Statistics: centre and spread

IB Mathematics AI SL | Syllabus 4.3

50 questions ▪ Averages, variation and changing data

Learning goal	Calculate, compare and interpret measures of centre and spread.
Syllabus link	AI SL 4.3: measures of centre and spread; effects of constant changes to data.
Structure	Q1–15 short response; Q16–30 extended response; Q31–50 AI SL Paper 1-style; answer key at the back.
Technology	GDC available throughout, including Paper 1 practice. Use one-variable statistics for raw standard deviation, variance and discrete quartiles.

Instructions. Show your method and give units. Work in your exercise book. Give exact answers where practical; otherwise use 3 significant figures, unless stated differently. For descriptive statistics use σ_x (division by n), not s_x.
GDC check. For frequency tables, use the values (or class midpoints) with the frequency list. For variance, square the unrounded σ_x.
Quartiles. The answers use the medians of the lower and upper halves; for odd n, exclude the overall median. Different technology conventions may give different quartiles; identify your convention if it differs.

1. Short response

Questions 1–15 | Build fluency; GDC available.

1. Five students spend 12, 15, 13, 18 and 17 minutes completing a task. Find the mean time.

2. The waiting times at two service desks are recorded in seconds. Find the median for each desk. Show the ordered list you use.

(a) Desk A: 24, 31, 27, 18, 35.

(b) Desk B: 35, 42, 30, 51, 39, 45.

3. Three groups report the numbers of books read by their members in one month.

Group	Numbers of books
A	3, 4, 4, 5, 6, 6, 6
B	0, 1, 1, 2, 2, 4
C	1, 3, 5, 7

For each group, state the mode or modes, or state that there is no mode.

4. A café records the number of drinks bought in each of 25 orders.

Number of drinks, x	1	2	3	4	5
Frequency, f	4	8	7	4	2

Find the mean number of drinks per order and state the modal number.

5. Six customers spend the following amounts, in dollars, at a craft stall:

$$15, \quad 18, \quad 20, \quad 22, \quad 25, \quad 110.$$

Find the mean and median. Explain which better represents a typical customer's spending in this group.

6. For a group of 12 students, the mean volume of water drunk during one day is 1.75 litres. Find the total volume drunk by the group.

7. The mean time for five bus journeys is 23 minutes. Four of the journey times are 18, 21, 24 and 25 minutes. Find the time taken for the fifth journey.

8. The masses of 20 parcels are grouped in the table.

Mass, m (g)	$0 \le m < 200$	$200 \le m < 400$	$400 \le m < 600$	$600 \le m < 800$
Frequency	3	7	6	4

Write down the class midpoints, estimate the mean mass and state the modal class.

9. The lower quartile, median and upper quartile of the daily numbers of visitors to a small library are 48, 62 and 81, respectively. Find the interquartile range and explain what it measures in this context.

10. A delivery service records the following numbers of parcels collected in 11 one-hour periods:

$$9,\ 6,\ 13,\ 5,\ 18,\ 8,\ 10,\ 2,\ 15,\ 6,\ 12.$$

Use your GDC to find Q_1, the median, Q_3 and the interquartile range. State your quartile convention if it differs from the one on page 1.

11. Eight delivery times, in minutes, are

$$5, \quad 7, \quad 8, \quad 8, \quad 9, \quad 10, \quad 12, \quad 13.$$

Use your GDC to find the mean, standard deviation σ_x and variance. Include appropriate units.

12. Use the relationship between variance and standard deviation.

(a) The masses of some samples have standard deviation 1.8 g. Find their variance.

(b) Some race times have variance $12.25\,\text{s}^2$. Find their standard deviation.

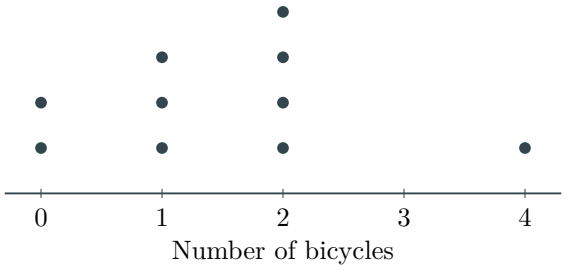
13. A balance adds 4 g to every mass reading. The recorded masses have mean 52 g, median 51 g, interquartile range 8 g and standard deviation 5 g. Each reading is corrected by subtracting 4 g. Find the corrected mean, median, interquartile range, standard deviation and variance.

14. Some daily walking distances, measured in kilometres, have the following summary.

Measure	Mean	Median	IQR	Standard deviation
Value (km)	2.5	2.4	0.8	0.6

Find the mean, median, interquartile range, standard deviation and variance when these distances are recorded in metres.

15. The dot plot shows the number of bicycles in each of 10 households. Each dot represents one household.



(a) Find the mean number of bicycles per household.

(b) State the median and modal number of bicycles.

2. Extended response

Questions 16–30 | Apply, compare and interpret; GDC available.

16. A cafe records the waiting times, in minutes, of eight customers.

8	9	10	11	12	12	12	14

(a) Calculate the mean and standard deviation of the recorded waiting times.

(b) The first value was entered incorrectly: it should be 18, not 8. Calculate the corrected mean, median and standard deviation.

(c) Does the increase in the calculated mean show that the cafe's service actually became slower? Explain your answer.

