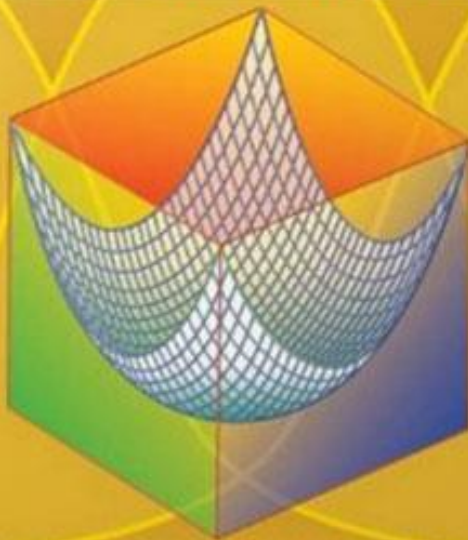


SOLUTIONS MANUAL



Fourth Edition

ENGINEERING STATISTICS



MONTGOMERY RUNGER HUBELE

Chapter 2

Section 2-1

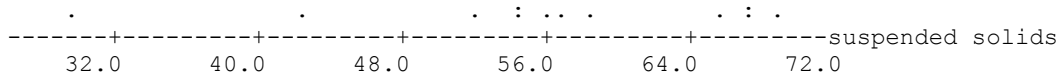
2-1. **Sample average:** $\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{\sum_{i=1}^{12} x_i}{12} = \frac{673.1}{12} = 56.09$

Sample standard deviation:

$$\sum_{i=1}^{12} x_i = 673.10 \quad \sum_{i=1}^{12} x_i^2 = 39168$$

$$s = \sqrt{\frac{\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n}}{n-1}} = \sqrt{\frac{39168 - \frac{(673.10)^2}{12}}{12-1}} = \sqrt{\frac{1412.70}{11}} = \sqrt{128.43} = 11.33$$

Dot diagram:



2-2. **Sample average:** $\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{\sum_{i=1}^{19} x_i}{19} = \frac{272.82}{19} = 14.36 \text{ min}$

Sample standard deviation:

$$\sum_{i=1}^{19} x_i = 272.82 \quad \sum_{i=1}^{19} x_i^2 = 10334$$

$$s = \sqrt{\frac{\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n}}{n-1}} = \sqrt{\frac{10334 - \frac{(272.82)^2}{19}}{19-1}} = \sqrt{\frac{6416.59}{18}} = \sqrt{356.48 \text{ (min)}^2} = 18.88 \text{ min}$$

Dot diagram



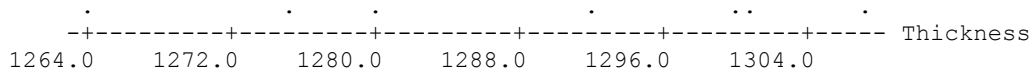
2-3. **Sample average:** $\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{\sum_{i=1}^7 x_i}{7} = \frac{9019}{7} = 1288.43 \text{ angstroms}$

Sample standard deviation:

$$\sum_{i=1}^7 x_i = 9019 \quad \sum_{i=1}^7 x_i^2 = 11621835$$

$$s = \sqrt{\frac{\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n}}{n-1}} = \sqrt{\frac{11621835 - \frac{(9019)^2}{7}}{7-1}} = \sqrt{\frac{1497.71}{6}} = \sqrt{249.62 \text{ (angstroms)}^2} = 15.80 \text{ angstroms}$$

Dot diagram:



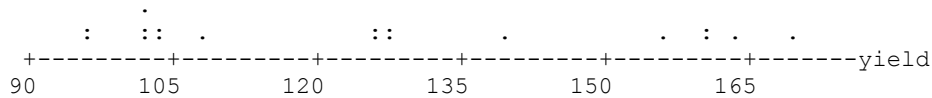
2-4. **Sample average:** $\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{\sum_{i=1}^{18} x_i}{18} = \frac{2272}{18} = 126.22 \text{ kN}$

Sample standard deviation:

$$\sum_{i=1}^{18} x_i = 2272 \quad \sum_{i=1}^{18} x_i^2 = 298392$$

$$s^2 = \sqrt{\frac{\sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n}}{n-1}} = \sqrt{\frac{298392 - \frac{(2272)^2}{18}}{18-1}} = \sqrt{\frac{11615.11}{17}} = \sqrt{683.24 \text{ (kN)}^2} = 26.14 \text{ kN}$$

Dot Diagram:



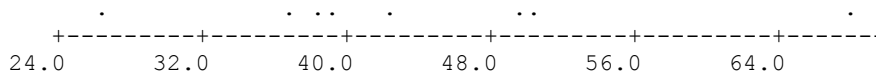
2-5. **Sample average:** $\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{\sum_{i=1}^8 x_i}{8} = \frac{351.8}{8} = 43.98$

Sample standard deviation:

$$\sum_{i=1}^8 x_i = 351.8 \quad \sum_{i=1}^8 x_i^2 = 16528.40$$

$$s = \sqrt{\frac{\sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n}}{n-1}} = \sqrt{\frac{16528.40 - \frac{(351.8)^2}{8}}{8-1}} = \sqrt{\frac{1058}{7}} = \sqrt{151.143} = 12.29$$

Dot diagram:



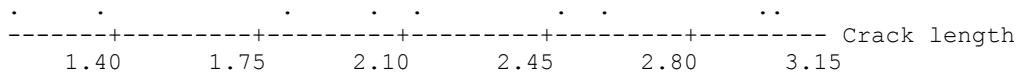
2-6. **Sample average:** $\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{\sum_{i=1}^9 x_i}{9} = \frac{19.56}{9} = 2.173 \text{ mm}$

Sample standard deviation:

$$\sum_{i=1}^9 x_i = 19.56 \quad \sum_{i=1}^9 x_i^2 = 45.953$$

$$s = \sqrt{\frac{\sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n}}{n-1}} = \sqrt{\frac{45.953 - \frac{(19.56)^2}{9}}{9-1}} = \sqrt{\frac{3.443}{8}} = \sqrt{0.4304} = 0.6560 \text{ mm}$$

Dot Diagram:



2-7. **Sample average:**

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

Sample variance:

$$\sum_{i=1}^{35} x_i = 28368$$

$$\sum_{i=1}^{35} x_i^2 = 23552500$$

$$s^2 = \frac{\sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n}}{n-1} = \frac{23552500 - \frac{(28368)^2}{35}}{35-1} = \frac{559830.743}{34} = 16465.61 \text{ (watts/m}^2\text{)}^2$$

Sample standard deviation:

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

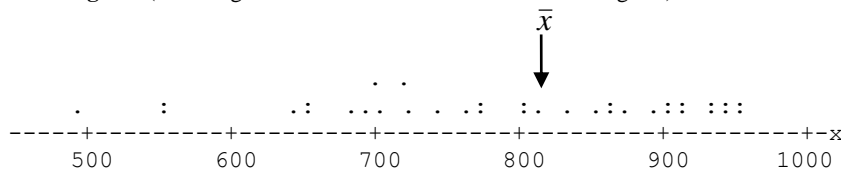
The sample standard deviation could also be found using

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

where

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

Dot Diagram (rounding of the data is used to create the dot diagram)



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

2-8. **High Dose Group:**

Sample average:

$$\bar{x} = \frac{\sum_{i=1}^{22} x_i}{22} = \frac{1158.2}{22} = 52.65$$

Sample variance:

$$\sum_{i=1}^{22} x_i = 1158.2$$

$$\sum_{i=1}^{22} x_i^2 = 92270.6$$

$$s^2 = \frac{\sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n}}{n-1} = \frac{92270.6 - \frac{(1158.2)^2}{22}}{22-1} = \frac{31296.63}{21} = 1490.32$$

Sample standard deviation:

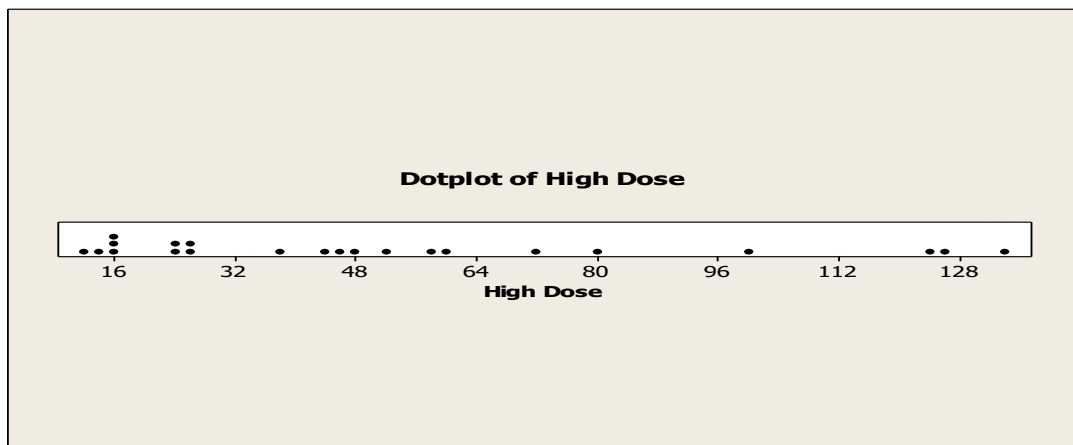
$$s = \sqrt{1490.32} = 38.60$$

The sample standard deviation could also be found using

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

where

$$\sum_{i=1}^{22} (x_i - \bar{x})^2 = 31296.6$$



Control Group:

Sample average:

$$\bar{x} = \frac{\sum_{i=1}^{22} x_i}{22} = \frac{8418.7}{22} = 382.67$$

Sample variance:

$$\sum_{i=1}^{22} x_i = 8418.7$$

$$\sum_{i=1}^{22} x_i^2 = 6901280$$

$$s^2 = \frac{\sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n}}{n-1} = \frac{6901280 - \frac{(8418.7)^2}{22}}{22-1} = \frac{3679711.38}{21} = 175224.35$$

Sample standard deviation:

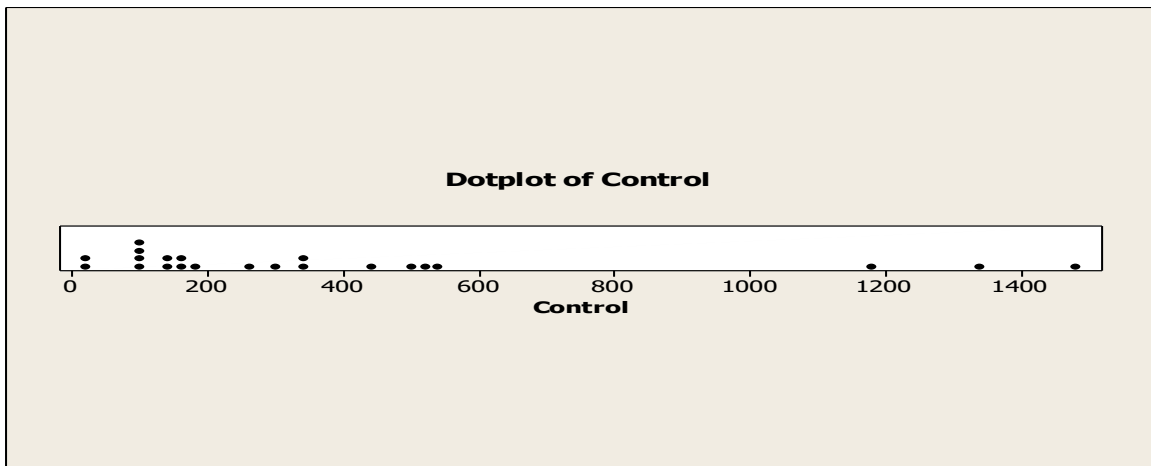
$$s = \sqrt{175224.35} = 418.60$$

The sample standard deviation could also be found using

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

where

$$\sum_{i=1}^{22} (x_i - \bar{x})^2 = 3679711.59$$



The control group has higher variance.

