

**TEST BANK**



McCLAVE | SINICICH

A FIRST COURSE IN  
**STATISTICS** TENTH EDITION



## Ch. 2 Methods for Describing Sets of Data

### 2.1 Describing Qualitative Data

#### 1 Identify Classes/Compute Class Frequencies/Relative Frequencies/Percentages

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

**Solve the problem.**

- 1) In an eye color study, 25 out of 50 people in the sample had brown eyes. In this situation, what does the number .50 represent?  
A) a class relative frequency  
B) a class  
C) a class frequency  
D) a class percentage
- 2) What class percentage corresponds to a class relative frequency of .37?  
A) 37%  
B) .37%  
C) .63%  
D) 63%

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

- 3) A sample of 100 e-mail users were asked whether their primary e-mail account was a free account, an institutional (school or work) account, or an account that they pay for personally. Identify the classes for the resulting data.

#### 2 Construct Frequency/Relative Frequency Table

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

**Solve the problem.**

- 1) What number is missing from the table?

Grades on Test	Frequency	Relative Frequency
A	6	.24
B	7	
C	9	.36
D	2	.08
F	1	.04

- A) .28                      B) .07                      C) .72                      D) .70

- 2) What number is missing from the table?

Year in College	Frequency	Relative Frequency
Freshman	600	.30
Sophomore	560	.28
Junior		.22
Senior	400	.20

- A) 440                      B) 220                      C) 480                      D) 520

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

3) Complete the frequency table for the data shown below.

green	blue	brown	orange	blue
brown	orange	blue	red	green
blue	brown	green	red	brown
blue	brown	blue	blue	red

Color	Frequency
Green	
Blue	
Brown	
Orange	

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Answer the question True or False.

4) A frequency table displays the proportion of observations falling into each class.

A) True

B) False

### 3 Construct, Interpret Bar Graph

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Solve the problem.

1)



The manager of a store conducted a customer survey to determine why customers shopped at the store. The results are shown in the figure. What proportion of customers responded that merchandise was the reason they shopped at the store?

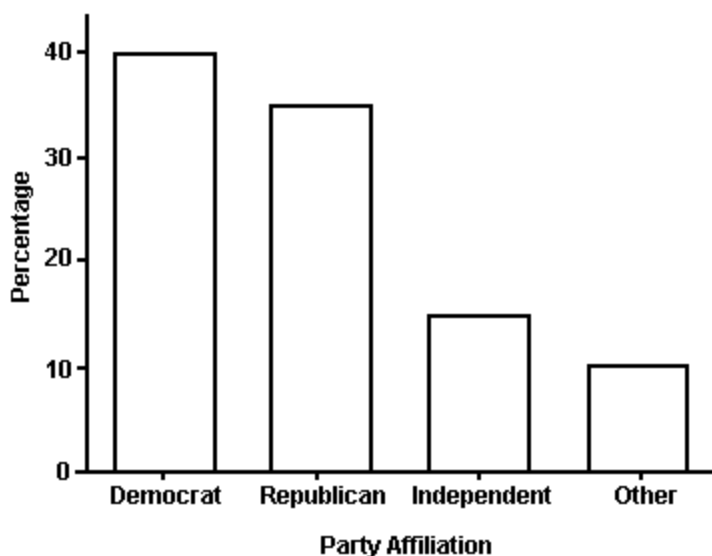
A)  $\frac{3}{7}$

B) 30

C)  $\frac{1}{2}$

D)  $\frac{2}{7}$

2)



The bar graph shows the political affiliation of 1,000 registered U.S. voters. What percentage of the voters belonged to one of the traditional two parties (Democratic or Republican)?

- A) 75%                                      B) 40%                                      C) 35%                                      D) 25%

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

3) The data below show the types of medals won by athletes representing the United States in the 2006 Winter Olympics.

gold      gold      silver      gold      bronze      silver      silver  
bronze    gold      silver      silver    bronze      silver      gold  
gold      silver      silver      bronze    bronze      gold      silver  
gold      gold      bronze      bronze

- a. Construct a frequency table for the data.
- b. Construct a relative frequency table for the data.
- c. Construct a frequency bar graph for the data.

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Answer the question True or False.

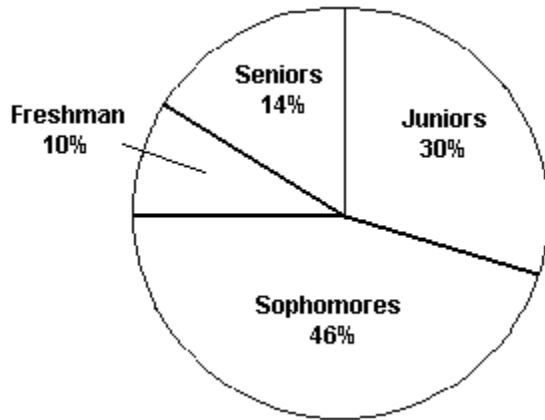
- 4) The bars in a bar graph can be arranged by height in ascending order from left to right.  
A) True                                      B) False
  
- 5) Either vertical or horizontal bars can be used when constructing a bar graph.  
A) True                                      B) False

#### 4 Construct, Interpret Pie Chart

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Solve the problem.

1)



The pie chart shows the classifications of students in a statistics class.

What percentage of the class consists of freshman, sophomores, and juniors?

- A) 86%                      B) 14%                      C) 44%                      D) 54%

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

2) The table shows the number of each type of book found at an online auction site during a recent search.

Type of Book	Number
Children's	51,033
Fiction	141,114
Nonfiction	253,074
Educational	67,252

- a. Construct a relative frequency table for the book data.  
b. Construct a pie chart for the book data.

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Answer the question True or False.

3) If 25% of your statistics class is sophomores, then in a pie chart representing classifications of the students in your statistics class the slice assigned to sophomores is  $90^\circ$ .

- A) True                                      B) False

4) The slices of a pie chart must be arranged from largest to smallest in a clockwise direction.

- A) True                                      B) False



## 2.2 Graphical Methods for Describing Quantitative Data

### 1 Construct, Interpret Histogram

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

**Solve the problem.**

- 1) The data show the total number of medals (gold, silver, and bronze) won by each country winning at least one gold medal in the 2006 Winter Olympics.

1 2 3 3 4 9 9 11 11

11 14 14 19 22 23 24 25 29

- a. Complete the class frequency table for the data.

Total Medals	Frequency
1-5	
6-10	
11-15	
16-20	
21-25	
26-30	

- b. Using the classes from the frequency table, construct a histogram for the data.

- 2) The total points scored by a basketball team for each game during its last season have been summarized in the table below.

Score	Frequency
41-60	3
61-80	8
81-100	12
101-120	7

- a. Explain why you cannot use the information in the table to construct a stem-and-leaf display for the data.  
b. Construct a histogram for the scores.

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

**Answer the question True or False.**

- 3) All class intervals in a histogram have the same width.  
A) True B) False
- 4) A histogram can be constructed using either class frequencies or class relative frequencies as the heights of the bars.  
A) True B) False
- 5) The bars in a histogram should be arranged by height in descending order from left to right.  
A) True B) False

## 2 Construct, Interpret Stem-and-Leaf Display

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

**Solve the problem.**

- 1) A survey was conducted to determine how people feel about the quality of programming available on television. Respondents were asked to rate the overall quality from 0 (no quality at all) to 100 (extremely good quality). The stem-and-leaf display of the data is shown below.

