



Averages and Spread Sheet 2

Q1.

Helen is studying the daily mean wind speed for Camborne using the large data set from 1987. The data for one month are summarised in Table 1 below.

Windspeed	n/a	6	7	8	9	11	12	13	14	16
Frequency	13	2	3	2	2	3	1	2	1	2

Table 1

- (a) Calculate the mean for these data. (1)
- (b) Calculate the standard deviation for these data and state the units. (2)

The means and standard deviations of the daily mean wind speed for the other months from the large data set for Camborne in 1987 are given in Table 2 below. The data are not in month order.

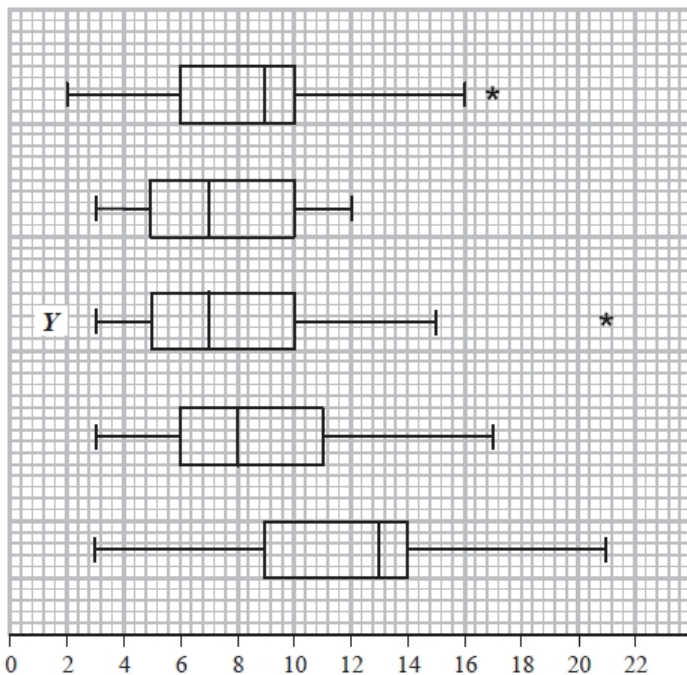
Month	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>
Mean	7.58	8.26	8.57	8.57	11.57
Standard Deviation	2.93	3.89	3.46	3.87	4.64

Table 2

- (c) Using your knowledge of the large data set, suggest, giving a reason, which month had a mean of 11.57 (2)

The data for these months are summarised in the box plots on the opposite page. They are not in month order or the same order as in Table 2.

- (d) (i) State the meaning of the * symbol on some of the box plots. (3)
- (ii) Suggest, giving your reasons, which of the months in Table 2 is most likely to be summarised in the box plot marked *Y*.



(Total for question = 8 marks)



Q2.

Sara is investigating the variation in daily maximum gust, t kn, for Camborne in June and July 1987.

She used the large data set to select a sample of size 20 from the June and July data for 1987. Sara selected the first value using a random number from 1 to 4 and then selected every third value after that.

(a) State the sampling technique Sara used.

(1)

(b) From your knowledge of the large data set, explain why this process may not generate a sample of size 20.

(1)

The data Sara collected are summarised as follows

$$n = 20 \quad \sum t = 374 \quad \sum t^2 = 7600$$

(c) Calculate the standard deviation.

(2)

(Total for question = 4 marks)



Q3.

A lake contains three different types of carp.

There are an estimated 450 mirror carp, 300 leather carp and 850 common carp.

Tim wishes to investigate the health of the fish in the lake.

He decides to take a sample of 160 fish.

(a) Give a reason why stratified random sampling cannot be used.

(1)

(b) Explain how a sample of size 160 could be taken to ensure that the estimated populations of each type of carp are fairly represented.

You should state the name of the sampling method used.

(2)

As part of the health check, Tim weighed the fish.

His results are given in the table below.

Weight (w kg)	Frequency (f)	Midpoint (m kg)
$2 \leq w < 3.5$	8	2.75
$3.5 \leq w < 4$	32	3.75
$4 \leq w < 4.5$	64	4.25
$4.5 \leq w < 5$	40	4.75
$5 \leq w < 6$	16	5.5

(You may use $\sum fm = 692$ and $\sum fm^2 = 3053$)

(c) Calculate an estimate for the standard deviation of the weight of the carp.

(2)

Tim realised that he had transposed the figures for 2 of the weights of the fish.

He had recorded in the table 2.3 instead of 3.2 and 4.6 instead of 6.4.

(d) Without calculating a new estimate for the standard deviation, state what effect

(i) using the correct figure of 3.2 instead of 2.3

(ii) using the correct figure of 6.4 instead of 4.6

would have on your estimated standard deviation.

Give a reason for each of your answers.

(2)

(Total for question = 7 marks)



Q4.

Joshua is investigating the daily total rainfall in Hurn for May to October 2015

Using the information from the large data set, Joshua wishes to calculate the mean of the daily total rainfall in Hurn for May to October 2015

(a) Using your knowledge of the large data set, explain why Joshua needs to clean the data before calculating the mean.

(1)

Using the information from the large data set, he produces the grouped frequency table below.

Daily total rainfall (r mm)	Frequency	Midpoint (x mm)
$0 \leq r < 0.5$	121	0.25
$0.5 \leq r < 1.0$	10	0.75
$1.0 \leq r < 5.0$	24	3.0
$5.0 \leq r < 10.0$	12	7.5
$10.0 \leq r < 30.0$	17	20.0

You may use $\sum fx = 539.75$ and $\sum fx^2 = 7704.1875$

(b) Use linear interpolation to calculate an estimate for the upper quartile of the daily total rainfall.

