**Chapter 3**

**Descriptive Statistics: Numerical Measures**

**Learning Objectives**

1. Understand the purpose of measures of location.

2. Be able to compute the mean, weighted mean, geometric 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. Understand the role of summary measures in data dashboards.

**Solutions:**

1. 

10, 12, 16, 17, 20

Median = 16 (middle value)

2. 

10, 12, 16, 17, 20, 21

Median =

3. a. 

b. 

4.

|  |  |
| --- | --- |
| **Period** | **Rate of Return (%)** |
| 1 | -6.0 |
| 2 | -8.0 |
| 3 | -4.0 |
| 4 | 2.0 |
| 5 | 5.4 |

The mean growth factor over the five periods is:



So the mean growth rate (0.9775 – 1)100% = –2.25%.

5. 15, 20, 25, 25, 27, 28, 30, 34

 2nd position = 20

 6th position = 28

6. 

Median = 57 6th item

Mode = 53 It appears 3 times

7. a. The mean commute time is 26.9 minutes.

b. The median commute time is 25.95 minutes.

c. The data are bimodal. The modes are 23.4 and 24.8.

d. The index for the third quartile is , so the third quartile is the mean of the values of the 36th and 37th observations in the sorted data, or 

8. a. 

b. 

c. of 3-point shots were made from the 20 feet, 9 inch line during the 19 games.

d. Moving the 3-point line back to 20 feet, 9 inches has reduced the number of 3-point shots taken per game from 19.07 to 18.42, or 19.07 – 18.42 = .65 shots per game. The percentage of 3-points made per game has been reduced from 35.2% to 34.3%, or only .9%. The move has reduced both the number of shots taken per game and the percentage of shots made per game, but the differences are small. The data support the Associated Press Sportsconclusion that the move has not changed the game dramatically.

The 2008-09 sample data shows 120 3-point baskets in the 19 games. Thus, the mean number of points scored from the 3-point line is 120(3)/19 = 18.95 points per game. With the previous 3-point line at 19 feet, 9 inches, 19.07 shots per game and a 35.2% success rate indicate that the mean number of points scored from the 3-point line was 19.07(.352)(3) = 20.14 points per game. There is only a mean of 20.14 – 18.95 = 1.19 points per game less being scored from the 20 feet, 9 inch 3-point line.

9. a. 

b. Order the data from low 6.7 to high 36.6

Median  Use 5th and 6th positions.



c. Mode = 7.2 (occurs 2 times)

d.  Use 3rd position. *Q*1 = 7.2

 Use 8th position. *Q*3 = 17.2

e. Σ*xi* = $148 billion

The percentage of total endowments held by these 2.3% of colleges and universities is (148/413)(100) = 35.8%.

f. A decline of 23% would be a decline of .23(148) = $34 billion for these 10 colleges and universities. With this decline, administrators might consider budget cutting strategies such as

* Hiring freezes for faculty and staff
* Delaying or eliminating construction projects
* Raising tuition
* Increasing enrollments

10. a. 

Order the data from low 100 to high 360

Median  Use 10th and 11th positions

Median = 

Mode = 120 (occurs 3 times)

b.  Use 5th and 6th positions



 Use 15th and 16th positions



c.  Use 18th and 19th positions

90th percentile 

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

11. a. The median number of hours worked per week for high school science teachers is 54.

b. The median number of hours worked per week for high school English teachers is 47.

c. The median number of hours worked per week for high school science teachers is greater than the median number of hours worked per week for high school English teachers; the difference is 54 – 47 = 7 hours.

12. a. The minimum number of viewers that watched a new episode is 13.3 million, and the maximum number is 16.5 million.

b. The mean number of viewers that watched a new episode is 15.04 million or approximately 15.0 million; the median is also 15.0 million. The data is multimodal (13.6, 14.0, 16.1, and 16.2 million); in such cases the mode is usually not reported.

c. The data are first arranged in ascending order. The index for the first quartile is , so the first quartile is the value of the 6th observation in the sorted data, or 14.1. The index for the third quartile is , so the third quartile is the value of the 16th observation in the sorted data, or 16.0.

d. A graph showing the viewership data over the air dates follows. Period 1 corresponds to the first episode of the season, period 2 corresponds to the second episode, and so on.

This graph shows that viewership of *The Big Bang Theory* has been relatively stable over the 2011–2012 television season.

13. Using the mean we get =15.58, = 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

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

Mode: 18.6, 19.4

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

14. For March 2011:

The index for the first quartile is , so the first quartile is the value of the 13th observation in the sorted data, or 6.8.

The index for the median is , so the median (or second quartile) is the average of the values of the 25th and 26th observations in the sorted data, or 8.0.

The index for the third quartile is , so the third quartile is the value of the 38th observation in the sorted data, or 9.4.

For March 2012:

The minimum is 3.0

The index for the first quartile is , so the first quartile is the value of the 13th observation in the sorted data, or 6.8.

The index for the median is , so the median (or second quartile) is the average of the values of the 25th and 26th observations in the sorted data, or 7.35.

The index for the third quartile is , so the third quartile is the value of the 38th observation in the sorted data, or 8.6.