Stem	Leaf
3	0 9
4	0 3 4 7 8 9 9 9
5	0 1 1 2 3 4 5
6	1 2 5 6 6
7	2 6
8	
9	8

What percentage of the respondents rated overall television quality as very good (regarded as ratings of 80 and above)?

- A) 4%                                      B) 1%                                      C) 32%                                      D) 8%

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

- 2) The scores for a statistics test are as follows:

87 76 91 77 92 94 88 85 66 89  
79 98 51 98 83 88 82 55 15 69

Create a stem-and-leaf display for the data.

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

**Answer the question True or False.**

- 3) For large data sets, a stem-and-leaf display is a better choice than a histogram.  
A) True                                      B) False





- 4) A data set contains the observations 7, 4, 2, 5, 8. Find  $\sum(x - 1)$ .
- A) 21                                      B) 31                                      C) 27                                      D) 25
- 5) A data set contains the observations 4, 6, 5, 7, 1. Find  $\sum x^2 - \frac{(\sum x)^2}{5}$ .
- A) 21.2                                      B) 423.2                                      C) 101.6                                      D) 232.8
- 6) Which expression represents the sum of the squares of the observations in a data set?
- A)  $\sum x^2$                                       B)  $(\sum x^2)^2$                                       C)  $\sum \sqrt{x}$                                       D)  $\sqrt{\sum x}$

## 2.4 Numerical Measures of Central Tendency

### 1 Find Mean, Median, Mode

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

#### Solve the problem.

- 1) A sociologist recently conducted a survey of senior citizens who have net worths too high to qualify for Medicaid but have no private health insurance. The ages of the 25 uninsured senior citizens were as follows:

72	77	70	80	90
78	65	93	69	94
73	96	80	66	85
67	72	85	74	77
64	91	79	68	86

Find the median of the observations.

- A) 77                                      B) 74                                      C) 78                                      D) 77.5
- 2) The scores for a statistics test are as follows:

74	76	69	77	63	92	61	85	79	89
79	88	50	70	85	86	85	87	18	67

Compute the mean score.

- A) 74.00                                      B) 77.20                                      C) 64.95                                      D) 75

- 3) A shoe retailer keeps track of all types of information about sales of newly released shoe styles. One newly released style was marketed to tall people. Listed below are the shoe sizes of 12 randomly selected customers who purchased the new style. Find the mode of the shoe sizes.

$9\frac{1}{2}$       11      12       $11\frac{1}{2}$   
 $8\frac{1}{2}$        $10\frac{1}{2}$       8      11  
 10      11       $9\frac{1}{2}$       10

- A) 11                                      B)  $10\frac{1}{4}$                                       C)  $10\frac{1}{2}$                                       D)  $9\frac{1}{2}$

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

- 4) Each year advertisers spend billions of dollars purchasing commercial time on network television. In the first 6 months of one year, advertisers spent \$1.1 billion. Who were the largest spenders? In a recent article, the top 10 leading spenders and how much each spent (in million of dollars) were listed:

Company A	\$73	Company F	\$26.9
Company B	61.5	Company G	26.6
Company C	57.7	Company H	22.3
Company D	55.2	Company I	21.6
Company E	30.2	Company J	20.7

Calculate the mean and median for the data.

- 5) The data show the total number of medals (gold, silver, and bronze) won by each country winning at least one gold medal in the 2006 Winter Olympics. Find the mean, median, and mode of the numbers of medals won by these countries.

1    2    3    3    4    9    9    11    11  
 11   14   14   19   22   23   24   25   29

- 6) Calculate the mean of a sample for which  $\sum x = 196$  and  $n = 8$ .

- 7) The calculator screens summarize a data set.

```

1-Var Stats
x̄=79.95238095
Σx=1679
Σx²=138471
Sx=14.54467666
σx=14.19415101
↓n=21
  
```

```

1-Var Stats
↑n=21
minX=30
Q1=75
Med=82
Q3=90
maxX=97
  
```

- a. How many data items are in the set?  
 b. What is the sum of the data?  
 c. Identify the mean, median, and mode, if possible.

## 2 Interpret Measures of Central Tendency

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

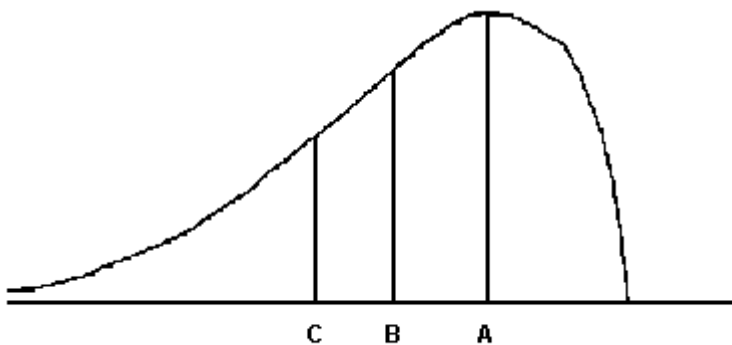
Solve the problem.

- 1) At the U.S. Open Tennis Championship a statistician keeps track of every serve that a player hits during the tournament. The statistician reported that the mean serve speed of a particular player was 98 miles per hour. Suppose that the statistician indicated that the serve speed distribution was skewed to the left. Which of the following values is most likely the value of the median serve speed?  
A) 104 mph                      B) 92 mph                      C) 86 mph                      D) 98 mph
- 2) During one recent year, U.S. consumers redeemed 6.36 billion manufacturers' coupons and saved themselves \$2.45 billion. Calculate and interpret the mean savings per coupon.  
A) The average savings was \$0.39 per coupon.  
B) The average savings was 259.6 cents per coupon.  
C) Half of all coupons were worth more than 259.6 cents in savings.  
D) Half of all coupons were worth more than \$0.39 in savings.
- 3) The output below displays the mean and median for the state high school dropout rates in 1998 and 2002.

	Drop 1998	Drop 2002
N	51	51
MEAN	28.92	26.19
MEDIAN	27.49	25.92

Interpret the 2002 median dropout rate of 25.92.

- A) Half of the 51 states had a dropout rate below 25.92%.
  - B) Most of the 51 states had a dropout rate close to 25.92%.
  - C) The most frequently observed dropout rate of the 51 states was 25.92%.
  - D) Half of the 51 states had a dropout rate of 25.92%.
- 4)



For the distribution drawn here, identify the mean, median, and mode.

