CHAPTER 6

Section 6-1

6-1. Sample average:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{592.035}{8} = 74.0044 \text{ mm}$$

Sample variance:

$$\sum_{i=1}^{8} x_i = 592.035$$

$$\sum_{i=1}^{8} x_i^2 = 43813.18031$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{43813.18031 - \frac{(592.035)^2}{8}}{8-1}$$

$$= \frac{0.0001569}{7} = 0.000022414 \quad (\text{mm})^2$$

Sample standard deviation:

$$s = \sqrt{0.000022414} = 0.00473 \quad \text{mm}$$

The sample standard deviation could also be found using

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \quad \text{where} \quad \sum_{i=1}^{8}(x_i - \bar{x})^2 = 0.0001569$$

Dot Diagram:

```
                    .. ...:            .
-------+---------+---------+---------+---------+---------diameter
 73.9920   74.0000   74.0080   74.0160   74.0240   74.0320
```

There appears to be a possible outlier in the data set.

6-2. Sample average:

$$\bar{x} = \frac{\sum_{i=1}^{19} x_i}{19} = \frac{272.82}{19} = 14.359 \text{ min}$$

Sample variance:

$$\sum_{i=1}^{19} x_i = 272.82$$

$$\sum_{i=1}^{19} x_i^2 = 10333.8964$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{10333.8964 - \frac{(272.82)^2}{19}}{19-1}$$

$$= \frac{6416.49}{18} = 356.47 \text{ (min)}^2$$

Sample standard deviation:

$$s = \sqrt{356.47} = 18.88 \text{ min}$$

The sample standard deviation could also be found using

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

where

$$\sum_{i=1}^{19} (x_i - \bar{x})^2 = 6416.49$$

6-3. Sample average:

$$\bar{x} = \frac{84817}{12} = 7068.1 \quad yards$$

Sample variance:

$$\sum_{i=1}^{12} x_i = 84817$$

$$\sum_{i=1}^{19} x_i^2 = 600057949$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{600057949 - \frac{(84817)^2}{12}}{12-1}$$

$$= \frac{564324.92}{11} = 51302.265 \quad (yards)^2$$

Sample standard deviation:

$$s = \sqrt{51302.265} = 226.5 \quad yards$$

The sample standard deviation could also be found using

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

where

$$\sum_{i=1}^{12} (x_i - \bar{x})^2 = 564324.92$$

Dot Diagram: (rounding was used to create the dot diagram)

```
        .      .  :  .. ..   :                            :
   -+---------+---------+---------+---------+---------+-----C1
  6750      6900      7050      7200      7350      7500
```

6-4. Sample mean:

$$\bar{x} = \frac{\sum_{i=1}^{18} x_i}{18} = \frac{2272}{18} = 126.22 \quad \text{kN}$$

Sample variance:

$$\sum_{i=1}^{18} x_i = 2272$$

$$\sum_{i=1}^{18} x_i^2 = 298392$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{298392 - \frac{(2272)^2}{18}}{18-1}$$

$$= \frac{11615.11}{17} = 683.24 \quad (\text{kN})^2$$

Sample standard deviation:

$$s = \sqrt{683.24} = 26.14 \quad \text{kN}$$

The sample standard deviation could also be found using

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

where

$$\sum_{i=1}^{18} (x_i - \bar{x})^2 = 11615.11$$

Dot Diagram:

```
         .
     :   ::  .           ::       .          .  : .   .
 +---------+---------+---------+---------+---------+-------yield
90       105       120       135       150       165
```

6-5. Sample average:

$$\bar{x} = \frac{351.8}{8} = 43.975$$

Sample variance:

$$\sum_{i=1}^{8} x_i = 351.8$$

$$\sum_{i=1}^{19} x_i^2 = 16528.403$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{16528.043 - \frac{(351.8)^2}{8}}{8-1}$$

$$= \frac{1057.998}{7} = 151.143$$

Sample standard deviation:

$$s = \sqrt{151.143} = 12.294$$

The sample standard deviation could also be found using

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

where

$$\sum_{i=1}^{8} (x_i - \bar{x})^2 = 1057.998$$

Dot Diagram:

```
      .            . ..   .        ..                     .
   +---------+---------+---------+---------+---------+-------
24.0      32.0      40.0      48.0      56.0      64.0
```

6-6. Sample average:

$$\bar{x} = \frac{\sum_{i=1}^{35} x_i}{35} = \frac{28368}{35} = 810.514 \quad watts/m^2$$

Sample variance:

$$\sum_{i=1}^{19} x_i = 28368$$

$$\sum_{i=1}^{19} x_i^2 = 23552500$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{23552500 - \frac{(28368)^2}{35}}{35-1} = \frac{559830.743}{34}$$

$$= 16465.61 \quad (watts/m^2)^2$$

Sample standard deviation:

$$s = \sqrt{16465.61} = 128.32 \quad watts/m^2$$

The sample standard deviation could also be found using

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

where

$$\sum_{i=1}^{35} (x_i - \bar{x})^2 = 559830.743$$

6-7. $\mu = \frac{6905}{1270} = 5.44$; The value 5.44 is the population mean since the actual physical population of all flight times during the operation is available.

6-8 a.) Sample average:

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{19.56}{9} = 2.173 \quad mm$$

b.) Sample variance:

$$\sum_{i=1}^{9} x_i = 19.56$$

$$\sum_{i=1}^{9} x_i^2 = 45.953$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{45.953 - \frac{(19.56)^2}{9}}{9-1} = \frac{3.443}{8}$$

$$= 0.4303 \quad (mm)^2$$

Sample standard deviation:

$$s = \sqrt{0.4303} = 0.6560 \quad mm$$

c.) Dot Diagram

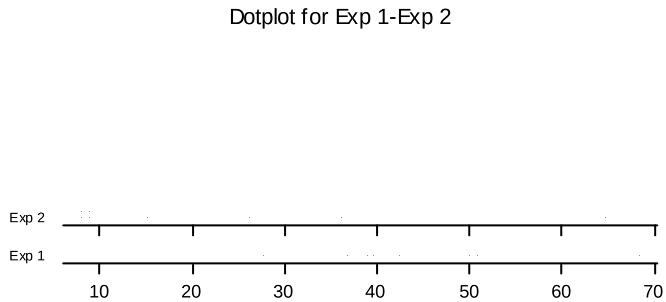
```
.     .            .     .  .         .  .          ..
-------+---------+---------+---------+---------+---------crack length
    1.40      1.75      2.10      2.45      2.80      3.15
```

6-9. Dot Diagram (rounding of the data is used to create the dot diagram)

```
                                    x̄
                                    ↓
                    . .
       .     :       .:  ... . . .: :. . .:. .:: :::
-----+---------+---------+---------+---------+---------+-x
    500       600       700       800       900      1000
```

The sample mean is the point at which the data would balance if it were on a scale.

a. Dot Diagram of CRT data in exercise 6-5 (Data were rounded for the plot)



The data are centered a lot lower in the second experiment. The lower CRT resolution reduces the visual accommodation.

6-11. a) $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{57.47}{8} = 7.184$

b) $s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{412.853 - \frac{(57.47)^2}{8}}{8-1} = \frac{0.00299}{7} = 0.000427$

$s = \sqrt{0.000427} = 0.02066$

c) Examples: repeatability of the test equipment, time lag between samples, during which the pH of the solution could change, and operator skill in drawing the sample or using the instrument.

6-12 sample mean $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{748.0}{9} = 83.11$ drag counts

sample variance $s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{62572 - \frac{(748.0)^2}{9}}{9-1}$

$= \frac{404.89}{8} = 50.61 \text{ drag counts}^2$

sample standard deviation $s = \sqrt{50.61} = 7.11$ drag counts

Dot Diagram

```
  .         . . . . . . .                         .
---+---------+---------+---------+---------+---------+---drag counts
  75.0      80.0      85.0      90.0      95.0     100.0
```

6-13. a) $\bar{x} = 65.86$ °F

$s = 12.16$ °F

b) Dot Diagram

```
                                                        :   :
   .         .     .    .   ..    .: .: .  .:..: .. :: .... ..
-+---------+---------+---------+---------+---------+-----temp
30        40        50        60        70        80
```

c) Removing the smallest observation (31), the sample mean and standard deviation become

$\bar{x} = 66.86$ °F

s = 10.74 °F

Section 6-3

6-14 Stem-and-leaf display of octane rating N = 83

Leaf Unit = 0.10 83|4 represents 83.4

```
  1   83|4
  3   84|33
  4   85|3
  7   86|777
 13   87|456789
 24   88|23334556679
 34   89|0233678899
(13)  90|0111344456789
 36   91|00011122256688
 22   92|22236777
 14   93|023347
  8   94|2247
  4   95|
  4   96|15
  2   97|
  2   98|8
  1   99|
  1  100|3
```

6-15 a.) Stem-and-leaf display for cycles to failure: unit = 100 1|2 represents 1200

```
  1    0T|3
  1    0F|
  5    0S|7777
 10    0o|88899
 22    1*|000000011111
 33    1T|22222223333
(15)   1F|444445555555555
 22    1S|66667777777
 11    1o|888899
  5    2*|011
  2    2T|22
```

b) No, only 5 out of 70 coupons survived beyond 2000 cycles.

6-16 Stem-and-leaf display of percentage of cotton N = 64
Leaf Unit = 0.10 32|1 represents 32.1%

```
   1   32|1
   6   32|56789
   9   33|114
  17   33|56666688
  24   34|0111223
 (14)  34|55666667777779
  26   35|001112344
  17   35|56789
  12   36|234
   9   36|6888
   5   37|13
   3   37|689
```

6-17. Stem-and-leaf display for Problem 2-4.yield: unit = 1 1|2 represents 12

```
   1    7o|8
   1    8*|
   7    8T|223333
  21    8F|44444444555555
  38    8S|66666666667777777
 (11)   8o|88888999999
  41    9*|00000000001111
  27    9T|22233333
  19    9F|444444445555
   7    9S|666677
   1    9o|8
```

6-18 Descriptive Statistics

Variable	N	Median	Q1	Q3
Octane Rating	83	90.400	88.600	92.200

6-19. Descriptive Statistics

Variable	N	Median	Q1	Q3
cycles	70	1436.5	1097.8	1735.0

6-20 median: $\tilde{x} = 34.700$ %
mode: 34.7 %
sample average: $\bar{x} = 34.798$ %

6-21. Descriptive Statistics

Variable	N	Median	Q1	Q3
yield	90	89.250	86.100	93.125

6-22 a.) sample mean: $\overline{x} = 65.811$ inches standard deviation $s = 2.106$ inches

b.) Stem-and-leaf display of female engineering student heights N = 37
Leaf Unit = 0.10 61|0 represents 61.0 inches

```
  1   61|0
  3   62|00
  5   63|00
  9   64|0000
 17   65|00000000
 (4)  66|0000
 16   67|00000000
  8   68|00000
  3   69|00
  1   70|0
```

c.) median: $\tilde{x} = 66.000$ inches

6-23 Stem-and-leaf display for Problem 6-23. Strength: unit = 1.0 1|2 represents 12

```
   1  532|9
   1  533|
   2  534|2
   4  535|47
   5  536|6
   9  537|5678
  20  538|12345778888
  26  539|016999
  37  540|11166677889
  46  541|123666688
 (13) 542|0011222357899
  41  543|011112556
  33  544|00012455678
  22  545|2334457899
  13  546|23569
   8  547|357
   5  548|11257
```

6-24 Stem-and-leaf of concentration N = 60 Leaf Unit = 1.0 2|2 represents 29
Note: Minitab has dropped the value to the right of the decimal to make this display.

```
   1    2|9
   2    3|1
   3    3|9
   8    4|22223
  12    4|5689
  20    5|01223444
 (13)   5|5666777899999
  27    6|11244
  22    6|556677789
  13    7|022333
   7    7|6777
   3    8|01
   1    8|9
```

The data have a symmetrical bell-shaped distribution, and therefore may be normally distributed.

Sample Mean $\bar{x} = \dfrac{\sum_{i=1}^{n} x_i}{n} = \dfrac{3592.0}{60} = 59.87$

Sample Standard Deviation

$$\sum_{i=1}^{60} x_i = 3592.0 \qquad \text{and} \qquad \sum_{i=1}^{60} x_i^2 = 224257$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{224257 - \frac{(3592.0)^2}{60}}{60-1} = \frac{9215.93}{59}$$

$$= 156.20$$

and

$$s = \sqrt{156.20} = 12.50$$

Sample Median $\tilde{x} = 59.45$

```
Variable          N     Median
concentration    60      59.45
```

6-25 Stem-and-leaf display for Problem 6-25. Yard: unit = 1.0

Note: Minitab has dropped the value to the right of the decimal to make this display.