It may be easier to compare these results if we place them in a table.

|  |  |  |
| --- | --- | --- |
|  | **March 2011** | **March 2012** |
| First Quartile | 6.8 | 6.8 |
| Median | 8.0 | 7.35 |
| Third Quartile | 9.4 | 8.6 |

The results show that in March 2012 approximately 25% of the states had an unemployment rate of 6.8% or less, the same as in March 2011. However, the median of 7.35% and the third quartile of 8.6% in March 2012 are both less than the corresponding values in March 2011, indicating that unemployment rates across the states are decreasing.

15. To calculate the average sales price we must compute a weighted mean. The weighted mean is



= 38.11

Thus, the average sales price per case is $38.11.

16. a.

|  |  |  |
| --- | --- | --- |
| Grade *xi* | Weight *Wi* |  |
| 4 (A) | 9 |  |
| 3 (B) | 15 |  |
| 2 (C) | 33 |  |
| 1 (D) | 3 |  |
| 0 (F) | 0 |  |
|  | 60 Credit Hours | |



b. Yes; satisfies the 2.5 grade point average requirement

17. a. 



The weighted average total return for the Morningstar funds is 7.81%.

b. If the amount invested in each fund was available, it would be better to use those amounts as weights. The weighted return computed in part (a) will be a good approximation, if the amount invested in the various funds is approximately equal.

c. Portfolio Return = 



The portfolio return would be 12.27%.

18.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Assessment | Deans | *f*i*M*i | Recruiters | *f*i*M*i |
| 5 | 44 | 220 | 31 | 155 |
| 4 | 66 | 264 | 34 | 136 |
| 3 | 60 | 180 | 43 | 129 |
| 2 | 10 | 20 | 12 | 24 |
| 1 | 0 | 0 | 0 | 0 |
| Total | 180 | 684 | 120 | 444 |

Deans: 

Recruiters: 

19. To calculate the mean growth rate we must first compute the geometric mean of the five growth factors:

|  |  |  |
| --- | --- | --- |
| **Year** | **% Growth** | **Growth Factor *xi*** |
| 2007 | 5.5 | 1.055 |
| 2008 | 1.1 | 1.011 |
| 2009 | -3.5 | 0.965 |
| 2010 | -1.1 | 0.989 |
| 2011 | 1.8 | 1.018 |



The mean annual growth rate is (1.007152 – 1)100 = 0.7152%.

20.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Stivers** | | **Trippi** | |
| **Year** | **End of Year Value** | **Growth Factor** | **End of Year Value** | **Growth Factor** |
| 2004 | $11,000 | 1.100 | $5,600 | 1.120 |
| 2005 | $12,000 | 1.091 | $6,300 | 1.125 |
| 2006 | $13,000 | 1.083 | $6,900 | 1.095 |
| 2007 | $14,000 | 1.077 | $7,600 | 1.101 |
| 2008 | $15,000 | 1.071 | $8,500 | 1.118 |
| 2009 | $16,000 | 1.067 | $9,200 | 1.082 |
| 2010 | $17,000 | 1.063 | $9,900 | 1.076 |
| 2011 | $18,000 | 1.059 | $10,600 | 1.071 |

For the Stivers mutual fund we have:

18000=10000, so =1.8 and



So the mean annual return for the Stivers mutual fund is (1.07624 – 1)100 = 7.624%

For the Trippi mutual fund we have:

10600=5000, so =2.12 and



So the mean annual return for the Trippi mutual fund is (1.09848 – 1)100 = 9.848%.

While the Stivers mutual fund has generated a nice annual return of 7.6%, the annual return of 9.8% earned by the Trippi mutual fund is far superior.

21. 5000=3500, so =1.428571, and so



So the mean annual growth rate is (1.040426 – 1)100 = 4.0404%

22. 25,000,000=10,000,000, so =2.50, and so



So the mean annual growth rate is (1.165 – 1)100 = 16.5%

23. Range 20 - 10 = 10

10, 12, 16, 17, 20



*Q*1 (2nd position) = 12



*Q*3 (4th position) = 17

IQR = *Q*3 – *Q*1 = 17 – 12 = 5

24. 





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

IQR = *Q*3 – *Q*1 = 29 – 22.5 = 6.5







26. a. Range = 190 – 168 = 22

b. 



c. 

d. 

27. a. The mean price for a round–trip flight into Atlanta is $356.73, and the mean price for a round–trip flight into Salt Lake City is $400.95. Flights into Atlanta are less expensive than flights into Salt Lake City. This possibly could be explained by the locations of these two cities relative to the 14 departure cities; Atlanta is generally closer than Salt Lake City to the departure cities.

b. For flights into Atlanta, the range is $290.0, the variance is 5517.41, and the standard Deviation is $74.28. For flights into Salt Lake City, the range is $458.8, the variance is 18933.32, and the standard deviation is $137.60.

The prices for round–trip flights into Atlanta are less variable than prices for round–trip flights into Salt Lake City. This could also be explained by Atlanta’s relative nearness to the 14 departure cities.

