3

Elementary Descriptive Statistics

Engineering data are always variable. Given precise enough measurement, even supposedly constant process conditions produce differing responses. Therefore, it is not individual data values that demand an engineer's attention as much as the pattern or **distribution** of those responses. The task of summarizing data is to describe their important distributional characteristics. This chapter discusses simple methods that are helpful in this task.

The chapter begins with some elementary graphical and tabular methods of data summarization. The notion of quantiles of a distribution is then introduced and used to make other useful graphical displays. Next, standard numerical summary measures of location and spread for quantitative data are discussed. Finally comes a brief look at some elementary methods for summarizing qualitative and count data.

3.1 Elementary Graphical and Tabular Treatment of Quantitative Data

Almost always, the place to begin in data analysis is to make appropriate graphical and/or tabular displays. Indeed, where only a few samples are involved, a good picture or table can often tell most of the story about the data. This section discusses the usefulness of dot diagrams, stem-and-leaf plots, frequency tables, histograms, scatterplots, and run charts.

3.1.1 Dot Diagrams and Stem-and-Leaf Plots

When an engineering study produces a small or moderate amount of univariate quantitative data, a **dot diagram**, easily made with pencil and paper, is often quite revealing. A dot diagram shows each observation as a dot placed at a position corresponding to its numerical value along a number line.

Example 1
(*Example 1, Chapter 1,*
revisited—p. 2)

Portraying Thrust Face Runouts

Section 1.1 considered a heat treating problem where distortion for gears laid and gears hung was studied. Figure 1.1 has been reproduced here as Figure 3.1. It consists of two dot diagrams, one showing thrust face runout values for gears laid and the other the corresponding values for gears hung, and shows clearly that the laid values are both generally smaller and more consistent than the hung values.
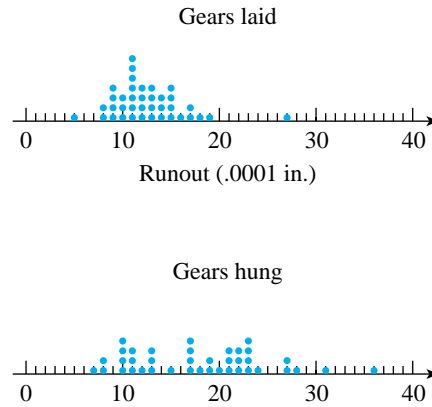
Gears laid

Runout (.0001 in.)

Gears hung

Figure 3.1   Dot diagrams of runouts

Example 2

Portraying Bullet Penetration Depths

Sale and Thom compared penetration depths for several types of .45 caliber bullets fired into oak wood from a distance of 15 feet. Table 3.1 gives the penetration depths (in mm from the target surface to the back of the bullets) for two bullet types. Figure 3.2 presents a corresponding pair of dot diagrams.

Table 3.1
Bullet Penetration Depths (mm)

230 Grain Jacketed Bullets	200 Grain Jacketed Bullets
40.50, 38.35, 56.00, 42.55,	63.80, 64.65, 59.50, 60.70,
38.35, 27.75, 49.85, 43.60,	61.30, 61.50, 59.80, 59.10,
38.75, 51.25, 47.90, 48.15,	62.95, 63.55, 58.65, 71.70,
42.90, 43.85, 37.35, 47.30,	63.30, 62.65, 67.75, 62.30,
41.15, 51.60, 39.75, 41.00	70.40, 64.05, 65.00, 58.00

Example 2
(*continued*)

230 Grain jacketed bullets
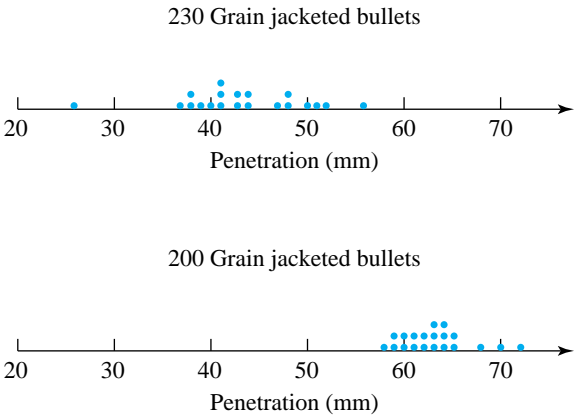
200 Grain jacketed bullets

Figure 3.2   Dot diagrams of penetration depths

The dot diagrams show the penetrations of the 200 grain bullets to be both larger and more consistent than those of the 230 grain bullets. (The students had predicted larger penetrations for the lighter bullets on the basis of greater muzzle velocity and smaller surface area on which friction can act. The different consistencies of penetration were neither expected nor explained.)

Dot diagrams give the general feel of a data set but do not always allow the recovery of exactly the values used to make them. A **stem-and-leaf plot** carries much the same visual information as a dot diagram while preserving the original values exactly. A stem-and-leaf plot is made by using the last few digits of each data point to indicate where it falls.
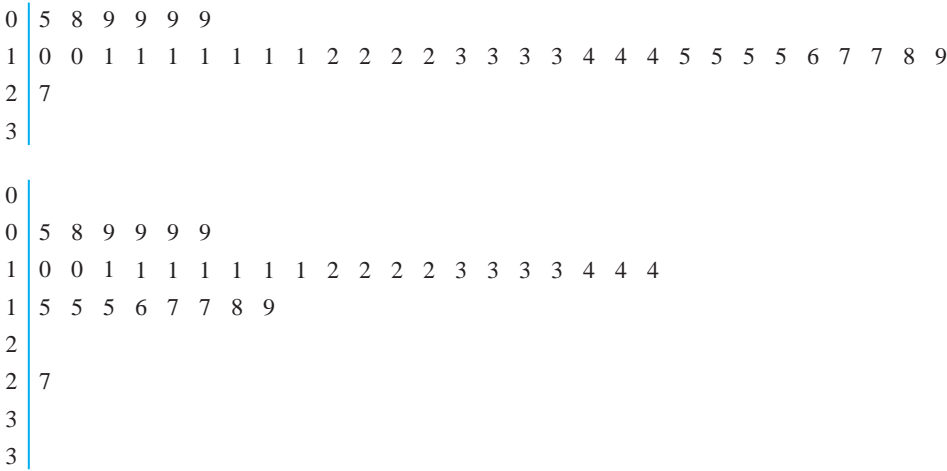
```
0 | 5 8 9 9 9 9
1 | 0 0 1 1 1 1 1 1 1 1 2 2 2 2 3 3 3 3 4 4 4 5 5 5 5 6 7 7 8 9
2 | 7
3 |
```

```
0 |
0 | 5 8 9 9 9 9
1 | 0 0 1 1 1 1 1 1 1 2 2 2 2 3 3 3 3 4 4 4
1 | 5 5 5 6 7 7 8 9
2 |
2 | 7
3 |
3 |
```

Figure 3.3   Stem-and-leaf plots of laid gear runouts (*Example 1*)

Example 1
(*continued*)

Figure 3.3 gives two possible stem-and-leaf plots for the thrust face runouts of laid gears. In both, the first digit of each observation is represented by the number to the left of the vertical line or "stem" of the diagram. The numbers to the right of the vertical line make up the "leaves" and give the second digits of the observed runouts. The second display shows somewhat more detail than the first by providing "0–4" and "5–9" leaf positions for each possible leading digit, instead of only a single "0–9" leaf for each leading digit.

Example 2
(*continued*)

Figure 3.4 gives two possible stem-and-leaf plots for the penetrations of 200 grain bullets in Table 3.1. On these, it was convenient to use two digits to the left of the decimal point to make the stem and the two following the decimal point to create the leaves. The first display was made by recording the leaf values directly from the table (from left to right and top to bottom). The second display is a better one, obtained by ordering the values that make up each leaf. Notice that both plots give essentially the same visual impression as the second dot diagram in Figure 3.2.

58	.65, .00	58	.00, .65
59	.50, .80, .10	59	.10, .50, .80
60	.70	60	.70
61	.30, .50	61	.30, .50
62	.95, .65, .30	62	.30, .65, .95
63	.80, .55, .30	63	.30, .55, .80
64	.65, .05	64	.05, .65
65	.00	65	.00
66		66	
67	.75	67	.75
68		68	
69		69	
70	.40	70	.40
71	.70	71	.70

Figure 3.4   Stem-and-leaf plots of the 200 grain penetration depths

When comparing two data sets, a useful way to use the stem-and-leaf idea is to make two plots **back-to-back**.
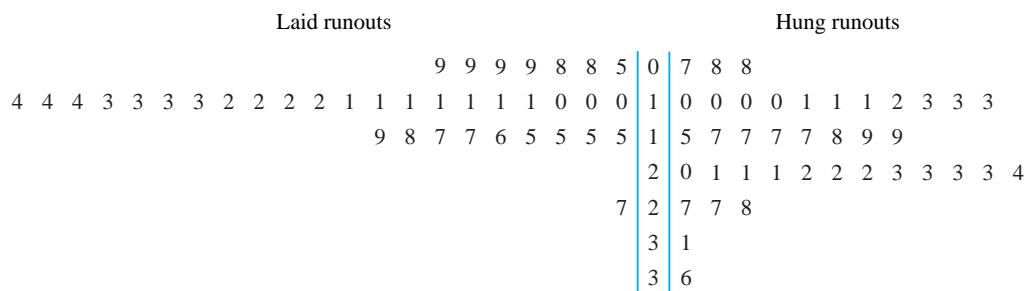
Laid runouts   Hung runouts

	9 9 9 9 8 8 5	0	7 8 8
4 4 4 3 3 3 3 2 2 2 2 1 1 1 1 1 1 1 0 0 0	1	0 0 0 0 1 1 1 2 3 3 3
9 8 7 7 6 5 5 5 5	1	5 7 7 7 7 8 9 9
	2	0 1 1 1 2 2 2 3 3 3 3 4
7	2	7 7 8
	3	1
	3	6

Figure 3.5   Back-to-back stem-and-leaf plots of runouts (*Example 1*)

<table>
<tr><td>Example 1
(<i>continued</i>)</td><td>Figure 3.5 gives back-to-back stem-and-leaf plots for the data of Table 1.1 (pg. 3). It shows clearly the differences in location and spread of the two data sets.</td></tr>
</table>

3.1.2   Frequency Tables and Histograms

Dot diagrams and stem-and-leaf plots are useful devices when mulling over a data set. But they are not commonly used in presentations and reports. In these more formal contexts, frequency tables and histograms are more often used.

A **frequency table** is made by first breaking an interval containing all the data into an appropriate number of smaller intervals of equal length. Then tally marks can be recorded to indicate the number of data points falling into each interval. Finally, frequencies, relative frequencies, and cumulative relative frequencies can be added.

Example 1
(*continued*)

Table 3.2 gives one possible frequency table for the laid gear runouts. The relative frequency values are obtained by dividing the entries in the frequency column

Table 3.2
Frequency Table for Laid Gear Thrust Face Runouts

Runout (.0001 in.)	Tally	Frequency	Relative Frequency	Cumulative Relative Frequency
5–8	|||	3	.079	.079
9–12	⊞ ⊞ ⊞ |||	18	.474	.553
13–16	⊞ ⊞ ||	12	.316	.868
17–20	||||	4	.105	.974
21–24		0	0	.974
25–28			1	.026	1.000
		38	1.000	

by 38, the number of data points. The entries in the cumulative relative frequency column are the ratios of the totals in a given class and all preceding classes to the total number of data points. (Except for round-off, this is the sum of the relative frequencies on the same row and above a given cumulative relative frequency.) The tally column gives the same kind of information about distributional shape that is provided by a dot diagram or a stem-and-leaf plot.

Choosing intervals for a frequency table   The choice of intervals to use in making a frequency table is a matter of judgment. Two people will not necessarily choose the same set of intervals. However, there are a number of simple points to keep in mind when choosing them. First, in order to avoid visual distortion when using the tally column of the table to gain an impression of distributional shape, intervals of equal length should be employed. Also, for aesthetic reasons, round numbers are preferable as interval endpoints. Since there is usually aggregation (and therefore some loss of information) involved in the reduction of raw data to tallies, the larger the number of intervals used, the more detailed the information portrayed by the table. On the other hand, if a frequency table is to have value as a summarization of data, it can't be cluttered with too many intervals.

After making a frequency table, it is common to use the organization provided by the table to create a **histogram**. A (frequency or relative frequency) histogram is a kind of bar chart used to portray the shape of a distribution of data points.

Example 2
(*continued*)   Table 3.3 is a frequency table for the 200 grain bullet penetration depths, and Figure 3.6 is a translation of that table into the form of a histogram.

Table 3.3
Frequency Table for 200 Grain Penetration Depths

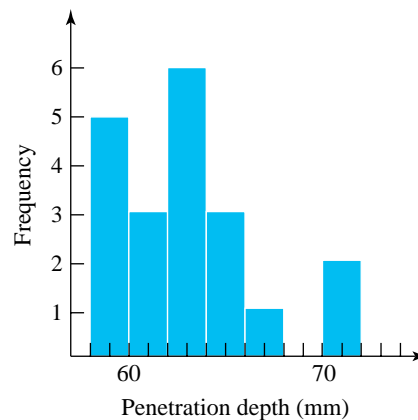| Penetration Depth (mm)	Tally	Frequency	Relative Frequency	Cumulative Relative Frequency
58.00–59.99	ⵕⵕ	5	.25	.25
60.00–61.99	|||	3	.15	.40
62.00–63.99	ⵕⵕ		6	.30	.70
64.00–65.99	|||	3	.15	.85
66.00–67.99			1	.05	.90
68.00–69.99		0	0	.90
70.00–71.99	||	2	.10	1.00
		20	1.00	

Example 2
(*continued*)

Figure 3.6   Histogram of the 200 grain
penetration depths

The vertical scale in Figure 3.6 is a frequency scale, and the histogram is a **frequency histogram**. By changing to relative frequency on the vertical scale, one can produce a **relative frequency histogram**. In making Figure 3.6, care was taken to

Guidelines for making histograms

1. (continue to) use intervals of equal length,
2. show the entire vertical axis beginning at zero,
3. avoid breaking either axis,
4. keep a uniform scale across a given axis, and
5. center bars of appropriate heights at the midpoints of the (penetration depth) intervals.

Following these guidelines results in a display in which *equal enclosed areas correspond to equal numbers of data points*. Further, data point positioning is clearly indicated by bar positioning on the horizontal axis. If these guidelines are not followed, the resulting bar chart will in one way or another fail to faithfully represent its data set.

Figure 3.7 shows terminology for common **distributional shapes** encountered when making and using dot diagrams, stem-and-leaf plots, and histograms.

The graphical and tabular devices discussed to this point are deceptively simple methods. When routinely and intelligently used, they are powerful engineering tools. The information on location, spread, and shape that is portrayed so clearly on a histogram can give strong hints as to the functioning of the physical process that is generating the data. It can also help suggest physical mechanisms at work in the process.

Examples of engineering interpretations of distribution shape

For example, if data on the diameters of machined metal cylinders purchased from a vendor produce a histogram that is decidedly **bimodal** (or **multimodal**, having several clear humps), this suggests that the machining of the parts was done
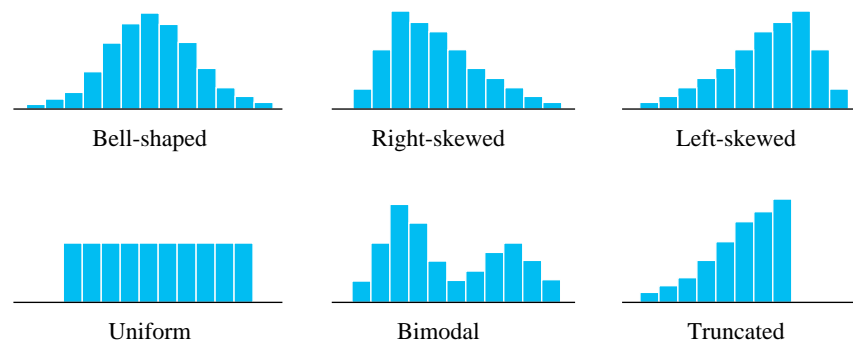
Figure 3.7 Distributional shapes

on more than one machine, or by more than one operator, or at more than one time. The practical consequence of such multichannel machining is a distribution of diameters that has more variation than is typical of a production run of cylinders from a single machine, operator, and setup. As another possibility, if the histogram is **truncated**, this might suggest that the lot of cylinders has been 100% inspected and sorted, removing all cylinders with excessive diameters. Or, upon marking engineering specifications (requirements) for cylinder diameter on the histogram, one may get a picture like that in Figure 3.8. It then becomes obvious that the lathe turning the cylinders needs adjustment in order to increase the typical diameter. But it also becomes clear that the basic process variation is so large that this adjustment will fail to bring essentially all diameters into specifications. Armed with this realization and a knowledge of the economic consequences of parts failing to meet specifications, an engineer can intelligently weigh alternative courses of action: sorting of all incoming parts, demanding that the vendor use more precise equipment, seeking a new vendor, etc.

