SOUTH-WESTERN
CENGAGE Learning

Chapter 3
Descriptive Statistics: Numerical Measures

Learning Objectives

1. Understand the purpose of measures of location.

2. Be able to compute the mean, median, mode, quartiles, and various percentiles.

3. Understand the purpose of measures of variability.

4. Be able to compute the range, interquartile range, variance, standard deviation, and coefficient of variation.

5. Understand skewness as a measure of the shape of a data distribution. Learn how to recognize when a data distribution is negatively skewed, roughly symmetric, and positively skewed.

6. Understand how z scores are computed and how they are used as a measure of relative location of a data value.

7. Know how Chebyshev's theorem and the empirical rule can be used to determine the percentage of the data within a specified number of standard deviations from the mean.

8. Learn how to construct a 5-number summary and a box plot.

9. Be able to compute and interpret covariance and correlation as measures of association between two variables.

10. Be able to compute a weighted mean.

Chapter 3
Solutions:

1. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{75}{5} = 15$

10, 12, 16, 17, 20

Median = 16 (middle value)

2. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{96}{6} = 16$

10, 12, 16, 17, 20, 21

Median $= \frac{16+17}{2} = 16.5$

3. 15, 20, 25, 25, 27, 28, 30, 34

$i = \frac{20}{100}(8) = 1.6$ 2nd position = 20

$i = \frac{25}{100}(8) = 2$ $\frac{20+25}{2} = 22.5$

$i = \frac{65}{100}(8) = 5.2$ 6th position = 28

$i = \frac{75}{100}(8) = 6$ $\frac{28+30}{2} = 29$

4. Mean $= \frac{\Sigma x_i}{n} = \frac{657}{11} = 59.727$

Median = 57 6th item

Mode = 53 It appears 3 times

5. a. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{3181}{20} = \$159$

b. Median 10th \$160 Los Angeles
11th \$162 Seattle

Median $= \frac{160+162}{2} = \$161$

c. Mode = \$167 San Francisco and New Orleans

d. $i = \left(\frac{25}{100}\right)20 = 5$

5th $134
6th $139

$$Q_1 = \frac{134+139}{2} = \$136.50$$

e. $i = \left(\frac{75}{100}\right)20 = 15$

15th $167
16th $173

$$Q_3 = \frac{167+173}{2} = \$170$$

6. a. Marketing Majors

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{363}{10} = 36.3$$

Data in order

28.4 30.6 34.2 34.2 35.2 35.8 37.7 39.5 42.4 45.0

Median (5^{th} and 6^{th}) = $\frac{35.2+35.8}{2} = 35.5$

Mode – 34.2 (2 times)

Accounting Majors

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{731.2}{16} = 45.7$$

Data in order

33.5 38.0 40.2 40.8 41.1 41.7 43.5 44.2 45.2 47.8 49.1 49.7 49.9 53.9 55.5 57.1

Median (8^{th} and 9^{th}) = $\frac{44.2+45.2}{2} = 44.7$

There is no mode – all values occur one time.

b. Marketing

Q_1: i = .25(10) = 2.5

Round to 3^{rd} position $Q_1 = 34.2$
Q_3: i = .75(10) = 7.5

Round to 8th position $Q_3 = 39.5$

Accounting

Q_1: $i = .25(16) = 4$

Use 4th and 5th positions

$$Q_1 = \frac{40.8 + 41.1}{2} = 40.95$$

Q_3: $i = .75(16) = 12$

Use 12th and 13th positions

$$Q_3 = \frac{49.7 + 49.9}{2} = 49.8$$

c. The difference between the sample means is 45.7 – 36.3 = 9.4. The difference between the sample medians is 44.7 – 35.5 = 9.2. On the basis of both of these measures of central location, accounting majors have an average salary that is slightly over $9,000 more per year than marketing majors. The highest accounting major salary was $57,1000, while the highest marketing major salary was $45,000. The top 25% of accounting majors made $49,800 or more per year. The top 25% of marketing majors made $39,500 or more per year.

7. a. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{871.74}{24} = 36.32$

Median is average of 10th and 11th values after arranging in ascending order.

$$\text{Median} = \frac{39.00 + 39.95}{2} = 39.48$$

Data are multimodal

b. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{491.14}{24} = 20.46$

$$\text{Median} = \frac{19.75 + 19.95}{2} = 19.85$$

Data are bimodal: 19.95 (3 brokers), 29.95 (3 brokers)

c. Comparing the measures of central location, we conclude that it costs more to trade 100 shares in a broker assisted trade than 500 shares online.

d. From the data we have here it is more related to whether the trade is broker-assisted or online. The amount of the online transaction is 5 times as great but the cost of the transaction is less.

However, if the comparison was restricted to broker-assisted or online trades, we would probably find that larger transactions cost more.

8. a. $\bar{x} = \frac{\sum x_i}{n} = \frac{3200}{20} = 160$

Order the data from low 100 to high 360

Median $i = \left(\frac{50}{100}\right)20 = 10$ Use 10th and 11th positions

Median = $\left(\frac{130+140}{2}\right) = 135$

Mode = 120 (occurs 3 times)

b. $i = \left(\frac{25}{100}\right)20 = 5$ Use 5th and 6th positions

$Q_1 = \left(\frac{115+115}{2}\right) = 115$

$i = \left(\frac{75}{100}\right)20 = 15$ Use 15th and 16th positions

$Q_3 = \left(\frac{180+195}{2}\right) = 187.5$

c. $i = \left(\frac{90}{100}\right)20 = 18$ Use 18th and 19th positions

90th percentile $= \left(\frac{235+255}{2}\right) = 245$

90% of the tax returns cost $245 or less. 10% of the tax returns cost $245 or more.

9. a. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{6330}{15} = 422$ minutes

b. Median $i = \left(\frac{50}{100}\right)15 = 7.5$

8th position 380 minutes

c. 85th percentile

$i = \left(\frac{85}{100}\right)15 = 12.75$

13th position 690 minutes

d. Using the mean $\bar{x} = 422$, cell-phone subscribers are using 422/750 = 56% of the capacity of their plans. Part (c) shows 85% of the subscribers are using 690 minutes or less. In general, cell-phone users are not coming close to using the 750 minute capacity of their plans.

10. a. Minimum = .4%; Maximum = 3.5%

b. $\Sigma x_i = 69$

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{69}{30} = 2.3\%$$

Median is average of 15th and 16th items.

Both are 2.5%, so the median is 2.5%.

The mode is 2.7%; forecast by 4 economists.

c. For Q_1,

$$i = \left(\frac{25}{100}\right)30 = 7.5; \text{ round up and use the } 8^{\text{th}} \text{ item}$$

$$Q_1 = 2.0\%$$

For Q_3,

$$i = \left(\frac{75}{100}\right)30 = 22.5; \text{ round up and use the 23rd item}$$

$$Q_3 = 2.8\%$$

d. Generally, the 2% to 3% growth should be considered optimistic.

11. Using the mean we get $\bar{x}_{\text{city}} = 15.58$, $\bar{x}_{\text{highway}} = 18.92$

For the samples we see that the mean mileage is better on the highway than in the city.

City

13.2 14.4 15.2 15.3 15.3 15.3 15.9 16 16.1 16.2 16.2 16.7 16.8

↑ Median (15.9)

Mode: 15.3

Highway

17.2 17.4 18.3 18.5 18.6 18.6 18.7 19.0 19.2 19.4 19.4 20.6 21.1

↑ Median (18.7)

Mode: 18.6, 19.4

The median and modal mileages are also better on the highway than in the city.