28. a. The mean serve speed is 180.95, the variance is 21.42, and the standard deviation is 4.63.

b. Although the mean serve speed for the twenty Women's Singles serve speed leaders for the 2011 Wimbledon tournament is slightly higher, the difference is very small. Furthermore, given the variation in the twenty Women's Singles serve speed leaders from the 2012 Australian Open and the twenty Women's Singles serve speed leaders from the 2011 Wimbledon tournament, the difference in the mean serve speeds is most likely due to random variation in the players’ performances.

29. a. Range = 60 – 28 = 32

IQR = *Q*3 – *Q*1 = 55 – 45 = 10

b. 







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

30. Dawson Supply: Range = 11 – 9 = 2



J.C. Clark: Range = 15 – 7 = 8



31. a.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **18–34** | **35–44** | **45+** |
| mean | 1368.0 | 1330.1 | 1070.4 |
| median | 1423.0 | 1382.5 | 1163.5 |
| standard deviation | 540.8 | 431.7 | 334.5 |

b. The 45+ group appears to spend less on coffee than the other two groups, and the 18–34 and 35–44 groups spend similar amounts of coffee.

32. a. Freshmen 

Seniors 

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

 *Q*1 = 1079 (7th item)

 *Q*3 = 1475 (19th item)

IQR = *Q*3 – *Q*1 = 1479 – 1075 = 404

Seniors

IQR = *Q*3 – *Q*1 = 502 – 370.5 = 131.5

d. 

Freshmen 

Seniors 

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

33. a. For 2011





For 2012





b. The mean score is 76 for both years, but there is an increase in the standard deviation for the scores in 2012. The golfer is not as consistent in 2012 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 2012. The optimism in 2012 is that three of the eight scores were better than any score reported for 2011. If the golfer can work for consistency, eliminate the high score rounds, and reduce the standard deviation, golf scores should show improvement.

34. Quarter milers

*s* = 0.0564

Coefficient of Variation = (*s*/)100% = (0.0564/0.966)100% = 5.8%

Milers

*s* = 0.1295

Coefficient of Variation = (*s*/)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.

35. 



10 

20 

12 

17 

16 

36. 









37. a.  At least 75%

b.  At least 89%

c.  At least 61%

d.  At least 83%

e.  At least 92%

38. a. Approximately 95%

b. Almost all

c. Approximately 68%

39. a. This is from 2 standard deviations below the mean to 2 standard deviations above the mean.

With *z* = 2, Chebyshev’s theorem gives:



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:



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.

40. a.   $3.33 is one standard deviation below the mean and $3.53 is one standard deviation above the mean. The empirical rule says that approximately 68% of gasoline sales are in this price range.

b.   Part (a) shows that approximately 68% of the gasoline sales are between $3.33 and $3.53. Since the bell-shaped distribution is symmetric, approximately half of 68%, or 34%, of the gasoline sales should be between $3.33 and the mean price of $3.43. $3.63 is two standard deviations above the mean price of $3.43. 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 $3.43 and $3.63. The percentage of gasoline sales between $3.33 and $3.63 should be approximately 34% + 47.5% = 81.5%.

  c.   $3.63 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 $3.63.

41. a. 615 is one standard deviation above the mean. Approximately 68% of the scores are between 415 and 615 with half of 68%, or 34%, of the scores between the mean of 515 and 615. Also, since the distribution is symmetric, 50% of the scores are above the mean of 515. With 50% of the scores above 515 and with 34% of the scores between 515 and 615, 50% – 34% = 16% of the scores are above 615.

b. 715 is two standard deviations above the mean. Approximately 95% of the scores are between 315 and 715 with half of 95%, or 47.5%, of the scores between the mean of 515 and 715. Also, since the distribution is symmetric, 50% of the scores are above the mean of 515. With 50% of the scores above 515 and with 47.5% of the scores between 515 and 715, 50%– 47.5% = 2.5% of the scores are above 715.

c. Approximately 68% of the scores are between 415 and 615 with half of 68%, or 34%, of the scores between 415 and the mean of 515.

d. Approximately 95% of the scores are between 315 and 715 with half of 95%, or 47.5%, of the scores between 315 and the mean of 515. Approximately 68% of the scores are between 415 and 615 with half of 68%, or 34%, of the scores between the mean of 515 and 615. Thus, 47.5% + 34% = 81.5% of the scores are between 315 and 615.

42. a. 

b. 

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

d. 

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

43. a. 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



c. 

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.

44. a. 



b. 

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.



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. 



Largest margin 24: . No outliers.

45. a. 

b. $75.00 – $72.20 = $2.80

$2.80/$72.20 = .0388 Ticket price increased 3.88% during the one-year period.

c. 7th position – Green Bay Packers 63

8th position – Pittsburgh Steelers 67

Median =

d.  Use 4th position

*Q*1  = 61 (Tennessee Titans)

 Use 11th position

*Q*3  = 83 (Indianapolis Colts)

e. 

f. Dallas Cowboys: 

With *z*> 3, this is an outlier. The Dallas Cowboys have an unusually high ticket price compared to the other NFL teams.

46. 15, 20, 25, 25, 27, 28, 30, 34

Smallest = 15



Largest = 34

47.



