Probability & Statistics

Chapter 2
Measures of central tendency

Cambridge AS level

NEW ERA
International school

2.2 THE MEAN

The mean is referred to more precisely as the arithmetic mean and it is the most commonly known average. The sum of a set of data values can be found from the mean.

We use the upper case Greek letter 'sigma', written Σ, to represent 'sum' and $\bar{x}$ to represent the mean, where x represents our data values.

2.2 THE MEAN

The notation used for ungrouped and for grouped data are shown on separate rows in the following table.

	Sum	Data values	Frequency of data values	Number of data values	Sum of data values	Mean
Ungrouped	Σ	x	–	n	Σx	$\overline{x} = \dfrac{\Sigma x}{n}$
Grouped	Σ	x	f	Σf	$\Sigma x f$ or $\Sigma f x$	$\overline{x} = \dfrac{\Sigma x f}{\Sigma f}$

TIP: $n = \sum f$ for grouped data.

TIP: $\Sigma x f = \Sigma f x$ indicates the sum of the products of each value and its frequency. For example, the sum of five 10s and six 20s is $(10 \times 5) + (20 \times 6) = (5 \times 10) + (6 \times 20) = 170$.

Example 1:

Five labourers, whose mean mass is 70.2kg, wish to go to the top of a building in a lift with some cement. Find the greatest mass of cement they can take if the lift has a maximum weight allowance of 500kg.

$$\sum x = \bar{x} \times n$$

$$= 70.2 \times 5$$

$$= 351 \ kg$$

KEY POINT:

For ungrouped data, the mean is $\bar{x} = \dfrac{\sum x}{n}$

For grouped data, $\bar{x} = \dfrac{\sum xf}{\sum f} = \dfrac{\sum fx}{\sum f}$

$$351 \ kg + y = 500$$

$$y = 149$$

The greatest mass of cement they can take is 149kg.

Find the mean of the 40 values, given in the following table.

x	31	32	33	34	35
f	5	7	9	8	11

x	31	32	33	34	35	
f	5	7	9	8	11	$\Sigma f = 40$
xf	155	224	297	272	385	$\Sigma xf = 1333$

$$\overline{x} = \frac{\sum xf}{\sum f}$$

$$= \frac{1333}{40} \qquad = 33.325$$

COMBINED SETS OF DATA

There are many different ways to combine sets of data. However, here we do this by simply considering all of their values together.

To find the mean of two combined sets, we divide the sum of all their values by the total number of values in the two sets.

For example, by combining the dataset 1, 2, 3, 4 with the dataset 4, 5, 6 we obtain a new set of data that has seven values in it: 1, 2, 3, 4, 4, 5, 6. Note that the value 4 appears twice.

COMBINED SETS OF DATA

For example, by combining the dataset 1, 2, 3, 4 with the dataset 4, 5, 6 we obtain a new set of data that has seven values in it: 1, 2, 3, 4, 4, 5, 6. Note that the value 4 appears twice.

$$\bar{x} = \frac{1+2+3+4}{4} = 2.5$$

$$\bar{y} = \frac{4+5+6}{3} = 5$$

$$\bar{z} = \frac{1+2+3+4+4+5+6}{7} = 3\frac{4}{7}$$

TIP: Note that the mean of the combined sets is not $\frac{2.5+5}{2}$.

A large bag of sweets claims to contain 72 sweets, having a total mass of 852.4g.

A small bag of sweets claims to contain 24 sweets, having a total mass of 282.8g.

What is the mean mass of all the sweets together?

Total number of sweets $= 72 + 24 = 96$

Total mass of sweets $= 852.4 + 282.8 = 1135.2g$

$$\bar{x} = \frac{1135.2}{96} = 11.825g$$

Example 4: A family has 38 films on DVD with a mean playing time of 1 hour 32 minutes. They also have 26 films on video cassette, with a mean playing time of 2 hours 4 minutes. Find the mean playing time of all the films in their collection.

$$38 + 26 = 64 \text{ films}$$

$$92 \times 38 + 124 \times 26 = 6720 \text{ min}$$

$$\bar{x} = \frac{6720}{64} = 105 \text{ min}$$

1 hour 45 min

MEANS FROM GROUPED FREQUENCY TABLES

When data are presented in a grouped frequency table or illustrated in a histogram or cumulative frequency graph, we lose information about the raw values. For this reason we cannot determine the mean exactly but we can calculate an estimate of the mean. We do this by using mid-values to represent the values in each class.

