



eClassroom

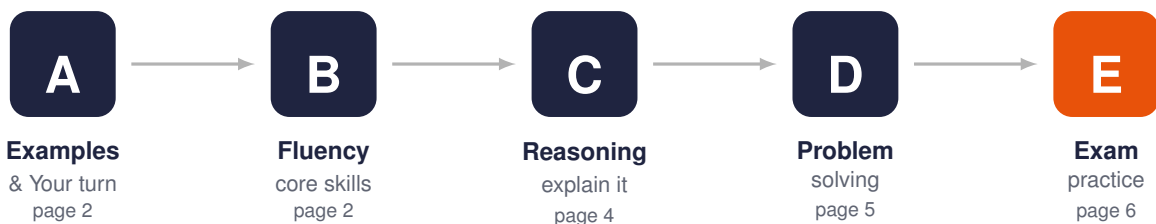
A-LEVEL MATHEMATICS • STATISTICS 1

Representations of Data

Worked Solutions

Pearson Edexcel International A Level (WST01) & A Level (9MA0) • Calculator and statistical tables allowed: show every step

Full worked solutions for the free Representations of Data worksheet, A-Level and International A-Level Maths



Tap a section to jump straight to it

Full worked solutions for every question on the Statistics 1 Representations of Data worksheet, in the same order and with the same question numbers.

How to mark your work. Tick each step you got right, not just the final answer. At A-level most marks are for method, so a wrong final answer can still earn most of the marks. If you lost marks, find the *first* line where your working is different from ours. That is the thing to fix.

In Section E: **M** = method mark, **dm** = method mark that depends on the previous M mark, **A** = accuracy mark (needs the M mark before it), **B** = independent mark.

Question sheet

Exam questions by topic

Statistics 1 exam questions

More worksheets



Join students already acing their exams.

Book your FREE lesson at www.eclassroom.org.uk

Page 1 of 7

A**Your turn answers**

Compare each line of your working with the solution.

CORE**A1** Densities: 0–4 has 3, 4–10 has 4.5. Width 0.5 cm per unit; height 2 cm per unit of density**3 cm wide, 9 cm tall****A2** $\frac{13}{4} = 3.25$: 4th; 7th; 9.75: 10th: **$Q_1 = 41, Q_2 = 50, Q_3 = 58$** **A3** $1.5 \times 12 = 18$: limits 28 and 76. 25 is below 28: **25 is an outlier****STRETCH****A4** $\frac{2}{5}(15) + 24 + \frac{5}{15}(45) = 6 + 24 + 15$: **45****A5** $Q_2 - Q_1 = 8 > Q_3 - Q_2 = 5$: **negative skew**; $\frac{43 - 76 + 30}{13} =$ **-0.231****A6** **D higher on average (31 vs 25); C more consistent (IQR 6 vs 11)****A7** 1 cm² stands for $\frac{24}{3} = 8$ items: $13.5 \times 8 =$ **108****B****Fluency solutions**

One mark for each correct part. The notes in italics explain why each sequence was chosen.

CORE**B1** (a) $\frac{36}{4}$: **9** (b) 3.2×5 : **16** (c) $\frac{45}{1.5}$: **30**(d) 0.25 cm per unit and 0.5 cm per unit of density; 14–20 has density 5: **1.5 cm wide, 2.5 cm tall***Frequency = density $\times$ width; scale the width and the height separately.*

- B2** (a) 11th: **37** (b) 5.25: 6th: **31** (c) 15.75: 16th: **43** (d) **12**

$n = 21$: $\frac{n}{4}$ and $\frac{3n}{4}$ are not whole, so round up.

- B3** (a) $12 - 12$ and $20 + 12$: **0 and 32** (b) $31 < 32$: **not an outlier**

(c) limits 50 ± 12 : 38 and 62; $64 > 62$: **yes** (d) **the mean**: it uses every value; the median depends only on position

Use the rule the question gives.

- B4** (a) **32** (b) **12** (c) $Q_3 - Q_2 = 8 > Q_2 - Q_1 = 4$: **positive skew** (d) **75%**

Each quarter of the data lies between two marks on the box plot.

- B5** (a) $\frac{2}{5}(25) + \frac{6}{10}(30)$: **28** (b) $\frac{8}{20}(15)$: **6** (c) **2, 5, 3, 0.75**

A part of a class holds the same fraction of its frequency.