17. A warehouse counts the damaged items in each of 25 deliveries.

Damaged items, x	0	1	2	3	4
Number of deliveries	4	9	8	3	1

(a) Calculate the mean and standard deviation of the number of damaged items per delivery.

(b) Find the median and mode.

(c) A report says, "Every delivery contained the mean number of damaged items." Explain why this statement is incorrect, and write an accurate sentence using the mean.

18. The daily electricity use, e kWh, of a household is recorded for 30 days.

Electricity use (kWh)	$4 \le e < 8$	$8 \le e < 12$	$12 \le e < 16$	$16 \le e < 20$
Number of days	5	8	11	6

(a) Write down the four class midpoints and the modal class.

(b) Use the midpoints to estimate the mean daily electricity use.

(c) Explain why the exact mean cannot be found from this table. State the assumption made about each class when calculating the estimate.

19. Two machines fill packets labelled 500 g. The masses, in grams, of seven packets from each machine are shown.

Machine A	495	498	499	500	501	502	505
Machine B	490	494	498	500	502	506	510

(a) Calculate the mean and standard deviation for each machine.

(b) Find the median and interquartile range for each machine.

(c) Which machine produced more consistent masses in these samples? Support your answer using two measures of spread. Explain why these samples do not guarantee that every future packet will have mass close to 500 g.

20. The table shows the number of library visits made by a group of students in one week. The mean is 1.5 visits.

Visits	0	1	2	3
Number of students	3	4	k	2

(a) Form an equation in k and find its value.

(b) Find the median and mode.

(c) One further student, who made no visits, is added to the data. Find the new mean and median.

21. A fitness centre records how long visitors stay on a particular weekday.

Session	Number of visitors	Mean stay (minutes)
Morning	18	42
Evening	27	52

(a) Calculate the mean stay of all 45 visitors.

(b) Explain why taking the mean of 42 and 52 does not give the mean stay of all 45 visitors.

(c) Data for 15 weekend visitors are added. The mean stay of all 60 visitors is 50 minutes. Find the mean stay of the weekend visitors.

22. A charity receives nine donations, in euros:

$$8,\quad 10,\quad 10,\quad 12,\quad 12,\quad 14,\quad 14,\quad 16,\quad 120.$$

The donation of 120 euros is a genuine individual gift.

(a) Calculate the mean, median and interquartile range of all nine donations.

(b) Find the mean and median of the other eight donations. Describe the effect of the gift of 120 euros on the mean and median.

(c) Which of the original mean and median would you use to describe a typical individual donation? Justify your choice, and explain why the large gift should not automatically be removed from the charity's complete data.

23. A weather station summarises temperatures in degrees Celsius.

Statistic	Mean	Median	Q_1	Q_3	σ_x
Value (°C)	19.5	19	17.5	21	2.4

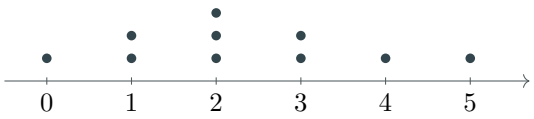
Temperatures are converted using $F = 1.8C + 32$.

(a) Find the mean and median in degrees Fahrenheit.

(b) Find the interquartile range, standard deviation and variance in Fahrenheit, giving appropriate units.

(c) Explain why adding 32 in the conversion has no effect on the variance.

24. The dot plot shows the number of books returned to a small reading club on each of several days. Each dot represents one day.



Number of books returned

(a) State the number of days recorded and calculate the mean number of books returned per day.

(b) Find the median and interquartile range.

(c) On the next day, five books are returned. Find the new mean and median, and explain why the mean increases while the median is unchanged.

25. A technician tests two chargers. The time, t minutes, to charge each of 20 identical batteries is recorded for each charger.

Time (minutes)	$30 \le t < 40$	$40 \le t < 50$	$50 \le t < 60$	$60 \le t < 70$
Charger A: frequency	2	6	10	2
Charger B: frequency	4	5	5	6

(a) Use the class midpoints to estimate the mean and standard deviation of the charging times for each charger.

(b) State the modal class for each charger.

(c) The technician wants a low mean charging time and consistent results. Which charger do these estimates favour? Give two numerical reasons, and state one limitation of using grouped data for this comparison.

26. Eleven people report their weekly podcast listening time, in hours:

$$3,\quad 4,\quad 6,\quad 8,\quad 9,\quad 11,\quad 13,\quad 16,\quad 20,\quad 22,\quad 28.$$

(a) Find the median, Q_1, Q_3 and interquartile range. Use the medians of the lower and upper halves, excluding the overall median.

(b) Another app reports $Q_1 = 7$ and $Q_3 = 18$ for the same data. Find its interquartile range and explain how both sets of quartiles can arise. State what should be checked before comparing interquartile ranges produced by different apps.

27. A walking app has recorded the following step counts for eight days:

$$4200, \quad 5100, \quad 5600, \quad 5900, \quad 6200, \quad 6500, \quad 6800, \quad 7700.$$

One day's count is missing. The mean over all nine days is 6200 steps.

(a) Find the missing step count.

(b) For the complete data, calculate the median and standard deviation.

(c) The user's daily target is at least 6000 steps. How many days met the target? Explain why a mean above 6000 does not show that the target was met every day.

28. Sensors in two classrooms produce the following temperature summaries.

Room	Mean (°C)	Median (°C)	IQR (°C)	σ_x (°C)
A	21.4	21.3	1.2	0.8
B	20.8	20.5	2.0	1.3

Every reading from room B is 0.6°C too low; room A's readings are correct.