1	22 | 6
5	23 | 2334
8	23 | 677
16	24 | 00112444
20	24 | 5578
33	25 | 0111122334444
46	25 | 5555556677899
(15)	26 | 000011123334444
39	26 | 56677888
31	27 | 0000112222233333444
12	27 | 66788999
4	28 | 003
1	28 | 5

Sample Mean $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{100} x_i}{100} = \frac{26030.2}{100} = 260.3 \quad yards$

Sample Standard Deviation

$$\sum_{i=1}^{100} x_i = 26030.2 \qquad \text{and} \qquad \sum_{i=1}^{100} x_i^2 = 6793512$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{6793512 - \frac{((26030.2)^2}{100}}{100-1} = \frac{17798.42}{99}$$

$$= 179.782 \quad yards^2$$

and

$$s = \sqrt{179.782} = 13.41 \quad yards$$

Sample Median

```
Variable            N     Median
yards             100     260.85
```

6-26 Stem-and-leaf of speed (in megahertz) N = 120
Leaf Unit = 1.0 63|4 represents 634 megahertz

```
   2   63|47
   7   64|24899
  16   65|223566899
  35   66|0000001233455788899
  48   67|0022455567899
 (17)  68|00001111233333458
  55   69|0000112345555677889
  36   70|011223444556
  24   71|0057889
  17   72|000012234447
   5   73|59
   3   74|68
   1   75|
   1   76|3
```

35/120= 29% exceed 700 megahertz.

Sample Mean $\bar{x} = \frac{\sum_{i=1}^{120} x_i}{120} = \frac{82413}{120} = 686.78 \quad mhz$

Sample Standard Deviation

$$\sum_{i=1}^{120} x_i = 82413 \qquad \text{and} \qquad \sum_{i=1}^{120} x_i^2 = 56677591$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{56677591 - \frac{(82413)^2}{120}}{120-1} = \frac{78402.925}{119}$$

$$= 658.85 \quad mhz^2$$

and

$$s = \sqrt{658.85} = 25.67 \quad mhz$$

Sample Median $\tilde{x} = 683.0 \, mhz$

```
Variable              N      Median
speed               120      683.00
```

6-27 a.)Stem-and-leaf display of Problem 6-27. Rating: unit = 0.10 1|2 represents 1.2

```
  1   83|0
  2   84|0
  5   85|000
  7   86|00
  9   87|00
 12   88|000
 18   89|000000
(7)   90|0000000
 15   91|0000000
  8   92|0000
  4   93|0
  3   94|0
  2   95|00
```

b.) Sample Mean

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{40} x_i}{40} = \frac{3578}{40} = 89.45$$

Sample Standard Deviation

$$\sum_{i=1}^{40} x_i = 3578 \qquad \text{and} \qquad \sum_{i=1}^{40} x_i^2 = 320366$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{320366 - \frac{(3578)^2}{40}}{40-1} = \frac{313.9}{39}$$

$$= 8.05$$

and

$$s = \sqrt{8.05} = 2.8$$

Sample Median

```
Variable              N        Median
rating               40        90.000
```

c.) 22/40 or 55% of the taste testers considered this particular Pinot Noir truly exceptional.

6-28 a.) Stem-and-leaf diagram of $NbOCl_3$ N = 27
Leaf Unit = 100 0|4 represents 40 gram-mole/liter x 10^{-3}

```
  6     0|444444
  7     0|5
 (9)    1|001122233
 11     1|5679
  7     2|
  7     2|5677
  3     3|124
```

b.) sample mean $\bar{x} = \frac{\sum_{i=1}^{27} x_i}{27} = \frac{41553}{27} = 1539$ gram – mole/liter x10^{-3}

Sample Standard Deviation

$$\sum_{i=1}^{27} x_i = 41553 \quad \text{and} \quad \sum_{i=1}^{27} x_i^2 = 87792869$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{87792869 - \frac{(41553)^2}{27}}{27-1} = \frac{23842802}{26} = 917030.85$$

and $s = \sqrt{917030.85} = 957.62$ gram - mole/liter x 10^{-3}

Sample Median $\tilde{x} = 1256$ gram – mole/liter x10^{-3}

```
Variable              N       Median
NbOCl3               40         1256
```

6-29 a.)Stem-and-leaf display for Problem 6-29. Height: unit = 0.10 1|2 represents 1.2

```
   Female Students|| Male Students
         0|61   1  |
        00|62   3  |
        00|63   5  |
      0000|64   9  |
  00000000|65  17  |   2   65|00
      0000|66  (4) |   3   66|0
  00000000|67  16  |   7   67|0000
     00000|68   8  |  17   68|0000000000
        00|69   3  | (15)  69|000000000000000
         0|70   1  |  18   70|0000000
                   |  11    71|00000
                   |   6    72|00
                   |   4   73|00
                   |   2  74|0
                   |   1   75|0
```

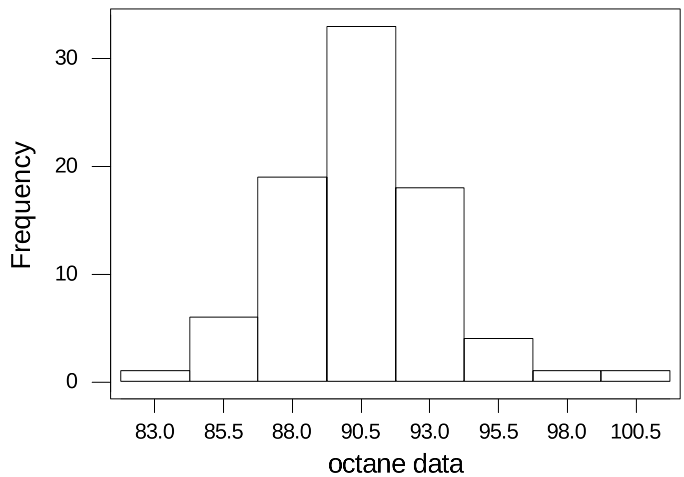
b.) The male engineering students are taller than the female engineering students. Also there is a slightly wider range in the heights of the male students.

Section 6-4

6-30 Frequency Tabulation for Exercise 6-14.Octane Data

Class	Lower Limit	Upper Limit	Midpoint	Frequency	Relative Frequency	Cumulative Frequency	Cum. Rel. Frequency
at or below		81.75		0	.0000	0	.0000
1	81.75	84.25	83.0	1	.0120	1	.0120
2	84.25	86.75	85.5	6	.0723	7	.0843
3	86.75	89.25	88.0	19	.2289	26	.3133
4	89.25	91.75	90.5	33	.3976	59	.7108
5	91.75	94.25	93.0	18	.2169	77	.9277
6	94.25	96.75	95.5	4	.0482	81	.9759
7	96.75	99.25	98.0	1	.0120	82	.9880
8	99.25	101.75	100.5	1	.0120	83	1.0000
above	101.75			0	.0000	83	1.0000

Mean = 90.534 Standard Deviation = 2.888 Median = 90.400

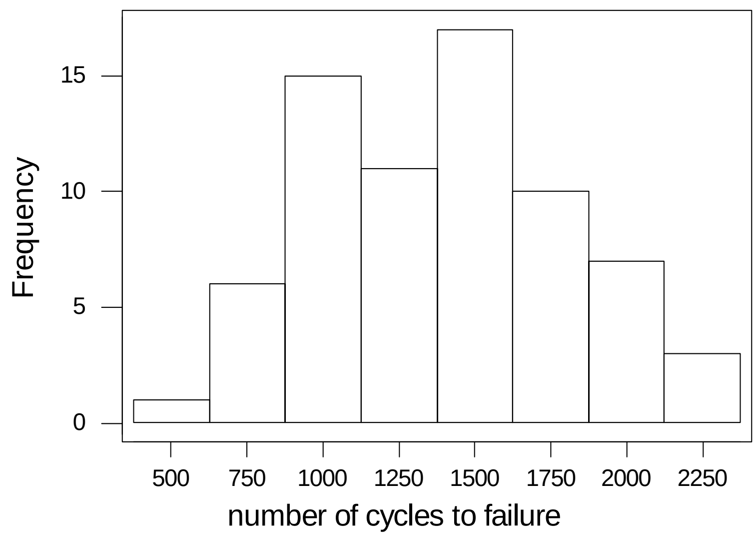


6-31.

Frequency Tabulation for Exercise 6-15.Cycles

Class	Lower Limit	Upper Limit	Midpoint	Frequency	Relative Frequency	Cumulative Frequency	Cum. Rel. Frequency
at or below		.000		0	.0000	0	.0000
1	.000	266.667	133.333	0	.0000	0	.0000
2	266.667	533.333	400.000	1	.0143	1	.0143
3	533.333	800.000	666.667	4	.0571	5	.0714
4	800.000	1066.667	933.333	11	.1571	16	.2286
5	1066.667	1333.333	1200.000	17	.2429	33	.4714
6	1333.333	1600.000	1466.667	15	.2143	48	.6857
7	1600.000	1866.667	1733.333	12	.1714	60	.8571
8	1866.667	2133.333	2000.000	8	.1143	68	.9714
9	2133.333	2400.000	2266.667	2	.0286	70	1.0000
above	2400.000			0	.0000	70	1.0000

Mean = 1403.66 Standard Deviation = 402.385 Median = 1436.5

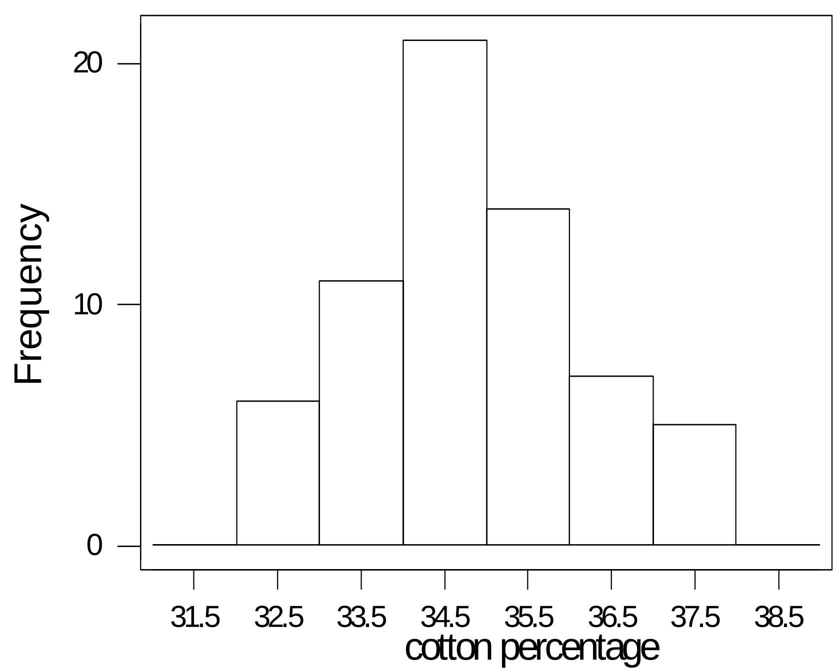


6-32

Frequency Tabulation for Exercise 6-16.Cotton content

Class	Lower Limit	Upper Limit	Midpoint	Frequency	Relative Frequency	Cumulative Frequency	Cum. Rel. Frequency
at or below		31.0		0	.0000	0	.0000
1	31.0	32.0	31.5	0	.0000	0	.0000
2	32.0	33.0	32.5	6	.0938	6	.0938
3	33.0	34.0	33.5	11	.1719	17	.2656
4	34.0	35.0	34.5	21	.3281	38	.5938
5	35.0	36.0	35.5	14	.2188	52	.8125
6	36.0	37.0	36.5	7	.1094	59	.9219
7	37.0	38.0	37.5	5	.0781	64	1.0000
8	38.0	39.0	38.5	0	.0000	64	1.0000
above	39.0			0	.0000	64	1.0000

Mean = 34.798 Standard Deviation = 1.364 Median = 34.700

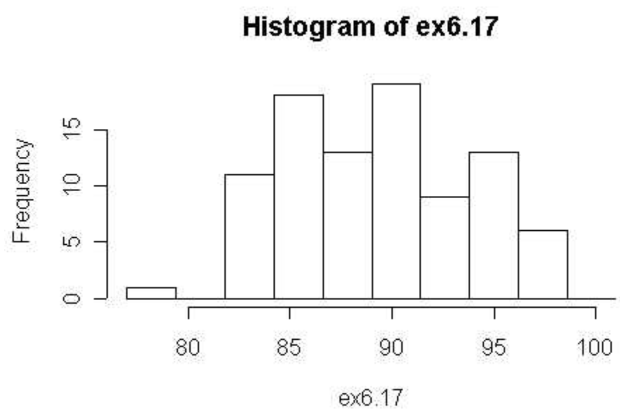


6-33.