- A) A = mode, B = median, C = mean
- B) A = median, B = mode, C = mean
- C) A = mode, B = mean, C = median
- D) A = mean, B = mode, C = median

- 5) In a distribution that is skewed to the right, what is the relationship of the mean, median, and mode?
- A) mean > median > mode  
 B) median > mean > mode  
 C) mode > median > mode  
 D) mode > mean > median
- 6) Many firms use on-the-job training to teach their employees computer programming. Suppose you work in the personnel department of a firm that just finished training a group of its employees to program, and you have been requested to review the performance of one of the trainees on the final test that was given to all trainees. The mean of the test scores is 71. Additional information indicated that the median of the test scores was 79. What type of distribution most likely describes the shape of the test scores?
- A) skewed to the left  
 B) symmetric  
 C) skewed to the right  
 D) unable to determine with the information given
- 7) A shoe company reports the mode for the shoe sizes of men's shoes is 12. Interpret this result.
- A) The most frequently occurring shoe size for men is size 12  
 B) Most men have shoe sizes between 11 and 13.  
 C) Half of the shoes sold to men are larger than a size 12  
 D) Half of all men's shoe sizes are size 12
- 8) Which of the following is *not* a measure of central tendency?
- A) range  
 B) median  
 C) mode  
 D) mean
- 9) The distribution of salaries of professional basketball players is skewed to the right. Which measure of central tendency would be the best measure to determine the location of the center of the distribution?
- A) median  
 B) mode  
 C) mean  
 D) range

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

- 10) Parking at a university has become a problem. University administrators are interested in determining the average time it takes a student to find a parking spot. An administrator inconspicuously followed 270 students and recorded how long it took each of them to find a parking spot. The times had a distribution that was skewed to the left. Based on this information, discuss the relationship between the mean and the median for the 270 times collected.
- 11) The output below displays the mean and median for the state high school dropout rates in 1998 and 2002.

	Drop 1998	Drop 2002
N	51	51
MEAN	28.48	26.35
MEDIAN	27.17	25.54

Use the information to determine the shape of the distributions of the high school dropout rates in 1998 and 2002.

- 12) The total points scored by a basketball team for each game during its last season have been summarized in the table below. Identify the modal class of the distribution of scores.

Score	Frequency
41-60	3
61-80	8
81-100	12
101-120	7

- 13) The calculator screens summarize a data set.

```

1-Var Stats
x̄=73.65217391
Σx=1694
Σx²=138696
Sx=25.16239744
σx=24.60931018
↓n=23
  
```

```

1-Var Stats
↑n=23
minX=0
Q1=73
Med=81
Q3=90
maxX=97
  
```

- Identify the mean and the median.
- Based only on the mean and the median, do you expect that the data set is skewed to the right, symmetric, or skewed to the left? Explain.

**MULTIPLE CHOICE.** Choose the one alternative that best completes the statement or answers the question.

**Answer the question True or False.**

- The mean and the median are useful measures of central tendency for both qualitative and quantitative data.
  - True
  - False
- In a symmetric and mound shaped distribution, we expect the values of the mean, median, and mode to differ greatly from one another.
  - True
  - False
- In symmetric distributions, the mean and the median will be approximately equal.
  - True
  - False
- In skewed distributions, the mean is the best measure of the center of the distribution since it is least affected by extreme observations.
  - True
  - False
- In practice, the population mean  $\mu$  is used to estimate the sample mean  $\bar{x}$ .
  - True
  - False
- In general, the sample mean is a better estimator of the population mean for larger sample sizes.
  - True
  - False

## 2.5 Numerical Measures of Variability

### 1 Calculate Range, Variance, Standard Deviation

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

#### Solve the problem.

- 1) Each year advertisers spend billions of dollars purchasing commercial time on network television. In the first 6 months of one year, advertisers spent \$1.1 billion. Who were the largest spenders? In a recent article, the top 10 leading spenders and how much each spent (in million of dollars) were listed:

Company A	\$72.4	Company F	\$27.9
Company B	60.6	Company G	25.7
Company C	56.9	Company H	23.9
Company D	54.9	Company I	22
Company E	29	Company J	20.4

Calculate the sample variance.

- A) 379.671                      B) 3824.109                      C) 2106.264                      D) 1887.327
- 2) Calculate the range of the following data set:
- 9, 4, 5, 1, 9, 14, 9, 8, 6
- A) 13                              B) 14                              C) 1                              D) 15
- 3) The top speeds for a sample of five new automobiles are listed below. Calculate the standard deviation of the speeds.
- 140, 125, 180, 115, 105
- A) 29.2831                      B) 212.3205                      C) 151.4447                      D) 121.04

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

- 4) The ages of five randomly chosen professors are 65, 41, 48, 58, and 50. Calculate the sample variance of these ages.
- 5) The data show the total number of medals (gold, silver, and bronze) won by each country winning at least one gold medal in the 2006 Winter Olympics. Find the range, sample variance, and sample standard deviation of the numbers of medals won by these countries.

1	2	3	3	4	9	9	11	11
11	14	14	19	22	23	24	25	29

6) The calculator screens summarize a data set.

```

1-Var Stats
x̄=79.95238095
Σx=1679
Σx²=138471
Sx=14.54467666
σx=14.19415101
↓n=21
    
```

```

1-Var Stats
↑n=21
minX=30
Q1=75
Med=82
Q3=90
maxX=97
    
```

- Identify the smallest measurement in the data set.
- Identify the largest measurement in the data set.
- Calculate the range of the data set.

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

7) Calculate the variance of a sample for which  $n = 5$ ,  $\sum x^2 = 1320$ ,  $\sum x = 80$ .

- A) 10.00                      B) 8.00                      C) 326.00                      D) 3.16

8) Calculate the standard deviation of a sample for which  $n = 6$ ,  $\sum x^2 = 830$ ,  $\sum x = 60$ .

- A) 6.78                      B) 46.00                      C) 6.19                      D) 164.00

9) Compute  $s^2$  and  $s$  for the data set: -2, 1, -4, -2, -1, -2

- A) 2.67; 1.63                      B) 1.67; 1.29                      C) 19; 4.36                      D) 2; 1.41

10) Compute  $s^2$  and  $s$  for the data set:  $\frac{3}{10}, \frac{7}{10}, \frac{3}{10}, \frac{3}{10}, \frac{1}{5}, \frac{1}{10}$ .

- A) 0.042; 0.204                      B) 0.015; 0.121                      C) 0.695; 0.834                      D) 4.167; 2.041

## 2 Interpret Measures of Variability

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

### Solve the problem.

1) The range of scores on a statistics test was 42. The lowest score was 57. What was the highest score?

- A) 99                                      B) 78  
 C) 70.5                                      D) cannot be determined

2) The temperature fluctuated between a low of 73°F and a high of 89°F. Which of the following could be calculated using just this information?

- A) range                                      B) variance  
 C) standard deviation                      D) median

3) Which of the following is a measure of the variability of a distribution?

- A) range                                      B) skewness                                      C) median                                      D) sample size



SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

- 4) Various state and national automobile associations regularly survey gasoline stations to determine the current retail price of gasoline. Suppose one such national association contacts 200 stations in the United States to determine the price of regular unleaded gasoline at each station. In the context of this problem, define the following descriptive measures:  $\mu$ ,  $\sigma$ ,  $\bar{x}$ ,  $s$ .
- 5) Given the sample variance of a distribution, explain how to find the standard deviation.
- 6) Which is expressed in the same units as the original data, the variance or the standard deviation?
- 7) Which measures variability about the mean, the range or the standard deviation?
- 8) For a given data set, which is typically greater, the range or the standard deviation?