(a) Find the corrected mean and median for room B.

(b) Find the corrected interquartile range, standard deviation and variance for room B, including units.

(c) After correction, compare the mean temperatures and decide which room had more stable temperatures. Use two measures of spread in your explanation.

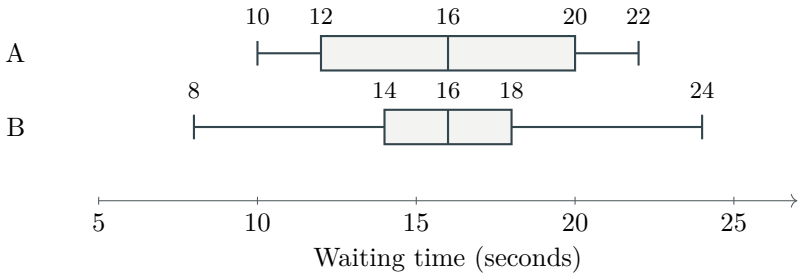
29. A cycling app converts a daily distance x km into activity points p using $p = ax + b$, where $a > 0$. For one rider, the distances have mean 8.4 km and standard deviation 2.5 km. The app wants the corresponding points to have mean 60 and standard deviation 15.

(a) Find a and b.

(b) The lower and upper quartiles of the distances are 6.5 km and 10.5 km. Find the interquartile range of the activity points.

(c) How many points are awarded for an 11 km ride? Find the variance of the activity points and give its units.

30. The box plots summarise waiting times at two ticket kiosks. The five labelled values on each plot are the minimum, Q_1, median, Q_3 and maximum. There are no outliers.



(a) For each kiosk, state the median and find the interquartile range and range.

(b) A customer says, "Kiosk B has less variable waiting times." Evaluate this claim using both measures of spread from part (a).

(c) Can the exact mean and standard deviation be calculated from either box plot alone? Explain your answer.

3. AI SL exam-style / Paper 1

Questions 31–50 | GDC available. Original questions with mark allocations.

31. A household records its daily water use, in litres, on nine days:

$$118,\quad 124,\quad 121,\quad 130,\quad 121,\quad 126,\quad 119,\quad 121,\quad 154.$$

(a) Find the mean and the standard deviation of the daily water use. [2 marks]

(b) Write down the median and the mode. [2 marks]

(c) Explain why the median may give a better description of a typical day's water use than the mean. [2 marks]

32. The number of food tokens bought by each of 25 visitors at a festival is recorded.

Number of tokens	0	1	2	3	4
Frequency	2	7	8	5	3

(a) Find the mean and the standard deviation of the number of tokens bought. [3 marks]

(b) Write down the mode. [1 mark]

(c) Each token costs \$2.50. Find the mean amount spent on tokens per visitor. [2 marks]

33. The heights, in centimetres, of ten seedlings are

$$8.2,\quad 9.1,\quad 9.4,\quad 10.0,\quad 10.3,\quad 10.8,\quad 11.2,\quad 11.5,\quad 12.4,\quad 14.1.$$

(a) Find the lower quartile, the upper quartile and the interquartile range. [3 marks]

(b) Find the standard deviation of the heights. [2 marks]

(c) A model predicts that every seedling will grow by exactly 1.8 cm. Write down the standard deviation of the predicted heights. [1 mark]

34. The waiting times, t minutes, of 30 passengers at a ferry terminal are summarized below.

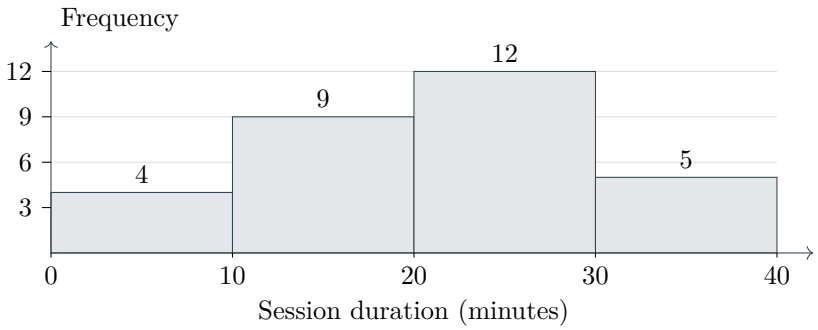
t (minutes)	$0 \le t < 4$	$4 \le t < 8$	$8 \le t < 12$	$12 \le t < 16$	$16 \le t < 20$
Frequency	5	11	9	4	1

(a) Write down the modal class. [1 mark]

(b) Use class midpoints to estimate the mean waiting time. [3 marks]

(c) Explain why the value found in part (b) is an estimate. [2 marks]

35. The histogram shows the duration, t minutes, of internet sessions at a library. The class intervals are $0 \le t < 10$, $10 \le t < 20$, $20 \le t < 30$ and $30 \le t < 40$.



