



S08-1hour Comparing Data Sets: 1-hour Pack

EDEXCEL 1ST0 · Calculator allowed · about 60 minutes

Total Marks

Name: _____ Date: ____ / ____ / ____

Answer ALL questions. Show all your working.

- 1** Class X has test scores with a range of 18 marks. Class Y has test scores with a range of 7 marks. Write down which class is more consistent.

(Total for Question 1 is 1 mark)

- 2** Route 1 has a median journey time of 22 minutes and an interquartile range of 4 minutes. Route 2 has a median journey time of 22 minutes and an interquartile range of 11 minutes. Write down which route has the more consistent journey times.

(Total for Question 2 is 1 mark)

- 3** Two garages record the time taken, in minutes, to service a car. Garage X: median 42, IQR 10. Garage Y: median 38, IQR 16. Work out the difference between the two medians, and state which garage has more variable service times.

(Total for Question 3 is 2 marks)

- 4** Write down whether the following statement is True or False.
When comparing two data sets, a smaller range or interquartile range always means the data is more consistent.

(Total for Question 4 is 1 mark)

- 5** Café P served drinks with a median wait time of 6 minutes and a range of 9 minutes. Café Q served drinks with a median wait time of 6 minutes and a range of 3 minutes. Work out how much smaller Café Q's range is than Café P's range, and state which cafe has more consistent wait times.

(Total for Question 5 is 2 marks)

- 6** Two swim squads recorded their fastest 100m times, in seconds. Squad 1: mean 68.4, range 5.2. Squad 2: mean 71.1, range 2.6. Work out the difference in the mean times, and state which squad is faster and which is more consistent.

(Total for Question 6 is 2 marks)

- 7 Two schools recorded how long (in minutes) their Year 10 students spent on homework one evening.

School P

Time, t (minutes)	$0 \leq t < 20$	$20 \leq t < 40$	$40 \leq t < 60$	$60 \leq t < 80$	$80 \leq t < 100$
Frequency	4	10	14	8	4

School Q

Time, t (minutes)	$0 \leq t < 20$	$20 \leq t < 40$	$40 \leq t < 60$	$60 \leq t < 80$	$80 \leq t < 100$
Frequency	8	14	10	5	3

- (a) Calculate an estimate of the mean homework time for School P. (3)

- (b) Calculate an estimate of the mean homework time for School Q. (3)

- (c) Compare the two schools' typical homework times. (1)

(Total for Question 7 is 7 marks)

- 8** Two groups of volunteers each timed how many minutes it took them to solve the same puzzle. The times, in order, were:

Group P (9 people): 8, 10, 11, 12, 12, 13, 14, 15, 19

Group Q (9 people): 9, 9, 10, 10, 11, 12, 15, 17, 21

- (a) For Group P, find the lower quartile, the upper quartile and the interquartile range. (3)

-
- (b) For Group Q, find the lower quartile, the upper quartile and the interquartile range. (3)

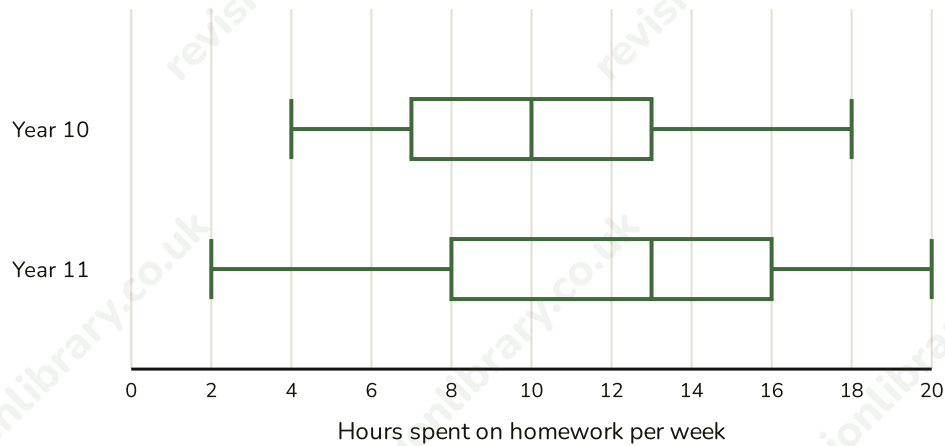
-
- (c) Compare the consistency of the two groups using the interquartile range, and write a conclusion. (2)

(Total for Question 8 is 8 marks)

-
- 9** Two farms recorded the mass, in kg, of lambs at birth. Farm A: mean 4.2, standard deviation 0.6. Farm B: mean 4.2, standard deviation 0.9. Work out the difference between the two standard deviations, and state which farm's lambs have more variable birth masses.

(Total for Question 9 is 2 marks)

- 10** The box plots show the number of hours per week that Year 10 and Year 11 students spent on homework, based on a survey at a school.



