

SI



REVISION PACK 1

AVERAGES, SPREAD & DIAGRAMS

1. A group of 10 married couples and 3 single men found that the mean age of the 10 women was 41.2 years and the standard deviation of the women's ages was 15.1 years. For the 13 men, the mean age \bar{x}_m was 46.3 years and the standard deviation was 12.7 years.

(i) Find the mean age of the whole group of 23 people.

[2]

(ii) The individual women's ages are denoted by x_w and the individual men's ages by x_m . By first finding $\sum x_w^2$ and $\sum x_m^2$, find the standard deviation for the whole group.

[5]

2. Rachel measured the lengths in millimetres of some of the leaves on a tree. Her results are recorded below.

32 35 45 37 38 44 33 39 36 45

Find the mean and standard deviation of the lengths of these leaves.

[3]

3. A summary of 24 observations of x gave the following information:

$$\Sigma(x - a) = -73.2 \quad \text{and} \quad \Sigma(x - a)^2 = 2115.$$

The mean of these values of x is 8.95.

- (i) Find the value of the constant a .

[2]

- (ii) Find the standard deviation of these values of x .

[2]

4. The length of time, t minutes, taken to do the crossword in a certain newspaper was observed on 12 occasions. The results are summarised below.

$$\Sigma(t - 35) = -15 \quad \Sigma(t - 35)^2 = 82.23$$

Calculate the mean and standard deviation of these times taken to do the crossword.

[4]

5. The following table shows the results of a survey to find the average daily time, in minutes, that a group of schoolchildren spent in internet chat rooms.

Time per day (t minutes)	Frequency
$0 \leq t < 10$	2
$10 \leq t < 20$	f
$20 \leq t < 40$	11
$40 \leq t < 80$	4

The mean time was calculated to be 27.5 minutes.

- (i) Form an equation involving f and hence show that the total number of children in the survey was 26.

[4]

- (ii) Find the standard deviation of these times.

[2]

6. As part of a data collection exercise, members of a certain school year group were asked how long they spent on their Mathematics homework during one particular week. The times are given to the nearest 0.1 hour. The results are displayed in the following table.

Time spent (t hours)	$0.1 \leq t \leq 0.5$	$0.6 \leq t \leq 1.0$	$1.1 \leq t \leq 2.0$	$2.1 \leq t \leq 3.0$	$3.1 \leq t \leq 4.5$
Frequency	11	15	18	30	21

- (i) Draw, on graph paper, a histogram to illustrate this information. [5]
- (ii) Calculate an estimate of the mean time spent on their Mathematics homework by members of this year group. [3]

7. The lengths of time in minutes to swim a certain distance by the members of a class of twelve 9-year-olds and by the members of a class of eight 16-year-olds are shown below.