(a) Find the total number of sessions. [1 mark]

(b) Write down the modal class. [1 mark]

(c) Use class midpoints to estimate the mean and the standard deviation of the session durations. [4 marks]

36. Two robots each complete the same assembly task six times. Their completion times, in seconds, are recorded.

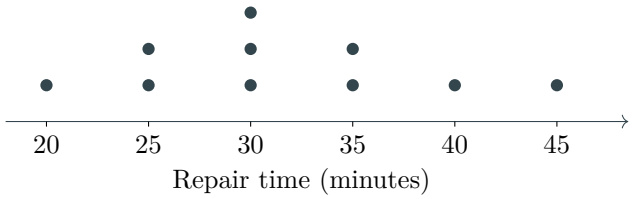
Run	1	2	3	4	5	6
Robot A	42	44	45	45	46	48
Robot B	39	43	45	45	47	51

(a) Find the mean completion time for each robot. [2 marks]

(b) Find the standard deviation of the completion times for each robot. [2 marks]

(c) Compare the robots' average completion times and the consistency of their performance. [2 marks]

37. The dot plot shows the time taken for each of ten scooter repairs at Garage P. Each dot represents one repair.

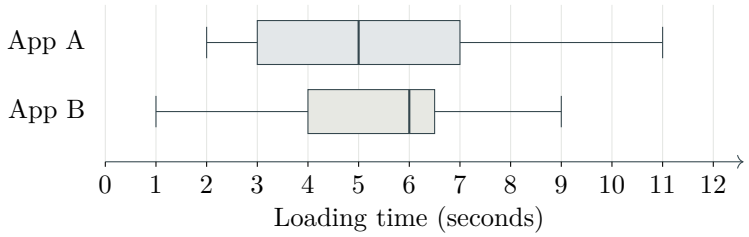


(a) Find the mean and the standard deviation of these repair times. [3 marks]

(b) Find the median and the interquartile range. [2 marks]

(c) Garage Q has a mean repair time of 31.5 minutes and a standard deviation of 10 minutes. State which garage has more consistent repair times. Justify your answer. [2 marks]

38. The box plots summarize the loading times of two apps. The whiskers show the minimum and maximum. App B has an upper quartile of 6.5 seconds.



(a) Write down the median loading time for each app. [2 marks]

(b) Find the interquartile range for each app. [2 marks]

(c) An update multiplies every loading time for App A by 0.75. Find the new median and interquartile range for App A. [2 marks]

39. Eight donations, in dollars, were recorded as

$$35, \quad 40, \quad 42, \quad 48, \quad 50, \quad 55, \quad 60, \quad 70.$$

(a) Calculate the mean of the recorded donations. [2 marks]

(b) The donation recorded as \$70 was actually \$170. Find the corrected mean and standard deviation. [3 marks]

(c) Explain why the median remains \$49 after this correction. [1 mark]

40. An exhibition records the visit durations of two groups of visitors.

Group	Number of visitors	Mean duration (minutes)
School group	18	37.5
Walking club	12	46.0

(a) Find the total visit duration for the school group. [2 marks]

(b) Calculate the mean visit duration for all 30 visitors. [2 marks]

(c) Five further visitors have a mean visit duration of 45 minutes. Find the mean duration for all 35 visitors. [2 marks]

41. The number of reusable cups returned by each participant at an event is shown below. The mean number returned is 2.2.

Number of cups	0	1	2	3	4
Frequency	2	3	k	5	4

(a) Find the value of k. [3 marks]

(b) Write down the mode. [1 mark]

(c) Find the standard deviation of the number of cups returned. [2 marks]

42. Eight volunteers collect the following masses of litter, in kilograms:

$$2.4, \quad 2.9, \quad 3.2, \quad 3.6, \quad 3.8, \quad 4.1, \quad 4.5, \quad 5.1.$$

(a) Find the mean mass collected by these volunteers. [2 marks]

(b) A ninth volunteer joins the group. Find the mass this volunteer must collect for the mean of all nine masses to be 4 kg. [2 marks]

(c) For the nine masses in part (b), write down the median. [1 mark]

43. The temperatures recorded by a weather station have a mean of 17.2 °C and a standard deviation of 2.4 °C. Celsius temperatures, C, are converted to Fahrenheit temperatures, F, using

$$F = 1.8C + 32.$$

(a) Calculate the mean temperature in degrees Fahrenheit. [2 marks]

(b) Calculate the standard deviation in degrees Fahrenheit. [2 marks]

(c) Hence, find the variance in $(°\text{F})^2$. [1 mark]

44. A coffee company compares the volumes dispensed by two machines. The available summary statistics are shown below; 1 L = 1000 mL.

Statistic	Machine A	Machine B
Mean	0.250 L	250 mL
Standard deviation	—	1.5 mL
Variance	$4.0 \times 10^{-6}\,\text{L}^2$	—

(a) Find the mean and standard deviation for Machine A, both in millilitres. [2 marks]

(b) Find the variance for Machine B in L^2. [2 marks]

(c) State which machine dispenses more consistent volumes. [1 mark]

45. The masses, m grams, of 40 apples are grouped as follows.

Mass, m (grams)	Frequency
$100 \le m < 120$	4
$120 \le m < 140$	9
$140 \le m < 160$	p
$160 \le m < 180$	10
$180 \le m < 200$	5

(a) Find p and write down the modal class. [2 marks]

(b) Use class midpoints to estimate the mean mass. [2 marks]

(c) Use class midpoints to estimate the standard deviation of the masses. [2 marks]

46. Two online shops summarize the number of orders they receive per day. Standard deviations are given to 3 significant figures.

