Chapter 1

Section 1.1

1. (a) The population consists of all the times the process could be run. It is conceptual.

   (b) The population consist of all the registered voters in the state. It is tangible.

   (c) The population consist of all people with high cholesterol levels. It is tangible.

   (d) The population consist of all concrete specimens that could be made from the new formulation. It is conceptual.

   (e) The population consist of all bolts manufactured that day. It is tangible.

2. (iii). It is very unlikely that students whose names happen to fall at the top of a page in the phone book will differ systematically in height from the population of students as a whole. It is somewhat more likely that engineering majors will differ, and very likely that students involved with basketball intramurals will differ.

3. (a) False

   (b) True

4. (a) False

   (b) True

5. (a) No. What is important is the population proportion of defectives; the sample proportion is only an approximation. The population proportion for the new process may in fact be greater or less than that of the old process.

   (b) No. The population proportion for the new process may be 0.12 or more, even though the sample proportion was only 0.11.

   (c) Finding 2 defective circuits in the sample.

6. (a) False

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

(b) True

(c) True

7. A good knowledge of the process that generated the data.

8. (a) An observational study

   (b) It is not well-justified. Because the study is observational, there could be differences between the groups other than the level of exercise. These other differences (confounders) could cause the difference in blood pressure.

9. (a) A controlled experiment

   (b) It is well-justified, because it is based on a controlled experiment rather than an observational study.

Section 1.2

1. False

2. No. In the sample 1, 2, 4, the mean is 7/3, which does not appear at all.

3. No. In the sample 1, 2, 4, the mean is 7/3, which does not appear at all.

4. No. The median of the sample 1, 2, 4, 5 is 3.

5. The sample size can be any odd number.

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

6. Yes. For example, the list 1, 2, 12 has an average of 5 and a standard deviation of 6.08.

7. Yes. If all the numbers in the list are the same, the standard deviation will equal 0.

8. The mean increases by $50; the standard deviation is unchanged.

9. The mean and standard deviation both increase by 5%.

10. (a) Let $X_1, ..., X_{100}$ denote the 100 numbers of occupants.

$$\sum_{i=1}^{100} X_i = 70(1) + 15(2) + 10(3) + 3(4) + 2(5) = 152$$

$$\overline{X} = \frac{\sum_{i=1}^{100} X_i}{100} = \frac{152}{100} = 1.52$$

(b) The sample variance is

$$\begin{aligned} s^2 &= \frac{1}{99}\left(\sum_{i=1}^{100} X_i^2 - 100\overline{X}^2\right) \\ &= \frac{1}{99}[(70)1^2 + (15)2^2 + (10)3^2 + (3)4^2 + (2)5^2 - 100(1.52^2)] \\ &= 0.87838 \end{aligned}$$

The standard deviation is $s = \sqrt{s^2} = 0.9372$.
Alternatively, the sample variance can be computed as

$$\begin{aligned} s^2 &= \frac{1}{99}\sum_{i=1}^{100}(X_i - \overline{X})^2 \\ &= \frac{1}{99}[70(1-1.52)^2 + 15(2-1.52)^2 + 10(3-1.52)^2 + 3(4-1.52)^2 + 2(5-1.52)^2] \\ &= 0.87838 \end{aligned}$$

(c) The sample median is the average of the 50th and 51st value when arranged in order. Both these values are equal to 1, so the median is 1.

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

(d) The first quartile is the average of the 25th and 26th value when arranged in order. Both these values are equal to 1, so the first quartile is 1. The third quartile is the average of the 75th and 76th value when arranged in order. Both these values are equal to 2, so the first quartile is 2.

(e) Of the 100 cars, 15 + 10 + 3 + 2 = 30 had more than the mean of 1.52 occupants, so the proportion is 30/100 = 0.3.

(f) The quantity that is one standard deviation greater than the mean is 1.52 + 0.9372 = 2.5472. Of the 100 cars, 10 + 3 + 2 = 15 had more than 2.8652 children, so the proportion is 15/100 = 0.15.

(g) The region within one standard deviation of the mean is $1.52 \pm 0.9372 = (0.5828, 2.4572)$. Of the 100 cars, 70 + 15 = 85 are in this range, so the proportion is 85/100 = 0.85.

11. The total height of the 20 men is $20 \times 178 = 3560$. The total height of the 30 women is $30 \times 164 = 4920$. The total height of all 50 people is $3560 + 4920 = 8480$. There are $20 + 30 = 50$ people in total. Therefore the mean height for both groups put together is $8480/50 = 169.6$ cm.

12. (a) The mean for A is

$$(18.0+18.0+18.0+20.0+22.0+22.0+22.5+23.0+24.0+24.0+25.0+25.0+25.0+25.0+26.0+26.4)/16 = 22.744$$

The mean for B is

$$(18.8+18.9+18.9+19.6+20.1+20.4+20.4+20.4+20.4+20.5+21.2+22.0+22.0+22.0+22.0+23.6)/16 = 20.700$$

The mean for C is

$$(20.2+20.5+20.5+20.7+20.8+20.9+21.0+21.0+21.0+21.0+21.0+21.5+21.5+21.5+21.5+21.6)/16 = 20.013$$

The mean for D is

$$(20.0+20.0+20.0+20.0+20.2+20.5+20.5+20.7+20.7+20.7+21.0+21.1+21.5+21.6+22.1+22.3)/16 = 20.806$$

(b) The median for A is (23.0 + 24.0)/2 = 23.5. The median for B is (20.4 + 20.4)/2 = 20.4. The median for C is (21.0 + 21.0)/2 = 21.0. The median for D is (20.7 + 20.7)/2 = 20.7.

