Elementary Statistics Using Excel 7th edition Triola Solutions Manual

SKU: 9780136921721-SOLUTIONS

Chapter 2: Exploring Data with Tables and Graphs

Section 2-1: Frequency Distributions for Organizing and Summarizing Data

1. The table summarizes 1000 commute times. It is not possible to identify the exact values of all of the original times.
2. The classes of 0–30, 30–60, …, 120–150 overlap, so it is not always clear which class we should put a value in. For example, the commute time value of 30 minutes could go in the first class or the second class. The classes should be mutually exclusive, meaning that there is no overlap.
3.

Daily Commute Time in Boston (minutes)	Relative Frequency
0–29	46.8%
30–59	42.2%
60–89	9.2%
90–119	1.0%
120–149	0.8%

4. The sum of the relative frequencies is 125%, but it should be 100%, with a small round off error. All of the relative frequencies appear to be roughly the same, but if they are from a normal distribution, they should start low, reach a maximum, and then decrease.
5. Class width: 10
   Class midpoints: 24.5, 34.5, 44.5, 54.5, 64.5, 74.5, 84.5
   Class boundaries: 19.5, 29.5, 39.5, 49.5, 59.5, 69.5, 79.5, 89.5
   Number: 91
6. Class width: 10
   Class midpoints: 24.5, 34.5, 44.5, 54.5, 64.5, 74.5
   Class boundaries: 19.5, 29.5, 39.5, 49.5, 59.5, 69.5, 79.5
   Number: 91
7. Class width: 100
   Class midpoints: 49.5, 149.5, 249.5, 349.5, 449.5, 549.5, 649.5
   Class boundaries: –0.5, 99.5, 199.5, 299.5, 399.5, 499.5, 599.5, 699.5
   Number: 153
8. Class width: 100
   Class midpoints: 149.5, 249.5, 349.5, 449.5, 549.5
   Class boundaries: 99.5, 199.5, 299.5, 399.5, 499.5, 599.5
   Number: 147
9. No. The maximum frequency is in the second class instead of being near the middle, so the frequencies below the maximum do not mirror those above the maximum.
10. Yes. The frequencies start low, reach a maximum of 38, and then decrease. The values below the maximum are very roughly a mirror image of those above it.
11. Yes. Except for the single value that lies between 600 and 699, the frequencies start low, reach a maximum of 90, and then decrease. The values below the maximum are very roughly a mirror image of those above it. (That single value between 600 and 699 is an outlier that makes the determination of a normal distribution somewhat questionable, but using a *loose interpretation* of the criteria for normality, it is reasonable to conclude that the distribution is normal.)
12. Yes. Except for two values that lie between 500 and 599, there is a low frequency of 25, then a maximum frequency of 92, and then a low frequency of 28. The values below and above the maximum are roughly a mirror image. (Those two values between 500 and 599 are outliers that make the determination of a normal distribution somewhat questionable, but using a *loose interpretation* of the criteria for normality, it is reasonable to conclude that the distribution is normal.)

Copyright © 2022 Pearson Education, Inc.

Elementary Statistics Using Excel 7th edition Triola Solutions Manual

13. The data amounts do not appear to have a normal distribution. The distribution does not appear to be symmetric because the frequencies preceding the maximum frequency of 16 are far outweighed by the frequencies following the maximum.

Daily Commute Time in Chicago (Minutes)	Frequency
0–14	5
15–29	16
30–44	14
45–59	9
60–74	5
75–89	1

14. The ages do appear to have a normal distribution.

Age (years)	Frequency
40–44	2
45–49	7
50–54	10
55–59	10
60–64	6
65–69	3
70–74	1

15.

Duration (sec)	Frequency
125–149	1
150–174	0
175–199	0
200–224	3
225–249	34
250–274	12

16. The intensities do not appear to have a normal distribution.

Tornado F-Scale	Frequency
0	24
1	16
2	2
3	2
4	1

17.

Burger King Lunch Service Times (sec)	Frequency
70–109	11
110–149	23
150–189	7
190–229	6
230–269	3
230–269	6

Copyright © 2022 Pearson Education, Inc.

Elementary Statistics Using Excel 7th edition Triola Solutions Manual

18.

Burger King Dinner Service Times (sec)	Frequency
30–69	1
70–109	6
110–149	26
150–189	7
190–229	3
230–269	6
270–309	1

19. The distribution does appear to be a normal distribution.

Weight (kg)	Frequency
40–49	2
50–59	22
60–69	23
70–79	13
80–89	3
90–99	4

20. The distribution does appear to be a normal distribution.

Weight (g)	Frequency
4.300–4.399	11
4.400–4.499	17
4.500–4.599	28
4.600–4.699	16
4.700–4.799	3

21. Because there are disproportionately more 0s and 5s, it appears that the heights were reported instead of measured. Consequently, it is likely that the results are not very accurate.