Statistic	Shop A	Shop B
Mean	310	265
Median	240	260
Lower quartile	190	245
Upper quartile	270	280
Standard deviation	226	40.0

(a) Find the interquartile range for each shop. [2 marks]

(b) Explain why the median may be more suitable than the mean for describing a typical day's orders at Shop A. [2 marks]

(c) Shop B gains a contract for exactly 45 additional orders every day. Find its new mean and standard deviation. [2 marks]

47. A museum records the number of visitors entering during six one-hour periods.

Period	1	2	3	4	5	6
Visitors	52	55	55	60	63	69

(a) Find the mean and standard deviation of the six counts. [3 marks]

(b) A software error lists each count twice, giving 12 entries. Write down the mean and standard deviation of these 12 entries. [2 marks]

(c) Starting from the original six counts, a recalibration adds 10 to every count. Find the new mean and standard deviation. [2 marks]

48. The numbers of plastic bottles collected from seven classrooms, in ascending order, are

$$8, \quad 10, \quad 12, \quad a, \quad 16, \quad b, \quad 24.$$

The mean is 15 and the median is 14.

(a) Find a and b. [3 marks]

(b) Find the standard deviation of the seven numbers. [2 marks]

(c) Find the interquartile range. [1 mark]

49. Customers from two groups make purchases at an aquarium gift shop.

Group	Number of customers	Mean amount spent ($)
A	n	8.40
B	12	14.70

The mean amount spent by all the customers is $10.20.

(a) Find n. [3 marks]

(b) Find the total amount spent by all the customers. [2 marks]

(c) A fee of $1.50 is added to every customer's purchase. Write down the new mean amount paid. [1 mark]

50. Two machines process 12 items each. Their processing times, x seconds, are recorded below.

x (seconds)	17	18	20	22	23
Frequency: Machine A	0	3	6	3	0
Frequency: Machine B	2	0	8	0	2

(a) Find the mean processing time for each machine. [2 marks]

(b) Find the standard deviation of the processing times for each machine. [2 marks]

(c) An upgrade changes every processing time to $y = 0.8x + 2$. To meet a standard, the new mean must be at most 19 seconds and the new standard deviation at most 1.2 seconds. Determine which machine meets both requirements. Justify your answer. [3 marks]

Answer key and solution outlines

Use the working to check your reasoning as well as your final answer. Decimal answers are given to 3 significant figures where appropriate. Quartiles follow the convention on page 1; σ_x uses division by n. For calculator questions, retain unrounded values until the final answer.

Section 1: Short response (1–15)

1. $\bar{x} = \dfrac{12+15+13+18+17}{5} = \dfrac{75}{5} = \boxed{15 \text{ minutes}}$.

2.

(a) Ordered list: $18, 24, 27, 31, 35$. Median $= \boxed{27\,\text{s}}$.

(b) Ordered list: $30, 35, 39, 42, 45, 51$. Median $= \dfrac{39+42}{2} = \boxed{40.5\,\text{s}}$.

3. A: $\boxed{6}$ books (three occurrences). B: $\boxed{1 \text{ and } 2}$ books (two occurrences each). C: $\boxed{\text{no mode}}$; every value occurs once.

4. $\sum f = 25$ and $\sum fx = 4+16+21+16+10 = 67$. Mean $= \dfrac{67}{25} = \boxed{2.68}$ drinks per order. Modal number $= \boxed{2}$ drinks.

5. Mean $= \dfrac{210}{6} = \boxed{35 \text{ dollars}}$; median $= \dfrac{20+22}{2} = \boxed{21 \text{ dollars}}$. The median better represents typical spending: the unusually large 110-dollar purchase pulls the mean upwards.

6. Total $= 12(1.75) = \boxed{21 \text{ litres}}$.

7. Total time $= 5(23) = 115$ minutes. Fifth time $= 115 - (18+21+24+25) = \boxed{27 \text{ minutes}}$.

8. Midpoints: $100, 300, 500, 700$ g.

$$\bar{m} \approx \frac{3(100)+7(300)+6(500)+4(700)}{20} = \frac{8200}{20} = \boxed{410\,\text{g}}.$$

Modal class: $\boxed{200 \le m < 400}$, since its frequency is the largest.

9. IQR $= 81 - 48 = \boxed{33 \text{ visitors}}$. It measures the spread of the middle 50% of the daily visitor counts, between 48 and 81 visitors.

10. Ordered data: $2, 5, 6, 6, 8, 9, 10, 12, 13, 15, 18$. Using the stated convention, $\boxed{Q_1 = 6}$, $\boxed{\text{median} = 9}$, $\boxed{Q_3 = 13}$ and $\boxed{\text{IQR} = 7}$ parcels. Exclude the overall median 9 before finding the medians of the two halves. A correctly stated alternative technology convention may give different quartiles.

11. GDC one-variable statistics give $\boxed{\bar{x} = 9 \text{ minutes}}$ and $\boxed{\sigma_x = 2.45 \text{ minutes}}$ (3 s.f.). Squaring the unrounded standard deviation gives $\boxed{\sigma_x^2 = 6\,\text{min}^2}$.

12.

(a) Variance $= (1.8)^2 = \boxed{3.24\,\text{g}^2}$.

(b) Standard deviation $= \sqrt{12.25} = \boxed{3.5\,\text{s}}$.

13. Corrected mean $= 52 - 4 = \boxed{48\,\text{g}}$; median $= 51 - 4 = \boxed{47\,\text{g}}$; IQR $= \boxed{8\,\text{g}}$; standard deviation $= \boxed{5\,\text{g}}$; variance $= 5^2 = \boxed{25\,\text{g}^2}$. Subtracting a constant shifts the centre and leaves the measures of spread unchanged.