12. Disney

Total Revenue: $3,321 million (13 movies)

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{3321}{13} = \$255.5$$

104 110 136 169 249 250 253 273 304 325 346 354 448

Median 7th position

Median = $253

Q_1: i = .25(13) = 3.25 4th position

Q_1 = $169

Q_3: i = .75(13) = 9.75 10th position

Q_3 = $325

Pixar

Total Revenue: $3,231 million (6 movies)

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{3231}{6} = \$538.5$$

362 363 485 525 631 865

Median (3rd and 4th positions)

$$\text{Median} = \frac{485+525}{2} = \$505$$

Q_1: i = .25(6) = 1.5 2nd position

Q_1 = $363

Q_3: i = .75(6) = 4.5 5th position

Q_3 = $631

The total box office revenues for the two companies over the 10 year period are approximately the same: Disney $3321 million; Pixar $3231 million. But Disney generated its revenue with 13 films while Pixar generated its revenue with only 6 films.

	Disney	Pixar
Mean	$225.5	$538.5
Median	$253	$505

The first quartiles show 75% of Disney films do better than $169 million while 75% of Pixar films do better than $363 million. The third quartiles show 25% of Disney films do better than $325 million while 25% of Pixar films do better than $631. In all of these comparisons, Pixar films are about twice as successful as Disney films when it comes to box office revenue. In buying Pixar, Disney looks to acquire Pixar's ability to make higher revenue films.

13. Range 20 - 10 = 10

10, 12, 16, 17, 20

$$i = \frac{25}{100}(5) = 1.25$$

Q_1 (2nd position) = 12

$$i = \frac{75}{100}(5) = 3.75$$

Q_3 (4th position) = 17

IQR = Q_3 - Q_1 = 17 - 12 = 5

14. $$\bar{x} = \frac{\Sigma x_i}{n} = \frac{75}{5} = 15$$

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1} = \frac{64}{4} = 16$$

$$s = \sqrt{16} = 4$$

15. 15, 20, 25, 25, 27, 28, 30, 34 Range = 34 - 15 = 19

$$i = \frac{25}{100}(8) = 2 \qquad Q_1 = \frac{20+25}{2} = 22.5$$

$$i = \frac{75}{100}(8) = 6 \qquad Q_3 = \frac{28+30}{2} = 29$$

IQR = Q_3 - Q_1 = 29 - 22.5 = 6.5

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{204}{8} = 25.5$$

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1} = \frac{242}{7} = 34.57$$

$$s = \sqrt{34.57} = 5.88$$

16. a. Range = 190 - 168 = 22

b. $\Sigma(x_i - \bar{x})^2 = 376$

$$s^2 = \frac{376}{5} = 75.2$$

c. $s = \sqrt{75.2} = 8.67$

d. Coefficient of Variation $= \left(\frac{8.67}{178}\right)100\% = 4.87\%$

17. a. With DVD $\bar{x} = \frac{\Sigma x_i}{n} = \frac{2050}{5} = 410$

Without DVD $\bar{x} = \frac{\Sigma x_i}{n} = \frac{1550}{5} = 310$

With DVD $410 - $310 = $100 more expensive

b. With DVD Range = 500 - 300 = 200

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1} = \frac{22000}{4} = 5500$$

$$s = \sqrt{5500} = 74.2$$

Without DVD Range = 360 - 290 = 70

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1} = \frac{3200}{4} = 800$$

$$s = \sqrt{800} = 28.3$$

Models with DVD players have the greater variation in prices. The price range is $300 to $500. Models without a DVD player have less variation in prices. The price range is $290 to $360.

18. a. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{266}{7} = \$38$ per day

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1} = \frac{582}{6} = 97$$

$$s = \sqrt{97} = \$9.85$$

b. The mean car-rental rate per day is $38 for both Eastern and Western cities. However, Eastern cities show a greater variation in rates per day. This greater variation is most likely due to the inclusion of the most expensive city (New York) in the Eastern city sample.

19. a. Range = 60 - 28 = 32

IQR = Q_3 - Q_1 = 55 - 45 = 10

b. $\bar{x} = \frac{435}{9} = 48.33$

$\Sigma(x_i - \bar{x})^2 = 742$

$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1} = \frac{742}{8} = 92.75$

$s = \sqrt{92.75} = 9.63$

c. The average air quality is about the same. But, the variability is greater in Anaheim.

20. Dawson Supply: Range = 11 - 9 = 2

$s = \sqrt{\frac{4.1}{9}} = 0.67$

J.C. Clark: Range = 15 - 7 = 8

$s = \sqrt{\frac{60.1}{9}} = 2.58$

21. a. Cities:

$\bar{x} = \frac{\Sigma x_i}{n} = \frac{198}{6} = \$33$

$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1} = \frac{72}{6-1} = 14.40$

$s = \sqrt{14.40} = 3.79$

Retirement Areas:

$\bar{x} = \frac{\Sigma x_i}{n} = \frac{192}{6} = \$32$

$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1} = \frac{18}{6-1} = 3.60$

$s = \sqrt{3.60} = 1.90$

b. Mean cost of the market basket is roughly the same with the retirement areas sample mean $1 less. However, there is more variation in the cost in cities than in retirement areas.

22. a. Freshmen $\bar{x} = \frac{\Sigma x_i}{n} = \frac{32125}{25} = 1285$

Seniors $\bar{x} = \frac{\Sigma x_i}{n} = \frac{8660}{20} = \$433$

Freshmen spend almost three times as much on back-to-school items as seniors.

b. Freshmen Range = 2094 – 374 = 1720

Seniors Range = 632 – 280 = 352

c. Freshmen

$i = \left(\frac{25}{100}\right)25 = 6.25$ $Q_1 = 1079$ (7th item)

$i = \left(\frac{75}{100}\right)25 = 18.75$ $Q_3 = 1475$ (19th item)

IQR = Q_3 - Q_1 = 1479 – 1075 = 404

Seniors

$i = \left(\frac{25}{100}\right)20 = 5$ $Q_1 = \frac{368 + 373}{2} = 370.5$

$i = \left(\frac{75}{100}\right)20 = 15$ $Q_1 = \frac{489 + 515}{2} = 502$

IQR = Q_3 - Q_1 = 502 – 370.5 = 131.5

d. $s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}}$

Freshmen $s = \sqrt{\frac{3233186}{24}} = 367.04$

Seniors $s = \sqrt{\frac{178610}{19}} = 96.96$

e. All measures of variability show freshmen have more variation in back-to-school expenditures.

23. a. For 2005

$\bar{x} = \frac{\Sigma x_i}{n} = \frac{608}{8} = 76$

$s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{30}{7}} = 2.07$

For 2006

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{608}{8} = 76$$

$$s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{194}{7}} = 5.26$$

b. The mean score is 76 for both years, but there is an increase in the standard deviation for the scores in 2006. The golfer is not as consistent in 2006 and shows a sizeable increase in the variation with golf scores ranging from 71 to 85. The increase in variation might be explained by the golfer trying to change or modify the golf swing. In general, a loss of consistency and an increase in the standard deviation could be viewed as a poorer performance in 2006. The optimism in 2006 is that three of the eight scores were better than any score reported for 2005. If the golfer can work for consistency, eliminate the high score rounds, and reduce the standard deviation, golf scores should show improvement.