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

(c) $0.20(16) = 3.2 \approx 3$. Trim the 3 highest and 3 lowest observations.
The 20% trimmed mean for A is

$$(20.0 + 22.0 + 22.0 + 22.5 + 23.0 + 24.0 + 24.0 + 25.0 + 25.0 + 25.0)/10 = 23.25$$

The 20% trimmed mean for B is

$$(19.6 + 20.1 + 20.4 + 20.4 + 20.4 + 20.4 + 20.5 + 21.2 + 22.0 + 22.0)/10 = 20.70$$

The 20% trimmed mean for C is

$$(20.7 + 20.8 + 20.9 + 21.0 + 21.0 + 21.0 + 21.0 + 21.0 + 21.5 + 21.5)/10 = 21.04$$

The 20% trimmed mean for D is

$$(20.0 + 20.2 + 20.5 + 20.5 + 20.7 + 20.7 + 20.7 + 21.0 + 21.1 + 21.5)/10 = 20.69$$

(d) $0.25(17) = 4.25$. Therefore the first quartile is the average of the numbers in positions 4 and 5. $0.75(17) = 12.75$. Therefore the third quartile is the average of the numbers in positions 12 and 13.
A: $Q_1 = 21.0$, $Q_3 = 25.0$; B: $Q_1 = 19.85$, $Q_3 = 22.0$; C: $Q_1 = 20.75$, $Q_3 = 21.5$; D: $Q_1 = 20.1$, $Q_3 = 21.3$

(e) The variance for A is

$$\begin{aligned} s^2 &= \frac{1}{15}[18.0^2 + 18.0^2 + 18.0^2 + 20.0^2 + 22.0^2 + 22.0^2 + 22.5^2 + 23.0^2 + 24.0^2 \\ &\quad + 24.0^2 + 25.0^2 + 25.0^2 + 25.0^2 + 25.0^2 + 26.0^2 + 26.4^2 - 16(22.744^2)] = 8.2506 \end{aligned}$$

The standard deviation for A is $s = \sqrt{8.2506} = 2.8724$.
The variance for B is

$$\begin{aligned} s^2 &= \frac{1}{15}[18.8^2 + 18.9^2 + 18.9^2 + 19.6^2 + 20.1^2 + 20.4^2 + 20.4^2 + 20.4^2 + 20.4^2 \\ &\quad + 20.5^2 + 21.2^2 + 22.0^2 + 22.0^2 + 22.0^2 + 22.0^2 + 23.6^2 - 16(20.700^2)] = 1.8320 \end{aligned}$$

The standard deviation for B is $s = \sqrt{1.8320} = 1.3535$.
The variance for C is

$$\begin{aligned} s^2 &= \frac{1}{15}[20.2^2 + 20.5^2 + 20.5^2 + 20.7^2 + 20.8^2 + 20.9^2 + 21.0^2 + 21.0^2 + 21.0^2 \\ &\quad + 21.0^2 + 21.0^2 + 21.5^2 + 21.5^2 + 21.5^2 + 21.5^2 + 21.6^2 - 16(20.013^2)] = 0.17583 \end{aligned}$$

The standard deviation for C is $s = \sqrt{0.17583} = 0.4193$.
The variance for D is

$$\begin{aligned} s^2 &= \frac{1}{15}[20.0^2 + 20.0^2 + 20.0^2 + 20.0^2 + 20.2^2 + 20.5^2 + 20.5^2 + 20.7^2 + 20.7^2 \\ &\quad + 20.7^2 + 21.0^2 + 21.1^2 + 21.5^2 + 21.6^2 + 22.1^2 + 22.3^2 - 16(20.806^2)] = 0.55529 \end{aligned}$$

The standard deviation for D is $s = \sqrt{0.55529} = 0.7542$.

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

(f) Method A has the largest standard deviation. This could be expected, because of the four methods, this is the crudest. Therefore we could expect to see more variation in the way in which this method is carried out, resulting in more spread in the results.

(g) Other things being equal, a smaller standard deviation is better. With any measurement method, the result is somewhat different each time a measurement is made. When the standard deviation is small, a single measurement is more valuable, since we know that subsequent measurements would probably not be much different.

13. (a) All would be divided by 2.54.

(b) Not exactly the same, because the measurements would be a little different the second time.

14. (a) We will work in units of \$1000. Let S_0 be the sum of the original 10 numbers and let S_1 be the sum after the change. Then $S_0/10 = 70$, so $S_0 = 700$. Now $S_1 = S_0 - 100 + 1000 = 1600$, so the new mean is $S_1/10 = 160$.

(b) The median is unchanged at 55.