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

- 9) The total points scored by a basketball team for each game during its last season have been summarized in the table below. Which statement following the table must be true?

Score	Frequency
41-60	3
61-80	8
81-100	12
101-120	7

- A) The range is at least 41 but at most 79.
  - B) The range is 79.
  - C) The range is at least 41 but at most 120.
  - D) The range is at least 81 but at most 100.
- 10) Which number on the screen below is the sample standard deviation of the data?

```
1-Var Stats
x̄=5.8
Σx=58
Σx²=408
Sx=2.82055944
σx=2.675817632
↓n=10
```

- A) 2.82
- B) 408
- C) 5.8
- D) 2.67

Answer the question True or False.

- 11) The range is an insensitive measure of data variation for large data sets because two data sets can have the same range but be vastly different with respect to data variation.
  - A) True
  - B) False
- 12) For any quantitative data set,  $\sum(x - \bar{x}) = 0$ .
  - A) True
  - B) False



5) The calculator screens summarize a data set.

```

1-Var Stats
x̄=5.5
Σx=55
Σx²=385
Sx=3.027650354
σx=2.872281323
↓n=10

```

```

1-Var Stats
↑n=10
minX=1
Q1=3
Med=5.5
Q3=8
maxX=10

```

- Identify the mean and the sample standard deviation. Round to one place after the decimal, where necessary.
- Find the interval that corresponds to measurements within two standard deviations of the mean.

## 2 Use Empirical Rule

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Solve the problem.

- At the U.S. Open Tennis Championship a statistician keeps track of every serve that a player hits during the tournament. The statistician reported that the mean serve speed of a particular player was 97 miles per hour (mph) and the standard deviation of the serve speeds was 11 mph. Assume that the statistician also gave us the information that the distribution of the serve speeds was mound-shaped and symmetric. What proportion of the player's serves was between 108 mph and 130 mph?

- A) 0.1585                      B) 0.317                      C) 0.997                      D) 130

- The amount of television viewed by today's youth is of primary concern to Parents Against Watching Television (PAWT). 300 parents of elementary school-aged children were asked to estimate the number of hours per week that their child watches television. The mean and the standard deviation for their responses were 17 and 5, respectively. PAWT constructed a stem-and-leaf display for the data that showed that the distribution of times was a symmetric, mound-shaped distribution. Give an interval where you believe approximately 95% of the television viewing times fell in the distribution.

- A) between 7 and 27 hours per week                      B) less than 12 and more than 22 hours per week  
 C) between 2 and 32 hours per week                      D) less than 27

- A sociologist recently conducted a survey of citizens over 60 years of age who have net worths too high to qualify for Medicaid but have no private health insurance. The ages of the 25 uninsured senior citizens were as follows:

68 73 66 76 86 74 61 89 65 90 69 92 76  
 62 81 63 68 81 70 73 60 87 75 64 82

Suppose the mean and standard deviation are 74.04 and 9.75, respectively. If we assume that the distribution of ages is mound-shaped and symmetric, what percentage of the respondents will be between 64.29 and 93.54 years old?

- A) approximately 81.5%                      B) approximately 68%  
 C) approximately 95%                      D) approximately 84%





- 2) By law, a box of cereal labeled as containing 36 ounces must contain at least 36 ounces of cereal. The machine filling the boxes produces a distribution of fill weights with a mean equal to the setting on the machine and with a standard deviation equal to 0.03 ounce. To ensure that most of the boxes contain at least 36 ounces, the machine is set so that the mean fill per box is 36.09 ounces. Assuming nothing is known about the shape of the distribution, what can be said about the proportion of cereal boxes that contain less than 36 ounces.
- A) The proportion is at most 11%.  
 B) The proportion is at least 89%.  
 C) The proportion is at most 5.5%.  
 D) The proportion is less than 2.5%.
- 3) A study was designed to investigate the effects of two variables — (1) a student's level of mathematical anxiety and (2) teaching method — on a student's achievement in a mathematics course. Students who had a low level of mathematical anxiety were taught using the traditional expository method. These students obtained a mean score of 320 with a standard deviation of 40 on a standardized test. Assuming no information concerning the shape of the distribution is known, what percentage of the students scored between 240 and 400?
- A) at least 75%  
 B) approximately 95%  
 C) at least 89%  
 D) approximately 68%
- 4) A study was designed to investigate the effects of two variables — (1) a student's level of mathematical anxiety and (2) teaching method — on a student's achievement in a mathematics course. Students who had a low level of mathematical anxiety were taught using the traditional expository method. These students obtained a mean score of 430 with a standard deviation of 30 on a standardized test. Assuming a non-mound-shaped distribution, what percentage of the students scored over 520?
- A) at most 11%  
 B) approximately 2.5%  
 C) at least 89%  
 D) at most 5.5%
- 5) A recent survey was conducted to compare the cost of solar energy to the cost of gas or electric energy. Results of the survey revealed that the distribution of the amount of the monthly utility bill of a 3-bedroom house using gas or electric energy had a mean of \$103 and a standard deviation of \$15. If nothing is known about the shape of the distribution, what percentage of homes will have a monthly utility bill of less than \$73?
- A) at most 25%  
 B) at least 75%  
 C) at most 11.1%  
 D) at least 88.9%
- 6) Many firms use on-the-job training to teach their employees computer programming. Suppose you work in the personnel department of a firm that just finished training a group of its employees to program, and you have been requested to review the performance of one of the trainees on the final test that was given to all trainees. The mean and standard deviation of the test scores are 82 and 5, respectively. Assuming nothing is known about the distribution, what percentage of test-takers scored above 97?
- A) at most 11%  
 B) at least 89%  
 C) approximately 0.15%  
 D) approximately 99.85%
- 7) If nothing is known about the shape of a distribution, what percentage of the observations fall within 2 standard deviations of the mean?
- A) at least 75%  
 B) at most 25%  
 C) approximately 95%  
 D) approximately 5%

- 8) Fill in the blank. \_\_\_\_\_ gives us a method of interpreting the standard deviation of any data set, regardless of the shape of the distribution.
- A) Chebyshev's Rule  
B) The Empirical Rule  
C) both A and B  
D) neither A nor B
- 9) Fill in the blank. \_\_\_\_\_ is a method of interpreting the standard deviation of data that have a mound-shaped, symmetric distribution.
- A) The Empirical Rule  
B) Chebyshev's Rule  
C) both A and B  
D) neither A nor B
- 10) Given a data set, which of the following is most likely to be the percentage of data within three standard deviations of the mean?
- A) 95%                      B) 65%                      C) 70%                      D) 85%

**Answer the question True or False.**

- 11) Both Chebyshev's rule and the empirical rule guarantee that no data item will be more than four standard deviations from the mean.
- A) True                                      B) False
- 12) Chebyshev's rule applies to qualitative data sets, while the empirical rule applies to quantitative data sets.
- A) True                                      B) False
- 13) Chebyshev's rule applies to large data sets, while the empirical rule applies to small data sets.
- A) True                                      B) False
- 14) Your teacher announces that the scores on a test have a mean of 83 points with a standard deviation of 4 points, so it is reasonable to expect that you scored at least 70 on the test.
- A) True                                      B) False