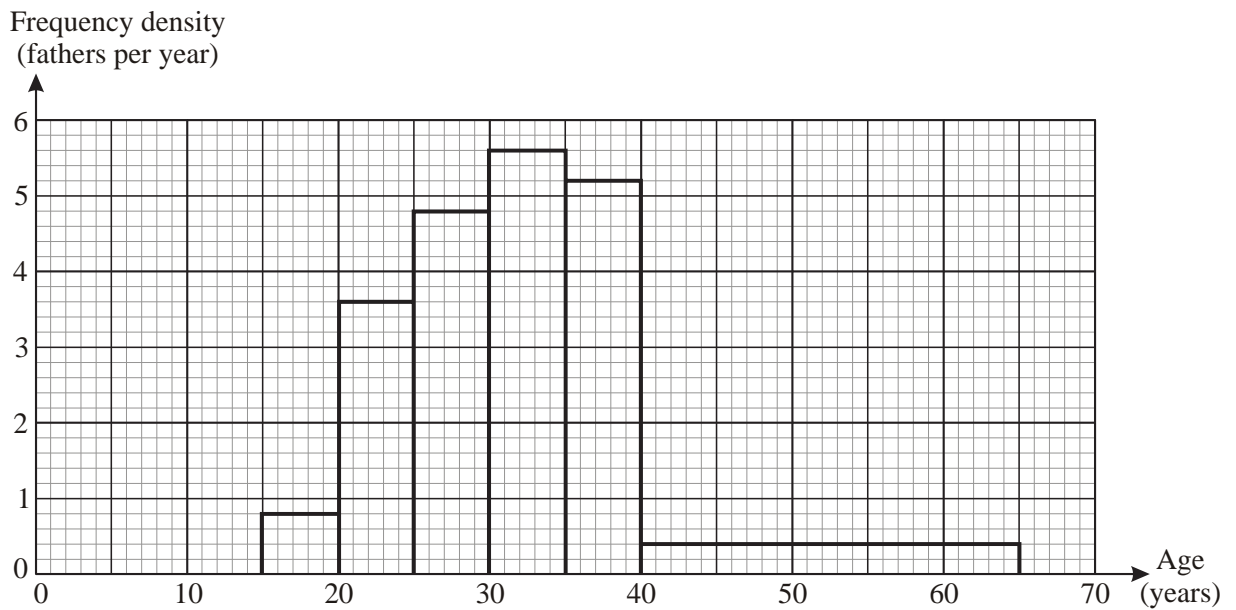
9-year-olds: 13.0 16.1 16.0 14.4 15.9 15.1 14.2 13.7 16.7
16.4 15.0 13.2

16-year-olds: 14.8 13.0 11.4 11.7 16.5 13.7 12.8 12.9

- (i) Draw a back-to-back stem-and-leaf diagram to represent the information above. [4]

- (ii) A new pupil joined the 16-year-old class and swam the distance. The mean time for the class of nine pupils was now 13.6 minutes. Find the new pupil's time to swim the distance. [3]

8. Each father in a random sample of fathers was asked how old he was when his first child was born. The following histogram represents the information.



- (i) What is the modal age group? [1]
- (ii) How many fathers were between 25 and 30 years old when their first child was born? [2]
- (iii) How many fathers were in the sample? [2]
- (iv) Find the probability that a father, chosen at random from the group, was between 25 and 30 years old when his first child was born, given that he was older than 25 years. [2]

9. The weights of 30 children in a class, to the nearest kilogram, were as follows.

50	45	61	53	55	47	52	49	46	51
60	52	54	47	57	59	42	46	51	53
56	48	50	51	44	52	49	58	55	45

Construct a grouped frequency table for these data such that there are five equal class intervals with the first class having a lower boundary of 41.5 kg and the fifth class having an upper boundary of 61.5 kg.

[4]

10. A study of the ages of car drivers in a certain country produced the results shown in the table.

Percentage of drivers in each age group

	Young	Middle-aged	Elderly
Males	40	35	25
Females	20	70	10

Illustrate these results diagrammatically.

[4]

11. The pulse rates, in beats per minute, of a random sample of 15 small animals are shown in the following table.

115	120	158	132	125
104	142	160	145	104
162	117	109	124	134

- (i) Draw a stem-and-leaf diagram to represent the data. [3]
- (ii) Find the median and the quartiles. [2]
- (iii) On graph paper, using a scale of 2 cm to represent 10 beats per minute, draw a box-and-whisker plot of the data. [3]

12. The following back-to-back stem-and-leaf diagram shows the cholesterol count for a group of 45 people who exercise daily and for another group of 63 who do not exercise. The figures in brackets show the number of people corresponding to each set of leaves.

