Form 4: Chapter 8
Measures of Dispersion for Ungrouped Data
Fully-worked Solutions

UPSKILL 8.1

1

	Set P	*Set Q*	*Set R*
	90	93	8
	90	92	10
	90	90	412
	90	88	11
	90	87	9
Sum	450	450	450
Mean $(\bar{x})$	$\frac{450}{5} = 90$	$\frac{450}{5} = 90$	$\frac{450}{5} = 90$

(b) The means of the three sets of data are the same, i.e. 90.

(c) For the set of data *P*, each value is the same. The dispersion of data is zero.
For the set of data *Q*, the values of data gather around the value of mean. The dispersion of data is not wide.
For the set of data *R*, there is an extreme value (412) such that the value is unusually large compared to the other four values. The dispersion of data is very wide.

2 *Interpretation*
The heights of students in class *A* disperse from 143 cm to 173 cm. Many values of data gather in the range 151 cm to 159 cm.
The heights of students in class *B* disperse from 134 cm to 174 cm. There is no gathering of values of data.
Conclusion
The heights of students in class *B* are more widely spread compared to that of class *A*.

3 *Interpretation*
The numbers of goals scored by team *H* disperse from 0 goal to 5 goals. Many values of data gather at 2 goals.
The numbers of goals scored by team *K* disperse from 0 goal to 7 goals. There is no gathering of values of data.
Conclusion
The numbers of goals scored by team *K* are more widely spread compared to that team *H*.

UPSKILL 8.2

1 (a) $5\quad 7\quad \underset{\underset{Q_1}{\uparrow}}{8}\quad 10\quad 12\quad \underset{\underset{Q_2}{\uparrow}}{13}\quad 15\quad 19\quad \underset{\underset{Q_3}{\uparrow}}{21}\quad 22\quad 23$

Range = 23 – 5 = 18
Interquartile range = 21 – 8 = 13

(b) $1\quad 3\quad \underset{Q_1}{\uparrow}\quad 5\quad 7\quad \underset{Q_2}{\uparrow}\quad 8\quad 11\quad \underset{Q_3}{\uparrow}\quad 14\quad 15$

Range = 15 – 1 = 14
Interquartile range = 12.5 – 4 = 8.5

(c) $2\quad 5\quad \underset{Q_1}{\uparrow}\quad 6\quad 7\quad \underset{\underset{Q_2}{\uparrow}}{8}\quad 9\quad 10\quad \underset{Q_3}{\uparrow}\quad 13\quad 16$

Range =16 – 2 = 14
Interquartile range = 11.5 – 5.5 = 6

2 (a) Mean = $\bar{x} = \frac{\sum x}{n} = \frac{108}{5} = 21.6$

(b)

x	15	17	21	24	31	$\sum x = 108$
x^2	225	289	441	576	961	$\sum x^2 = 2\ 492$

Variance = $\frac{\sum x^2}{n} - (\bar{x})^2 = \frac{2\ 492}{5} - 21.6^2 = 31.84$

(c) Standard deviation = $\sqrt{31.84} = 5.643$

3 (a) Mean = $\bar{x} = \frac{\sum x}{n} = \frac{54}{9} = 6$

(b)

x	1	4	5	8	9	4	8	8	7	$\sum x = 54$
x^2	1	16	25	64	81	16	64	64	49	$\sum x^2 = 380$

Variance = $\frac{\sum x^2}{n} - (\bar{x})^2 = \frac{380}{9} - 6^2 = 6\frac{2}{9}$

(c) Standard deviation = $\sqrt{6\frac{2}{9}} = 2.494$

4 (a) Range = 10 – 1 = 9

(b) $1 \quad \underset{\underset{Q_1}{\uparrow}}{7} \quad 8 \quad \underset{\underset{Q_2}{\uparrow}}{8} \quad 9 \quad \underset{\underset{Q_3}{\uparrow}}{9} \quad 10$