## 2.7 Numerical Measures of Relative Standing

### 1 Compute, Interpret z-Score

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

**Solve the problem.**

- 1) Many firms use on-the-job training to teach their employees computer programming. Suppose you work in the personnel department of a firm that just finished training a group of its employees to program, and you have been requested to review the performance of one of the trainees on the final test that was given to all trainees. The mean and standard deviation of the test scores are 74 and 3, respectively, and the distribution of scores is mound-shaped and symmetric. Suppose the trainee in question received a score of 72. Compute the trainee's z-score.
- A)  $z = -0.67$                       B)  $z = -2$                       C)  $z = -6$                       D)  $z = 0.93$

- 2) A recent survey was conducted to compare the cost of solar energy to the cost of gas or electric energy. Results of the survey revealed that the distribution of the amount of the monthly utility bill of a 3-bedroom house using gas or electric energy had a mean of \$138 and a standard deviation of \$12. Three solar homes reported monthly utility bills of \$101, \$92, and \$99. Which of the following statements is true?
- A) Homes using solar power may have lower utility bills than homes using only gas and electricity.  
 B) The utility bills for homes using solar power are about the same as those for homes using only gas and electricity.  
 C) Homes using solar power may actually have higher utility bills than homes using only gas and electricity.  
 D) Homes using solar power always have lower utility bills than homes using only gas and electricity.
- 3) A radio station claims that the amount of advertising each hour has a mean of 14 minutes and a standard deviation of 1.3 minutes. You listen to the radio station for 1 hour and observe that the amount of advertising time is 16 minutes. Calculate the z-score for this amount of advertising time.
- A)  $z = 1.54$                       B)  $z = -1.54$                       C)  $z = 2.6$                       D)  $z = 1.05$
- 4) On a given day, gasoline prices in the state of Colorado had a mean price of \$2.28/gallon with a standard deviation of \$0.09. A particular Colorado gas station had gasoline for \$2.19/gallon. Interpret the z-score for this gas station.
- A) The gas price of this station falls 1 standard deviation below the mean gas price of all Colorado stations.  
 B) The gas price of this station falls 1 standard deviation above the mean gas price of all Colorado stations.  
 C) The gas price of this Colorado station falls 9 standard deviations below the mean gas price of all Colorado stations.  
 D) The gas price of this Colorado station falls 9 standard deviations above the mean gas price of all Colorado stations.
- 5) Which of the following is a measure of relative standing?
- A) z-score                      B) mean                      C) variance                      D) pie chart

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

- 6) A study was designed to investigate the effects of two variables — (1) a student's level of mathematical anxiety and (2) teaching method — on a student's achievement in a mathematics course. Students who had a low level of mathematical anxiety were taught using the traditional expository method. These students obtained a mean score of 420 and a standard deviation of 20 on a standardized test. Find and interpret the z-score of a student who scored 490 on the standardized test.
- 7) A recent survey was conducted to compare the cost of solar energy to the cost of gas or electric energy. Results of the survey revealed that the distribution of the amount of the monthly utility bill of a 3-bedroom house using gas or electric energy had a mean of \$112.00 and a standard deviation of \$15.00. Assuming the distribution is mound-shaped and symmetric, would you expect to see a 3-bedroom house using gas or electric energy with a monthly utility bill of \$217.00? Explain.

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

- 8) Find the z-score for the value 91, when the mean is 59 and the standard deviation is 1.
- A)  $z = 32.00$                       B)  $z = 31.00$                       C)  $z = -1.53$                       D)  $z = 1.53$





## 2 Find, Interpret Percentile

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

**Solve the problem.**

- 1) When Scholastic Achievement Test scores (SATs) are sent to test-takers, the percentiles associated with scores are also given. Suppose a test-taker scored at the 84th percentile on the verbal part of the test and at the 25th percentile on the quantitative part. Interpret these results.
- A) This student performed better than 84% of the other test-takers on the verbal part and better than 25% on the quantitative part.
  - B) This student performed better than 84% of the other test-takers on the verbal part and better than 75% on the quantitative part.
  - C) This student performed better than 16% of the other test-takers on the verbal part and better than 75% on the quantitative part.
  - D) This student performed better than 16% of the other test-takers on the verbal part and better than 25% on the quantitative part.

- 2) Summary information is given for the weights (in pounds) of 1000 randomly sampled tractor trailers.

MIN:	4013	25%:	5613
MAX:	10,613	75%:	8613
AVE:	7013	Std. Dev.:	1400

Find the percentage of tractor trailers with weights between 5613 and 8613 pounds.

- A) 50%
  - B) 75%
  - C) 25%
  - D) 100%
- 3) The test scores of 30 students are listed below. Which number could be the 30th percentile?

31 41 45 48 52 55 56 56 63 65  
67 67 69 70 70 74 75 78 79 79  
80 81 83 85 85 87 90 92 95 99

- A) 64
- B) 56
- C) 67
- D) 90

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

- 4) A retail store's customer satisfaction rating is at the 88<sup>th</sup> percentile. What percentage of retail stores has higher customer satisfaction ratings than this store?
- 5) In a summary of recent real estate sales, the median home price is given as \$325,000. What percentile corresponds to a home price of \$325,000?

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

**Answer the question True or False.**

- 6) The mean of a data set is at the 50<sup>th</sup> percentile.
- A) True
  - B) False
- 7) Percentile rankings are of practical value only with large data sets.
- A) True
  - B) False

8) The process for finding a percentile is similar to the process for finding the median.

A) True

B) False

## 2.8 Methods for Detecting Outliers (Optional)

### 1 Determine if Datum is an Outlier

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

#### Solve the problem.

1) At the U.S. Open Tennis Championship a statistician keeps track of every serve that a player hits during the tournament. The statistician reported that the mean serve speed of a particular player was 95 miles per hour (mph) and the standard deviation of the serve speeds was 8 mph. Using the  $z$ -score approach for detecting outliers, which of the following serve speeds would represent outliers in the distribution of the player's serve speeds?

Speeds: 67 mph, 103 mph, and 111 mph

A) 67 is the only outlier.

B) 67 and 103 are both outliers, but 111 is not.

C) 67, 103, and 111 are all outliers.

D) None of the three speeds is an outlier.

2) The speeds of the fastballs thrown by major league baseball pitchers were measured by radar gun. The mean speed was 87 miles per hour. The standard deviation of the speeds was 5 mph. Which of the following speeds would be classified as an outlier?

A) 103 mph

B) 77 mph

C) 95 mph

D) 82 mph

3) Which of the following statements concerning the box plot and  $z$ -score methods for detecting outliers is false?

A) The  $z$ -score method is less affected by an extreme observation in the data set.

B) The box plot method uses the quartiles as a basis for detecting outliers.

C) The  $z$ -score method uses the mean and standard deviation as a basis for detecting outliers.

D) The box plot method is less affected by an extreme observation in the data set.

4) Which of the following statements could be an explanation for the presence of an outlier in the data?