Investigating the shape of a data set is useful not only because it can lend insight into physical mechanisms but also because shape can be important when determining the appropriateness of methods of formal statistical inference like those discussed later in this book. A methodology appropriate for one distributional shape may not be appropriate for another.
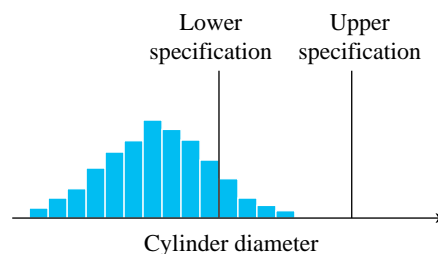

Figure 3.8 Histogram marked with engineering specifications

3.1.3   Scatterplots and Run Charts

Dot diagrams, stem-and-leaf plots, frequency tables, and histograms are univariate tools. But engineering data are often multivariate and *relationships between the variables* are then usually of interest. The familiar device of making a two-dimensional **scatterplot** of data pairs is a simple and effective way of displaying potential relationships between two variables.

Example 3

Bolt Torques on a Face Plate

Brenny, Christensen, and Schneider measured the torques required to loosen six distinguishable bolts holding the front plate on a type of heavy equipment component. Table 3.4 contains the torques (in ft lb) required for bolts number 3 and 4, respectively, on 34 different components. Figure 3.9 is a scatterplot of the bivariate data from Table 3.4. In this figure, where several points must be plotted at a single location, the number of points occupying the location has been plotted instead of a single dot.

The plot gives at least a weak indication that large torques at position 3 are accompanied by large torques at position 4. In practical terms, this is comforting;

Table 3.4
Torques Required to Loosen Two Bolts on Face Plates (ft lb)

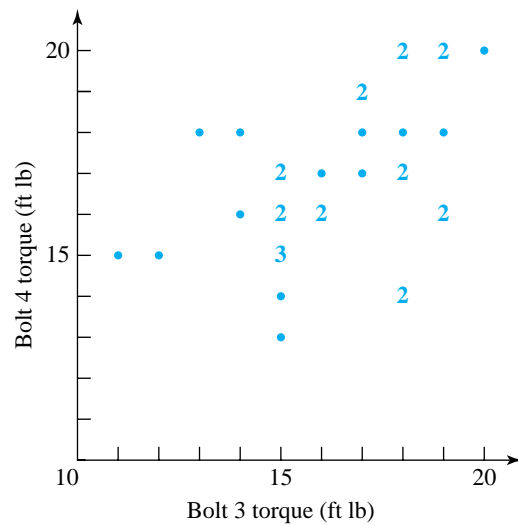| Component	Bolt 3 Torque	Bolt 4 Torque	Component	Bolt 3 Torque	Bolt 4 Torque
1	16	16	18	15	14
2	15	16	19	17	17
3	15	17	20	14	16
4	15	16	21	17	18
5	20	20	22	19	16
6	19	16	23	19	18
7	19	20	24	19	20
8	17	19	25	15	15
9	15	15	26	12	15
10	11	15	27	18	20
11	17	19	28	13	18
12	18	17	29	14	18
13	18	14	30	18	18
14	15	15	31	18	14
15	18	17	32	15	13
16	15	17	33	16	17
17	18	20	34	16	16

Figure 3.9 Scatterplot of bolt 3 and bolt 4 torques

otherwise, unwanted differential forces might act on the face plate. It is also quite reasonable that bolt 3 and bolt 4 torques be related, since the bolts were tightened by different heads of a single pneumatic wrench operating off a single source of compressed air. It stands to reason that variations in air pressure might affect the tightening of the bolts at the two positions similarly, producing the big-together, small-together pattern seen in Figure 3.9.

The previous example illustrates the point that relationships seen on scatterplots suggest a common physical cause for the behavior of variables and can help reveal that cause.

In the most common version of the scatterplot, the variable on the horizontal axis is a time variable. A scatterplot in which univariate data are plotted against time order of observation is called a **run chart** or **trend chart**. Making run charts is one of the most helpful statistical habits an engineer can develop. Seeing patterns on a run chart leads to thinking about what process variables were changing in concert with the pattern. This can help develop a keener understanding of how process behavior is affected by those variables that change over time.

Example 4

Diameters of Consecutive Parts Turned on a Lathe

Williams and Markowski studied a process for rough turning of the outer diameter on the outer race of a constant velocity joint. Table 3.5 gives the diameters (in inches above nominal) for 30 consecutive joints turned on a particular automatic

Example 4
(*continued*)

Table 3.5
30 Consecutive Outer Diameters Turned on a Lathe

Joint	Diameter (inches above nominal)	Joint	Diameter (inches above nominal)
1	−.005	16	.015
2	.000	17	.000
3	−.010	18	.000
4	−.030	19	−.015
5	−.010	20	−.015
6	−.025	21	−.005
7	−.030	22	−.015
8	−.035	23	−.015
9	−.025	24	−.010
10	−.025	25	−.015
11	−.025	26	−.035
12	−.035	27	−.025
13	−.040	28	−.020
14	−.035	29	−.025
15	−.035	30	−.015
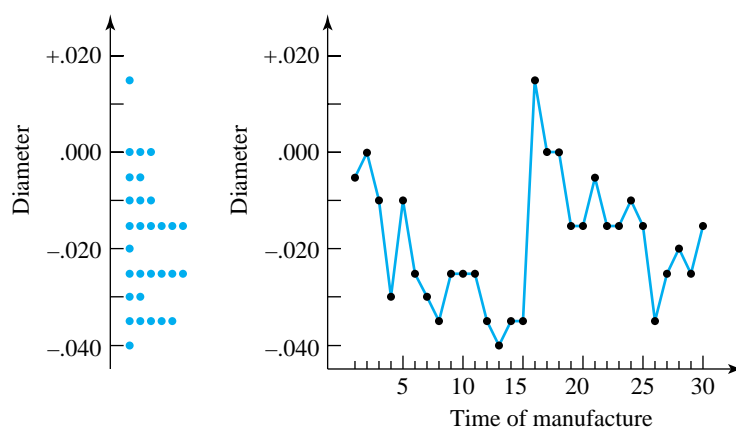
Figure 3.10   Dot diagram and run chart of consecutive outer diameters

lathe. Figure 3.10 gives both a dot diagram and a run chart for the data in the table. In keeping with standard practice, consecutive points on the run chart have been connected with line segments.

Here the dot diagram is not particularly suggestive of the physical mechanisms that generated the data. But the time information added in the run chart is revealing. Moving along in time, the outer diameters tend to get smaller until

part 16, where there is a large jump, followed again by a pattern of diameter generally decreasing in time. In fact, upon checking production records, Williams and Markowski found that the lathe had been turned off and allowed to cool down between parts 15 and 16. The pattern seen on the run chart is likely related to the behavior of the lathe's hydraulics. When cold, the hydraulics probably don't do as good a job pushing the cutting tool into the part being turned as when they are warm. Hence, the turned parts become smaller as the lathe warms up. In order to get parts closer to nominal, the aimed-for diameter might be adjusted up by about .020 in. and parts run only after warming up the lathe.

Section 1 Exercises

1. The following are percent yields from 40 runs of a chemical process, taken from J. S. Hunter's article "The Technology of Quality" (*RCA Engineer*, May/June 1985):

   65.6, 65.6, 66.2, 66.8, 67.2, 67.5, 67.8, 67.8, 68.0, 68.0, 68.2, 68.3, 68.3, 68.4, 68.9, 69.0, 69.1, 69.2, 69.3, 69.5, 69.5, 69.5, 69.8, 69.9, 70.0, 70.2, 70.4, 70.6, 70.6, 70.7, 70.8, 70.9, 71.3, 71.7, 72.0, 72.6, 72.7, 72.8, 73.5, 74.2

   Make a dot diagram, a stem-and-leaf plot, a frequency table, and a histogram of these data.

2. Make back-to-back stem-and-leaf plots for the two samples in Table 3.1.

3. Osborne, Bishop, and Klein collected manufacturing data on the torques required to loosen bolts holding an assembly on a piece of heavy machinery. The accompanying table shows part of their data concerning two particular bolts. The torques recorded (in ft lb) were taken from 15 different pieces of equipment as they were assembled.
   (a) Make a scatterplot of these paired data. Are there any obvious patterns in the plot?

(b) A trick often employed in the analysis of paired data such as these is to reduce the pairs to differences by subtracting the values of one of the variables from the other. Compute differences (top bolt–bottom bolt) here. Then make and interpret a dot diagram for these values.

Piece	Top Bolt	Bottom Bolt
1	110	125
2	115	115
3	105	125
4	115	115
5	115	120
6	120	120
7	110	115
8	125	125
9	105	110
10	130	110
11	95	120
12	110	115
13	110	120
14	95	115
15	105	105

3.2 Quantiles and Related Graphical Tools

Most readers will be familiar with the concept of a *percentile*. The notion is most famous in the context of reporting scores on educational achievement tests. For example, if a person has scored at the 80th percentile, roughly 80% of those taking

the test had worse scores, and roughly 20% had better scores. This concept is also useful in the description of engineering data. However, because it is often more convenient to work in terms of fractions between 0 and 1 rather than in percentages between 0 and 100, slightly different terminology will be used here: "Quantiles," rather than percentiles, will be discussed. After the quantiles of a data set are carefully defined, they are used to create a number of useful tools of descriptive statistics: quantile plots, boxplots, Q-Q plots, and normal plots (a type of theoretical Q-Q plot).

3.2.1 Quantiles and Quantile Plots

Roughly speaking, for a number p between 0 and 1, the p **quantile** of a distribution is a number such that a fraction p of the distribution lies to the left and a fraction $1 - p$ of the distribution lies to the right. However, because of the discreteness of finite data sets, it is necessary to state *exactly* what will be meant by the terminology. Definition 1 gives the precise convention that will be used in this text.

Definition 1

For a data set consisting of n values that when ordered are $x_1 \leq x_2 \leq \cdots \leq x_n$,

1. if $p = \frac{i-.5}{n}$ for a positive integer $i \leq n$, the p **quantile** of the data set is

$$Q(p) = Q\left(\frac{i - .5}{n}\right) = x_i$$

(The ith smallest data point will be called the $\frac{i-.5}{n}$ quantile.)

2. for any number p between $\frac{.5}{n}$ and $\frac{n-.5}{n}$ that is not of the form $\frac{i-.5}{n}$ for an integer i, the p **quantile** of the data set will be obtained by linear interpolation between the two values of $Q(\frac{i-.5}{n})$ with corresponding $\frac{i-.5}{n}$ that bracket p.

In both cases, the notation $Q(p)$ will be used to denote the p quantile.

Definition 1 identifies $Q(p)$ for all p between $.5/n$ and $(n - .5)/n$. To find $Q(p)$ for such a value of p, one may solve the equation $p = (i - .5)/n$ for i, yielding

Index (i) of the ordered data point that is $Q(p)$

$$i = np + .5$$

and locate the "$(np + .5)$th ordered data point."

Example 5

Quantiles for Dry Breaking Strengths of Paper Towel

Lee, Sebghati, and Straub did a study of the dry breaking strength of several brands of paper towel. Table 3.6 shows ten breaking strengths (in grams) reported by the students for a generic towel. By ordering the strength data and computing values of $\frac{i-.5}{10}$, one can easily find the .05, .15, .25, ..., .85, and .95 quantiles of the breaking strength distribution, as shown in Table 3.7.

Since there are $n = 10$ data points, each one accounts for 10% of the data set. Applying convention (1) in Definition 1 to find (for example) the .35 quantile,

Table 3.6
Ten Paper Towel Breaking
Strengths

Test	Breaking Strength (g)
1	8,577
2	9,471
3	9,011
4	7,583
5	8,572
6	10,688
7	9,614
8	9,614
9	8,527
10	9,165

Table 3.7
Quantiles of the Paper Towel Breaking Strength
Distribution

i	$\frac{i-.5}{10}$	ith Smallest Data Point, $x_i = Q\left(\frac{i-.5}{10}\right)$
1	.05	$7,583 = Q(.05)$
2	.15	$8,527 = Q(.15)$
3	.25	$8,572 = Q(.25)$
4	.35	$8,577 = Q(.35)$
5	.45	$9,011 = Q(.45)$
6	.55	$9,165 = Q(.55)$
7	.65	$9,471 = Q(.65)$
8	.75	$9,614 = Q(.75)$
9	.85	$9,614 = Q(.85)$
10	.95	$10,688 = Q(.95)$

Example 5
(*continued*)

the smallest 3 data points and half of the fourth smallest are counted as lying to the left of the desired number, and the largest 6 data points and half of the seventh largest are counted as lying to the right. Thus, the fourth smallest data point must be the .35 quantile, as is shown in Table 3.7.

To illustrate convention (2) of Definition 1, consider finding the .5 and .93 quantiles of the strength distribution. Since .5 is $\frac{.5-.45}{.55-.45} = .5$ of the way from .45 to .55, linear interpolation gives

▶ $Q(.5) = (1 - .5)\, Q(.45) + .5\, Q(.55) = .5(9{,}011) + .5(9{,}165) = 9{,}088$ g

Then, observing that .93 is $\frac{.93-.85}{.95-.85} = .8$ of the way from .85 to .95, linear interpolation gives

$Q(.93) = (1 - .8)\, Q(.85) + .8Q(.95) = .2(9{,}614) + .8(10{,}688) = 10{,}473.2$ g

Particular round values of p give quantiles $Q(p)$ that are known by special names.

Definition 2

$Q(.5)$ is called the **median** of a distribution.

Definition 3

$Q(.25)$ and $Q(.75)$ are called the **first (or lower) quartile** and **third (or upper) quartile** of a distribution, respectively.

Example 5
(*continued*)

Referring again to Table 3.7 and the value of $Q(.5)$ previously computed, for the breaking strength distribution

$$\text{Median} = Q(.5) = 9{,}088 \text{ g}$$

▶ $$\text{1st quartile} = Q(.25) = 8{,}572 \text{ g}$$

▶ $$\text{3rd quartile} = Q(.75) = 9{,}614 \text{ g}$$

A way of representing the quantile idea graphically is to make a **quantile plot**.

Definition 4

A **quantile plot** is a plot of $Q(p)$ versus p. For an ordered data set of size n containing values $x_1 \le x_2 \le \cdots \le x_n$, such a display is made by first plotting the points $(\frac{i-.5}{n}, x_i)$ and then connecting consecutive plotted points with straight-line segments.

It is because convention (2) in Definition 1 calls for linear interpolation that straight-line segments enter the picture in making a quantile plot.

Example 5
(*continued*)

Referring again to Table 3.7 for the $\frac{i-.5}{10}$ quantiles of the breaking strength distribution, it is clear that a quantile plot for these data will involve plotting and then connecting consecutive ones of the following ordered pairs.

$$
\begin{array}{lll}
(.05,\,7{,}583) & (.15,\,8{,}527) & (.25,\,8{,}572) \\
(.35,\,8{,}577) & (.45,\,9{,}011) & (.55,\,9{,}165) \\
(.65,\,9{,}471) & (.75,\,9{,}614) & (.85,\,9{,}614) \\
(.95,\,10{,}688) & &
\end{array}
$$

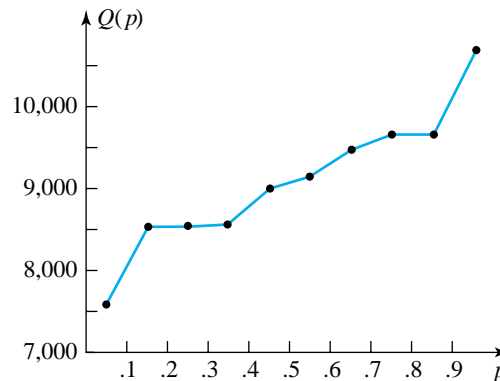Figure 3.11 gives such a plot.