2-9. The only two data sets that may have resulted from a designed experiment is in Exercise 2-4 and 2-8.

Section 2-2

2-10. The stem and leaf display for weld strength $N = 100$
Leaf Unit = 1.0

```
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 01111556
33 544 00012455678
22 545 233447899
13 546 23569
8  547 357
5  548 11257
```

2-11. a) Stem-and-leaf display for cycles: 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|44444555555555
22  1S|66667777777
11  1o|888899
5   2*|011
2   2T|22
```

b) No, only 5/70 survived beyond 2000 cycles.

2-12. Stem-and-leaf of Suspended solids $N = 60$
Leaf Unit = 1.0

```
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
```

2-13. Stem-and-leaf display for yield: unit = 1 |2 represents 12

```

1    70|8
1    8*|
7    8T|223333
21   8F|44444444555555
38   8S|6666666666777777
(11) 80|88888999999
41   9*|00000000001111
27   9T|22233333
19   9F|444444445555
7    9S|666677
1    90|8

```

2-14. Stem-and-leaf of High Dose N = 22
Leaf Unit = 1.0

```

5    1    24566
9    2    3456
10   3     8
(3)  4    367
9    5     28
7    6     0
6    7    29
4    8
4    9     9
3   10
3   11
3   12    46
1   13     4

```

It's not symmetric – right skewed.

Stem-and-leaf of Control N = 22
Leaf Unit = 100

```

11   0    000011111111
11   0    2233
7    0    4455
3    0
3    0
3    1    1
2    1    3
1    1    4

```

It's not symmetric – right skewed.

Their shapes are similar.

2-15. Stem-and-leaf of solar intensity measurements N = 35
Leaf Unit = 10

```

1    4    9
1    5
3    5    56
3    6
7    6    5569
10   7    003
14   7    5677
(4)  8    0023
17   8    56779
12   9    00113344
4    9    5556