(c) Let $X_1, ..., X_{10}$ be the original 10 numbers. Let $T_0 = \sum_{i=1}^{10} X_i^2$. Then the variance is $(1/9)[T_0 - 10(70^2)] = 20^2 = 400$, so $T_0 = 52{,}600$. Let T_1 be the sum of the squares after the change. Then $T_1 = T_0 - 100^2 + 1000^2 = 1{,}042{,}600$. The new standard deviation is $\sqrt{(1/9)[T_1 - 10(160^2)]} = 295.63$.

15. (a) The sample size is $n = 16$. The tertiles have cutpoints $(1/3)(17) = 5.67$ and $(2/3)(17) = 11.33$. The first tertile is therefore the average of the sample values in positions 5 and 6, which is $(44 + 46)/2 = 45$. The second tertile is the average of the sample values in positions 11 and 12, which is $(76 + 79)/2 = 77.5$.

(b) The sample size is $n = 16$. The quintiles have cutpoints $(i/5)(17)$ for $i = 1, 2, 3, 4$. The quintiles are therefore the averages of the sample values in positions 3 and 4, in positions 6 and 7, in positions 10 and 11, and in positions 13 and 14. The quintiles are therefore $(23 + 41)/2 = 32$, $(46 + 49)/2 = 47.5$, $(74 + 76)/2 = 75$, and $(82 + 89)/2 = 85.5$.

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

16. (a) Seems certain to be an error.

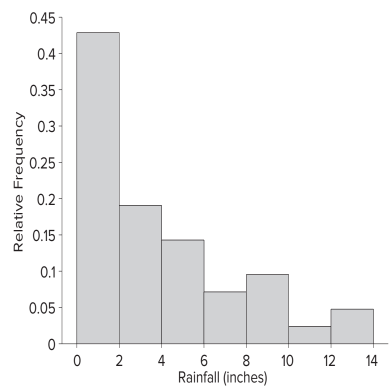
(b) Could be correct.

Section 1.3

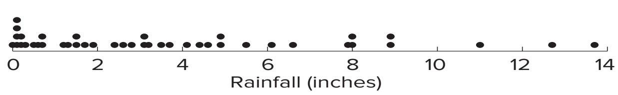
1. (a)

Stem	Leaf
0	011112235677
1	235579
2	468
3	11257
4	14699
5	5
6	16
7	9
8	0099
9	
10	
11	0
12	7
13	7

(b) Here is one histogram. Other choices for the endpoints are possible.

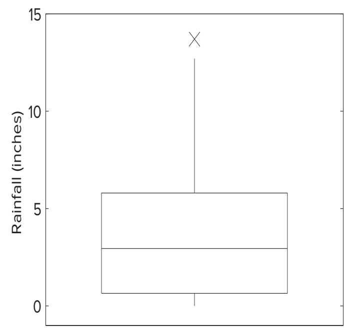


(c)



PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

(d)

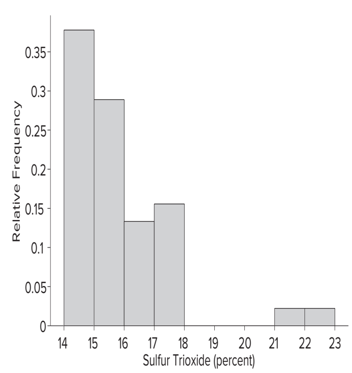


The boxplot shows one outlier.

2. (a)

Stem	Leaf
14	12333444467788889
15	0023344567799
16	124456
17	2222338
18	
19	
20	
21	9
22	4

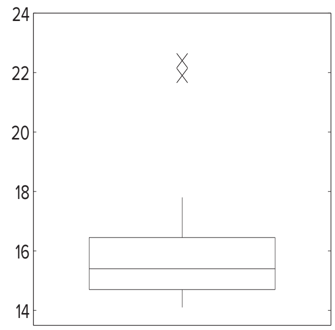
(b) Here is one histogram. Other choices for the endpoints are possible.



PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

(c)

(d)



The boxplot shows 2 outliers.

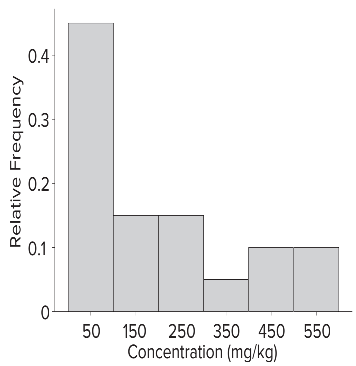
3.

