GCSE Statistics: Grade 7

Worked solutions

GCSE Statistics practice

1 **STANDARD DEVIATION**

The waiting times, in minutes, of six patients at Clinic A are shown below.

12 15 9 14 10 12

a. Calculate the standard deviation of these waiting times. Give your answer correct to 2 decimal places.

b. At Clinic B the mean waiting time is 12 minutes and the standard deviation is 4.5 minutes. Compare the consistency of the waiting times at the two clinics.

(4 marks)

a) Find the mean first

$$\text{sum of } x = 12 + 15 + 9 + 14 + 10 + 12$$

$$\text{sum of } x = 72$$

$$\text{mean} = 72 \div 6$$

$$\text{mean} = 12$$

Square each value and add

$$\text{sum of } x^2 = 144 + 225 + 81 + 196 + 100 + 144$$

$$\text{sum of } x^2 = 890$$

$$\text{sd} = \sqrt{\left(\frac{\text{sum of } x^2}{n} - \text{mean}^2\right)}$$

$$\text{sd} = \sqrt{\left(\frac{890}{6} - 12^2\right)}$$

$$\text{sd} = \sqrt{(148.333\ldots - 144)}$$

$$\text{sd} = \sqrt{4.333\ldots}$$

$$\text{sd} = 2.08 \text{ minutes}$$

b)

2.08 is less than 4.5

Clinic A has the lower standard deviation

so its waiting times are more consistent

a) 2.08 minutes b) Clinic A more consistent

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

2 SPEARMAN'S RANK CORRELATION COEFFICIENT

Two judges, Judge X and Judge Y, each rank eight cakes at a village show. Rank 1 is the best cake.

Cake	A	B	C	D	E	F	G	H
Judge X	1	2	3	4	5	6	7	8
Judge Y	3	1	2	6	4	8	5	7

a. Calculate Spearman's rank correlation coefficient for these ranks. Give your answer correct to 3 decimal places.

b. Interpret your answer to part (a) in context.

(4 marks)

a) d = rank from X − rank from Y

d −2, 1, 1, −2, 1, −2, 2, 1

d^2 4, 1, 1, 4, 1, 4, 4, 1

sum of d^2 = 4 + 1 + 1 + 4 + 1 + 4 + 4 + 1

sum of d^2 = 20

$$r = 1 - \frac{6 \times \text{sum of } d^2}{n(n^2 - 1)}$$

$$r = 1 - \frac{6 \times 20}{8 \times (64 - 1)}$$

$$r = 1 - \frac{120}{504}$$

$$r = 1 - 0.238...$$

$$r = 0.762$$

b)

0.762 is close to 1

Strong positive correlation

The two judges mostly agree on the order of the cakes

a) 0.762 b) strong positive, judges mostly agree

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

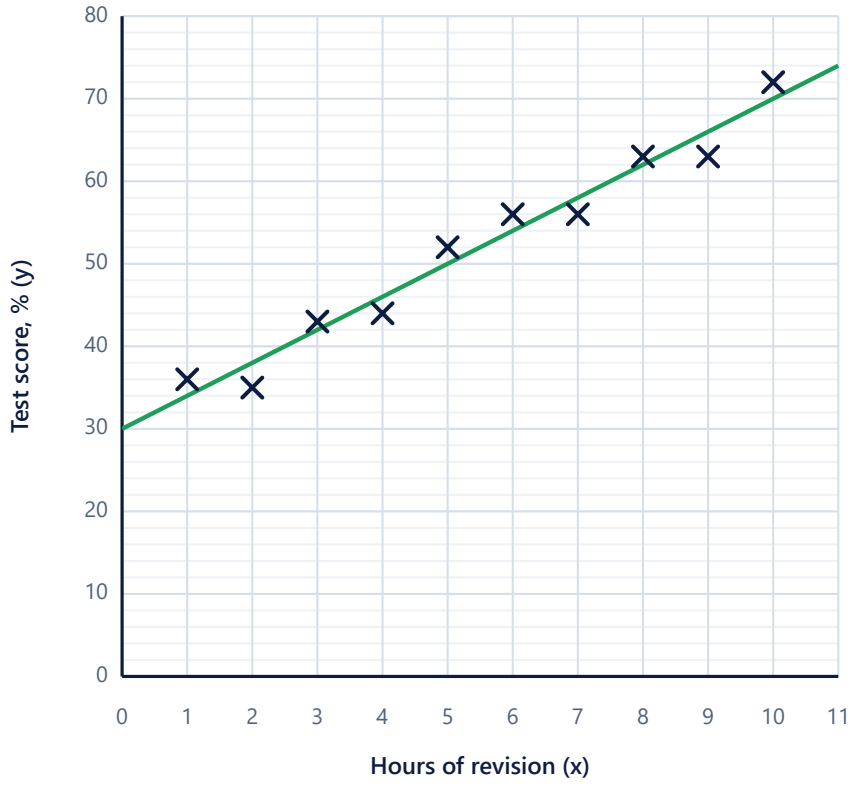
3 **THE REGRESSION LINE Y = A + BX**

A teacher records the number of hours of revision, x, and the test score, y %, for each of 10 students.
The scatter diagram shows this information.
The equation of the regression line of y on x is $y = 30 + 4x$

a. Draw the regression line on the scatter diagram.

b. Interpret the value 4 in the equation of the regression line.

c. Use the equation to estimate the test score of a student who revises for 6.5 hours.