Last Digit	Frequency
0	9
1	2
2	1
3	3
4	1
5	15
6	2
7	0
8	3
9	1

Copyright © 2022 Pearson Education, Inc.

Elementary Statistics Using Excel 7th edition Triola Solutions Manual

22. Because there are disproportionately more 0s and 5s, it appears that the heights were reported instead of measured. There does appear to be a gap due to the tendency of respondents to round their heights to values ending in 0 or 5. Because the results appear to be reported instead of measured, it is likely that the results are not very accurate.

Last Digit	Frequency
0	26
1	1
2	1
3	2
4	2
5	12
6	1
7	0
8	4
9	1

23. The actresses appear to be generally younger than the actors.

Age When Oscar Was Won	Actresses	Actors
20–29	34.1%	1.1%
30–39	37.4%	31.9%
40–49	16.5%	41.8%
50–59	3.3%	17.6%
60–69	6.6%	6.6%
70–79	1.1%	1.1%
80–89	1.1%	0.0%

24. There do appear to be differences, but overall they are not very substantial differences.

Blood Platelet Count	Males	Females
0–99	0.7%	0.0%
100–199	33.3%	17.0%
200–299	58.8%	62.6%
300–399	6.5%	19.0%
400–499	0.0%	0.0%
500–599	0.0%	1.4%
600–699	0.7%	0.0%

25.

Age (years) of Best Actress When Oscar Was Won	Cumulative Frequency
Less than 30	31
Less than 40	65
Less than 50	80
Less than 60	83
Less than 70	89
Less than 80	90
Less than 90	91

Copyright © 2022 Pearson Education, Inc.

Elementary Statistics Using Excel 7th edition Triola Solutions Manual

26.

Age (years) of Best Actor When Oscar Was Won	Cumulative Frequency
Less than 30	1
Less than 40	30
Less than 50	68
Less than 60	84
Less than 70	90
Less than 80	91

27. No. The United States has 37.1% of the cost of piracy for only the five countries listed, not the total cost of piracy for all countries. Because only the top five costs of piracy are listed, we know only that any other country must have a cost less than $1.9 billion.

Country	Relative Frequency
United States	37.1%
China	35.5%
India	11.0%
France	8.6%
United Kingdom	7.8%

28. Yes, it appears that births occur on the days of the week with frequencies that are about the same.

Day	Relative Frequency
Monday	13.0%
Tuesday	16.5%
Wednesday	18.0%
Thursday	14.3%
Friday	14.3%
Saturday	10.8%
Sunday	13.3%

29. It is very similar to Table 2-2. Both frequency distributions begin with a low frequency in the first class, followed by the maximum frequency in the second class, and the frequencies are generally lower as you progress from top to bottom in the table. (TI data: Frequencies are 79, 148, 157, 43, 49, 6, 9, 0, 0, 9.)

Daily Commute Time in Los Angeles (minutes)	Frequency
0–14	157
15–29	324
30–44	282
45–59	103
60–74	98
75–89	7
90–104	18
105–119	0
120–134	0
135–149	11

Copyright © 2022 Pearson Education, Inc.

Elementary Statistics Using Excel 7th edition Triola Solutions Manual

30. It is very similar to the frequency distribution for Exercise 1. Both frequency distributions begin with the maximum frequency, and then the frequencies decrease moving from top to bottom in the table. (TI data: Frequencies are 289, 154, 43, 11, 0, 3.)

Daily Commute Time in Dallas (minutes)	Frequency
0–29	547
30–59	339
60–89	92
90–119	15
120–149	0
150–179	7

31. Yes, the frequency distribution appears to be a normal distribution.

Systolic Blood Pressure (mm Hg)	Frequency
80–99	11
100–119	116
120–139	131
140–159	34
160–179	7
180–199	1

32. Yes, the frequency distribution appears to be a normal distribution.

Diastolic Blood Pressure (mm Hg)	Frequency
40–54	27
55–69	107
70–84	133
85–99	31
100–114	2

33. Yes, the frequency distribution appears to be a normal distribution.

Magnitude	Frequency
1.00–1.49	19
1.50–1.99	97
2.00–2.49	187
2.50–2.99	147
3.00–3.49	100
3.50–3.99	38
4.00–4.49	8
4.50–4.99	4

34. No, the frequency distribution does not appear to be a normal distribution.

Depth (km)	Frequency
0.0–9.9	539
10.0–19.9	49
20.0–29.9	10
30.0–39.9	1
40.0–49.9	1

Copyright © 2022 Pearson Education, Inc.