48. 5, 6, 8, 10, 10, 12, 15, 16, 18

Smallest = 5

 *Q*1 = 8 (3rd position)

Median = 10

 *Q*3 = 15 (7th position)

Largest = 18



49. 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

50. a. The first place runner in the men’s group finished minutes ahead of the first place runner in the women’s group. Lauren Wald would have finished in 11th place for the combined groups.

b. Men: . Use the 11th and 12th place finishes.

Median = 

Women: . Use the 16th place finish. Median = 131.67.

Using the median finish times, the men’s group finished minutes ahead of the women’s group.

Also note that the fastest time for a woman runner, 109.03 minutes, is approximately equal to the median time of 109.64 minutes for the men’s group.

c. Men: Lowest time = 65.30; Highest time = 148.70

*Q*1:  Use 6th position. *Q*1 = 87.18

*Q*3:  Use 17th position. *Q*3 = 128.40

Five number summary for men: 65.30, 87.18, 109.64, 128.40, 148.70

Women: Lowest time = 109.03; Highest time = 189.28

*Q*1:  Use 8th position. *Q*1 = 122.08

*Q*3:  Use 24th position. *Q*3 = 147.18

Five number summary for women: 109.03, 122.08, 131.67, 147.18, 189.28

d. Men: IQR = 

Lower Limit = 

Upper Limit = 

There are no outliers in the men’s group.

Women: IQR = 

Lower Limit = 

Upper Limit = 

The two slowest women runners with times of 189.27 and 189.28 minutes are outliers in the women’s group.

e.



The box plots show the men runners with the faster or lower finish times. However, the box plots show the women runners with the lower variation in finish times. The interquartile ranges of 41.22 minutes for men and 25.10 minutes for women support this conclusion.

51. a. Median (11th position) = 4019



*Q*1 (6th position) = 1872



*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 (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.



52. a. Median *n* = 20; 10th and 11th positions

Median =

b. Smallest 68

*Q*1: ; 5th and 6th positions



*Q*3: ; 15th and 16th positions



Largest 77

5- number summary: 68, 71.5, 73.5, 74.5, 77

c. IQR = *Q*3 – *Q*1 = 74.5 – 71.5 = 3

Lower Limit = *Q*1 – 1.5(IQR)

= 71.5 – 1.5(3) = 67

Upper Limit = *Q*3 + 1.5(IQR)

= 74.5 + 1.5(3) = 79

All ratings are between 67 and 79. There are no outliers for the T-Mobile service.

d. Using the solution procedures shown in parts a, b, and c, the five number summaries and outlier limits for the other three cell-phone services are as follows.

AT&T 66, 68, 71, 73, 75 Limits: 60.5 and 80.5

Sprint 63, 65, 66, 67.5, 69 Limits: 61.25 and 71.25

Verizon 75, 77, 78.5, 79.5, 81 Limits: 73.25 and 83.25

There are no outliers for any of the cell-phone services.

e.



The box plots show that Verizon is the best cell-phone service provider in terms of overall customer satisfaction. Verizon’s lowest rating is better than the highest AT&T and Sprint ratings and is better than 75% of the T-Mobile ratings. Sprint shows the lowest customer satisfaction ratings among the four services.

53. a. Total Salary for the Philadelphia Phillies = $96,870,000

Median *n* = 28; 14th and 15th positions

Median =

Smallest 390

*Q*1: ; 7th and 8th positions



*Q*3: ; 21st and 22nd positions



Largest 14250

5– number summary for the Philadelphia Phillies: 390, 432.5, 1300, 6175, 14250

Using the 5-number summary, the lower quartile shows salaries closely bunched between 390 and 432.5. The median is 1300. The most variation is in the upper quartile where the salaries are spread between 6175 and 14250, or between $6,175,000 and $14,250,000.

b. IQR = *Q*3 – *Q*1 = 6175 – 432.5 = 5742.5

Lower Limit = *Q*1 – 1.5(IQR)

= 432.5 –1.5(5742.5) = – 8181.25; Use 0

Upper Limit = *Q*3 + 1.5(IQR)

= 6175 + 1.5(5742.5) = 14788.75

All salaries are between 0 and 14788.75. There are no salary outliers for the Philadelphia Phillies.

c. Using the solution procedures shown in parts a and b, the total salary, the five-number summaries, and the outlier limits for the other teams are as follows.

Los Angeles Dodgers $136,373,000

390, 403, 857.5, 9125, 19000 Limits: 0 and 22208

Tampa Bay Rays $ 42,334,000

390, 399, 415, 2350, 6000 Limits: 0 and 5276.5

Boston Red Sox $120,460,000

396, 439.5, 2500, 8166.5, 14000 Limits: 0 and 19757

The Los Angeles Dodgers had the highest payroll while the Tampa Bay Rays clearly had the lowest payroll among the four teams. With the lower salaries, the Rays had two outlier salaries compared to other salaries on the team. But these top two salaries are substantially below the top salaries for the other three teams. There are no outliers for the Phillies, Dodgers and Red Sox.

d.



The box plots show that the lowest salaries for the four teams are very similar. The Red Sox have the highest median salary. Of the four teams the Dodgers have the highest upper end salaries and highest total payroll, while the Rays are clearly the lowest paid team.