A) The measurement may be correct and from the same population as the rest but represents a rare event. Generally, we accept this explanation only after carefully ruling out all others.

B) The measurement is incorrect. It may have been observed, recorded, or entered into the computer incorrectly.

C) The measurement belongs to a population different from that from which the rest of the sample was drawn.

D) All of the above are explanations for outliers.

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

5) A radio station claims that the amount of advertising each hour has a mean of 13 minutes and a standard deviation of 2 minutes. You listen to the radio station for 1 hour and observe that the amount of advertising time is 7.4 minutes. Based on your observation, what would you infer about the radio station's claim?

- 6) The following data represent the scores of 50 students on a statistics exam. The mean score is 80.02, and the standard deviation is 11.9.

39	51	59	63	66	68	68	69	70	71
71	71	73	74	76	76	76	77	78	79
79	79	79	80	80	82	83	83	83	85
85	86	86	88	88	88	88	89	89	89
90	90	91	91	92	95	96	97	97	98

Use the z-score method to identify potential outliers among the scores.

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

**Answer the question True or False.**

- 7) The z-score uses the quartiles to identify outliers in a data set.

A) True

B) False

- 8) An outlier is defined as any observation that falls within the outer fences of a box plot.

A) True

B) False

- 9) Box plots are used to detect outliers in qualitative data sets, while z-scores are used to detect outliers in quantitative data sets.

A) True

B) False

- 10) An outlier in a data set may have a simple explanation such as a scale was not working properly or the researcher inverted the digits of a number when recording a measurement.

A) True

B) False

- 11) An outlier may be caused by accidentally including the height of a six-year-old boy in a set of data representing the heights of 12-year-old boys.

A) True

B) False

- 12) The outer fences of a box plot are three standard deviations from the mean.

A) True

B) False

## 2 Calculate Quartiles and IQR

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

**Solve the problem.**

- 1) At the U.S. Open Tennis Championship a statistician keeps track of every serve that a player hits during the tournament. The lower quartile of a particular player's serve speeds was reported to be 93 mph. Which of the following interpretations of this information is correct?

A) 75% of the player's serves were hit at speeds greater than 93 mph.

B) 25% of the player's serves were hit at 93 mph.

C) 75% of the player's serves were hit at speeds less than 93 mph.

D) 93 serves traveled faster than the lower quartile.

- 2) A sociologist recently conducted a survey of citizens over 60 years of age who have net worths too high to qualify for Medicaid but have no private health insurance. The ages of the 25 uninsured senior citizens were as follows:

68 73 66 76 86 74 61 89 65 90 69 92 76  
62 81 63 68 81 70 73 60 87 75 64 82

Find the upper quartile of the data.

- A) 81.5                                      B) 65.5                                      C) 73                                      D) 92

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

- 3) The amount of television viewed by today's youth is of primary concern to Parents Against Watching Television (PAWT). Three hundred parents of elementary school-aged children were asked to estimate the number of hours per week that their child watches television. The upper quartile for the distribution was given as 20 hours. Interpret this value.
- 4) For a given data set, the lower quartile is 45, the median is 50, and the upper quartile is 57. The minimum value in the data set is 32, and the maximum is 81.
- Find the interquartile range.
  - Find the inner fences.
  - Find the outer fences.
  - Is either of the minimum or maximum values considered an outlier? Explain.

- 5) The calculator screens summarize a data set.

```
1-Var Stats
x̄=79.95238095
Σx=1679
Σx²=138471
Sx=14.54467666
σx=14.19415101
↓n=21
```

```
1-Var Stats
↑n=21
minX=30
Q1=75
Med=82
Q3=90
maxX=97
```

- Identify the lower and upper quartiles of the data set.
- Find the interquartile range.
- Is there reason to suspect that the data may contain an outlier? Explain.



## 2.9 Graphing Bivariate Relationships (Optional)

### 1 Construct, Interpret Scatterplot

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

**Solve the problem.**

- 1) The data below represent the numbers of absences and the final grades of 15 randomly selected students from a statistics class. Construct a scattergram for the data. Do you detect a trend?

Student	Number of Absences	Final Grade as a Percent
1	5	79
2	6	78
3	2	86
4	12	56
5	9	75
6	5	90
7	8	78
8	15	48
9	0	92
10	1	78
11	9	81
12	3	86
13	10	75
14	3	89
15	11	65

- 2) The scores of nine members of a women's golf team in two rounds of tournament play are listed below.

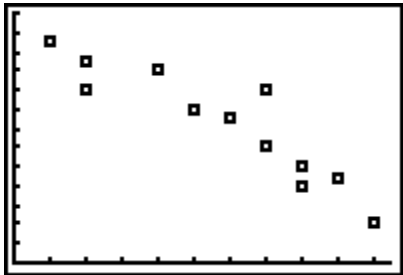
Player	1	2	3	4	5	6	7	8	9
Round 1	85	90	87	78	92	85	79	93	86
Round 2	90	87	85	84	86	78	77	91	82

Construct a scattergram for the data.

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

**Answer the question True or False.**

- 3) Scatterplots are useful for both qualitative and quantitative data.  
A) True B) False
- 4) The scatterplot below shows a negative relationship between two variables.



- A) True B) False

## 2.10 Distorting the Truth with Descriptive Techniques

### 1 Understand Misleading Statistics

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

#### Solve the problem.

- 1) Explain how stretching the vertical axis of a histogram can be misleading.
- 2) Explain how using a scale break on the vertical axis of a histogram can be misleading.
- 3) Explain how it can be misleading to draw the bars in a histogram so that the width of each bar is proportional to its height rather than have all bars the same width.
- 4) Explain how it can be misleading to report only the mean of a distribution without any measure of the variability.



## Ch. 2 Methods for Describing Sets of Data

### Answer Key

#### 2.1 Describing Qualitative Data

##### 1 Identify Classes/Compute Class Frequencies/Relative Frequencies/Percentages

- 1) A
- 2) A
- 3) free account, institutional account, account paid for personally

##### 2 Construct Frequency/Relative Frequency Table

- 1) A
- 2) A
- 3)

Color	Frequency
Green	3
Blue	7
Brown	5
Orange	2
Red	3

- 4) B

##### 3 Construct, Interpret Bar Graph

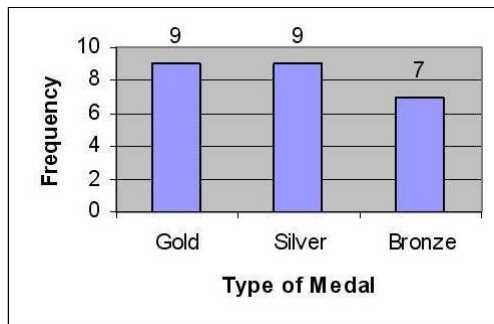
- 1) A
- 2) A
- 3) a.

Medal	Frequency
Gold	9
Silver	9
Bronze	7

- b.

Medal	Relative Frequency
Gold	.36
Silver	.36
Bronze	.28

- c.