24. Quarter milers

$s = 0.0564$

Coefficient of Variation = $(s/\bar{x})100\% = (0.0564/0.966)100\% = 5.8\%$

Milers

$s = 0.1295$

Coefficient of Variation = $(s/\bar{x})100\% = (0.1295/4.534)100\% = 2.9\%$

Yes; the coefficient of variation shows that as a percentage of the mean the quarter milers' times show more variability.

25. $$\bar{x} = \frac{\Sigma x_i}{n} = \frac{75}{5} = 15$$

$$s^2 = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{64}{4}} = 4$$

10 $\quad z = \frac{10-15}{4} = -1.25$

20 $\quad z = \frac{20-15}{4} = +1.25$

12 $\quad z = \frac{12-15}{4} = -.75$

17 $\quad z = \frac{17-15}{4} = +.50$

16 $\quad z = \frac{16-15}{4} = +.25$

26. $z = \frac{520-500}{100} = +.20$

$z = \frac{650-500}{100} = +1.50$

$z = \frac{500-500}{100} = 0.00$

$z = \frac{450-500}{100} = -.50$

$z = \frac{280-500}{100} = -2.20$

27. a. $z = \frac{40-30}{5} = 2$ $1-\frac{1}{2^2} = 0.75$ At least 75%

b. $z = \frac{45-30}{5} = 3$ $1-\frac{1}{3^2} = 0.89$ At least 89%

c. $z = \frac{38-30}{5} = 1.6$ $1-\frac{1}{1.6^2} = 0.61$ At least 61%

d. $z = \frac{42-30}{5} = 2.4$ $1-\frac{1}{2.4^2} = 0.83$ At least 83%

e. $z = \frac{48-30}{5} = 3.6$ $1-\frac{1}{3.6^2} = 0.92$ At least 92%

28. a. Approximately 95%

b. Almost all

c. Approximately 68%

29. a. This is from 2 standard deviations below the mean to 2 standard deviations above the mean. With $z = 2$, Chebyshev's theorem gives:

$$1-\frac{1}{z^2} = 1-\frac{1}{2^2} = 1-\frac{1}{4} = \frac{3}{4}$$

Therefore, at least 75% of adults sleep between 4.5 and 9.3 hours per day.

b. This is from 2.5 standard deviations below the mean to 2.5 standard deviations above the mean. With $z = 2.5$, Chebyshev's theorem gives:

$$1-\frac{1}{z^2} = 1-\frac{1}{2.5^2} = 1-\frac{1}{6.25} = .84$$

Therefore, at least 84% of adults sleep between 3.9 and 9.9 hours per day.

c. With z = 2, the empirical rule suggests that 95% of adults sleep between 4.5and 9.3 hours per day. The percentage obtained using the empirical rule is greater than the percentage obtained using Chebyshev's theorem.

30. a. $2.20 is one standard deviation below the mean and $2.40 is one standard deviation above the mean. The empirical rule says that approximately 68% of gasoline sales are in the price range.

b. Part (a) shows that approximately 68% of the gasoline sales are between $2.20 and $2.40. Since the bell-shaped distribution is symmetric, approximately half of 68%, or 34%, of the gasoline sales should be between $2.20 and the mean price of $2.30. $2.50 is two standard deviations above the mean price of $2.30. The empirical rule says that approximately 95% of the gasoline sales should be within two standard deviations of the mean. Thus, approximately half of 95%, or 47.5%, of the gasoline sales should be between the mean price of $2.30 and $2.50. The percentage of gasoline sales between $2.20 and $2.50 should be approximately 34% + 47.5% = 81.5%.

c. $2.50 is two standard deviations above the mean and the empirical rule says that approximately 95% of the gasoline sales should be within two standard deviations of the mean. Thus, 1 - 95% = 5% of the gasoline sales should be more than two standard deviations from the mean. Since the bell-shaped distribution is symmetric, we expected half of 5%, or 2.5%, would be more than $2.50.

31. a. 607 is one standard deviation above the mean. Approximately 68% of the scores are between 407 and 607 with half of 68%, or 34%, of the scores between the mean of 507 and 607. Also, since the distribution is symmetric, 50% of the scores are above the mean of 507. With 50% of the scores above 507 and with 34% of the scores between 507 and 607, 50% - 34% = 16% of the scores are above 607.

b. 707 is two standard deviations above the mean. Approximately 95% of the scores are between 307 and 707 with half of 95%, or 47.5%, of the scores between the mean of 507 and 707. Also, since the distribution is symmetric, 50% of the scores are above the mean of 507. With 50% of the scores above 507 and with 47.5% of the scores between 507 and 707, 50%- 47.5% = 2.5% of the scores are above 707.

c. Approximately 68% of the scores are between 407 and 607 with half of 68%, or 34%, of the scores between 407 and the mean of 507.

d. Approximately 95% of the scores are between 307 and 707 with half of 95%, or 47.5%, of the scores between 307 and the mean of 507. Approximately 68% of the scores are between 407 and 607 with half of 68%, or 34%, of the scores between the mean of 507 and 607. Thus, 47.5% + 34% = 81.5% of the scores are between 307 and 607.

32. a. $$z = \frac{x-\mu}{\sigma} = \frac{2300-3100}{1200} = -.67$$

b. $$z = \frac{x-\mu}{\sigma} = \frac{4900-3100}{1200} = 1.50$$

c. $2300 is .67 standard deviations below the mean. $4900 is 1.50 standard deviations above the mean. Neither is an outlier.

d. $$z = \frac{x-\mu}{\sigma} = \frac{13000-3100}{1200} = 8.25$$

$13,000 is 8.25 standard deviations above the mean. This cost is an outlier.

33. a. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{64}{7} = 9.14 \text{ days}$

Median: with $n = 7$, use 4th position

2, 3, 8, 8, 12, 13, 18

Median = 8 days

Mode: 8 days (occurred twice)

b. Range = Largest value – Smallest value
= 18 – 2 = 16

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

$$\Sigma(x_i - \bar{x})^2 = (13-9.14)^2 + (12-9.14)^2 + (8-9.14)^2 + (3-9.14)^2 + (8-9.14)^2 + (2-9.14)^2 + (18-9.14)^2 = 192.86$$

$$s = \sqrt{\frac{192.86}{6}} = 5.67$$

c. $z = \frac{x - \bar{x}}{s} = \frac{18-9.14}{5.67} = 1.56$

The 18 days required to restore service after hurricane Wilma is not an outlier.

d. Yes, FP&L should consider ways to improve its emergency repair procedures. The mean, median and mode show repairs requiring an average of 8 to 9 days can be expected if similar hurricanes are encountered in the future. The 18 days required to restore service after hurricane Wilma should not be considered unusual if FP&L continues to use its current emergency repair procedures. With the number of customers affected running into the millions, plans to shorten the number of days to restore service should be undertaken by the company.