For this data, we would conclude that paying higher salaries do not always bring championships. In the National League Championship, the lower paid Phillies beat the higher paid Dodgers. In the American League Championship, the lower paid Rays beat the higher paid Red Sox. The biggest surprise was how the Tampa Bay Rays over achieved based on their salaries and made it to the World Series. Teams with the highest salaries do not always win the championships.

54. a. 

Median 23rd position 15.1

24th position 15.6

Median =

b. *Q*1: 

12th position: *Q*1 = 11.7

*Q*3: 

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



55. a.

b. Negative relationship

c/d. 





There is a strong negative linear relationship.

56. a.

b. Positive relationship

c/d. 











A positive linear relationship

57. a.

b. The scatter diagram shows a positive relationship with higher predicted point margins associated with higher actual point margins.

c. Let *x* = predicted point margin and *y* = actual point margin







A positive covariance shows a positive relationship between predicted point margins and actual point margins.

d.



The modest positive correlation shows that the Las Vegas predicted point margin is a general, but not a perfect, indicator of the actual point margin in college football bowl games.

Note: The Las Vegas odds makers set the point margins so that someone betting on a favored team has to have the team *win by more than the point margin* to win the bet. For example, someone betting on Auburn to win the Outback Bowl would have to have Auburn win by more than five points to win the bet. Since Auburn beat Northwestern by only three points, the person betting on Auburn would have lost the bet.

A review of the predicted and actual point margins shows that the favorites won by more than the predicted point margin in five bowl games: Gator, Sugar, Cotton, Alamo, and the Championship bowl game. The underdog either won its game or kept the actual point margin less than the predicted point margin in the other five bowl games. In this case, betting on the underdog would have provided winners in the Outback, Capital One, Rose, Fiesta and Orange bowls. In this example, the Las Vegas odds point margins made betting on the favored team a 50-50 probability of winning the bet.

58. Let *x* = miles per hour and *y* = miles per gallon







A strong negative linear relationship exists. For driving speeds between 25 and 60 miles per hour,

higher speeds are associated with lower miles per gallon.

59. a.  

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
| 7.1 | 7.02 | 0.2852 | 0.6893 | 0.0813 | 0.4751 | 0.1966 |
| 5.2 | 5.31 | -1.6148 | -1.0207 | 2.6076 | 1.0419 | 1.6483 |
| 7.8 | 5.38 | 0.9852 | -0.9507 | 0.9706 | 0.9039 | -0.9367 |
| 7.8 | 5.40 | 0.9852 | -0.9307 | 0.9706 | 0.8663 | -0.9170 |
| 5.8 | 5.00 | -1.0148 | -1.3307 | 1.0298 | 1.7709 | 1.3505 |
| 5.8 | 4.07 | -1.0148 | -2.2607 | 1.0298 | 5.1109 | 2.2942 |
| 9.3 | 6.53 | 2.4852 | 0.1993 | 6.1761 | 0.0397 | 0.4952 |
| 5.7 | 5.57 | -1.1148 | -0.7607 | 1.2428 | 0.5787 | 0.8481 |
| 7.3 | 6.99 | 0.4852 | 0.6593 | 0.2354 | 0.4346 | 0.3199 |
| 7.6 | 11.12 | 0.7852 | 4.7893 | 0.6165 | 22.9370 | 3.7605 |
| 8.2 | 7.56 | 1.3852 | 1.2293 | 1.9187 | 1.5111 | 1.7028 |
| 7.1 | 12.11 | 0.2852 | 5.7793 | 0.0813 | 33.3998 | 1.6482 |
| 6.3 | 4.39 | -0.5148 | -1.9407 | 0.2650 | 3.7665 | 0.9991 |
| 6.6 | 4.78 | -0.2148 | -1.5507 | 0.0461 | 2.4048 | 0.3331 |
| 6.2 | 5.78 | -0.6148 | -0.5507 | 0.3780 | 0.3033 | 0.3386 |
| 6.3 | 6.08 | -0.5148 | -0.2507 | 0.2650 | 0.0629 | 0.1291 |
| 7.0 | 10.05 | 0.1852 | 3.7193 | 0.0343 | 13.8329 | 0.6888 |
| 6.2 | 4.75 | -0.6148 | -1.5807 | 0.3780 | 2.4987 | 0.9719 |
| 5.5 | 7.22 | -1.3148 | 0.8893 | 1.7287 | 0.7908 | -1.1692 |
| 6.5 | 3.79 | -0.3148 | -2.5407 | 0.0991 | 6.4554 | 0.7999 |
| 6.0 | 3.62 | -0.8148 | -2.7107 | 0.6639 | 7.3481 | 2.2088 |
| 8.3 | 9.24 | 1.4852 | 2.9093 | 2.2058 | 8.4638 | 4.3208 |
| 7.5 | 4.40 | 0.6852 | -1.9307 | 0.4695 | 3.7278 | -1.3229 |
| 7.1 | 6.91 | 0.2852 | 0.5793 | 0.0813 | 0.3355 | 0.1652 |
| 6.8 | 5.57 | -0.0148 | -0.7607 | 0.0002 | 0.5787 | 0.0113 |
| 5.5 | 3.87 | -1.3148 | -2.4607 | 1.7287 | 6.0552 | 3.2354 |
| 7.5 | 8.42 | 0.6852 | 2.0893 | 0.4695 | 4.3650 | 1.4315 |
|  |  |  | Total | 25.77407 | 130.0594 | 25.5517 |



There is evidence of a modest positive linear association between the jobless rate and the delinquent housing loan percentage. If the jobless rate were to increase, it is likely that an increase in the percentage of delinquent housing loans would also occur.

b.