Coconuts are packed into 75 crates, with 40 of a similar size in each crate.

46 crates contain coconuts with a total mass from 20 up to but not including 25kg.

22 crates contain coconuts with a total mass from 25 up to but not including 40kg.

7 crates contain coconuts with a total mass from 40 up to but not including 54kg.

a. Calculate an estimate of the mean mass of a crate of coconuts.

b. Use your answer to part **a** to estimate the mean mass of a coconut.

Mass (kg)	20−	25−	40 − 54	
No. crates (f)	46	22	7	$\Sigma f = 75$
Mid-value (x)	22.5	32.5	47.0	
xf	1035	715	329	$\Sigma xf = 2079$

TIP: When gaps appear between classes of grouped data, class boundaries should be used to find class mid-values. The following example shows a situation in which using incorrect boundaries leads to an incorrect estimate of the mean.

a. $\bar{x} = \dfrac{2079}{75} = 27.72$ kg

b. $\bar{c} = \dfrac{27.72}{40} = 0.693\ kg$

Calculate an estimate of the mean age of a group of 50 students, where there are sixteen 18-year-olds, twenty 19-year-olds and fourteen who are either 20 or 21 years old.

Age (A years)	Mid-value (x)	No. students (f)	xf
$18 \leqslant A < 19$	18.5	16	296
$19 \leqslant A < 20$	19.5	20	390
$20 \leqslant A < 22$	21.0	14	294
		$\Sigma f = 50$	$\Sigma xf = 980$

$$\bar{x} = \frac{980}{50} = 19.6 \ years$$

TIP: Incorrect mid-values of $18, 19$ and 20.5 give an incorrect estimated mean of 19.1 years.

EXERCISE 2B

1 Calculate the mean of the following sets of numbers:

a   28, 16, 83, 72, 105, 55, 6 and 35

b   7.3, 8.6, 11.7, 9.1, 1.7 and 4.2

c $3\frac{1}{2}, 5\frac{1}{4}, 9\frac{3}{4}, -4\frac{1}{4}$ and $7\frac{3}{8}$.

2 a   The mean of 15, 31, 47, 83, 97, 119 and p^2 is 63. Find the possible values of p.

b   The mean of 6, 29, 3, 14, q, $(q+8)$, q^2 and $(10-q)$ is 20. Find the possible values of q.

3 Given that:

a $n = 14$ and $\Sigma x = 325.5$, find $\bar{x}$.

b $n = 45$ and $\bar{y} = 23.6$, find the value of Σy.

c $\Sigma z = 4598$ and $\bar{z} = 52.25$, find the number of values in the set of data.

d $\Sigma xf = 86$ and $\bar{x} = 7\frac{1}{6}$, find the value of Σf.

e $\Sigma f = 135$ and $\bar{x} = 0.842$, find the value of Σxf.

4 Find the mean of x and of y given in the following tables.

a

x	18.0	18.5	19.0	19.5	20.0
f	8	10	17	24	1

b

y	3.62	3.65	3.68	3.71	3.74
f	127	209	322	291	251

5 For the data given in the following table, it is given that $\bar{q} = 8\frac{5}{9}$.

q	7	8	9	10
f	9	13	a	11

Calculate the value of a.

6 Calculate an estimate of the mean of x and of y given in the following tables.

a

x	$0 \le x < 2$	$2 \le x < 4$	$4 \le x < 8$	$8 \le x < 14$
f	8	9	11	2

b

y	$13 \le y < 16$	$16 \le y < 21$	$21 \le y < 28$	$28 \le y < 33$	$33 \le y < 36$
f	7	17	29	16	11

7 An examination was taken by 50 students. The 22 boys scored a mean of 71% and the girls scored a mean of 76%. Find the mean score of all the students.