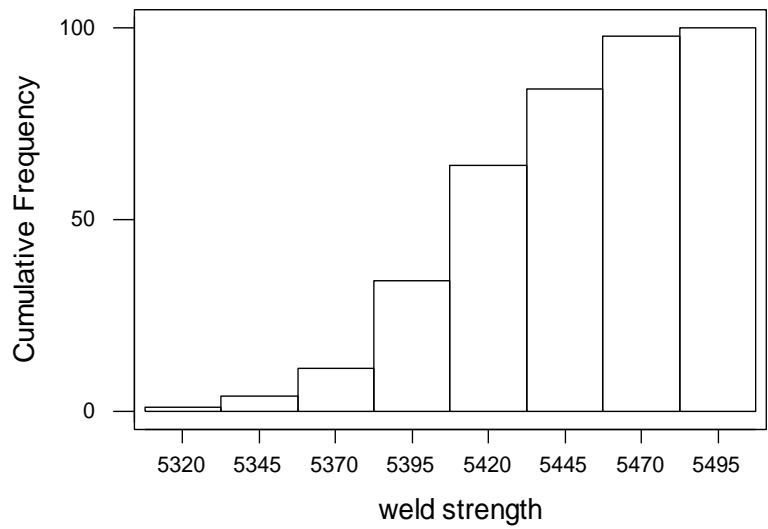
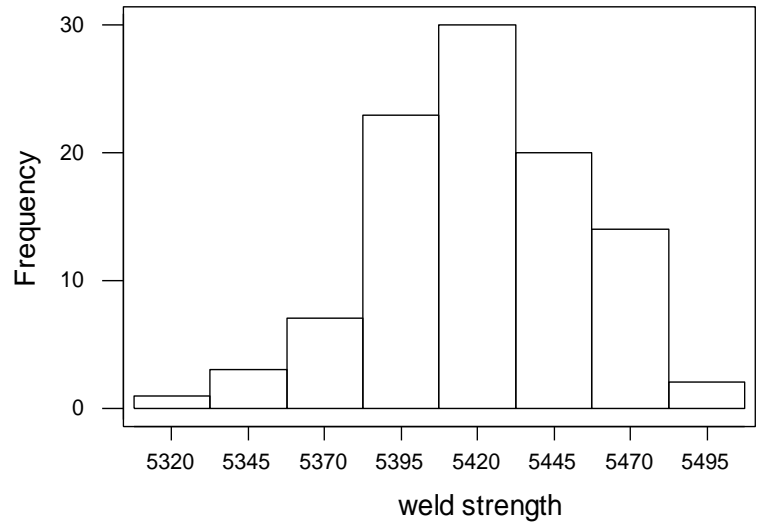
```

It's not symmetric – left skewed.

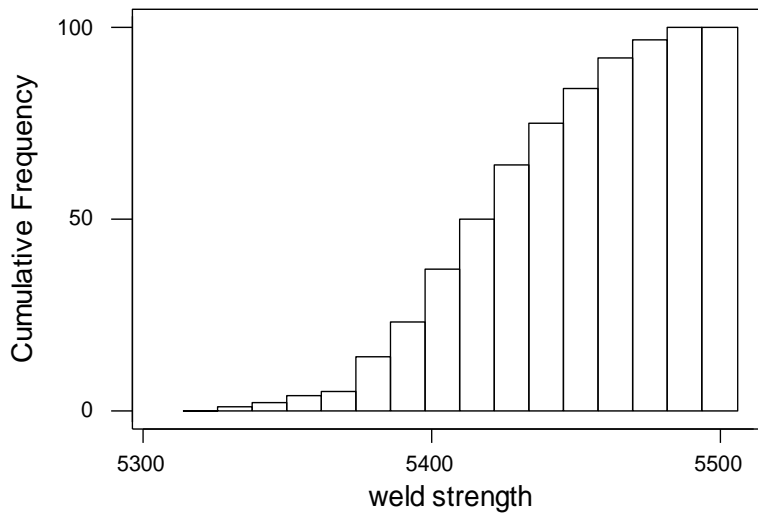
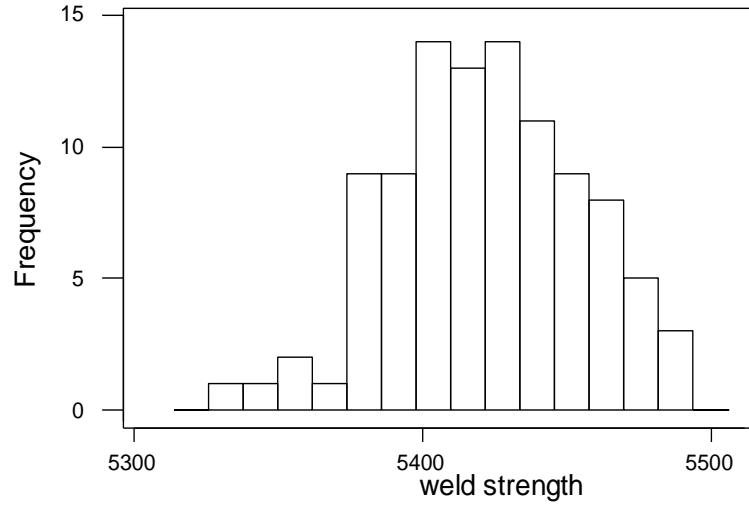
2-16.	Variable Weld strength	N 100	Median 5421.5	Q1 5399.0	Q3 5445.8	5 th 5366.45	95 th 5480.8
2-17.	Variable Cycles	N 70	Median 1436.5	Q1 1097.8	Q3 1735.0	5 th 772.85	95 th 2113.5
2-18.	Variable Solids	N 60	Median 59.45	Q1 52.03	Q3 68.35	5 th 39.455	95 th 79.965
2-19.	Variable Yield	N 90	Median 89.25	Q1 86.10	Q3 93.125	5 th 83.055	95 th 96.58

Section 2-3

2-20. a) 8 bins

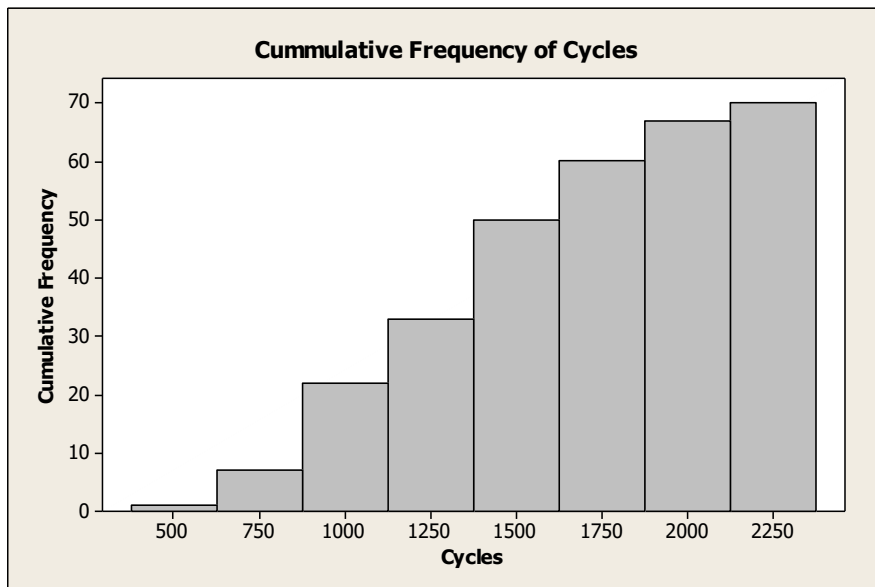
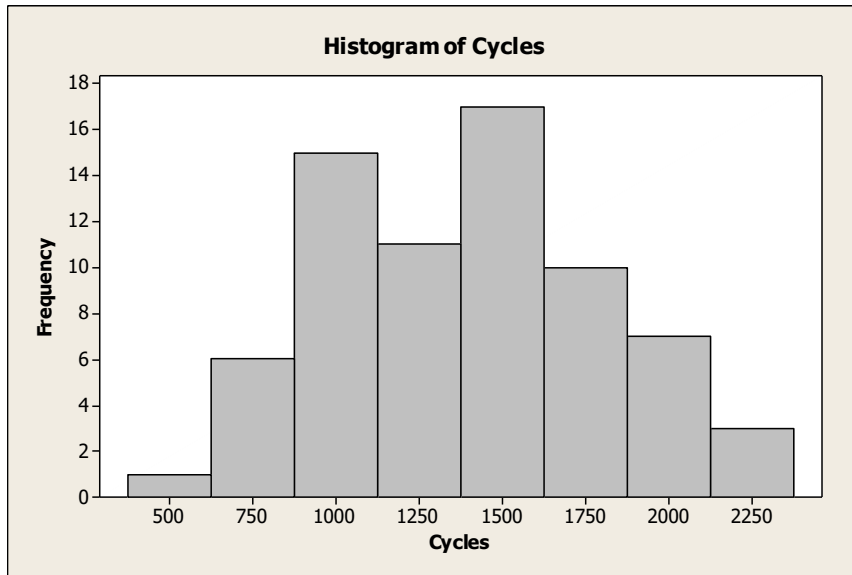


b) 16 bins

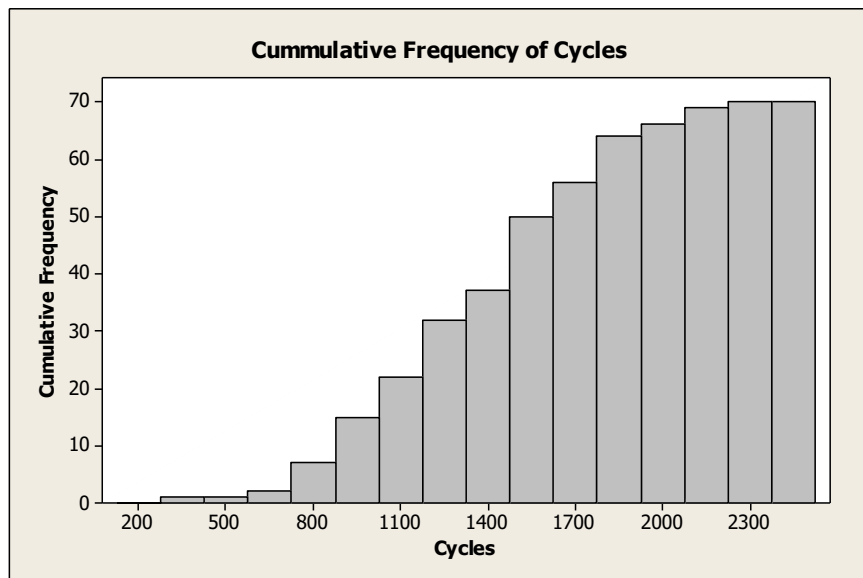
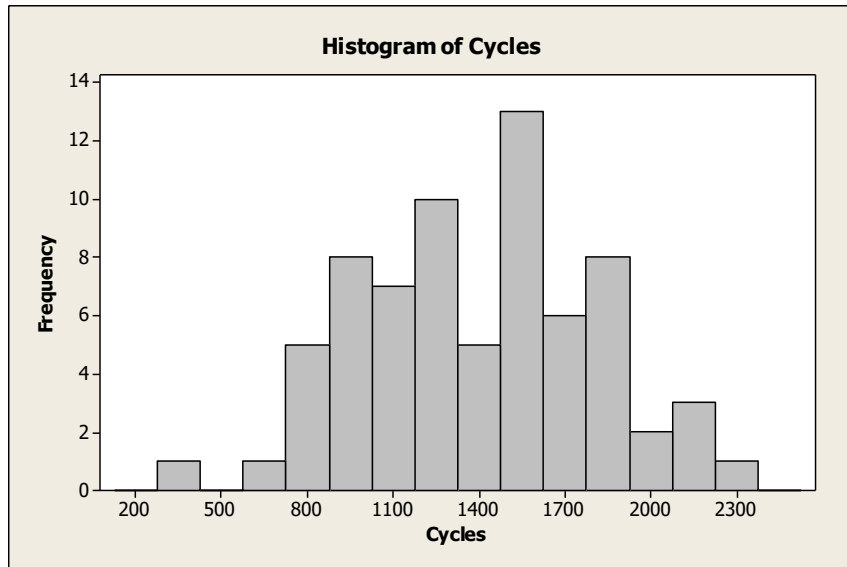


Yes, both histograms display similar information based on this dataset.

2-21. a) 8 bins

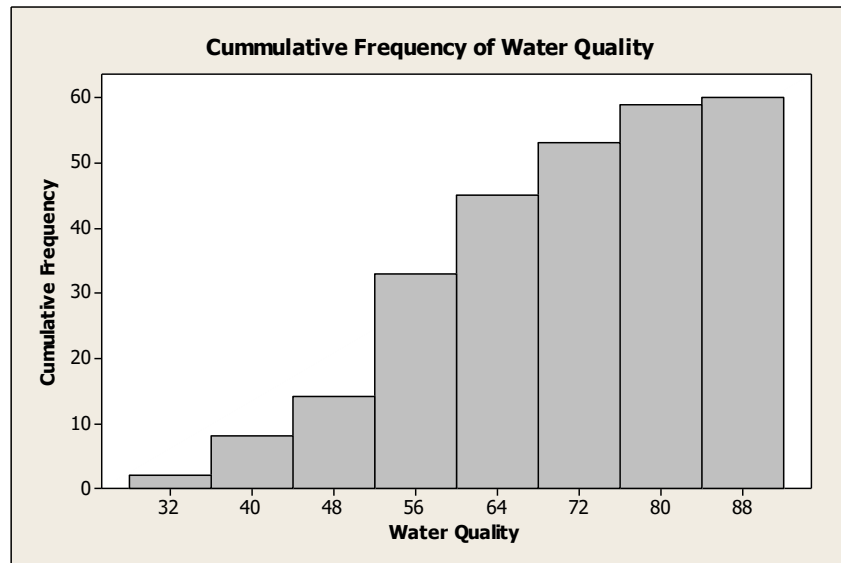
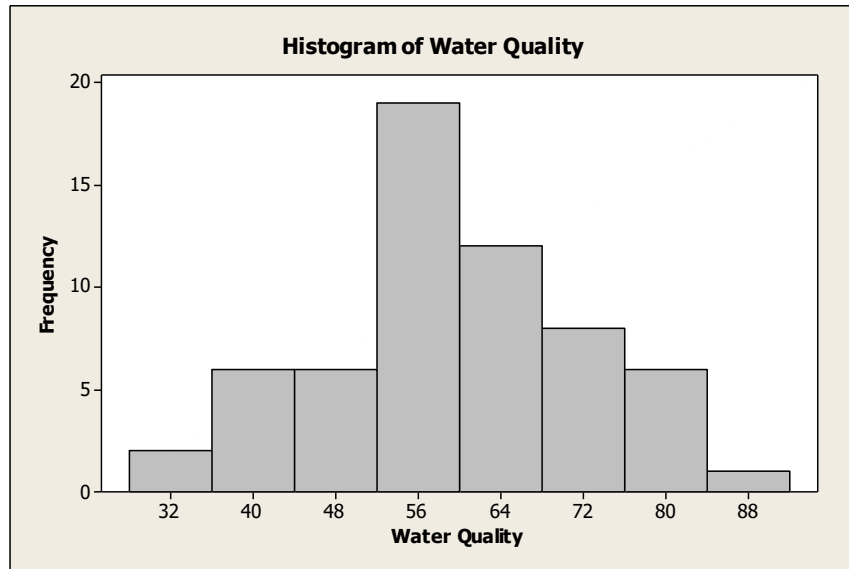


b) 16 bins

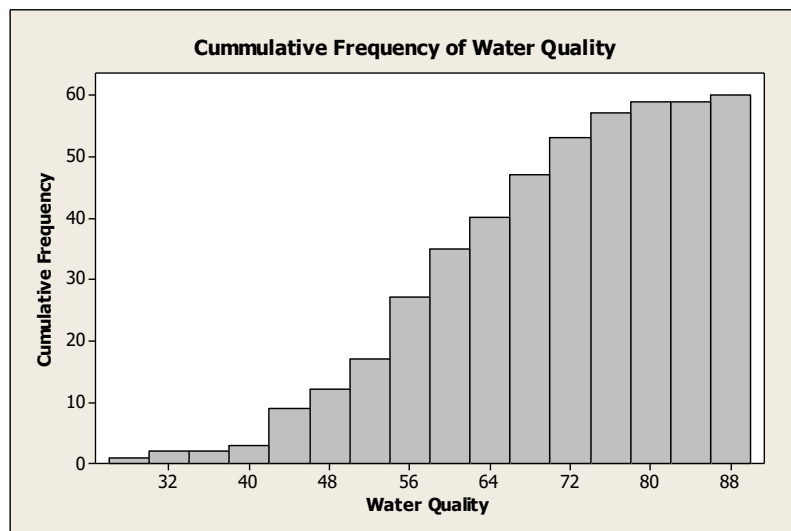
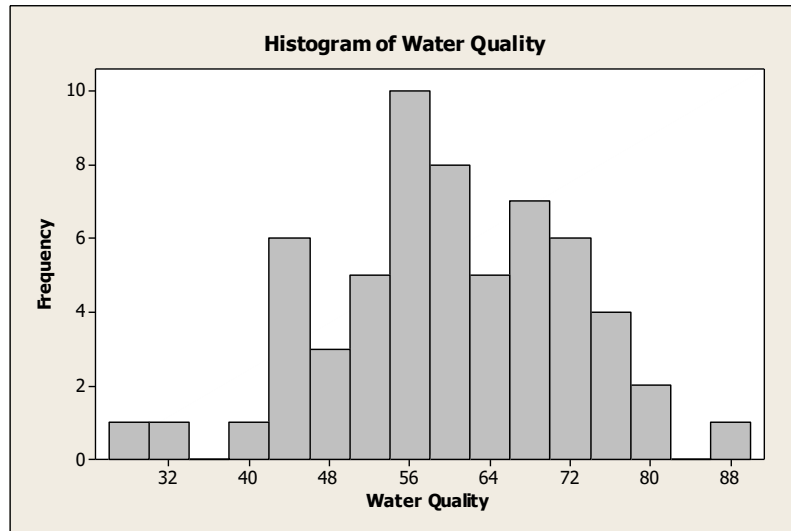


Yes, both histograms display similar information based on this dataset.

2-22. a) 8 bins

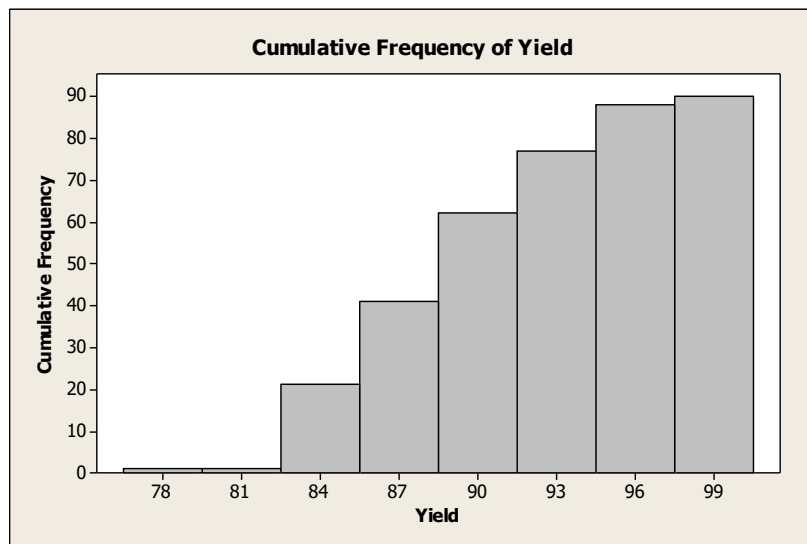
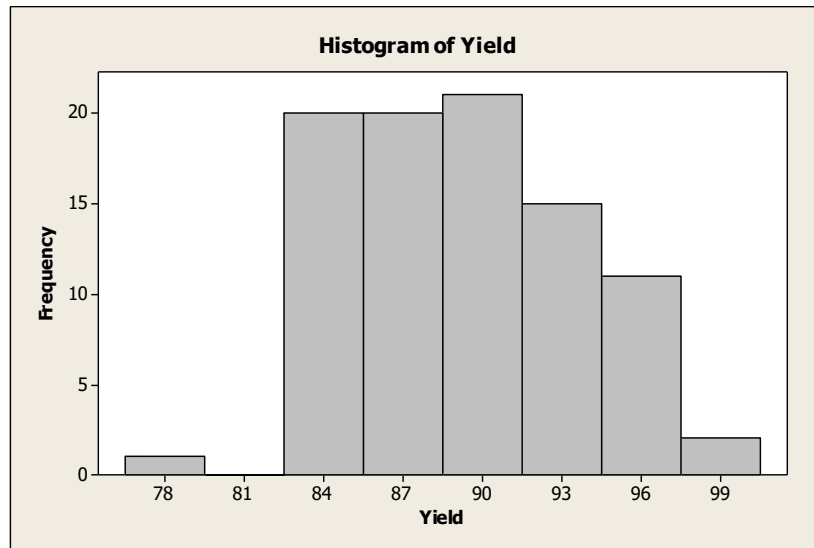


b) 16 bins

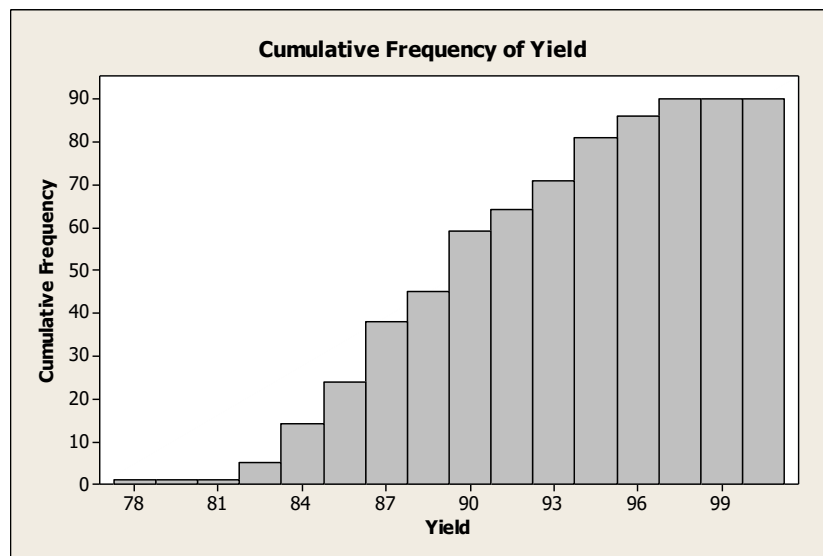
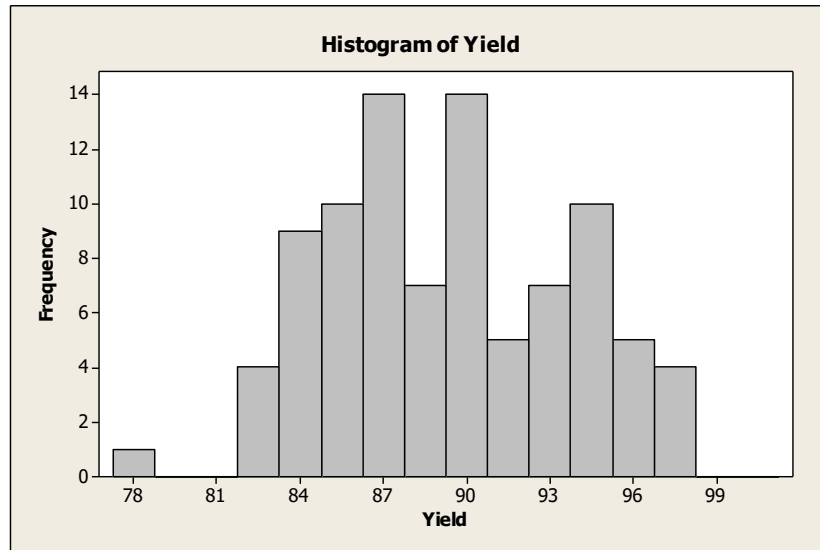


Yes, both histograms display similar information based on this dataset.

2-23. a) 8 bins

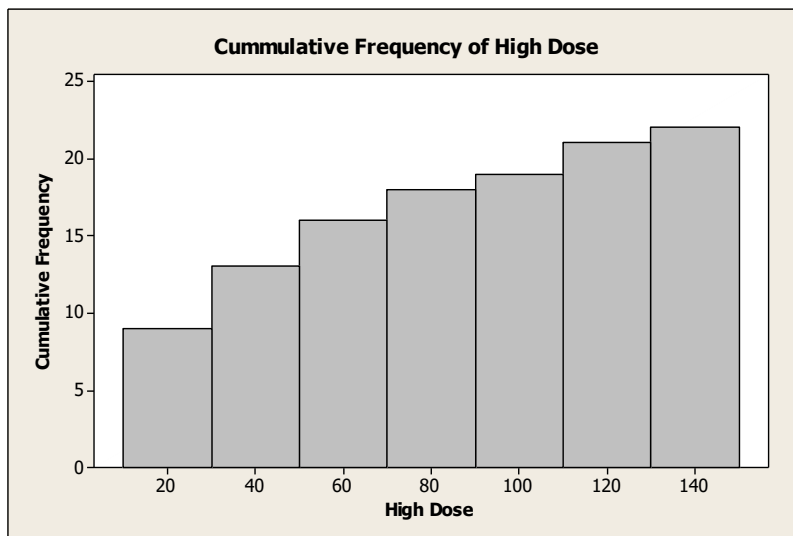
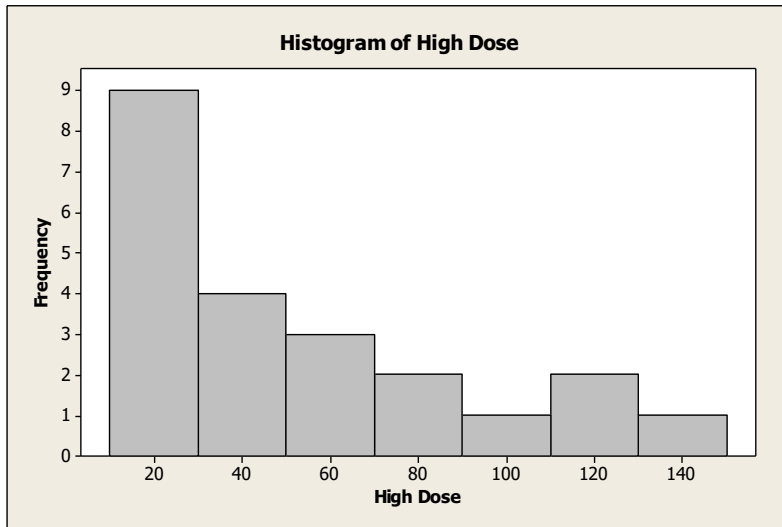


b) 16 bins

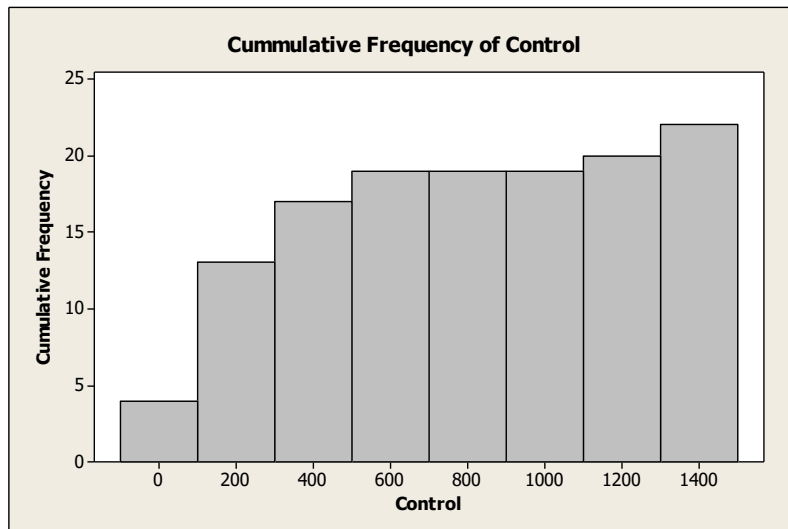
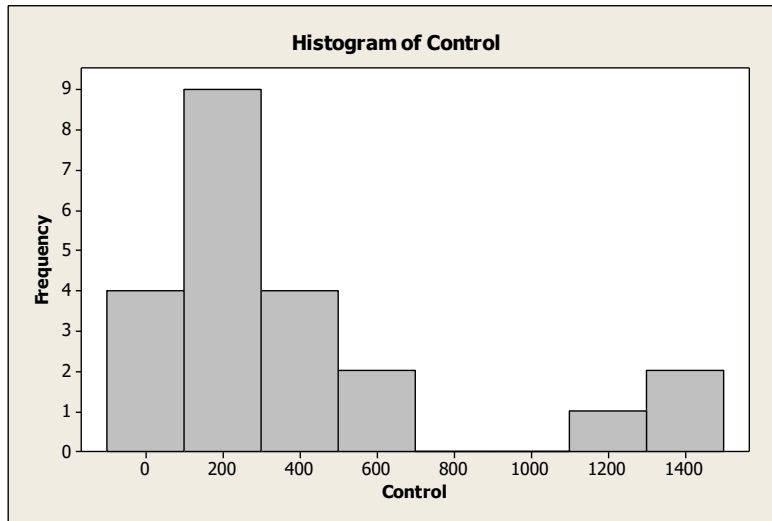


Yes, both histograms display similar information based on this dataset.

2-24. High Dose

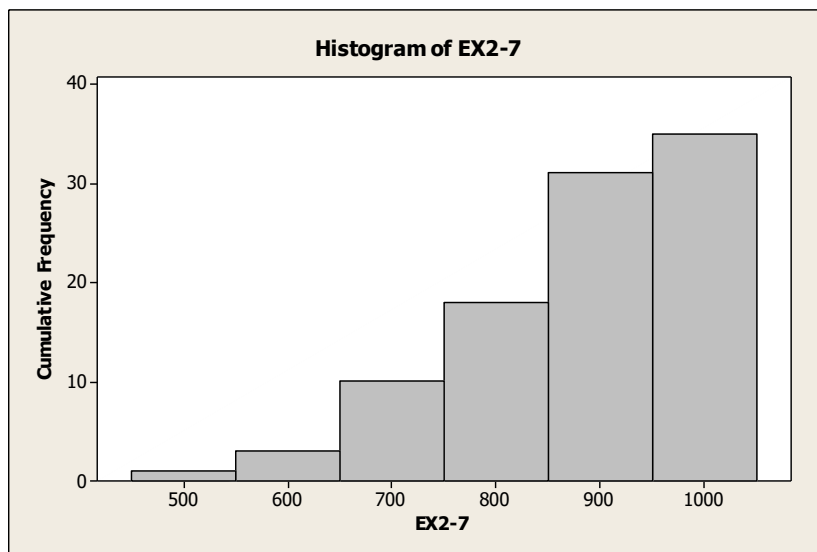
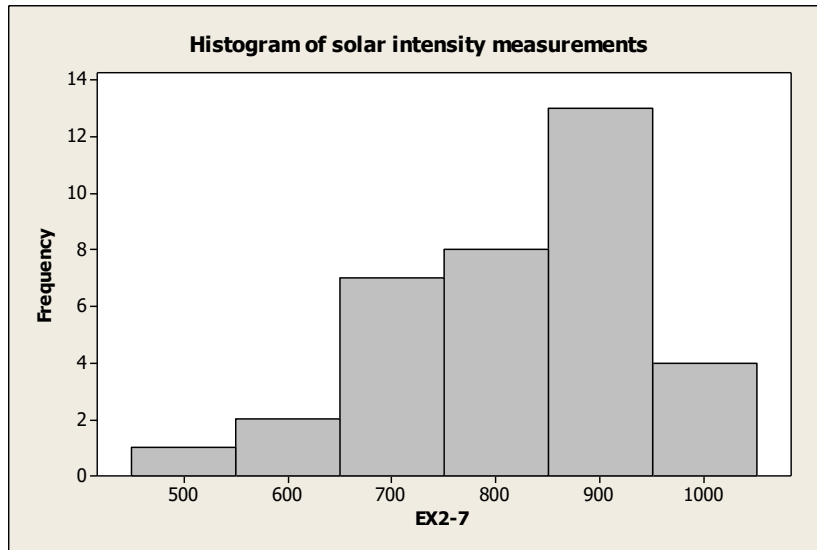


Control group



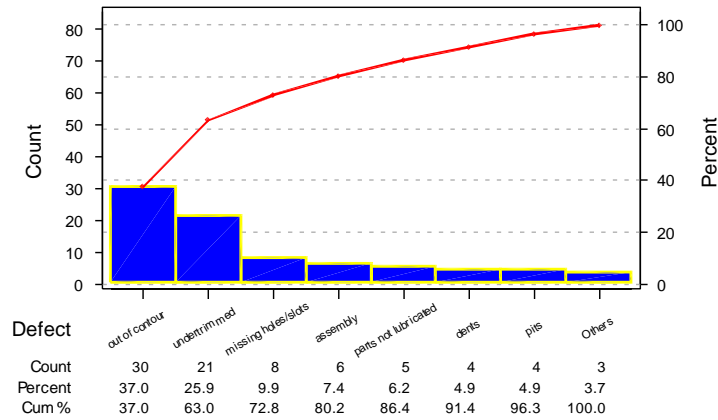
