Estimating the Mean from Grouped Data

Maths Level 2 Exam-style practice · 5 questions · 23 marks · Calculator allowed

Remember

The method: 1) find the midpoint of each group; 2) multiply midpoint × frequency (call it *fx*); 3) add the *fx* column; 4) add the frequency column; 5) divide *fx* total by frequency total.

Midpoint: (lower + upper) ÷ 2. For $10 \leq t < 20$ the midpoint is 15.

Why an estimate? You assume every value in a group sits at its midpoint. The real values are unknown.

Modal class: the group with the highest frequency. **Median class:** the group containing the middle value – count up the frequencies to the halfway point.

Section A – The method

1. The table shows how long 40 customers waited at a bank.

Time, *t* (min)	Frequency	Midpoint	Midpoint × frequency
$0 \leq t < 4$	9		
$4 \leq t < 8$	16		
$8 \leq t < 12$	11		
$12 \leq t < 16$	4		

(a) Complete the table. (b) Estimate the mean waiting time. (c) Write down the modal class. (d) Which class contains the median? (7 marks)

2. Explain why your answer to (b) is an estimate. (1 mark)

Section B – In context

3. Ages of 60 people at a fitness class:

Age (years)	16–25	26–35	36–45	46–55	56–65
Frequency	8	19	17	11	5

(a) Estimate the mean age. (Use midpoints 20.5, 30.5, …) (b) The instructor claims 'most members are under 40'. Is that supported? Explain. (5 marks)

4. Amount spent by 50 shoppers (£):

Spend, *s*	0 ≤ s < 20	20 ≤ s < 40	40 ≤ s < 60	60 ≤ s < 100
Frequency	12	18	14	6

(a) Estimate the mean spend. (b) The shop's target is an average spend of £40. Has it been met? (c) The last group is wider than the others. How does this affect the reliability of the estimate? (6 marks)

5. A delivery company records journey lengths for 80 deliveries: 0–5 km: 22; 5–10 km: 31; 10–20 km: 19; 20–40 km: 8. Estimate the mean journey length and the total distance driven. (4 marks)

Estimating the Mean from Grouped Data – Answers

Level 2 · Working shown so you can check your method, not just the final number.

Section A

1. (a) Midpoints 2, 6, 10, 14. Products 18, 96, 110, 56. (b) Total fx = 280; 280 ÷ 40 = **7 minutes**. (c) **4 ≤ t < 8**. (d) 40 values → 20th/21st; cumulative 9, 25 → the median is in **4 ≤ t < 8**.
2. The exact times are unknown – each customer is assumed to be at the midpoint of their group.

Section B

3. (a) fx: 164 + 579.5 + 688.5 + 555.5 + 302.5 = 2,290. ÷ 60 = **38.2 years**. (b) Under 36: 27 people; the 36–45 group (17) straddles 40, so 'under 40' is at most 27 + some of 17 – around 27–35 of 60, i.e. roughly half. The claim is **not clearly supported**.
4. (a) Midpoints 10, 30, 50, 80. fx: 120 + 540 + 700 + 480 = 1,840 ÷ 50 = **£36.80**. (b) **No** – £36.80 is below £40. (c) A wide group means the midpoint (80) could be a long way from the real values, so the estimate is less reliable; the 6 big spenders could be anywhere from £60 to £100.
5. Midpoints 2.5, 7.5, 15, 30. fx: 55 + 232.5 + 285 + 240 = 812.5 ÷ 80 = **10.2 km** (10.16). Total ≈ **812.5 km**.

Where Level 2 students lose marks on this topic

- Dividing by the number of groups (4) instead of the total frequency (40).
- Using the group's upper or lower bound instead of the midpoint.
- Giving the modal class as the highest frequency (16) instead of the group (4 ≤ t < 8).