Figure 3.11   Quantile plot of paper towel strengths

A quantile plot allows the user to do some informal visual smoothing of the plot to compensate for any jaggedness. (The tacit assumption is that the underlying data-generating mechanism would itself produce smoother and smoother quantile plots for larger and larger samples.)

3.2.2 Boxplots

Familiarity with the quantile idea is the principal prerequisite for making **boxplots**, an alternative to dot diagrams or histograms. The boxplot carries somewhat less information, but it has the advantage that many can be placed side-by-side on a single page for comparison purposes.

    There are several common conventions for making boxplots. The one that will be used here is illustrated in generic fashion in Figure 3.12. A box is made to extend from the first to the third quartiles and is divided by a line at the median. Then the **interquartile range**

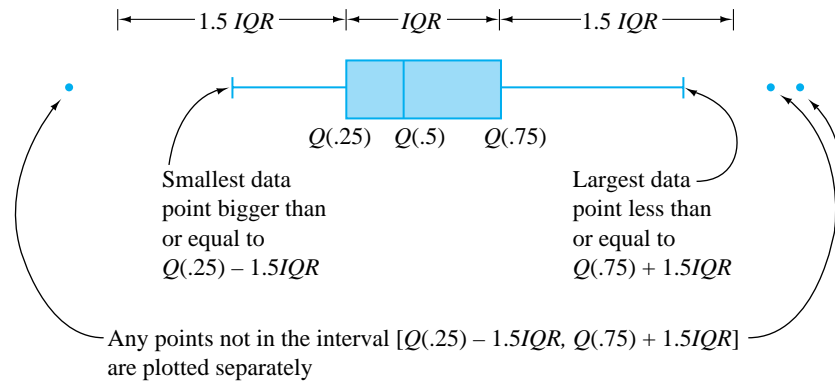*Interquartile range*

$$
IQR = Q(.75) - Q(.25)
$$

Figure 3.12   Generic boxplot

is calculated and the smallest data point within $1.5IQR$ of $Q(.25)$ and the largest data point within $1.5IQR$ of $Q(.75)$ are determined. Lines called **whiskers** are made to extend out from the box to these values. Typically, most data points will be within the interval $[Q(.25) - 1.5IQR, Q(.75) + 1.5IQR]$. Any that are not then get plotted individually and are thereby identified as *outlying* or unusual.

Example 5
(*continued*)

Consider making a boxplot for the paper towel breaking strength data. To begin,

$$Q(.25) = 8{,}572 \text{ g}$$
$$Q(.5) = 9{,}088 \text{ g}$$
$$Q(.75) = 9{,}614 \text{ g}$$

So

$$IQR = Q(.75) - Q(.25) = 9{,}614 - 8{,}572 = 1{,}042 \text{ g}$$

and

$$1.5IQR = 1{,}563 \text{ g}$$

Then

$$Q(.75) + 1.5IQR = 9{,}614 + 1{,}563 = 11{,}177 \text{ g}$$

and

$$Q(.25) - 1.5IQR = 8{,}572 - 1{,}563 = 7{,}009 \text{ g}$$

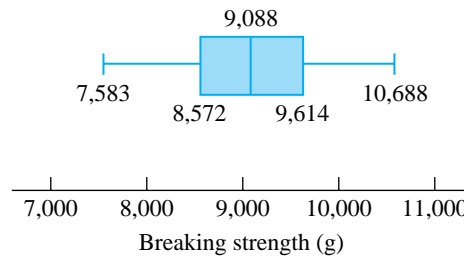Since all the data points lie in the range 7,009 g to 11,177 g, the boxplot is as shown in Figure 3.13.

Figure 3.13 Boxplot of the paper towel strengths

A boxplot shows distributional location through the placement of the box and whiskers along a number line. It shows distributional spread through the extent of the box and the whiskers, with the box enclosing the middle 50% of the distribution. Some elements of distributional shape are indicated by the symmetry (or lack thereof) of the box and of the whiskers. And a gap between the end of a whisker and a separately plotted point serves as a reminder that no data values fall in that interval.

Two or more boxplots drawn to the same scale and side by side provide an effective way of comparing samples.

Example 6
(*Example 2, page 67, revisited*)

More on Bullet Penetration Depths

Table 3.8 contains the raw information needed to find the $\frac{i-.5}{20}$ quantiles for the two distributions of bullet penetration depth introduced in the previous section. For the 230 grain bullet penetration depths, interpolation yields

$$Q(.25) = .5Q(.225) + .5Q(.275) = .5(38.75) + .5(39.75) = 39.25 \text{ mm}$$

$$Q(.5) = .5Q(.475) + .5Q(.525) = .5(42.55) + .5(42.90) = 42.725 \text{ mm}$$

$$Q(.75) = .5Q(.725) + .5Q(.775) = .5(47.90) + .5(48.15) = 48.025 \text{ mm}$$

So

$$IQR = 48.025 - 39.25 = 8.775 \text{ mm}$$

$$1.5IQR = 13.163 \text{ mm}$$

$$Q(.75) + 1.5IQR = 61.188 \text{ mm}$$

$$Q(.25) - 1.5IQR = 26.087 \text{ mm}$$

Example 6
(*continued*)

Similar calculations for the 200 grain bullet penetration depths yield

$$Q(.25) = 60.25 \text{ mm}$$
$$Q(.5) = 62.80 \text{ mm}$$
$$Q(.75) = 64.35 \text{ mm}$$
$$Q(.75) + 1.5IQR = 70.50 \text{ mm}$$
$$Q(.25) - 1.5IQR = 54.10 \text{ mm}$$

Table 3.8
Quantiles of the Bullet Penetration Depth Distributions

i	$\frac{i-.5}{20}$	ith Smallest 230 Grain Data Point $= Q(\frac{i-.5}{20})$	ith Smallest 200 Grain Data Point $= Q(\frac{i-.5}{20})$
1	.025	27.75	58.00
2	.075	37.35	58.65
3	.125	38.35	59.10
4	.175	38.35	59.50
5	.225	38.75	59.80
6	.275	39.75	60.70
7	.325	40.50	61.30
8	.375	41.00	61.50
9	.425	41.15	62.30
10	.475	42.55	62.65
11	.525	42.90	62.95
12	.575	43.60	63.30
13	.625	43.85	63.55
14	.675	47.30	63.80
15	.725	47.90	64.05
16	.775	48.15	64.65
17	.825	49.85	65.00
18	.875	51.25	67.75
19	.925	51.60	70.40
20	.975	56.00	71.70

Figure 3.14 then shows boxplots placed side by side on the same scale. The plots show the larger and more consistent penetration depths of the 200 grain bullets. They also show the existence of one particularly extreme data point in the 200 grain data set. Further, the relative lengths of the whiskers hint at some skewness (recall the terminology introduced with Figure 3.7) in the data. And all of this is done in a way that is quite uncluttered and compact. Many more of
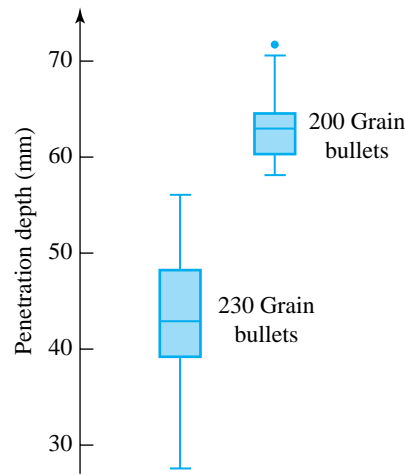
Figure 3.14   Side-by-side boxplots for the bullet penetration depths

these boxes could be added to Figure 3.14 (to compare other bullet types) without visual overload.

3.2.3 Q-Q Plots and Comparing Distributional Shapes

It is often important to compare the shapes of two distributions. Comparing histograms is one rough way of doing this. A more sensitive way is to make a single plot based on the quantile functions for the two distributions and exploit the fact that "equal shape" is equivalent to "linearly related quantile functions." Such a plot is called a **quantile-quantile plot** or, more briefly, a **Q-Q plot**.

Consider the two small artificial data sets given in Table 3.9. Dot diagrams of these two data sets are given in Figure 3.15. The two data sets have the same shape. But why is this so? One way to look at the equality of the shapes is to note that

$$i\text{th smallest value in data set } 2 = 2(i\text{th smallest value in data set } 1) + 1 \quad \textbf{(3.1)}$$

Then, recognizing ordered data values as quantiles and letting Q_1 and Q_2 stand for the quantile functions of the two respective data sets, it is clear from display (3.1) that

$$Q_2(p) = 2Q_1(p) + 1 \quad \textbf{(3.2)}$$

Table 3.9
Two Small Artificial Data Sets

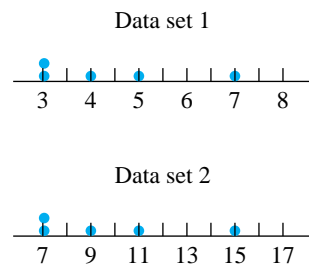| Data Set 1	Data Set 2
3, 5, 4, 7, 3	15, 7, 9, 7, 11

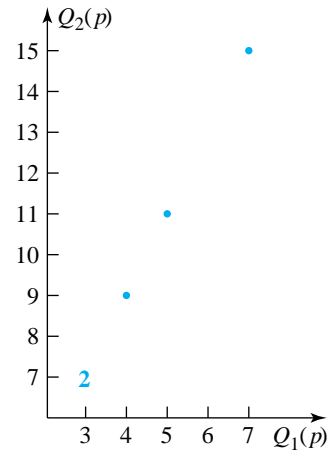Figure 3.15   Dot diagrams for two small data sets

Figure 3.16   Q-Q plot for the data of Table 3.9

That is, the two data sets have quantile functions that are linearly related. Looking at either display (3.1) or (3.2), it is obvious that a plot of the points

$$\left(Q_1\left(\frac{i - .5}{5}\right), \quad Q_2\left(\frac{i - .5}{5}\right)\right)$$

(for $i = 1, 2, 3, 4, 5$) should be exactly linear. Figure 3.16 illustrates this—in fact Figure 3.16 is a Q-Q plot for the data sets of Table 3.9.

Definition 5

A **Q-Q plot** for two data sets with respective quantile functions Q_1 and Q_2 is a plot of ordered pairs $(Q_1(p), Q_2(p))$ for appropriate values of p. When two data sets of size n are involved, the values of p used to make the plot will be $\frac{i-.5}{n}$ for $i = 1, 2, \ldots, n$. When two data sets of unequal sizes are involved, the values of p used to make the plot will be $\frac{i-.5}{n}$ for $i = 1, 2, \ldots, n$, where n is the size of the smaller set.

To make a Q-Q plot *for two data sets of the same size,*

Steps in making a Q-Q plot

1. order each from the smallest observation to the largest,

2. pair off corresponding values in the two data sets, and

3. plot ordered pairs, with the horizontal coordinates coming from the first data set and the vertical ones from the second.

When data sets of unequal size are involved, the ordered values from the smaller data set must be paired with quantiles of the larger data set obtained by interpolation.

A *Q-Q* plot that is reasonably linear indicates the two distributions involved have similar shapes. When there are significant departures from linearity, the character of those departures reveals the ways in which the shapes differ.

Example 6
(*continued*)

Returning again to the bullet penetration depths, Table 3.8 (page 84) gives the raw material for making a *Q-Q* plot. The depths on each row of that table need only be paired and plotted in order to make the plot given in Figure 3.17.

The scatterplot in Figure 3.17 is not terribly linear when looked at as a whole. However, the points corresponding to the 2nd through 13th smallest values in each data set do look fairly linear, indicating that (except for the *extreme* lower ends) the lower ends of the two distributions have similar shapes.

The horizontal jog the plot takes between the 13th and 14th plotted points indicates that the gap between 43.85 mm and 47.30 mm (for the 230 grain data) is out of proportion to the gap between 63.55 and 63.80 mm (for the 200 grain data). This hints that there was some kind of basic physical difference in the mechanisms that produced the smaller and larger 230 grain penetration depths. Once this kind of indication is discovered, it is a task for ballistics experts or materials people to explain the phenomenon.

Because of the marked departure from linearity produced by the 1st plotted point (27.75, 58.00), there is also a drastic difference in the shapes of the extreme lower ends of the two distributions. In order to move that point back on line with the rest of the plotted points, it would need to be moved to the right or down (i.e., increase the smallest 230 grain observation or decrease the smallest 200 grain observation). That is, *relative to the 200 grain distribution, the 230 grain distribution is long-tailed to the low side.* (Or to put it differently, relative to the 230 grain distribution, the 200 grain distribution is short-tailed to the low side.) Note that the difference in shapes was already evident in the boxplot in Figure 3.14. Again, it would remain for a specialist to explain this difference in distributional shapes.
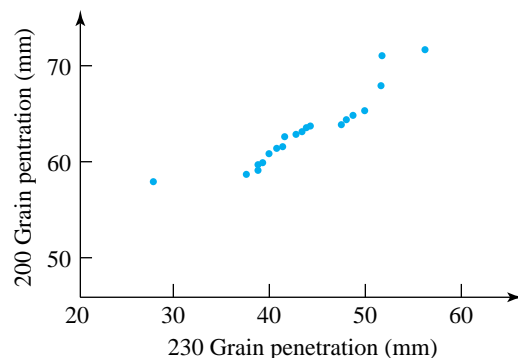

Figure 3.17   *Q-Q* plot for the bullet penetration depths

The Q-Q plotting idea is useful when applied to two data sets, and it is easiest to explain the notion in such an "empirical versus empirical" context. But its greatest usefulness is really when it is applied to one quantile function that represents a data set and a second that represents a *theoretical distribution*.

Definition 6

A **theoretical Q-Q plot** or **probability plot** for a data set of size n and a theoretical distribution, with respective quantile functions Q_1 and Q_2, is a plot of ordered pairs $(Q_1(p), Q_2(p))$ for appropriate values of p. In this text, the values of p of the form $\frac{i-.5}{n}$ for $i = 1, 2, \ldots, n$ will be used.

Recognizing $Q_1(\frac{i-.5}{n})$ as the ith smallest data point, one sees that a theoretical Q-Q plot is a plot of points with horizontal plotting positions equal to observed data and vertical plotting positions equal to quantiles of the theoretical distribution. That is, with ordered data $x_1 \le x_2 \le \cdots \le x_n$, the points

Ordered pairs making a probability plot

$$\left(x_i, Q_2\left(\frac{i - .5}{n} \right) \right)$$

are plotted. Such a plot allows one to ask, "Does the data set have a shape similar to the theoretical distribution?"

Normal plotting

The most famous version of the theoretical Q-Q plot occurs when quantiles for the **standard normal** or **Gaussian** distribution are employed. This is the familiar bell-shaped distribution. Table 3.10 gives some quantiles of this distribution. In order to find $Q(p)$ for p equal to one of the values .01, .02, $\ldots$, .98, .99, locate the entry in the row labeled by the first digit after the decimal place and in the column labeled by the second digit after the decimal place. (For example, $Q(.37) = -.33$.) A simple numerical approximation to the values given in Table 3.10 adequate for most plotting purposes is

Approximate standard normal quantiles

$$Q(p) \approx 4.9(p^{.14} - (1 - p)^{.14}) \tag{3.3}$$

The origin of Table 3.10 is not obvious at this point. It will be explained in Section 5.2, but for the time being consider the following crude argument to the effect that the quantiles in the table correspond to a bell-shaped distribution. Imagine that each entry in Table 3.10 corresponds to a data point in a set of size $n = 99$. A possible frequency table for those 99 data points is given as Table 3.11. The tally column in Table 3.11 shows clearly the bell shape.

The standard normal quantiles can be used to make a theoretical Q-Q plot as a way of assessing how bell-shaped a data set looks. The resulting plot is called a **normal (probability) plot**.