(b) Interquartile range = 9 – 7 = 2

(c)

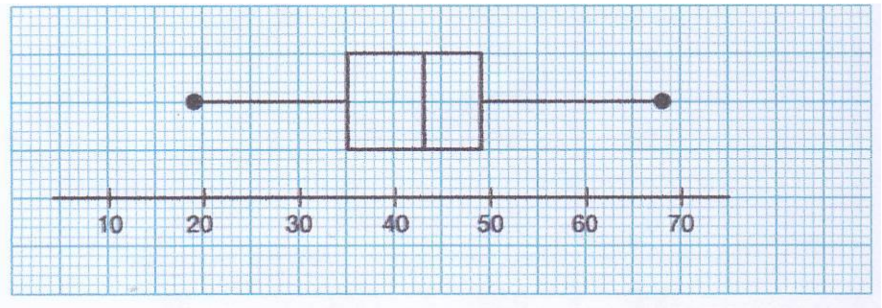
x	8	9	7	8	9	10	1	$\sum x = 52$
x^2	64	81	49	64	81	100	1	$\sum x^2 = 440$

Variance = $\dfrac{\sum x^2}{n} - (\bar{x})^2 = \dfrac{440}{7} - \left(\dfrac{52}{7}\right)^2 = 7\dfrac{33}{49}$

(c) Standard deviation = $\sqrt{7\dfrac{3}{7}} = 2.770$

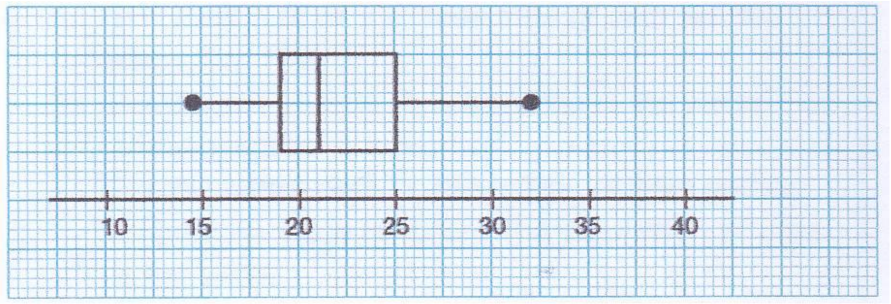
Standard deviation takes into consideration each value of data in the set of data and it is stated in a unit similar to mean. Even though variance takes into consideration each value of data in the set of data but its unit is the square of the unit of mean. The interquartile range only takes into consideration 50% of the values of data which are at the centre of the set of data after the values of data are arranged in ascending order. The range will be affected by the extreme value (i.e. 1).

5 $19 \quad 28 \quad 34 \quad \underset{Q_1}{\uparrow} \quad 36 \quad 36 \quad 38 \quad \underset{\underset{Q_2}{\uparrow}}{43} \quad 45 \quad 48 \quad 48 \quad \underset{Q_3}{\uparrow} \quad 50 \quad 65 \quad 68$



6 14 15 19 19 20 22 22 25 30 32

↑ Q_1 ↑ Q_2 ↑ Q_3



7 (a) 20 24 28 32 34 38 40 40 42 46

↑ Q_1 ↑ Q_2 ↑ Q_3

Range = 46 – 20 = 26

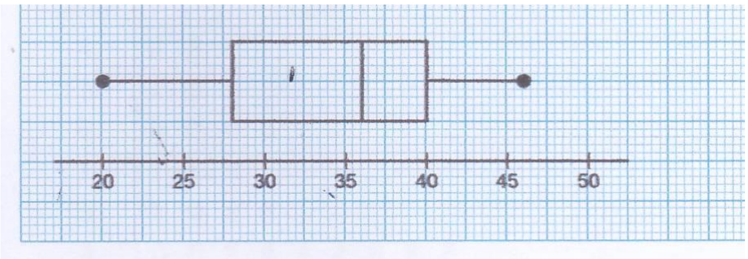
Interquartile range = 40 – 28 = 12

x	20	24	28	32	34	38	40	40	42	46	$\sum x = 344$
x^2	400	576	784	1 024	1 156	1 444	1 600	1 600	1 764	2 116	$\sum x^2 = 12\ 464$

$$\text{Standard deviation} = \sqrt{\frac{12\ 464}{10} - \left(\frac{344}{10}\right)^2} = \sqrt{63.04} = 7.940$$