(2)

(c) Calculate an estimate for the standard deviation of the daily total rainfall in Hurn for May to October 2015

(2)

(d) (i) State the assumption involved with using class midpoints to calculate an estimate of a mean from a grouped frequency table.

(ii) Using your knowledge of the large data set, explain why this assumption does not hold in this case.

(iii) State, giving a reason, whether you would expect the actual mean daily total rainfall in Hurn for May to October 2015 to be larger than, smaller than or the same as an estimate based on the grouped frequency table.

(3)

(Total for question = 8 marks)



Q5.

Stav is studying the large data set for September 2015

He codes the variable Daily Mean Pressure, x , using the formula $y = x - 1010$

The data for all 30 days from Hurn are summarised by

$$\sum y = 214 \quad \sum y^2 = 5912$$

- (a) State the units of the variable x (1)
- (b) Find the mean Daily Mean Pressure for these 30 days. (2)
- (c) Find the standard deviation of Daily Mean Pressure for these 30 days. (3)

Stav knows that, in the UK, winds circulate

- in a **clockwise** direction around a region of **high** pressure
- in an **anticlockwise** direction around a region of **low** pressure

The table gives the Daily Mean Pressure for 3 locations from the large data set on 26/09/2015

Location	Heathrow	Hurn	Leuchars
Daily Mean Pressure	1029	1028	1028
Cardinal Wind Direction			

The Cardinal Wind Directions for these 3 locations on 26/09/2015 were, in random order,

W NE E

You may assume that these 3 locations were under a single region of pressure.

- (d) Using your knowledge of the large data set, place each of these Cardinal Wind Directions in the correct location in the table.
Give a reason for your answer.

(2)

(Total for question = 8 marks)



Q6.

The following table summarises the times, t minutes to the nearest minute, recorded for a group of students to complete an exam.

Time (minutes) t	11 – 20	21 – 25	26 – 30	31 – 35	36 – 45	46 – 60
Number of students f	62	88	16	13	11	10

[You may use $\sum ft^2 = 134281.25$]

- (a) Estimate the mean and standard deviation of these data. (5)
- (b) Use linear interpolation to estimate the value of the median. (2)
- (c) Show that the estimated value of the lower quartile is 18.6 to 3 significant figures. (1)
- (d) Estimate the interquartile range of this distribution. (2)
- (e) Give a reason why the mean and standard deviation are not the most appropriate summary statistics to use with these data. (1)

The person timing the exam made an error and each student actually took 5 minutes less than the times recorded above. The table below summarises the actual times.

Time (minutes) t	6 – 15	16 – 20	21 – 25	26 – 30	31 – 40	41 – 55
Number of students f	62	88	16	13	11	10

- (f) Without further calculations, explain the effect this would have on each of the estimates found in parts (a), (b), (c) and (d). (3)

(Total 14 marks)



Q7.

The mark, x , scored by each student who sat a statistics examination is coded using

$$y = 1.4x - 20$$

The coded marks have mean 60.8 and standard deviation 6.60

Find the mean and the standard deviation of x .

(4)

(Total 4 marks)

Q8.

Keith records the amount of rainfall, in mm, at his school, each day for a week. The results are given below.

2.8 5.6 2.3 9.4 0.0 0.5 1.8

Jenny then records the amount of rainfall, x mm, at the school each day for the following 21 days. The results for the 21 days are summarised below.

$$\sum x = 84.6$$

(a) Calculate the mean amount of rainfall during the whole 28 days.

(2)

Keith realises that he has transposed two of his figures. The number 9.4 should have been 4.9 and the number 0.5 should have been 5.0

Keith corrects these figures.

(b) State, giving your reason, the effect this will have on the mean.

(2)

(Total 4 marks)

Q9.

Dian uses the large data set to investigate the Daily Total Rainfall, r mm, for Camborne.

(a) Write down how a value of $0 < r \leq 0.05$ is recorded in the large data set.

(1)

Dian uses the data for the 31 days of August 2015 for Camborne and calculates the following statistics

$$n = 31 \quad \sum r = 174.9 \quad \sum r^2 = 3523.283$$

(b) Use these statistics to calculate

- (i) the mean of the Daily Total Rainfall in Camborne for August 2015,
- (ii) the standard deviation of the Daily Total Rainfall in Camborne for August 2015.

(3)

Dian believes that the mean Daily Total Rainfall in August is less in the South of the UK than in the North of the UK. The mean Daily Total Rainfall in Leuchars for August 2015 is 1.72 mm to 2 decimal places.

(c) State, giving a reason, whether this provides evidence to support Dian's belief.

(2)

Dian uses the large data set to estimate the proportion of days with no rain in Camborne for 1987 to be 0.27 to 2 decimal places.

(d) Explain why the distribution $B(14, 0.27)$ might **not** be a reasonable model for the number of days without rain for a 14-day summer event.

(1)

Subscribe To The Ultimate Study Tool For A-Level Maths At ALEvelMathsRevision.com/UST

Need more help with this topic? Go to [ALevelMathsRevision.com](https://www.ALevelMathsRevision.com) for video tutorials and more!



(Total for question = 7 marks)

**Subscribe To The Ultimate Study Tool For A-Level
Maths At [ALevelMathsRevision.com/UST](https://www.ALevelMathsRevision.com/UST)**