3x - 12,$ which has gradient $M_2 = 3.$ Therefore, these lines are neither parallel nor-perpendicular.

4 The line $x = -5$ has an undefined gradient and the line $x - \frac{y}{2} = -3$ has gradient $M = 2.$ These two lines are not parallel, so must have a point of intersection. Putting $x = -5$ in $x - \frac{y}{2} = -3$ gives $y = -4$: the intersection point is $(-5, -4)$.

5 Using the gradient formula $M = \frac{1}{3} = \frac{-5-3}{s-2} = \frac{-2}{s-2}$. The solution of this equation is $s = -4$.

6 The line of the new street has gradient $M = \frac{-1}{\left(\frac{2}{7}\right)} = \frac{-7}{2}$. The new street has a point $B(-1,-0.2)$,

so the line has equation $\frac{-7}{2} = \frac{y+0.2}{x+1}$ (or $y = \frac{-7}{2}x - \frac{37}{10}$ in gradient-intercept form)

7 The gradient of the sides are $M_{AB} = \frac{-1-0}{-3-2} = \frac{1}{5}, M_{BC} = \frac{3-0}{5-2} = 1, M_{CD} = \frac{2-3}{0-5} = \frac{1}{5}, M_{DA} = \frac{2+1}{0+3} = 1$.

Therefore, the lines between AB an CD are parallel and those between BC and DA are parallel: the quadrilateral is a parallelogram.

Exercise 4H

1 a Bernard is at a point with $x = 50, y = -75$. Since $\frac{1}{2}x - 100 = 25 - 100 = -75 = y$, then Bernard is on the line with equation $y = \frac{1}{2}x - 100$.

b At the point of intersection (x, y), we have $y = -x + 410 = \frac{1}{2}x - 100$. The solution to the second equality is $x = 340$, from which we use the first equality to find $y = -340 + 410 = 70$.

2 a The line joining

$A(2,2)$ and $B(4,6)$ has gradient $M = \frac{6-2}{4-2} = 2$ and midpoint $m\left(\frac{2+4}{2}, \frac{2+6}{2}\right) = m(3,4)$. The perpendicular bisector therefore has equation $y - 4 = \frac{-1}{2}(x-3)$ or (equivalently)

$y = \frac{-1}{2}x + \frac{11}{2}$.

b The line through

C and D has gradient -1, so the equation of the perpendicular bisector through $(2,1)$ is

$y - 1 = y - 1 = \frac{-1}{-1}(x-2) = x - 1$

3 The point gradient form of the line is $y = \frac{-3}{5}x - \frac{8}{5}$; the gradient is $\frac{-3}{5}$. Therefore, the line through $(2,4)$ intersecting $3x + 5y + 8 = 0$ at a right angle has equation $y - 4 = \frac{5}{3}(x-2)$. Lines intersect at (−1, −1) and distance to hotel is $\sqrt{(2--1)^2 + (4--1)^2} = \sqrt{34} \approx 5.83$.

4 The gradient of the line between A and B is $m = \frac{7}{6}$, so the perpendicular bisector has gradient $m_\perp = -\frac{6}{7}$. The midpoint of A and B is $M = \left(5, \frac{3}{2}\right)$. Therefore, the perpendicular bisector of A and B has equation $y - \frac{3}{2} = -\frac{6}{7}(x-5)$. Since every point on this line is equidistant to A and B, then we find the location of the school finding the intersection of this line with the line on which the school sits: $7y = x - 4$. By solving these equations simultaneously, one finds the point $(6.36, 0.337)$.

5 Every point on the perpendicular bisector of AB is equidistant to points A and B, and every point on the perpendicular bisector of AC is equidistant to points A and C. Therefore, the intersection of these bisectors is equidistant to points A, B and C.

The equation of the perpendicular bisector of line through AB has gradient

$M_{AB\perp} = -\frac{1}{M_{AB}} = -1$, (as $M_{AB} = \frac{4+1}{4+1} = 1$). The midpoint of AB is $m_{AB}\left(\frac{3}{2}, \frac{3}{2}\right)$. Therefore, the perpendicular bisector of AB has equation $y - \frac{3}{2} = -\left(x - \frac{3}{2}\right) \Rightarrow y = -x + 3$.