Variance = 63.04

(b)



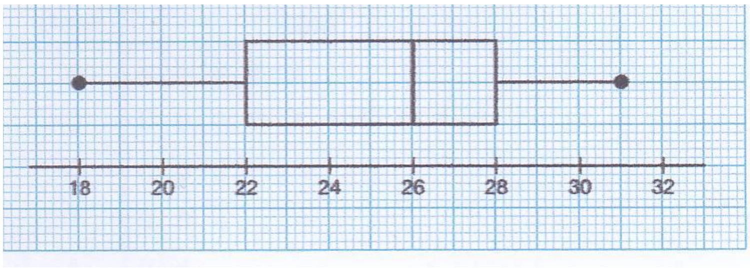
(c) (i) New range = $\frac{1}{2} \times 26 = 13$

New interquartile range = $\frac{1}{2} \times 12 = 6$

New standard deviation = $\frac{1}{2} \times 7.940 = 3.97$

New variance = $\left(\frac{1}{2}\right)^2 \times 63.04 = 15.76$

(ii) (a) 18 20 $\underset{\underset{Q_1}{\uparrow}}{22}$ 25 25 $\underset{\underset{Q_2}{\uparrow}}{}$ 27 28 $\underset{\underset{Q_3}{\uparrow}}{28}$ 29 31



(d) The range of the set of data (c) is $\frac{1}{2}$ times of the range of the set of data (a).

The interquartile range of the set of data (c) is $\frac{1}{2}$ times of the interquartile range of the set data (a).

The standard deviation of the set of data (c) is $\frac{1}{2}$ times of the standard deviation of the set of data (a).

The variance of the set of data (c) is $\left(\frac{1}{2}\right)^2$ times of the variance of the set of data (a).

The addition of 8 does not have any effect on each measure of deviation.

8 (a)

x	9	11	12	13	14	14	15	16	17	$\sum x = 121$
x^2	81	121	144	169	196	196	225	256	289	$\sum x^2 = 1\,674$

Mean = $\frac{121}{9} = 13\frac{4}{9}$

Standard deviation = $\sqrt{\frac{1\,677}{9} - \left(\frac{121}{9}\right)^2} = 2.362$

(b) (i)

x	9	11	12	13	14	14	15	16	40	$\sum x = 144$
x^2	81	121	144	169	196	196	225	256	289	$\sum x^2 = 2\,988$

Mean = $\frac{144}{9} = 16$

Standard deviation = $\sqrt{\frac{2\,988}{9} - .16^2} = 8.718$

(ii) The extreme value will cause the standard deviation to be larger.

(c) (i)

x	14	11	12	13	14	14	15	16	17	$\sum x = 126$
x^2	196	121	144	169	196	196	225	256	289	$\sum x^2 = 1\,792$

Mean = $\frac{126}{9} = 14$

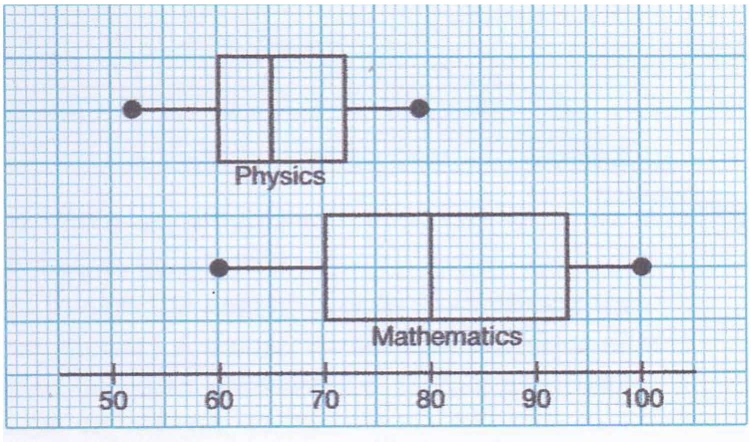
Standard deviation = $\sqrt{\frac{1\,792}{9} - 14^2} = 1.764$

