

KEY TERMS

A **variable** describes some characteristic of interest that can vary in value. Some variables are **categorical** (soldiers' gender – male or female). Others are **quantitative** (soldiers' head circumference or foot length).

The **distribution** of a variable describes the possible values the variable takes and how often it takes these values. Stemplots are one way to graph the distribution of a quantitative variable.

Shape, center, and spread describe the overall pattern of the distribution of a quantitative variable. Some distributions have simple shapes, such as **unimodal** (single peak) or **symmetric** (one side is the mirror image of the other).

Outliers are data values that lie outside the overall pattern of the distribution. Always look for gaps in the data and outliers and try to explain them.

A **stemplot** (or **stem-and-leaf plot**) is a useful tool for conveying the shape of relatively small datasets and identifying outliers. It consists of two columns, one for the stems and the other for the leaves (often separated by a vertical line).