34. a. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{765}{10} = 76.5$

$$s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{442.5}{10-1}} = 7$$

b. $z = \frac{x - \bar{x}}{s} = \frac{84-76.5}{7} = 1.07$

Approximately one standard deviation above the mean. Approximately 68% of the scores are within one standard deviation. Thus, half of (100-68), or 16%, of the games should have a winning score of 84 or more points.

$$z = \frac{x - \bar{x}}{s} = \frac{90 - 76.5}{7} = 1.93$$

Approximately two standard deviations above the mean. Approximately 95% of the scores are within two standard deviations. Thus, half of (100-95), or 2.5%, of the games should have a winning score of more than 90 points.

c. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{122}{10} = 12.2$

$$s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{559.6}{10-1}} = 7.89$$

Largest margin 24: $z = \frac{x - \bar{x}}{s} = \frac{24 - 12.2}{7.89} = 1.50$. No outliers.

35. a. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{79.86}{20} = 3.99$

Median $= \frac{4.17 + 4.20}{2} = 4.185$ (average of 10th and 11th values)

b. $Q_1 = 4.00$ (average of 5th and 6th values)

$Q_3 = 4.50$ (average of 15th and 16th values)

c. $s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{12.51}{19}} = 0.81$

d. The distribution is significantly skewed to the left.

e. Allison One: $z = \frac{4.12 - 3.99}{0.81} = 0.16$

Omni Audio SA 12.3: $z = \frac{2.32 - 3.99}{0.81} = -2.06$

f. The lowest rating is for the Bose 501 Series. Its *z*-score is:

$$z = \frac{2.14 - 3.99}{0.81} = -2.28$$

This is not an outlier so there are no outliers.

36. 15, 20, 25, 25, 27, 28, 30, 34

Smallest = 15

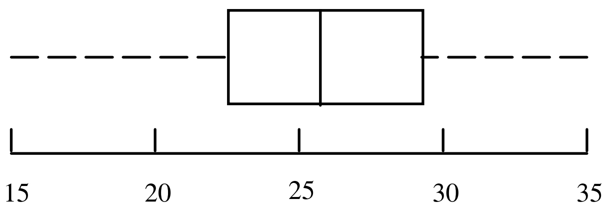
$i = \frac{25}{100}(8) = 2$ $\qquad Q_1 = \frac{20 + 25}{2} = 22.5$

$$\text{Median} = \frac{25+27}{2} = 26$$

$$i = \frac{75}{100}(8) = 8 \qquad Q_3 = \frac{28+30}{2} = 29$$

Largest = 34

37.



38. 5, 6, 8, 10, 10, 12, 15, 16, 18

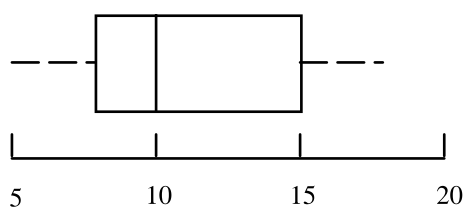
Smallest = 5

$i = \frac{25}{100}(9) = 2.25$ Q_1 = 8 (3rd position)

Median = 10

$i = \frac{75}{100}(9) = 6.75$ Q_3 = 15 (7th position)

Largest = 18



39. IQR = 50 - 42 = 8

Lower Limit: Q_1 - 1.5 IQR = 42 - 12 = 30

Upper Limit: Q_3 + 1.5 IQR = 50 + 12 = 62

65 is an outlier

40. a. Smallest 619

Q_1: $i = \left(\frac{25}{100}\right)n = .25(22) = 5.5$

6th position: Q_1 = 725

Median: 1/2 way between 11th and 12th

$$\text{Median} = \frac{912+1120}{2} = 1016$$

Q_3: $i = \left(\frac{75}{100}\right)n = .75(22) = 16.5$

17th position: Q_3 = 1699

Largest 4450

619, 725, 1016, 1699, 4450

b. IQR = 1699 - 725 = 974

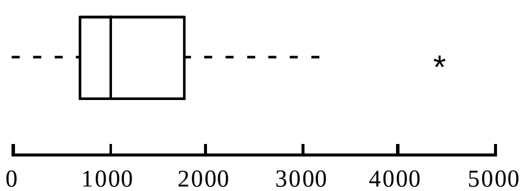
Lower Limit = Q_1 - 1.5(IQR)
= 725 - 1.5(975) = -736

Use Lower Limit = 0

Upper Limit = Q_3 + 1.5(IQR)
= 1699 + 1.5(975) = 3160

c. Yes; larger than upper limit of $3,160,000

d. No; less than upper limit of $3,160,000



41. a. Median (11th position) 4019

$$i = \frac{25}{100}(21) = 5.25$$

Q_1 (6th position) = 1872

$$i = \frac{75}{100}(21) = 15.75$$

Q_3 (16th position) = 8305

608, 1872, 4019, 8305, 14138

b. Limits:

IQR = Q_3 - Q_1 = 8305 - 1872 = 6433

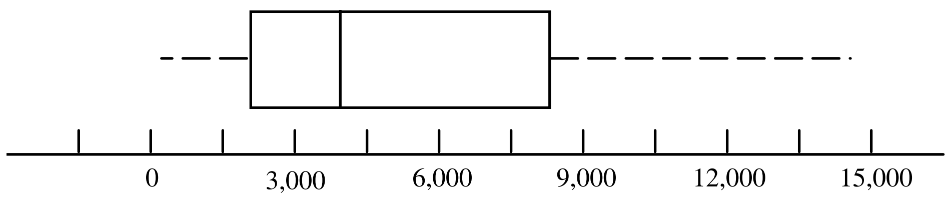
Lower Limit: Q_1 - 1.5 (IQR) = -7777

Upper Limit: $Q_3 + 1.5\,(\text{IQR}) = 17955$

c. There are no outliers, all data are within the limits.

d. Yes, if the first two digits in Johnson and Johnson's sales were transposed to 41,138, sales would have shown up as an outlier. A review of the data would have enabled the correction of the data.

e.



42. a. Median $n = 30$; 15th and 16th positions

$$\text{Median} = \frac{63+69}{2} = 66$$

b. Smallest 30

Q_1: $i = \left(\frac{25}{100}\right)30 = 7.5$

8th position: $Q_1 = 49$

Q_3: $i = \left(\frac{75}{100}\right)30 = 22.5$

23rd position: $Q_3 = 88$

Largest 208

30, 49, 66, 88, 208

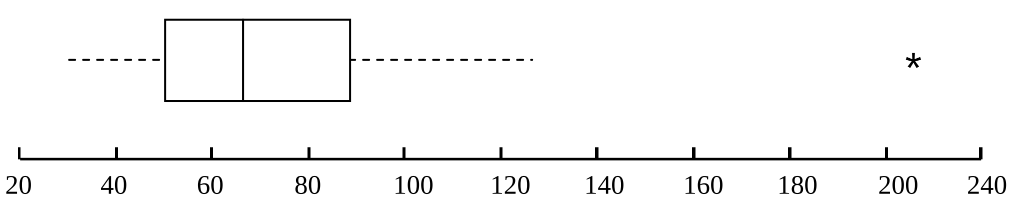
c. $\text{IQR} = Q_3 - Q_1 = 88 - 49 = 39$

Upper Limit $= Q_3 + 1.5(\text{IQR})$

$= 88 + 1.5(39) = 146.5$

New York Yankees payroll is an outlier.

d.