(ii) When the smallest value is replaced by a value which is nearby the mean, the standard deviation will become smaller.

Conclusion

Replacing an extreme value with a value close to the mean will make the set of data to have a narrower dispersion.

9



Interpretation

(i) The median (i.e. 80) of the Mathematics marks is larger than the median (i.e. 65) of the Physics marks.

(ii) The range (i.e. 40) of the of the Mathematics marks is larger than the range (i.e. 27) of the Physics marks.

(iii) The interquartile range (i.e. 23) of the Mathematics marks is larger than the interquartile range (i.e. 12) of the Physics marks.

Conclusion

The Mathematics marks are more widely dispersed compared to the Physics marks.

10 Student *X*

x	86	84	96	71	94	79	$\sum x = 510$
x^2	7 396	7 056	9 216	5 041	8 836	6 241	$\sum x^2 = 43\ 786$

Mean = $\frac{510}{6} = 85$

Standard deviation = $\sqrt{\frac{43\ 786}{6} - 85^2} = 8.524$

Student *Y*

y	84	76	89	91	76	94	$\sum y = 510$
y^2	7 056	5 776	7 921	8 281	5 776	8 836	$\sum y^2 = 43\ 786$

Mean = $\frac{510}{6} = 85$

Standard deviation = $\sqrt{\frac{43\ 645}{6} - 85^2} = 7.024$

Although the mean of both students, *X* and *Y*, are the same but the standard deviation of student *Y* is smaller than the standard deviation of student *X*. Hence, the performance of student *Y* is more consistent and should be selected.

11 Standard deviation = 6

$$\sqrt{\frac{260}{5} - \left(\bar{x}\right)^2} = 6$$

$$52 - \left(\bar{x}\right)^2 = 36$$

$$\left(\bar{x}\right)^2 = 16$$

$$\bar{x} = 4$$

12 Standard deviation = 2.5

$$\sqrt{\frac{\sum x^2}{10} - \left(\frac{55}{10}\right)^2} = 2.5$$

$$\frac{\sum x^2}{10} - 5.5^2 = 6.25$$

$$\sum x^2 = 365$$

13 $\frac{600}{8} - \left(\frac{\sum x}{8}\right)^2 = 50$

$$\left(\frac{\sum x}{8}\right)^2 = 75 - 50$$

$$\left(\frac{\sum x}{8}\right)^2 = 25$$

$$\frac{\sum x}{8} = 5$$

$$\sum x = 40$$

Summative Practice 8

Multiple-Choice Questions

1 New standard deviation = $\frac{1}{4} \times 6.24 = 1.56$

Answer: A

2 32 32 34 34 36 36 36 38 40 42

$\uparrow$ Q_1 (34) $\uparrow$ Q_2 (between 36 and 36) $\uparrow$ Q_3 (38)

Interquartile range = 38 – 34 = 4

Answer: D

3 Variance = $\frac{43\ 294}{10} - \left(\frac{586}{10}\right)^2 = 895.44$

Answer: C

4 The standard deviation is better than variance because it has a unit which is the same as the measures of central tendency.

Answer: D

5 Interquartile range = 90 – 70 = 20

Answer: A

Structured Questions

1

	Set P	*Set Q*	*Set R*
	190	182	8
	190	202	2
	190	196	4
	190	190	921
	190	180	6
Sum	950	950	950
Mean $(\bar{x})$	190	190	190

(b) The means of all the three sets of data are the same, i.e. 190.