14. Multiply each distance by 1000. Mean $= \boxed{2500\,\text{m}}$, median $= \boxed{2400\,\text{m}}$, IQR $= \boxed{800\,\text{m}}$ and standard deviation $= \boxed{600\,\text{m}}$. Variance $= (600)^2 = \boxed{360000\,\text{m}^2}$; variance scales by 1000^2.

15.

(a) Frequencies for $0, 1, 2, 3, 4$ bicycles are $2, 3, 4, 0, 1$. Mean $= \dfrac{2(0)+3(1)+4(2)+1(4)}{10} = \boxed{1.5}$ bicycles per household.

(b) The fifth and sixth values are 1 and 2, so the median is $\boxed{1.5}$ bicycles. Modal number $= \boxed{2}$ bicycles.

Section 2: Extended response (16–30)

16.

(a) GDC: mean = 11 minutes; $\sigma_x = 1.80$ minutes.

(b) Replace the 8 by 18 in the data list. The corrected total is $88 - 8 + 18 = 98$, so the mean is $98/8 = 12.25$ minutes. Median = 12 minutes; GDC: $\sigma_x = 2.59$ minutes.

(c) No. The same eight waiting times are being described more accurately after a data-entry error is corrected. There is no new evidence of a change in service speed over time.

17.

(a) With values $0, 1, 2, 3, 4$ and frequencies $4, 9, 8, 3, 1$, GDC: mean = 1.52 damaged items per delivery; $\sigma_x = 1.02$ items.

(b) Median = 1 item; mode = 1 item.

(c) The mean describes the deliveries collectively; individual counts differ and must be whole numbers. For example, "Across the 25 deliveries, there were on average 1.52 damaged items per delivery."

18.

(a) Midpoints: $6, 10, 14, 18$ kWh. Modal class: $12 \le e < 16$ kWh.

(b) $\overline{e} \approx \dfrac{5(6) + 8(10) + 11(14) + 6(18)}{30} = 12.4$ kWh.

(c) The individual daily values within each interval are unknown. The estimate represents every value in a class by its midpoint, equivalently assuming that the mean within each class is its midpoint.

19.

(a) GDC: machine A has mean 500 g and $\sigma_x = 2.93$ g; machine B has mean 500 g and $\sigma_x = 6.32$ g.

(b) A: median 500 g, $Q_1 = 498$ g, $Q_3 = 502$ g, so IQR = 4 g. B: median 500 g, $Q_1 = 494$ g, $Q_3 = 506$ g, so IQR = 12 g.

(c) A has the smaller standard deviation ($2.93 < 6.32$ g) and smaller IQR ($4 < 12$ g), so its sampled masses are more consistent. Only seven packets from each machine were observed; this does not establish how every future packet will behave.

20.

(a) $\dfrac{4 + 2k + 3(2)}{3 + 4 + k + 2} = 1.5$, so $10 + 2k = 1.5(9 + k)$ and $k = 7$.

(b) There are 16 students. The 8th and 9th values are both 2, so median = 2 visits. Mode = 2 visits.

(c) The total remains 24 visits and the number of students becomes 17. New mean $= 24/17 = 1.41$ visits; the 9th value is 2, so the median remains 2 visits.

21.

(a) $\overline{x} = \dfrac{18(42) + 27(52)}{45} = \dfrac{2160}{45} = 48$ minutes.

(b) The two sessions have different numbers of visitors. Their means must be weighted by 18 and 27; $(42 + 52)/2 = 47$ incorrectly gives the sessions equal weight.

(c) The weekend total is $60(50) - 2160 = 840$ minutes. Weekend mean $= 840/15 = 56$ minutes.

22.

(a) Mean $= 216/9 = 24$ euros; median = 12 euros. The quartiles are 10 and 15 euros, so IQR = 5 euros.

(b) For the other eight gifts, mean $= 96/8 = 12$ euros and median = 12 euros. Including the large gift raises the mean from 12 to 24 euros and leaves the median unchanged.

(c) The original median, 12 euros, better represents a typical gift among these donations because the large gift pulls the mean upward. The gift is valid, so it belongs in a complete record; omitting it would understate the charity's receipts. Any separate summary excluding it should be clearly labelled.

23.

(a) Mean $= 1.8(19.5) + 32 = 67.1°$F; median $= 1.8(19) + 32 = 66.2°$F.

(b) IQR $= 1.8(21 - 17.5) = 6.3°$F. Standard deviation $= 1.8(2.4) = 4.32°$F. Variance $= 4.32^2 = 18.7\,(°\text{F})^2$ to 3 significant figures.

(c) Adding the same constant to every temperature also adds it to the mean, so each deviation from the mean is unchanged. The factor 1.8 scales the standard deviation, and its square scales the variance.

Section 2: Extended response (16–30) (continued)

24.

(a) There are 10 days. Total books = 23, so mean = $23/10 = 2.3$ books per day.

(b) The ordered data are $0, 1, 1, 2, 2, 2, 3, 3, 4, 5$. Median = 2 books, $Q_1 = 1$ book and $Q_3 = 3$ books; IQR = 2 books.

(c) New mean = $(23 + 5)/11 = 28/11 = 2.55$ books per day; new median = 2 books. The added value is above the previous mean, so the mean rises. The middle value of the new ordered list is still 2.

25.