43. a. Median n = 14; 7th and 8th positions

$$\text{Median} = \frac{2.1+3.5}{2} = 2.8 \text{ millions}$$

b. Smallest .8

Q_1: $i = \left(\frac{.25}{100}\right)14 = 3.5$

4th position: Q_1 = 1.0

Q_3: $i = \left(\frac{75}{100}\right)14 = 10.5$

11th position: Q_3 = 4.3

Largest 8.0

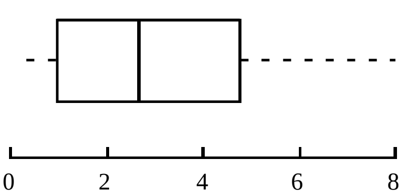
.8, 1.0, 2.8, 4.3, 8.0

c. IQR = Q_3 - Q_1 = 4.3 - 1.0 = 3.3

Upper Limit = Q_3 + 1.5(IQR)
= 4.3 + 1.5(3.3) = 9.25

Grasso's $8.5 million is larger than other executives, but it is not considered an outlier. Merrill Lynch's $7.7 million and Wells Fargo's $8.0 million show other executives close to Grasso's salary plus bonus.

d.



44. a. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{837.5}{46} = 18.2$

Median 23rd position 15.1
24th position 15.6

$$\text{Median} = \frac{15.1+15.6}{2} = 15.35$$

b. Q_1: $i = \left(\frac{25}{100}\right)n = .25(46) = 11.5$

12th position: Q_1 = 11.7

Q_3: $i = \left(\frac{75}{100}\right)n = .75(46) = 34.5$

35th position: $Q_3 = 23.5$

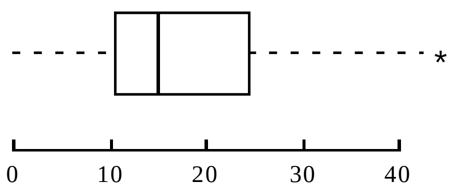
c. 3.4, 11.7, 15.35, 23.5, 41.3

d. IQR = 23.5 - 11.7 = 11.8

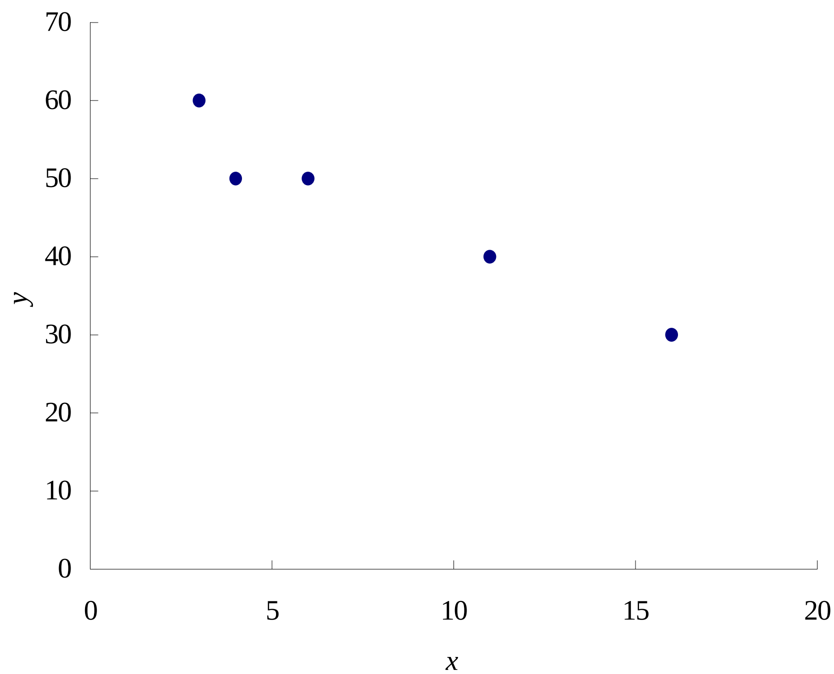
Lower Limit = Q_1 - 1.5(IQR)
= 11.7 - 1.5(11.8) = -6 Use 0

Upper Limit = Q_3 + 1.5(IQR)
= 23.5 + 1.5(11.8) = 41.2

Yes, one: Alger Small Cap 41.3



45. a.



b. Negative relationship

c/d. $\Sigma x_i = 40$ $\bar{x} = \frac{40}{5} = 8$ $\Sigma y_i = 230$ $\bar{y} = \frac{230}{5} = 46$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = -240$ $\Sigma(x_i - \bar{x})^2 = 118$ $\Sigma(y_i - \bar{y})^2 = 520$

$$s_{xy} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{n-1} = \frac{-240}{5-1} = -60$$

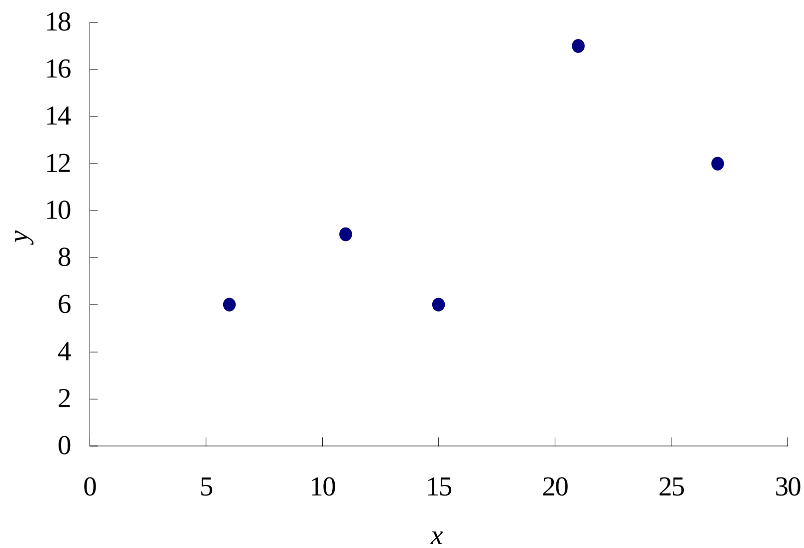
$$s_x = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{118}{5-1}} = 5.4314$$

$$s_y = \sqrt{\frac{\Sigma(y_i - \bar{y})^2}{n-1}} = \sqrt{\frac{520}{5-1}} = 11.4018$$

$$r_{xy} = \frac{s_{xy}}{s_x s_y} = \frac{-60}{(5.4314)(11.4018)} = -0.969$$

There is a strong negative linear relationship.

46. a.



b. Positive relationship

c/d. $\Sigma x_i = 80 \quad \bar{x} = \frac{80}{5} = 16 \quad \Sigma y_i = 50 \quad \bar{y} = \frac{50}{5} = 10$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 106 \quad \Sigma(x_i - \bar{x})^2 = 272 \quad \Sigma(y_i - \bar{y})^2 = 86$

$$s_{xy} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{n-1} = \frac{106}{5-1} = 26.5$$

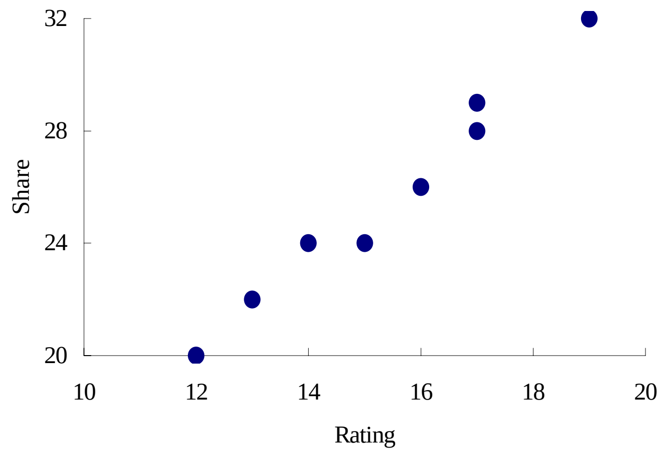
$$s_x = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{272}{5-1}} = 8.2462$$

$$s_y = \sqrt{\frac{\Sigma(y_i - \bar{y})^2}{n-1}} = \sqrt{\frac{86}{5-1}} = 4.6368$$

$$r_{xy} = \frac{s_{xy}}{s_x s_y} = \frac{26.5}{(8.2462)(4.6368)} = 0.693$$

A positive linear relationship

47. a.