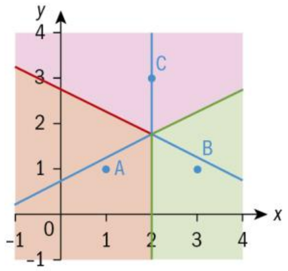
Similarly, for AC: Note that A and C have the same y co-ordinate, so the perpendicular bisector has the form

$x = k = \frac{-1+7}{2} = 3.$

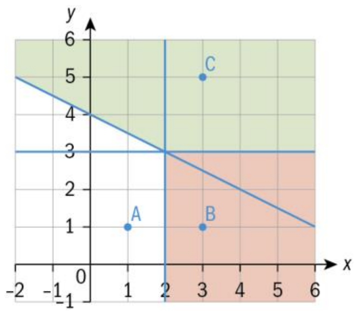
These two perpendicular bisectors intersect at $D(3,0)$, which is equidistant to A, B and C.

Exercise 4I

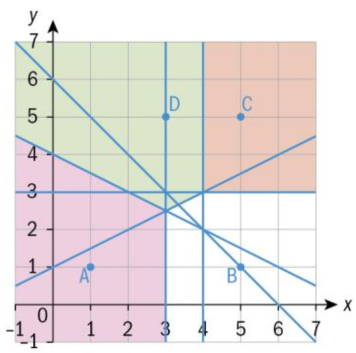
1 a



b



2 a and c



b Bookstore A

d Bookstore D

Exercise 4J

1 a The intersection of the perpendicular bisectors will be a vertex in the Voronoi diagram. It will be equally distant from A, B, and C and the centre of a circle passing through the three points. Hence the position of the solution to 'the toxic waste problem'.

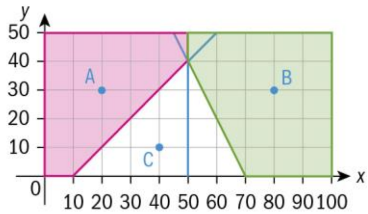
b $y = -5x + 21$ and $y = 2.5$

c (3.7, 2.5)

d 2.75km

2 a i $y = x - 10$ **ii** $y = -2x + 140$

b



c i 0.24 **ii** 0.33

d i (50,40) **ii** 31.6m

Chapter Review

1 a For example, $-2x+2y=4, y=2+x, -3x+3y-6=0$

b For example, $y+2x-19=0, 2y-6=-4(x-8), y=-2x+19$

c For example, $2x+4y+\frac{1}{3}=0, 12x+24y=-2, y=\frac{1}{2}x-\frac{1}{12}$

2 a At the intersection point (x,y), we have $y=-\frac{1}{2}x-7=-5x+10$. Re-arranging the second equality gives

$\frac{9}{2}x=17 \Rightarrow x=\frac{34}{9}$. Then $y=-\frac{1}{2}\times\frac{34}{9}-7=\frac{-17}{9}-7=-\frac{80}{7}$

b Setting $y=0$ in $y-1=-(x-4)$ re-arranges to give $x=5$: the intersection is at $(5,0)$.

c The equation of the second line is equivalent to (multiplying whole expression by -4): $x+4y=-8$. Adding this expression to the equation of the first line gives

$-x+3y+x+4y=-2-8=-10 \Rightarrow 7y=-10 \Rightarrow y=-\frac{10}{7}$. The corresponding value of

x is $x=-8-4y=-8-4\times\left(-\frac{10}{7}\right)=-\frac{16}{7}$

3 Calculate gradients of the edges:

$m_{ML}=\frac{3+1}{1+2}=\frac{4}{3}, m_{LT}=\frac{0-3}{5-1}=\frac{-3}{4}, m_{TH}=\frac{-4-0}{2-5}=\frac{4}{3}, m_{HM}=\frac{-4+1}{2+2}=\frac{-3}{4}$. This quadrilateral has two sets of parallel edges which are perpendicular to one another, so it's a square or a rectangle. The lengths of the sides are $ML=\sqrt{(1+2)^2+(3+1)^2}=\sqrt{25}=5, LT=\sqrt{(0-3)^2+(5-1)^2}=5$: two perpendicular edges have equal length so this is a square.