Frequency Tabulation for Exercise 6-17.Yield

Class	Lower Limit	Upper Limit	Midpoint	Frequency	Relative Frequency	Cumulative Frequency	Cum. Rel. Frequency
at or below		77.000		0	.0000	0	.0000
1	77.000	79.400	78.200	1	.0111	1	.0111
2	79.400	81.800	80.600	0	.0000	1	.0111
3	81.800	84.200	83.000	11	.1222	12	.1333
4	84.200	86.600	85.400	18	.2000	30	.3333
5	86.600	89.000	87.800	13	.1444	43	.4778
6	89.000	91.400	90.200	19	.2111	62	.6889
7	91.400	93.800	92.600	9	.1000	71	.7889
8	93.800	96.200	95.000	13	.1444	84	.9333
9	96.200	98.600	97.400	6	.0667	90	1.0000
10	98.600	101.000	99.800	0	.0000	90	1.0000
above	101.000			0	.0000	90	1.0000

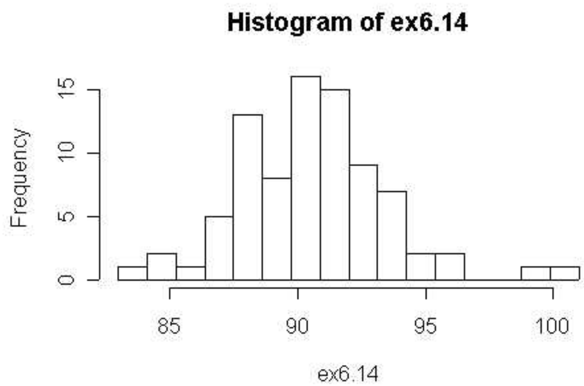
Mean = 89.3756 Standard Deviation = 4.31591 Median = 89.25



6-34 Frequency Tabulation for Exercise 6-14.Octane Data

Class	Lower Limit	Upper Limit	Midpoint	Frequency	Relative Frequency	Cumulative Frequency	Cum. Rel. Frequency
at or below		83.000		0	.0000	0	.0000
1	83.000	84.125	83.5625	1	.0120	1	.0120
2	84.125	85.250	84.6875	2	.0241	3	.0361
3	85.250	86.375	85.8125	1	.0120	4	.0482
4	86.375	87.500	86.9375	5	.0602	9	.1084
5	87.500	88.625	88.0625	13	.1566	22	.2651
6	88.625	89.750	89.1875	8	.0964	30	.3614
7	89.750	90.875	90.3125	16	.1928	46	.5542
8	90.875	92.000	91.4375	15	.1807	61	.7349
9	92.000	93.125	92.5625	9	.1084	70	.8434
10	93.125	94.250	93.6875	7	.0843	77	.9277
11	94.250	95.375	94.8125	2	.0241	79	.9518
12	95.375	96.500	95.9375	2	.0241	81	.9759
13	96.500	97.625	97.0625	0	.0000	81	.9759
14	97.625	98.750	98.1875	0	.0000	81	.9759
15	98.750	99.875	99.3125	1	.0120	82	.9880
16	99.875	101.000	100.4375	1	.0120	83	1.0000
above	101.000			0	.0000	83	1.0000

Mean = 90.534 Standard Deviation = 2.888 Median = 90.400



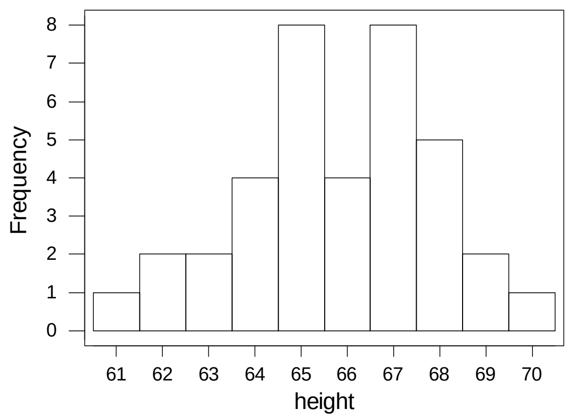
The histograms have the same shape. Not much information is gained by doubling the number of bins.

6-35

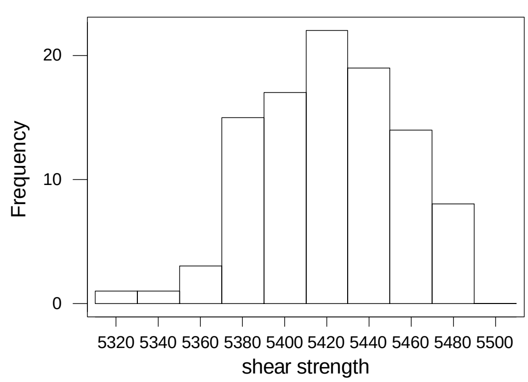
Frequency Tabulation for Problem 6-22. Height Data

Class	Lower Limit	Upper Limit	Midpoint	Frequency	Relative Frequency	Cumulative Frequency	Cum. Rel. Frequency
at or below		60.500		0	.0000	0	.0000
1	60.500	61.500	61.000	1	.0270	1	.0270
2	61.500	62.500	62.000	2	.0541	3	.0811
3	62.500	63.500	63.000	2	.0541	5	.1351
4	63.500	64.500	64.000	4	.1081	9	.2432
5	64.500	65.500	65.000	8	.2162	17	.4595
6	65.500	66.500	66.000	4	.1081	21	.5676
7	66.500	67.500	67.000	8	.2162	29	.7838
8	67.500	68.500	68.000	5	.1351	34	.9189
9	68.500	69.500	69.000	2	.0541	36	.9730
10	69.500	70.500	70.000	1	.0270	37	1.0000
above	70.500			0	.0000	37	1.0000

Mean = 65.811   Standard Deviation = 2.106   Median = 66.0



6-36 The histogram for the spot weld shear strength data shows that the data appear to be normally distributed (the same shape that appears in the stem-leaf-diagram).

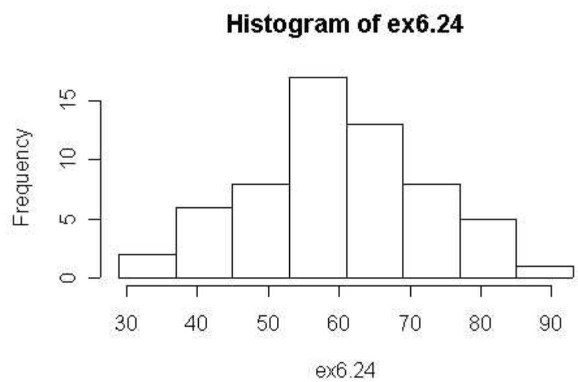


6-37

Frequency Tabulation for exercise 6-24. Concentration data

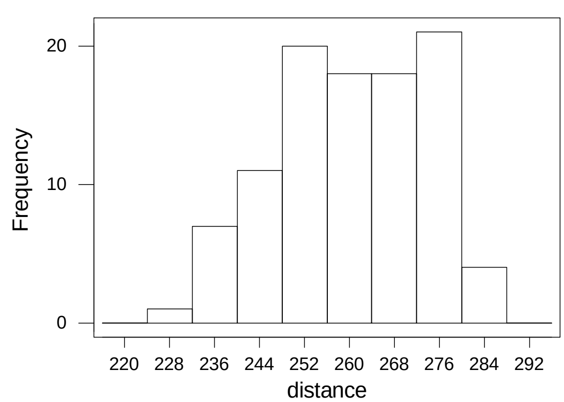
Class	Lower Limit	Upper Limit	Midpoint	Frequency	Relative Frequency	Cumulative Frequency	Cum. Rel. Frequency
at or below		29.000		0	.0000	0	.0000
1	29.0000	37.000	33.000	2	.0333	2	.0333
2	37.0000	45.000	41.000	6	.1000	8	.1333
3	45.0000	53.000	49.000	8	.1333	16	.2667
4	53.0000	61.000	57.000	17	.2833	33	.5500
5	61.0000	69.000	65.000	13	.2167	46	.7667
6	69.0000	77.000	73.000	8	.1333	54	.9000
7	77.0000	85.000	81.000	5	.0833	59	.9833
8	85.0000	93.000	89.000	1	.0167	60	1.0000
above	93.0000			0	.0800	60	1.0000

Mean = 59.87 Standard Deviation = 12.50 Median = 59.45

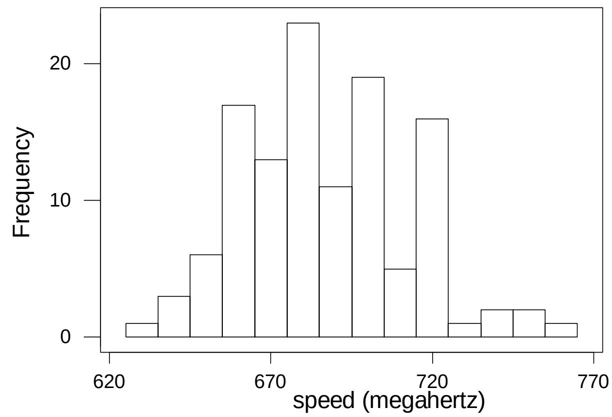


Yes, the histogram shows the same shape as the stem-and-leaf display.

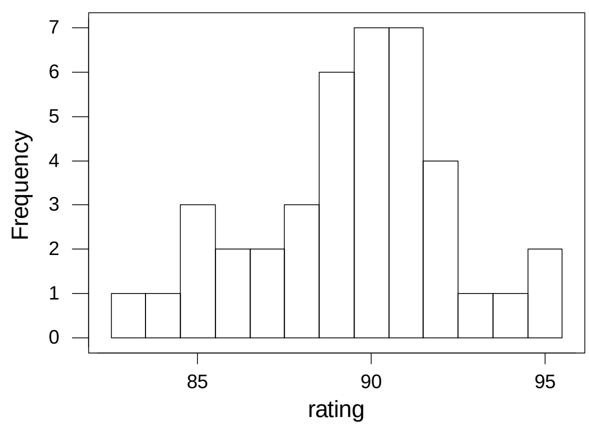
6-38 Yes, the histogram of the distance data shows the same shape as the stem-and-leaf display in exercise 6-25.



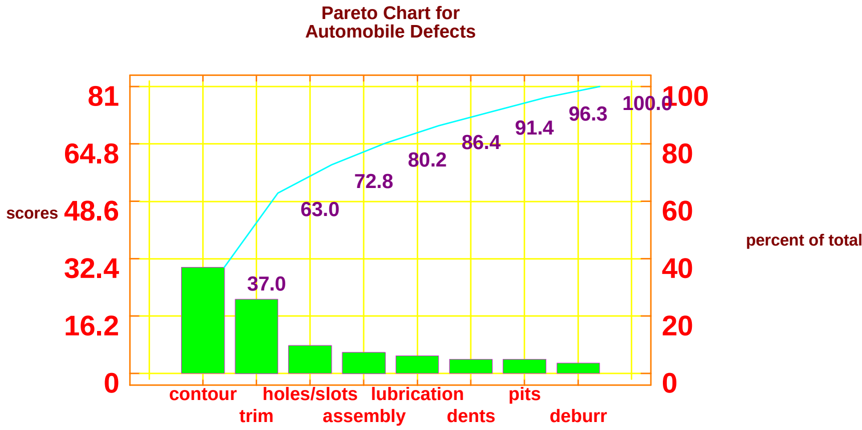
6-39 Histogram for the speed data in exercise 6-26. Yes, the histogram of the speed data shows the same shape as the stem-and-leaf display in exercise 6-26



6-40 Yes, the histogram of the wine rating data shows the same shape as the stem-and-leaf display in exercise 6-27.



6-41



Roughly 63% of defects are described by parts out of contour and parts under trimmed.