They both look similar.

2-25. 6 bins



2-26.

Pareto Chart for Defect



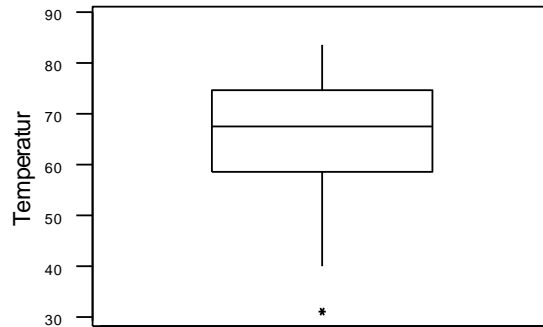
Roughly 63% of defects are described by parts out of contour and parts under trimmed.

Section 2-4

- 2-27. a) Sample Mean: 65.86, Sample Standard Deviation: 12.16
 b) Q₁: 58.5, Q₃: 75
 c) Median: 67.5
 d) Sample Mean: 66.86, Sample Standard Deviation: 10.74, Q₁: 60, Q₃: 75, Median: 68

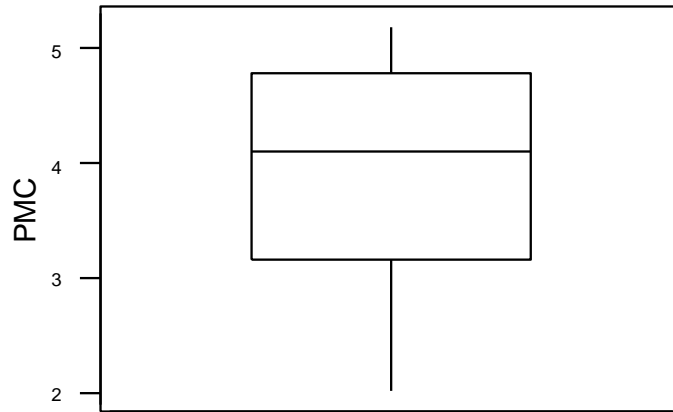
The mean has increased while the sample standard deviation has decreased. The lower quartile has increased while the upper quartile has remained unchanged. The median has increased slightly due to the removal of the data point. The smallest value appears quite different than the other temperature values.

- e) Using the entire data set, the box plot is

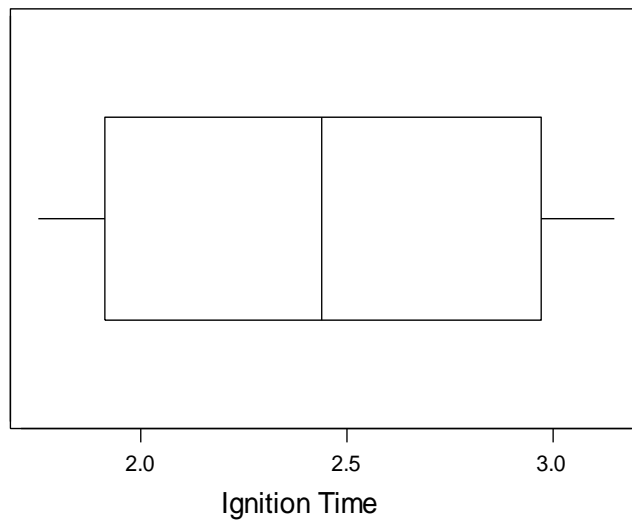


The value of 31 appears to be one possible outlier.

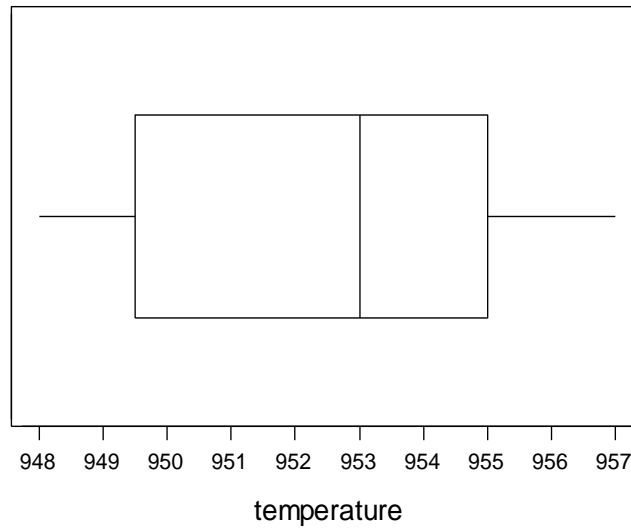
- 2-28. a) Sample Mean: 4
b) Sample Variance: 0.867, Sample Standard Deviation: 0.931
c)



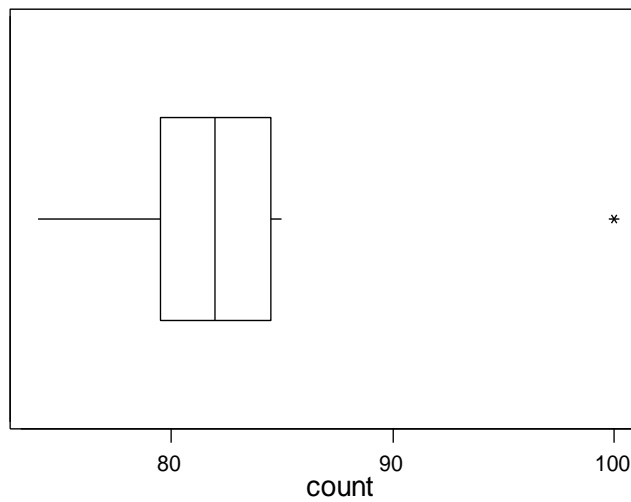
- 2-29. a) Sample mean = 2.415, Sample standard deviation = 0.534
b)



- 2-30. a) Sample mean = 952.44, Sample standard deviation = 3.09
 b) Median = 953. The largest temperature could take on any value as long as it is the larger than the current largest value.
 c)

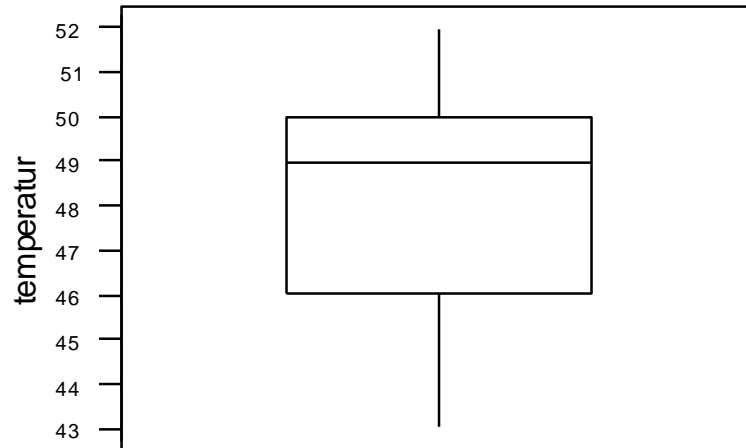


- 2-31. a) Sample mean: 83.11, sample variance = 50.55, sample standard deviation = 7.11
 b) $Q_1 = 79.5$, $Q_3 = 84.50$
 c)



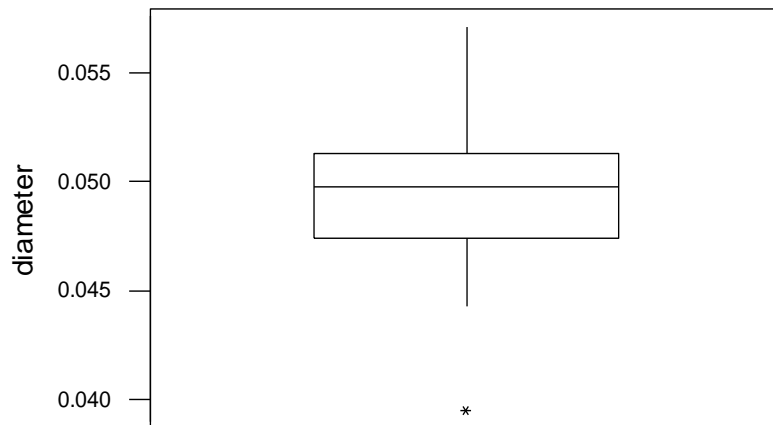
- d) Sample mean = 81, sample standard deviation = 3.46, $Q_1 = 79.25$, $Q_3 = 83.75$. The sample mean and the sample standard deviation have decreased. The lower quartile has decreased slightly while the upper quartile has decreased.

