3.1.

The content of liquid detergent bottles is being analyzed. Twelve bottles, randomly selected from the process, are measured, and the results are as follows (in fluid ounces): 16.05, 16.03, 16.02, 16.04, 16.05, 16.01, 16.02, 16.02, 16.03, 16.01, 16.00, 16.07.

(a) Calculate the sample average.

$$\overline{x} = \sum_{i=1}^{n} x_i \Big/ n = (16.05 + 16.03 + \cdots + 16.07) \big/ 12 = 16.029 \text{ oz}$$

(b) Calculate the sample standard deviation.

$$s = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2 \Big/ n}{n-1}} = \sqrt{\frac{(16.05^2 + \cdots + 16.07^2) - (16.05 + \cdots + 16.07)^2 / 12}{12-1}} = 0.0202 \text{ oz}$$

MTB > Stat > Basic Statistics > Display Descriptive Statistics

Descriptive Statistics: Ex3-1

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3
Ex3-1	12	0	16.029	0.00583	0.0202	16.000	16.012	16.025	16.047

Variable	Maximum
Ex3-1	16.070

3.5.

The nine measurements that follow are furnace temperatures recorded on successive batches in a semiconductor manufacturing process (units are °F): 953, 955, 948, 951, 957, 949, 954, 950, 959

(a) Calculate the sample average.

$$\overline{x} = \sum_{i=1}^{n} x_i \Big/ n = (953 + 955 + \cdots + 959) \big/ 9 = 952.9 \text{ °F}$$

(b) Calculate the sample standard deviation.

$$s = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2 \Big/ n}{n-1}} = \sqrt{\frac{(953^2 + \cdots + 959^2) - (953 + \cdots + 959)^2 / 9}{9-1}} = 3.7 \text{ °F}$$

MTB > Stat > Basic Statistics > Display Descriptive Statistics

Descriptive Statistics: Ex3-5

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3
Ex3-5	9	0	952.89	1.24	3.72	948.00	949.50	953.00	956.00

Variable	Maximum
Ex3-5	959.00

3.6.

Consider the furnace temperature data in Exercise 3.5.

(a) Find the sample median of these data.

In ranked order, the data are {948, 949, 950, 951, **953**, 954, 955, 957, 959}. The sample median is the middle value.

(b) How much could the largest temperature measurement increase without changing the sample median?

Since the median is the value dividing the ranked sample observations in half, it remains the same regardless of the size of the largest measurement.

3.9.

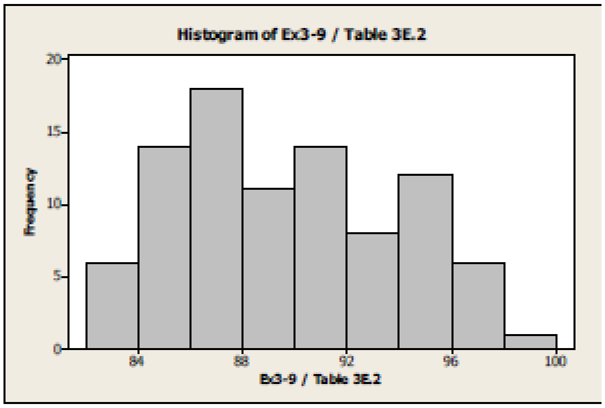
The data shown in Table 3E.2 are chemical process yield readings on successive days (read down, then across). Construct a histogram for these data. Comment on the shape of the histogram. Does it resemble any of the distributions that we have discussed in this chapter?

TABLE 3E.2

Process Yield

94.1	87.3	94.1	92.4	84.6	85.4
93.2	84.1	92.1	90.6	83.6	86.6
90.6	90.1	96.4	89.1	85.4	91.7
91.4	95.2	88.2	88.8	89.7	87.5
88.2	86.1	86.4	86.4	87.6	84.2
86.1	94.3	85.0	85.1	85.1	85.1
95.1	93.2	84.9	84.0	89.6	90.5
90.0	86.7	87.3	93.7	90.0	95.6
92.4	83.0	89.6	87.7	90.1	88.3
87.3	95.3	90.3	90.6	94.3	84.1
86.6	94.1	93.1	89.4	97.3	83.7
91.2	97.8	94.6	88.6	96.8	82.9
86.1	93.1	96.3	84.1	94.4	87.3
90.4	86.4	94.7	82.6	96.1	86.4
89.1	87.6	91.1	83.1	98.0	84.5