- (a) Write down the median number of hours spent on homework per week for each year group. (2)
-
- (b) Find the interquartile range for each year group. (2)
-
- (c) Compare the two year groups using the median and the interquartile range, and write a conclusion. (2)

(Total for Question 10 is 6 marks)

11 For each statement about comparing data sets, write down whether it is True or False.

(i) If two data sets have the same range, they must have the same spread for most of their values. (1)

(ii) The interquartile range ignores extreme outliers, so it can give a fairer picture of consistency than the range. (1)

(iii) You can compare the averages of two data sets by comparing their medians, even if one data set has an outlier. (1)

(iv) A larger sample size always makes a data set more consistent. (1)

(Total for Question 11 is 4 marks)

12 Two production lines record the mass, in grams, of bags of crisps. Line 1: mean 152.4, range 6.0. Line 2: mean 148.9, range 2.5. A quality target requires a mean mass of at least 150 g with low variability.

(a) Work out the difference in mean mass between the two lines.

(b) State, with a reason, which line better meets the quality target.

(Total for Question 12 is 3 marks)

13 Two hospitals record patient waiting times, in minutes, in A&E. Hospital A: median 38, lower quartile 25, upper quartile 52. Hospital B: median 44, lower quartile 30, upper quartile 46.

(a) Work out the interquartile range for each hospital.

(b) Compare the typical waiting time and the consistency of waiting times at the two hospitals.

(Total for Question 13 is 3 marks)

- 14** Two orchards record the mass, in kg, of apples picked from a random sample of trees. Orchard A: median 42, lower quartile 36, upper quartile 50. Orchard B: median 45, lower quartile 41, upper quartile 49.
- (a) Work out the interquartile range for each orchard.
- (b) State which orchard produces a more consistent yield per tree, giving a reason based on your answer to part (a).

(Total for Question 14 is 3 marks)

- 15** A teacher compares two classes' end-of-year exam marks (out of 100). Class P: mean 61, standard deviation 8. Class Q: mean 61, standard deviation 15. The pass mark is 40.
- (a) Give one similarity between the two classes' results.
- (b) Explain which class is more likely to have students scoring below the pass mark, giving a reason.

(Total for Question 15 is 3 marks)

- 16** Three sprinters' race times over a season are summarised: Runner X: mean 11.85 s, standard deviation 0.12. Runner Y: mean 11.72 s, standard deviation 0.28. Runner Z: mean 11.85 s, standard deviation 0.05. A coach wants to select the runner who is both fast and consistent for a relay.
- (a) Identify the fastest runner on average.
- (b) Identify the most consistent runner.
- (c) Explain why the coach's choice is not straightforward, referring to both runners identified in (a) and (b).

(Total for Question 16 is 3 marks)

- 17** Three sprinters' 100m race times over a season are summarised below.

Runner	Mean (s)	Median (s)	Range (s)	IQR (s)	Standard deviation (s)
Amara	12.1	12.0	1.8	0.9	0.31
Beth	12.4	12.3	0.6	0.5	0.18
Cara	12.6	12.5	2.4	1.6	0.52

- (a) Which runner has the fastest average 100m time? Justify your answer using the table. (2)
- (b) Which runner is the most consistent? Justify your answer using the table. (2)
- (c) A coach says: "Beth is the best runner because she has the smallest range." (2)
Comment on this statement, referring to the other statistics in the table.

(Total for Question 17 is 6 marks)

Mark scheme · S08-1hour Comparing Data Sets: 1-hour Pack

Question 1

- **B1** Class Y cao
- **Answer: Class Y**

Question 2

- **B1** Route 1 cao
- **Answer: Route 1**

Question 3

- **M1** correct method, 42 - 38
- **A1** 4 minutes; Garage Y is more variable (larger IQR), both cao
- **Answer: 4 minutes; Garage Y has more variable service times**

Question 4

- **B1** True cao
- **Answer: True**

Question 5

- **M1** correct method, 9 - 3
- **A1** 6 minutes smaller, and Cafe Q is more consistent, both cao
- **Answer: 6 minutes smaller; Cafe Q has more consistent wait times**

Question 6

- **M1** correct method, 71.1 - 68.4
- **A1** 2.7 seconds; Squad 1 is faster (lower mean); Squad 2 is more consistent (smaller range), all cao
- **Answer: 2.7 seconds; Squad 1 is faster; Squad 2 is more consistent**

Question 7

- (a) **M1** uses correct midpoints (10, 30, 50, 70, 90)
- (a) **M1** $Sfx / Sf = 1960 / 40$, oe
- (a) **A1** 49 (minutes) cao
- (a) **Answer: 49 minutes**
- (b) **M1** uses correct midpoints (10, 30, 50, 70, 90)
- (b) **M1** $Sfx / Sf = 1620 / 40$, oe
- (b) **A1** 40.5 (minutes) cao
- (b) **Answer: 40.5 minutes**
- (c) **B1** School P students spend longer on homework on average than School Q (49 minutes compared with 40.5 minutes), oe
- (c) **Answer: School P spends longer on homework on average (49 vs 40.5 minutes).**