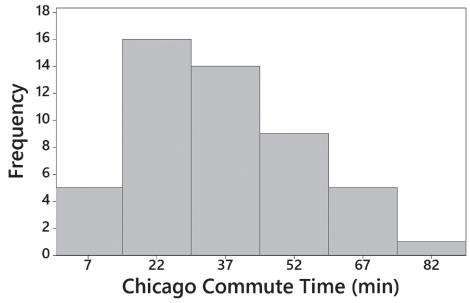
Elementary Statistics Using Excel 7th edition Triola Solutions Manual

35. An outlier can dramatically increase the number of classes.

Weight (lb)	With Outlier	Without Outlier
200–219	6	6
220–239	5	5
240–259	12	12
260–279	36	36
280–299	87	87
300–319	28	28
320–339	0	
340–359	0	
360–379	0	
380–399	0	
400–419	0	
420–439	0	
440–459	0	
460–479	0	
480–499	0	
500–519	1	

Section 2-2: Histograms

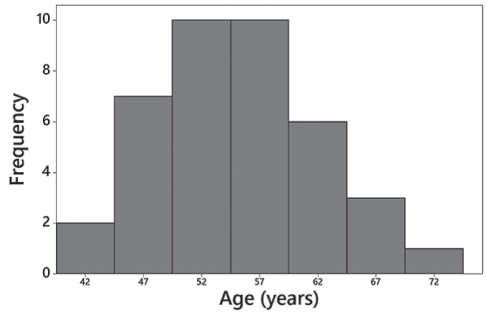
1. The histogram should be bell-shaped.
2. Not necessarily. Because the sample subjects themselves chose to be included, the voluntary response sample might not be representative of the population.
3. With a data set that is so small, the true nature of the distribution cannot be seen with a histogram.
4. The outlier will result in a single bar that is far away from all of the other bars in the histogram, and the height of that bar will correspond to a frequency of 1.
5. 40
6. Approximate values: Class width: 0.1 gram, lower limit of first class: 5.5 grams, upper limit of first class: 5.6 grams
7. The shape of the graph would not change. The vertical scale would be different, but the relative heights of the bars would be the same.
8. 40 of the quarters are "pre-1964" made with 90% silver and 10% copper, and the other 40 quarters are "post-1964" made with a copper-nickel alloy. The histogram depicts weights from two different populations of quarters.
9. Because the data are skewed to the right, the histogram does not appear to depict data from a population with a normal distribution.



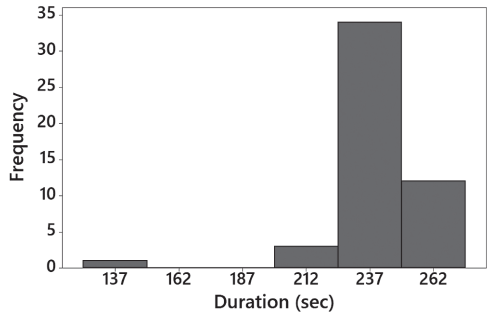
Copyright © 2022 Pearson Education, Inc.

Elementary Statistics Using Excel 7th edition Triola Solutions Manual

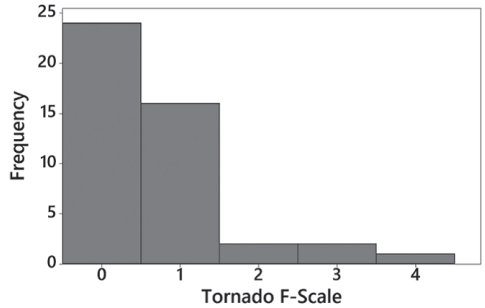
10. The histogram does appear to be the graph of data from a population with a normal distribution.



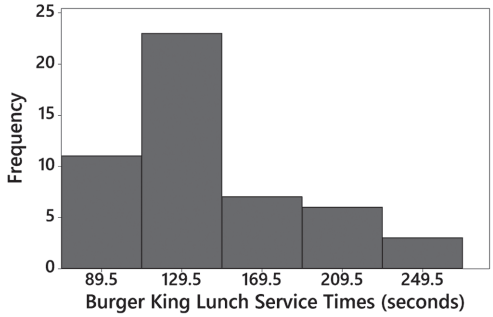
11. Because it is far from being bell-shaped, the histogram does not appear to depict data from a population with a normal distribution.



12. The histogram appears to be skewed to the right (or positively skewed).



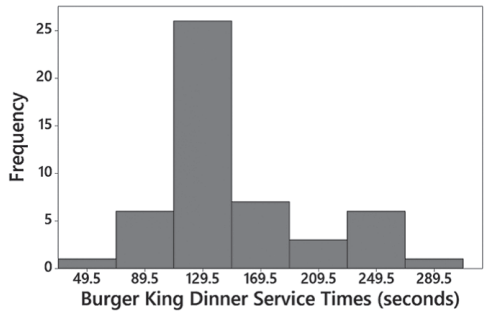
13. Using a loose interpretation of "normal," the histogram appears to be approximately normal. The histogram does show skewness to the right.