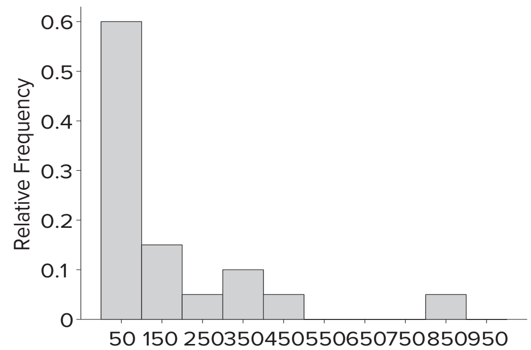
Stem	Leaf
1	1588
2	00003468
3	0234588
4	0346
5	2235666689
6	00233459
7	113558
8	568
9	1225
10	1
11	
12	2
13	06
14	
15	
16	
17	1
18	6
19	9
20	
21	
22	
23	3

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

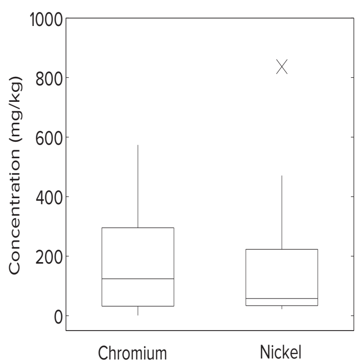
There are 23 stems in this plot. An advantage of this plot over the one in Figure 1.6 is that the values are given to the tenths digit instead of to the ones digit. A disadvantage is that there are too many stems, and many of them are empty.

4. (a) Here are histograms for each group. Other choices for the endpoints are possible.





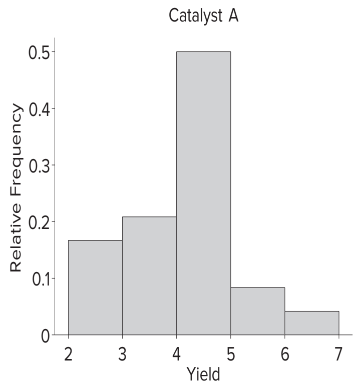
(b)

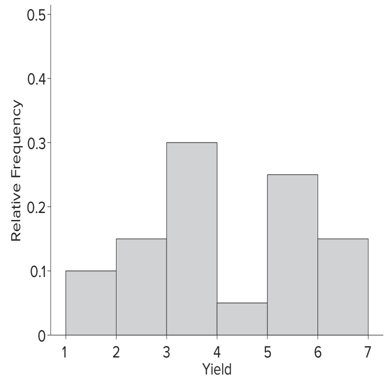


(c) The concentrations of nickel are on the whole lower than the concentrations of chromium. The nickel concentrations are highly skewed to the right, which can be seen from the median being much closer to the first quartile than the third. The chromium concentrations are somewhat less skewed. Finally, the nickel concentrations include an outlier.

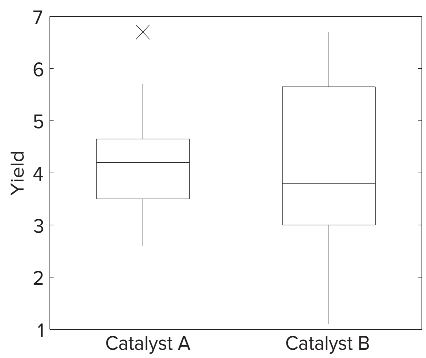
PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

5. (a) Here are histograms for each group. Other choices for the endpoints are possible.



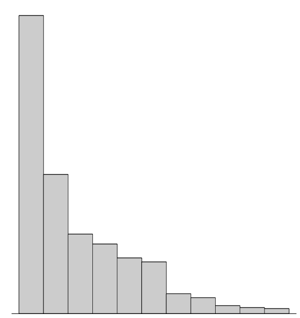


(b)



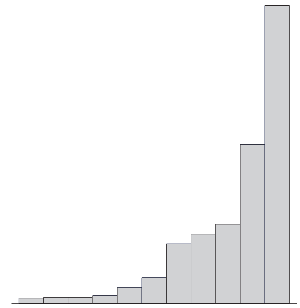
(c) The yields for catalyst B are considerably more spread out than those for catalyst A.The median yield for catalyst A is greater than the median for catalyst B. The median yield for B is closer to the first quartile than the third, but the lower whisker is longer than the upper one, so the median is approximately equidistant from the extremes of the data. Thus the yields for catalyst B are approximately symmetric. The largest yield for catalyst A is an outlier; the remaining yields for catalyst A are approximately symmetric.

6. (a) The histogram should be skewed to the right. Here is an example.

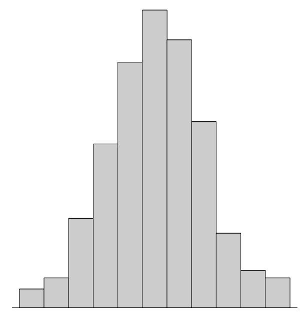


PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

(b) The histogram should be skewed to the left. Here is an example.



(c) The histogram should be approximately symmetric. Here is an example.



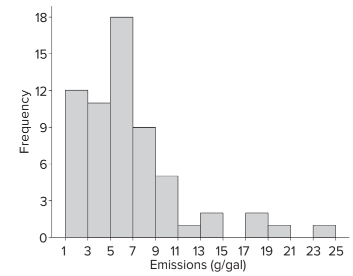
7. (a) The proportion is the sum of the relative frequencies (heights) of the rectangles above 240. This sum is approximately 0.14 + 0.10 + 0.05 + 0.01 + 0.02. This is closest to 30%.

(b) The height of the rectangle over the interval 240–260 is greater than the sum of the heights of the rectangles over the interval 280–340. Therefore there are more men in the interval 240–260 mg/dL.