Section 6-5

6-42 Descriptive Statistics of O-ring joint temperature data

Variable	N	Mean	Median	TrMean	StDev	SE Mean
Temp	36	65.86	67.50	66.66	12.16	2.03

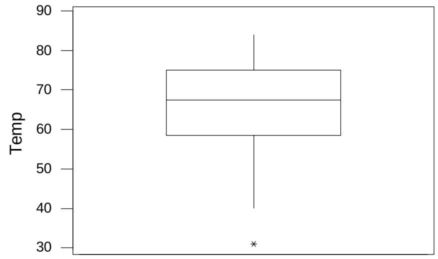
Variable	Minimum	Maximum	Q1	Q3
Temp	31.00	84.00	58.50	75.00

a.)
Lower Quartile: Q_1=58.50
Upper Quartile: Q_3=75.00
b.) Median = 67.50
c.) Data with lowest point removed

Variable	N	Mean	Median	TrMean	StDev	SE Mean
Temp	35	66.86	68.00	67.35	10.74	1.82

Variable	Minimum	Maximum	Q1	Q3
Temp	40.00	84.00	60.00	75.00

The mean and median have increased and the standard deviation and difference between the upper and lower quartile has decreased.
d.)Box Plot - The box plot indicates that there is an outlier in the data.

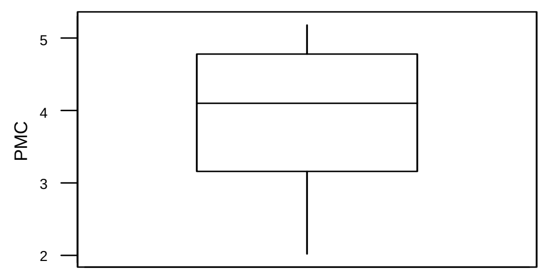


6-43. Descriptive Statistics

Variable	N	Mean	Median	Tr Mean	StDev	SE Mean
PMC	20	4.000	4.100	4.044	0.931	0.208

Variable	Min	Max	Q1	Q3
PMC	2.000	5.200	3.150	4.800

a) Sample Mean: 4
b) Sample Variance: 0.867
Sample Standard Deviation: 0.931
c)

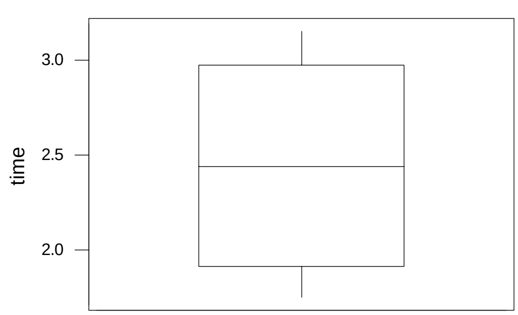


6-44 Descriptive Statistics

Variable	N	Mean	Median	TrMean	StDev	SE Mean
time	8	2.415	2.440	2.415	0.534	0.189

Variable	Minimum	Maximum	Q1	Q3
time	1.750	3.150	1.912	2.973

a.)Sample Mean: 2.415
Sample Standard Deviation: 0.543
b.) Box Plot – There are no outliers in the data.



6-45. Descriptive Statistics

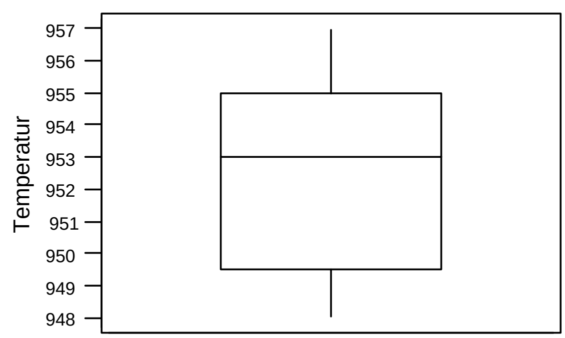
Variable	N	Mean	Median	Tr Mean	StDev	SE Mean
Temperat	9	952.44	953.00	952.44	3.09	1.03

Variable	Min	Max	Q1	Q3
Temperat	948.00	957.00	949.50	955.00

a) Sample Mean: 952.44
Sample Variance: 9.55
Sample Standard Deviation: 3.09

b) Median: 953; Any increase in the largest temperature measurement will not affect the median.

c)



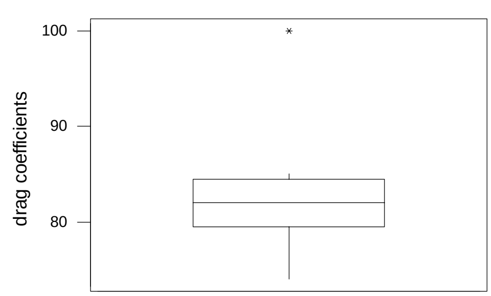
6-46 Descriptive statistics

Variable	N	Mean	Median	TrMean	StDev	SE Mean
drag coefficients	9	83.11	82.00	83.11	7.11	2.37

Variable	Minimum	Maximum	Q1	Q3
drag coefficients	74.00	100.00	79.50	84.50

a.) Upper quartile: Q_1 =79.50
Lower Quartile: Q_3=84.50

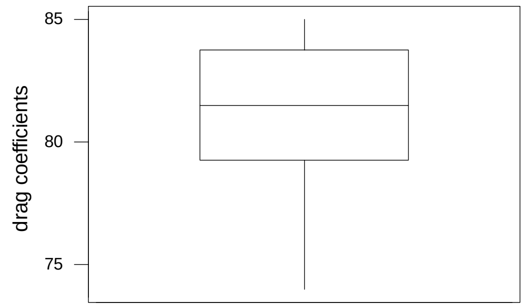
b.)



c.) Variable N Mean Median TrMean StDev SE Mean

drag coefficients	8	81.00	81.50	81.00	3.46	1.22

Variable	Minimum	Maximum	Q1	Q3
drag coefficients	74.00	85.00	79.25	83.75



Removing the largest observation (100) lowers the mean and median. Removing this "outlier also greatly reduces the variability as seen by the smaller standard deviation and the smaller difference between the upper and lower quartiles.

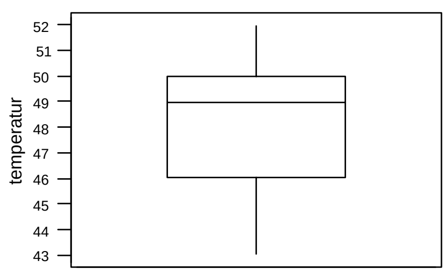
6-47.

Descriptive Statistics

Variable	N	Mean	Median	Tr Mean	StDev	SE Mean
temperat	24	48.125	49.000	48.182	2.692	0.549

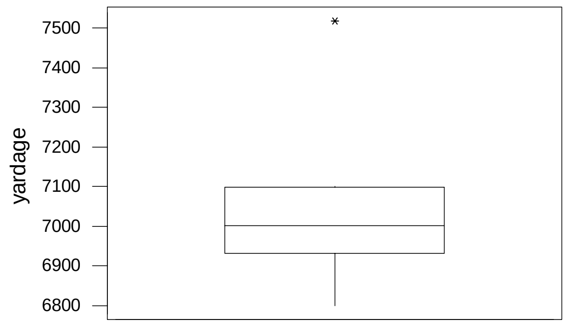
Variable	Min	Max	Q1	Q3
temperat	43.000	52.000	46.000	50.000

a) Sample Mean: 48.125
   Sample Median: 49
b) Sample Variance: 7.246
   Sample Standard Deviation: 2.692
c)



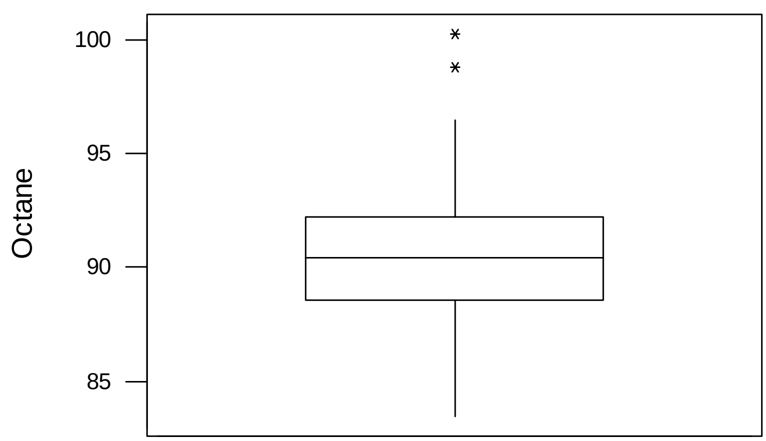
The data appear to be slightly skewed.

4-48 The golf course yardage data appear to be skewed. Also, there is an outlying data point above 7500 yards.



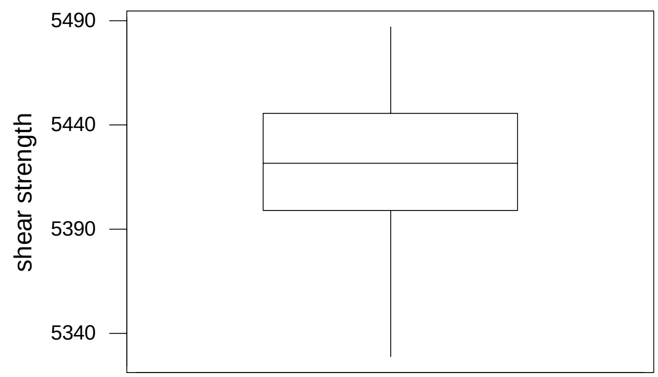
6-49

Boxplot for problem 6-49

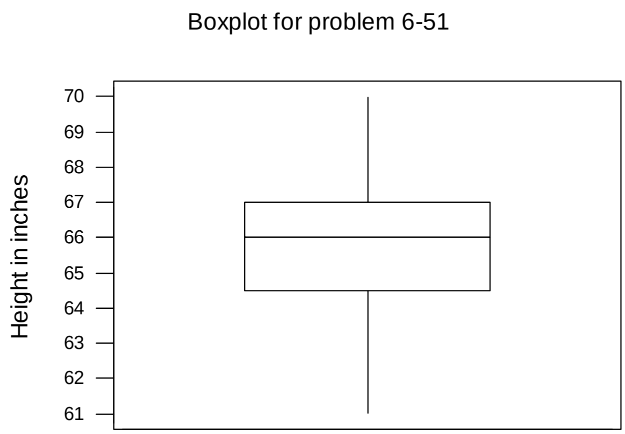


This plot conveys the same basic information as the stem and leaf plot but in a different format. The outliers that were separated from the main portion of the stem and leaf plot are shown here separated from the whiskers.

6-50 The box plot shows that the data are symmetrical about the mean. It also shows that there are no outliers in the data. These are the same interpretations seen in the stem-leaf-diagram.

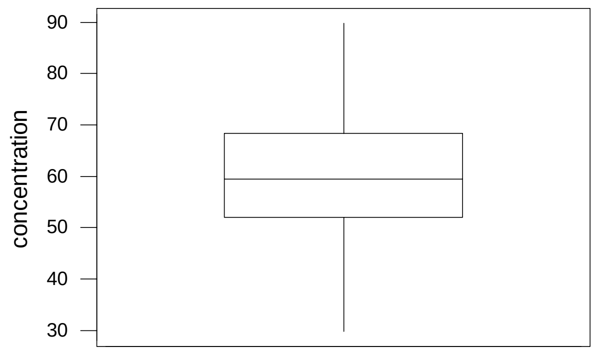


6-51

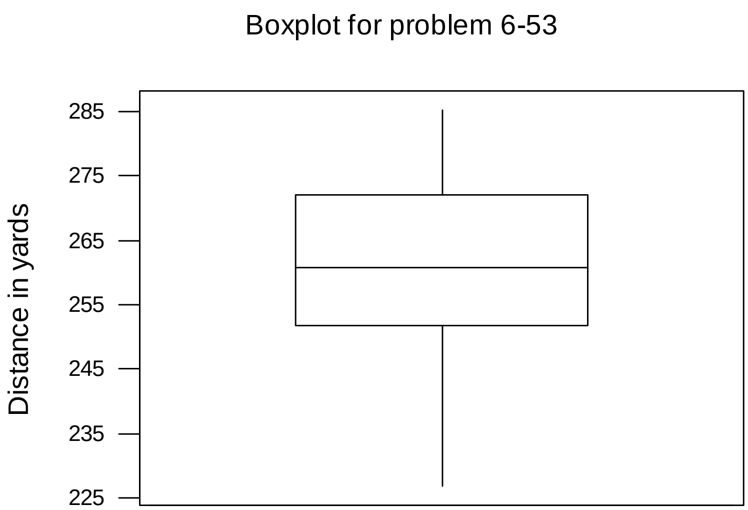


This plot, as the stem and leaf one, indicates that the data fall mostly in one region and that the measurements toward the ends of the range are more rare.