	People who exercise		People who do not exercise	
(9)	9 8 7 6 4 3 2 2 1	3	1 5 7 7	(4)
(12)	9 8 8 8 7 6 6 5 3 3 2 2	4	2 3 4 4 5 8	(6)
(9)	8 7 7 7 6 5 3 3 1	5	1 2 2 2 3 4 4 5 6 7 8 8 9	(13)
(7)	6 6 6 6 4 3 2	6	1 2 3 3 3 4 5 5 5 7 7 8 9 9	(14)
(3)	8 4 1	7	2 4 5 5 6 6 7 8 8	(9)
(4)	9 5 5 2	8	1 3 3 4 6 7 9 9 9	(9)
(1)	4	9	1 4 5 5 8	(5)
(0)		10	3 3 6	(3)

Key: 2 | 8 | 1 represents a cholesterol count of 8.2 in the group who exercise and 8.1 in the group who do not exercise

- (i) Give one useful feature of a stem-and-leaf diagram. [1]
- (ii) Find the median and the quartiles of the cholesterol count for the group who do not exercise. [3]

You are given that the lower quartile, median and upper quartile of the cholesterol count for the group who exercise are 4.25, 5.3 and 6.6 respectively.

- (iii) On a single diagram on graph paper, draw two box-and-whisker plots to illustrate the data. [4]

13. During January the numbers of people entering a store during the first hour after opening were as follows.

Time after opening, x minutes	Frequency	Cumulative frequency
$0 < x \leq 10$	210	210
$10 < x \leq 20$	134	344
$20 < x \leq 30$	78	422
$30 < x \leq 40$	72	a
$40 < x \leq 60$	b	540

- (i) Find the values of a and b . [2]
- (ii) Draw a cumulative frequency graph to represent this information. Take a scale of 2 cm for 10 minutes on the horizontal axis and 2 cm for 50 people on the vertical axis. [4]
- (iii) Use your graph to estimate the median time after opening that people entered the store. [2]
- (iv) Calculate estimates of the mean, m minutes, and standard deviation, s minutes, of the time after opening that people entered the store. [4]
- (v) Use your graph to estimate the number of people entering the store between $\left(m - \frac{1}{2}s\right)$ and $\left(m + \frac{1}{2}s\right)$ minutes after opening. [2]

14. The arrival times of 204 trains were noted and the number of minutes, t , that each train was late was recorded. The results are summarised in the table.

Number of minutes late (t)	$-2 \leq t < 0$	$0 \leq t < 2$	$2 \leq t < 4$	$4 \leq t < 6$	$6 \leq t < 10$
Number of trains	43	51	69	22	19

- (i) Explain what $-2 \leq t < 0$ means about the arrival times of trains. [1]
- (ii) Draw a cumulative frequency graph, and from it estimate the median and the interquartile range of the number of minutes late of these trains. [7]

15. In a survey, people were asked how long they took to travel to and from work, on average. The median time was 3 hours 36 minutes, the upper quartile was 4 hours 42 minutes and the interquartile range was 3 hours 48 minutes. The longest time taken was 5 hours 12 minutes and the shortest time was 30 minutes.

(i) Find the lower quartile.

[2]

(ii) Represent the information by a box-and-whisker plot, using a scale of 2 cm to represent 60 minutes.

[4]

16. The stem-and-leaf diagram below represents data collected for the number of hits on an internet site on each day in March 2007. There is one missing value, denoted by x .

0		0 1 5 6	(4)
1		1 3 5 6 6 8	(6)
2		1 1 2 3 4 4 4 8 9	(9)
3		1 2 2 2 x 8 9	(7)
4		2 5 6 7 9	(5)

Key: 1 | 5 represents 15 hits

- (i) Find the median and lower quartile for the number of hits each day

[2]

- (ii) The interquartile range is 19. Find the value of x .

[2]

17. The salaries, in thousands of dollars, of 11 people, chosen at random in a certain office, were found to be:

40, 42, 45, 41, 352, 40, 50, 48, 51, 49, 47.

Choose and calculate an appropriate measure of central tendency (mean, mode or median) to summarise these salaries. Explain briefly why the other measures are not suitable.