(a) Use midpoints $35, 45, 55, 65$. GDC with the respective frequency lists gives A: estimated mean 51 minutes, estimated $\sigma_x = 8$ minutes; B: estimated mean 51.5 minutes, estimated $\sigma_x = 11.1$ minutes.

(b) A: $50 \leq t < 60$ minutes. B: $60 \leq t < 70$ minutes.

(c) The estimates favour A: its mean is lower ($51 < 51.5$ minutes) and its standard deviation is lower ($8 < 11.1$ minutes). Individual charging times within each class are unknown, so the exact means and standard deviations may differ from these midpoint estimates.

26.

(a) Median = 11 hours. The lower half is $3, 4, 6, 8, 9$ and the upper half is $13, 16, 20, 22, 28$. Thus $Q_1 = 6$ hours, $Q_3 = 20$ hours and IQR = 14 hours.

(b) The app's IQR is $18 - 7 = 11$ hours. Different quartile conventions can give different results. For example, an interpolating method averages 6 and 8 for Q_1, and 16 and 20 for Q_3. Check that the same quartile convention is used before comparing IQRs from different apps.

27.

(a) The recorded total is 48000 steps. Missing count = $9(6200) - 48000 = 7800$ steps.

(b) GDC for all nine values: median = 6200 steps; $\sigma_x = 1.10 \times 10^3$ steps (3 significant figures).

(c) Five days met the target: $6200, 6500, 6800, 7700, 7800$. A mean combines all nine counts, so higher counts can offset days below the target; four of these days were below 6000.

28.

(a) Add 0.6°C to every reading. Corrected mean = 21.4°C; corrected median = 21.1°C.

(b) The IQR and standard deviation are unchanged: IQR = 2.0°C, $\sigma_x = 1.3$°C. Variance = $1.3^2 = 1.69\,(°C)^2$.

(c) The corrected means are equal, both 21.4°C. Room A is more stable in these data: its IQR is 1.2°C compared with 2.0°C, and its standard deviation is 0.8°C compared with 1.3°C.

29.

(a) Since $a > 0$, $2.5a = 15$, giving $a = 6$. Then $8.4(6) + b = 60$, so $b = 9.6$.

(b) IQR = $6(10.5 - 6.5) = 24$ points. Adding 9.6 shifts both quartiles equally.

(c) $p = 6(11) + 9.6 = 75.6$ points. Variance = $15^2 = 225$ points^2.

30.

(a) A: median 16 s, IQR = $20 - 12 = 8$ s, range = $22 - 10 = 12$ s. B: median 16 s, IQR = $18 - 14 = 4$ s, range = $24 - 8 = 16$ s.

(b) B's middle half has less spread, since its IQR is smaller. However, B has the larger overall range. The claim needs a stated measure of variation; these two measures favour different kiosks.

(c) No. A box plot gives five summary values but not all the individual waiting times or their frequencies. These summaries are insufficient to determine the exact mean or standard deviation.

Section 3: AI SL exam-style / Paper 1 (31–50)

31.

(a) GDC: $\bar{x} = 126$ L and $\sigma_x = 10.498677\ldots \approx 10.5$ L.

(b) Median = 121 L; mode = 121 L.

(c) The high value 154 pulls the mean upwards. The median is less affected by this unusually high use and is close to the main cluster of values.

Section 3: AI SL exam-style / Paper 1 (31–50) (continued)

32.

(a) Enter values $0, 1, 2, 3, 4$ with frequencies $2, 7, 8, 5, 3$. GDC: $\bar{x} = 2$, $\sigma_x = 1.131370\ldots \approx 1.13$ tokens.

(b) Mode $= 2$ tokens.

(c) Mean expenditure $= 2.50 \times 2 = \$5.00$.

33.

(a) $Q_1 = 9.4\,\text{cm}$, $Q_3 = 11.5\,\text{cm}$ and IQR $= 11.5 - 9.4 = 2.1\,\text{cm}$.

(b) GDC: $\sigma_x = 1.627882\ldots \approx 1.63\,\text{cm}$.

(c) $1.63\,\text{cm}$; adding a constant does not change the standard deviation.

34.

(a) $4 \le t < 8$ minutes.

(b) Midpoints are $2, 6, 10, 14, 18$; $\bar{t} \approx \dfrac{5(2) + 11(6) + 9(10) + 4(14) + 1(18)}{30} = 8$ minutes.

(c) The exact times within each class are unknown. The calculation represents all times in a class by its midpoint.

35.

(a) $4 + 9 + 12 + 5 = 30$ sessions.

(b) $20 \le t < 30$ minutes.

(c) Enter midpoints $5, 15, 25, 35$ with frequencies $4, 9, 12, 5$. Estimated mean $= 21$ minutes; estimated $\sigma_t = \sqrt{84} \approx 9.17$ minutes.

36.

(a) Both means are 45 seconds.

(b) Robot A: $\sigma_A = 1.825741\ldots \approx 1.83$ seconds. Robot B: $\sigma_B = 3.651483\ldots \approx 3.65$ seconds.

(c) Both take the same time on average. Robot A is more consistent because its standard deviation is smaller.

37.

(a) Values $20, 25, 30, 35, 40, 45$ have frequencies $1, 2, 3, 2, 1, 1$. GDC: $\bar{x} = 31.5$ minutes and $\sigma_x = 7.088723\ldots \approx 7.09$ minutes.

(b) Median $= 30$ minutes; $Q_1 = 25$, $Q_3 = 35$, so IQR $= 10$ minutes.