Table 3.10
Standard Normal Quantiles

	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0		−2.33	−2.05	−1.88	−1.75	−1.65	−1.55	−1.48	−1.41	−1.34
.1	−1.28	−1.23	−1.18	−1.13	−1.08	−1.04	−.99	−.95	−.92	−.88
.2	−.84	−.81	−.77	−.74	−.71	−.67	−.64	−.61	−.58	−.55
.3	−.52	−.50	−.47	−.44	−.41	−.39	−.36	−.33	−.31	−.28
.4	−.25	−.23	−.20	−.18	−.15	−.13	−.10	−.08	−.05	−.03
.5	0.00	.03	.05	.08	.10	.13	.15	.18	.20	.23
.6	.25	.28	.31	.33	.36	.39	.41	.44	.47	.50
.7	.52	.55	.58	.61	.64	.67	.71	.74	.77	.81
.8	.84	.88	.92	.95	.99	1.04	1.08	1.13	1.18	1.23
.9	1.28	1.34	1.41	1.48	1.55	1.65	1.75	1.88	2.05	2.33

Table 3.11
A Frequency Table for the Standard Normal Quantiles

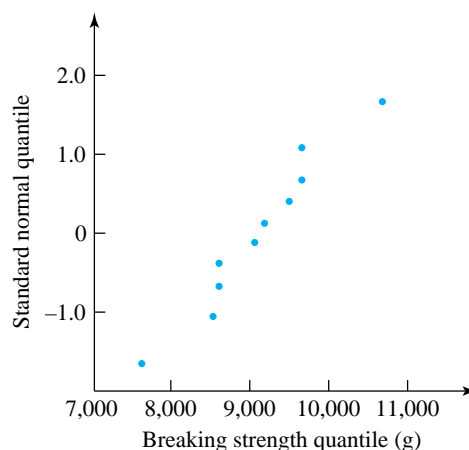| Value	Tally	Frequency
−2.80 to −2.30	|	1
−2.29 to −1.79	||	2
−1.78 to −1.28	⊞⊞ ||	7
−1.27 to −.77	⊞⊞ ⊞⊞ ||	12
−.76 to −.26	⊞⊞ ⊞⊞ ⊞⊞ ||	17
−.25 to .25	⊞⊞ ⊞⊞ ⊞⊞ ⊞⊞ |	21
.26 to .76	⊞⊞ ⊞⊞ ⊞⊞ ||	17
.77 to 1.27	⊞⊞ ⊞⊞ ||	12
1.28 to 1.78	⊞⊞ ||	7
1.79 to 2.29	||	2
2.30 to 2.80	|	1

Example 5
(*continued*)

Consider again the paper towel strength testing scenario and now the issue of how bell-shaped the data set in Table 3.6 (page 79) is. Table 3.12 was made using Tables 3.7 (page 79) and 3.10; it gives the information needed to produce the theoretical Q-Q plot in Figure 3.18.

Considering the small size of the data set involved, the plot in Figure 3.18 is fairly linear, and so the data set is reasonably bell-shaped. As a practical consequence of this judgment, it is then possible to use the normal probability models discussed in Section 5.2 to describe breaking strength. These could be employed to make breaking strength predictions, and methods of formal statistical inference based on them could be used in the analysis of breaking strength data.

Example 5
(*continued*)

Table 3.12
Breaking Strength and Standard Normal Quantiles

i	$\frac{i-.5}{10}$	$\frac{i-.5}{10}$ Breaking Strength Quantile	$\frac{i-.5}{10}$ Standard Normal Quantile
1	.05	7,583	-1.65
2	.15	8,527	-1.04
3	.25	8,572	$-.67$
4	.35	8,577	$-.39$
5	.45	9,011	$-.13$
6	.55	9,165	.13
7	.65	9,471	.39
8	.75	9,614	.67
9	.85	9,614	1.04
10	.95	10,688	1.65

Figure 3.18   Theoretical *Q-Q* plot for the paper towel strengths

Special graph paper, called **normal probability paper** (or just *probability paper*), is available as an alternative way of making normal plots. Instead of plotting points on regular graph paper using vertical plotting positions taken from Table 3.10, points are plotted on probability paper using vertical plotting positions of the form $\frac{i-.5}{n}$. Figure 3.19 is a normal plot of the breaking strength data from Example 5 made on probability paper. Observe that this is virtually identical to the plot in Figure 3.18.

Normal plots are not the only kind of theoretical *Q-Q* plots useful to engineers. Many other types of theoretical distributions are of engineering importance, and each can be used to make theoretical *Q-Q* plots. This point is discussed in more
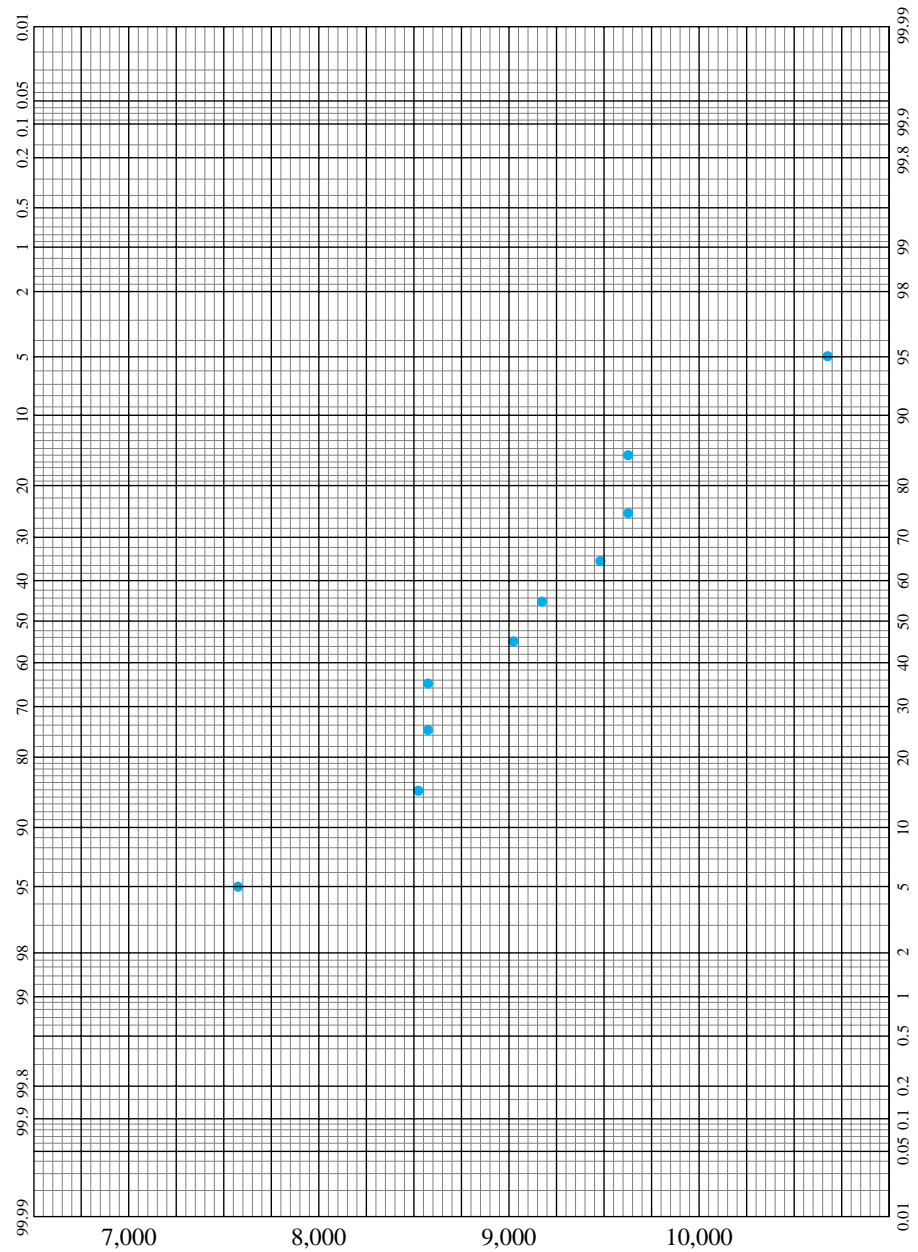
Figure 3.19 Normal plot for the paper towel strengths (made on probability paper, used with permission of the Keuffel and Esser Company)

detail in Section 5.3, but the introduction of theoretical Q-Q plotting here makes it possible to emphasize the relationship between probability plotting and (empirical) Q-Q plotting.

1. The following are data (from *Introduction to Contemporary Statistical Methods* by L. H. Koopmans) on the impact strength of sheets of insulating material cut in two different ways. (The values are in ft lb.)

Lengthwise Cuts	Crosswise Cuts
1.15	.89
.84	.69
.88	.46
.91	.85
.86	.73
.88	.67
.92	.78
.87	.77
.93	.80
.95	.79

(a) Make quantile plots for these two samples. Find the medians, the quartiles, and the .37 quantiles for the two data sets.

(b) Draw (to scale) carefully labeled side-by-side boxplots for comparing the two cutting methods. Discuss what these show about the two methods.

(c) Make and discuss the appearance of a Q-Q plot for comparing the shapes of these two data sets.

2. Make a Q-Q plot for the two small samples in Table 3.13 in Section 3.3.

3. Make and interpret a normal plot for the yield data of Exercise 1 of Section 3.1.

4. Explain the usefulness of theoretical Q-Q plotting.

....

3.3   Standard Numerical Summary Measures

The smooth functioning of most modern technology depends on the reduction of large amounts of data to a few informative numerical summary values. For example, over the period of a month, a lab doing compressive strength testing for a manufacturer's concrete blocks may make hundreds or even thousands of such measurements. But for some purposes, it may be adequate to know that those strengths average 4,618 psi with a range of 2,521 psi (from smallest to largest).

In this section, several standard summary measures for quantitative data are discussed, including the mean, median, range, and standard deviation. Measures of location are considered first, then measures of spread. There follows a discussion of the difference between sample statistics and population parameters and then illustrations of how numerical summaries can be effectively used in simple plots to clarify the results of statistical engineering studies. Finally, there is a brief discussion of the use of personal computer software in elementary data summarization.

3.3.1   Measures of Location

Most people are familiar with the concept of an "average" as being representative of, or in the center of, a data set. Temperatures may vary between different locations in a blast furnace, but an average temperature tells something about a middle or representative temperature. Scores on an exam may vary, but one is relieved to score at least above average.

The word *average*, as used in colloquial speech, has several potential technical meanings. One is the median, $Q(.5)$, which was introduced in the last section. The median divides a data set in half. Roughly half of the area enclosed by the bars of a well-made histogram will lie to either side of the median. As a measure of center, it is completely insensitive to the effects of a few extreme or outlying observations. For example, the small set of data

$$2, 3, 6, 9, 10$$

has median 6, and this remains true even if the value 10 is replaced by 10,000,000 and/or the value 2 is replaced by $-200,000$.

The previous section used the median as a center value in the making of boxplots. But the median is not the technical meaning most often attached to the notion of average in statistical analyses. Instead, it is more common to employ the (arithmetic) *mean*.

Definition 7	The **(arithmetic) mean** of a sample of quantitative data (say, $x_1, x_2, \ldots, x_n$) is $$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

The mean is sometimes called the *first moment* or *center of mass* of a distribution, drawing on an analogy to mechanics. Think of placing a unit mass along the number line at the location of each value in a data set—the balance point of the mass distribution is at $\bar{x}$.

Example 7	**Waste on Bulk Paper Rolls**

Hall, Luethe, Pelszynski, and Ringhofer worked with a company that cuts paper from large rolls purchased in bulk from several suppliers. The company was interested in determining the amount of waste (by weight) on rolls obtained from the various sources. Table 3.13 gives percent waste data, which the students obtained for six and eight rolls, respectively, of paper purchased from two different sources.

The medians and means for the two data sets are easily obtained. For the supplier 1 data,

$$Q(.5) = .5(.65) + .5(.92) = .785\% \text{ waste}$$

and

$$\bar{x} = \frac{1}{6}(.37 + .52 + .65 + .92 + 2.89 + 3.62) = 1.495\% \text{ waste}$$

Example 7
(*continued*)

Table 3.13
Percent Waste by Weight on Bulk Paper Rolls

| Supplier 1	Supplier 2
.37, .52, .65,	.89, .99, 1.45, 1.47,
.92, 2.89, 3.62	1.58, 2.27, 2.63, 6.54

For the supplier 2 data,

$$Q(.5) = .5(1.47) + .5(1.58) = 1.525\% \text{ waste}$$

and

$$\bar{x} = \frac{1}{8}(.89 + .99 + 1.45 + 1.47 + 1.58 + 2.27 + 2.63 + 6.54)$$

$$= 2.228\% \text{ waste}$$

Figure 3.20 shows dot diagrams with the medians and means marked. Notice that a comparison of either medians or means for the two suppliers shows the supplier 2 waste to be larger than the supplier 1 waste. But there is a substantial difference between the median and mean values for a given supplier. In both cases, the mean is quite a bit larger than the corresponding median. This reflects the right-skewed nature of both data sets. In both cases, the center of mass of the distribution is pulled strongly to the right by a few extremely large values.
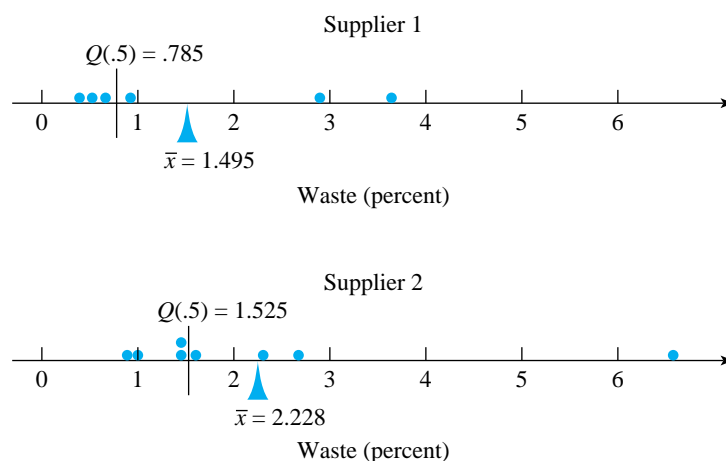
Figure 3.20 Dot diagrams for the waste percentages

Example 7 shows clearly that, in contrast to the median, the mean is a measure of center that can be strongly affected by a few extreme data values. People sometimes say that because of this, one or the other of the two measures is "better." Such statements lack sense. Neither is better; they are simply measures with different properties. And the difference is one that intelligent consumers of statistical information do well to keep in mind. The "average" income of employees at a company paying nine workers each $10,000/year and a president $110,000/year can be described as $10,000/year or $20,000/year, depending upon whether the median or mean is being used.

3.3.2 Measures of Spread

Quantifying the variation in a data set can be as important as measuring its location. In manufacturing, for example, if a characteristic of parts coming off a particular machine is being measured and recorded, the spread of the resulting data gives information about the intrinsic precision or **capability** of the machine. The location of the resulting data is often a function of machine setup or settings of adjustment knobs. Setups can fairly easily be changed, but improvement of intrinsic machine precision usually requires a capital expenditure for a new piece of equipment or overhaul of an existing one.

Although the point wasn't stressed in Section 3.2, the interquartile range, $IQR = Q(.75) - Q(.25)$, is one possible measure of spread for a distribution. It measures the spread of the middle half of a distribution. Therefore, it is insensitive to the possibility of a few extreme values occurring in a data set. A related measure is the **range**, which indicates the spread of the entire distribution.

Definition 8

The **range** of a data set consisting of ordered values $x_1 \leq x_2 \leq \cdots \leq x_n$ is

$$R = x_n - x_1$$

Notice the word usage here. The word *range* could be used as a verb to say, "The data range from 3 to 21." But to use the word as a noun, one says, "The range is $(21 - 3) = 18$." Since the range depends only on the values of the smallest and largest points in a data set, it is necessarily highly sensitive to extreme (or outlying) values. Because it is easily calculated, it has enjoyed long-standing popularity in industrial settings, particularly as a tool in statistical quality control.

However, most methods of formal statistical inference are based on another measure of distributional spread. A notion of "mean squared deviation" or "root mean squared deviation" is employed to produce measures that are called the **variance** and the **standard deviation**, respectively.

Definition 9	The **sample variance** of a data set consisting of values $x_1, x_2, \ldots, x_n$ is

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

The **sample standard deviation**, s, is the nonnegative square root of the sample variance.

Apart from an exchange of $n - 1$ for n in the divisor, s^2 is an average squared distance of the data points from the central value $\bar{x}$. Thus, s^2 is nonnegative and is 0 only when all data points are exactly alike. The units of s^2 are the squares of the units in which the original data are expressed. Taking the square root of s^2 to obtain s then produces a measure of spread expressed in the original units.