[3]

SOLUTIONS



Correct answer obtained using Σx and Σa

(ii) standard deviation = $\sqrt{\frac{2115}{24} - (-3.05)^2}$ M1

For $\frac{2115}{24} - (\pm \text{ their coded mean})^2$

= 8.88 A1
Correct answer

OR

sd = $\sqrt{\frac{3814.2}{24} - 8.95^2}$ M1

For $\frac{\text{their } \Sigma x^2}{24} - 8.95^2$ where Σx^2 is obtained

from expanding $\Sigma(x - a)^2$ with $2a\Sigma x$ seen

= 8.88 A12
Correct answer

[4]

4. mean = $35 - 15/12$ M1
For $-15/12$ seen

= 33.75 (33.8) minutes A1
Correct answer

sd = $\sqrt{82.23/12 - (-15/12)^2}$ M1

$82.23/12 - (\pm \text{ their coded mean})^2$

= 2.3 minutes A1
Correct answer

[4]

5. (i) $5 \times 2 + 15f + 30 \times 11 + 60 \times 4 = 27.5(17 + f)$ M1
For attempt at LHS, accept end points or cl width

For attempt at RHS, must have $17 + f$ M1

$f = 9$ A1
For correct f

total = 26 AG A14
For correct answer given, ft if previous answer rounds to 9

(ii) $\sigma = 16.1$ M1
For method including sq rt and mean squared (can be implied if using calculator,

must be x^2f on mid-points) or $\sum \frac{f(x-\bar{x})^2}{26}$

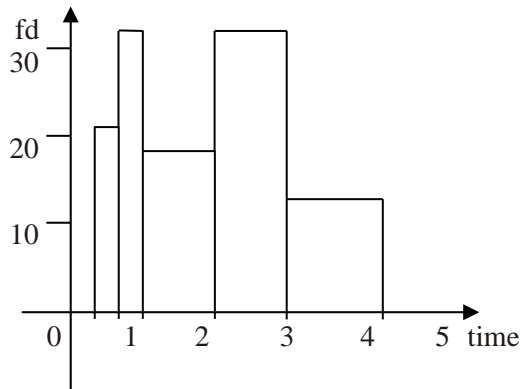
A12

For correct answer

[6]

6. (i) Fd: 22, 30, 18, 30, 14
Attempt at freq density or scaling

M1



Bar lines correctly located at 0.55, 1.05, 2.05, 3.05, no gaps

B1

correct widths of bars

B1

both axes uniform from at least 0 to 15 or 30, and 0.05 to 4.5 and labelled, (fd, or freq per half hour, time, hours, t)

B15

- (ii) mid-points 0.3, 0.8, 1.55, 2.55, 3.8
an attempt at mid-points (not class widths)

M1

= 199.5 / 95
using (Σ their fx) / their 95

M1

mean = 2.1 hours
correct answer from 199.5 in num

A13

[8]

7. (i)

16 yr olds		9 year olds
7, 4	11	
9, 8,	12	
7, 0	13	0, 2, 7,
8	14	2, 4,

	15	0, 1, 9,
5	16	0, 1, 4, 7,

3 columns including an integer stem in the middle, single digits in leaves. Can go downwards B1

One leaf column correct, ordering not necessary B1

Other leaf column correct (ordering not nec) and both leaves labelled correctly (could be in key) B1

Key $7 \mid 13 \mid 2$ means 13.7 minutes and 13.2 minutes B14
Key correct both ways or two keys one each way, must have minutes

(ii) \sum (8 pupils) = 106.8 B1
106.8 seen or implied

\sum (9 pupils) = 13.6×9 (= 122.4) B1
for 13.6×9

New pupil's time = 15.6 min B1ft3
Ft on $122.4 - \sum 8$

[7]

8. (i) 30-35 years B1 1

(ii) 4.8×5 M1
Multiplying by 5

= 24 A12
Correct answer

(iii) $4 + 18 + 24 + 28 + 26 + 10$ M1
Summing their 6 attempts at frequencies

= 110 A12
Correct answer

(iv) $24 / 88$ M1
Dividing their (ii) by their attempt at > 25 group

= 0.273 A1 ft2
Correct answer, ft on above

[7]

9.