60. a.

b. 



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
| 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 |



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.

61. a. 

b. 

c.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
| 68 | 50 | .5 | -.4286 | .25 | .1837 | -.2143 |
| 70 | 49 | 2.5 | -1.4286 | 6.25 | 2.0408 | -3.5714 |
| 65 | 44 | -2.5 | -6.4286 | 6.25 | 41.3265 | 16.0714 |
| 96 | 64 | 28.5 | 13.5714 | 812.25 | 184.1837 | 386.7857 |
| 57 | 46 | -10.5 | -4.4286 | 110.25 | 19.6122 | 46.5000 |
| 70 | 45 | 2.5 | -5.4286 | 6.25 | 29.4694 | -13.5714 |
| 80 | 73 | 12.5 | 22.5714 | 156.25 | 509.4694 | 282.1429 |
| 67 | 45 | -.5 | -5.4286 | .25 | 29.4694 | 2.7143 |
| 44 | 29 | -23.5 | -21.4286 | 552.25 | 459.1837 | 503.5714 |
| 69 | 44 | 1.5 | -6.4286 | 2.25 | 41.3265 | -9.6429 |
| 76 | 69 | 8.5 | 18.5714 | 72.25 | 344.8980 | 157.8571 |
| 69 | 51 | 1.5 | .5714 | 2.25 | .3265 | .8571 |
| 70 | 58 | 2.5 | 7.5714 | 6.25 | 57.3265 | 18.9286 |
| 44 | 39 | -23.5 | -11.4286 | 552.25 | 130.6122 | 268.5714 |
|  |  |  | Total | 2285.5 | 1849.4286 | 1657.0000 |









High positive correlation as should be expected.

62. a. The mean is 2.95 and the median is 3.0.

b. The index for the first quartile is , so the first quartile is the mean of the values of the 5th and 6th observations in the sorted data, or .

The index for the third quartile is , so the third quartile is the mean of the values of the 15th and 16th observations in the sorted data, or .

c. The range is 7 and the interquartile range is 4.5 – 1 = 3.5.

d. The variance is 4.37 and standard deviation is 2.09.

e. Because most people dine out a relatively few times per week and a few families dine out very frequently, we would expect the data to be positively skewed. The skewness measure of 0.34 indicates the data are somewhat skewed to the right.

f. The lower limit is –4.25 and the upper limit is 9.75. No values in the data are less than the lower limit or greater than the upper limit, so the Minitab boxplot indicates there are no outliers.



63. 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 9th and 10th 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 8th position

Median = 25

b. Men Women

|  |  |
| --- | --- |
| *Q*1 *i* = .25(18) = 4.5 | *i* = .25(15) = 3.75 |
| Use 5th position | Use 4th position |
| *Q*1= 25 | *Q*1= 22 |
|  |  |
| *Q*3 *i* = .75(18) = 13.5 | *i* = .75(15) = 11.25 |
| Use 14th position | Use 12th position |
| *Q*3= 29 | *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.

64. a. The mean and median patient wait times for offices with a wait tracking system are 17.2 and 13.5, respectively. The mean and median patient wait times for offices without a wait tracking system are 29.1 and 23.5, respectively.

b. The variance and standard deviation of patient wait times for offices with a wait tracking system are 86.2 and 9.3, respectively. The variance and standard deviation of patient wait times for offices without a wait tracking system are 275.7 and 16.6, respectively.

c. Offices with a wait tracking system have substantially shorter patient wait times than offices without a wait tracking system.

d. 

e. 

As indicated by the positive *z*–scores, both patients had wait times that exceeded the means of their respective samples. Even though the patients had the same wait time, the *z*–score for the sixth patient in the sample who visited an office with a wait tracking system is much larger because that patient is part of a sample with a smaller mean and a smaller standard deviation.

f. The *z*–scores for all patients follow.

|  |  |
| --- | --- |
| **Without Wait Tracking System** | **With Wait Tracking System** |
| -0.31 | 1.49 |
| 2.28 | -0.67 |
| -0.73 | -0.34 |
| -0.55 | 0.09 |
| 0.11 | -0.56 |
| 0.90 | 2.13 |
| -1.03 | -0.88 |
| -0.37 | -0.45 |
| -0.79 | -0.56 |
| 0.48 | -0.24 |

The *z*–scores do not indicate the existence of any outliers in either sample.

65. a. 

Mean debt upon graduation is $10,000.

b. 



66. a. 

b. 

c. 

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.

67. a. Public Transportation: 

Automobile: 

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.

68. 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 =

Median home price = $215,900

b. 

55% increase over the five-year period

c. *n* = 14



Use the 4th position

*Q*1 = 175.0



Use the 11th position

*Q*3 = 628.3

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. 