(c) For the set of data *P*, each value is the same. The dispersion is zero.
For the set of data *Q*, the values of data gather around the value of mean. The dispersion of data is not wide.
For the set of data *R*, there is an extreme value (921) such that its value is extra ordinary large compared to the other four values. The dispersion of data is extra ordinary wide.

2 5 Cerdas

Stem	Leaf
4	6 8
5	2 0
6	0 0 4 6
7	2 6 6
8	2 6 8 8
9	0 0 2 4 4 4 8 8 8

Key: 4 | 6 means 46 marks

5 Bestari

Stem	Leaf
4	4 6 8 8 8
5	0 0 0 2 2 4 6 6 6 8
6	6 6 8
7	0 0 4
8	0 4
9	0

Key: 4 | 6 means 46 marks

The median of the marks of the students of 5 Cerdas (76) is higher than the median of the marks of the students of 5 Bestari (56). The marks of the class 5 Cerdas disperse from 46 to 98 such that the marks gather in the range 90 to 98. The marks of the class 5 Bestari disperse from 46 to 90 such that the marks gather in the range 50 to 58.

3

x	30	32	34	36	38	40	42	
f	2	4	4	5	2	2	1	$\sum f = 20$
fx	60	128	136	180	76	80	42	$\sum fx = 702$
fx^2	1 800	4 096	4 624	6 480	2 888	3 200	1 764	$\sum fx^2 = 24\ 852$

(a) Mean = $\dfrac{\sum fx}{\sum f} = \dfrac{702}{20} = 35.1$

Variance = $\dfrac{24\ 852}{20} - (35.1)^2 = 10.59$

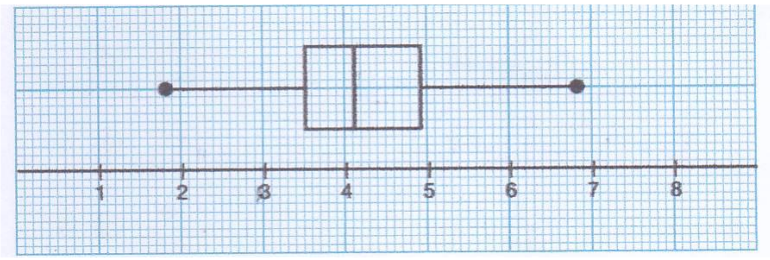
Standard deviation = $\sqrt{10.59} = 3.254$

(b) (i) Takes into consideration each value of data.
(ii) Its unit is the same as the mean.

4 1.8 2.5 2.6 2.7 3.4 ↑ 3.6 3.7 3.8 3.8 4.0 ↑ 4.2 4.3 4.4 4.6 4.8 ↑ 5.0 5.1 5.4 5.6 6.8

Q_1 Q_2 Q_3

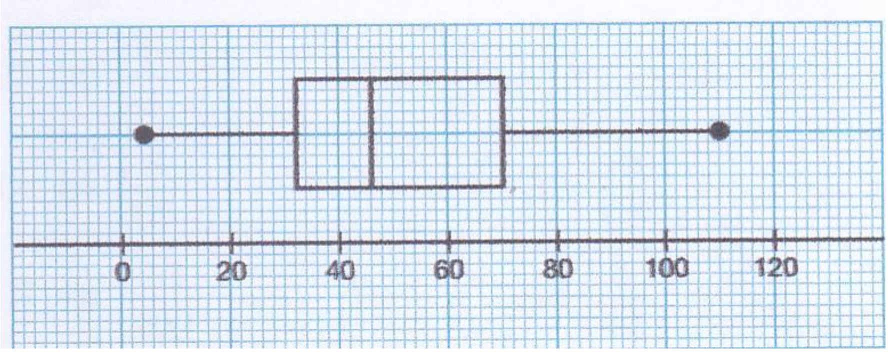
Median = 4.1, First quartile = 3.5, Third quartile = 4.9,
Minimum value = 1.8, Maximum value = 6.8