b. Scatter diagram shows a positive relationship between rating and share. Higher shares are associated with higher ratings.

c. $$s_{xy} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{n-1} = \frac{80}{9-1} = 10$$

Sample covariance at +10 shows a positive relationship.

d. $$s_x = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{48}{9-1}} = 2.45$$

$$s_y = \sqrt{\frac{\Sigma(y_i - \bar{y})^2}{n-1}} = \sqrt{\frac{136}{9-1}} = 4.12$$

$$r_{xy} = \frac{s_{xy}}{s_x s_y} = \frac{10}{2.45(4.12)} = +.99$$

$r_{xy} = +.99$ shows a very strong positive relationship.

48. Let x = driving speed and y = mileage

$$\Sigma x_i = 420 \quad \bar{x} = \frac{420}{10} = 42 \quad \Sigma y_i = 270 \quad \bar{y} = \frac{270}{10} = 27$$

$$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = -475 \quad \Sigma(x_i - \bar{x})^2 = 1660 \quad \Sigma(y_i - \bar{y})^2 = 164$$

$$s_{xy} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{n-1} = \frac{-475}{10-1} = -52.7778$$

$$s_x = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{1660}{10-1}} = 13.5810$$

$$s_y = \sqrt{\frac{\Sigma(y_i - \bar{y})^2}{n-1}} = \sqrt{\frac{164}{10-1}} = 4.2687$$

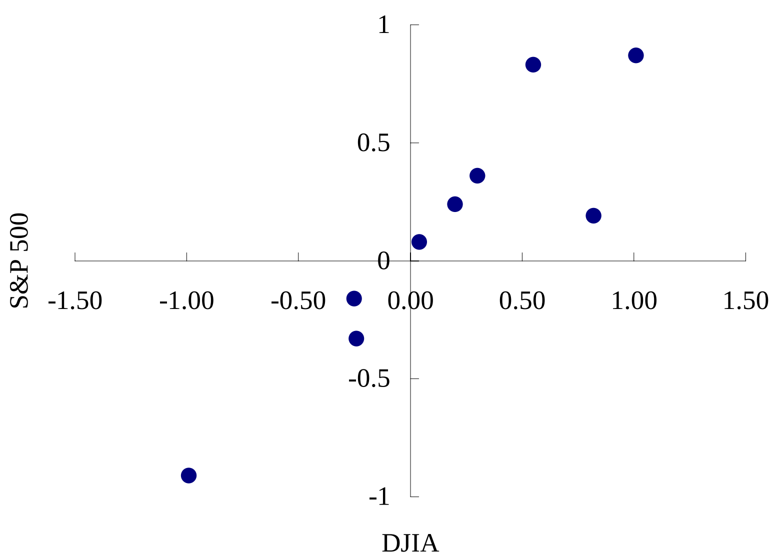
$$r_{xy} = \frac{s_{xy}}{s_x s_y} = \frac{-52.7778}{(13.5810)(4.2687)} = -.91$$

A strong negative linear relationship

49. a. The sample correlation coefficient is .78.

b. There is a positive linear relationship between the performance score and the overall rating.

50. a.



b. $$\bar{x} = \frac{\Sigma x_i}{n} = \frac{1.44}{9} = .16$$

$$\bar{y} = \frac{\Sigma x_i}{n} = \frac{1.17}{9} = .13$$

x_i	y_i	$(x_i - \bar{x})$	$(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
0.20	0.24	0.04	0.11	0.0016	0.0121	0.0044
0.82	0.19	0.66	0.06	0.4356	0.0036	0.0396
-0.99	-0.91	-1.15	-1.04	1.3225	1.0816	1.1960
0.04	0.08	-0.12	-0.05	0.0144	0.0025	0.0060
-0.24	-0.33	-0.40	-0.46	0.1600	0.2166	0.1840
1.01	0.87	0.85	0.74	0.7225	0.5476	0.6290
0.30	0.36	0.14	0.23	0.0196	0.0529	0.0322
0.55	0.83	0.39	0.70	0.1521	0.4900	0.2730
-0.25	-0.16	-0.41	-0.29	0.1681	0.0841	0.1189
			Total	2.9964	2.4860	2.4831

$$s_{xy} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{n-1} = \frac{2.4831}{8} = .3104$$

$$s_x = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{2.9964}{8}} = .6120$$

$$s_y = \sqrt{\frac{\Sigma(y_i - \bar{y})^2}{n-1}} = \sqrt{\frac{2.4860}{8}} = .5574$$

$$r_{xy} = \frac{s_{xy}}{s_x s_y} = \frac{.3104}{(.6120)(.5574)} = .9098$$

c. There is a strong positive linear association between DJIA and S&P 500. If you know the change in either, you will have a good idea of the stock market performance for the day.

51. a. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{420}{12} = 35$

b. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{288}{12} = 24$

c. $s_{xy} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{n-1} = \frac{3446}{12-1} = 313.27$

$$s_x = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{4058}{12-1}} = 19.21$$

$$s_y = \sqrt{\frac{\Sigma(y_i - \bar{y})^2}{n-1}} = \sqrt{\frac{3448}{12-1}} = 17.70$$

$$r_{xy} = \frac{s_{xy}}{s_x s_y} = \frac{313.27}{19.21(17.70)} = +.92$$

High positive correlation.

52. a. $\bar{x} = \frac{\Sigma w_i x_i}{\Sigma w_i} = \frac{6(3.2)+3(2)+2(2.5)+8(5)}{6+3+2+8} = \frac{70.2}{19} = 3.69$

b. $\frac{3.2+2+2.5+5}{4} = \frac{12.7}{4} = 3.175$

53.

f_i	M_i	$f_i M_i$
4	5	20
7	10	70
9	15	135
5	20	100
25		325

$\bar{x} = \frac{\Sigma f_i M_i}{n} = \frac{325}{25} = 13$

f_i	M_i	$M_i - \bar{x}$	$(M_i - \bar{x})^2$	$f_i(M_i - \bar{x})^2$
4	5	-8	64	256
7	10	-3	9	63
9	15	+2	4	36
5	20	+7	49	245
				600

$s^2 = \frac{\Sigma f_i (M_i - \bar{x})^2}{n-1} = \frac{600}{24} = 25$

$s = \sqrt{25} = 5$

54. a.