- 4) A
- 5) A

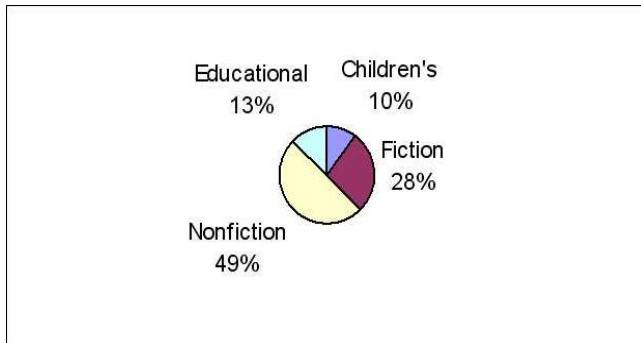
##### 4 Construct, Interpret Pie Chart

- 1) A

2) a.

Type of Book	Relative Frequency
Children's	.10
Fiction	.28
Nonfiction	.49
Educational	.13

b.



3) A

4) B

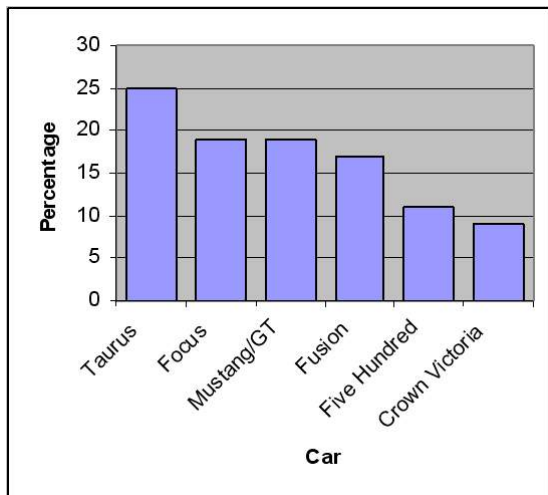
**5 Construct, Interpret Pareto Diagram**

1) In a Pareto diagram, the bars are arranged by height in a descending order from left to right.

2) a.

Car	Relative Frequency
Crown Victoria	.09
Five Hundred	.11
Taurus	.25
Fusion	.17
Focus	.19
Mustang/GT	.19

b.

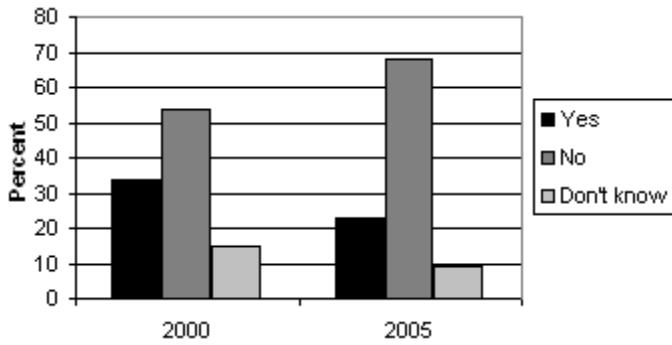


3) B

4) B

## 6 Construct, Interpret Side-by-Side Bar Chart

1)



Losses due to employee theft have decreased from 2000 to 2005.

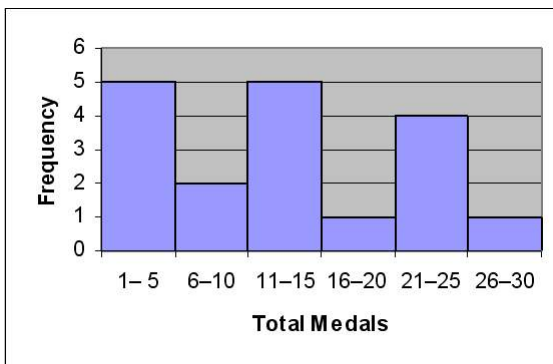
## 2.2 Graphical Methods for Describing Quantitative Data

### 1 Construct, Interpret Histogram

1) a.

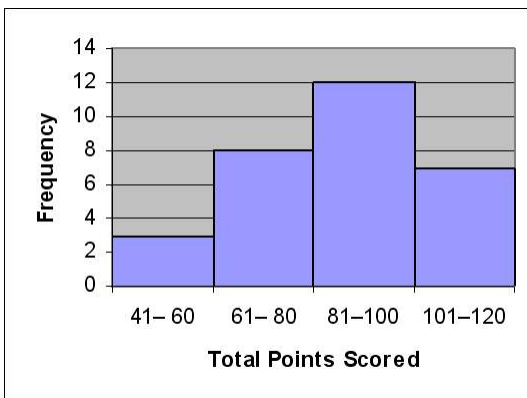
Total Medals	Frequency
1-5	5
6-10	2
11-15	5
16-20	1
21-25	4
26-30	1

b.



2) a. The exact scores would be needed to construct a stem-and-leaf display but the exact scores are not available in the table given.

b.



3) A

4) A

5) B

## 2 Construct, Interpret Stem-and-Leaf Display

1) A

2)

Stem	Leaf
1	5
2	
3	
4	
5	1 5
6	6 9
7	6 7 9
8	2 3 5 7 8 8 9
9	1 2 4 8 8

3) B

## 3 Construct, Interpret Dot-Plot

1) A

2) A

3) A

4) A

## 2.3 Summation Notation

### 1 Use Summation Notation

1) A

2) A

3) A

4) A

5) A

6) A

## 2.4 Numerical Measures of Central Tendency

### 1 Find Mean, Median, Mode

1) A

2) A

3) A

4) The mean of the data is  $x = \frac{\sum x}{n}$

$$\frac{73 + 61.5 + 57.7 + 55.2 + 30.2 + 26.9 + 26.6 + 22.3 + 21.6 + 20.7}{10}$$

$$= \frac{395.7}{10}$$

$$= 39.57 \Rightarrow \$39.57 \text{ million}$$

The median is the average of the middle two observations.

$$M = \frac{30.2 + 26.9}{2} = 28.55 \Rightarrow \$28.55 \text{ million}$$

5) The mean is the sum of the numbers divided by 18:

$$\frac{1 + 2 + 3 + 3 + 4 + 9 + 9 + 11 + 11 + 11 + 14 + 14 + 19 + 22 + 23 + 24 + 25 + 29}{18}$$
$$= \frac{234}{18} = 13 \text{ medals.}$$

The median is the mean of the two middle numbers:  $\frac{11 + 11}{2} = 11$  medals.

The mode is the most frequent number of medals: 11 medals.

6) The mean is divided by n:

$$\frac{\sum x}{n} = \frac{196}{8} = 24.5.$$

7) a.  $n = 21$

b.  $\sum x = 1679$

c. mean:  $\bar{x} \approx 79.95$ ; median: Med=82; mode: not possible

## 2 Interpret Measures of Central Tendency

1) A

2) A

3) A

4) A

5) A

6) A

7) A

8) A

9) A

10) Since the distribution is skewed to the left, we know that the median time will exceed the mean time.

11) In both 1998 and 2002, the mean dropout rates exceed the median dropout rates. This indicates that both the 1998 and 2002 high school dropout rates have distributions that are skewed to the right.

12) The modal class is the class with the greatest frequency: 81–100 points.

13) a. mean:  $\bar{x} \approx 73.65$ ; median: Med=81

b. We expect the data to be skewed to the left because the mean is less than the median.