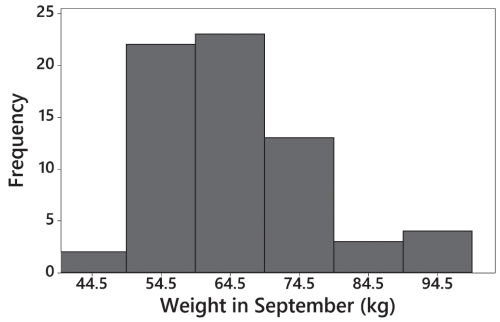
Copyright © 2022 Pearson Education, Inc.

Elementary Statistics Using Excel 7th edition Triola Solutions Manual

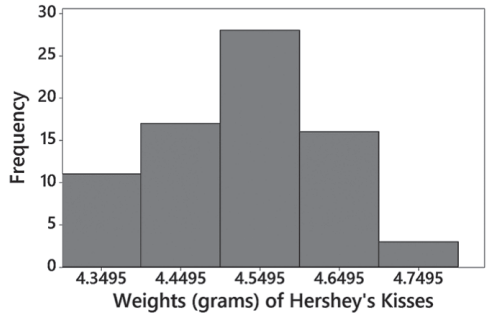
14. The histogram isn't very close to being bell-shaped, so it does not appear to depict data from a population with a normal distribution. The departure from a normal distribution is not too substantial.



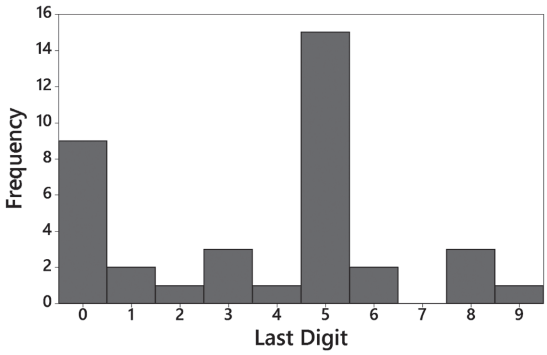
15. Because the histogram is roughly bell-shaped, it does appear to depict a normal distribution.



16. Because the histogram isn't close enough to being bell-shaped, it does not appear to depict data from a population with a normal distribution.



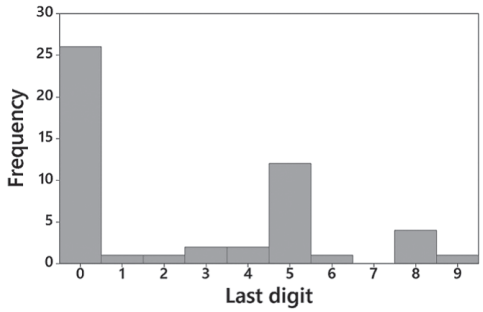
17. The digits 0 and 5 appear to occur more often than the other digits, so it appears that the heights were reported and not actually measured. This suggests that the data might not be very useful.



Copyright © 2022 Pearson Education, Inc.

Elementary Statistics Using Excel 7th edition Triola Solutions Manual

18. The digits 0 and 5 appear to occur more often than the other digits, so it appears that the weights were reported and not actually measured. This suggests that the data might not be very useful.



19. Only part (c) appears to represent data from a normal distribution. Part (a) has a systematic pattern that is not that of a straight line, part (b) has points that are not close to a straight-line pattern, and part (d) is really bad because it shows a systematic pattern and points that are not close to a straight-line pattern.

20. The histogram for the movie lengths suggests that the data have a distribution that is approximately normal with an outlier of 120 minutes. The histogram for the time of tobacco use appears to be similar to the histogram for the times of alcohol use, but both distributions appear to be skewed to the right, with each histogram having an outlier.

Section 2-3: Graphs That Enlighten and Graphs That Deceive

1. The data set is too small for a dotplot to reveal important characteristics of the data. Because the data are listed in order for each of the last several years, a time-series graph would be most effective for these data.

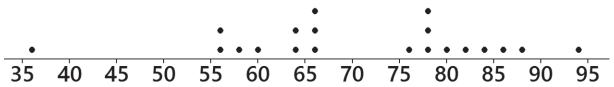
2. Yes, the original data values can be found from the stemplot.

```
5 | 4
6 | 8
7 | 0369
8 | 123
9 | 8
```

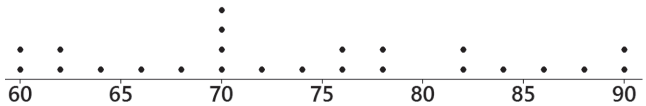
3. No. Graphs should be constructed in a way that is fair and objective. The readers should be allowed to make their own judgments, instead of being manipulated by misleading graphs.

4. No. If the sample is a bad sample, such as one obtained from voluntary responses, there are no graphs or statistical methods that can be used to salvage the data.

