The large data set

96 marks

Hints online, answers free with an account: mathswariz.uk/worksheets/a-level-statistics-large-data-set/

Date

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

A Do Now

Quick questions to start. Some use GCSE skills or an earlier lesson; some set slips that the examiner reports say cost marks.

A1 Needed today [3 marks]

Find the median and the mean of these 8 values:
$0,\ 2.4,\ 0,\ 5.1,\ 0.6,\ 0,\ 8.3,\ 1.2$

Answer .....

A2 Needed today [2 marks]

State the type of data, quantitative or qualitative, discrete or continuous, for (i) daily total rainfall in mm, (ii) daily mean cloud cover in oktas, a whole number from 0 to 8.

Answer .....

A3 Common slip [2 marks]

In the large data set a value of daily total rainfall is written "tr", and another value is written "n/a". State what each means.

Answer .....

A4 Common slip [2 marks]

On one day the daily mean wind direction is recorded as W, a bearing of 270°. Which way is the wind blowing?

Answer .....

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

PART 1 The stations, the variables and their units

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

EXAMPLE 1 · I DO

The large data set contains weather readings from eight stations.
(a) Name the five UK stations and the three overseas stations.
(b) Which station is in the southern hemisphere? Which UK station is furthest north?
(c) Which months and years does the large data set cover?
(d) Ella wants to use the large data set to estimate the mean daily temperature at Heathrow for the whole of 2015. Explain why she cannot.

Hint Learn the stations and the months. The data are for May to October only, so they cannot describe a whole year. In Perth the seasons are the other way round: May to October is autumn and winter there.

YOUR TURN · WE DO

B1 Answer these about the large data set.

(a) Name the overseas station in the southern hemisphere. Camborne is the UK station furthest south: in which part of the UK is it? [2 marks]

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

(b) Jo wants to compare temperatures at Perth and at Heathrow in July 2015. She says: "July is summer, so the two should be similar." Explain why she is wrong. [2 marks]

(c) Explain why the large data set cannot be used to estimate the mean daily rainfall at Camborne for a whole year. [2 marks]

Total for B1: 6 marks

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

EXAMPLE 2 · I DO

Beaufort scale: Calm below 1 knot; Light 1 to 10 knots; Moderate 11 to 16 knots; Fresh 17 to 21 knots. 1 knot $= 1.15$ mph.

(a) A daily mean windspeed of 14 knots is recorded. Give its Beaufort description and its speed in mph.

(b) The daily maximum relative humidity is 98%. What does this suggest about the weather that day?

(c) A daily mean cloud cover of 9 oktas is recorded. Explain why this must be an error.

(d) Explain why daily total sunshine cannot be compared for Beijing and Hurn using the large data set.

Hint The overseas stations record only temperature, rainfall, pressure and windspeed. Humidity above 95% means mist or fog. The Beaufort description is a word, so it is qualitative.

YOUR TURN · WE DO

B2 Beaufort scale: Calm below 1 knot; Light 1 to 10 knots; Moderate 11 to 16 knots; Fresh 17 to 21 knots. 1 knot $= 1.15$ mph.

(a) A daily mean windspeed of 8 knots is recorded. Give its Beaufort description and its speed in mph. [3 marks]

Answer .....

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

(b) A daily maximum relative humidity of 97% is recorded. What does this suggest, and why can daily total sunshine at Jacksonville not be found from the large data set? [2 marks]

PRACTICE · YOU DO

C1 A student copies some values from the large data set. Each value below is impossible. Explain why. [4 marks]

(i) Daily mean cloud cover: 10 oktas.

(ii) Daily maximum relative humidity: 104%.

(iii) Daily total rainfall: -0.4 mm.

(iv) Daily mean wind direction: a bearing of 410°.