- 2-32. a) Sample Mean: 48.125, Sample Median: 49
 b) Sample Variance: 7.247, Sample Standard Deviation: 2.692
 c) The data appear to be skewed.



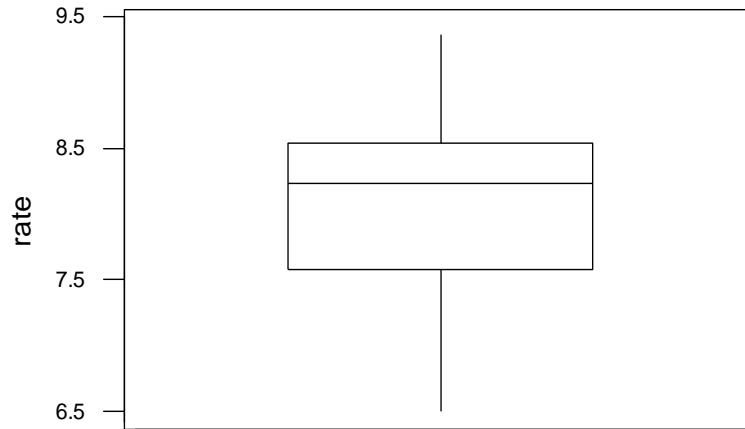
- d) 5th Percentile: 43.25, 95th Percentile: 52

- 2-33. a) Sample Mean: 0.04939, Sample Variance: 0.00001568
 b) Q_1 : 0.04738, Q_3 : 0.0513
 c) Sample Median: 0.04975
 d)



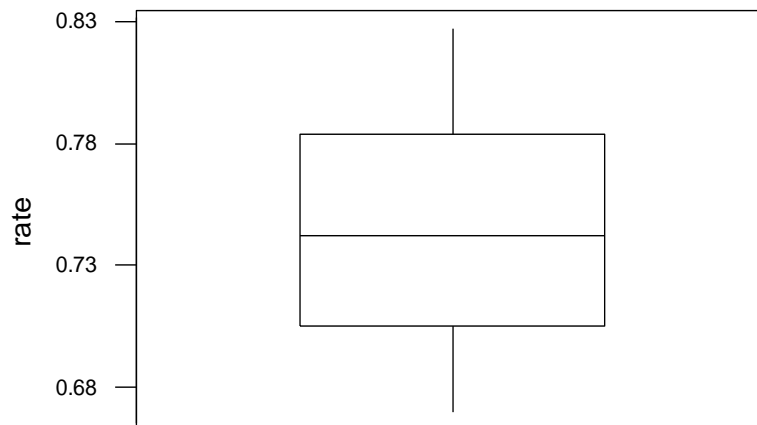
- e) 5th Percentile: 0.03974, 95th Percentile: 0.057

- 2-34. a) Sample Mean: 8.059, Sample Variance: 0.661
b) Q_1 : 7.575, Q_3 : 8.535
c) Sample Median: 8.235
d)



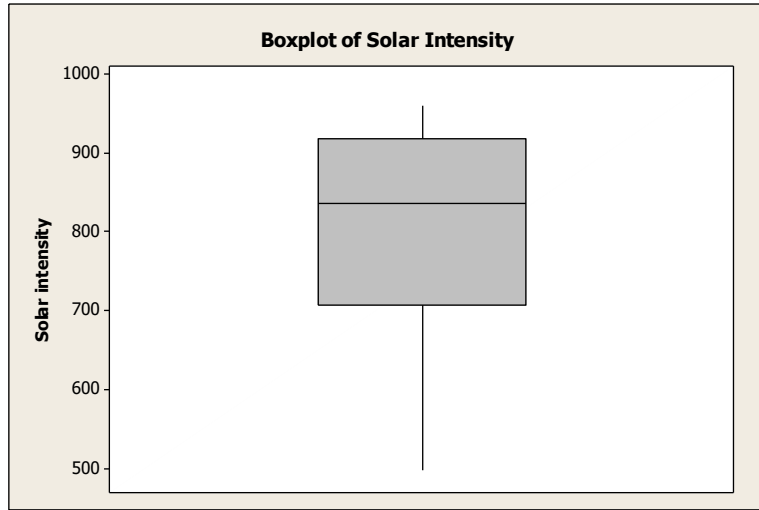
- e) 5th Percentile: 6.175, 95th Percentile: 9.3315

- 2-35. a) Sample Mean: 0.7481, Sample Variance: 0.00226
b) Q_1 : 0.7050, Q_3 : 0.7838
c) Sample Median: 0.742
d)



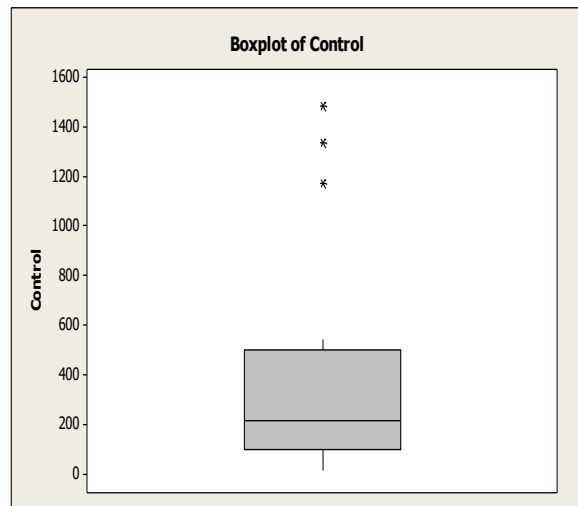
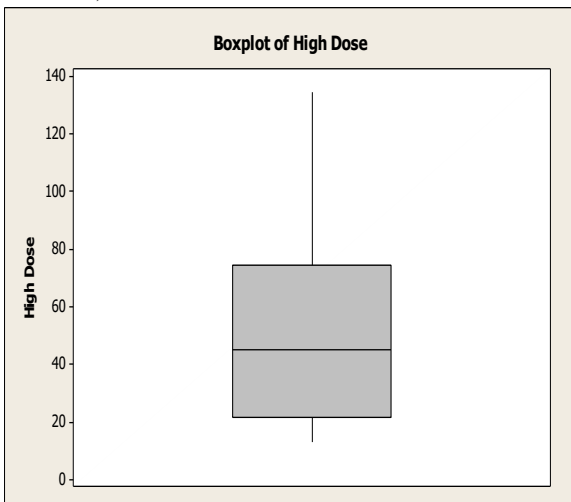
- e) 5th Percentile: 0.5025, 95th Percentile: 0.821

- 2-36. a) Sample Mean: 810.5, Sample Variance: 16465.61
 b) Q_1 : 708, Q_3 : 918
 c) Sample Median: 835
 d)



- e) 5th Percentile: 546, 95th Percentile: 957.6

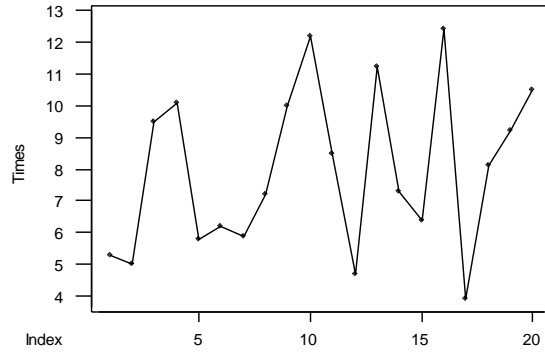
- 2-37. a) High Dose: Sample Mean: 52.65, Sample Variance: 1490.32
 Control: Sample Mean: 382.7, Sample Variance: 175224.35
 b) High Dose: Q_1 : 21.70, Q_3 : 74.38
 Control: Q_1 : 101.9, Q_3 : 501.1
 c) High Dose: Sample Median: 45
 Control: Sample Median: 215.4
 d)



- e) High Dose: 5th Percentile: 13.125, 95th Percentile: 133.67
 Control: 5th Percentile: 17.045, 95th Percentile: 1460.23
 All summary statistics are larger for the control group.

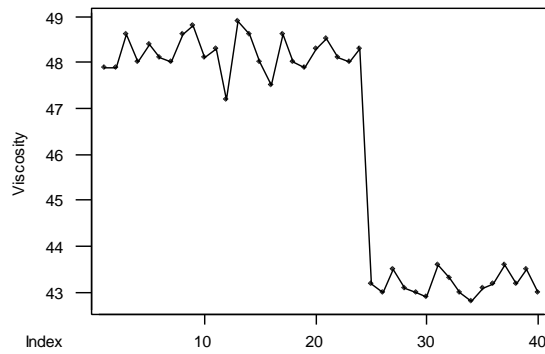
Section 2-5

2-38.



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

2-39. a)



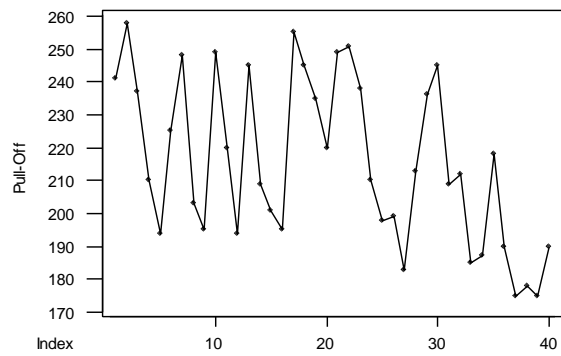
Stem-and-leaf display for Problem 2-32. Viscosity: unit = 0.1 1|2 represents 1.2