5. The pulse rate of 36 beats per minute appears to be an outlier.



6. There do not appear to be any outliers.



7. The data are arranged in order from lowest to highest, as 36, 56, 56, and so on.

```
3 | 6
4 |
5 | 668
6 | 044666
7 | 6888
8 | 02468
9 | 4
```

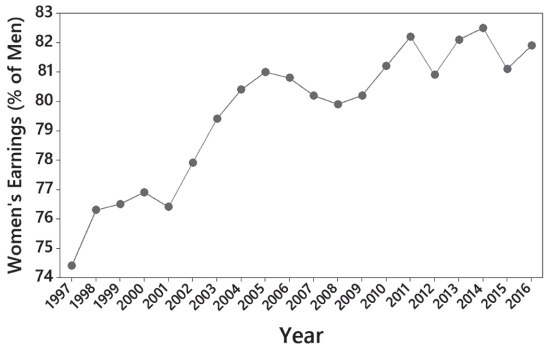
Copyright © 2022 Pearson Education, Inc.

Elementary Statistics Using Excel 7th edition Triola Solutions Manual

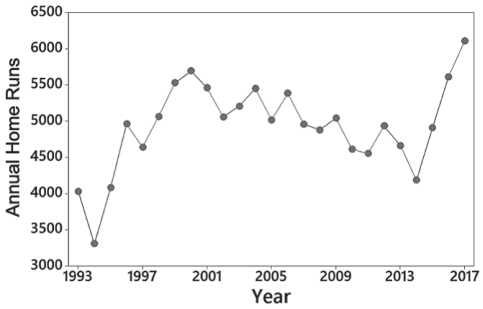
8. The two values closest to the middle are 72 mm Hg and 74 mm Hg.

6 | 0022468
7 | 0000246688
8 | 22468
9 | 00

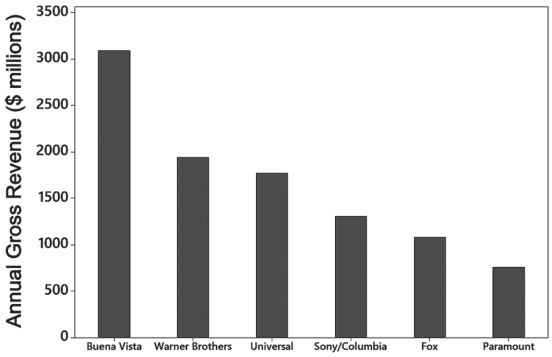
9. There is a gradual upward trend that appears to be leveling off in recent years. An upward trend would be helpful to women so that their earnings become equal to those of men.



10. The numbers of home runs rose from 1994 to 2000, then there was a gradual decline from 2000 to 2014, and there was a steady increase after 2014.



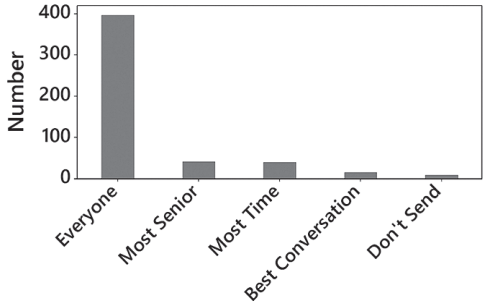
11. Given that box office receipts are closely tracked and widely reported, these data are very accurate.



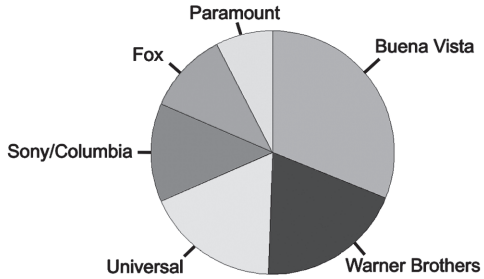
Copyright © 2022 Pearson Education, Inc.

Elementary Statistics Using Excel 7th edition Triola Solutions Manual

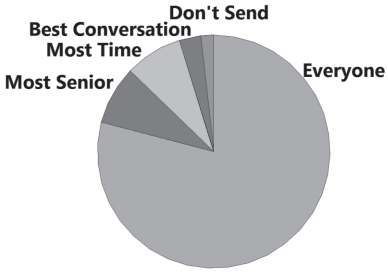
12. The overwhelming response was that thank-you notes should be sent to everyone who is met during a job interview. Given what is at stake, that seems like a wise strategy.



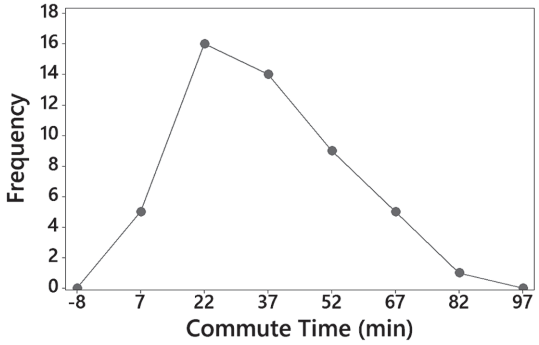
13.



