Key Notation

x data value

$\bar{x}$ mean

f frequency

$\sum x$ sum of the data

n number of pieces of data

Mean

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

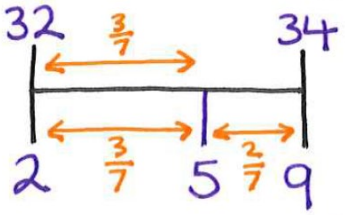
if group frequency table, then x is the midpoints.

Estimate Median

(continuous data)

$30 < t \leq 32$	$32 < t \leq 34$	$34 < t \leq 35$
2	7	1

$n = 10$ median = $10 \div 2 = 5^{th}$ number (no need to do $\frac{n+1}{2}$)



Median = $32 + (2 \times \frac{3}{7}) = 32.9$

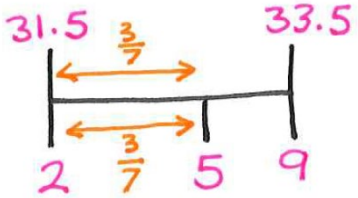
Estimate Median

(continuous data rounded)

Watch out for gaps between the intervals.

Interval 30 – 31 should be rewritten as 29.5 – 31.5 etc.

30 – 31	32 – 33	34 – 35
2	7	1



Median = $31.5 + (2 \times \frac{3}{7}) = 32.4$

Standard Deviation

$$\sigma = \sqrt{\frac{\sum x^2}{n} - (\bar{x})^2}$$

Standard Deviation = σ

Variance = σ^2

For continuous data the x is the mid-point

Short cut if you are given S_{xx}

$$\sigma = \sqrt{\frac{S_{xx}}{n}}$$

Formula Booklet

Standard deviation = √(Variance)

Interquartile range = $\text{IQR} = Q_3 - Q_1$

For a set of n values $x_1, x_2, \ldots x_i, \ldots x_n$

$$S_{xx} = \Sigma(x_i - \bar{x})^2 = \Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}$$

Standard deviation = $\sqrt{\frac{S_{xx}}{n}}$ or $\sqrt{\frac{\sum x^2}{n} - \bar{x}^2}$

Coding

ORIGINAL DATA ⟷ USER FRIENDLY DATA

Mean
is affected by all calculations.

Standard Deviation
is affected by multiplication and division only.

Original Mean	coding	New Mean ($\bar{y}$)
$\bar{x}$	$\times 3$	$3\bar{x}$
$\bar{x}$	$\div 3$	$\frac{\bar{x}}{3}$
$\bar{x}$	$\times 3 + 2$	$3\bar{x} + 2$
$\bar{x}$	$\div 2 - 4$	$\frac{\bar{x}}{2} - 4$

Original Standard Deviation	coding	Original Mean (σ_y)
σ_x	$\times 3$	$3\sigma_x$
σ_x	$+ 3$	σ_x
σ_x	$\times 3 - 2$	$3\sigma_x$
σ_x	$\div 3 + 2$	$\frac{\sigma_x}{3}$