

2.

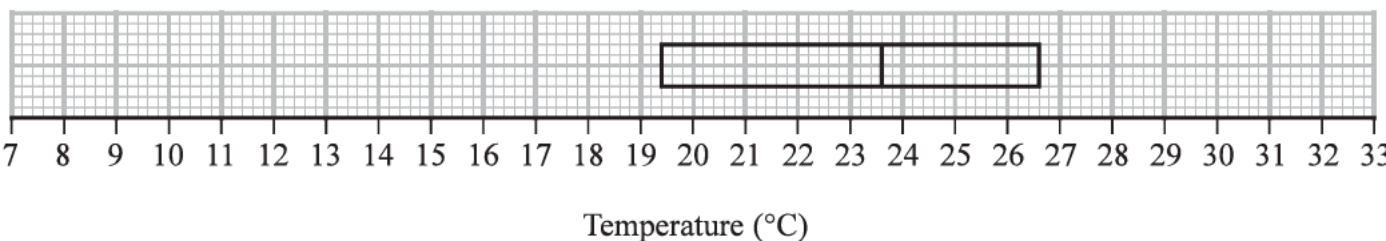


Figure 1

The partially completed box plot in Figure 1 shows the distribution of daily mean air temperatures using the data from the large data set for Beijing in 2015

An outlier is defined as a value
 more than $1.5 \times \text{IQR}$ below Q_1 or
 more than $1.5 \times \text{IQR}$ above Q_3

The three lowest air temperatures in the data set are 7.6°C , 8.1°C and 9.1°C
 The highest air temperature in the data set is 32.5°C

- (a) Complete the box plot in Figure 1 showing clearly any outliers. (4)
- (b) Using your knowledge of the large data set, suggest from which month the two outliers are likely to have come. (1)

Using the data from the large data set, Simon produced the following summary statistics for the daily mean air temperature, $x^\circ\text{C}$, for Beijing in 2015

$$n = 184 \quad \sum x = 4153.6 \quad S_{xx} = 4952.906$$

- (c) Show that, to 3 significant figures, the standard deviation is 5.19°C (1)

Simon decides to model the air temperatures with the random variable

$$T \sim N(22.6, 5.19^2)$$

- (d) Using Simon's model, calculate the 10th to 90th interpercentile range. (3)

Simon wants to model another variable from the large data set for Beijing using a normal distribution.

- (e) State two variables from the large data set for Beijing that are **not** suitable to be modelled by a normal distribution. Give a reason for each answer. (2)