5 4 5 18 21 28 ↑ 36 38 39 43 44 ↑ 48 50 56 62 70 ↑ 70 81 99 100 110

Q_1 Q_2 Q_3

Median = 46, First quartile = 32, Third quartile = 70,
Minimum value = 4, Maximum value = 110



6 Student *A*

x	12.1	12.2	12.0	12.0	11.9	11.9	12.0	11.9	12.0	12.0	$\sum x = 120$
x^2	146.41	148.84	144	144	141.61	141.61	144	141.61	144	144	$\sum x^2 = 1\,440.08$

$$\text{Mean} = \frac{120}{10} = 12 \text{ seconds}$$

$$\text{Standard deviation} = \sqrt{\frac{1\,440.03}{10} - 12^2} = 0.0894 \text{ seconds}$$

Student *B*

y	11.6	11.5	12.1	12.3	12.4	12.3	12.0	11.9	11.9	12.0	$\sum y =$ 120
y^2	134.56	132.25	146.41	151.29	153.76	151.29	144	141.61	141.61	144	$\sum y^2 =$ 1 440.78

$$\text{Mean} = \frac{120}{10} = 12 \text{ seconds}$$

$$\text{Standard deviation} = \sqrt{\frac{1\,440.78}{10} - 12^2} = 0.2793 \text{ seconds}$$

Although the means of both students are the same, student *A* should be selected because of his smaller standard deviation (i.e. more consistent).

7

x	20	22	32	36	38	42	46	48	52	58	$\sum x$ = 394
x^2	400	484	1 024	1 296	1 444	1 764	2 116	2 304	2 704	3 364	$\sum x^2$ = 16 900

(a) $20 \quad 22 \quad \underset{\underset{Q_1}{\uparrow}}{32} \quad 36 \quad 38 \quad \underset{\underset{Q_2}{\uparrow}}{} \quad 42 \quad 46 \quad \underset{\underset{Q_3}{\uparrow}}{48} \quad 52 \quad 58$

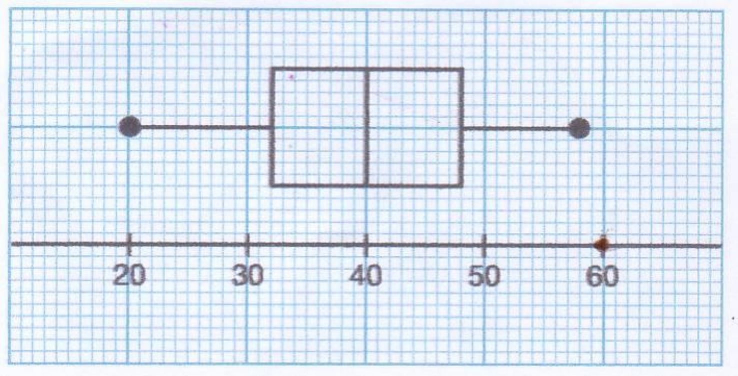
(i) Range = 58 – 20 = 38

(ii) Interquartile range = 48 – 32 = 16

(iii) Variance = $\dfrac{16\,900}{10} - \left(\dfrac{394}{10}\right)^2 = 137.64$

(iv) Standard deviation = $\sqrt{137.64} = 11.73$

(b)



(c) New interquartile range = $16 \times \frac{3}{2} = 24$

New variance = $137.64 \times \left(\frac{3}{2}\right)^2 = 309.69$

(d)

x	3	20	22	32	36	38	42	46	48	52	58	$\sum x =$ 397
x^2	9	400	484	1 024	1 296	1 444	1 764	2 116	2 304	2 704	3 364	$\sum x^2 =$ 16 909