6-52 The box plot and the stem-leaf-diagram show that the data are very symmetrical about the mean. It also shows that there are no outliers in the data.

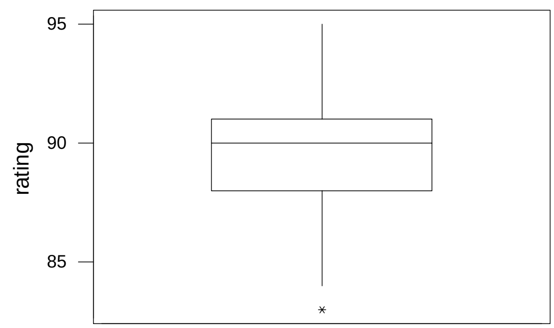


6-53

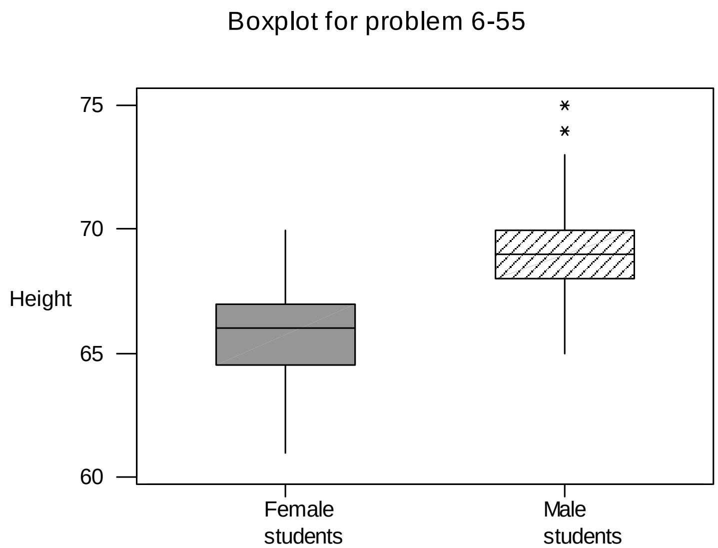


The plot indicates that most balls will fall somewhere in the 250-275 range. In general, the population is grouped more toward the high end of the region. This same type of information could have been obtained from the stem and leaf graph of problem 6-25.

6-54 The box plot shows that the data are not symmetrical about the mean. The data are skewed to the right and have a longer right tail (at the lower values). It also shows that there is an outlier in the data. These are the same interpretations seen in the stem-leaf-diagram.

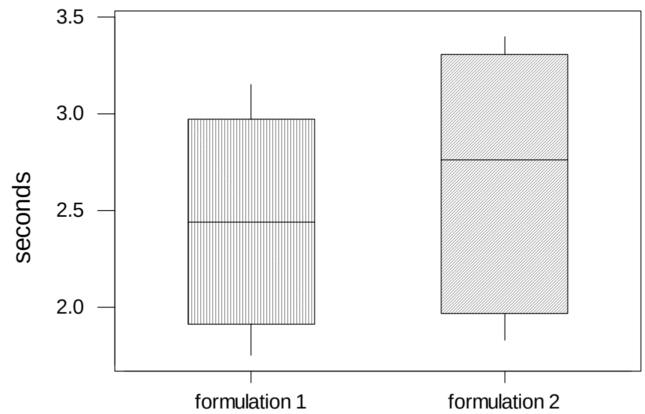


6-55



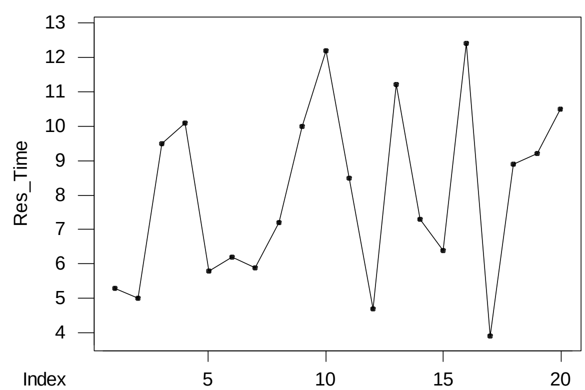
We can see that the two distributions seem to be centered at different values.

6-56 The box plot shows that there is a difference between the two formulations. Formulation 2 has a higher mean cold start ignition time and a larger variability in the values of the start times. The first formulation has a lower mean cold start ignition time and is more consistent. Care should be taken, though since these box plots for formula 1 and formula 2 are made using 8 and 10 data points respectively. More data should be collected on each formulation to get a better determination.



Section 6-6

6-57. Time Series Plot



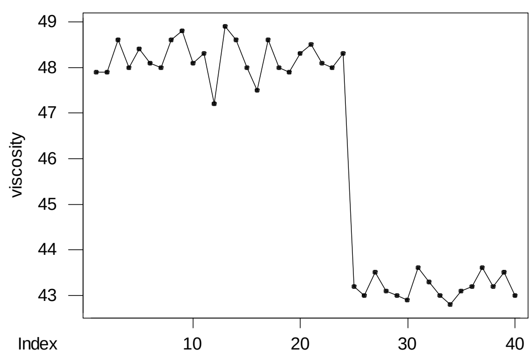
Computer response time appears random. No trends or patterns are obvious.

6-58 a.) Stem-leaf-plot of viscosity N = 40
Leaf Unit = 0.10

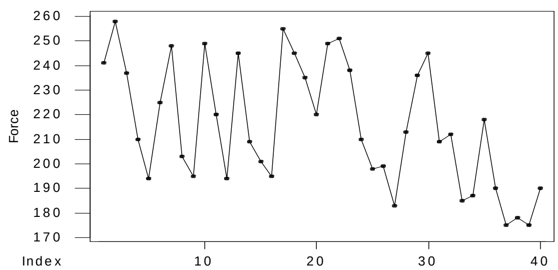
```
  2   42 89
 12   43 0000112223
 16   43 5566
 16   44
 16   44
 16   45
 16   45
 16   46
 16   46
 17   47 2
 (4)  47 5999
 19   48 000001113334
  7   48 5666689
```

The stem-leaf-plot shows that there are two "different" sets of data. One set of data is centered about 43 and the second set is centered about 48. The time series plot shows that the data starts out at the higher level and then drops down to the lower viscosity level at point 24. Each plot gives us a different set of information.

b.) If the specifications on the product viscosity are 48.0±2, then there is a problem with the process performance after data point 24. An investigation needs to take place to find out why the location of the process has dropped from around 48.0 to 43.0. The most recent product is not within specification limits.



6-59. a)



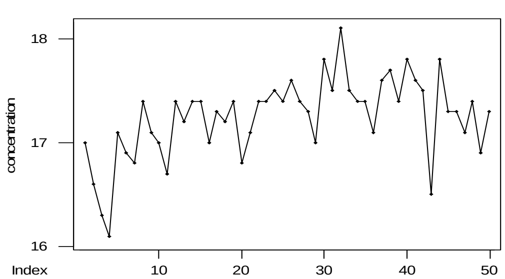
b) Stem-and-leaf display for Problem 2-23.Force: unit = 1   1|2 represents 12

```
  3   17|558
  6   18|357
 14   19|00445589
 18   20|1399
 (5)  21|00238
 17   22|005
 14   23|5678
 10   24|1555899
  3   25|158
```

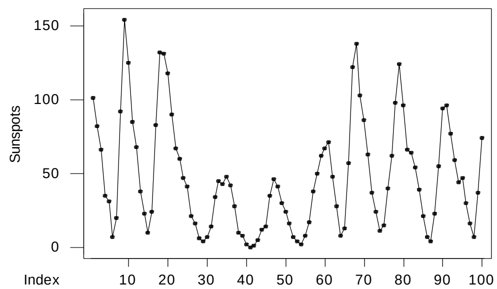
In the time series plot there appears to be a downward trend beginning after time 30. The stem and leaf plot does not reveal this.

6-60

```
1    18 | 1
4    17 | 888
8    17 | 6667
25   17 | 444444444444445555
(7)  17 | 2233333
18   17 | 000011111
9    16 | 8899
5    16 | 567
2    16 | 3
1    16 | 1
```



6-61 a)



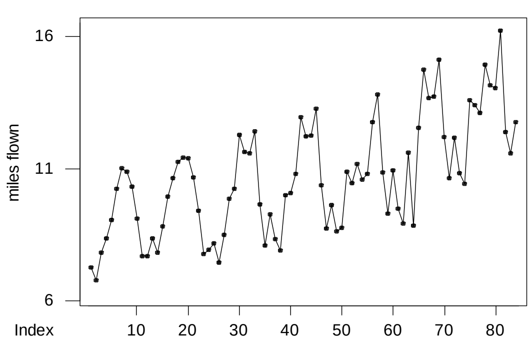
b) Stem-and-leaf display for Problem 2-25.Sunspots: unit = 1 1|2 represents 12

```
17    0|01224445677777888
29    1|001234456667
39    2|0113344488
50    3|00145567789
50    4|011234567788
38    5|04579
33    6|0223466778
23    7|147
20    8|2356
16    9|024668
10   10|13
 8   11|8
 7   12|245
 4   13|128

     HI|154
```

The data appears to decrease between 1790 and 1835, the stem and leaf plot indicates skewed data.

6-62 a.) Time Series Plot



Each year the miles flown peaks during the summer hours. The number of miles flown increased over the years 1964 to 1970.

b.) `Stem-and-leaf of miles fl  N  = 84`

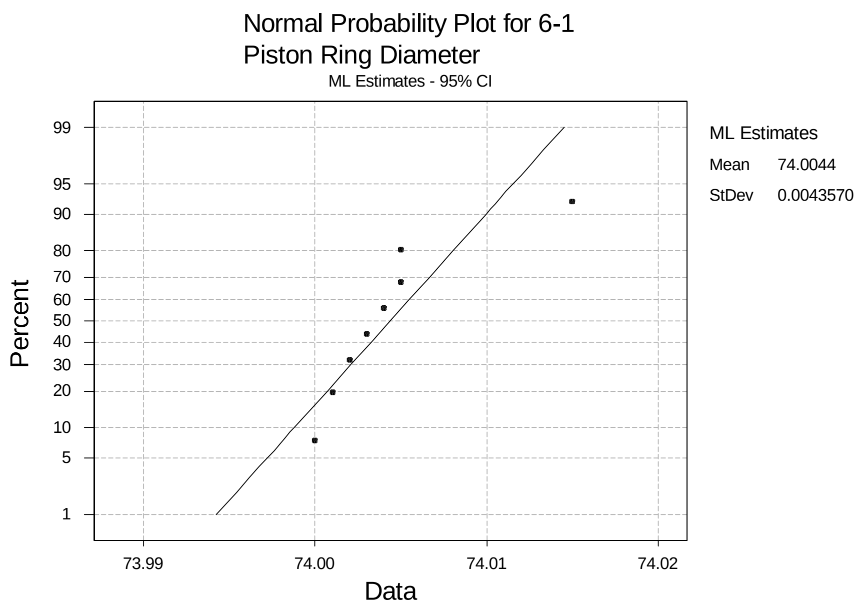
```
   Leaf Unit = 0.10

  1    6 7
 10    7 246678889
 22    8 013334677889
 33    9 01223466899
(18)  10 022334456667888889
 33   11 012345566
 24   12 11222345779
 13   13 1245678
  6   14 0179
  2   15 1
  1   16 2
```

When grouped together, the yearly cycles in the data are not seen. The data in the stem-leaf-diagram appear to be nearly normally distributed.