4 a Lines with equations $y=-x+3, y=\frac{1}{8}x-\frac{21}{8}$ intersect where $-x+3=\frac{1}{8}x-\frac{21}{8} \Rightarrow x=5$. The corresponding y value is $y=-2$: these lines intersect at $P_1(5,-2)$.

Lines with equations $y=-x+3, y=\frac{5}{6}x-\frac{1}{2}$ intersect where $-x+3=\frac{5}{6}x-\frac{1}{2} \Rightarrow x=\frac{21}{11}$. The corresponding y value is $y=\frac{12}{11}$: these lines intersect at $P_2\left(\frac{21}{11},\frac{12}{11}\right)$.

Lines with equations $y=\frac{5}{6}x-\frac{1}{2}, y=\frac{1}{8}x-\frac{21}{8}$ intersect where $\frac{5}{6}x-\frac{1}{2}=\frac{1}{8}x-\frac{21}{8} \Rightarrow x=-3$. The corresponding y value is $y=-3$: these lines intersect at $P_3(-3,-3)$.

b Using Pythagoras' theorem: the length $P_1P_2 = \sqrt{\left(5-\frac{21}{11}\right)^2 + \left(-2-\frac{12}{11}\right)^2} = \sqrt{\left(\frac{34}{11}\right)^2 + \left(\frac{34}{11}\right)^2} = 4.37$ (3 s.f)

The length $P_1P_3 = \sqrt{(5+3)^2 + (-2+3)^2} = \sqrt{8^2+1^2} = 8.06$ (3 s.f)

The length $P_2P_3 = \sqrt{\left(-3-\frac{21}{11}\right)^2 + \left(-3-\frac{12}{11}\right)^2} = \sqrt{\left(\frac{54}{11}\right)^2 + \left(\frac{45}{11}\right)^2} = 6.39$ (3 s.f)

c The perimeter of the triangle is the sum of the lengths of the sides: $4.37 + 8.06 + 6.39 = 18.8$ (3 s.f.)

5 a km per hour = $\frac{\text{rise}}{\text{run}} = \frac{6-3}{8-4} = 0.75$ (using the points $(8,6)$ and $(4,3)$)

b Maria walks at $\frac{5}{1.6} = 3.125$ km per hour; Maria walks faster as she covers more distance per unit hour.

c Petya's line has gradient $\frac{3}{4}$ and passes through

$B(4,3)$, so the equation of the line is $y = \frac{3}{4}x$.

Maria's line has gradient 3.125 and also passes through B, so the equation of the line is $y = 3.125x - 9.5$.

6 a Incorrect. The line has y-intercept of -3, but gradient $M = \frac{\text{rise}}{\text{run}} = \frac{-2}{10} = -\frac{1}{5}$. The correct equation is $y = -\frac{1}{5}x - 3$.

b Incorrect. The correct equation is $y = 5$.

c The gradient intercept form of the equation is $y = \frac{4}{11}x + \frac{2}{5}$. The line in the figure has gradient $m = \frac{8+3}{4+1} = 5$, which is consistent with the given equation. The sketched line also has a y-intercept of $\frac{2}{5}$, so the equation $-20x + 55y = 22$ does describe the line in the figure.

7 a line has gradient of

$-\frac{1}{4}$, and passes through $(0,0)$, so has equation $y - 0 = -\frac{1}{4}(x-0) \Rightarrow y = -\frac{1}{4}x$.

b Line has gradient

$\frac{0.5}{3} = \frac{1}{6}$, and goes through $(-3,-1)$, so has equation $y + 1 = \frac{1}{6}(x+3) \Rightarrow y = \frac{x}{6} - \frac{1}{2}$.

8 a The line $x + 3y = 0$ has gradient

$M = -\frac{1}{3}$. Therefore, any line perpendicular has gradient $-\frac{1}{M} = 3$. The specific perpendicular line through $(-3,2)$ has equation $y - 2 = 3(x+3) \Rightarrow y = 3x + 11$

b Any line perpendicular to $y = -3x + 0.75$ has gradient $M_\perp = -\frac{1}{-3} = \frac{1}{3}$. A line with this gradient passing through $(1,1.5)$ has equation $y - 1.5 = \frac{1}{3}(x-1) \Rightarrow y = \frac{1}{3}x + \frac{7}{6}$.