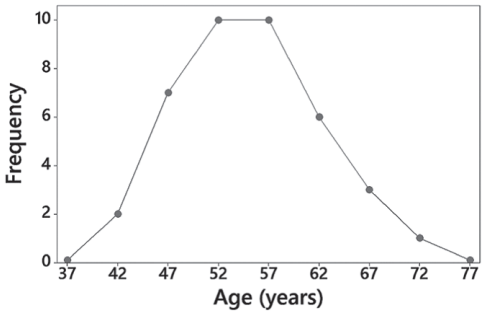
14.



15. The distribution does not have much skewness.



16. The distribution does not appear to be skewed.



17. Because the vertical scale starts with a frequency of 200 instead of 0, the difference between the "no" and "yes" responses is greatly exaggerated. The graph makes it appear that about five times as many respondents said "no," when the ratio is actually a little less than 2.5 to 1.

18. The fare increased from $1 to $2.50, so it increased by a factor of 2.5. But when the larger bill is drawn so that the width is 2.5 times that of the smaller bill and the height is 2.5 times that of the smaller bill, the larger bill has an area that is 6.25 times that of the smaller bill (instead of being 2.5 times its size, as it should be). The illustration greatly exaggerates the increase in the fare.

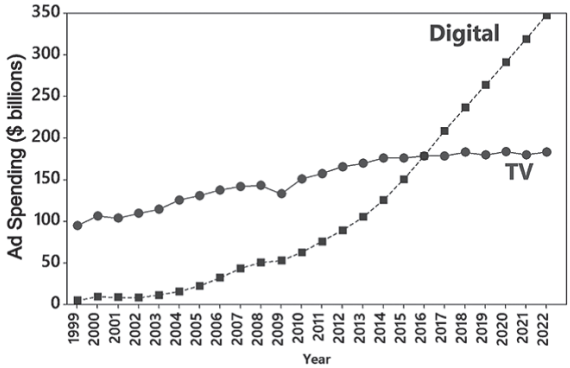
Copyright © 2022 Pearson Education, Inc.

Elementary Statistics Using Excel 7th edition Triola Solutions Manual

19. The two costs are one-dimensional in nature, but the baby bottles are three-dimensional objects. The $4500 cost isn't even twice the $2600 cost, but the baby bottles make it appear that the larger cost is about five times the smaller cost.
20. The graph is misleading because it depicts one-dimensional data with three-dimensional boxes. See the first and last boxes in the graph. Workers with advanced degrees have annual incomes that are roughly 3 times the incomes of those with no high school diplomas, but the graph exaggerates this difference by making it appear that workers with advanced degrees have incomes that are roughly 27 times the amounts for workers with no high school diploma.
21.

```
96 |
96 | 59
97 | 0001112333444
97 | 55666666788888999
98 | 555566666666666666677777788888889
99 | 001244
99 | 56
```

22. Around 2016, digital ad spending surpassed TV ad spending. TV ad spending appears to be relatively stable with only slight increases over time, but digital ad spending is growing at a much faster rate than TV ad spending.



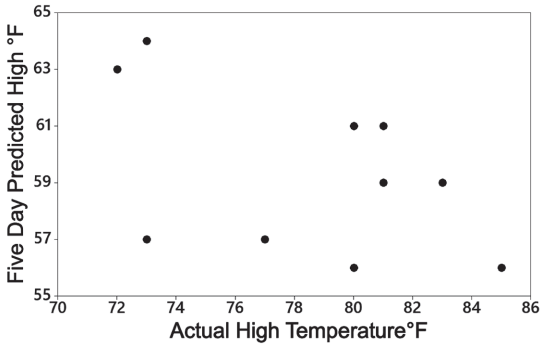
Section 2-4: Scatterplots, Correlation, and Regression

1. The term *linear* refers to a straight line, and r measures how well a scatterplot of the sample paired data fits a straight-line pattern.
2. No. Finding the presence of a statistical correlation between two variables does not justify any conclusion that one of the variables is a cause of the other.
3. A scatterplot is a graph of paired (x, y) quantitative data. It helps us by providing a visual image of the data plotted as points, and such an image is helpful in enabling us to see patterns in the data and to recognize that there may be a correlation between the two variables.
4. a. 1
   b. 0
   c. 0
   d. -1

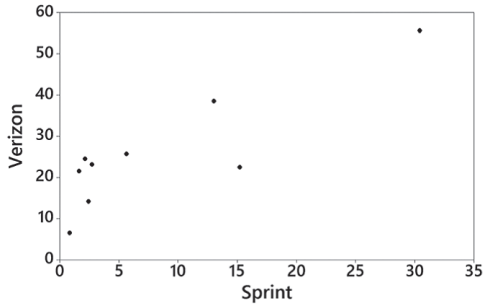
Copyright © 2022 Pearson Education, Inc.