Total for B2: 5 marks

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

C2 Beaufort scale: Calm below 1 knot; Light 1 to 10 knots; Moderate 11 to 16 knots; Fresh 17 to 21 knots. 1 knot $= 1.15$ mph.
Tom says: "In the large data set, the Beaufort description of the windspeed is discrete data." Explain why Tom is wrong. Then give the Beaufort description for daily mean windspeeds of 0.5 knots, 10 knots and 17 knots. [4 marks]

C3 The mean daily mean temperatures at Perth in two months of 2015 were 18.0 °C and 13.1 °C. The values are for this question only.
One of the months is May and the other is July. Kim says: "13.1 °C must be May, because May is in spring." Explain why Kim is wrong, and say which month had which temperature. [4 marks]

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

PART 2 Missing values in the large data set

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

EXAMPLE 3 · I DO

The table is set out like an extract from the large data set for Leeming. The values are for this question only.

(a) Find the mean daily total rainfall, taking each tr as 0.

(b) Find the median daily total rainfall.

(c) Find the mean and the range of the daily maximum gust.

Day	Rainfall (mm)	Max gust (kn)
1	3.2	24
2	tr	31
3	0	n/a
4	6.8	18
5	tr	27
6	0	n/a
7	1.4	22
8	12.6	35
9	tr	20
10	0.2	29

Hint A tr value is a number between 0 and 0.05: keep the day and use 0 (or a value up to 0.05). An n/a value is missing: leave the day out and divide by the number of values you have.

YOUR TURN · WE DO

B3 The table is set out like an extract from the large data set for Camborne. The values are for this question only.

Day	Rainfall (mm)	Sunshine (h)
1	0.4	6.2
2	tr	9.8
3	5.6	n/a
4	0	1.3
5	tr	11.4
6	2.2	n/a
7	9.8	0.5
8	0	7.4

(a) Find the mean daily total rainfall, taking each tr as 0. [2 marks]

Answer .....

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

(b) Find the median daily total rainfall. [2 marks]

Answer .....

(c) Find the mean daily total sunshine. [2 marks]

Answer .....

Total for B3: 6 marks

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

C4 For the 31 days of one month at a UK station, 6 days of daily total rainfall are recorded as tr. The other 25 days have a total of 62.0 mm. Ben deletes the tr days and says: "The mean daily rainfall is $\frac{62.0}{25} = 2.48$ mm." Explain Ben's mistake. Find the mean taking tr as 0, and the largest the mean could be. [5 marks]

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

C5 Beaufort scale: Calm below 1 knot; Light 1 to 10 knots; Moderate 11 to 16 knots; Fresh 17 to 21 knots. 1 knot $= 1.15$ mph. [6 marks]
The daily mean windspeeds, in knots, at a UK station on 10 days were

$$6,\ 12,\ 9,\ 17,\ 4,\ 11,\ 8,\ 21,\ 13,\ 9$$

The values are for this question only. Find the mean daily mean windspeed in mph, and the number of days with each Beaufort description.

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

PART 3 Sampling from the large data set

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

EXAMPLE 4 · I DO

Hassan wants a systematic sample of 20 days of the daily mean temperature at Hurn in 2015.
(a) Describe how he could take it.
(b) One of the days he picks has the temperature recorded as n/a. What should he do?

Hint Number the 184 days from 1 May. Interval $k = 184 \div n$, rounded down if needed; a random start from 1 to k. An n/a value cannot be used: leave it out, or replace it with another randomly chosen day.

YOUR TURN · WE DO

B4 Lucy wants a systematic sample of 23 days of the daily mean windspeed at Leuchars in 1987.

(a) Describe how she could take it. [3 marks]

(b) Two of the days in her sample have the windspeed recorded as n/a. Explain what she should do, and the effect on her sample. [2 marks]

Total for B4: 5 marks

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

EXAMPLE 5 · I DO

In one month, the mean daily mean windspeeds at three UK stations were: Leuchars 11.2 knots, Leeming 8.4 knots and Heathrow 7.9 knots. The values are for this question only.

Use these figures and the locations of the stations to compare the windspeeds.

Hint A comparison needs three things: the figures, a reason from where the stations are (coast or inland, north or south, hemisphere), and a conclusion.

YOUR TURN · WE DO

B5 In one month, the mean daily total rainfall was 2.9 mm at Camborne and 1.6 mm at Heathrow. The values are for this question only.

