Mean average from grouped data

Starter

1. **(Review of last lesson)**

   The table shows the number of goals scored in a series of football matches.

Number of goals	1	2	3
Number of matches	8	8	x

   (a)   If the mean number of goals is 2.04, find x.

   (b)   If the modal number of goals is 3, find the smallest possible value of x.

   (c)   If the median number of goals is 2, find the largest possible value of x.

2. **(Review of last lesson)**

   The marks obtained by the members of a class are summarised in the table.

Mark	a	b	c
Frequency	x	y	z

   Calculate the mean mark in terms of a, b, c, x, y and z.

Notes

Last lesson we found the mean from a frequency table. In this lesson, we will use a similar method to *estimate* the mean from *grouped data*.

Consider the data in the table below.

Time (t) in sec	$50 \le t < 60$	$60 \le t < 70$	$70 \le t < 80$	$80 \le t < 90$	$90 \le t < 100$	$100 \le t < 120$
Frequency	2	3	6	4	4	1

The first column of data tells us there are 2 values that are $50 \le t < 60$, the 2nd column says there are 3 values that are $60 \le t < 70$.

Using our method from the previous lesson we want to multiply the frequency by a number from the first row. Since we do not know what the precise values in each interval were, we have to guess what they might have been. Our best guess is that the mean of the values in each interval is equal to the *mid-point*. i.e. for the 1st interval, 55; for the 2nd interval 65 etc.

Since the *mid-point is a best guess*, the value we calculate for the mean is an *estimate*.

Working:

$$\text{Mean} \approx \frac{(55 \times 2) + (65 \times 3) + (75 \times 6) + (85 \times 4) + (95 \times 4) + (110 \times 1)}{20} = 79.25 \text{ sec}$$

N.B.   We use the approximation symbol, $\approx$, since the calculation is an *estimation* of the mean.

Median and mode for grouped data

We can use the same method to find the median and mode for grouped data as with a frequency table. In such cases though, we call it the *median class* and *modal class*.

For the table above:
Modal class = $70 \le t < 80$ seconds since this interval has the highest frequency (6).

Median class: the frequency is 20 so we want the $\dfrac{20+1}{2} = 10.5$th value

i.e. the mean of the 10th and 11th values
The running total goes: $2, 5, 11, 15, 19, 20$
So the 10th and 11th values are both in the 3rd interval i.e. $70 \le t < 80$
Therefore, the median class is $70 \le t < 80$ seconds

**N.B.**   Be careful when the intervals are of different widths.

E.g. 1 The table shows the number of letters delivered to the houses in a street.

Mark	$0-2$	$3-4$	$5-7$	$8-12$
Frequency	10	8	5	3

(a)   Calculate an estimate of the mean number of letters delivered per house.
(b)   What is the modal class?
(c)   Find the median class.

E.g. 2 The bar chart shows the heights of 60 athletes.
(a)   Calculate an estimate for the mean height of the 60 athletes.
(b)   Explain why your answer is an **estimate** for the mean height.

Heights of athletes
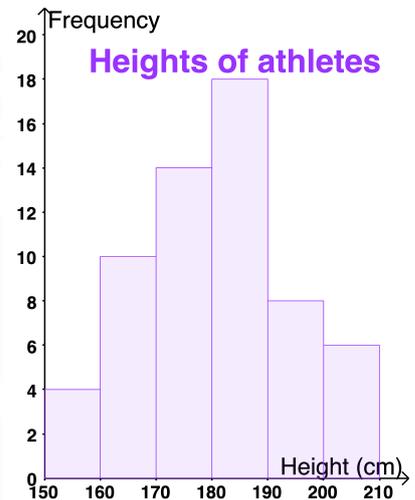
Video: **Estimating the mean from grouped data**

Solutions to Starter and E.g.s

Exercise
9-1 class textbook: p475 M14.4 Qu 1-6
A*-G class textbook: p429 M14.4 Qu 1-6
9-1 homework book: p163 M14.4 Qu 1-3
A*-G homework book: p120 M14.4 Qu 1-3

Summary

$$\text{Mean from grouped data} = \frac{\text{Sum of (mid-interval value} \times \text{frequency)}}{\text{Total frequency}}$$

Since the *mid-point is a best guess*, the value we calculate for the mean is an *estimate*.