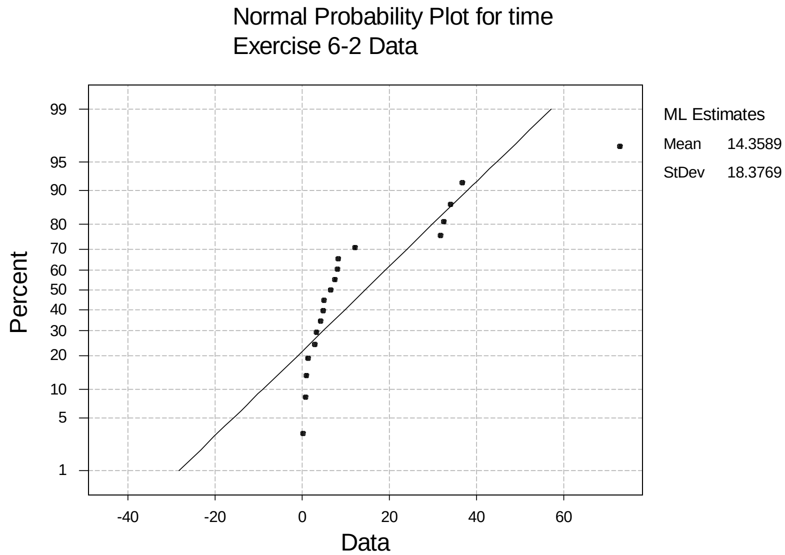
Section 6-7

6-63



The pattern of the data indicates that the sample may not come from a normally distributed population or that the largest observation is an outlier. Note the slight bending downward of the sample data at both ends of the graph.

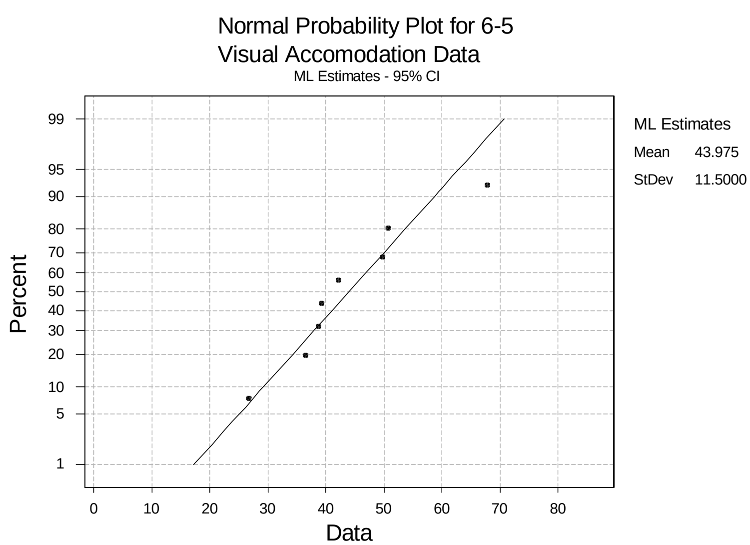
6-64



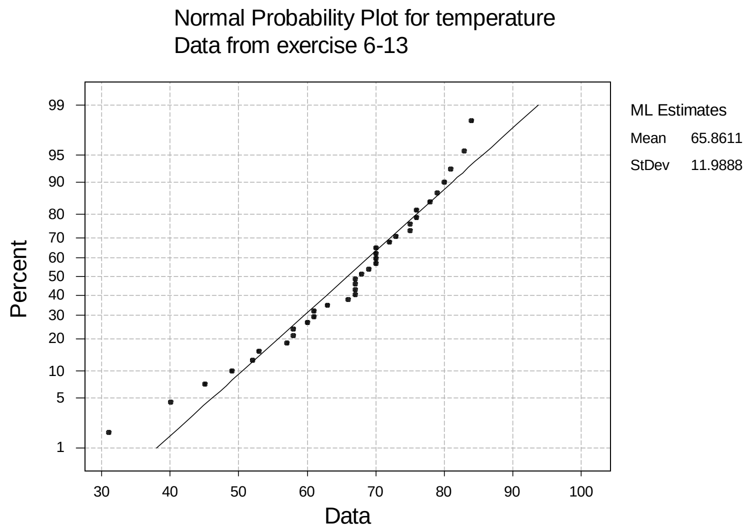
It appears that the data do not come from a normal distribution. Very few of the data points fall on the line.

6-65

There is no evidence to doubt that data are normally distributed

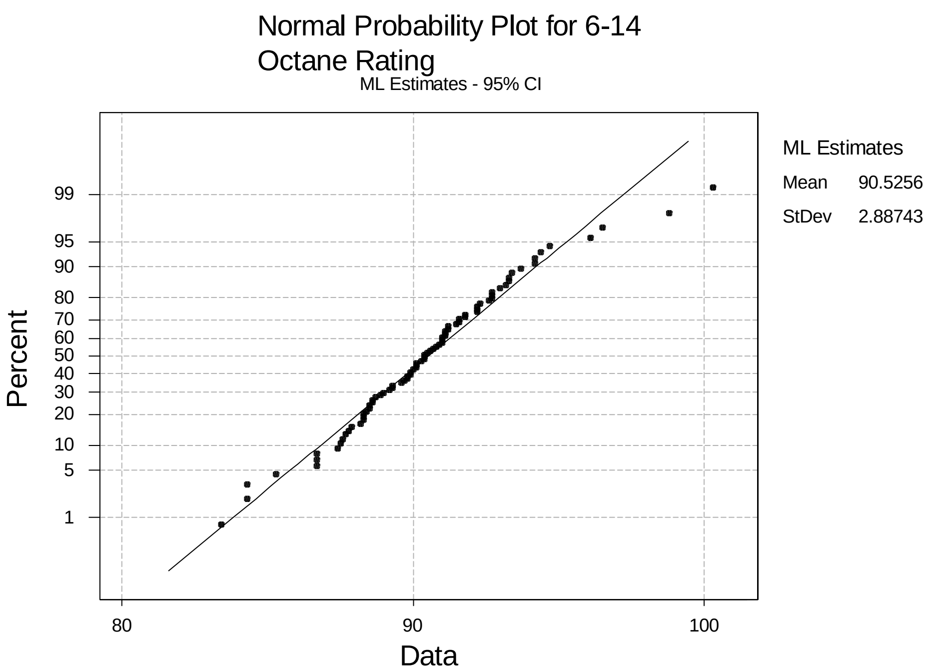


6-66



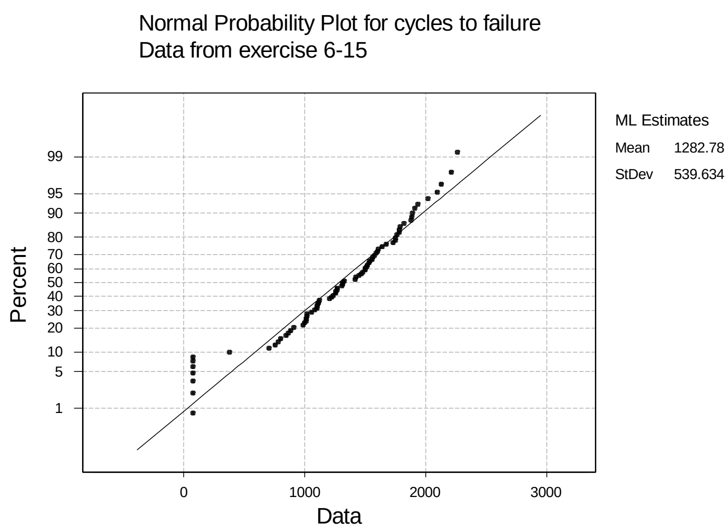
The data appear to be normally distributed. Although, there are some departures from the line at the ends of the distribution.

6-67



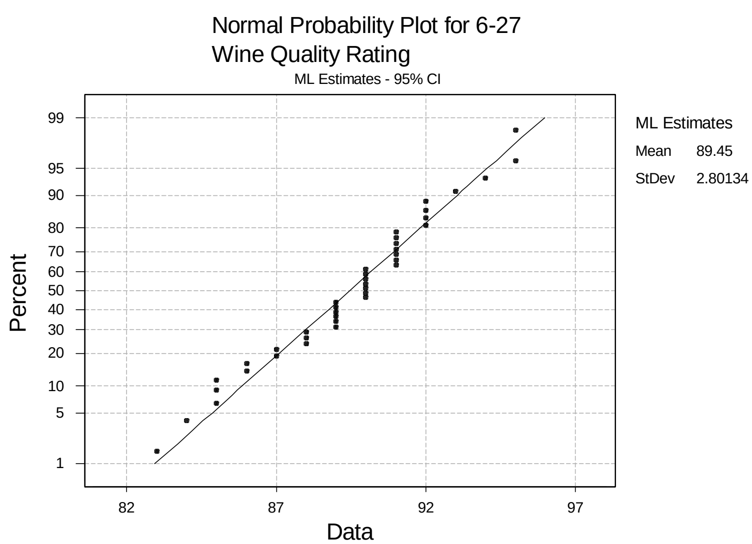
There are a few points outside the confidence limits, indicating that the sample is not perfectly normal. These deviations seem to be fairly small though.

6-68



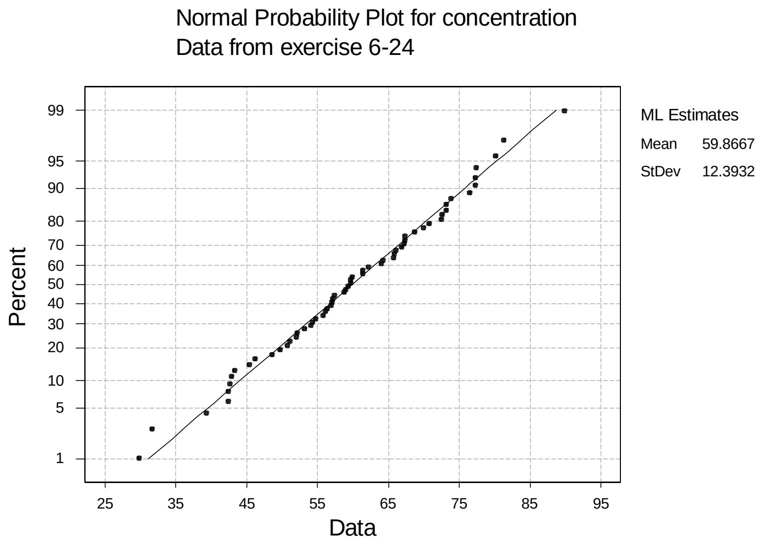
The data appear to be normally distributed. Although, there are some departures from the line at the ends of the distribution.

6-69



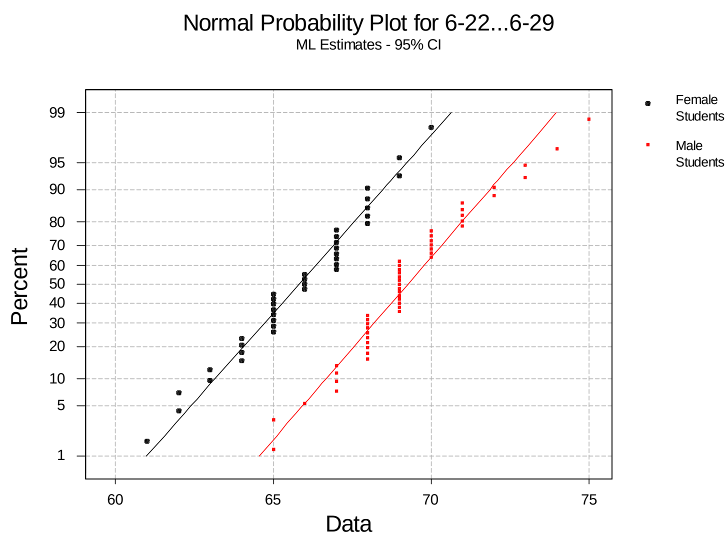
The data seem to be normally distributed. Notice that there are clusters of observations because of the discrete nature of the ratings.

6-70



The data appear to be normally distributed. Nearly all of the data points fall very close to, or on the line.

6-71



Both populations seem to be normally distributed, moreover, the lines seem to be roughly parallel indicating that the populations may have the same variance and differ only in the value of their mean.

6-72 Yes, it is possible to obtain an estimate of the mean from the 50th percentile value of the normal probability plot. The fiftieth percentile point is the point at which the sample mean should equal the population mean and 50% of the data would be above the value and 50% below. An estimate of the standard deviation would be to subtract the 50th percentile from the 64th percentile These values are based on the values from the z-table that could be used to estimate the standard deviation.

Supplemental Exercises

6-73. a) Sample Mean = 65.083
The sample mean value is close enough to the target value to accept the solution as conforming. There is a slight difference due to inherent variability.

b) $s^2 = 1.86869$ $s = 1.367$

c) A major source of variability might be variability in the reagent material. Furthermore, if the same setup is used for all measurements it is not expected to affect the variability. However, if each measurement uses a different setup, then setup differences could also be a major source of variability.