8 A company employs 12 drivers. Their mean monthly salary is $1950. A new driver is employed and the mean monthly salary falls by $8. Find the monthly salary of the new driver.

M 9 The mean age of the 16 members of a karate club is 26 years and 3 months. One member leaves the club and the mean age of those remaining is 26 years. Find the age of the member who left the club. Give a reason why your answer might not be very accurate.

10 The following table shows the hourly rates of pay, in dollars, of a company's employees.

Hourly rate ($)	6	7	8	109
No. employees (f)	8	11	17	1

a   Is the mean a good average to use here? Give a reason for your answer.

b   Find the mean rate of pay for the majority of the employees.

11 A train makes a non-stop journey from one city to another and back again each day. Over a period of 30 days, the mean number of passengers per journey is exactly 61.5. Exact one-way ticket prices paid by these passengers are given by percentage in the following table.

Price ($)	34	38	45
Passengers (%)	30	41	29

a   Calculate the total revenue from ticket sales, and explain why your answer is an approximation.

b   The minimum and maximum possible revenues differ by k. Find the value of k.

12 The heights, in centimetres, of 54 children are represented in the following diagram.

The children are split into two equal-sized groups: a 'tall half' and a 'short half'.

Calculate an estimate of the difference between the mean heights of these two groups of children.
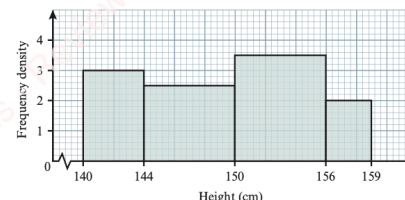
13 The following table summarises the number of tomatoes produced by the plants in the plots on a farm.

No. tomatoes	20–29	30–49	50–79	80–100
No. plots (f)	329	413	704	258

a   Calculate an estimate of the mean number of tomatoes produced by these plots.

b   The tomatoes are weighed accurately and their mean mass is found to be 156.50 grams. At market they are sold for $3.20 per kilogram and the total revenue is $50 350. Find the actual mean number of tomatoes produced per plot.

c   Why could your answer to part **b** be inaccurate?

14 Twenty boys and girls were each asked how many aunts and uncles they have. The entry 4/5 in the following table, for example, shows that 4 boys and 5 girls each have 3 aunts and 2 uncles.

a   Find the mean number of uncles that the boys have.

b   For the boys and girls together, calculate the mean number of:

   i   aunts   ii   aunts and uncles.

c   Suggest an alternative way of presenting the data so that the calculations in parts **a** and **b** would be simpler to make.

B / G		Aunts		
	0	1	2	3
Uncles 0	1/0	0/2	2/1	1/1
Uncles 1	0/0	3/4	0/4	0/0
Uncles 2	0/0	1/0	7/11	4/5
Uncles 3	1/0	0/0	0/1	0/1

15 A calculated estimate of the mean capacity of 120 refrigerators stored at a warehouse is 348 litres. The capacities are given in the following table.

Capacity (litres)	160–	200–	320–	400–p
No. refrigerators (f)	12	28	48	32

A delivery of n new refrigerators, all with capacities between 200 and 320 litres arrives at the warehouse. This causes the mean capacity to decrease by 8 litres. Find the value of n and state what assumptions you are making in your calculations.

16 A carpet fitter is employed to fit carpet in each of the 72 guest bedrooms at a new hotel. The following table shows how many rooms were completed during the first 10 days of work.

No. rooms completed	5	6 or 7
No. days (f)	2	8

Based on these figures, estimate how many more days it will take to finish the job. What assumptions are you making in your calculations?

17 In the figure opposite, a square of side 8 cm is joined edge-to-edge to a semicircle, with centre O. P is 2 cm from O on the figure's axis of symmetry.

Points X and Y are fixed but the position of Z is variable on the shape's perimeter.

a Find the mean distance from P to X,Y and Z when angle POZ is equal to:

  i 180°   **ii** 135°.

b Find obtuse angle POZ, so that the mean distance from P to X,Y and Z is identical to the mean distance from P to X and Y.