- B6** (a) $8 > 4$: **positive skew** (b) $\frac{34 - 52 + 22}{12}$: **0.333** (c) $\frac{31 - 25}{8}$: **0.75** (d) $\frac{3(31 - 26)}{8}$: **1.875**

All three measures are positive: they agree on the direction.

STRETCH

- B7** Area of 0–10 is 12 for 18, so 1.5 per unit of area: frequencies 18, 27, 30, 27, 12

- (a) **27** (b) **114** (c) $\frac{3}{5}(27) + 30 + \frac{4}{10}(27)$: **57** (d) 57th: $15 + \frac{12}{30} \times 5$: **17**

One known frequency fixes the scale for every bar.

- B8** (a) **$Q_1 = 17, Q_2 = 21, Q_3 = 25$** (b) $17 - 12$ and $25 + 12$: **5 and 37; 41 is an outlier**

- (c) **at 29**, the largest value that is not an outlier (d) **21.7 and 20.3; median 20.5**

The mean moves by more than the median when the outlier goes.

- B9** (a) **P higher (502 g vs 498 g)** (b) **P more spread (IQR 6 g vs 3 g)** (c) **Q**

- (d) **P positive skew, Q symmetrical**

Each comparison names the statistic, gives both values and refers to the context.



B10 (a) $\sum x = 20 \times 15 = \boxed{300}$; $\sum x^2 = 20(9 + 225) = \boxed{4680}$

(b) $\frac{300 - 27}{19} = \boxed{14.4}$ (c) $\sqrt{\frac{4680 - 729}{19}} - 14.37^2 = \boxed{1.22}$

(d) **27 is 4 sd from the mean**, so it contributed most of $\sum (x - \bar{x})^2$

Remove the value from both totals, then recalculate with $n = 19$

C

Reasoning solutions

Your wording can be different. Check that your reason makes the same point.

CORE

C1 Height is frequency density, not frequency: 0–10 has $\frac{20}{10} = 2$ and 10–30 has $\frac{30}{20} = 1.5$:

the 0–10 bar is taller

C2 On the left, leaves read outwards from the stem: the stem comes first. **32, 35, 37**

C3 She must use the rule given: $52 + 1.0 \times 12 = 64$, and $66 > 64$: **66 is an outlier**

C4 The whisker ends at the largest value that is not an outlier, not at the limit.

Limit $30 + 15 = 45$: **whisker to 44; 52 marked as an outlier (×)**

STRETCH

C5 It needs named statistics, values and context.

Class A scored higher on average (median 61 vs 55) and its marks were more varied (IQR 18 vs 12)

C6 $Q_2 - Q_1 = 8 > Q_3 - Q_2 = 4$ means the lower half is more spread out: **negative skew**

C7 Only $\frac{5}{20}$ of the 10–30 class lies below 15: $20 + \frac{5}{20}(30) = 27.5$: **about 27 or 28**

C8 (a) The numerator is $(Q_3 - Q_2) - (Q_2 - Q_1)$. Since $Q_1 \leq Q_2 \leq Q_3$, each bracket lies between 0 and $Q_3 - Q_1$, so the numerator lies between $-(Q_3 - Q_1)$ and $Q_3 - Q_1$: **$-1 \leq \text{measure} \leq 1$**

(b) $Q_2 = \frac{Q_1 + Q_3}{2}$ gives $Q_3 - 2Q_2 + Q_1 = 0$: **the quartiles are symmetrical about the median**



D**Problem solving solutions**

There is often more than one correct method. Any complete, correct method scores full marks.

CORE

- D1** (a) 6×5 and 2.5×10 : **30 and 25**
 (b) $\frac{3}{5}(30) + 40 + \frac{6}{10}(25) = 18 + 40 + 15$: **73**
- D2** (a) $n = 23$: 12th, 6th and 18th values: **median 34, $Q_1 = 30$, $Q_3 = 40$**
 (b) $1.5 \times 10 = 15$: limits 15 and 55. 56 and 58 are above 55; 18 is above 15:
exactly two outliers, 56 and 58
- D3** (a) 2.2 is halfway between $Q_2 = 1.8$ and $Q_3 = 2.6$: $12.5\% + 25\% = 37.5\%$ of 120: **45**
 (b) $2.6 + 1.5 \times 1.4 = 4.7$; the heaviest is 4.2: **limit 4.7 kg, no outliers**
- D4** $n = 11$: 3rd, 6th and 9th values: $Q_1 = 44$, $Q_2 = 48$, $Q_3 = 55$
 $55 + 1.5 \times 11 = 71.5$, and $79 > 71.5$: **K's mark is an outlier**