```

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

b) The plots indicates that the process is not stable and not capable of meeting the specifications.

2-40. a)

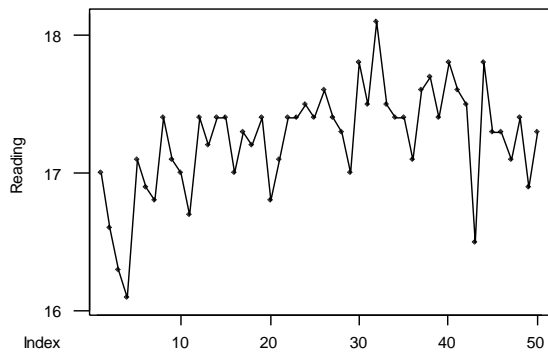


b) Stem-and-leaf display for Problem 2-33. Force: unit = 1 1|2 represents 12

```
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.

2-41.



Stem-and-leaf display for Concentration: unit = 0.01 1|2 represents 0.12
 LO|1610,1630

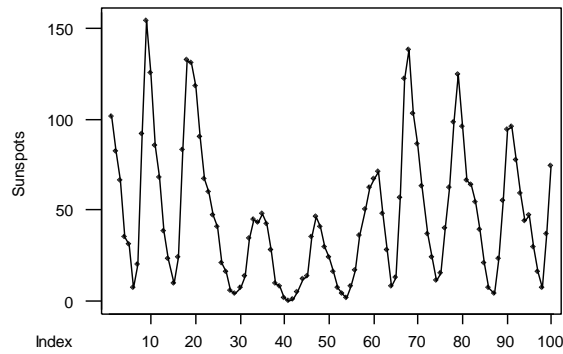
```

3 165|0
4 166|0
5 167|0
7 168|00
9 169|00
13 170|0000
18 171|00000
20 172|00
25 173|00000
25 174|0000000000000000
12 175|0000
8 176|000
5 177|0
4 178|000

```

HI|1810
 The data appear skewed.

2-42. a)



b) Stem-and-leaf display for Problem 2-35. Sunspots: unit = 1 1|2 represents 12

```

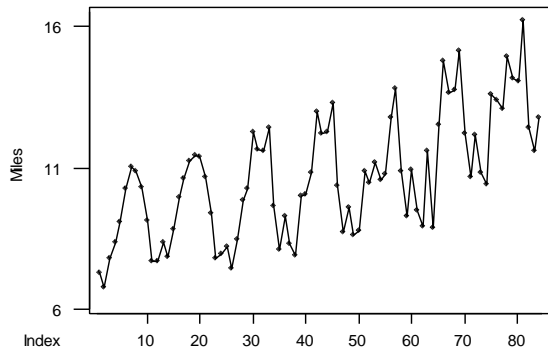
17 0|0122444567777888
29 1|001234456667
39 2|0113344488
50 3|00145567789
50 4|011234567788
38 5|04579
33 6|0223466778
23 7|147
20 8|2356
16 9|024668
10 10|13
8 11|8
7 12|245
4 13|128

```

HI |154

The data appears to decrease between 1790 and 1835, and after 1839 the stem and leaf plot indicates skewed data.

2-43. a)



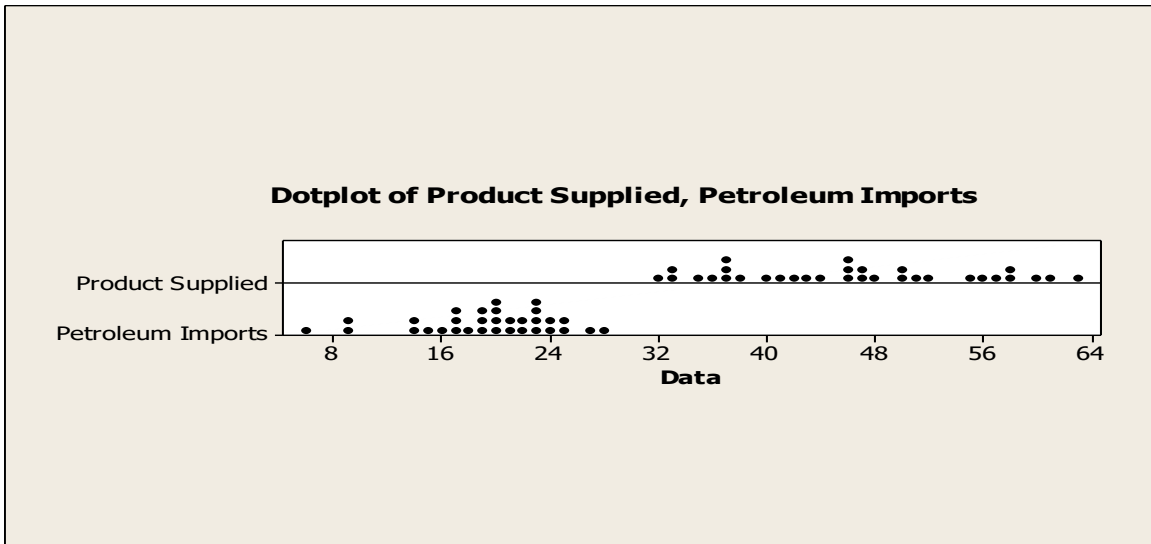
b) Stem-and-leaf display for Problem 2-36. Miles: unit = 0.1 1|2 represents 1.2