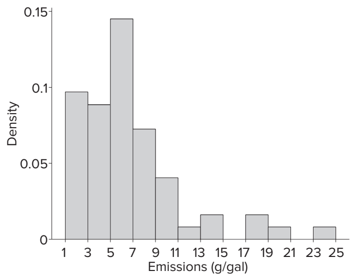
8. The relative frequencies of the rectangles shown are 0.05, 0.1, 0.15, 0.25, 0.2, and 0.1. The sum of these relative frequencies is 0.85. Since the sum of all the relative frequencies must be 1, the missing rectangle has a height of 0.15.

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

9. (a)

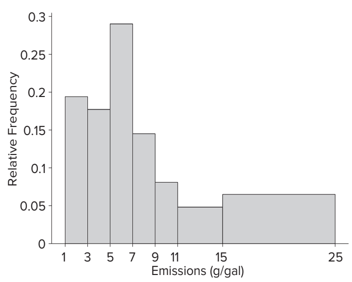


(b)



(c) Yes, the shapes of the histograms are the same.

10. (a)



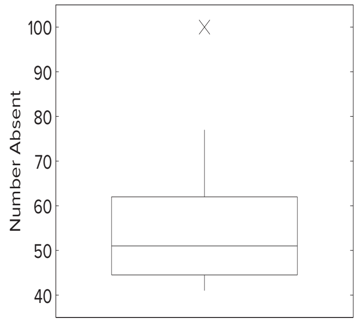
(b) No

(c) The class interval widths are unequal.

(d) The classes 11−<15 and 15−<25

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

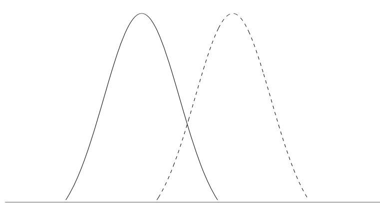
11. (a)

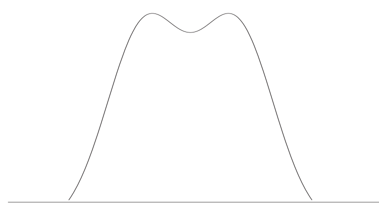


(b) Yes. The value 100 is an outlier.

12. The mean. The median, and the first and third quartiles are indicated directly on a boxplot, and the interquartile range can be computed as the difference between the first and third quartiles.

13. The figure on the left is a sketch of separate histograms for each group. The histogram on the right is a sketch of a histogram for the two groups combined. There is more spread in the combined histogram than in either of the separate ones. Therefore the standard deviation of all 200 resistances is greater than 5Ω. The answer is (ii).





14. (a) True

(b) False

(c) True

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

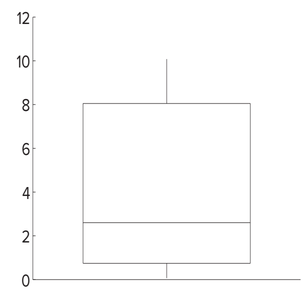
(d) False

(e) False

(f) True

15. (a) IQR = 3rd quartile − 1st quartile. A: IQR = 6.02 − 1.42 = 4.60, B: IQR = 9.13 − 5.27 = 3.86

(b) Yes, since the minimum is within 1.5 IQR of the first quartile and the maximum is within 1.5 IQR of the third quartile, there are no outliers, and the given numbers specify the boundaries of the box and the ends of the whiskers.



(c) No. The minimum value of −2.235 is an "outlier," since it is more than 1.5 times the interquartile range below the first quartile. The lower whisker should extend to the smallest point that is not an outlier, but the value of this point is not given.

16. (a) (4)

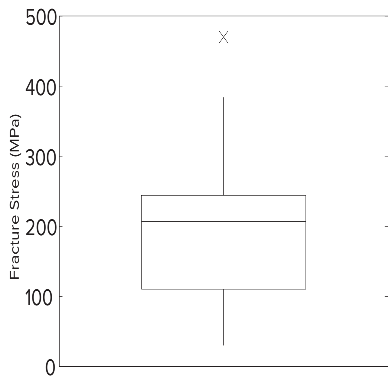
(b) (2)

(c) (1)

(d) (3)

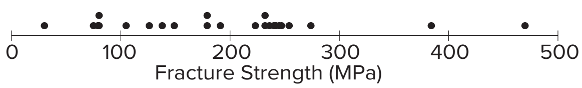
PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

17. (a)



(b) The boxplot indicates that the value 470 is an outlier.

(c)



(d) The dotplot indicates that the value 384 is detached from the bulk of the data, and thus could be considered to be an outlier.

18.

Catalyst B	Stem	Catalyst A
61	1	
965	2	6669
887541	3	14689
3	4	011344556789
98520	5	27
743	6	7

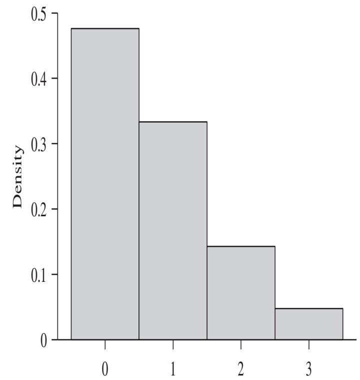
The distribution of yields for Catalyst A is unimodal, while the distribution of yields for Catalyst B is bimodal.