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.

69. a. Median for *n* = 50; Use 25th and 26th positions

25th – South Dakota 16.8

26th – Pennsylvania 16.9

Median =

b. *Q*1: 

13th position: *Q*1 = 13.7% (Iowa)

*Q*3: 

38th position: *Q*3 = 20.2% (North Carolina & Georgia)

25% of the states have a poverty level less than or equal to 13.7% and 25% of the states have a poverty level greater than or equal to 20.2%

c. IQR = *Q*3 – *Q*1 = 20.2 – 13.7 = 6.5

Upper Limit = *Q*3 + 1.5(IQR)

= 20.2 + 1.5(6.5) = 29.95

Lower Limit = *Q*1 – 1.5(IQR)

= 13.7 – 1.5(6.5) = 3.95



The Minitab box plot shows the distribution of poverty levels is skewed to the right (positive). There are no states considered outliers. Mississippi with 29.5% is closest to being an outlier on the high poverty rate side. New Hampshire has the lowest poverty level with 9.6%. The five-number summary is 9.6, 13.7, 16.85, 20.2 and 29.95.

d. The states in the lower quartile are the states with the lowest percentage of children who have lived below the poverty level in the last 12 months. These states are as follows.

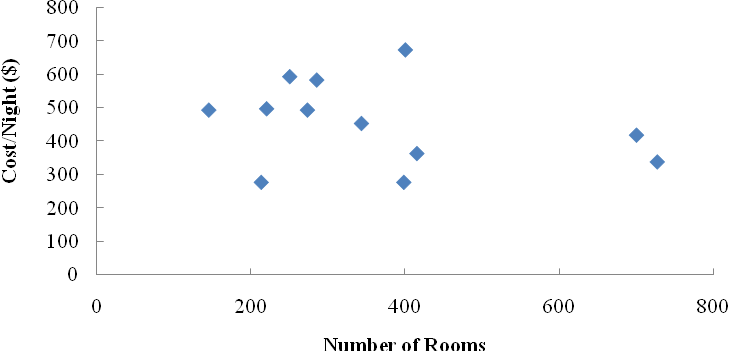
|  |  |  |
| --- | --- | --- |
| **State** | **Region** | **Poverty %** |
| New Hampshire | NE | 9.6 |
| Maryland | NE | 9.7 |
| Connecticut | NE | 11.0 |
| Hawaii | W | 11.4 |
| New Jersey | NE | 11.8 |
| Utah | W | 11.9 |
| Wyoming | W | 12.0 |
| Minnesota | MW | 12.2 |
| Virginia | SE | 12.2 |
| Massachusetts | NE | 12.4 |
| North Dakota | MW | 13.0 |
| Vermont | NE | 13.2 |

Generally, these states are the states with better economic conditions and less poverty. The Northeast region with 6 of the 12 states in this quartile appears to be the best economic region of the country. The West region was second with 3 of the 12 states in this group.

70. a. rooms

b. 

c.



It is difficult to see much of a relationship. When the number of rooms becomes larger, there is no indication that the cost per night increases. The cost per night may even decrease slightly.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
|  | 499 | -144 | 42 | 20.736 | 1,764 | -6,048 |
| 727 | 340 | 363 | -117 | 131,769 | 13,689 | -42,471 |
| 285 | 585 | -79 | 128 | 6,241 | 16,384 | -10,112 |
| 273 | 495 | -91 | 38 | 8,281 | 1,444 | -3,458 |
| 145 | 495 | -219 | 38 | 47,961 | 1,444 | -8,322 |
| 213 | 279 | -151 | -178 | 22,801 | 31,684 | 26,878 |
| 398 | 279 | 34 | -178 | 1,156 | 31,684 | -6,052 |
| 343 | 455 | -21 | -2 | 441 | 4 | 42 |
| 250 | 595 | -114 | 138 | 12,996 | 19,044 | -15,732 |
| 414 | 367 | 50 | -90 | 2,500 | 8,100 | -4,500 |
| 400 | 675 | 36 | 218 | 1,296 | 47,524 | 7,848 |
| 700 | 420 | 336 | -37 | 112,896 | 1,369 | -12,432 |
|  |  |  | Total | 69,074 | 174,134 | -74,359 |

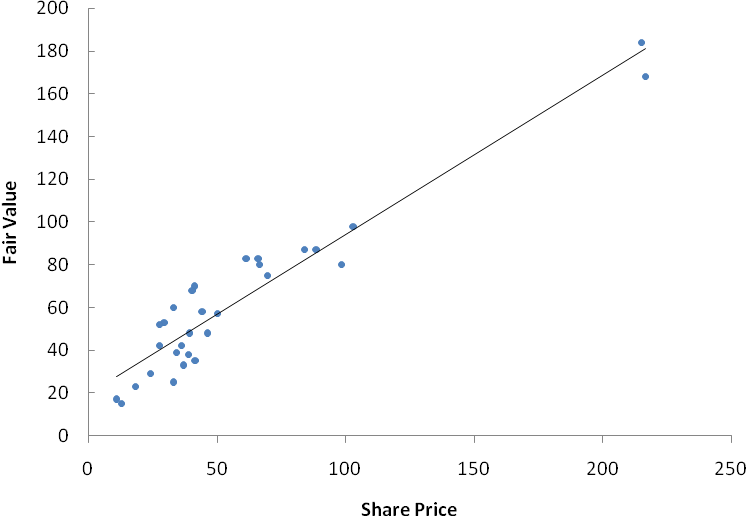
d.



