SCAN ME

Variance and Standard Deviation

REVISE THIS TOPIC

1 The table below shows the times spent waiting on the phone *w*, to the nearest minute, for 16 patients who phoned their doctor's surgery.

Patient	A	B	C	D	E	F	G	H	I	J	K	L
Wait time (*w* mins)	34	3	7	17	9	24	23	6	1	92	36	6

(a) Calculate the mean wait time. **(1)**

(b) Calculate the variance of the wait times and state the units. **(2)**

(c) Calculate the standard deviation of the wait times and state the units. **(2)**

(d) Show clearly that all of the wait times are within three standard deviations of the mean. **(2)**

(a) $\overline{w}$ = 21.5 minutes (from calculator)

(b) σ^2 = 581.25 minutes^2 (from calculator)

(c) σ = 24.1 minutes (from calculator)

(d) $\overline{w} + 3\sigma = 21.5 + 3 \times 24.1\ldots$ The maximum wait time was 92 minutes.

$= 93.827\ldots$ $92 < 93.827\ldots$

$\overline{w} - 3\sigma = 21.5 - 3 \times 24.1\ldots$ The minimum wait time was 1 minute.

$= -50.827$ $1 > -50.827$

(Total for Question 1 is 7 marks)

2 The data below shows the number of points scored by a basketball team p, in the first half of a season.

67 71 75 79 85 85 88 88 91

(a) Calculate the mean number of points scored. **(1)**

(b) Calculate the standard deviation of the points scored. **(1)**

The basketball team played a 9 more games in the second half of the season.
The points for these games, q, are summarised below.

$\sum q = 774 \qquad \sum q^2 = 66\,836$

(c) Calculate the mean and standard deviation of the points second in the second half of the season. **(3)**

(d) Use your answers to parts (a), (b) and (c) to compare the points scored by the team in the first half of the season to the points scored in the second half of the season. **(2)**

(a) $\bar{p} = 81$ points (from calculator)

(b) $\sigma_p = 7.93$ points (from calculator)

(c) $\bar{q} = \frac{774}{9} = 86$ points

$\sigma_q = \sqrt{\frac{66836}{9} - 86^2}$

$= 5.497$ points

(d) The mean number of points scored in the second 9 games (86) is higher than that for the first 9 games (81) so the team scored more points on average in the second half of the season.

The standard deviation for the second half of the season (5.497) is lower than that for the first half of the season (7.93) so the points were less varied in the second half of the season.

(Total for Question 2 is 7 marks)

3 A sample of 8 students from Year 7 and 8 students from Year 10 were taken.

The heights of the Year 7 students (x) and Year 10 students (y), in cm, are summarised below.

$$\sum x = 1228 \qquad \sum x^2 = 189\,060 \qquad \sum y = 1368 \qquad \sum y^2 = 235\,032$$

(a) Calculate the mean and standard deviation of the heights of Year 7 students. **(3)**

(b) Calculate the mean and standard deviation of the heights of Year 10 students. **(3)**

(c) Compare the heights of Year 7 and Year 10 students. **(2)**

(d) Calculate the mean and standard deviation of the heights of all 16 students. **(3)**

(a) $\bar{x} = \frac{1228}{8} = 153.5$ cm

$\sigma_x = \sqrt{\frac{189060}{8} - 153.5^2}$

$= 8.38$ cm

(b) $\bar{y} = \frac{1368}{8} = 171$ cm

$\sigma_y = \sqrt{\frac{235032}{8} - 171^2}$

$= 11.75$ cm

(c) The mean height of Year 7 students (153.5) was lower than the mean for Year 10 (171) so Year 10 students were taller on average.

The standard deviation for Year 7 height (8.38) was lower than that for Year 10 (11.75) so Year 7 heights were less varied.

(d) Mean $= \frac{1228 + 1368}{16} = 162.25$ cm

Standard Deviation $= \sqrt{\frac{189060 + 235032}{16} - 162.25^2}$

$= 13.442$ cm

(Total for Question 3 is 11 marks)

4 8 different teams were timed to escape from an escape room.
Their times to escape x, to the nearest minute, are summarised below.

$$\sum x = 434 \qquad \sum (x - \bar{x})^2 = 187.5$$

(a) Calculate the mean time taken to escape. **(1)**

(b) Calculate the standard deviation of the times taken to escape. **(2)**

Another team is timed to escape from an escape room.
This extra team escapes in 55 minutes.
The mean and standard deviation are recalculated for all 9 times.