19. (a) The mean is 0.7619; the median is 1.

(b) Skewed to the left.

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

(c) The histogram is skewed to the right.



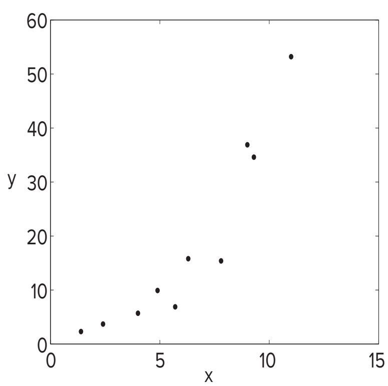
20. (a) iii

(b) i

(c) iv

(d) ii

21. (a)

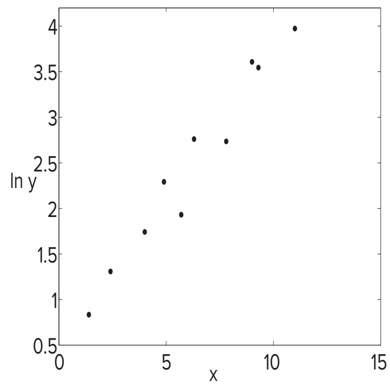


The relationship is non-linear.

(b)

x	1.4	2.4	4.0	4.9	5.7	6.3	7.8	9.0	9.3	11.0
$\ln y$	0.83	1.31	1.74	2.29	1.93	2.76	2.73	3.61	3.54	3.97

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.



The relationship is approximately linear.

(c) It would be easier to work with x and $\ln y$, because the relationship is approximately linear.

Supplementary Exercises for Chapter 1

1. (a) The mean will be divided by 2.2.

   (b) The standard deviation will be divided by 2.2.

2. (a) The mean will increase by 50 g.

   (b) The standard deviation will be unchanged.

3. (a) False. The true percentage could be greater than 5%, with the observation of 4 out of 100 due to sampling variation.

   (b) True

   (c) False. If the result differs greatly from 5%, it is unlikely to be due to sampling variation.

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

(d) True. If the result differs greatly from 5%, it is unlikely to be due to sampling variation.

4. (a) No. This could well be sampling variation.

   (b) Yes. It is virtually impossible for sampling variation to be this large.

5. (a) It is not possible to tell by how much the mean changes, because the sample size is not known.

   (b) If there are more than two numbers on the list, the median is unchanged. If there are only two numbers on the list, the median is changed, but we cannot tell by how much.

   (c) It is not possible to tell by how much the standard deviation changes, both because the sample size is unknown and because the original standard deviation is unknown.

6. (a) The sum of the numbers decreases by $12.9 - 1.29 = 11.61$, so the mean decreases by $11.61/15 = 0.774$.

   (b) No, it is not possible to determine the value of the mean after the change, since the original mean is unknown.

   (c) The median is the eighth number when the list is arranged in order, and this is unchanged.

   (d) It is not possible to tell by how much the standard deviation changes, because the original standard deviation is unknown.

7. (a) The mean decreases by 0.774.

   (b) The value of the mean after the change is $25 - 0.774 = 24.226$.

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

(c) The median is unchanged.

(d) It is not possible to tell by how much the standard deviation changes, because the original standard deviation is unknown.

8. (a) The sum of the numbers 284.34, 173.01, 229.55, 312.95, 215.34, 188.72, 144.39, 172.79, 139.38, 197.81, 303.28, 256.02, 658.38, 105.14, 295.24, 170.41 is 3846.75. The mean is therefore 3846.75/16 = 240.4219.

(b) The 16 values arranged in increasing order are:

105.14, 139.38, 144.39, 170.41, 172.79, 173.01, 188.72, 197.81,

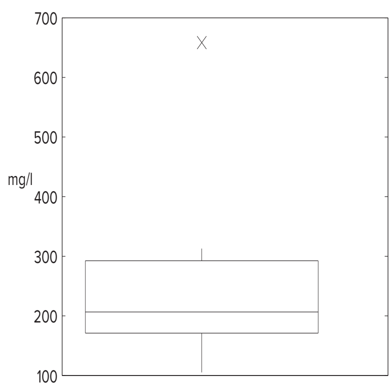
215.34, 229.55, 256.02, 284.34, 295.24, 303.28, 312.95, 658.38

The median is the average of the 8th and 9th numbers, which is $(197.81 + 215.34)/2 = 206.575$.

(c) $0.25(17) = 4.25$, so the first quartile is the average of the 4th and 5th numbers, which is $(170.41 + 172.79)/2 = 171.60$.

(d) $0.75(17) = 12.75$, so the third quartile is the average of the 12th and 13th numbers, which is $(284.34 + 295.24)/2 = 289.79$.

(e)



The median is closer to the first quartile than to the third quartile, which indicates that the sample is skewed a bit to the right. In addition, the sample contains an outlier.

9. Statement (i) is true. The sample is skewed to the right.

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

10. (a) False. The length of the whiskers is at most 1.5 IQR.

    (b) False. The length of the whiskers is at most 1.5 IQR.

    (c) True. A whisker extends to the most extreme data point that is within 1.5 IQR of the first or third quartile.