Example 7
(continued)

The spreads in the two sets of percentage wastes recorded in Table 3.13 can be expressed in any of the preceding terms. For the supplier 1 data,

$$Q(.25) = .52$$

$$Q(.75) = 2.89$$

and so

$$IQR = 2.89 - .52 = 2.37\% \text{ waste}$$

Also,

$$R = 3.62 - .37 = 3.25\% \text{ waste}$$

Further,

$$s^2 = \frac{1}{6-1}((.37 - 1.495)^2 + (.52 - 1.495)^2 + (.65 - 1.495)^2 + (.92 - 1.495)^2$$

$$+ (2.89 - 1.495)^2 + (3.62 - 1.495)^2)$$

$$= 1.945(\% \text{ waste})^2$$

so that

$$s = \sqrt{1.945} = 1.394\% \text{ waste}$$

Similar calculations for the supplier 2 data yield the values

$$IQR = 1.23\% \text{ waste}$$

and

$$R = 6.54 - .89 = 5.65\% \text{ waste}$$

Further,

$$s^2 = \frac{1}{8-1}((.89 - 2.228)^2 + (.99 - 2.228)^2 + (1.45 - 2.228)^2 + (1.47 - 2.228)^2$$

$$+ (1.58 - 2.228)^2 + (2.27 - 2.228)^2 + (2.63 - 2.228)^2 + (6.54 - 2.228)^2)$$

$$= 3.383(\% \text{ waste})^2$$

so

$$s = 1.839\% \text{ waste}$$

Supplier 2 has the smaller IQR but the larger R and s. This is consistent with Figure 3.20. The central portion of the supplier 2 distribution is tightly packed. But the single extreme data point makes the overall variability larger for the second supplier than for the first.

The calculation of sample variances just illustrated is meant simply to reinforce the fact that s^2 is a kind of mean squared deviation. Of course, the most sensible way to find sample variances in practice is by using either a handheld electronic calculator with a preprogrammed variance function or a statistical package on a personal computer.

The measures of variation, IQR, R, and s, are not directly comparable. Although it is somewhat out of the main flow of this discussion, it is worth interjecting at this point that it is possible to "put R and s on the same scale." This is done by dividing R by an appropriate conversion factor, known to quality control engineers as d_2. Table B.2 contains *control chart constants* and gives values of d_2 for various sample sizes n. For example, to get R and s on the same scale for the supplier 1 data, division of R by 2.534 is in order, since $n = 6$.

Students often have some initial difficulty developing a feel for the meaning of the standard deviation. One possible help in this effort is a famous theorem of a Russian mathematician.

| Proposition 1 (*Chebyschev's Theorem*) | For any data set and any number k larger than 1, a fraction of at least $1 - (1/k^2)$ of the data are within ks of $\bar{x}$. |

This little theorem says, for example, that at least $\frac{3}{4}$ of a data set is within 2 standard deviations of its mean. And at least $\frac{8}{9}$ of a data set is within 3 standard deviations of its mean. So the theorem promises that if a data set has a small standard deviation, it will be tightly packed about its mean.

Example 7
(*continued*)

Returning to the waste data, consider illustrating the meaning of Chebyschev's theorem with the supplier 1 values. For example, taking $k = 2$, at least $\frac{3}{4} = 1 - (\frac{1}{2})^2$ of the 6 data points (i.e., at least 4.5 of them) must be within 2 standard deviations of $\bar{x}$. In fact

$$\bar{x} - 2s = 1.495 - 2(1.394) = -1.294\% \text{ waste}$$

and

$$\bar{x} + 2s = 1.495 + 2(1.394) = 4.284\% \text{ waste}$$

so simple counting shows that all (a fraction of 1.0) of the data are between these two values.

3.3.3   Statistics and Parameters

At this point, it is important to introduce some more basic terminology. Jargon and notation for distributions of samples are somewhat different than for population distributions (and theoretical distributions).

Definition 10

Numerical summarizations of sample data are called (sample) **statistics**. Numerical summarizations of population and theoretical distributions are called (population or model) **parameters**. Typically, Roman letters are used as symbols for statistics, and Greek letters are used to stand for parameters.

As an example, consider the mean. Definition 7 refers specifically to a calculation for a sample. If a data set represents an entire population, then it is common to use the lowercase Greek letter mu (μ) to stand for the **population mean** and to write

Population mean

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i \qquad \textbf{(3.4)}$$

Comparing this expression to the one in Definition 7, not only is a different symbol used for the mean but also N is used in place of n. It is standard to denote a population size as N and a sample size as n. Chapter 5 gives a definition for the

mean of a theoretical distribution. But it is worth saying now that the symbol μ will be used in that context as well as in the context of equation (3.4).

As another example of the usage suggested by Definition 10, consider the variance and standard deviation. Definition 9 refers specifically to the *sample* variance and standard deviation. If a data set represents an entire population, then it is common to use the lowercase Greek sigma squared (σ^2) to stand for the **population variance** and to define

<div style="text-align: left; color: gray; font-style: italic;">Population variance</div>

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2 \qquad (3.5)$$

The nonnegative square root of σ^2 is then called the **population standard deviation**, σ. (The division in equation (3.5) is by N, and not the $N-1$ that might be expected on the basis of Definition 9. There are reasons for this change, but they are not accessible at this point.) Chapter 5 defines a variance and standard deviation for theoretical distributions, and the symbols σ^2 and σ will be used there as well as in the context of equation (3.5).

On one point, this text will deviate from the Roman/Greek symbolism convention laid out in Definition 10: the notation for quantiles. $Q(p)$ will stand for the pth quantile of a distribution, whether it is from a sample, a population, or a theoretical model.

3.3.4 Plots of Summary Statistics

Plots against time

Plotting numerical summary measures in various ways is often helpful in the early analysis of engineering data. For example, plots of summary statistics against time are frequently revealing.

Example 8
(*Example 8, Chapter 1, revisited—p. 18*)

Monitoring a Critical Dimension of Machined Parts

Cowan, Renk, Vander Leest, and Yakes worked with a company that makes precision metal parts. A critical dimension of one such part was monitored by occasionally selecting and measuring five consecutive pieces and then plotting the sample mean and range. Table 3.14 gives the $\bar{x}$ and R values for 25 consecutive samples of five parts. The values reported are in .0001 in.

Figure 3.21 is a plot of both the means and ranges against order of observation. Looking first at the plot of ranges, no strong trends are obvious, which suggests that the basic short-term variation measured in this critical dimension is stable. The combination of process and measurement precision is neither improving nor degrading with time. The plot of means, however, suggests some kind of physical change. The average dimensions from the second shift on October 27 (samples 9 through 15) are noticeably smaller than the rest of the means. As discussed in Example 8, Chapter 1, it turned out to be the case that the parts produced on that

Table 3.14
Means and Ranges for a Critical Dimension on Samples of $n = 5$ Parts

| Sample	Date	Time		$\bar{x}$	R	Sample	Date	Time		$\bar{x}$	R
1	10/27	7:30	AM	3509.4	5	14		10:15		3504.4	4
2		8:30		3509.2	2	15		11:15		3504.6	3
3		9:30		3512.6	3	16	10/28	7:30	AM	3513.0	2
4		10:30		3511.6	4	17		8:30		3512.4	1
5		11:30		3512.0	4	18		9:30		3510.8	5
6		12:30	PM	3513.6	6	19		10:30		3511.8	4
7		1:30		3511.8	3	20		6:15	PM	3512.4	3
8		2:30		3512.2	2	21		7:15		3511.0	4
9		4:15		3500.0	3	22		8:45		3510.6	1
10		5:45		3502.0	2	23		9:45		3510.2	5
11		6:45		3501.4	2	24		10:45		3510.4	2
12		8:15		3504.0	2	25		11:45		3510.8	3
13		9:15		3503.6	3						

Example 8
(*continued*)
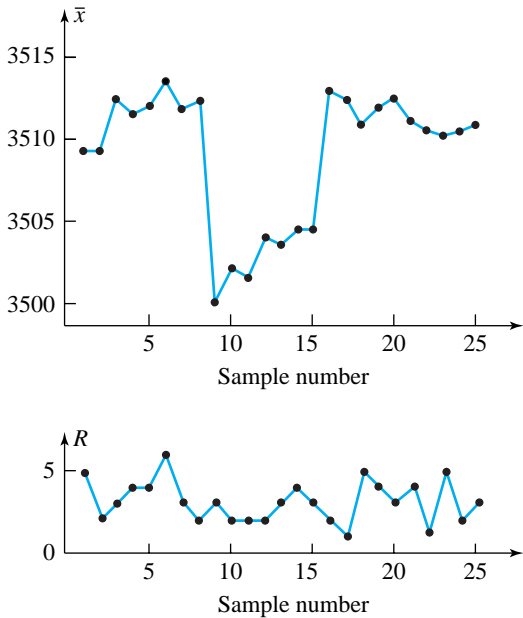
Figure 3.21   Plots of $\bar{x}$ and R over time

shift were not really systematically any different from the others. Instead, the person making the measurements for samples 9 through 15 used the gauge in a fundamentally different way than other employees. The pattern in the $\bar{x}$ values was caused by this change in measurement technique.

Terminology and causes for patterns on plots against Time

    Patterns revealed in the plotting of sample statistics against time ought to alert an engineer to look for a physical cause and (typically) a cure. **Systematic variations** or **cycles** in a plot of means can often be related to process variables that come and go on a more or less regular basis. Examples include seasonal or daily variables like ambient temperature or those caused by rotation of gauges or fixtures. **Instability** or variation in excess of that related to basic equipment precision can sometimes be traced to mixed lots of raw material or overadjustment of equipment by operators. **Changes in level** of a process mean can originate in the introduction of new machinery, raw materials, or employee training and (for example) tool wear. **Mixtures** of several patterns of variation on a single plot of some summary statistic against time can sometimes (as in Example 8) be traced to changes in measurement calibration. They are also sometimes produced by consistent differences in machines or streams of raw material.

Plots against process variables

    Plots of summary statistics against time are not the only useful ones. Plots against process variables can also be quite informative.

Example 9
(*Example 6, Chapter 1, revisited—p. 15*)

Plotting the Mean Shear Strength of Wood Joints

In their study of glued wood joint strength, Dimond and Dix obtained the values given in Table 3.15 as mean strengths (over three shear tests) for each combination of three woods and three glues. Figure 3.22 gives a revealing plot of these $3 \times 3 = 9$ different $\bar{x}$'s.

Table 3.15
Mean Joint Strengths for Nine Wood/Glue Combinations

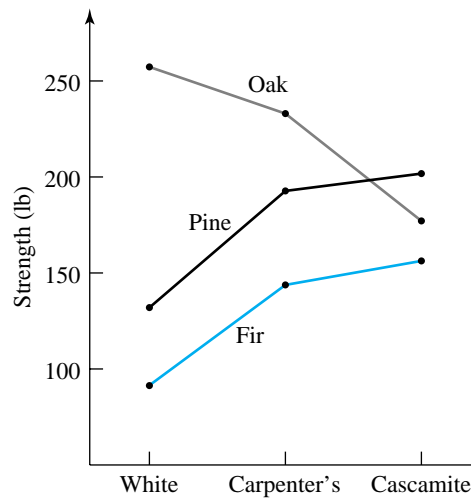| Wood	Glue	$\bar{x}$ Mean Joint Shear Strength (lb)
pine	white	131.7
pine	carpenter's	192.7
pine	cascamite	201.3
fir	white	92.0
fir	carpenter's	146.3
fir	cascamite	156.7
oak	white	257.7
oak	carpenter's	234.3
oak	cascamite	177.7

Example 9
(*continued*)

Figure 3.22   Plot of mean joint strength vs. glue type for three woods

From the plot, it is obvious that the gluing properties of pine and fir are quite similar, with pine joints averaging around 40–45 lb stronger. For these two soft woods, cascamite appears slightly better than carpenter's glue, both of which make much better joints than white glue. The gluing properties of oak (a hardwood) are quite different from those of pine and fir. In fact, the glues perform in exactly the opposite ordering for the strength of oak joints. All of this is displayed quite clearly by the simple plot in Figure 3.22.

The two previous examples have illustrated the usefulness of plotting sample statistics against time and against levels of an experimental variable. Other possibilities in specific engineering situations can potentially help the working engineer understand and manipulate the systems on which he or she works.

3.3.5   Summary Statistics and Personal Computer Software

The numerical data summaries introduced in this chapter are relatively simple. For small data sets they can be computed quite easily using only a pocket calculator. However, for large data sets and in cases where subsequent additional calculations or plotting may occur, statistical or spreadsheet software can be convenient.

Printout 1 illustrates the use of the MINITAB statistical package to produce summary statistics for the percent waste data sets in Table 3.13. (The appropriate MINITAB routine is found under the "Stat/Basic Statistics/Display Descriptive Statistics" menu.) The mean, median, and standard deviation values on the printout agree with those produced in Example 7. However, the first and third quartile

Printout 1   Descriptive Statistics for the Percent Waste Data of Table

```
Descriptive Statistics


Variable          N       Mean     Median     TrMean      StDev    SE Mean
Supply 1          6      1.495      0.785      1.495      1.394      0.569
Supply 2          8      2.228      1.525      2.228      1.839      0.650

Variable    Minimum    Maximum         Q1         Q3
Supply 1      0.370      3.620      0.483      3.073
Supply 2      0.890      6.540      1.105      2.540
```

figures on the printout do not match exactly those found earlier. MINITAB simply uses slightly different conventions for those quantities than the ones introduced in Section 3.2.

High-quality statistical packages like MINITAB (and JMP, SAS, SPSS, SYSTAT, SPLUS, etc.) are widely available. One of them should be on the electronic desktop of every working engineer. Unfortunately, this is not always the case, and engineers often assume that standard spreadsheet software (perhaps augmented with third party plug-ins) provides a workable substitute. Often this is true, but sometimes it is not.

The primary potential problem with using a spreadsheet as a substitute for statistical software concerns numerical accuracy. Spreadsheets can and do on occasion return catastrophically wrong values for even simple statistics. Established vendors of statistical software have many years of experience dealing with subtle numerical issues that arise in the computation of even simple summaries of even small data sets. Most vendors of spreadsheet software seem unaware of or indifferent to these matters. For example, consider the very small data set

$$0, 1, 2$$

The sample variance of these data is easily seen to be 1.0, and essentially any statistical package or spreadsheet will reliably return this value. However, suppose 100,000,000 is added to each of these $n = 3$ values, producing the data set

$$100000000, 100000001, 100000002$$

The actual sample variance is unchanged, and high-quality statistical software will reliably return the value 1.0. However, as of late 1999, the current version of the leading spreadsheet program returned the value 0 for this second sample variance. This is a badly wrong answer to an apparently very simple problem.

So at least until vendors of spreadsheet software choose to integrate an established statistical package into their products, we advise extreme caution in the use of spreadsheets to do statistical computations. A good source of up-to-date information on this issue is the AP Statistics electronic bulletin board found at http://forum.swarthmore.edu/epigone/apstat-l.

Section 3 Exercises ....

1. Calculate and compare the means, medians, ranges, interquartile ranges, and standard deviations of the two data sets introduced in Exercise 1 of Section 3.2. Discuss the interpretation of these values in the context of comparing the two cutting methods.

2. Are the numerical values you produced in Exercise 1 above most naturally thought of as statistics or as parameters? Explain.

3. Use a statistical package to compute basic summary statistics for the two data sets introduced in

Exercise 1 of Section 3.2 and thereby check your answers to Exercise 1 here.

4. Add 1.3 to each of the lengthwise cut impact strengths referred to in Exercise 1 and then recompute the values of the mean, median, range, interquartile range, and standard deviation. How do these compare with the values obtained earlier? Repeat this exercise after multiplying each lengthwise cut impact strength by 2 (instead of adding 1.3).

....

3.4   Descriptive Statistics for Qualitative and Count Data (*Optional*)

The techniques presented thus far in this chapter are primarily relevant to the analysis of measurement data. As noted in Section 1.2, conventional wisdom is that where they can be obtained, measurement data (or *variables data*) are generally preferable to count and qualitative data (or *attributes data*). Nevertheless, qualitative or count data will sometimes be the primary information available. It is therefore worthwhile to consider their summarization.