Variance = $\frac{16\,909}{11} - \left(\frac{397}{11}\right)^2 = 15.32$

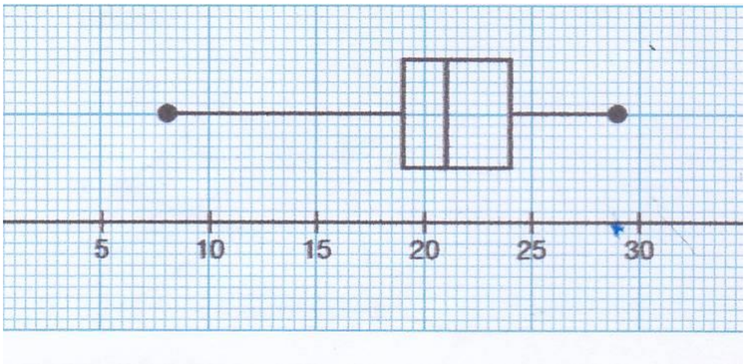
(d) The new standard deviation (i.e. 15.32) is larger than the original standard deviation (i.e. 11.73).
This is because the inclusion of 3 marks (which is an extreme value) will cause the degree of dispersion of data to be higher. Hence, the standard deviation becomes larger.

8 (a)

x	23	16	27	8	19	29	24	19	20	21	23	$\sum x =$ 229
x^2	529	256	729	64	361	841	576	361	400	441	529	$\sum x^2 =$ 5 087

8 16 19 19 20 21 23 23 24 27 29

$\uparrow Q_1$ (19) $\quad \uparrow Q_2$ (21) $\quad \uparrow Q_3$ (24)



(b) Original mean = $\frac{229}{11} = 20.82$

Original of variance = $\frac{5\ 087}{11} - \left(\frac{229}{11}\right)^2 = 29.06$

New mean = $\left(20.82 \times \frac{5}{2}\right) + 10 = 62.05$

New variance = $29.06 \times \left(\frac{5}{2}\right)^2 = 181.61$

(c)

x	23	16	27	19	29	24	19	20	21	23	$\sum x = 221$
x^2	529	256	729	361	841	576	361	400	441	529	$\sum x^2 =$ 5 023

Standard deviation = $\sqrt{\frac{5\ 023}{10} - \left(\frac{221}{10}\right)^2} = 3.727$

The value of 8 marks is an extreme value. When it is removed, the standard deviation becomes smaller.

SPM SPOT

1 10 13 18 24 26 26 33 37 42 45

$$\begin{array}{cccccccccc} 10 & 13 & 18 & 24 & 26 & 26 & 33 & 37 & 42 & 45 \\ & & \uparrow & & \uparrow & & & \uparrow & & \\ & & Q_1 & & m & & & Q_3 & & \end{array}$$

The Q_3 is incorrectly drawn. It should be 37 and not 42.
Answer: D

2 (a) (i)

$$\bar{x} = \frac{19}{3}$$

$$\frac{k + k + 4 + 3k - 4 + 2k + 3 + k - 1 + 4k}{6} = \frac{19}{3}$$

$$12k + 2 = \frac{19}{3} \times 6$$

$$12k + 2 = 38$$

$$12k = 36$$

$$k = 3$$

(ii) When $k = 3$,

3 7 5 9 2 12

Rearrange:

$$\begin{array}{cccccc} 2 & 3 & 5 & 7 & 9 & 12 \\ & \uparrow & & \uparrow & \uparrow & \\ & Q_1 & & m & Q_3 & \end{array}$$

Interquartile range = 9 – 3 = 6

(iii)

x	2	3	5	7	9	12	$\sum x = 38$
x^2	4	9	25	49	81	144	$\sum x^2 = 312$

Standard deviation = $= \sqrt{\frac{\sum x^2}{N} - \left(\bar{x}\right)^2} = \sqrt{\frac{312}{6} - \left(\frac{19}{3}\right)^2}$ = 3.448

(b) If 4*k* (which is 12) is replaced by (*k* + 4) (which is 7), the standard deviation will become smaller because an extreme value is replaced by a number near the mean.