(a) Compare the rainfall at the two stations, giving a reason from their locations. [2 marks]

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

(b) Kai says this shows that it rains more at Camborne than at Heathrow all year round. Give two reasons why this conclusion is not justified. [2 marks]

PRACTICE · YOU DO

C6 Sam wants to compare the daily total sunshine at Beijing and at Hurn using the large data set. [4 marks]

(i) Explain why he cannot.

(ii) He then takes a simple random sample of 20 days of daily total sunshine at Hurn in 2015, and 3 of the values are n/a. Describe what he should do.

Total for B5: 4 marks

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

C7 Mia takes two samples of the daily maximum gust at Hurn in 1987. The values are for this question only. Her sample of 28 days from May has mean 24.0 knots. Her sample of 31 days from October has mean 27.5 knots.

(a) Find the mean of all 59 values. [3 marks]

Answer .....

(b) Explain why she cannot compare these gusts with gusts at Perth using the large data set. [2 marks]

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

(c) Mia says 25.8 knots is a good estimate of the mean daily maximum gust at Hurn for May to October 1987. Give three reasons why it may not be. [3 marks]

Total for C7: 8 marks

D Exam-style questions

Edexcel exam style

D1 The table is set out like an extract from the large data set for Heathrow for 12 days. The values are for this question only.

(a) Explain what tr and n/a mean, and state how many values of rainfall can be used in calculations. [2 marks]

Day	Rainfall (mm)	Max humidity (%)
1	0	88
2	tr	92
3	4.6	98
4	12.2	99
5	0.8	94
6	tr	90
7	0	85
8	0	87
9	3.4	96
10	n/a	n/a
11	1.0	93
12	7.0	97

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

(b) Find the mean daily total rainfall, taking each tr as 0. [3 marks]

Answer .....

(c) Find the median daily total rainfall. [2 marks]

Answer .....

(d) On how many of the days was it likely to be misty or foggy? Comment on the rainfall on those days. [2 marks]

Total for D1: 9 marks

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

D2

> Lewis says all three of his statements are correct.

Lewis writes three statements about the large data set.
(1) "Daily total sunshine in Beijing in 2015 was higher than at Heathrow."
(2) "I deleted the tr values before working out the mean rainfall, so my mean is accurate."
(3) "A daily mean windspeed of 18 knots is Moderate on the Beaufort scale."
Beaufort scale: Calm below 1 knot; Light 1 to 10 knots; Moderate 11 to 16 knots; Fresh 17 to 21 knots. 1 knot $= 1.15$ mph.

(a) Explain why statement (1) cannot be checked using the large data set. [2 marks]

(b) Explain the error in statement (2), and the effect it has on his mean. [2 marks]

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

(c) Correct statement (3). [2 marks]

(d) Convert 18 knots to mph. [2 marks]

Answer .....

Total for D2: 8 marks

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

E Extension

Stretch

No route is given. The hints are at the end of the sheet. Use one at a time.

E1 The 184 days from 1 May to 31 October are numbered 1 (1 May) to 184 (31 October). A systematic sample of the daily mean windspeed at one station is taken with interval 7: a start s is chosen at random from 1 to 7, then every 7th day is taken.

(a) Find the number of days in the sample for each value of s. [3 marks]

Answer .....

(b) The windspeeds for 9 June to 14 June are n/a. Show that these are days 40 to 45. [2 marks]

Answer .....

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

(c) Find the probability that the sample contains at least 26 windspeeds that can be used. [4 marks]

Answer .....

Total for E1: 9 marks

Extension hints

Cover these with paper. Uncover one at a time.

E1

Hint 1 Write 184 as a multiple of 7 plus a remainder. Which starts get an extra day at the end?

Hint 2 Count the days in May to find the number of 1 June, then of 9 June.

Hint 3 A day d is in the sample with start s when d and s leave the same remainder on division by 7. Find which start meets which n/a day.

How did it go?

Colour one circle on each line: not yet, nearly, yes.

I can name the stations and months in the large data set, and say which variables are recorded where.	○ ○ ○
I can work with tr and n/a in an extract, and convert a windspeed to a Beaufort description and to mph.	○ ○ ○
I can take a sample from the large data set and say what to do about missing values.	○ ○ ○
I can compare two stations using figures, their locations and a conclusion.	○ ○ ○

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.