(c) Without further calculation, state what effect, if any, including the extra time of 55 minutes will have on

(i) the mean
(ii) the standard deviation **(2)**

(a) $\bar{x} = \frac{434}{8} = 54.25$ minutes

(b) $\sigma_x = \sqrt{\frac{187.5}{8}}$

$= 4.84$ minutes

(c) (i) The extra time of 55 minutes is greater than the mean so the mean will increase.

(c) (ii) $\bar{x} + \sigma_x = 59.09$ so the extra time of 55 minutes is within 1 standard deviation of the mean.
This means the standard deviation will decrease.

(Total for Question 4 is 5 marks)

5 Martin wants to research the number of bedrooms and bathrooms there were in properties in his town.

Martin goes into town and asks the people he meets their address and how many bedrooms and bathrooms are in their home. [Side note don't tell strangers your address].

He continues until he has a sample of size 40 ensuring each address is different.

(a) Write down the name of this sampling technique. **(1)**

The tables below summarise the number of bedrooms (x) and bathrooms (y) for the 40 properties.

Number of Bedrooms (x)	1	2	3	4	5
Frequency (f)	1	6	22	9	2

Number of Bathrooms (y)	1	2	3
Frequency (f)	28	9	3

(b) Calculate the mean number of bedrooms. **(1)**

(c) Calculate the standard deviation for the number of bedrooms. **(1)**

(d) Calculate the mean number of bathrooms. **(1)**

(e) Calculate the standard deviation for the number of bathrooms. **(1)**

(f) Are the number of bedrooms or number of bathrooms more varied? **(1)**
Give a reason for your answer.

(a) Opportunity sampling.

(b) $\bar{x}$ = 3.125 (from calculator)

(c) σ_x = 0.812 (from calculator)

(d) $\bar{y}$ = 1.375 (from calculator)

(e) σ_y = 0.620 (from calculator)

(f) The number of bedrooms is more varied because it has a higher standard deviation.

(Total for Question 5 is 6 marks)

6 A headteacher wants to know which day of the week is the worst for student lateness to school. To test this, they record the number of late students each weekday for 4 weeks.

The headteacher works out the mean number and standard deviation of the number of late students for each weekday.
The table below summarises the headteacher's calculations.

Day	**Mon**	**Tues**	**Weds**	**Thurs**	**Fri**
Mean	12	12.5	12.6	10.25	23
Standard Deviation	0	1.66	3.04	–2.28	2.55

As maths teacher checks the data and realises that two of the values were calculated incorrectly.

(a) State which two values were calculated incorrectly. Give reasons for your answers. **(2)**

The headteacher also wants to know about the number of detentions being set.

The number of detentions received in a one-month period for 30 Year 10 students (x) and 30 Year 11 students (y) are summarised below.

$\sum x = 126$ $\sum x^2 = 828$ $\sum y = 93$ $\sum (y-\bar{y})^2 = 392.7$

(b) By working out the standard deviation determine which of the year groups had more variation in their number of detentions. **(5)**

(a) The mean for Wednesdays cannot be 12.6

If there were 4 Wednesdays in the sample the value must end in .0 or .25 or .5 or .75

The standard deviation for Thursday cannot be –2.28

The standard deviation cannot be negative.

(b) $\sigma_x = \sqrt{\frac{828}{30} - \left(\frac{126}{30}\right)^2}$ $\sigma_y = \sqrt{\frac{392.7}{30}}$

= 3.156 detentions. = 3.618 detentions.

There was more variation in the detentions for Year 11 students as they had a higher standard deviation.

(Total for Question 6 is 7 marks)

7 A lottery draws 6 winning numbers.
Players win a prize if their ticket matches 2 or more of the winning numbers.

The prize breakdown for a lottery draw is shown below.

Numbers Matched	Prize Value (£)	Number of Prizes
6	11 449 068	1
5	1750	186
4	140	12 142
3	30	222 468
2	2	1 678 011

(a) Calculate the mean prize value to the nearest pound. **(1)**

(b) Calculate the percentage of the prizes awarded that were below the mean prize value. **(2)**

(c) Calculate the standard deviation of the prize values to the nearest pound. **(1)**

(d) State a reason why the standard deviation may not be a suitable measure of spread to describe these data. **(1)**

The prize values (£x) for a different lottery draw are summarised below.