Grade x_i	Weight W_i
4 (A)	9
3 (B)	15
2 (C)	33
1 (D)	3
0 (F)	0
	60 Credit Hours

$\bar{x} = \frac{\Sigma w_i x_i}{\Sigma w_i} = \frac{9(4)+15(3)+33(2)+3(1)}{9+15+33+3} = \frac{150}{60} = 2.50$

b. Yes; satisfies the 2.5 grade point average requirement

55. a. $\bar{x} = \frac{\Sigma w_i x_i}{\Sigma w_i} = \frac{3000(15)+5500(14)+4200(12)+3000(25)+3000(20)+3800(12)+2500(35)}{3000+5500+4200+3000+3000+3800+2500}$

$= \frac{440{,}500}{25{,}000} = 17.62\%$

b. $\bar{x} = \frac{\Sigma w_i x_i}{\Sigma w_i} = \frac{3000(1.21) + 5500(1.48) + 4200(1.72) + 3000(0) + 3000(.96) + 3800(2.48) + 2500(0)}{3000 + 5500 + 4200 + 3000 + 3000 + 3800 + 2500}$

$= \frac{31,298}{25,000} = 1.25\%$

56. a.

Class	f_i	M_i	$f_i M_i$
0	15	0	0
1	10	1	10
2	40	2	80
3	85	3	255
4	350	4	1400
Totals	500		1745

$\bar{x} = \frac{\Sigma_i fM_i}{n} = \frac{1745}{500} = 3.49$

b.

$M_i - \bar{x}$	$(M_i - \bar{x})^2$	$f_i(M_i - \bar{x})^2$
-3.49	12.18	182.70
-2.49	6.20	62.00
-1.49	2.22	88.80
-0.49	0.24	20.41
+0.51	0.26	91.04
	Total	444.95

$s^2 = \frac{\Sigma(M_i - \bar{x})^2 f_i}{n-1} = \frac{444.95}{499} = 0.8917$ $\quad s = \sqrt{0.8917} = 0.9443$

57.

Price per Share	Frequency	Midpoint	$f_i M_i$
\$20-29	7	24.5	171.5
\$30-39	6	34.5	207.0
\$40-49	6	44.5	267.0
\$50-59	3	54.5	163.5
\$60-69	4	64.5	258.0
\$70-79	3	74.5	223.5
\$80-89	1	84.5	84.5
Total	30		1375.0

$\bar{x} = \frac{\Sigma_i fM_i}{n} = \frac{1375.0}{30} = 45.83$

Price per Share	Frequency	Midpoint	$M_i - \bar{x}$	$(M_i - \bar{x})^2$	$f_i(M_i - \bar{x})^2$
$20-29	7	24.5	-21.33	455.11	3185.7778
$30-39	6	34.5	-11.33	128.44	770.6667
$40-49	6	44.5	-1.33	1.78	10.6667
$50-59	3	54.5	8.67	75.11	225.3333
$60-69	4	64.5	18.67	348.44	1393.7778
$70-79	3	74.5	28.67	821.78	2465.3333
$80-89	1	84.5	38.67	1495.11	1495.1111
				Total	9546.6667

$$s = \sqrt{\frac{\Sigma f_i (M_i - \bar{x})^2}{n-1}} = \sqrt{\frac{9546.6667}{29}} = 18.14$$

58. a. Let x = media expenditures ($ millions) and y = shipments in barrels (millions)

$$\Sigma x_i = 404.1 \quad \bar{x} = \frac{404.1}{10} = 40.41 \quad \Sigma y_i = 119.9 \quad \bar{y} = \frac{119.9}{10} = 11.99$$

$$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 3763.481 \quad \Sigma(x_i - \bar{x})^2 = 19,248.469 \quad \Sigma(y_i - \bar{y})^2 = 939.349$$

$$s_{xy} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{n-1} = \frac{3763.481}{10-1} = 418.1646$$

A positive relationship

b. $$s_x = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{19,248.469}{10-1}} = 46.2463$$

$$s_y = \sqrt{\frac{\Sigma(y_i - \bar{y})^2}{n-1}} = \sqrt{\frac{939.349}{10-1}} = 10.2163$$

$$r_{xy} = \frac{s_{xy}}{s_x s_y} = \frac{418.1646}{(46.2463)(10.2163)} = 0.885$$

Note: The same value can also be obtained using Excel's CORREL function

59. a. $$\bar{x} = \frac{\Sigma x_i}{n} = \frac{100}{10} = \$10,000$$

Mean debt upon graduation is $10,000.

b. $$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1} = \frac{221.78}{9} = 24.64$$

$$s = \sqrt{24.64} = 4.96$$

60. a. Arrange the data in ascending order

48.8 92.6 111.0 …….. 958.0 995.9 2325.0

With n = 14, the median is the average of home prices in position 7 and 8.

Median home price = $\dfrac{212.9+218.9}{2} = 215.9$

Median home price = $215,900

b. $\dfrac{215,900-139,300}{139,300} = .55$

55% increase over the five-year period

c. $n = 14$

$i = \dfrac{25}{100}(n) = 3.5$

Use the 4th position

$Q_1 = 175.0$

$i = \dfrac{75}{100}(n) = 10.5$

Use the 11th position

$Q_3 = 628.3$

$Q_3 = \dfrac{362.5+628.3}{2} = 495.4$

d. Lowest price = 48.8 and highest price = 2324.0.

Five-number summary: 48.8, 175.0, 215.9, 628.3, 2325.0

e. IQR = Q_3 - Q_1 = 628.3 – 175.0 = 453.3

Upper limit = Q_3 + 1.5IQR = 628.3 + 1.5(679.95) = 1308.25

Any price over $1,308,250 is an outlier.

Yes, the price $2,325,000 is an outlier.

f. $\bar{x} = \dfrac{\Sigma x_i}{n} = \dfrac{6749.4}{14} = 482.1$

The mean is sensitive to extremely high home prices and tends to overstate the more typical midrange home price. The sample mean of $482,100 has 79% of home prices below this value and 21% of the home prices above this value while the sample median $215,900 has 50% above and 50% below. The median is more stable and not influenced by the extremely high home prices. Using the sample mean $482,100 would overstate the more typical or middle home price.

61.

f_i	M_i	$f_i M_i$	$M_i - \bar{x}$	$(M_i - \bar{x})^2$	$f_i(M_i - \bar{x})^2$
10	47	470	-13.68	187.1424	1871.42
40	52	2080	-8.68	75.3424	3013.70
150	57	8550	-3.68	13..5424	2031.36
175	62	10850	+1.32	1.7424	304.92
75	67	5025	+6.32	39.9424	2995.68
15	72	1080	+11.32	128.1424	1922.14
10	77	770	+16.32	266.3424	2663.42
475		28,825			14,802.64

a. $\bar{x} = \dfrac{28,825}{475} = 60.68$

b. $s^2 = \dfrac{14,802.64}{474} = 31.23$

$s = \sqrt{31.23} = 5.59$

62. a. Public Transportation: $\bar{x} = \dfrac{320}{10} = 32$

Automobile: $\bar{x} = \dfrac{320}{10} = 32$

b. Public Transportation: $s = 4.64$

Automobile: $s = 1.83$

c. Prefer the automobile. The mean times are the same, but the auto has less variability.

d. Data in ascending order:

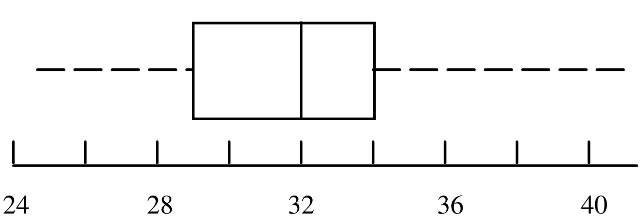
Public:	25	28	29	29	32	32	33	34	37	41
Auto:	29	30	31	31	32	32	33	33	34	35

Five number Summaries

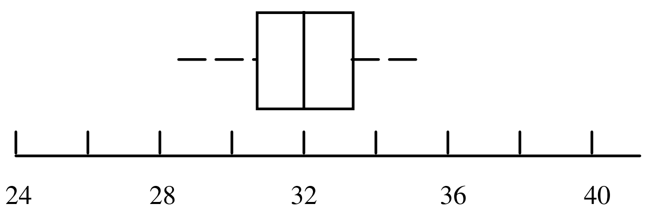
Public:	25	29	32	34	41
Auto:	29	31	32	33	35

Box Plots:

Public:



Auto:



The box plots do show lower variability with automobile transportation and support the conclusion in part c.