Use $\sqrt{n} = \sqrt{90} \cong 9$ bins



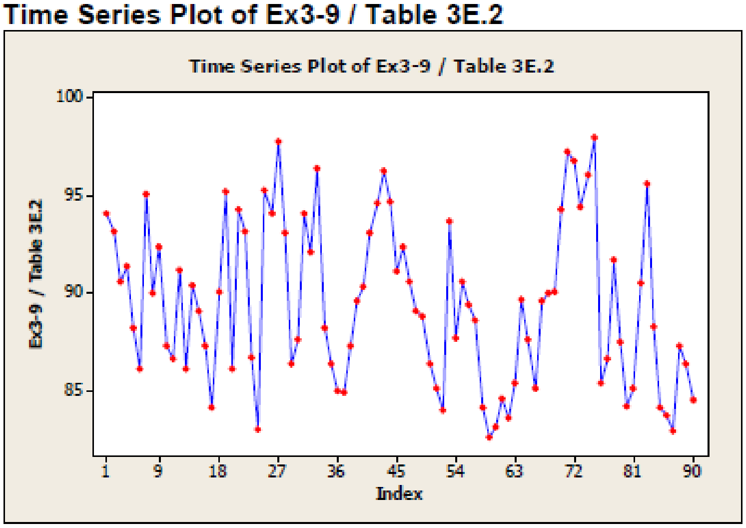
The distribution seems to be heavy tail in the right hand side. It is temping to suspect that there are two modes, however, the 2nd mode is rather minor than significant.

3.21.

Reconsider the yield data in Exercise 3.9. Construct a time-series plot for these data. Interpret the plot.

MTB > Graph > Time Series Plot > Simple

Time Series Plot of Ex3-9 / Table 3E.2



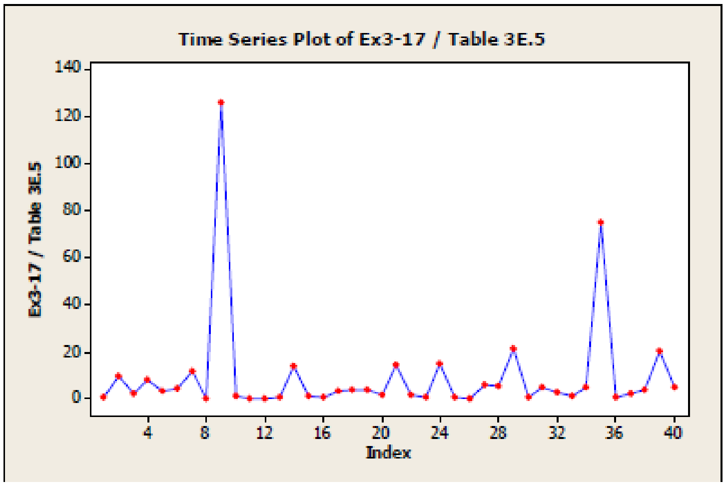
Time may be an important source of variability, as evidenced by potentially cyclic behavior.

3.22.

Consider the concentration of suspended solids from Exercise 3.17. Assume that reading across, then down, gives the data in time order. Construct and interpret a time-series plot.

MTB > Graph > Time Series Plot > Simple

Time Series Plot of Ex3-17 / Table 3E.5



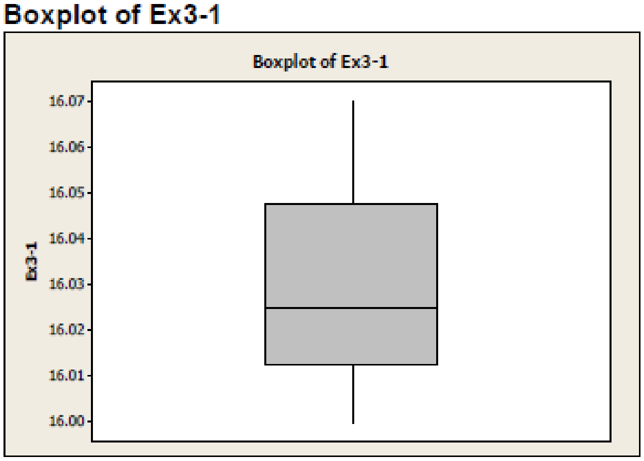
Although most of the readings are between 0 and 20, there are two unusually large readings (9, 35), as well as occasional "spikes" around 20. The order in which the data were collected may be an important source of variability.

3.25.

Construct a box plot for the data in Exercise 3.1.

MTB > Graph > Boxplot > Simple

Boxplot of Ex3-1



In a box plot of normally distributed data, the median line is in the middle of the box, and the two whiskers are the same length. This data is approximately normally distributed.