Elementary Statistics Using Excel 7th edition Triola Solutions Manual

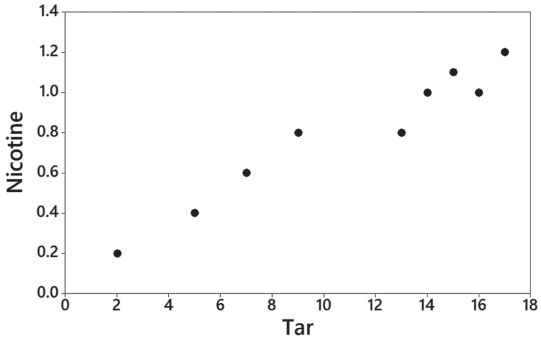
5. There does not appear to be a correlation. The given data suggest that five-day predicted high temperatures are not very accurate.



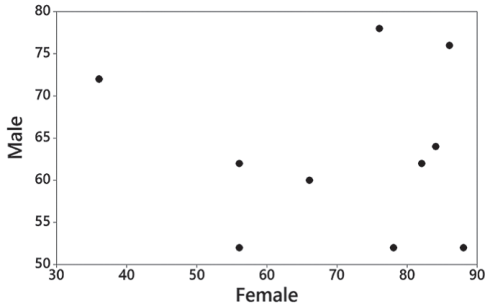
6. There does appear to be a linear correlation between the Sprint and Verizon airport data speeds. The correlation suggests that the speeds of the two companies are somehow related to the airports. One possibility is that the Sprint and Verizon cell towers are in similar locations at these airports.



7. There does appear to be a linear correlation between the amounts of tar and the amounts of nicotine in cigarettes.



8. There does not appear to be a correlation between pulse rates of females and males. The major flaw with this exercise is that the data are not paired as required. The results are therefore meaningless.



Copyright © 2022 Pearson Education, Inc.

Elementary Statistics Using Excel 7th edition Triola Solutions Manual

9. With $n = 10$ pairs of data, the critical values are ± 0.632. Because $r = -0.475$ is between –0.632 and 0.632, there is not sufficient evidence to conclude that there is a linear correlation.
10. With $n = 9$ pairs of data, the critical values are ± 0.666. Because $r = 0.866$ is in the right tail region beyond 0.666, there is sufficient evidence to conclude that there is a linear correlation.
11. With $n = 9$ pairs of data, the critical values are ± 0.666. Because $r = 0.971$ is in the right tail region beyond 0.666, there is sufficient evidence to conclude that there is a linear correlation.
12. With $n = 10$ pairs of data, the critical values are ± 0.632. Because $r = -0.076$ is between –0.632 and 0.632, there is not sufficient evidence to conclude that there is a linear correlation. The data are not paired, so the results are meaningless.
13. Because the *P*-value is 0.166, which is not small (such as 0.05 or less), there is a high chance (16.6%) of getting the sample results when there is no correlation, so there is not sufficient evidence to conclude that there is linear correlation.
14. Because the *P*-value is 0.003, which is small (such as 0.05 or less), there is a small chance (0.3%) of getting the sample results when there is no correlation, so there is sufficient evidence to conclude that there is linear correlation.
15. Because the *P*-value of 0.000 is small (such as 0.05 or less), there is a small chance of getting the sample results when there is no correlation, so there is sufficient evidence to conclude that there is a linear correlation.
16. Because the *P*-value of 0.835 is not small (such as 0.05 or less), there is a high chance (83.5% chance) of getting the sample results when there is no correlation, so there is not sufficient evidence to conclude that there is a linear correlation. But the data are not paired, so the *P*-value is meaningless here.

Quick Quiz

1. Class width: 20. It is not possible to identify the original data values.
2. Class limits: 0 and 19.
   Class boundaries: –0.5 and 19.5.
3. 69
4.

Annual Tornadoes in Oklahoma	Relative Frequency
0–19	4.3%
20–39	26.1%
40–59	30.4%
60–79	21.7%
80–99	8.7%
100–119	7.2%
120–139	0.0%
140–159	1.4%

5. 30, 30, 30, 31, 34, 34, 36, 36, 39
6. pareto chart
7. scatterplot
8. No, the term “normal distribution” has a different meaning than the term “normal” that is used in ordinary speech. A normal distribution has a bell shape, but the randomly selected lottery digits will have a uniform or flat shape.
9. variation
10. Parts a, d, and e describe normally distributed data.

Copyright © 2022 Pearson Education, Inc.

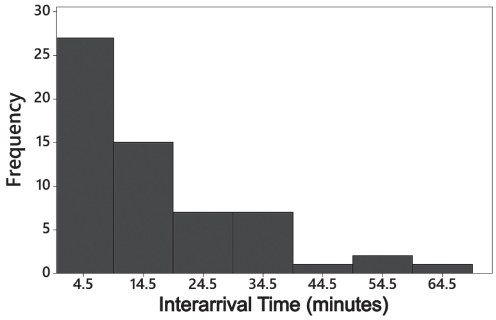
Elementary Statistics Using Excel 7th edition Triola Solutions Manual

Review Exercises

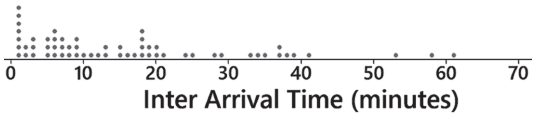
1. Both distributions are skewed to the right.