$n = 1\,800\,000$ $\sum x = 22\,752\,692$ $S_{xx} = 96\,868\,911\,115\,690$

(e) Calculate, to the nearest pound, the mean and standard deviation of the prize values for this lottery draw. **(3)**

(a) Mean = £12 (from calculator)

(b) $\frac{1\,678\,011}{1\,912\,808} \times 100 = 87.7\%$

(c) Standard Deviation = £8278 (from calculator)

(d) The standard deviation is affected by extreme values/outliers e.g. the prize for matching 6 numbers

(e) $\bar{x} = \frac{22\,752\,692}{1\,800\,000} = £12.64$ $\sigma_x = \sqrt{\frac{96\,868\,911\,115\,690}{1\,800\,000}} = £7335.94$

£13 (nearest pound) £7336 (nearest pound)

(Total for Question 7 is 8 marks)

8 The speeds S, to the nearest mph, of 40 cars travelling through a village are shown below.

Vehicle speed (S mph)	**Frequency (f)**	**Midpoint (x mph)**
$0 \leq S < 10$	1	5
$10 \leq S < 20$	3	15
$20 \leq S < 30$	31	25
$30 \leq S < 40$	4	35
$40 \leq S < 50$	1	45

You may use $\sum fx = 1010$ and $\sum fx^2 = 27\,000$

(a) Calculate, to 2 decimal places, an estimate for the mean vehicle speed. **(1)**

(b) Calculate, to 2 decimal places, an estimate for the standard deviation of the vehicle speeds. **(2)**

(c) Explain why your answers to parts (a) and (b) are only estimates. **(1)**

It is found that one of the vehicle speeds was incorrectly recorded as 2 mph but was in fact 25 mph.

(d) Without calculate a new estimates state what effect, if any, using the correct speed will have on

(i) your estimate for the mean
(ii) your estimate for the standard deviation

Give reasons for your answers. **(2)**

(a) $\bar{x} = \frac{1010}{40} = 25.25$ mph

(b) $\sigma_x = \sqrt{\frac{27000}{40} - 25.25^2} = 6.12$ mph

(c) The values used were the midpoints of the class intervals rather than the actual speeds.

(d) (i) The correct speed will result in an increase of the mean since 25 > 2

(d) (ii) $\bar{x} - \sigma_x = 19.13$ so the corrected speed is within 1 standard deviation of the mean.
The mean will have also increased by a small amount but since 25 is very close to the mean the corrected speed will still be within one standard deviation of the mean so the standard deviation will decrease.

(Total for Question 8 is 6 marks)

9 At a jigsaw puzzle competition entrants are timed to complete a jigsaw puzzle.
There are 50 entrants, some of which are professionals, and the rest are amateurs.

The times of the amateurs (x) and the professionals (y), to the nearest minute are summarised below.

$\sum(x-\bar{x})^2 = 5082$ $\quad \bar{y} = 43$ $\quad \sigma_x = 11$ $\quad \sigma_y = 3$

The total time for all 50 entrants was 3578 minutes.

Work out the standard deviation of the times all 50 of the entrants. **(6)**

$$11 = \sqrt{\frac{5082}{n_x}}$$

$$n_x = \frac{5082}{11^2} = 42$$

$$n_y = 50 - 42$$

$$n_y = 8$$

$$43 = \frac{\sum y}{8}$$

$$\sum y = 43 \times 8$$

$$\sum y = 344$$

$$\sum x = 3578 - 344$$

$$\sum x = 3234$$

$$\bar{x} = \frac{3234}{42} = 77$$

$$11 = \sqrt{\frac{\sum x^2}{42} - 77^2}$$

$$\sum x^2 = 42(11^2 + 77^2)$$

$$\sum x^2 = 254100$$

$$3 = \sqrt{\frac{\sum y^2}{8} - 43^2}$$

$$\sum y^2 = 8(3^2 + 43^2)$$

$$\sum y^2 = 14864$$

$$\sigma = \sqrt{\frac{254100 + 14864}{50} - \left(\frac{3234 + 344}{50}\right)^2}$$

$$= 16.076 \text{ minutes}$$

(Total for Question 9 is 6 marks)