    (d) True. A whisker extends to the most extreme data point that is within 1.5 IQR of the first or third quartile.

11. (a) Incorrect, the total area is greater than 1.

    (b) Correct. The total area is equal to 1.

    (c) Incorrect. The total area is less than 1.

    (d) Correct. The total area is equal to 1.

12. (i) It would be skewed to the right. The mean is greater than the median. Also note that half the values are between 0 and 0.10, so the left-hand tail is very short.

13. (a) Skewed to the left. The 85th percentile is much closer to the median (50th percentile) than the 15th percentile is. Therefore the histogram is likely to have a longer left-hand tail than right-hand tail.

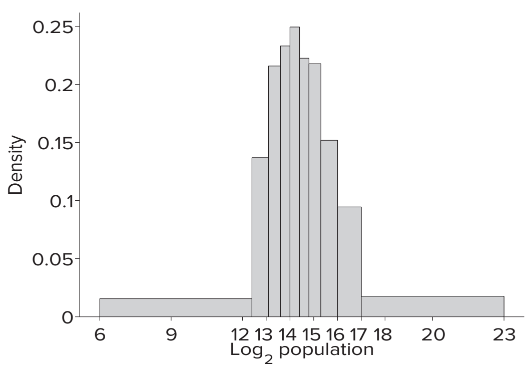
    (b) Skewed to the right. The 15th percentile is much closer to the median (50th percentile) than the 85th percentile is. Therefore the histogram is likely to have a longer right-hand tail than left-hand tail.

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

14.

Class Interval	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
0–< 1	12	0.2857	12	0.2857
1–< 2	6	0.1429	18	0.4286
2–< 3	3	0.0714	21	0.5000
3–< 4	5	0.1190	26	0.6190
4–< 5	5	0.1190	31	0.7381
5–< 6	1	0.0238	32	0.7619
6–< 7	2	0.0476	34	0.8095
7–< 8	1	0.0238	35	0.8333
8–< 9	4	0.0952	39	0.9286
9–< 10	0	0.0000	39	0.9286
10–< 11	0	0.0000	39	0.9286
11–< 12	1	0.0238	40	0.9524
12–< 13	1	0.0238	41	0.9762
13–< 14	1	0.0238	42	1.0000

15. (a)

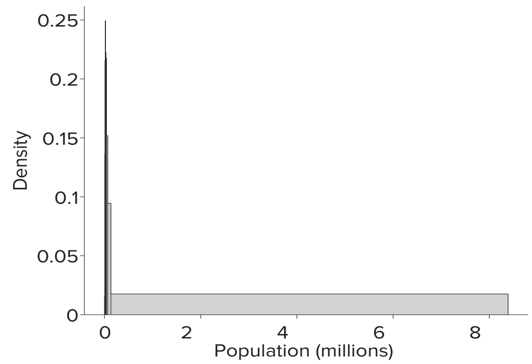


(b) 0.14

(c) Approximately symmetric

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

(d)



The data on the raw scale are skewed so much to the right that it is impossible to see the features of the histogram.

16. (a) The mean is

$$\frac{1}{23}(2099 + 528 + 2030 + 1350 + 1018 + 384 + 1499 + 1265 + 375 + 424 + 789 + 810 + 522 + 513 + 488 + 200 + 215 + 486 + 257 + 557 + 260 + 461 + 500) = 740.43$$

(b) The variance is

$$s^2 = \frac{1}{22}[2099^2 + 528^2 + 2030^2 + 1350^2 + 1018^2 + 384^2 + 1499^2 + 1265^2 + 375^2 + 424^2 + 789^2 + 810^2 + 522^2 + 513^2 + 488^2 + 200^2 + 215^2 + 486^2 + 257^2 + 557^2 + 260^2 + 461^2 + 500^2 - 23(740.43^2)] = 302320.26$$

The standard deviation is $s = \sqrt{302320.26} = 549.84$.

(c) The 23 values, arranged in increasing order, are:

200, 215, 257, 260, 375, 384, 424, 461, 486, 488, 500, 513, 522, 528, 557, 789, 810, 1018, 1265, 1350, 1499, 2030, 2099

The median is the 12th value, which is 513.

(d)

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

(e) Since $(0.1)(23) \approx 2$, the 10% trimmed mean is computed by deleting the two highest and two lowest values, and averaging the rest.

The 10% trimmed mean is

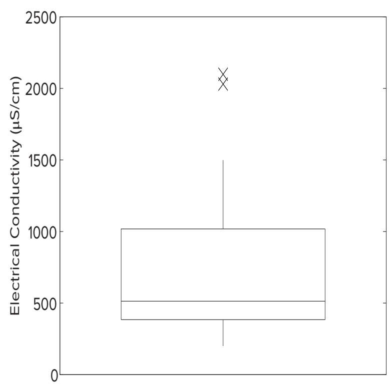
$$\frac{1}{19}(257 + 260 + 375 + 384 + 424 + 461 + 486 + 488 + 500 + 513 + 522 + 528 + 557 + 789 + 810 + 1018 + 1265 + 1350 + 1499) = 657.16$$

(f) $0.25(24) = 6$. Therefore, when the numbers are arranged in increasing order, the first quartile is the number in position 6, which is 384.