(c) Garage P: its standard deviation is about 7.09 minutes, smaller than Garage Q's 10 minutes. The means are equal.

38.

(a) App A: 5 seconds; App B: 6 seconds.

(b) App A: $7 - 3 = 4$ seconds. App B: $6.5 - 4 = 2.5$ seconds.

(c) New median $= 0.75(5) = 3.75$ seconds; new IQR $= 0.75(4) = 3$ seconds.

39.

(a) $\bar{x} = 400/8 = \$50$.

(b) Corrected total $= 400 - 70 + 170 = 500$, giving $\bar{x} = \$62.50$. GDC on the corrected data gives $\sigma_x = 41.334005\ldots \approx \$41.3$.

(c) The fourth and fifth ordered values remain 48 and 50; their mean remains \$49.

40.

(a) $18(37.5) = 675$ minutes.

(b) $\bar{t} = \dfrac{675 + 12(46)}{30} = \dfrac{1227}{30} = 40.9$ minutes.

(c) $\bar{t} = \dfrac{1227 + 5(45)}{35} = \dfrac{1452}{35} = 41.485714\ldots \approx 41.5$ minutes.

41.

(a) $\dfrac{34 + 2k}{14 + k} = 2.2 \Rightarrow 34 + 2k = 30.8 + 2.2k \Rightarrow k = 16$.

(b) Mode $= 2$ cups.

(c) Enter values $0, 1, 2, 3, 4$ with frequencies $2, 3, 16, 5, 4$. GDC: $\sigma_x = 1.013245\ldots \approx 1.01$ cups.

Section 3: AI SL exam-style / Paper 1 (31–50) (continued)

42.

(a) $\bar{x} = 29.6/8 = 3.7\,\text{kg}$.

(b) Required total $= 9(4) = 36\,\text{kg}$. The ninth volunteer must collect $36 - 29.6 = 6.4\,\text{kg}$.

(c) Median $= 3.8\,\text{kg}$, the fifth ordered value.

43.

(a) $\bar{F} = 1.8(17.2) + 32 = 62.96\,^\circ\text{F}$, or $63.0\,^\circ\text{F}$ to 3 significant figures.

(b) $\sigma_F = 1.8(2.4) = 4.32\,^\circ\text{F}$.

(c) $\sigma_F^2 = 4.32^2 = 18.6624\,(^\circ\text{F})^2 \approx 18.7\,(^\circ\text{F})^2$.

44.

(a) Mean $= 1000(0.250) = 250\,\text{mL}$. Standard deviation $= 1000\sqrt{4.0 \times 10^{-6}} = 2\,\text{mL}$.

(b) $\sigma_B = 0.0015\,\text{L}$, so variance $= (0.0015)^2 = 2.25 \times 10^{-6}\,\text{L}^2$.

(c) Machine B, since $1.5\,\text{mL} < 2\,\text{mL}$.

45.

(a) $p = 40 - (4 + 9 + 10 + 5) = 12$. Modal class: $140 \le m < 160$ grams.

(b) Enter midpoints $110, 130, 150, 170, 190$ with frequencies $4, 9, 12, 10, 5$. Estimated mean $= 6060/40 = 151.5\,\text{g}$ ($152\,\text{g}$ to 3 significant figures).

(c) GDC with the same midpoints and frequencies: estimated $\sigma_m = 23.404059\ldots \approx 23.4\,\text{g}$.

46.

(a) Shop A: $270 - 190 = 80$ orders. Shop B: $280 - 245 = 35$ orders.

(b) Shop A's mean is well above its median, suggesting that days with very large order counts pull the mean upwards. The median is less affected by such high counts.

(c) New mean $= 265 + 45 = 310$ orders. Standard deviation remains 40.0 orders.

47.

(a) GDC: $\bar{x} = 59$ visitors, $\sigma_x = \sqrt{33} = 5.744562\ldots \approx 5.74$ visitors.

(b) Mean $= 59$ visitors; standard deviation ≈ 5.74 visitors. Repeating every observation equally leaves both statistics unchanged.

(c) Mean $= 59 + 10 = 69$ visitors; standard deviation remains approximately 5.74 visitors.

48.

(a) The fourth value is the median, so $a = 14$. The total is $7(15) = 105$; hence $b = 105 - (8 + 10 + 12 + 14 + 16 + 24) = 21$.

(b) GDC on $8, 10, 12, 14, 16, 21, 24$: $\sigma_x = 5.371884\ldots \approx 5.37$ bottles.

(c) $Q_1 = 10$, $Q_3 = 21$, so IQR $= 11$ bottles.

49.

(a) $\dfrac{8.40n + 12(14.70)}{n + 12} = 10.20$. Thus $8.40n + 176.40 = 10.20n + 122.40$, giving $n = 30$.

(b) There are 42 customers; total amount spent $= 42(10.20) = \$428.40$.

(c) New mean $= 10.20 + 1.50 = \$11.70$.

50.

(a) Both means are 20 seconds.

(b) Machine A: $\sigma_A = \sqrt{2} \approx 1.41$ seconds. Machine B: $\sigma_B = \sqrt{3} \approx 1.73$ seconds.

(c) Both new means are $0.8(20) + 2 = 18$ seconds. New standard deviations are $0.8\sqrt{2} \approx 1.13$ seconds for A and $0.8\sqrt{3} \approx 1.39$ seconds for B. Only Machine A meets both requirements.