```

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
    
```

There is an increasing trend in the data.

2-44. Digidot plot



Stem-and-leaf of Product Supplied N = 32
 Leaf Unit = 1.0

```

4  3  2334
9  3  66778
14 4  00124
(7) 4  5566779
11 5  0014
7  5  5688
3  6  013
  
```

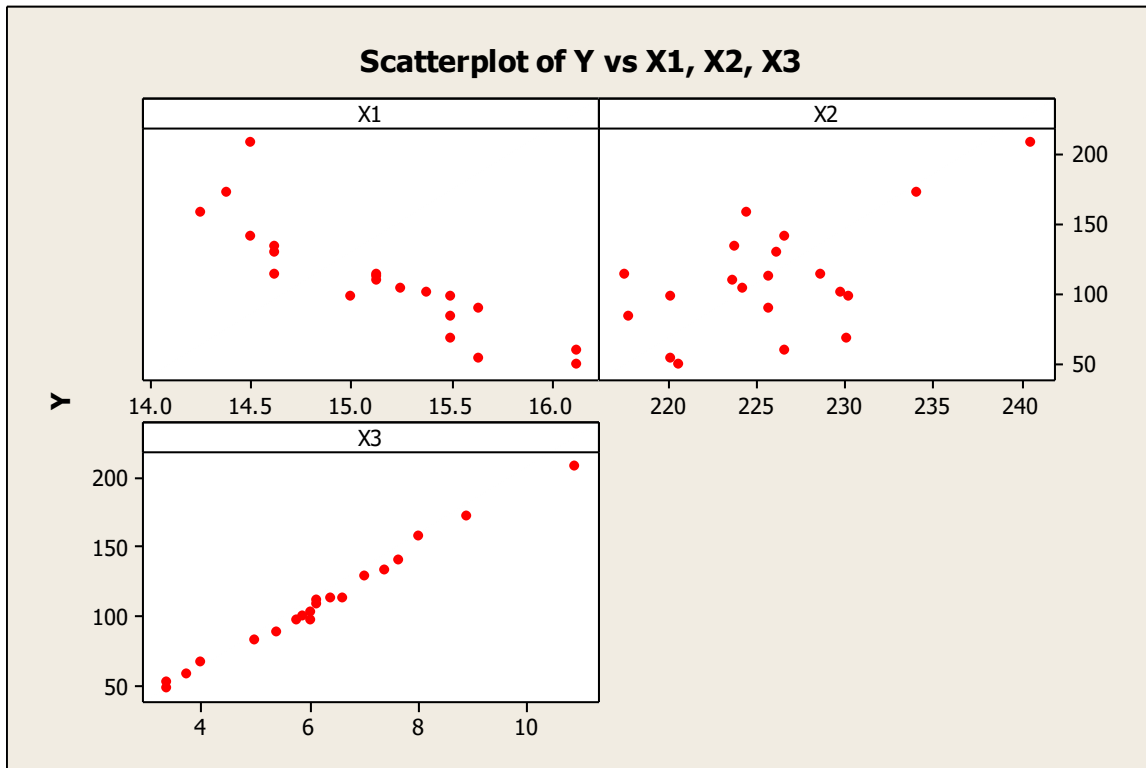
Stem-and-leaf of Petroleum Imports N = 32
 Leaf Unit = 1.0