9 a i The line through $A(60,20)$ and $B(220,120)$ has gradient $M_{AB} = \frac{120-20}{220-60} = \frac{100}{160} = \frac{5}{8}$. A line with this gradient passing through A has equation $y - 20 = \frac{5}{8}(x-60) \Rightarrow y = \frac{5}{8}x - \frac{35}{2}$.

ii $m_{AB}\left(\frac{60+220}{2},\frac{20+120}{2}\right)=m_{AB}(140,70)$

iii The perpendicular bisector has gradient $\frac{-1}{M_{AB}}=-\frac{8}{5}$ and passes through m_{AB}, so has equation

$y-70=-\frac{8}{5}(x-140)\Rightarrow y=-\frac{8}{5}x+294$

b i The line through $A(60,20)$ and $C(240,40)$ has gradient $M_{AC}=\frac{40-20}{240-60}=\frac{20}{180}=\frac{1}{9}$. A line with this gradient passing through A has equation $y-20=\frac{1}{9}(x-60)\Rightarrow y=\frac{1}{9}x+\frac{40}{3}$.

ii $m_{AC}\left(\frac{60+240}{2},\frac{20+40}{2}\right)=m_{AC}(150,30)$

iii The perpendicular bisector has gradient $\frac{-1}{M_{AC}}=-9$ and passes through m_{AC}, so has equation

$y-30=-9(x-150)\Rightarrow y=-9x+1380$

c i The line through $B(220,120)$ and $C(240,40)$ has gradient $M_{AC}=\frac{40-120}{240-220}=\frac{-80}{20}=-4$. A line with this gradient passing through A has equation
$y-40=-4(x-240)\Rightarrow y=-4x+1000$.

c ii $m_{AC}\left(\frac{220+240}{2},\frac{120+40}{2}\right)=m_{AC}(230,80)$

c iii The perpendicular bisector has gradient $\frac{-1}{M_{AC}}=\frac{1}{4}$ and passes through m_{AC}, so has equation

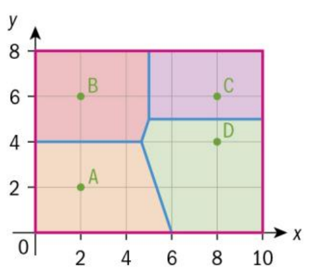
$y-80=\frac{1}{4}(x-230)\Rightarrow y=\frac{1}{4}x+\frac{45}{2}$

d T sits at the intersection of any two of the perpendicular bisectors. Solving $y=-9x+1380=\frac{1}{4}x+\frac{45}{2}$ gives $x=\frac{5430}{37}$, and $y=-9\times\frac{5430}{37}+1380=\frac{2190}{37}$

e At T, a wind turbine is equidistant to A, B and C. The length $M_{AT}=\sqrt{\left(60-\frac{5430}{37}\right)^2+\left(20-\frac{2190}{37}\right)^2}=95.2$ m (3.s.f). Therefore, a wind turbine at point T would meet the regulations.

f Area of function = $\pi\times(3\times25)^2=17671\approx17700$ m^2.

10 a



b Line between A and D has gradient $M = \frac{4-2}{8-2} = \frac{1}{3}$. The perpendicular bisector of A and D has gradient $-\frac{1}{M} = -3$ and passes through $m\left(\frac{8+2}{2}, \frac{4+2}{2}\right) = m(5,3)$. Therefore, the perpendicular bisector has equation $y - 3 = -3(x-5)$. Setting $y = 4$ in this gives $4 - 3 = -3x + 15 \Rightarrow x = \frac{14}{3}$.

c **i** $21\frac{1}{3}$ **ii** $19\frac{5}{6}$ **iii** 15 **iv** $23\frac{5}{6}$

Yes, B can support

11 Let $B = (x, y, z)$ M1

$\frac{x+4}{2} = 7 \Rightarrow x = 10$, $\frac{y-6}{2} = 3 \Rightarrow y = 12$, $\frac{z+10}{2} = -5 \Rightarrow z = -20$ A1A1A1