Question 8

- (a) **B1** lower quartile = 10.5
- (a) **B1** upper quartile = 14.5
- (a) **B1** interquartile range = 4
- **(a) Answer: Q1 = 10.5, Q3 = 14.5, IQR = 4**
- (b) **B1** lower quartile = 9.5
- (b) **B1** upper quartile = 16
- (b) **B1** interquartile range = 6.5
- **(b) Answer: Q1 = 9.5, Q3 = 16, IQR = 6.5**
- (c) **B1** Group P has the smaller interquartile range (4 compared with 6.5)
- (c) **B1** so Group P's times were more consistent/less spread out than Group Q's, oe
- **(c) Answer: Group P was more consistent (smaller IQR of 4 vs 6.5).**

Question 9

- **M1** correct method, 0.9 - 0.6
- **A1** 0.3; Farm B is more variable, both cao
- **Answer: 0.3 kg; Farm B has more variable birth masses**

Question 10

- (a) **B1** Year 10 median = 10 hours
- (a) **B1** Year 11 median = 13 hours
- **(a) Answer: Year 10: 10 hours; Year 11: 13 hours**
- (b) **B1** Year 10 IQR = 6 (13 - 7)
- (b) **B1** Year 11 IQR = 8 (16 - 8)
- **(b) Answer: Year 10: IQR = 6 hours; Year 11: IQR = 8 hours**
- (c) **B1** Year 11 has a higher median (13 compared with 10 hours), so typically spent longer on homework
- (c) **B1** Year 11 also has a larger IQR (8 compared with 6 hours), so their homework times were more spread out/less consistent than Year 10's, oe
- **(c) Answer: Year 11 typically spent longer on homework but was less consistent than Year 10.**

Question 11

- (i) **B1** False cao
- **(i) Answer: False**
- (ii) **B1** True cao
- **(ii) Answer: True**
- (iii) **B1** True cao
- **(iii) Answer: True**
- (iv) **B1** False cao
- **(iv) Answer: False**

Question 12

- **M1** correct method, 152.4 - 148.9
- **A1** 3.5 g cao
- **B1** Line 1, with valid reason, e.g. only Line 1 has a mean above 150 g, and although its range is larger it still meets the mass target oe
- **Answer: (a) 3.5 g. (b) Line 1, because its mean (152.4 g) meets the 150 g target while Line 2's mean (148.9 g) does not.**

Question 13

- **M1** correct method for at least one IQR, e.g. 52 - 25
- **A1** Hospital A: IQR 27; Hospital B: IQR 16, both cao
- **B1** correct comparison, e.g. Hospital B has a higher median (longer typical wait) but Hospital A has a larger IQR (less consistent) oe
- **Answer: (a) Hospital A: IQR = 27. Hospital B: IQR = 16. (b) Hospital B has the longer typical wait (higher median, 44 vs 38), but Hospital A's waiting times are less consistent (larger IQR, 27 vs 16).**

Question 14

- **M1** correct method for at least one IQR, e.g. 50 - 36
- **A1** Orchard A: IQR 14; Orchard B: IQR 8, both cao
- **B1** Orchard B, with reason that its IQR (8) is smaller than Orchard A's (14), so its yields are less spread out
- **Answer: (a) Orchard A: IQR = 14. Orchard B: IQR = 8. (b) Orchard B, because its smaller IQR shows the yields per tree are more tightly grouped together.**

Question 15

- **B1** correct similarity, e.g. both classes have the same mean mark (61) oe
- **B1** correctly identifies Class Q
- **B1** valid reason, e.g. Class Q has a much larger standard deviation, so its marks are more spread out and more likely to include low scores below 40 oe
- **Answer: (a) Both classes have the same mean (61). (b) Class Q, because its larger standard deviation (15 vs 8) means marks are more spread out, making very low scores below the pass mark more likely.**

Question 16

- **B1** Runner Y (lowest mean) cao
- **B1** Runner Z (lowest standard deviation) cao
- **B1** valid explanation, e.g. Runner Y is fastest on average but Runner Z is the most consistent, so no single runner is best on both measures oe
- **Answer: (a) Runner Y. (b) Runner Z. (c) Runner Y has the fastest average time but is not the most consistent, while Runner Z is the most consistent but not the fastest, so the coach must choose which quality matters more for the relay.**

Question 17

- (a) **B1** Amara
- (a) **B1** has the lowest mean (12.1s) and lowest median (12.0s) of the three runners, oe
- (a) **Answer: Amara**
- (b) **B1** Beth
- (b) **B1** has the smallest range (0.6s), smallest IQR (0.5s) and smallest standard deviation (0.18s) - all three measures of spread agree, oe
- (b) **Answer: Beth**
- (c) **B1** the range does correctly show Beth is the most consistent runner (also confirmed by her IQR and standard deviation)
- (c) **B1** however 'best' likely means fastest overall, and Amara has the lower mean and median time, so Amara is actually faster on average - the coach's claim conflates consistency with speed and is not fully justified, oe
- (c) **Answer: Beth is the most consistent, but Amara is faster on average, so the coach's claim is only partly justified.**