```

1  0  6
3  0  89
3  1
5  1  33
6  1  4
11 1  66777
16 1  89999
16 2  000011
10 2  2233
6  2  4445
2  2  67
  
```

Section 2-6

2-45. a) X1 has negative correlation with Y, X2 and X3 have positive correlation with Y.



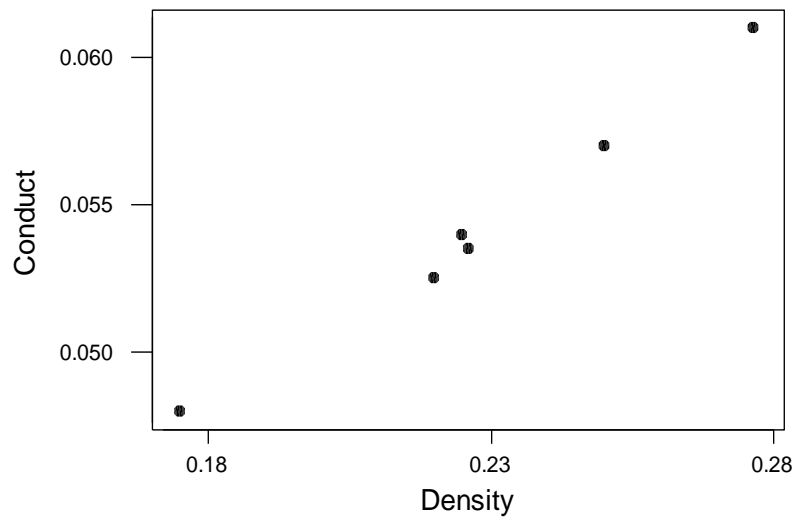
$$b) r_{x_1} = \frac{S_{x_1y}}{\sqrt{S_{x_1x_1}S_{yy}}} = \frac{-375.185}{\sqrt{5.875 \times 30725.23}} = -0.883$$

$$r_{x_2} = \frac{S_{x_2y}}{\sqrt{S_{x_2x_2}S_{yy}}} = \frac{2486.19}{\sqrt{591.778 \times 30725.23}} = 0.585$$

$$r_{x_3} = \frac{S_{x_3y}}{\sqrt{S_{x_3x_3}S_{yy}}} = \frac{1415.5}{\sqrt{65.778 \times 30725.23}} = 0.995$$

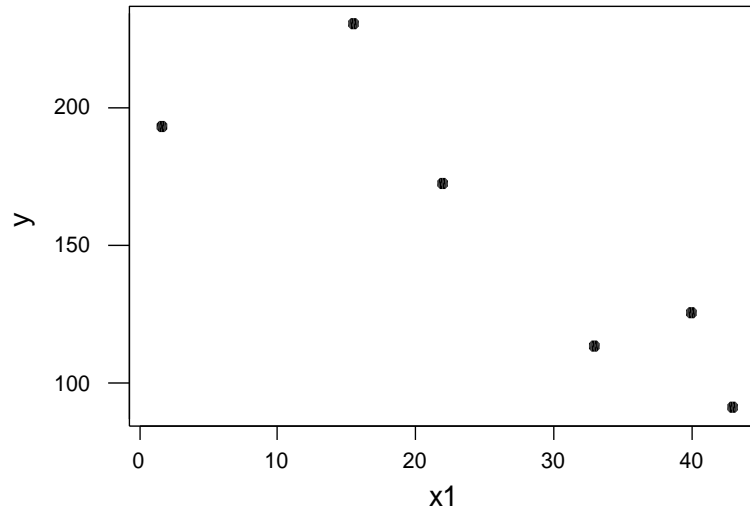
X1 has a strong negative correlation with Y, X3 has a strong positive correlation with Y and X2 has a moderate positive correlation with Y. The correlation coefficients agree with the scatter plot in part (a).

2-46. a) Positive sign

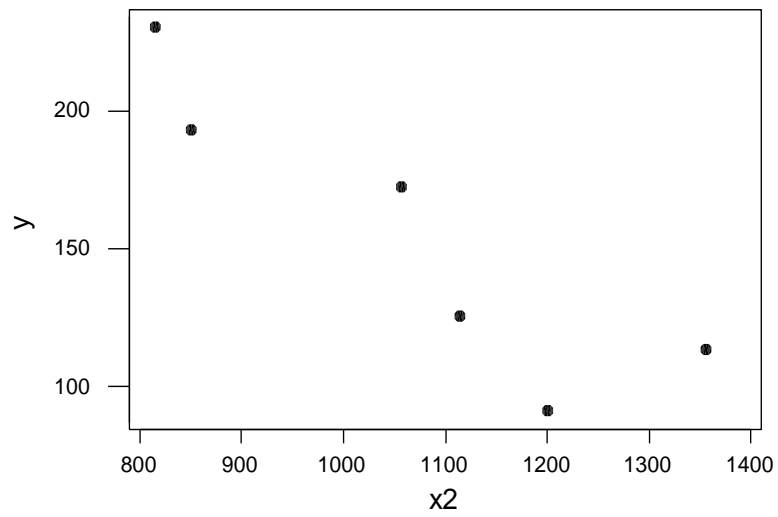


b) 0.993. X has a strong positive correlation with Y

- 2-47. a) Both sample correlations will be negative.
y versus x_1

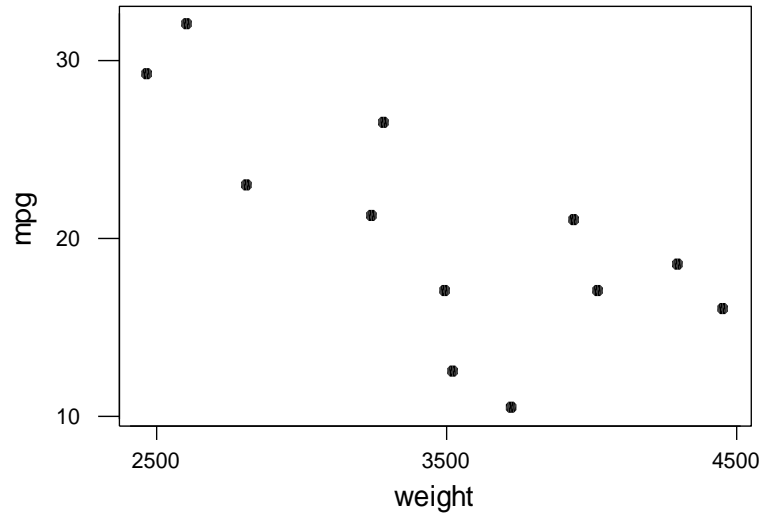


y versus x_2

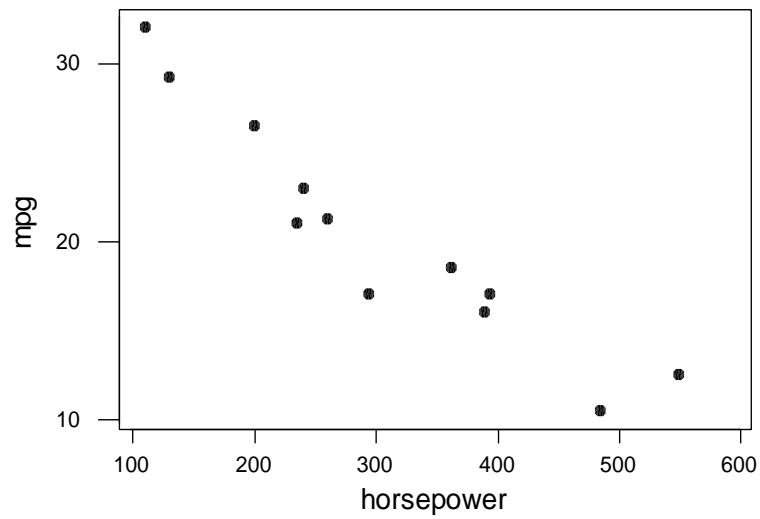


- b) y versus x_1 : -0.852; y versus x_2 : -0.898. X_1 and X_2 have a moderately strong negative correlation with Y

- 2-48. a) Both sample correlations will be negative.
MPG versus weight

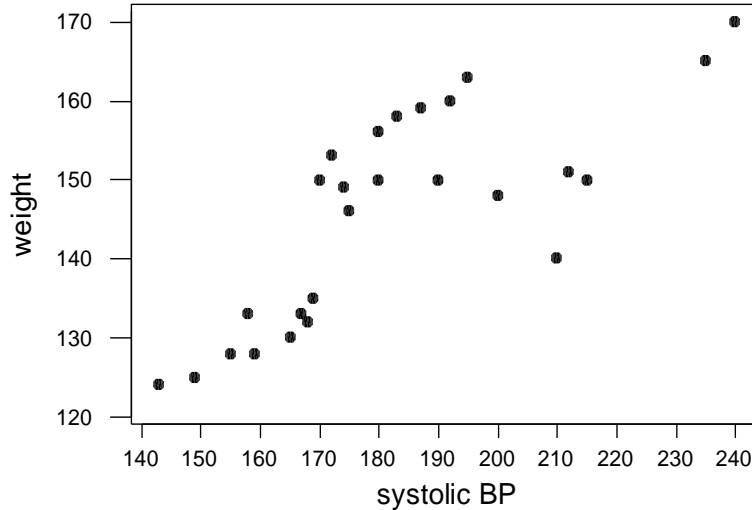


MPG versus horsepower



- b) MPG versus weight: -0.709; MPG versus horsepower: -0.947. Weight has a strong negative correlation with MPG while weight has a moderate negative correlation with MPG

- 2-49. a) The correlation coefficient will be positive



- b) 0.773. Weight has a moderate positive correlation with systolic BP.

Supplemental Exercises

- 2-50. a) Sample Mean = 7.1838; The sample mean value is close enough to the target value to accept the solution as conforming. There is slight difference due to inherent variability.
 b) $s^2 = 0.000427$, $s = 0.02066$; A major source of variability would include measurement to measurement error. A low variance is desirable since it may indicate consistency from measurement to measurement.

2-51. a) $\sum_{i=1}^6 x_i^2 = 10,433$ $\left(\sum_{i=1}^6 x_i\right)^2 = 62,001$ $n = 6$

$$s^2 = \frac{\sum_{i=1}^6 x_i^2 - \frac{\left(\sum_{i=1}^6 x_i\right)^2}{n}}{n-1} = \frac{10,433 - \frac{62,001}{6}}{6-1} = 19.9\Omega^2 \quad s = \sqrt{19.9\Omega^2} = 4.46\Omega$$

b) $\bar{x} = \frac{246}{6} = 41.5$ $n = 6$

$$s^2 = \frac{\sum_{i=1}^6 (x_i - \bar{x})^2}{n-1} = \frac{99.5}{5} = 19.9\Omega^2; \quad s = \sqrt{19.9\Omega^2} = 4.46\Omega$$

- c) $s^2 = 19.9\Omega^2$ $s = 4.46\Omega$; Shifting the data from the sample by a constant amount has no effect on the sample variance or standard deviation.
 d) 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).

- 2-52. a) Sample Range = 3.2, $s^2 = 0.866$, $s = 0.931$
 b) Sample Range = 3.2, $s^2 = 0.866$, $s = 0.931$;
 These are the same as in part a). Any constant would produce the same results.

$$2-53. \quad \text{a) } \bar{x}_{n+1} = \frac{\sum_{i=1}^{n+1} x_i}{n+1} = \frac{\sum_{i=1}^n x_i + x_{n+1}}{n+1}; \quad \bar{x}_{n+1} = \frac{n\bar{x}_n + x_{n+1}}{n+1}; \quad \bar{x}_{n+1} = \frac{n}{n+1} \bar{x}_n + \frac{x_{n+1}}{n+1}$$

$$\begin{aligned} \text{b) } ns_{n+1}^2 &= \sum_{i=1}^n x_i^2 + x_{n+1}^2 - \frac{\left(\sum_{i=1}^n x_i + x_{n+1}\right)^2}{n+1} \\ &= \sum_{i=1}^n x_i^2 + x_{n+1}^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n+1} - \frac{2x_{n+1} \sum_{i=1}^n x_i}{n+1} - \frac{x_{n+1}^2}{n+1} \\ &= \sum_{i=1}^n x_i^2 + \frac{n}{n+1} x_{n+1}^2 - \frac{n}{n+1} 2x_{n+1} \bar{x}_n - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n+1} \\ &= \left[\sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n+1} \right] + \frac{n}{n+1} [x_{n+1}^2 - 2x_{n+1} \bar{x}_n] \\ &= \sum_{i=1}^n x_i^2 + \left[-\frac{\left(\sum_{i=1}^n x_i\right)^2}{n} + \frac{\left(\sum_{i=1}^n x_i\right)^2}{n} \right] - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n+1} + \frac{n}{n+1} [x_{n+1}^2 - 2x_{n+1} \bar{x}_n] \\ &= \sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n} + \frac{(n+1)\left(\sum_{i=1}^n x_i\right)^2 - n\left(\sum_{i=1}^n x_i\right)^2}{n(n+1)} + \frac{n}{n+1} [x_{n+1}^2 - 2x_{n+1} \bar{x}_n] \\ &= (n-1)s_n^2 + \frac{\left(\sum_{i=1}^n x_i\right)^2}{n(n+1)} + \frac{n}{n+1} [x_{n+1}^2 - 2x_{n+1} \bar{x}_n] \\ &= (n-1)s_n^2 + \frac{n\bar{x}_n^2}{n+1} + \frac{n}{n+1} [x_{n+1}^2 - 2x_{n+1} \bar{x}_n] \\ &= (n-1)s_n^2 + \frac{n}{n+1} (x_{n+1} - 2x_{n+1} \bar{x}_n + \bar{x}_n^2) \\ &= (n-1)s_n^2 + \frac{n}{n+1} (x_{n+1} - \bar{x}_n)^2 \end{aligned}$$

$$\text{c) } \bar{x}_n = 41.5 \quad x_{n+1} = 46 \quad sn^2 = 19.9 \quad n = 6$$

$$\begin{aligned} \bar{x}_{n+1} &= \frac{6(41.5) + 46}{6+1} \\ &= 42.14 \end{aligned}$$

$$\begin{aligned} s_{n+1} &= \frac{\sqrt{(6-1)19.9 + \frac{6}{6+1} 46 - 41.5}}{6} \\ &= 4.41 \end{aligned}$$

2-54. The trimmed mean is pulled toward the median by eliminating outliers.

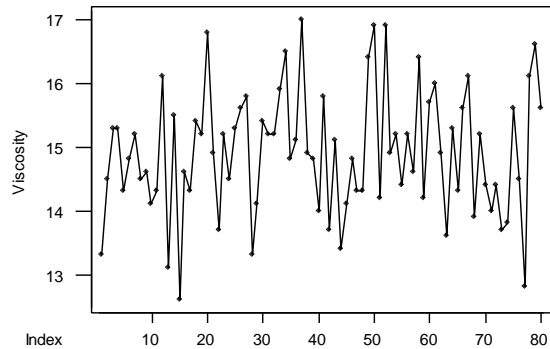
a) 10% Trimmed Mean = 89.29

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.

d) 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.

- 2-55. a) Sample 1: 4; Sample 2: 4 Yes, the two appear to exhibit the same variability.
 b) Sample 1: 1.604, Sample 2: 1.852 No, sample 2 has a larger standard deviation.
 c) The sample range is a crude estimate 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.
- 2-56. a)



- The data appears to vary between 12.5 and 17, with no obvious pattern.
- b) The plot indicates that the two processes generate similar results. This is evident since the data appear to be centered around the same mean.
- c) 1st 40 observations: Sample Mean = 14.87, Sample Variance = 0.899
 2nd 40 observations: Sample Mean = 14.92, Sample Variance = 1.05
 The quantities indicate the processes do yield the same mean level. The variability also appears to be about the same, with the sample variance for the 2nd 40 observations being slightly larger than that for the 1st 40.

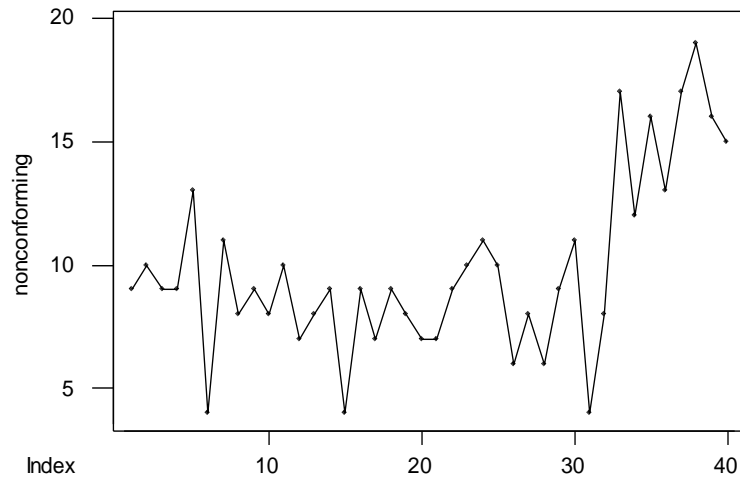
- 2-57. a) Stem-and-leaf of nonconforming; N = 40, Leaf Unit = 0.10

```

3 4 000
3 5
5 6 00
9 7 0000
15 8 000000
(9) 9 000000000
16 10 0000
12 11 000
9 12 0
8 13 00
6 14
6 15 0
5 16 00
3 17 00
1 18
1 19 0
  
```

- b) Sample Mean: 9.8; Sample Standard deviation: 3.611

c) There appears to be an increase in the average number of nonconforming springs made during the 40 days.



2-58. a) Stem-and-leaf of errors $N = 20$, Leaf Unit = 0.10

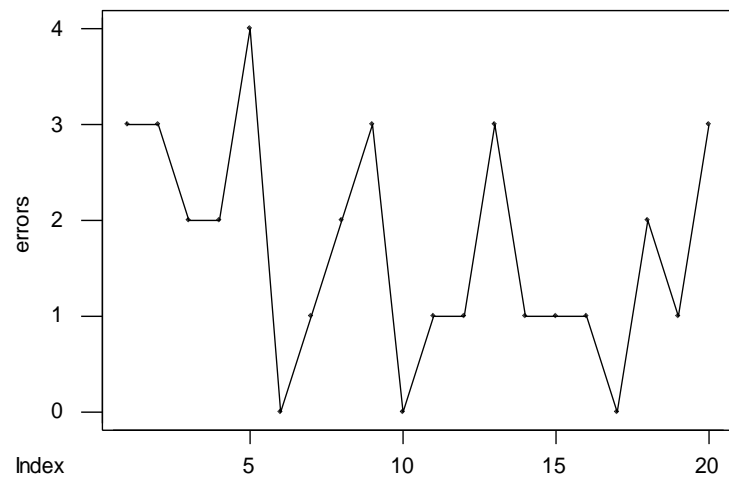
```

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

```

b) Sample average: 1.7, Sample Standard deviation: 1.174

c)



The time series plot indicates a slight decrease in the number of errors.