A low variance is desirable because it indicates consistency from measurement to measurement. This implies the measurement error has low variability.

6-74 a) $\sum_{i=1}^{6} x_i^2 = 10{,}433$ $\left(\sum_{i=1}^{6} x_i\right)^2 = 62{,}001$ $n = 6$

$$s^2 = \frac{\sum_{i=1}^{6} x_i^2 - \frac{\left(\sum_{i=1}^{6} x_i\right)^2}{n}}{n-1} = \frac{10{,}433 - \frac{62{,}001}{6}}{6-1} = 19.9\Omega^2$$

$$s = \sqrt{19.9\Omega^2} = 4.46\Omega$$

b) $\sum_{i=1}^{6} x_i^2 = 353$ $\left(\sum_{i=1}^{6} x_i\right)^2 = 1521$ $n = 6$

$$s^2 = \frac{\sum_{i=1}^{6} x_i^2 - \frac{\left(\sum_{i=1}^{6} x_i\right)^2}{n}}{n-1} = \frac{353 - \frac{1{,}521}{6}}{6-1} = 19.9\Omega^2$$

$$s = \sqrt{19.9\Omega^2} = 4.46\Omega$$

Shifting the data from the sample by a constant amount has no effect on the sample variance or standard deviation.

c) $\sum_{i=1}^{6} x_i^2 = 1043300$ $\left(\sum_{i=1}^{6} x_i\right)^2 = 6200100$ $n = 6$

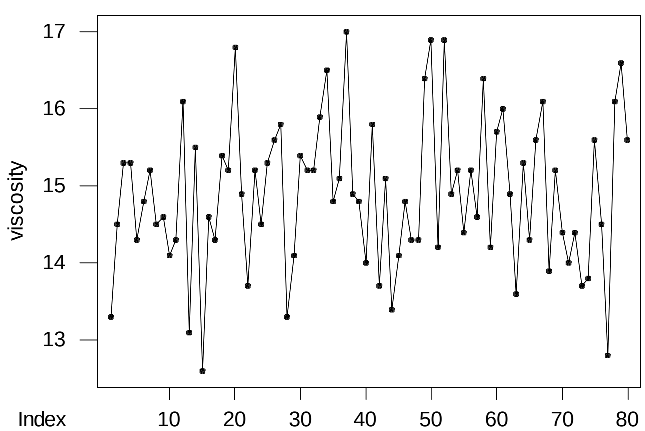
$$s^2 = \frac{\sum_{i=1}^{6} x_i^2 - \frac{\left(\sum_{i=1}^{6} x_i\right)^2}{n}}{n-1} = \frac{1043300 - \frac{6200100}{6}}{6-1} = 1990\Omega^2$$

$$s = \sqrt{1990\Omega^2} = 44.61\Omega$$

Yes, the rescaling is by a factor of 10. Therefore, s^2 and s would be rescaled by multiplying s^2 by 10^2 (resulting in $1990\Omega^2$) and s by 10 (44.6Ω).

6-75 a) Sample 1 Range = 4
Sample 2 Range = 4
Yes, the two appear to exhibit the same variability
b) Sample 1 s = 1.604
Sample 2 s = 1.852
No, sample 2 has a larger standard deviation.
c) The sample range is a relatively crude measure of the sample variability as compared to the sample standard deviation since the standard deviation uses the information from every data point in the sample whereas the range uses the information contained in only two data points - the minimum and maximum.

6-76 a.) It appears that the data may shift up and then down over the 80 points.



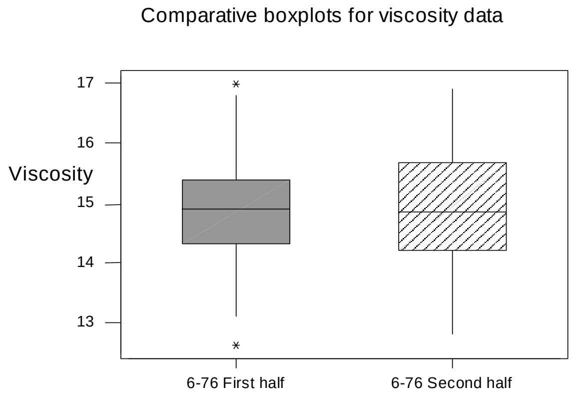
b.)It appears that the mean of the second set of 40 data points may be slightly higher than the first set of 40.

c.) Descriptive Statistics: viscosity 1, viscosity 2

Variable	N	Mean	Median	TrMean	StDev	SE Mean
Viscosity1	40	14.875	14.900	14.875	0.948	0.150
Viscosity2	40	14.923	14.850	14.914	1.023	0.162

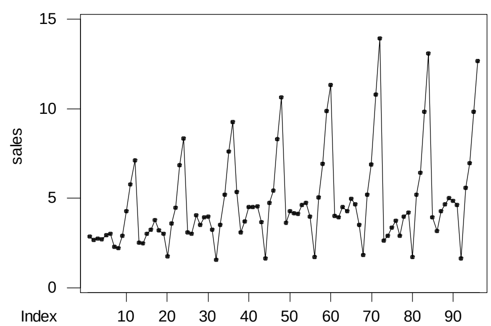
There is a slight difference in the mean levels and the standard deviations.

6-77



Both sets of data appear to have the same mean although the first half seem to be concentrated a little more tightly. Two data points appear as outliers in the first half of the data.

6-78



There appears to be a cyclic variation in the data with the high value of the cycle generally increasing. The high values are during the winter holiday months.

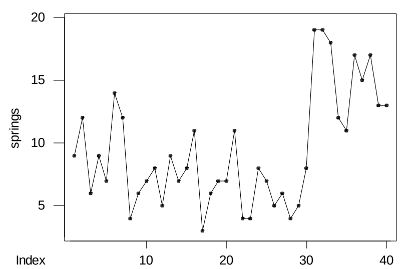
b) We might draw another cycle, with the peak similar to the last year's data (1969) at about 12.7 thousand bottles.

6-79 a) Stem-and-leaf display for Problem 2-35: unit = 1   1|2 represents 12

```
  1    0T|3
  8    0F|4444555
 18    0S|6666777777
 (7)   0o|8888999
 15    1*|111
 12    1T|22233
  7    1F|45
  5    1S|77
  3    1o|899
```

b) Sample Average = 9.325
Sample Standard Deviation = 4.4858

c)



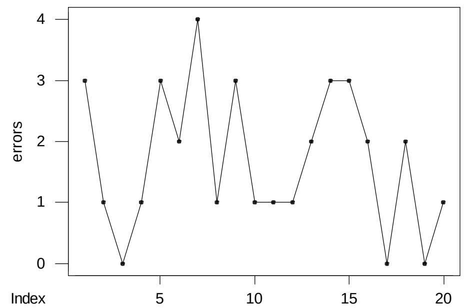
The time series plot indicates there was an increase in the average number of nonconforming springs made during the 40 days. In particular, the increase occurs during the last 10 days.

6-80 a.) Stem-and-leaf of errors   N = 20
Leaf Unit = 0.10

```
  3    0 000
 10    1 0000000
 10    2 0000
  6    3 00000
  1    4 0
```

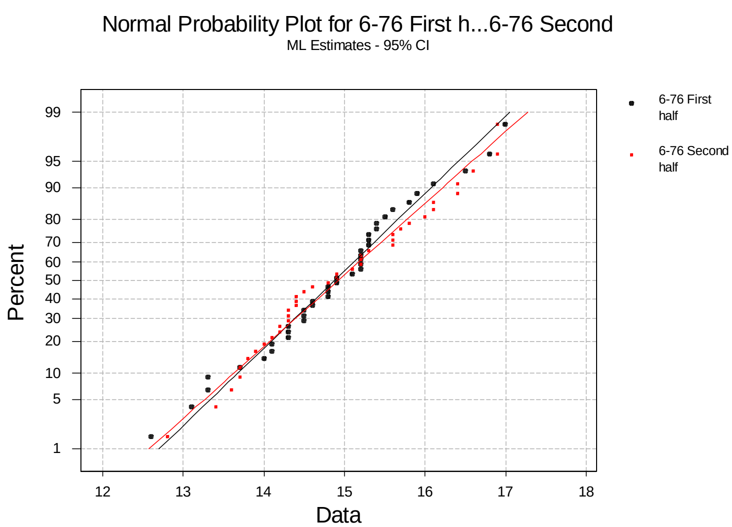
b.) Sample Average = 1.700
Sample Standard Deviation = 1.174

c.)



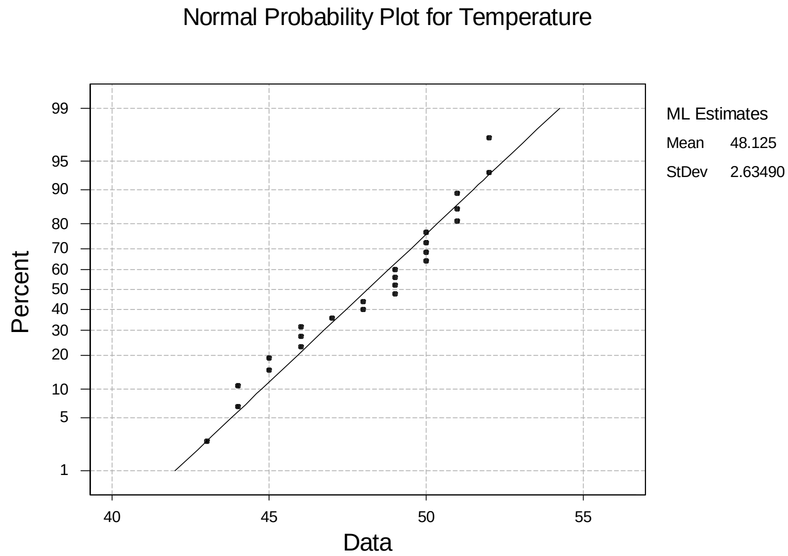
The time series plot indicates a slight decrease in the number of errors for strings 16 - 20.

6-81



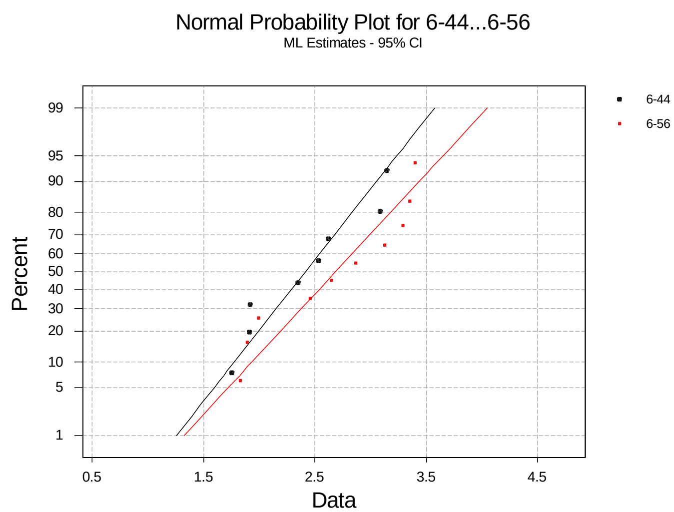
Both sets of data appear to be normally distributed and with roughly the same mean value. The difference in slopes for the two lines indicates that a change in variance might have occurred. This could have been the result of a change in processing conditions, the quality of the raw material or some other factor.

6-82



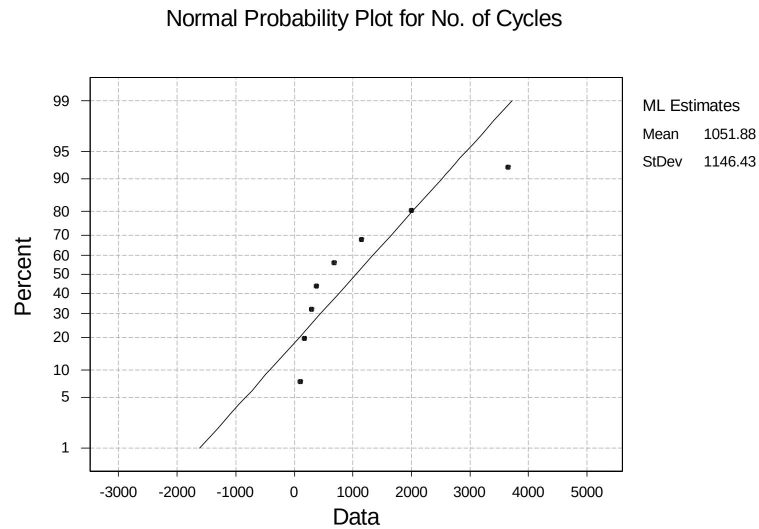
There appears to be no evidence that the data are not normally distributed. There are some repeat points in the data that cause some points to fall off the line.