Email Interarrival Time (minutes)	Frequency
0–9	27
10–19	15
20–29	7
30–39	7
40–49	1
50–59	2
60–69	1

2. Because the histogram has a shape that is far from being bell-shaped, it suggests that the data are from a population *not* having a normal distribution.



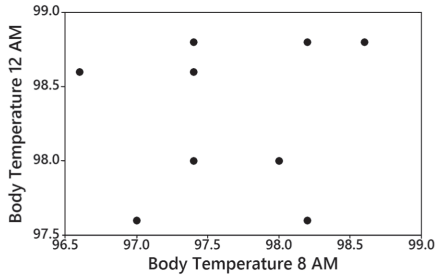
3. By using fewer classes, the histogram does a better job of illustrating the distribution.



4. There are no outliers.

```
0 | 111111122333555666677788999
1 | 012335567888899
2 | 0014589
3 | 3457789
4 | 1
5 | 38
6 | 1
```

5. No. There is no pattern suggesting that there is a relationship.

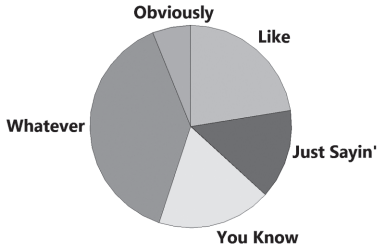


6. a. time-series graph
   b. scatterplot
   c. pareto chart

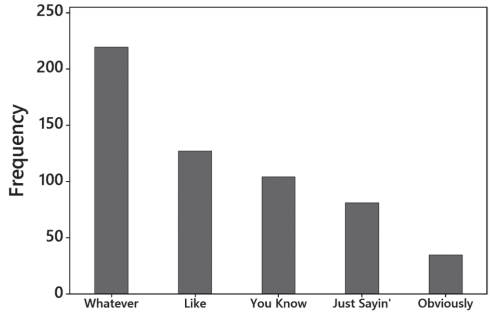
Copyright © 2022 Pearson Education, Inc.

Elementary Statistics Using Excel 7th edition Triola Solutions Manual

7. A pie chart wastes ink on components that are not data; pie charts lack an appropriate scale; pie charts don't show relative sizes of different components as well as some other graphs, such as a Pareto chart.



8. The Pareto chart does a better job. It draws attention to the most annoying words or phrases and shows the relative sizes of the different categories.



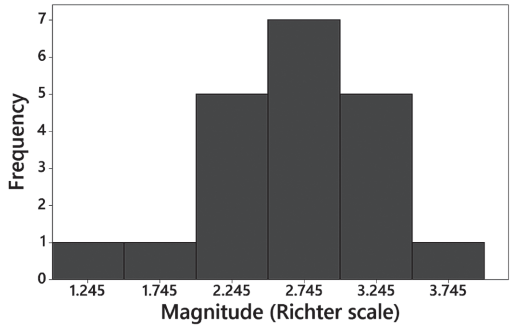
Cumulative Review Exercises

1.

Magnitude	Frequency
1.00–1.49	1
1.50–1.99	1
2.00–2.49	5
2.50–2.99	7
3.00–3.49	5
3.50–3.99	1

2. a. 1.00 and 1.49
   b. 0.995 and 1.495
   c. $\frac{1.00+1.49}{2}=1.245$

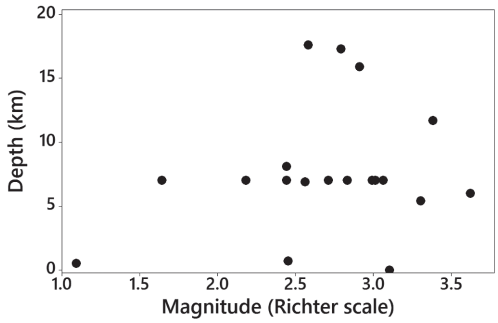
3. The distribution is closer to being a normal distribution than the others. Only the single lowest value of 1.09 prevents perfect symmetry, but that one value should not be a basis for stating that the distribution is skewed left.



Copyright © 2022 Pearson Education, Inc.

Elementary Statistics Using Excel 7th edition Triola Solutions Manual

4. a. continuous
   b. quantitative
   c. ratio
   d. sample
5. The scatterplot does not show any pattern. There does not appear to be correlation between magnitude and depth.



Copyright © 2022 Pearson Education, Inc.

Elementary Statistics Using Excel 7th edition Triola Solutions Manual