Weight	freq
41.5-45.5	4
45.5-49.5	7
49.5-53.5	10
53.5-57.5	5
57.5-61.5	4

Five groups

M1

Correct boundaries, accept 42-45, 46-49 etc

A1

Attempt to calculate frequencies Σ 29, 30 or 31.

M1

5 frequencies correct

A1

[4]

10. two pie charts or 2 bars (m and f)
3 lots of 2 or 2 lots of 3, bars, lines or sectors

M1

3 different age categories in each
group
one category touching, not superimposed, one
category not touching, bars equal width

A1

correct height or angle
accept pie chart visually correct

B1

labels m and f, percentage, drivers, y,m
elderly

B1

[4]

11. (i)

10	4 4 9
11	5 7
12	0 4 5
13	2 4
14	2 5
15	8
16	0 8

Correct stem

B1

Correct leaves, must be sorted and in columns
and give correct overall shape

B1

key 4 represents 104

B13

Key, must have vertical line in both

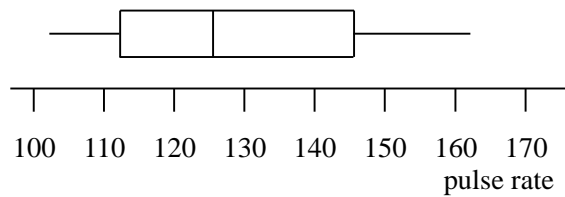
- (ii) median = 125
Any 2 correct values seen

B1

LQ = 115
UQ = 145
third correct value

B12

(iii)



B1

correct uniform scale from at least 110 to 160
with room for end points, and label or title

B1ft

correct median and quartiles on diagram fit their
values (must be box ends)

B13

correct whiskers, no line through box, touching
box in the middle not the top or bottom

[8]

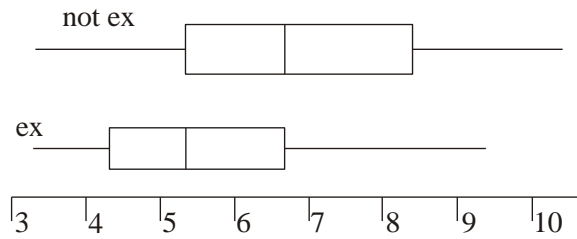
12. (i) shows all the data
Or other suitable advantage e.g. can see
the shape, mode etc.

B1 1

- (ii) Not exercise LQ = 5.4
 Median = 6.5
 UQ = 8.3
 ft on first answer missing the decimal point

B1
 B1ft
 B1ft3

(iii)



For one linear numbered scale from 3 to 9.5, or two identically positioned scales

B1

For not exercise all correct on linear scale

B1ft

For exercise correct on linear scale

B1

For two labels and cholesterol and scale labelled

B14

SR non linear scale max B0 B0 B0 B1

SR no graph paper lose one mark

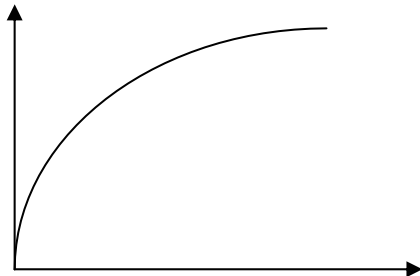
[8]

13. (i) $a = 494$
 $b = 46$

B1

B12

(ii)



B1
Correct linear scale minimum 0 to 540 and 0 to 60

Labels (cf or people or number of people) and (time, or minutes) and attempt at cf or cf step polygon

Attempt to plot points at (10, 210), (20, 344), (30, 422), (40, 494)

Correct graph through (0, 0) and (60, 540)

(iii) median is

Attempt to read from graph at line $y = 270$ or 270.5 13.5 to 14.6 min
Correct answer

(iv) $(5 \times 210 + 15 \times 134 + 25 \times 78 + 35 \times 72 + 50 \times 46) / 540$
 $= 9830 / 540$