$B = (10, 12, -20)$ A1

12 L has gradient of 3 M1

i Neither, gradient is $\frac{1}{3}$ A1

ii Parallel, gradient is 3 A1

iii Neither, gradient is 2 A1

iv Perpendicular, gradient is $\frac{-1}{3}$ A1

v Perpendicular, gradient is $\frac{-1}{3}$ A1

13 a $M = \left(\frac{-3+5}{2}, \frac{8+3}{2}\right) = (1, \frac{11}{2})$ M1A1

b gradient $= \frac{3-8}{5+3} = \frac{-5}{8}$ M1A1

c **i** $\frac{8}{5}$ A1

ii $y = \frac{8}{5}x + c$ through $(1, \frac{11}{2}) \Rightarrow \frac{11}{2} = \frac{8}{5} + c \Rightarrow c = \frac{39}{10}$, equation is $y = \frac{8}{5}x + \frac{39}{10}$ M1A1

14 a Length of lift BP= $\sqrt{500^2 + 400^2 + 300^2} = 707\,\text{m}\,(3\text{ s.f.})$ M1A1

b Length of lift PQ= $\sqrt{(900-500)^2 + (600-400)^2 + (700-300)^2} = 600\,\text{m}$ M1A1

So total distance is $1307 = 1.31 \times 10^3\,\text{m}\,(3\text{ s.f.})$ A1

15 a $V = 500000 - 50000t$ A1

b $125000 = 500000 - 50000t \Rightarrow t = 7.5$ M1A1

Time is 3:30 p.m. A1

c $500000 - 50000t = 800000 - 10000t \Rightarrow t = 6$ M1A1

Time is 2:00 p.m. A1

16 a L_1 is $3y = ax - 9$ with gradient of $\frac{a}{3}$ A1

Require $\frac{a}{3} = \frac{-3}{2} \Rightarrow a = \frac{-9}{2}$ R1A1

b Intersection is $(-3.23, 1.85)$ (3 s.f.) A1A1

17 a $2 \times 6 \times 5 + 2 \times 6 \times 3 + 2 \times 3 \times 5 = 126$ M1A1

b $\sqrt{3^2 + 5^2 + 6^2} = \sqrt{70} = 8.37\,(3\text{ s.f.})$ M1A1

c **i** $M = (1.5, 2.5, 3)$ A1

ii Triangle AMB is isosceles. Let Q be the midpoint of AB R1

$QB = 2.5$, $BM = \frac{\sqrt{70}}{2}$ A1A1

$\sin QMB = \dfrac{2.5}{\sqrt{70}/2} \Rightarrow QMB = 36.69...$ M1A1

$AMB = 73.4°$ (3 s.f.) A1

18 a $(1200 - a)r = 765$

$(500 - a)r = 315$ A1A1

$\dfrac{1200 - a}{500 - a} = \dfrac{765}{315} \Rightarrow 378000 - 315a = 382500 - 765a \Rightarrow 450a = 4500$ M1

$a = 10$ A1

$r = \dfrac{765}{1190} = \dfrac{9}{14} = 0.643$ (3 s.f.) M1A1

b $(1200 + 500 - 10) \times \dfrac{9}{14} = 1086.43$ (2 d.p.) Euros M1A1

19 a Minimum distance is the perpendicular distance R1

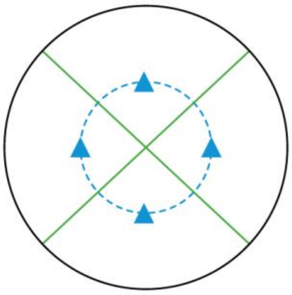
Gradient of road is 1 so line ST has gradient -1 R1

$y = -x + c$ through $(80, 140) \Rightarrow c = 220$ M1A1

Intersection of $y = -x + 220$ and $y = x - 80$ is $S = (150, 70)$. M1A1

b $ST = \sqrt{(150 - 80)^2 + (70 - 140)^2} = 99.0$ (3 s.f.) km M1A1

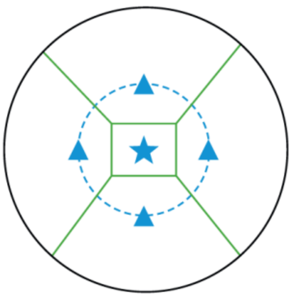
20 a



A2

b $\dfrac{\pi \times 10^2}{4} = 78.5$ m^2 (3 s.f.) M1A1

c



A4

d Square of side 5 has area of 25 m^2 M1A1

e $\dfrac{\pi \times 10^2}{4} - \dfrac{25}{4} = 72.3$ m^2 (3 s.f.) M1A1

f 4 A1