1. 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

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

(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	NE	E	W (1)

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

W NE E

entire pressure system

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)

1 hPa = 100 Pa , 1 Pa = 1 Nm^{-2}

a) hPa (1)

b) $x = y + 1010$

$\therefore \bar{x} = \bar{y} + 1010$ (1) ← $\bar{x}$ = mean value of x

$= \frac{214}{30} + 1010$

$= 1017.1333 = 1017$ hPa (1)

Question 1 continued.

c) $\sigma_x = \sigma_y$ standard deviation is not affected by this type of coding.

$$= \sqrt{\frac{\Sigma y^2}{n} - \bar{y}^2} \quad (1)$$

$$= \sqrt{\frac{5912}{30} - \left(\frac{214}{30}\right)^2}$$

$$= \sqrt{146.1822} \quad (1)$$

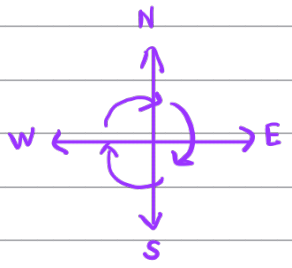
$$= 12.0905\ldots = 12.1 \text{ (3 s.f.)} \quad (1)$$

d) mean + standard deviation $\approx$ 1017 + 12.1

$$= 1029.1 \text{ hPa}$$

$\therefore$ entire pressure system is high pressure

so, wind moves in clockwise direction (1)



Locations from North to South: Leuchars, Heathrow, Hurn

- Heathrow (NE)
- Hurn (E)
- Leuchars (W)

- Leuchars (W)
- Heathrow (NE)
- Hurn (E)

(Total for Question 1 is 8 marks)

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

(a) Write down how a value of $0 < r \leqslant 0.05$ 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

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

(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)

a) r (1)

b) (i) $\text{mean} = \frac{\sum r}{n} = \frac{174.9}{31} = 5.64$ (3sf) (1)

(ii) $\text{sd} = \sqrt{\frac{\sum r^2}{n} - \left(\frac{\sum r}{n}\right)^2} = \sqrt{\frac{3523.283}{31} - \left(\frac{174.9}{31}\right)^2}$ (1)

$= 9.05$ (3sf) (1)

c) Leuchars is further North than Camborne, but the mean Daily Total Rainfall is smaller for Leuchars than for Camborne. Therefore there is no evidence that Dian's belief is true. (1) (1)

d) $p = 0.27$ is unlikely to be constant (1)

It's unlikely that there is a 27% chance of no rain every day for 14 days. A condition on the binomial distribution is that p must be constant.

3. Ben is studying the Daily Total Rainfall, x mm, in Leeming for 1987

He used all the data from the large data set and summarised the information in the following table.

x	0	0.1–0.5	0.6–1.0	1.1–1.9	2.0–4.0	4.1–6.9	7.0–12.0	12.1–20.9	21.0–32.0	tr
Frequency	55	18	18	21	17	9	9	6	2	29

(a) Explain how the data will need to be cleaned before Ben can start to calculate statistics such as the mean and standard deviation.

(2)

Using all 184 of these values, Ben estimates $\sum x = 390$ and $\sum x^2 = 4336$

(b) Calculate estimates for

$n = 184$

(i) the mean Daily Total Rainfall,

(ii) the standard deviation of the Daily Total Rainfall.

(3)

Ben suggests using the statistic calculated in part (b)(i) to estimate the annual mean Daily Total Rainfall in Leeming for 1987

(c) Using your knowledge of the large data set,

(i) give a reason why these data would not be suitable,

(ii) state, giving a reason, how you would expect the estimate in part (b)(i) to differ from the actual annual mean Daily Total Rainfall in Leeming for 1987

(2)

a) Replace 'tr' with a numerical value between 0 and 0.05. (1)

For example, 0.025. (1)

if the total amount of rainfall recorded is less than 0.05 mm, then it is recorded as 'tr'.

b) (i) mean, $\bar{x} = \frac{\sum x}{n}$

$= \frac{390}{184}$

$= 2.119\ldots = 2.12$ (3 s.f.) (1)

(ii) Standard deviation $= \sqrt{\frac{\Sigma x^2}{n} - \bar{x}^2}$

$\bar{x}^2$

$= \sqrt{\frac{4336}{184} - 2.119..^2}$ ✓ (1)

$= 4.367... = 4.37$ (3 s.f.) (1)

c) (i) The data only covers May to October. So, it is not a representative of the whole year. (1)

(ii) Winter months are missing when we'd expect more rain during this season. So, estimation in b(i) is expected to be an underestimate. (1)

4. Ming is studying the large data set for Perth in 2015

He intended to use all the data available to find summary statistics for the Daily Mean Air Temperature, x °C.
Unfortunately, Ming selected an incorrect variable on the spreadsheet.
This incorrect variable gave a mean of 5.3 and a standard deviation of 12.4

(a) Using your knowledge of the large data set, suggest which variable Ming selected.

(1)

The correct values for the Daily Mean Air Temperature are summarised as

$n = 184$ $\quad \sum x = 2801.2$ $\quad \sum x^2 = 44\,695.4$

(b) Calculate the mean and standard deviation for these data.

(3)

One of the months from the large data set for Perth in 2015 has

- mean $\bar{x} = 19.4$
- standard deviation $\sigma_x = 2.83$

for Daily Mean Air Temperature.

(c) Suggest, giving a reason, a month these data may have come from.

(2)

a) Rainfall ①
OR daily mean windspeed

there are only 5 measured variables for Perth:
- D.M. air temperature
- Rainfall
- D.M. pressure (always around 1000 hPa)
- D.M. windspeed
- D.M. windspeed Beaufort conversion.

b) $\bar{x} = \frac{\Sigma x}{n} = \frac{2801.2}{184} = 15.2239\ldots$
$= 15.2$ (3sf) ①

$\sigma_x = \sqrt{\frac{\Sigma x^2}{n} - \left[\frac{\Sigma x}{n}\right]^2} = \sqrt{\frac{44695.4}{184} - (15.22\ldots)^2}$ ①
$= 3.34$ (3sf) ①

c) $\bar{x} = 19.4$ which indicates a summer / spring month (1)

Perth is in the southern hemisphere so summer months start in October. (1)