STRETCH

- D5** (a) The 10–20 bar has area 10×2.4 for 24, so frequencies are 8, 24, 24, 20, 16, 8
 $\frac{5}{10}(24) + 24 + 20 + \frac{5}{10}(16) = 12 + 24 + 20 + 8$: **64**
 (b) 50th, in 20–25 (32 below): $20 + \frac{18}{24} \times 5$: **23.8 minutes**
- D6** (a) $n = 15$: 4th, 8th, 12th: **$Q_1 = 40$, $Q_2 = 46$, $Q_3 = 52$**
 (b) **Shop A customers spent longer on average (median 46 vs 44 minutes)**
and their times were less varied (IQR 12 vs 15 minutes)
- D7** With 13 values the quartiles are the 4th, 7th and 10th values, and they must stay 18, 25 and 33
 4th still 18: $a \leq 18$ and $2a \geq 18$; 7th still 25: $2a$ is not between 18 and 25, so $2a \geq 25$
 10th still 33: $2a \geq 33$; the maximum still 41: $2a \leq 41$: **$16.5 \leq a \leq 18$**



- D8** (a) Densities 3.6, 8.4, 5, 2.4, 0.3; tallest 8.4 is 12.6 cm, so 1.5 cm per unit; second tallest is 5: **7.5 cm**
- (b) $\frac{19 - 24 + 7.02}{19 - 7.02} \approx 0.169$: **positive skew**
- (c) **Not suitable**: the distances are skewed, but a normal distribution is symmetrical.

E**Exam practice solutions**

Mark-scheme notes are shown on the right of each line.

CORE

Exam questions for everyone

- E1** (a) 5.6×5 , 7.2×5 , 1×15 : **28, 36, 15** M1 A1
- (b) width: 5 minutes is 1 cm, so 15 minutes is 3 cm; height: density 5.6 is 5.6 cm, so density 1 is 1 cm M1
- 3 cm wide, 1 cm tall** A1
- (c) 65th: $15 + \frac{21}{36} \times 5 \approx 17.9$; 32.5th: $10 + \frac{16.5}{28} \times 5 \approx 12.9$; 97.5th: $20 + \frac{17.5}{30} \times 10 \approx 25.8$ M1 A1
- median 17.9, IQR 12.9 minutes** A1
- (d) $Q_3 - Q_2 = 7.9 > Q_2 - Q_1 = 5.0$: **positive skew** B1
- (e) $15 + 5 = 20$ waited over 30 minutes: $\frac{20}{130} \times \frac{19}{129}$ M1
- 0.0227** A1
-
- E2** (a) $n = 19$: 10th, 5th and 15th: **median 74, $Q_1 = 63$, $Q_3 = 85$** B1 B1
- (b) $1.5 \times 22 = 33$: limits 30 and 118; 138 is above 118, and 45 is above 30: **138 is the only outlier** M1 A1
- (c) **box from 63 to 85 with the median at 74; whiskers to 45 and 95; $\times$ at 138** B1 B1 B1
- (d) **group A higher on average (median 74 vs 70) and more varied (IQR 22 vs 12)** B1 B1
- (e) A: $85 - 74 = 74 - 63 = 11$: **symmetrical**; B: $78 - 70 = 8 > 70 - 66 = 4$: **positive skew** B1 B1



STRETCH

- E3 (a) 30 is $\frac{2}{8}$ of the way from $Q_2 = 28$ to $Q_3 = 36$: $50\% + \frac{1}{4} \times 25\% = 56.25\%$ of 160 M1
- 90 deliveries** A1
- (b) B: IQR 6, limits $26 - 9 = 17$ and $32 + 9 = 41$; $47 > 41$, while 18 and 38 lie inside M1 A1
- (c) B's IQR 6 is half of A's 12, but B's median 30 is higher than A's 28:
- B is more consistent, though slower on average** B1 B1
- (d) A: $\frac{36 - 56 + 24}{12} \approx 0.333$; B: $\frac{32 - 60 + 26}{6} \approx -0.333$ M1 A1
- A positive skew, B negative skew** A1
- (e) 38 equals the end of the upper whisker and is above Q_3 , so no quartile moves:
- the cross at 47 is removed; the box and whiskers are unchanged** B1 B1
- (f) **Not suitable**: A's times are skewed, and a normal distribution is symmetrical B1