6-83



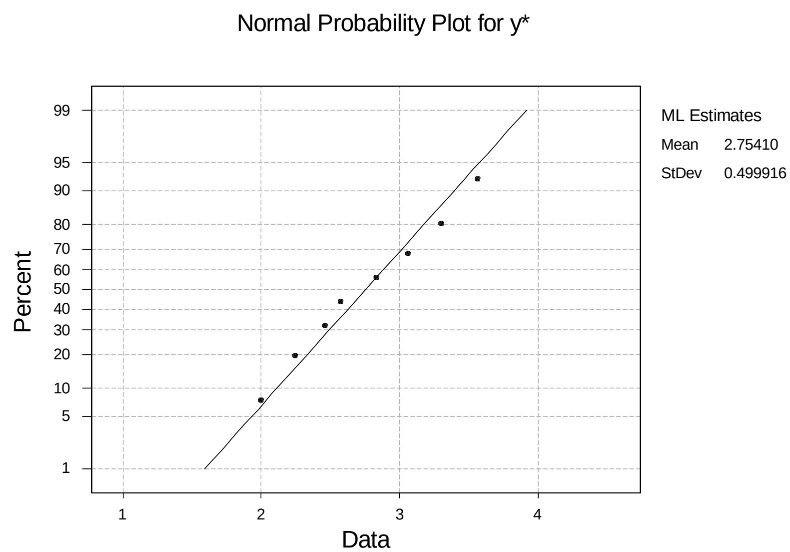
Although we do not have sufficient data points to really see a pattern, there seem to be no significant deviations from normality for either sample. The large difference in slopes indicates that the variances of the populations are very different.

6-84 a.)



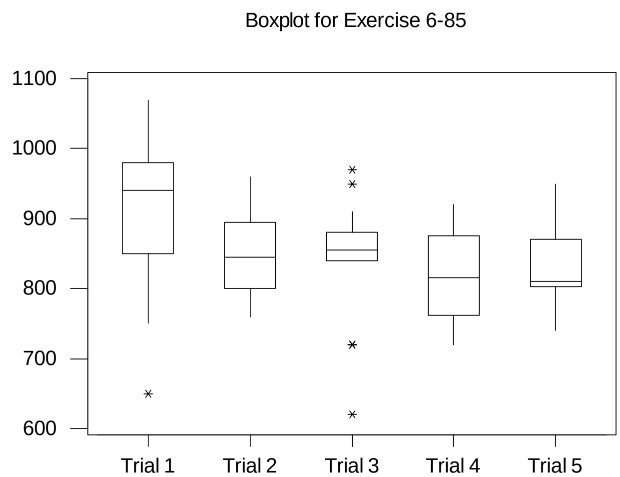
The data do not appear to be normally distributed. There is a curve in the line.

b.)



After the transformation y*=log(y), the normal probability plot shows no evidence that the data are not normally distributed.

6-85

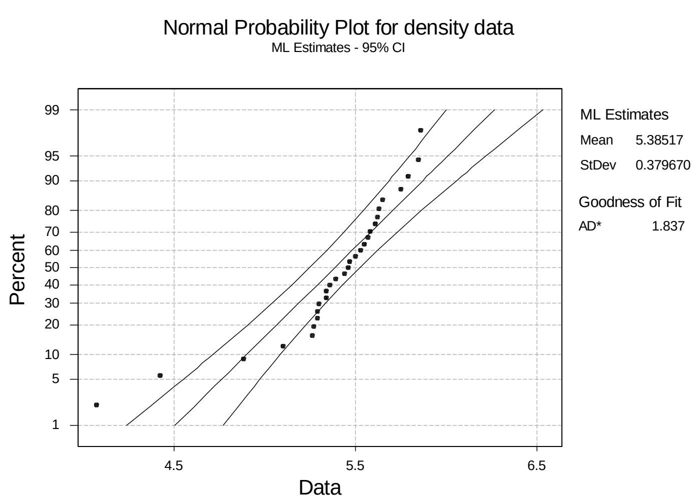


- There is a difference in the variability of the measurements in the trials. Trial 1 has the most variability in the measurements. Trial 3 has a small amount of variability in the main group of measurements, but there are four outliers. Trial 5 appears to have the least variability without any outliers.
- All of the trials except Trial 1 appear to be centered around 850. Trial 1 has a higher mean value
- All five trials appear to have measurements that are greater than the "true" value of 734.5.
- The difference in the measurements in Trial 1 may indicate a "start-up" effect in the data. There could be some bias in the measurements that is centering the data above the "true" value.

6-86 a.) Descriptive Statistics

Variable	N	Mean	Median	TrMean	StDev	SE Mean
density	29	5.4541	5.4600	5.4611	0.4072	0.0756

b.) There does appear to be a low outlier in the data.



c.)Due to the very low data point at 4.07, the mean may be lower than it should be. Therefore, the median would be a better estimate of the density of the earth. The median is not affected by outliers.

Mind Expanding Exercises

6-87 $\sum_{i=1}^{9} x_i^2 = 62572$ $\left(\sum_{i=1}^{9} x_i\right)^2 = 559504$ $n = 9$

$$s^2 = \frac{\sum_{i=1}^{9} x_i^2 - \frac{\left(\sum_{i=1}^{9} x_i\right)^2}{n}}{n-1} = \frac{62572 - \frac{559504}{9}}{9-1} = 50.61$$

$s = \sqrt{50.61} = 7.11$

Subtract 30 and multiply by 10

$\sum_{i=1}^{9} x_i^2 = 2579200$ $\left(\sum_{i=1}^{9} x_i\right)^2 = 22848400$ $n = 9$

$$s^2 = \frac{\sum_{i=1}^{9} x_i^2 - \frac{\left(\sum_{i=1}^{9} x_i\right)^2}{n}}{n-1} = \frac{2579200 - \frac{22848400}{9}}{9-1} = 5061.1$$

$s = \sqrt{5061.1} = 71.14$

Yes, the rescaling is by a factor of 10. Therefore, s^2 and s would be rescaled by multiplying s^2 by 10^2 (resulting in 5061.1) and s by 10 (71.14). Subtracting 30 from each value has no affect on the variance or standard deviation. This is because $V(aX + b) = a^2V(X)$.

6-88 $\sum_{i=1}^{n}(x_i - a)^2 = \sum_{i=1}^{n}(x_i - \bar{x}) + n(\bar{x} - a)^2$; The sum written in this form shows that the quantity is minimized when a = $\bar{x}$.

6-89 Of the two quantities $\sum_{i=1}^{n}(x_i - \bar{x})^2$ and $\sum_{i=1}^{n}(x_i - \mu)^2$, the quantity $\sum_{i=1}^{n}(x_i - \bar{x})^2$ will be smaller given that $\bar{x} \neq \mu$. This is because $\bar{x}$ is based on the values of the x_i 's. The value of μ may be quite different for this sample.

6-90 $y_i = a + bx_i$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \quad \text{and} \quad \bar{y} = \frac{\sum_{i=1}^{n}(a + bx_i)}{n} = \frac{na + b\sum_{i=1}^{n} x_i}{n} = a + b\bar{x}$$

$$s_x^{\ 2} = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n-1} \text{ and } s_x = \sqrt{s_x^{\ 2}}$$

$$s_y^{\ 2} = \frac{\sum_{i=1}^{n}(a + bx_i - a - b\bar{x})^2}{n-1} = \frac{\sum_{i=1}^{n}(bx_i - b\bar{x})^2}{n-1} = \frac{b^2\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = b^2 s_x^{\ 2}$$

Therefore, $s_y = bs_x$

6-91 $\bar{x} = 835.00$ °F $\quad s_x = 10.5$ °F

The results in °C:

$\bar{y} = -32 + 5/9\bar{x} = -32 + 5/9(835.00) = 431.89$ °C

$s_y^{\ 2} = b^2 s_x^{\ 2} = (5/9)^2(10.5)^2 = 34.028$ °C

6-92 Using the results found in Exercise 6-90 with a = $-\frac{\bar{x}}{s}$ and b = 1/*s*, the mean and standard deviation of the z_i are $\bar{z} = 0$ and $s_Z = 1$.

6-93. Yes, in this case, since no upper bound on the last electronic component is available, use a measure of central location that is not dependent on this value. That measure is the median.

Sample Median = $\frac{x_{(4)} + x_{(5)}}{2} = \frac{63 + 75}{2} = 69 \quad hours$

6-94 a) $\bar{x}_{n+1} = \dfrac{\sum_{i=1}^{n+1} x_i}{n+1} = \dfrac{\sum_{i=1}^{n} x_i + x_{n+1}}{n+1}$

$$\bar{x}_{n+1} = \frac{n\bar{x}_n + x_{n+1}}{n+1}$$

$$\bar{x}_{n+1} = \frac{n}{n+1}\bar{x}_n + \frac{x_{n+1}}{n+1}$$

b) $ns_{n+1}^2 = \sum_{i=1}^{n} x_i^2 + x_{n+1}^2 - \dfrac{\left(\sum_{i=1}^{n} x_i + x_{n+1}\right)^2}{n+1}$

$$= \sum_{i=1}^{n} x_i^2 + x_{n+1}^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n+1} - \frac{2x_{n+1}\sum_{i=1}^{n} x_i}{n+1} - \frac{x_{n+1}^2}{n+1}$$

$$= \sum_{i=1}^{n} x_i^2 + \frac{n}{n+1}x_{n+1}^2 - \frac{n}{n+1}2x_{n+1}\bar{x}_n - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n+1}$$

$$= \left[\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum x_i\right)^2}{n+1}\right] + \frac{n}{n+1}\left[x_{n+1}^2 - 2x_{n+1}\bar{x}_n\right]$$

$$= \sum_{i=1}^{n} x_i^2 + \left[\frac{\left(\sum x_i\right)^2}{n} - \frac{\left(\sum x_i\right)^2}{n}\right] - \frac{\left(\sum x_i\right)^2}{n+1} + \frac{n}{n+1}\left[x_{n+1}^2 - 2x_n\bar{x}_n\right]$$

$$= \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum x_i\right)^2}{n} + \frac{(n+1)\left(\sum x_i\right)^2 - n\left(\sum x_i\right)^2}{n(n+1)} + \frac{n}{n+1}\left[x_{n+1}^2 - 2x_n\bar{x}_n\right]$$

$$= (n-1)s_n^2 + \frac{\left(\sum x_i\right)^2}{n(n+1)} + \frac{n}{n+1}\left[x_{n+1}^2 - 2x_n\bar{x}_n\right]$$

$$= (n-1)s_n^2 + \frac{n\bar{x}^2}{n+1} + \frac{n}{n+1}\left[x_{n+1}^2 - 2x_n\bar{x}_n\right]$$

$$= (n-1)s_n^2 + \frac{n}{n+1}\left(x^2{}_{n+1} - 2x_n\bar{x}_n + \bar{x}_n^2\right)$$

$$= (n-1)s_n^2 + \frac{n}{n+1}\left(x_{n+1}^2 - \bar{x}_n\right)^2$$

c) $\bar{x}_n = 65.811$ inches $\quad x_{n+1} = 64$

$s_n^{\ 2} = 4.435 \quad n = 37 \ s_n = 2.106$

$$\bar{x}_{n+1} = \frac{37(65.81) + 64}{37 + 1} = 65.76$$

$$s_{n+1} = \sqrt{\frac{(37-1)4.435 + \frac{37}{37+1}(64 - 65.811)^2}{37}}$$

$$= 2.098$$

6-95. The trimmed mean is pulled toward the median by eliminating outliers.
a) 10% Trimmed Mean = 89.29
b) 20% Trimmed Mean = 89.19
Difference is very small
c) No, the differences are very small, due to a very large data set with no significant outliers.

6-96. If nT/100 is not an integer, calculate the two surrounding integer values and interpolate between the two. For example, if nT/100 = 2/3, one could calculate the mean after trimming 2 and 3 observations from each end and then interpolate between these two means.