There is evidence of a slightly negative linear association between the number of rooms and the cost per night for a double room. Although this is not a strong relationship, it suggests that the higher room rates tend to be associated with the smaller hotels.

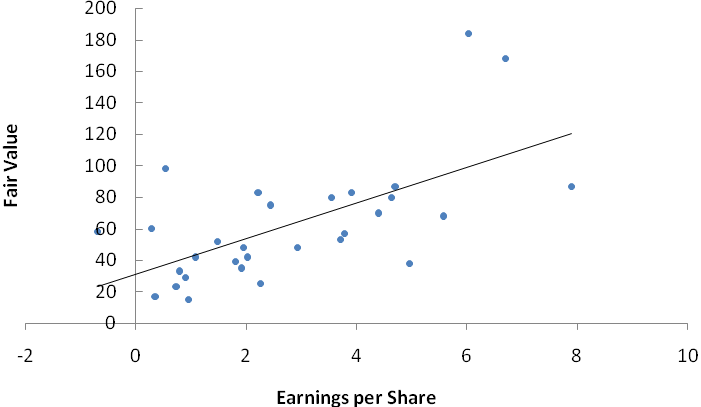
This tends to make sense when you think about the economies of scale for the larger hotels. Many of the amenities in terms of pools, equipment, spas, restaurants, and so on exist for all hotels in the *Travel + Leisure* top 50 hotels in the world. The smaller hotels tend to charge more for the rooms. The larger hotels can spread their fixed costs over many room and may actually be able to charge less per night and still achieve and nice profit. The larger hotels may also charge slightly less in an effort to obtain a higher occupancy rate. In any case, it appears that there is a slightly negative linear association between the number of rooms and the cost per night for a double room at the top hotels.

71. a. The scatter diagram is shown below.



The sample correlation coefficient is .954. This indicates a strong positive linear relationship between Morningstar’s Fair Value estimate per share and the most recent price per share for the stock.

b. The scatter diagram is shown below:



The sample correlation coefficient is .624. While not a strong of a relationship as shown in part a, this indicates a positive linear relationship between Morningstar’s Fair Value estimate per share and the earnings per share for the stock.

72. a.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
| .407 | .422 | -.1458 | -.0881 | .0213 | .0078 | .0128 |
| .429 | .586 | -.1238 | .0759 | .0153 | .0058 | -.0094 |
| .417 | .546 | -.1358 | .0359 | .0184 | .0013 | -.0049 |
| .569 | .500 | .0162 | -.0101 | .0003 | .0001 | -.0002 |
| .569 | .457 | .0162 | -.0531 | .0003 | .0028 | -.0009 |
| .533 | .463 | -.0198 | -.0471 | .0004 | .0022 | .0009 |
| .724 | .617 | .1712 | .1069 | .0293 | .0114 | .0183 |
| .500 | .540 | -.0528 | .0299 | .0028 | .0009 | -.0016 |
| .577 | .549 | .0242 | .0389 | .0006 | .0015 | .0009 |
| .692 | .466 | .1392 | -.0441 | .0194 | .0019 | -.0061 |
| .500 | .377 | -.0528 | -.1331 | .0028 | .0177 | .0070 |
| .731 | .599 | .1782 | .0889 | .0318 | .0079 | .0158 |
| .643 | .488 | .0902 | -.0221 | .0081 | .0005 | -.0020 |
| .448 | .531 | -.1048 | .0209 | .0110 | .0004 | -.0022 |
|  |  |  | Total | .1617 | .0623 | .0287 |









b. There is a low positive correlation between a major league baseball team’s winning percentage during spring training and its winning percentage during the regular season. The spring training record should not be expected to be a good indicator of how a team will play during the regular season.

Spring training consists of practice games between teams with the outcome as to who wins or who loses not counting in the regular season standings or affecting the chances of making the playoffs. Teams use spring training to help players regain their timing and evaluate new players. Substitutions are frequent with the regular or better players rarely playing an entire spring training game. Winning is not the primary goal in spring training games. A low correlation between spring training winning percentage and regular season winning percentage should be anticipated.

73. days

74.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *fi* | *Mi* | *fi Mi* |  |  |  |
| 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. 

b. 



75. a.

It appears the Panama Railroad Company outperformed the New York Stock Exchange annual average return of 8.4%, but the large drop in returns in 1870-71 makes it difficult to be certain.

b. The geometric mean is



So the mean annual return on Panama Railroad Company stock is 10.6%. During the period of 1853–1880, the Panama Railroad Company stock yielded a return superior to the 8.4% earned by the New York Stock Exchange.

Note that we could also calculate the geometric mean with Excel. If the growth factors for the individual years are in cells C2:C30, then typing =GEOMEAN(C2:C30) into an empty cell will yield the geometric mean.