



Luke finds the diagram above, showing a box plot with no axis.  
The diagram lies next to a set of data labelled ‘Large data set: Daily Total Rainfall – Hurn 1987’.

Luke thinks that the diagram should be rotated through 180° for it to represent these data.

(a) Using your knowledge of the large data set, explain why Luke is correct. (1)

A sample of data for the Daily Mean Visibility in Leeming in 1987 was taken and is summarised in the table below

Lower Quartile (dam)	Upper Quartile (dam)
1000	$x$

An outlier is defined as any value  
more than  $1.5 \times$  the interquartile range **above** the upper quartile  
or more than  $1.5 \times$  the interquartile range **below** the lower quartile

The last 6 values of the ordered data are 3300, 3400, 3400, 3600, 3600, 3700

Luke believes that there are 5 outliers at the upper end of the data.

(b) Find the range of possible values for  $x$  if Luke is correct. (4)

The first 6 values of the ordered data are 200, 200, 300, 400, 400, 600

Nish believes that there are 5 outliers at the lower end of the data, which gives the range of possible values for  $x$  as  $1267 \leq x < 1400$

Given that there are only 5 outliers **in total** in the raw data

(c) explain why Nish’s belief is **not** correct. (2)

Luke wants to investigate the Daily Mean Windspeed (Beaufort conversion) from the large data set for Leeming in 2015.

He plans to select data until he has 25 pieces of data from **each** of the three categories of windspeed (light, moderate and fresh).

(d) State the sampling technique Luke plans to use. (1)

(e) Using your knowledge of the large data set, explain clearly why Luke’s sampling technique may **not** be successful in obtaining his sample. (1)