(g) $0.75(24) = 18$. Therefore, when the numbers are arranged in increasing order, the third quartile is the number in position 18, which is 1018.

(h) IQR = 3rd quartile − 1st quartile = $1018 - 384 = 634$.

(i)

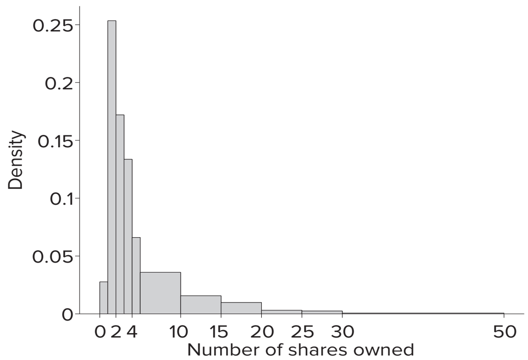


(j) The points 2030 and 2099 are outliers.

(k) skewed to the right

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

17. (a)

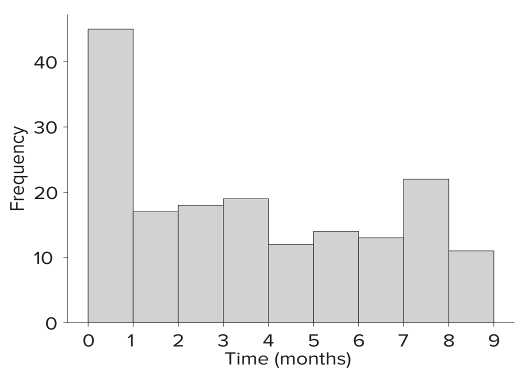


(b) The sample size is 651, so the median is approximated by the point at which the area to the left is 0.5 = 325.5/651. The area to the left of 3 is 295/651, and the area to the left of 4 is 382/651. The point at which the area to the left is 325.5/651 is $3 + (325.5 - 295)/(382 - 295) = 3.35$.

(c) The sample size is 651, so the first quartile is approximated by the point at which the area to the left is 0.25 = 162.75/651. The area to the left of 1 is 18/651, and the area to the left of 2 is 183/651. The point at which the area to the left is 162.75/651 is $1 + (162.75 - 18)/(183 - 18) = 1.88$.

(d) The sample size is 651, so the third quartile is approximated by the point at which the area to the left is 0.75 = 488.25/651. The area to the left of 5 is 425/651, and the area to the left of 10 is 542/651. The point at which the area to the left is 488.25/651 is $5 + (10 - 5)(488.25 - 425)/(542 - 425) = 7.70$.

18. (a)



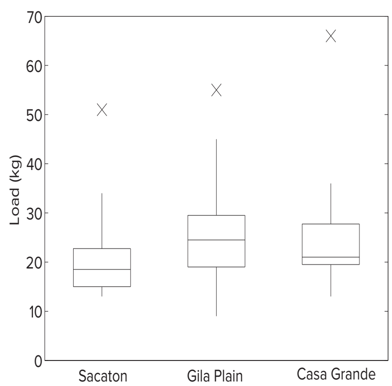
(b) The sample size is 171, so the median is the value in position $(171 + 1)/2 = 86$ when the values are arranged in order. There are $45 + 17 + 18 = 80$ values less than or equal to 3, and $80 + 19 = 99$ values less than or equal to 4. Therefore the class interval 3 – < 4 contains the median.

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.

(c) The sample size is 171, so the first quartile is the value in position $0.25(171 + 1) = 43$ when the values are arranged in order. There are 45 values in the first class interval $0 - < 1$. Therefore the class interval $0 - < 1$ contains the first quartile.

(d) The sample size is 171, so the third quartile is the value in position $0.75(171 + 1) = 129$ when the values are arranged in order. There are $45+17+18+19+12+14 = 125$ values less than or equal to 6, and $125+13 = 138$ values less than or equal to 7. Therefore the class interval $6 - < 7$ contains the third quartile.

19. (a)



(b) Each sample contains one outlier.

(c) In the Sacaton boxplot, the median is about midway between the first and third quartiles, suggesting that the data between these quartiles are fairly symmetric. The upper whisker of the box is much longer than the lower whisker, and there is an outlier on the upper side. This indicates that the data as a whole are skewed to the right. In the Gila Plain boxplot data, the median is about midway between the first and third quartiles, suggesting that the data between these quartiles are fairly symmetric. The upper whisker is slightly longer than the lower whisker, and there is an outlier on the upper side. This suggest that the data as a whole are somewhat skewed to the right. In the Casa Grande boxplot, the median is very close to the first quartile. This suggests that there are several values very close to each other about one-fourth of the way through the data. The two whiskers are of about equal length, which suggests that the tails are about equal, except for the outlier on the upper side.

PROPRIETARY MATERIAL. © The McGraw-Hill Companies, Inc. All rights reserved. No part of this Manual may be displayed, reproduced or distributed in any form or by any means, without the prior written permission of the publisher, or used beyond the limited distribution to teachers and educators permitted by McGraw-Hill for their individual course preparation. If you are a student using this Manual, you are using it without permission.