(4 marks)

a) Work out two points on the line

When $x = 0$

$y = 30 + 4 \times 0$

$y = 30$

When $x = 10$

$y = 30 + 4 \times 10$

$y = 70$

Plot (0, 30) and (10, 70)

Join them with a straight line

b) 4 is the gradient

For each extra hour of revision

the test score goes up by 4%, on average

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

c)

$y = 30 + 4 \times 6.5$

$y = 30 + 26$

$y = 56$

a) Line drawn b) 4% more per extra hour c) 56%

4 **WEIGHTED INDEX NUMBERS**

A council works out index numbers for the costs of running a leisure centre in 2025, using 2020 as the base year.

Cost	Index number (2020 = 100)	Weight
Staff	112	6
Energy	135	3
Maintenance	94	1

a. Calculate the weighted index number for the costs of running the leisure centre in 2025.

b. Interpret your answer to part (a).

(4 marks)

a) Multiply each index number by its weight

$112 \times 6 = 672$

$135 \times 3 = 405$

$94 \times 1 = 94$

$672 + 405 + 94 = 1171$

Add the weights

$6 + 3 + 1 = 10$

$\text{weighted index} = \frac{\text{sum of (index} \times \text{weight)}}{\text{sum of weights}}$

$\text{weighted index} = \frac{1171}{10}$

$\text{weighted index} = 117.1$

b)

$117.1 - 100 = 17.1$

Overall, running costs rose by 17.1% between 2020 and 2025

a) 117.1 b) costs up 17.1% since 2020

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

5 **STANDARDISED SCORES**

Amara and Ben are in the same year group. The table shows the mean and standard deviation of the year group's marks in two tests, and Amara's marks.

Test	Mean	Standard deviation	Amara's mark
Maths	56	8	68
Physics	52	5	61

a. Calculate Amara's standardised score for each test.

b. In which test did Amara do better, compared with the rest of the year group? Give a reason for your answer.

c. Ben's standardised score in the Maths test was −0.75. Work out Ben's mark in the Maths test. (5 marks)

a) standardised score = $\frac{\text{mark} - \text{mean}}{\text{standard deviation}}$

$\text{Maths} = \frac{68 - 56}{8}$

$\text{Maths} = \frac{12}{8}$

$\text{Maths} = 1.5$

$\text{Physics} = \frac{61 - 52}{5}$

$\text{Physics} = \frac{9}{5}$

$\text{Physics} = 1.8$

b)

1.8 is greater than 1.5

Amara did better in Physics

Her Physics mark is further above the mean, measured in standard deviations

c) mark = mean + (standardised score × sd)

$\text{mark} = 56 + (-0.75 \times 8)$

$\text{mark} = 56 - 6$

$\text{mark} = 50$

a) Maths 1.5, Physics 1.8 b) Physics c) 50

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

6 GEOMETRIC MEAN

The number of members of a climbing club went up by 50% in 2023, by 20% in 2024 and by 8% in 2025.

a. Calculate the geometric mean of the three yearly multipliers. Give your answer correct to 3 significant figures.

b. Interpret your answer to part (a) in context.

(4 marks)

a) Write each change as a multiplier

1.5, 1.2, 1.08

Multiply them together

$1.5 \times 1.2 \times 1.08 = 1.944$

3 values, so take the cube root

geometric mean = cube root of 1.944

geometric mean = 1.2480...

geometric mean = 1.25 (3 sf)

b)

A multiplier of 1.248 is a rise of 24.8%

On average, the number of members went up by about 25% each year

a) 1.25 b) average rise of about 25% a year

7 STANDARDISED DEATH RATES

The table shows the population and the number of deaths in one year for the town of Harwick, split into three age groups. It also shows the standard population.

Age group (years)	Population	Deaths	Standard population
0 to 39	12 000	18	500
40 to 69	8000	56	350
70 and over	5000	240	150
Total	25 000	314	1000

a. Calculate the standardised death rate for Harwick.

b. Explain why the standardised death rate, rather than the crude death rate, should be used to compare Harwick with another town.

(5 marks)

a) crude rate = $\frac{\text{deaths}}{\text{population}}$ x 1000 for each age group

$\frac{18}{12000} \times 1000 = 1.5$

$\frac{56}{8000} \times 1000 = 7$

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

$\frac{240}{5000} \times 1000 = 48$

standardised rate = $\frac{\text{crude rate}}{1000} \times$ standard population

$\frac{1.5}{1000} \times 500 = 0.75$

$\frac{7}{1000} \times 350 = 2.45$

$\frac{48}{1000} \times 150 = 7.2$

Add the three standardised rates

$0.75 + 2.45 + 7.2 = 10.4$

10.4 deaths per 1000

b)

The two towns may have different age structures

A town with more older people has more deaths

The standardised rate allows for this

a) 10.4 per 1000 b) allows for different age structures

8 **ESTIMATING THE MEDIAN BY LINEAR INTERPOLATION**

The table shows the times, t minutes, that 60 customers spent in a supermarket.

Time (t minutes)	Frequency
$0 < t \leq 10$	6
$10 < t \leq 20$	14
$20 < t \leq 30$	22
$30 < t \leq 40$	12
$40 < t \leq 60$	6

a. Use linear interpolation to estimate the median time. Give your answer correct to 1 decimal place.

b. Write down one assumption you made in part (a).

(4 marks)

a) Find where the median is

$60 \div 2 = 30$

The median is the 30th value

Cumulative frequencies

6, 20, 42, 54, 60

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

The 30th value is in the class $20 < t \leq 30$

It is 30 - 20 = 10 values into a class of 22

$$\text{median} = 20 + \frac{10}{22} \times 10$$

$$\text{median} = 20 + 4.545...$$

$$\text{median} = 24.5 \text{ minutes (1 dp)}$$

b)

The times are evenly spread

through the class $20 < t \leq 30$

a) 24.5 minutes b) times evenly spread in the class

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.