This section will cover the reduction of qualitative and count data to per-item or per-inspection-unit figures and the display of those ratios in simple bar charts and plots.

3.4.1   Numerical Summarization of Qualitative and Count Data

Recall from Definitions 8 and 9 in Chapter 1 that aggregation and counting are typically used to produce numerical values from qualitative data. Then, beginning with counts, it is often helpful to calculate rates on a per-item or per-inspection-unit basis.

When each item in a sample of n either does or does not have a characteristic of interest, the notation

Sample fraction of items with a characteristic

$$\hat{p} = \frac{\text{The number of items in the sample with the characteristic}}{n} \qquad (3.6)$$

will be used. A given sample can produce many such values of "p hat" if either a single characteristic has many possible categories or many different characteristics are being monitored simultaneously.

Example 10

Defect Classifications of Cable Connectors

Delva, Lynch, and Stephany worked with a manufacturer of cable connectors. Daily samples of 100 connectors of a certain design were taken over 30 production days, and each sampled connector was inspected according to a well-defined (operational) set of rules. Using the information from the inspections, each inspected connector could be classified as belonging to one of the following five mutually exclusive categories:

Category A: having "very serious" defects

Category B: having "serious" defects but no "very serious" defects

Category C: having "moderately serious" defects but no "serious" or "very serious" defects

Category D: having only "minor" defects

Category E: having no defects

Table 3.16 gives counts of sampled connectors falling into the first four categories (the four defect categories) over the 30-day period. Then, using the fact that $30 \times 100 = 3{,}000$ connectors were inspected over this period,

$$\hat{p}_A = 3/3000 = .0010$$
$$\hat{p}_B = 0/3000 = .0000$$
$$\hat{p}_C = 11/3000 = .0037$$
$$\hat{p}_D = 1/3000 = .0003$$

Notice that here $\hat{p}_E = 1 - (\hat{p}_A + \hat{p}_B + \hat{p}_C + \hat{p}_D)$, because categories A through E represent a set of nonoverlapping and exhaustive classifications into which an individual connector must fall, so that the $\hat{p}$'s must total to 1.

Table 3.16
Counts of Connectors Classified into Four Defect Categories

Category	Number of Sampled Connectors
A	3
B	0
C	11
D	1

Example 11

Pneumatic Tool Manufacture

Kraber, Rucker, and Williams worked with a manufacturer of pneumatic tools. Each tool produced is thoroughly inspected before shipping. The students collected some data on several kinds of problems uncovered at final inspection. Table 3.17 gives counts of tools having these problems in a particular production run of 100 tools.

Table 3.17
Counts and Fractions of Tools with Various Problems

Problem	Number of Tools	$\hat{p}$
Type 1 leak	8	.08
Type 2 leak	4	.04
Type 3 leak	3	.03
Missing part 1	2	.02
Missing part 2	1	.01
Missing part 3	2	.02
Bad part 4	1	.01
Bad part 5	2	.02
Bad part 6	1	.01
Wrong part 7	2	.02
Wrong part 8	2	.02

Table 3.17 is a summarization of highly multivariate qualitative data. The categories listed in Table 3.17 are not mutually exclusive; a given tool can fall into more than one of them. Instead of representing different possible values of a single categorical variable (as was the case with the connector categories in Example 10), the categories listed above each amount to 1 (present) of 2 (present and not present) possible values for a different categorical variable. For example, for type 1 leaks, $\hat{p} = .08$, so $1 - \hat{p} = .92$ for the fraction of tools *without* type 1 leaks. The $\hat{p}$ values do not necessarily total to the fraction of tools requiring rework at final inspection. A given faulty tool could be counted in several $\hat{p}$ values.

Another kind of per-item ratio, also based on counts, is sometimes confused with $\hat{p}$. Such a ratio arises when every item in a sample provides an opportunity for a phenomenon of interest to occur, but multiple occurrences are possible and counts are kept of the total number of occurrences. In such cases, the notation

Sample mean occurences per unit or item

$$\hat{u} = \frac{\text{The total number of occurrences}}{\text{The total number of inspection units or sampled items}} \quad \text{(3.7)}$$

is used. $\hat{u}$ is really closer in meaning to $\bar{x}$ than to $\hat{p}$, even though it can turn out to be a number between 0 and 1 and is sometimes expressed as a percentage and called a *rate*.

Although the counts totaled in the numerator of expression (3.7) must all be integers, the values totaled to create the denominator need not be. For instance, suppose vinyl floor tiles are being inspected for serious blemishes. If on one occasion inspection of 1 box yields a total of 2 blemishes, on another occasion .5 box yields 0 blemishes, and on still another occasion 2.5 boxes yield a total of 1 blemish, then

$$\hat{u} = \frac{2 + 0 + 1}{1 + .5 + 2.5} = .75 \text{ blemishes/box}$$

Depending on exactly how terms are defined, it may be appropriate to calculate either $\hat{p}$ values or $\hat{u}$ values or both in a single situation.

Example 10
(continued)

It was possible for a single cable connector to have more than one defect of a given severity and, in fact, defects of different severities. For example, Delva, Lynch, and Stephany's records indicate that in the 3,000 connectors inspected, 1 connector had exactly 2 moderately serious defects (along with a single very serious defect), 11 connectors had exactly 1 moderately serious defect (and no others), and 2,988 had no moderately serious defects. So the observed rate of moderately serious defects could be reported as

$$\hat{u} = \frac{2 + 11}{1 + 11 + 2988} = .0043 \text{ moderately serious defects/connector}$$

This is an occurrence rate for moderately serious defects($\hat{u}$), but not a fraction of connectors having moderately serious defects ($\hat{p}$).

The difference between the statistics $\hat{p}$ and $\hat{u}$ may seem trivial. But it is a point that constantly causes students confusion. Methods of formal statistical inference based on $\hat{p}$ are not the same as those based on $\hat{u}$. The distinction between the two kinds of rates must be kept in mind if those methods are to be applied appropriately.

To carry this warning a step further, note that not every quantity called a *percentage* is even of the form $\hat{p}$ or $\hat{u}$. In a laboratory analysis, a specimen may be declared to be "30% carbon." The 30% cannot be thought of as having the form of $\hat{p}$ in equation (3.6) or $\hat{u}$ in equation (3.7). It is really a single continuous measurement, not a summary statistic. Statistical methods for $\hat{p}$ or $\hat{u}$ have nothing to say about such rates.

3.4.2 Bar Charts and Plots for Qualitative and Count Data

Often, a study will produce several values of $\hat{p}$ or $\hat{u}$ that need to be compared. Bar charts and simple bivariate plots can be a great aid in summarizing these results.

Example 10
(*continued*)

Figure 3.23 is a bar chart of the fractions of connectors in the categories A through D. It shows clearly that most connectors with defects fall into category C, having moderately serious defects but no serious or very serious defects. This bar chart is a presentation of the behavior of a single categorical variable.
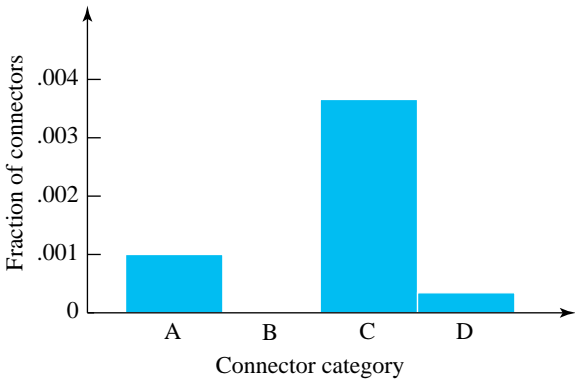
Figure 3.23   Bar chart of connector defects

Example 11
(*continued*)

Figure 3.24 is a bar chart of the information on tool problems in Table 3.17. It shows leaks to be the most frequently occurring problems on this production run.
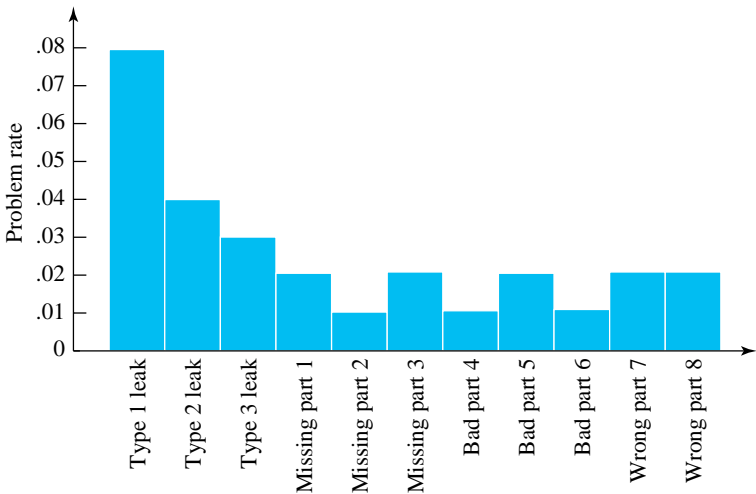
Figure 3.24   Bar chart for assembly problems

Figures 3.23 and 3.24 are both bar charts, but they differ considerably. The first concerns the behavior of a single (ordered) categorical variable—namely, Connector Class. The second concerns the behavior of 11 different present–not present categorical variables, like Type 1 Leak, Missing Part 3, etc. There may be some significance to the shape of Figure 3.23, since categories A through D are arranged in decreasing order of defect severity, and this order was used in the making of the figure. But the shape of Figure 3.24 is essentially arbitrary, since the particular ordering of the tool problem categories used to make the figure is arbitrary. Other equally sensible orderings would give quite different shapes.

The device of **segmenting bars** on a bar chart and letting the segments stand for different categories of a single qualitative variable can be helpful, particularly where several different samples are to be compared.

Example 12

Scrap and Rework in a Turning Operation

The article "Statistical Analysis: Roadmap for Management Action" by H. Rowe (*Quality Progress*, February 1985) describes a statistically based quality-improvement project in the turning of steel shafts. Table 3.18 gives the percentages of reworkable and scrap shafts produced in 18 production runs made during the study.

Figure 3.25 is a corresponding segmented bar graph, with the jobs ordered in time, showing the behavior of both the scrap and rework rates over time. (The total height of any bar represents the sum of the two rates.) The sharp reduction in both scrap and rework between jobs 10 and 11 was produced by overhauling one of the company's lathes. That lathe was identified as needing attention through engineering data analysis early in the plant project.

Table 3.18
Percents Scrap and Rework in a Turning Operation

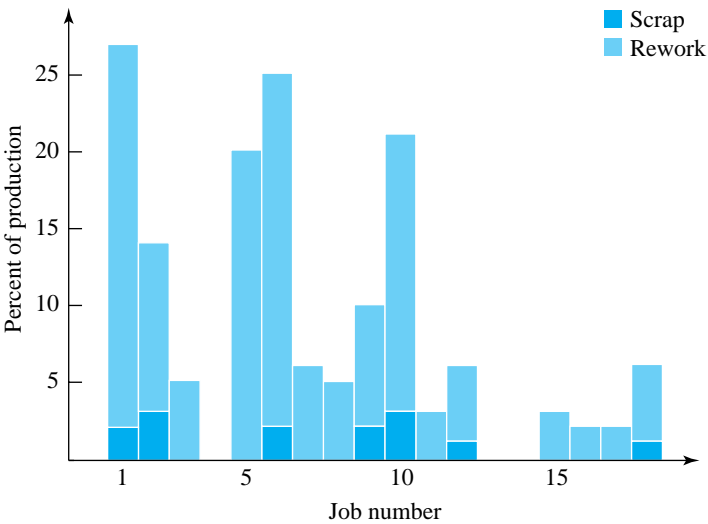| Job Number	Percent Scrap	Percent Rework	Job Number	Percent Scrap	Percent Rework
1	2	25	10	3	18
2	3	11	11	0	3
3	0	5	12	1	5
4	0	0	13	0	0
5	0	20	14	0	0
6	2	23	15	0	3
7	0	6	16	0	2
8	0	5	17	0	2
9	2	8	18	1	5

Example 12
(continued)

Figure 3.25   Segmented bar chart of scrap and rework rates

In many cases, the simple plotting of $\hat{p}$ or $\hat{u}$ values against time or process variables can make clear the essential message in a set of qualitative or count data.

Example 13

Defects per Truck Found at Final Inspection

In his text *Engineering Statistics and Quality Control*, I. W. Burr illustrates the usefulness of plotting $\hat{u}$ versus time with a set of data on defects found at final inspection at a truck assembly plant. From 95 to 130 trucks were produced daily at the plant; Table 3.19 gives part of Burr's daily defects/truck values. These statistics are plotted in Figure 3.26. The graph shows a marked decrease in quality (increase in $\hat{u}$) over the third and fourth weeks of December, ending with a rate

Table 3.19
Defects Per Truck on 26 Production Days

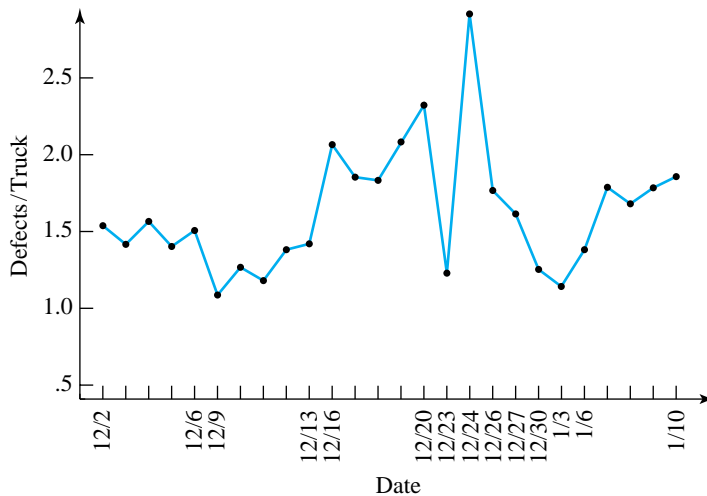| Date	$\hat{u}$ = Defects/Truck	Date	$\hat{u}$ = Defects/Truck	Date	$\hat{u}$ = Defects/Truck	Date	$\hat{u}$ = Defects/Truck
12/2	1.54	12/11	1.18	12/20	2.32	1/3	1.15
12/3	1.42	12/12	1.39	12/23	1.23	1/6	1.37
12/4	1.57	12/13	1.42	12/24	2.91	1/7	1.79
12/5	1.40	12/16	2.08	12/26	1.77	1/8	1.68
12/6	1.51	12/17	1.85	12/27	1.61	1/9	1.78
12/9	1.08	12/18	1.82	12/30	1.25	1/10	1.84
12/10	1.27	12/19	2.07				

Figure 3.26   Plot of daily defects per truck

of 2.91 defects/truck on Christmas Eve. Apparently, this situation was largely corrected with the passing of the holiday season.

Plots of $\hat{p}$ or $\hat{u}$ against levels of manipulated variables from an experiment are often helpful in understanding the results of that experiment.

Example 14

### Plotting Fractions of Conforming Pellets

Greiner, Grim, Larson, and Lukomski experimented with the same pelletizing machine studied by Cyr, Ellson, and Rickard (see Example 2 in Chapter 1). In one part of their study, they ran the machine at an elevated speed and varied the shot size (amount of powder injected into the dies) and the composition of that powder (in terms of the relative amounts of new and reground material). Table 3.20 lists the numbers of conforming pellets produced in a sample of 100 at each of $2 \times 2 = 4$ sets of process conditions. A simple plot of $\hat{p}$ values versus shot size is given in Figure 3.27.

The figure indicates that increasing the shot size is somewhat harmful, but that a substantial improvement in process performance happens when the amount of reground material used in the pellet-making mixture is increased. This makes sense. Reground material had been previously compressed into (nonconforming) pellets. In the process, it had been allowed to absorb some ambient humidity. Both the prior compression and the increased moisture content were potential reasons why this material improved the ability of the process to produce solid, properly shaped pellets.