Using mid points and frequencies
 $= 18.2$ min
Correct mean

$(5^2 \times 210 + 15^2 \times 134 + \dots) - 18.2^2$
Attempt at $\Sigma x^2 f / \Sigma f$ – their mean² numerically,
could use cfs, ucb, but not class widths

sd = 14.2 min
Correct answer

(v) $18.2 \pm 7.1 = 11.1, 25.3$
 $390 - 225$

Attempt to read their mean $\pm \frac{1}{2}$ sd from cf graph

= 155 to 170 people
Correct answer

14. (i) some trains were up to 2 minutes early
Or sensible equivalent, must use the idea ‘early’
2 not needed

(ii) cf table
NB All M marks are independent

Min Late, Less than	0	2	4	6	10
C freq	43	94	163	185	204

B1

M1

A14

M1

A12

M1

A1

M1

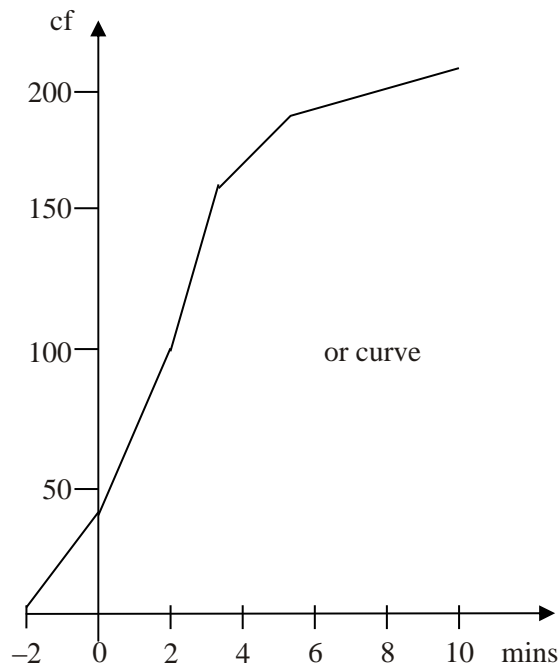
A14

M1

A12

[14]

B1 1



Attempt at C F table with upper limits no halves

M1

M1

Uniform linear scales from at least 0 to 10 and 0 to 204 and at least one axis labelled, CF or mins or t

M1

Attempt at graph their 5 points. $(-2, 0)$ not nec (could be midpoints or lower bounds not f d)

M1

Attempt at median along 102 or 102.5 line

M1

Attempt at LQ along 51/52 line and UQ along 153/154 line from graph

Median = rounding to 2.1 to 2.4 min
Correct median

B1

IQ range = rounding to 3.2 to 3.6 min
Correct IQ range allow from midpoints etc

A17

[8]

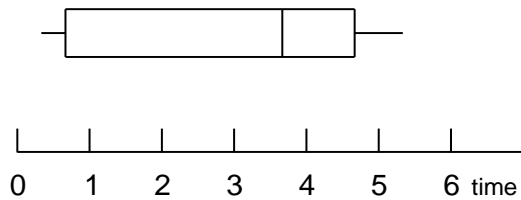
15. (i) LQ = 4 hr 42 min – 3 hr 48 min
Subtracting IQR from UQ

M1

= 54 min (0.9 hours)
Correct answer

A12

(ii)



	Correct whiskers (accept hour decimals or minutes)	B1	
	Correct median line, can be broken or extended	B1	
	Correct UQ and LQ fit on their (i), box ends	B1ft	
	correct uniform scale label hours or minutes, could be heading or key	B14	
			[6]
16.	(i) median = 16 th along = 24	B1	
	LQ = 16 not 15.5	B12	
	(ii) UQ = LQ + 19 = 35 For adding 19 to their LQ in whatever form	M1	
	$x = 5$ Must be 5 not 35. c.w.o.	A12	
			[4]
17.	median	B1	
	\$47000 Must have 47000	B1	
	data have an outlier, are skew etc Accept any equivalent reason	B1	
			[3]