

MTMG37 Example Solution to Class Exercise 1

Two histograms of daily mean temperatures, recorded in degrees Celsius ($^{\circ}\text{C}$) at the Department field site during 2003, are shown in Figures 1 and 2. Figure 1 suggests that the temperatures have a bi-modal distribution, with two peaks around 9 and 16 $^{\circ}\text{C}$. This feature disappears when the interval width is increased in Figure 2 and is probably just a result of sampling variation. The way in which data is plotted can strongly influence its interpretation, so be careful!

The histograms show that the distribution of daily mean temperatures is roughly symmetric and there are no obvious outliers. The boxplot provides no additional insight for these data and so is omitted. Summary statistics are displayed in Table 1. Note the number of significant figures: since the quartiles and ranges correspond to particular observations they should have the same precision (one-tenth of a degree) as the data; the mean and standard deviation are given two decimal places but any more is unnecessary. For symmetrically distributed data, the mean and standard deviation provide a reasonable summary. For these data, the mean temperature is 10.96 $^{\circ}\text{C}$, with quite a wide spread given by the standard deviation of 5.84 $^{\circ}\text{C}$.

Mean	Std Dev	1 st Quartile	Median	3 rd Quartile	IQR	Range
10.96	5.84	6.8	10.4	15.6	5.2	26.8

Table 1. Statistics of field-site daily mean temperatures ($^{\circ}\text{C}$) in 2003.

There are six missing values in the data record for 2003 and it is important to discover why this is so. For example, if missing values are recorded only when temperatures are very high (due perhaps to a fault in the thermometers) then our previous analysis of the data will be biased. Inspection of the data reveals that the missing values occur in three pairs that correspond to weekends, apparently randomly located through the year. It is unlikely, therefore, that the missing values are linked to the meteorological conditions, but are probably caused by technical faults that were not rectified until after the weekend.

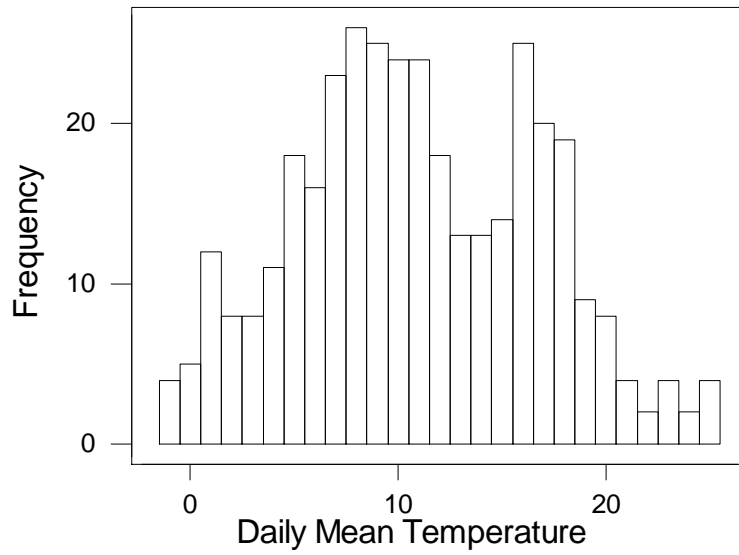


Figure 1. Histogram of field-site daily mean temperatures (°C) recorded in 2003.

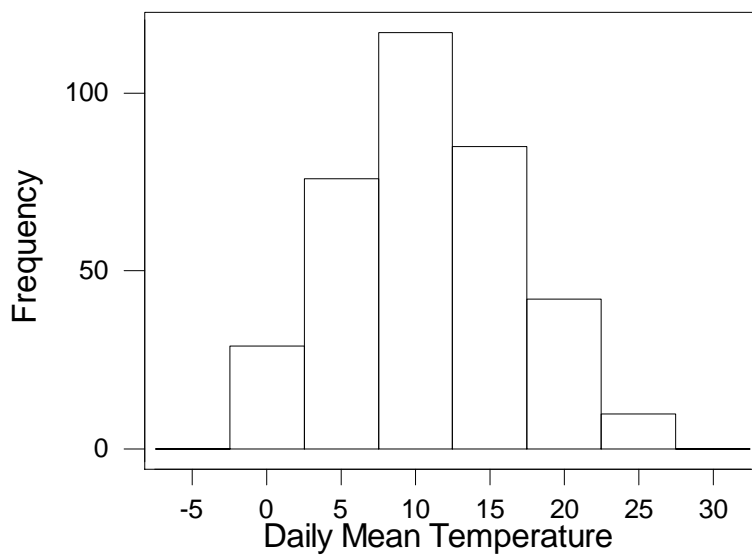


Figure 2. Histogram of field-site daily mean temperatures (°C) recorded in 2003.

This document was created with Win2PDF available at <http://www.daneprairie.com>.
The unregistered version of Win2PDF is for evaluation or non-commercial use only.