Example 14
(*continued*)

Table 3.20
Numbers of Conforming Pellets for Four Shot Size/Mixture Combinations

Sample	Shot Size	Mixture	Number Conforming
1	small	20% reground	38
2	small	50% reground	66
3	large	20% reground	29
4	large	50% reground	53
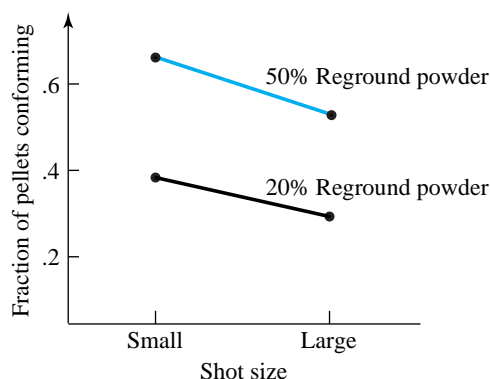
Figure 3.27   Plot of fraction conforming vs. shot size

Section 4 Exercises

1. From your field, give an example of a variable that is a rate (a) of the form $\hat{p}$, (b) of the form $\hat{u}$, and (c) of neither form.

2. Because gauging is easier, it is sometimes tempting to collect qualitative data related to measurements rather than the measurements themselves. For example, in the context of Example 1 in Chapter 1, if gears with runouts exceeding 15 were considered to be nonconforming, it would be possible to derive fractions nonconforming, $\hat{p}$, from simple "go–no go" checking of gears. For the two sets of gears represented in Table 1.1, what would have been the sample fractions nonconforming $\hat{p}$? Give a practical reason why having the values in Table 1.1 might be preferable to knowing only the corresponding $\hat{p}$ values.

3. Consider the measurement of the percentage copper in brass specimens. The resulting data will be a kind of rate data. Are the rates that will be obtained of the type $\hat{p}$, of the type $\hat{u}$, or of neither type? Explain.

Chapter 3 Exercises ....

1. The accompanying values are gains measured on 120 amplifiers designed to produce a 10 dB gain. These data were originally from the *Quality Improvement Tools* workbook set (published by the Juran Institute). They were then used as an example in the article "The Tools of Quality" (*Quality Progress*, September 1990).

8.1, 10.4, 8.8, 9.7, 7.8, 9.9, 11.7, 8.0, 9.3, 9.0, 8.2, 8.9, 10.1, 9.4, 9.2, 7.9, 9.5, 10.9, 7.8, 8.3, 9.1, 8.4, 9.6, 11.1, 7.9, 8.5, 8.7, 7.8, 10.5, 8.5, 11.5, 8.0, 7.9, 8.3, 8.7, 10.0, 9.4, 9.0, 9.2, 10.7, 9.3, 9.7, 8.7, 8.2, 8.9, 8.6, 9.5, 9.4, 8.8, 8.3, 8.4, 9.1, 10.1, 7.8, 8.1, 8.8, 8.0, 9.2, 8.4, 7.8, 7.9, 8.5, 9.2, 8.7, 10.2, 7.9, 9.8, 8.3, 9.0, 9.6, 9.9, 10.6, 8.6, 9.4, 8.8, 8.2, 10.5, 9.7, 9.1, 8.0, 8.7, 9.8, 8.5, 8.9, 9.1, 8.4, 8.1, 9.5, 8.7, 9.3, 8.1, 10.1, 9.6, 8.3, 8.0, 9.8, 9.0, 8.9, 8.1, 9.7, 8.5, 8.2, 9.0, 10.2, 9.5, 8.3, 8.9, 9.1, 10.3, 8.4, 8.6, 9.2, 8.5, 9.6, 9.0, 10.7, 8.6, 10.0, 8.8, 8.6

(a) Make a stem-and-leaf plot and a boxplot for these data. How would you describe the shape of this data set? Does the shape of your stem-and-leaf plot (or a corresponding histogram) give you any clue how a high fraction within specifications was achieved?

(b) Make a normal plot for these data and interpret its shape. (Standard normal quantiles for $p = .0042$ and $p = .9958$ are approximately -2.64 and 2.64, respectively.)

(c) Although the nominal gain for these amplifiers was to be 10 dB, the design allowed gains from 7.75 dB to 12.2 dB to be considered acceptable. About what fraction, p, of such amplifiers do you expect to meet these engineering specifications?

2. The article "The Lognormal Distribution for Modeling Quality Data When the Mean is Near Zero" by S. Albin (*Journal of Quality Technology*, April 1990) described the operation of a Rutgers University plastics recycling pilot plant. The most important material reclaimed from beverage bottles is PET plastic. A serious impurity is aluminum, which later can clog the filters in extruders when the recycled material is used. The following are the amounts (in ppm by weight of aluminum) found in bihourly samples of PET recovered at the plant over roughly a two-day period.

291, 222, 125, 79, 145, 119, 244, 118, 182, 63, 30, 140, 101, 102, 87, 183, 60, 191, 119, 511, 120, 172, 70, 30, 90, 115

(Apparently, the data are recorded in the order in which they were collected, reading left to right, top to bottom.)

(a) Make a run chart for these data. Are there any obvious time trends? What practical engineering reason is there for looking for such trends?

(b) Ignoring the time order information, make a stem-and-leaf diagram. Use the hundreds digit to make the stem and the other two digits (separated by commas to indicate the different data points) to make the leaves. After making an initial stem-and-leaf diagram by recording the data in the (time) order given above, make a second one in which the values have been ordered.

(c) How would you describe the shape of the stem-and-leaf diagram? Is the data set bell-shaped?

(d) Find the median and the first and third quartiles for the aluminum contents and then find the .58 quantile of the data set.

(e) Make a boxplot.

(f) Make a normal plot, using regular graph paper. List the coordinates of the 26 plotted points. Interpret the shape of the plot.

(g) Try transforming the data by taking natural logarithms and again assess the shape. Is the transformed data set more bell-shaped than the raw data set?

(h) Find the sample mean, the sample range, and the sample standard deviation for both the original data and the log-transformed values from (g). Is the mean of the transformed values equal to the natural logarithm of the mean of the original data?

3. The accompanying data are three hypothetical samples of size 10 that are supposed to represent measured manganese contents in specimens of 1045 steel (the units are points, or .01%). Suppose that these measurements were made on standard specimens having "true" manganese contents of 80, using three different analytical methods. (Thirty different specimens were involved.)

Method 1
87, 74, 78, 81, 78,
77, 84, 80, 85, 78

Method 2
86, 85, 82, 87, 85,
84, 84, 82, 82, 85

Method 3
84, 83, 78, 79, 85,
82, 82, 81, 82, 79

(a) Make (on the same coordinate system) side-by-side boxplots that you can use to compare the three analytical methods.

(b) Discuss the apparent effectiveness of the three methods in terms of the appearance of your diagram from (a) and in terms of the concepts of accuracy and precision discussed in Section 1.3.

(c) An alternative method of comparing two such analytical methods is to use both methods of analysis once on each of (say) 10 different specimens (10 specimens and 20 measurements). In the terminology of Section 1.2, what kind of data would be generated by such a plan? If one simply wishes to compare the average measurements produced by two analytical methods, which data collection plan (20 specimens and 20 measurements, or 10 specimens and 20 measurements) seems to you most likely to provide the better comparison? Explain.

4. Gaul, Phan, and Shimonek measured the resistances of 15 resistors of $2 \times 5 = 10$ different types. Two different wattage ratings were involved, and five different nominal resistances were used. All measurements were reported to three significant digits. Their data follow.

(a) Make back-to-back stem-and-leaf plots for comparing the $\frac{1}{4}$ watt and $\frac{1}{2}$ watt resistance distributions for each nominal resistance. In a few sentences, summarize what these show.

(b) Make pairs of boxplots for comparing the $\frac{1}{4}$ watt and $\frac{1}{2}$ watt resistance distributions for each nominal resistance.

(c) Make normal plots for the $\frac{1}{2}$ watt nominal 20 ohm and nominal 200 ohm resistors. Interpret these in a sentence or two. From the appearance of the second plot, does it seem that if the nominal 200 ohm resistances were treated as if they had a bell-shaped distribution, the tendency would be to overestimate or to underestimate the fraction of resistances near the nominal value?

$\frac{1}{4}$ Watt Resistors				
20 ohm	*75 ohm*	*100 ohm*	*150 ohm*	*200 ohm*
19.2	72.9	97.4	148	198
19.2	72.4	95.8	148	196
19.3	72.0	97.7	148	199
19.3	72.5	94.1	148	196
19.1	72.7	95.1	148	196
19.0	72.3	95.4	147	195
19.6	72.9	94.9	148	193
19.2	73.2	98.5	148	196
19.3	71.8	94.8	148	196
19.4	73.4	94.6	147	199
19.4	70.9	98.3	147	194
19.3	72.3	96.0	149	195
19.5	72.5	97.3	148	196
19.2	72.1	96.0	148	195
19.1	72.6	94.8	148	199

$\frac{1}{2}$ Watt Resistors				
20 ohm	*75 ohm*	*100 ohm*	*150 ohm*	*200 ohm*
20.1	73.9	97.2	152	207
19.7	74.2	97.9	151	205
20.2	74.6	96.8	155	214
24.4	72.1	99.2	146	195
20.2	73.8	98.5	148	202
20.1	74.8	95.5	154	211
20.0	75.0	97.2	149	197
20.4	68.6	98.7	150	197
20.3	74.0	96.6	153	199
20.6	71.7	102	149	196
19.9	76.5	103	150	207
19.7	76.2	102	149	210
20.8	72.8	102	145	192
20.4	73.2	100	147	201
20.5	76.7	100	149	257

(d) Compute the sample means and sample standard deviations for all 10 samples. Do these values agree with your qualitative statements made in answer to part (a)?

(e) Make a plot of the 10 sample means computed in part (d), similar to the plot in Figure 3.22. Comment on the appearance of this plot.

5. Blomquist, Kennedy, and Reiter studied the properties of three scales by each weighing a standard 5 g weight, 20 g weight, and 100 g weight twice on each scale. Their data are presented in the accompanying table. Using whatever graphical and numerical data summary methods you find helpful, make sense of these data. Write a several-page discussion of your findings. You will probably want to consider both accuracy and precision and (to the extent possible) make comparisons between scales and between students. Part of your discussion might deal with the concepts of repeatability and reproducibility introduced in Section 2.1. Are the pictures you get of the scale and student performances consistent across the different weights?

5-Gram Weighings		
Scale 1	*Scale 2*	*Scale 3*

	Scale 1	*Scale 2*	*Scale 3*
Student 1	5.03, 5.02	5.07, 5.09	4.98, 4.98
Student 2	5.03, 5.01	5.02, 5.07	4.99, 4.98
Student 3	5.06, 5.00	5.10, 5.08	4.98, 4.98

20-Gram Weighings		
Scale 1	*Scale 2*	*Scale 3*

	Scale 1	*Scale 2*	*Scale 3*
Student 1	20.04, 20.06	20.04, 20.04	19.94, 19.93
Student 2	20.02, 19.99	20.03, 19.93	19.95, 19.95
Student 3	20.03, 20.02	20.06, 20.03	19.91, 19.96

100-Gram Weighings		
Scale 1	*Scale 2*	*Scale 3*

	Scale 1	*Scale 2*	*Scale 3*
Student 1	100.06, 100.35	100.25, 100.08	99.87, 99.88
Student 2	100.05, 100.01	100.10, 100.02	99.87, 99.88
Student 3	100.00, 100.00	100.01, 100.02	99.88, 99.88

6. The accompanying values are the lifetimes (in numbers of 24 mm deep holes drilled in 1045 steel before tool failure) for $n = 12$ D952-II (8 mm) drills. These were read from a graph in "Computer-assisted Prediction of Drill-failure Using In-process Measurements of Thrust Force" by A. Thangaraj and P. K. Wright (*Journal of Engineering for Industry*, May 1988).

> 47, 145, 172, 86, 122, 110, 172, 52, 194, 116, 149, 48

Write a short report to your engineering manager summarizing what these data indicate about the lifetimes of drills of this type in this kind of application. Use whatever graphical and numerical data summary tools make clear the main features of the data set.

7. Losen, Cahoy, and Lewis purchased eight spanner bushings of a particular type from a local machine shop and measured a number of characteristics of these bushings, including their outside diameters. Each of the eight outside diameters was measured

once by each of two student technicians, with the following results (the units are inches):

Bushing	1	2	3	4
Student A	.3690	.3690	.3690	.3700
Student B	.3690	.3695	.3695	.3695

Bushing	5	6	7	8
Student A	.3695	.3700	.3695	.3690
Student B	.3695	.3700	.3700	.3690

A common device when dealing with paired data like these is to analyze the differences. Subtracting B measurements from A measurements gives the following eight values:

.0000, −.0005, −.0005, .0005, .0000, .0000, −.0005, .0000

(a) Find the first and third quartiles for these differences, and their median.
(b) Find the sample mean and standard deviation for the differences.
(c) Your mean in part (b) should be negative. Interpret this in terms of the original measurement problem.
(d) Suppose you want to make a normal plot of the differences on regular graph paper. Give the coordinates of the lower-left point on such a plot.

8. The accompanying data are the times to failure (in millions of cycles) of high-speed turbine engine bearings made out of two different compounds. These were taken from "Analysis of Single Classification Experiments Based on Censored Samples from the Two-parameter Weibull Distribution" by J. I. McCool (*The Journal of Statistical Planning and Inference*, 1979).

Compound 1
3.03, 5.53, 5.60, 9.30, 9.92,
12.51, 12.95, 15.21, 16.04, 16.84

Compound 2
3.19, 4.26, 4.47, 4.53, 4.67,
4.69, 5.78, 6.79, 9.37, 12.75

(a) Find the .84 quantile of the Compound 1 failure times.
(b) Give the coordinates of the two lower-left points that would appear on a normal plot of the Compound 1 data.
(c) Make back-to-back stem-and-leaf plots for comparing the life length properties of bearings made from Compounds 1 and 2.
(d) Make (to scale) side-by-side boxplots for comparing the life lengths for the two compounds. Mark numbers on the plots indicating the locations of their main features.
(e) Compute the sample means and standard deviations of the two sets of lifetimes.
(f) Describe what your answers to parts (c), (d), and (e) above indicate about the life lengths of these turbine bearings.

9. Heyde, Kuebrick, and Swanson measured the heights of 405 steel punches purchased by a company from a single supplier. The stamping machine in which these are used is designed to use .500 in. punches. Frequencies of the measurements they obtained are shown in the accompanying table.

Punch Height (.001 in.)	Frequency	Punch Height (.001 in.)	Frequency
482	1	496	7
483	0	497	13
484	1	498	24
485	1	499	56
486	0	500	82
487	1	501	97
488	0	502	64
489	1	503	43
490	0	504	3
491	2	505	1
492	0	506	0
493	0	507	0
494	0	508	0
495	6	509	2

(a) Summarize these data, using appropriate graphical and numerical tools. How would you describe the shape of the distribution of punch heights? The specifications for punch heights were in fact .500 in. to .505 in. Does this fact give you any insight as to the origin of the distributional shape observed in the data? Does it appear that the supplier has equipment capable of meeting the engineering specifications on punch height?

(b) In the manufacturing application of these punches, several had to be placed side-by-side on a drum to cut the same piece of material. In this context, why is having small variability in punch height perhaps even more important than having the correct mean punch height?

10. The article "Watch Out for Nonnormal Distributions" by D. C. Jacobs (*Chemical Engineering Progress*, November 1990) contains 100 measured daily purities of oxygen delivered by a single supplier. These are as follows, listed in the time order of their collection (read left to right, top to bottom). The values given are in hundredths of a percent purity above 99.00% (so 63 stands for 99.63%).

63, 61, 67, 58, 55, 50, 55, 56, 52, 64, 73, 57, 63,
81, 64, 54, 57, 59, 60, 68, 58, 57, 67, 56, 66, 60,
49, 79, 60, 62, 60, 49, 62, 56, 69, 75, 52, 56, 61,
58, 66, 67, 56, 55, 66, 55, 69, 60, 69, 70, 65, 56,
73, 65, 68, 59, 62, 58, 62, 66, 57, 60, 66, 54, 64,
62, 64, 64, 50, 50, 72, 85, 68, 58, 68, 80, 60, 60,
53, 49, 55, 80, 64, 59, 53, 73, 55, 54, 60, 60, 58,
50, 53, 48, 78, 72, 51, 60, 49, 67

You will probably want to use a statistical analysis package to help you do the following:

(a) Make a run chart for these data. Are there any obvious time trends? What would be the practical engineering usefulness of early detection of any such time trend?