63. $\bar{x} = \frac{\Sigma w_i x_i}{\Sigma w_i} = \frac{20(20)+30(12)+10(7)+15(5)+10(6)}{20+30+10+15+10} = \frac{965}{85} = 11.4 \text{ days}$

64. a. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{13,400}{20} = 670$

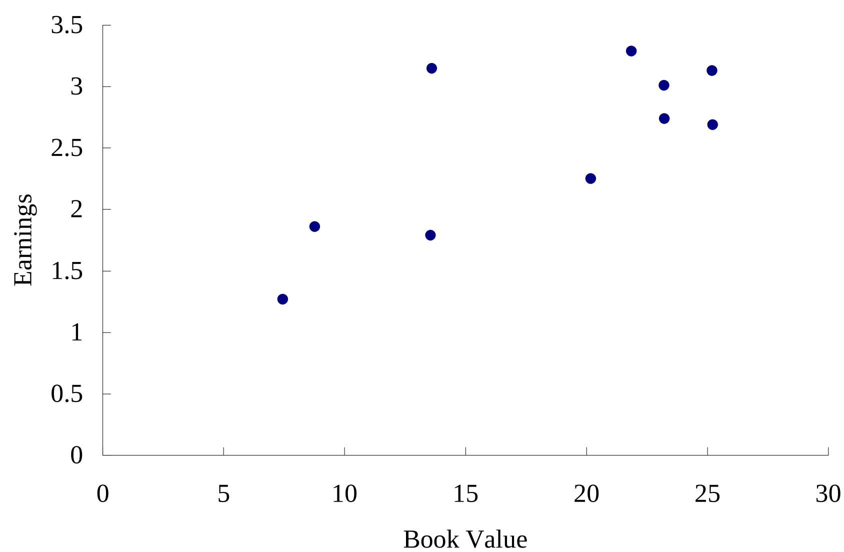
b. $s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{3,949,200}{20-1}} = \$456$

c. $z = \frac{x - \bar{x}}{s} = \frac{2040 - 670}{456} = 3.00$

Yes it is an outlier.

d. First of all, the employee payroll service will be up to date on tax regulations. This will save the small business owner the time and effort of learning tax regulations. This will enable the owner greater time to devote to other aspects of the business. In addition, a correctly filed employment tax return will reduce the potential of a tax penalty.

65. a. The scatter diagram is shown below:



b. The sample correlation coefficient is .75; this indicates a positive linear relationship between book value and earnings.

66. a. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{23}{10} = 2.3$

Median: 5th and 6th positions

$$\text{Median} = \frac{1.8+1.9}{2} = 1.85$$

b. $s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1} = \frac{17.08}{10-1} = 1.90$

$s = \sqrt{1.90} = 1.38$

c. Altria Group at 5%

d. $z = \frac{x-\bar{x}}{s} = \frac{1.6-2.3}{1.38} = -.51$

McDonald's is about 1/2 a standard deviation below the mean dividend yield.

e. $z = \frac{x-\bar{x}}{s} = \frac{3.7-2.3}{1.38} = +1.02$

General Motors is about one standard deviation above the mean dividend yield.

f. Altria Group $z = \frac{x-\bar{x}}{s} = \frac{5.0-2.3}{1.38} = +1.96$

Wal-Mart $z = \frac{x-\bar{x}}{s} = \frac{0.7-2.3}{1.38} = -1.16$

No outliers.

67. a. (800 + 750 + 900)/3 = 817

b.

Month	January	February	March
Weight	1	2	3

$$x = \frac{\Sigma w_i x_i}{\Sigma w_i} = \frac{1(800)+2(750)+3(900)}{1+2+3} = \frac{5000}{6} = 833$$

68. a. Arrange the data in order

Men

21 23 24 25 25 26 26 27 27 27 27 28 28 29 30 30 32 35

Median $i = .5(18) = 9$

Use 9^{th} and 10^{th} positions

Median = 27

Women

19 20 22 22 23 23 24 25 25 26 26 27 28 29 30

Median $i = .5(15) = 7.5$

Use 8^{th} position

Median = 25

b.

	Men	Women
Q_1	$i = .25(18) = 4.5$ Use 5^{th} position $Q_1 = 25$	$i = .25(15) = 3.75$ Use 4^{th} position $Q_1 = 22$
Q_3	$i = .75(18) = 13.5$ Use 14^{th} position $Q_3 = 29$	$i = .75(15) = 11.25$ Use 12^{th} position $Q_3 = 27$

c. Young people today are waiting longer to get married than young people did 25 years ago. The median age for men has increased from 25 to 27. The median age for women has increased from 22 to 25.

69. a. The scatter diagram indicates a positive relationship

b. $\Sigma x_i = 798 \quad \Sigma y_i = 11,688 \quad \Sigma x_i y_i = 1,058,019$

$\Sigma x_i^2 = 71,306 \quad \Sigma y_i^2 = 16,058,736$

$$r_{xy} = \frac{\Sigma x_i y_i - \left(\Sigma x_i \Sigma y_i\right)/n}{\sqrt{\Sigma x_i^2 - \left(\Sigma x_i\right)^2/n}\sqrt{\Sigma y_i^2 - \left(\Sigma y_i\right)^2/n}} = \frac{1,058,019 - (798)(11,688)/9}{\sqrt{71,306 - (798)^2/9}\sqrt{16,058,736 - (11,688)^2/9}} = .9856$$

Strong positive relationship

70. a. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{27000}{15} = 1800$

Median 8th position = 1351

b. Q_1: $i = \left(\frac{25}{100}\right)15 = 3.75$

4th position: $Q_1 = 387$

Q_3: $i = \left(\frac{75}{100}\right)15 = 11.25$

12th position: $Q_3 = 1710$

c. Range = 7450 - 170 = 7280

IQR = Q_3 - Q_1 = 1710 - 387 = 1323

d. $$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1} = \frac{51,454,242}{15-1} = 3,675,303$$

$$s = \sqrt{3,675,303} = 1917$$

e. High positive skewness. This seems reasonable. A relatively few people will have large monthly expenditures causing the right tail of the distribution to become longer.

f. $$z = \frac{x - \bar{x}}{s} = \frac{4135 - 1800}{1917} = 1.22$$

$$z = \frac{x - \bar{x}}{s} = \frac{7450 - 1800}{1917} = 2.95$$ do not indicate outliers.

These values of z do not indicate outliers.

However, the upper limit for outliers is

Q_3 + 1.5(IQR) = 1710 + 1.5(1323) = 3695

Thus, both \$4135 and \$7450 are outliers.