(b) Now ignore the time order of data collection and represent these data with a stem-and-leaf plot and a histogram. (Use .02% class widths in making your histogram.) Mark on these the supplier's lower specification limit of 99.50%

purity. Describe the shape of the purity distribution.

(c) The author of the article found it useful to reexpress the purities by subtracting 99.30 (remember that the preceding values are in units of .01% above 99.00%) and then taking natural logarithms. Do this with the raw data and make a second stem-and-leaf diagram and a second histogram to portray the shape of the transformed data. Do these figures look more bell-shaped than the ones you made in part (b)?

(d) Make a normal plot for the transformed values from part (c). What does it indicate about the shape of the distribution of the transformed values? (Standard normal quantiles for $p = .005$ and $p = .995$ are approximately -2.58 and 2.58, respectively.)

11. The following are some data taken from the article "Confidence Limits for Weibull Regression with Censored Data" by J. I. McCool (*IEEE Transactions on Reliability*, 1980). They are the ordered failure times (the time units are not given in the paper) for hardened steel specimens subjected to rolling contact fatigue tests at four different values of contact stress.

$.87 \times 10^6$ psi	$.99 \times 10^6$ psi	1.09×10^6 psi	1.18×10^6 psi
1.67	.80	.012	.073
2.20	1.00	.18	.098
2.51	1.37	.20	.117
3.00	2.25	.24	.135
3.90	2.95	.26	.175
4.70	3.70	.32	.262
7.53	6.07	.32	.270
14.7	6.65	.42	.350
27.8	7.05	.44	.386
37.4	7.37	.88	.456

(a) Make side-by-side boxplots for these data. Does it look as if the different stress levels produce life distributions of roughly the same shape? (Engineering experience suggests that

different stress levels often change the scale but not the basic shape of life distributions.)

(b) Make Q-Q plots for comparing all six different possible pairs of distributional shapes. Summarize in a few sentences what these indicate about the shapes of the failure time distributions under the different stress levels.

12. Riddle, Peterson, and Harper studied the performance of a rapid-cut industrial shear in a continuous cut mode. They cut nominally 2-in. and 1-in. strips of 14 gauge and 16 gauge steel sheet metal and measured the actual widths of the strips produced by the shear. Their data follow, in units of 10^{-3} in. above nominal.

| | | Material Thickness | |
		14 Gauge	16 Gauge
	1 in.	2, 1, 1, 1, 0, 0, −2, −10, −5, 1	−2, −6, −1, −2, −1, −2, −1, −1, −1, −5
Machine Setting			
	2 in.	10, 10, 8, 8, 8, 8, 7, 7, 9, 11	−4, −3, −4, −2, −3, −3, −3, −3, −4, −4

(a) Compute sample means and standard deviations for the four samples. Plot the means in a manner similar to the plot in Figure 3.22. Make a separate plot of this kind for the standard deviations.

(b) Write a short report to an engineering manager to summarize what these data and your summary statistics and plots show about the performance of the industrial shear. How do you recommend that the shear be set up in the future in order to get strips cut from these materials with widths as close as possible to specified dimensions?

13. The accompanying data are some measured resistivity values from *in situ* doped polysilicon specimens taken from the article "LPCVD Process Equipment Evaluation Using Statistical Methods"

by R. Rossi (*Solid State Technology*, 1984). (The units were not given in the article.)

   5.55, 5.52, 5.45, 5.53, 5.37, 5.22, 5.62, 5.69, 5.60, 5.58, 5.51, 5.53

(a) Make a dot diagram and a boxplot for these data and compute the statistics $\bar{x}$ and s.

(b) Make a normal plot for these data. How bell-shaped does this data set look? If you were to say that the shape departs from a perfect bell shape, in what specific way does it? (Refer to characteristics of the normal plot to support your answer.)

14. The article "Thermal Endurance of Polyester Enameled Wires Using Twisted Wire Specimens" by H. Goldenberg (*IEEE Transactions on Electrical Insulation*, 1965) contains some data on the lifetimes (in weeks) of wire specimens tested for thermal endurance according to AIEE Standard 57. Several different laboratories were used to make the tests, and the results from two of the laboratories, using a test temperature of 200°C, follow:

Laboratory 1	Laboratory 2
14, 16, 17, 18, 20, 22, 23, 25, 27, 28	27, 28, 29, 29, 29, 30, 31, 31, 33, 34

Consider first only the Laboratory 1 data.

(a) Find the median and the first and third quartiles for the lifetimes and then find the .64 quantile of the data set.

(b) Make and interpret a normal plot for these data. Would you describe this distribution as bell-shaped? If not, in what way(s) does it depart from being bell-shaped? Give the coordinates of the 10 points you plot on regular graph paper.

(c) Find the sample mean, the sample range, and the sample standard deviation for these data.

Now consider comparing the work of the two different laboratories (i.e., consider both data sets).

(d) Make back-to-back stem-and-leaf plots for these two data sets (use two leaves for observations 10–19, two for observations 20–29, etc.)

(e) Make side-by-side boxplots for these two data sets. (Draw these on the same scale.)

(f) Based on your work in parts (d) and (e), which of the two labs would you say produced the more precise results?

(g) Is it possible to tell from your plots in (d) and (e) which lab produced the more accurate results? Why or why not?

15. Agusalim, Ferry, and Hollowaty made some measurements on the thickness of wallboard during its manufacture. The accompanying table shows thicknesses (in inches) of 12 different 4 ft × 8 ft boards (at a single location on the boards) both before and after drying in a kiln. (These boards were nominally .500 in. thick.)

Board	1	2	3	4	5	6
Before Drying	.514	.505	.500	.490	.503	.500
After Drying	.510	.502	.493	.486	.497	.494

Board	7	8	9	10	11	12
Before Drying	.510	.508	.500	.511	.505	.501
After Drying	.502	.505	.488	.486	.491	.498

(a) Make a scatterplot of these data. Does there appear to be a strong relationship between after-drying thickness and before-drying thickness? How might such a relationship be of practical engineering importance in the manufacture of wallboard?

(b) Calculate the 12 before minus after differences in thickness. Find the sample mean and sample standard deviation of these values. How might the mean value be used in running the sheetrock manufacturing process? (Based on the mean value, what is an ideal before-drying thickness for the boards?) If somehow all variability in before-drying thickness could be eliminated, would substantial after-drying variability in thickness remain? Explain in terms of your calculations.

16. The accompanying values are representative of data summarized in a histogram appearing in the article "Influence of Final Recrystallization Heat Treatment on Zircaloy-4 Strip Corrosion" by Foster, Dougherty, Burke, Bates, and Worcester (*Journal of Nuclear Materials*, 1990). Given are $n = 20$ particle diameters observed in a bright-field TEM micrograph of a Zircaloy-4 specimen. The units are $10^{-2}\mu$m.

1.73, 2.47, 2.83, 3.20, 3.20, 3.57, 3.93, 4.30, 4.67, 5.03, 5.03, 5.40, 5.77, 6.13, 6.50, 7.23, 7.60, 8.33, 9.43, 11.27

(a) Compute the mean and standard deviation of these particle diameters.

(b) Make both a dot diagram and a boxplot for these data. Sketch the dot diagram on a ruled scale and make the boxplot below it.

(c) Based on your work in (b), how would you describe the shape of this data set?

(d) Make a normal plot of these data. In what specific way does the distribution depart from being bell-shaped?

(e) It is sometimes useful to find a scale of measurement on which a data set is reasonably bell-shaped. To that end, take the natural logarithms of the raw particle diameters. Normal-plot the log diameters. Does this plot appear to be more linear than your plot in (d)?

17. The data in the accompanying tables are measurements of the latent heat of fusion of ice taken from *Experimental Statistics* (NBS Handbook 91) by M. G. Natrella. The measurements were made (on specimens cooled to −.072°C) using two different methods. The first was an electrical method, and the second was a method of mixtures. The units are calories per gram of mass.

(a) Make side-by-side boxplots for comparing the two measurement methods. Does there appear to be any important difference in the precision of the two methods? Is it fair to say that at least one of the methods must be somewhat inaccurate? Explain.

Method A (Electrical)

79.98, 80.04, 80.02, 80.04, 80.03, 80.03, 80.04,
79.97, 80.05, 80.03, 80.02, 80.00, 80.02

Method B (Mixtures)

80.02, 79.94, 79.98, 79.97,
79.97, 80.03, 79.95, 79.97

(b) Compute and compare the sample means and the sample standard deviations for the two methods. How are the comparisons of these numerical quantities already evident on your plot in (a)?

18. T. Babcock did some fatigue life testing on specimens of 1045 steel obtained from three different heats produced by a single steel supplier. The lives till failure of 30 specimens tested on a rotary fatigue strength machine (units are 100 cycles) are

Heat 1

313, 100, 235, 250, 457,
11, 315, 584, 249, 204

Heat 2

349, 206, 163, 350, 189,
216, 170, 359, 267, 196

Heat 3

289, 279, 142, 334, 192,
339, 87, 185, 262, 194

(a) Find the median and first and third quartiles for the Heat 1 data. Then find the .62 quantile of the Heat 1 data set.
(b) Make and interpret a normal plot for the Heat 1 data. Would you describe this data set as bell-shaped? If not, in what specific way does the shape depart from the bell shape? (List the

coordinates of the points you plot on regular graph paper.)
(c) Find the sample mean and sample standard deviation of the Heat 1 data.
(d) Make a stem-and-leaf plot for the Heat 1 data using only the leading digits 0, 1, 2, 3, 4 and 5 to the left of the stem (and pairs of final digits to the right).
(e) Now make back-to-back stem-and-leaf plots for the Heat 1 and Heat 2 data. How do the two distributions of fatigue lives compare?
(f) Show the calculations necessary to make boxplots for each of the three data sets above. Then draw these side by side on the same scale to compare the three heats. How would you say that these three heats compare in terms of uniformity of fatigue lives produced? Do you see any clear differences between heats in terms of the average fatigue life produced?

19. Loveland, Rahardja, and Rainey studied a metal turning process used to make some (cylindrical) servo sleeves. Outside diameter measurements made on ten of these sleeves are given here. (Units are 10^{-5} inch above nominal. The "notch" axis of the sleeve was an identifiable axis and the non-notch axis was perpendicular to the notch axis. A dial bore gauge and an air spindler gauge were used.)

Sleeve	1	2	3	4	5
Notch/Dial Bore	130	160	170	310	200
Non-Notch/Dial Bore	150	150	210	160	160
Notch/Air Spindler	40	60	45	0	30
Sleeve	6	7	8	9	10
Notch/Dial Bore	130	200	150	200	140
Non-Notch/Dial Bore	140	220	150	220	160
Notch/Air Spindler	0	25	25	−40	65

(a) What can be learned from the dial bore data that could not be learned from data consisting of the given notch measurements above and ten non-notch measurements on a *different* ten servo sleeves?

(b) The dial bore data might well be termed "paired" data. A common method of analysis for such data is to take differences and study those. Compute the ten "notch minus non-notch" differences for the dial bore values. Make a dot diagram for these and then a boxplot. What physical interpretation does a nonzero mean for such differences have? What physical interpretation does a large variability in these differences have?

(c) Make a scatterplot of the air spindler notch measurements versus the dial bore notch measurements. Does it appear that the air spindler and dial bore measurements are strongly related?

(d) How would you suggest trying to determine which of the two gauges is most precise?

20. Duren, Leng and Patterson studied the drilling of holes in a miniature metal part using two different physical processes (laser drilling and electrical discharge machining). Blueprint specifications on these holes called for them to be drilled at an angle of 45° to the top surface of the part in question. The realized angles measured on 13 parts drilled using each process (26 parts in all) are

Laser (Hole A)
42.8, 42.2, 42.7, 43.1, 40.0, 43.5,
42.3, 40.3, 41.3, 48.5, 39.5, 41.1, 42.1

EDM (Hole A)
46.1, 45.3, 45.3, 44.7, 44.2, 44.6,
43.4, 44.6, 44.6, 45.5, 44.4, 44.0, 43.2

(a) Find the median and the first and third quartiles for the Laser data. Then find the .37 quantile of the Laser data set.

(b) Make and interpret a normal plot for the Laser data. Would you describe this distribution as bell-shaped? If not, in what way(s) does it depart from being bell-shaped?

(c) Find the sample mean, the sample range, and the sample standard deviation for the Laser data.

Now consider comparing the two different drilling methods.

(d) Make back-to-back stem-and-leaf plots for the two data sets.

(e) Make side-by-side boxplots for the two data sets. (Draw these on the same scale.)

(f) Based on your work in parts (d) and (e), which of the two processes would you say produced the most consistent results? Which process produced an "average" angle closest to the nominal angle (45°)?

As it turns out, each metal part actually had two holes drilled in it and their angles measured. Below are the measured angles of the second hole drilled in each of the parts made using the Laser process. (The data are listed in the same part order as earlier.)

Laser (Hole B)
43.1, 44.3, 44.5, 46.3, 43.9, 41.9,
43.4, 49.0, 43.5, 47.2, 44.8 ,44.0, 43.9

(g) Taking together the two sets of Laser measurements, how would you describe these values using the terminology of Section 1.2?

(h) Make a scatterplot of the Hole A and Hole B laser data. Does there appear to be a strong relationship between the angles produced in a single part by this drilling method?

(i) Calculate the 13 Hole A minus Hole B differences in measured angles produced using the Laser drilling process. Find the sample mean and sample standard deviation of these values. What do these quantities measure here?

21. Blad, Sobotka, and Zaug did some hardness testing of a metal specimen. They tested it on three different machines, a dial Rockwell tester, a digital Rockwell tester, and a Brinell tester. They made ten measurements with each machine and the values they obtained for Brinell hardness (after

conversion in the case of the Rockwell readings) were

Dial Rockwell
536.6, 539.2, 524.4, 536.6, 526.8, 531.6, 540.5, 534.0, 526.8, 531.6

Digital Rockwell
501.2, 522.0, 531.6, 522.0, 519.4, 523.2, 522.0, 514.2, 506.4, 518.1

Brinell
542.6, 526.0, 520.5, 514.0, 546.6, 512.6, 516.0, 580.4, 600.0, 601.0

Consider first only the Dial Rockwell data.

(a) Find the median and the first and third quartiles for the hardness measurements. Then find the .27 quantile of the data set.

(b) Make and interpret a normal plot for these data. Would you describe this distribution as bell-shaped? If not, in what way(s) does it depart from being bell-shaped?

(c) Find the sample mean, the sample range, and the sample standard deviation for these data.

Now consider comparing the readings from the different testers (i.e., consider all three data sets.)

(d) Make back-to-back stem-and-leaf plots for the two Rockwell data sets. (Use two "leaves" for observations 500–509, two for the observations 510–519, etc.)

(e) Make side-by-side boxplots for all three data sets. (Draw these on the same scale.)

(f) Based on your work in part (e), which of the three machines would you say produced the most precise results?

(g) Is it possible to tell from your plot (e) which machine produced the most accurate results? Why or why not?

22. Ritchey, Bazan, and Buhman did an experiment to compare flight times of several designs of paper helicopters, dropping them from the first to ground floors of the ISU Design Center. The flight times that they reported for two different designs were (the units are seconds)

Design 1	Design 2
2.47, 2.45, 2.43, 2.67, 2.69, 2.48, 2.44, 2.71, 2.84, 2.84	3.42, 3.50, 3.29, 3.51, 3.53, 2.67, 2.69, 3.47, 3.40, 2.87

(a) Find the median and the first and third quartiles for the Design 1 data. Then find the .62 quantile of the Design 1 data set.

(b) Make and interpret a normal plot for the Design 1 data. Would you describe this distribution as bell-shaped? If not, in what way(s) does it depart from being bell-shaped?

(c) Find the sample mean, the sample range, and the sample standard deviation for the Design 1 data. Show some work.

Now consider comparing the two different designs.

(d) Make back-to-back stem-and-leaf plots for the two data sets.

(e) Make side-by-side boxplots for the two data sets. (Draw these on the same scale.)

(f) Based on your work in parts (d) and (e), which of the two designs would you say produced the most consistent results? Which design produced the longest flight times?

(g) It is not really clear from the students' report whether the data came from the dropping of one helicopter of each design ten times, or from the dropping of ten helicopters of each design once. Briefly discuss which of these possibilities is preferable if the object of the study was to identify a superior design. (If necessary, review Section 2.3.4.)