14) B

15) B

16) A

17) B

18) B

19) A

## 2.5 Numerical Measures of Variability

### 1 Calculate Range, Variance, Standard Deviation

1) A

2) A

3) A

$$4) s^2 = \frac{\sum(x - \bar{x})^2}{n - 1}$$

$$\bar{x} = \frac{\sum x}{n} = \frac{65 + 41 + 48 + 58 + 50}{5} = 52.4$$

$$s^2 = \frac{(65 - 52.4)^2 + (41 - 52.4)^2 + (48 - 52.4)^2 + (58 - 52.4)^2 + (50 - 52.4)^2}{5 - 1}$$

$$= 86.30$$

5) The range is  $29 - 1 = 28$  medals.

$$\text{The variance is } s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n - 1} = \frac{4372 - \frac{(234)^2}{18}}{17} = \frac{1330}{17} \approx 78.24$$

$$\text{The standard deviation is } s = \sqrt{s^2} = \sqrt{\frac{1330}{17}} \approx 8.85$$

- 6) a.  $\min X = 30$   
 b.  $\max X = 97$   
 c.  $97 - 30 = 67$

- 7) A  
 8) A  
 9) A  
 10) A

## 2 Interpret Measures of Variability

- 1) A  
 2) A  
 3) A  
 4)  $\mu$  is the mean price of the regular unleaded gasoline prices of all retail gas stations in the United States.

$\sigma$  is the standard deviation of the regular unleaded gasoline prices of all retail gas stations in the United States.

$\bar{x}$  is the mean price of the regular unleaded gasoline prices collected from the 200 stations sampled.

$s$  is the standard deviation of the regular unleaded gasoline prices collected from the 200 stations sampled.

- 5) Take the square root of the sample variance to find the sample standard deviation.  
 6) standard deviation  
 7) standard deviation  
 8) range  
 9) A  
 10) A  
 11) A  
 12) A  
 13) B  
 14) B  
 15) A

## 2.6 Interpreting the Standard Deviation

### 1 Construct, Interpret Intervals About the Mean

- 1) A  
 2) A  
 3) A  
 4) measurements within three standard deviations of the mean

5) a. mean:  $\bar{x} = 5.5$ ; sample standard deviation:  $S_x \approx 3.0$

b.  $(5.5 - 2 \times 3.0, 5.5 + 2 \times 3.0) = (-.5, 11.5)$

## 2 Use Empirical Rule

1) A

2) A

3) A

4) A

5) A

6) A

7) A

8) A

9) We use the Empirical Rule to determine the percentage of serves with speeds faster than 77 mph. We do this by first finding the percentage of serves with speeds between 77 and 95 mph. The Empirical Rule states that approximately 34.0% (68%/2) fall between 77 and 95 mph. Because the distribution is symmetric about the mean speed of 95 mph, we know 50% of the serve speeds were faster than 95 mph. We add these findings together to determine that  $34.0\% + 50\% = 84.0\%$  of the serves were hit faster than 77 mph.

10) The value 80 falls one standard deviation above the mean in the distribution. Using the Empirical Rule, 68% of the days will have between 70 and 80 jobs submitted. Of the remaining 32% of the days, half, or  $32\%/2 = 16\%$ , of the days will have more than 80 jobs submitted.

11) The value of 20 ounces falls three standard deviations below the mean. The Empirical Rule states that approximately all of the boxes will contain cereal amounts between 20.00 ounces and 20.18 ounces. Therefore, approximately 100% of the boxes contain at least 20 ounces.

12) The Empirical Rule states that 95% of the data will fall between 73 and 85. Because the distribution is symmetric, half of the remaining 5%, or 2.5%, will have test scores above 85. Thus, 85 is the cutoff point that will identify the trainees who will receive the promotion.

13) 74% of the scores lie within one standard deviation of the mean, 96% within two standard deviations, and 98% within three standard deviations. These percentages are close to those given in the Empirical Rule, so the distribution is roughly mound-shaped and symmetric, though obviously skewed slightly to the left.

14) A

## 3 Use Chebyshev's Rule

1) A

2) A

3) A

4) A

5) A

6) A

7) A

8) A

9) A

10) A

11) B

12) B

13) B

14) A

## 2.7 Numerical Measures of Relative Standing

### 1 Compute, Interpret z-Score

1) A

2) A

3) A

4) A

5) A

6) The z-score is  $z = \frac{x - \mu}{\sigma}$ .

For a score of 49,  $z = \frac{490 - 420}{20} = 3.50$ .

This student's score falls 3.50 standard deviations above the mean score of 420.

7) The z-score for the value \$217.00 is:

$$z = \frac{x - \bar{x}}{s} = \frac{217 - 112}{15} = 7$$

An observation that falls 7 standard deviations above the mean is very unlikely. We would not expect to see a monthly utility bill of \$217.00 for this home.

8) A

9) history z-score = -2.00; physics z-score = 3.78; The student performed better on the physics test.

10) highest:  $z = 1.51$ ; lowest:  $z = -3.45$

11) The value of  $x$  lies 2.5 standard deviations below the mean.

12) mean: 65; standard deviation: 5

13) A

14) A

15) A

## 2 Find, Interpret Percentile

1) A

2) A

3) A

4) 12%

5) 50<sup>th</sup> percentile

6) B

7) A

8) A

## 2.8 Methods for Detecting Outliers (Optional)

### 1 Determine if Datum is an Outlier

1) A

2) A

3) A

4) D

5) The z-score for the value 7.4 is -2.8

Since the z-score would not indicate that 7.4 minutes represents an outlier, there is no evidence that the station's claim is incorrect.

6) The z-score of 39 is -3.46. Since this z-score is less than -3, the score of 39 is an outlier. All other scores have z-scores between -3 and 3, so there are no other outliers.

7) B

8) B

9) B

10) A

11) A

12) B

### 2 Calculate Quartiles and IQR

1) A

2) A

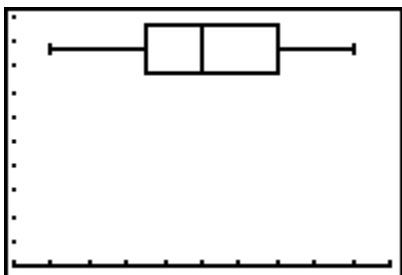
3) 75% of the TV viewing times are less than 20 hours per week. 25% of the times exceed 20 hours per week.



- 4) a. The interquartile range is  $57 - 45 = 12$ .
- b. The inner fences are  $45 - 1.5(12) = 27$  and  $57 + 1.5(12) = 75$ .
- c. The outer fences are  $45 - 3(12) = 9$  and  $57 + 3(12) = 93$ .
- d. The maximum of 81 is a potential outlier since it lies outside the inner fences. The minimum is within the inner fence and is not considered to be an outlier.
- 5) a. lower quartile:  $Q1=75$ ; upper quartile:  $Q3=90$
- b. interquartile range:  $90 - 75 = 15$
- c. Yes; the smallest measurement, 30, is three times the interquartile range less than the lower quartile, so it is a suspected outlier.

### 3 Construct, Interpret Boxplot

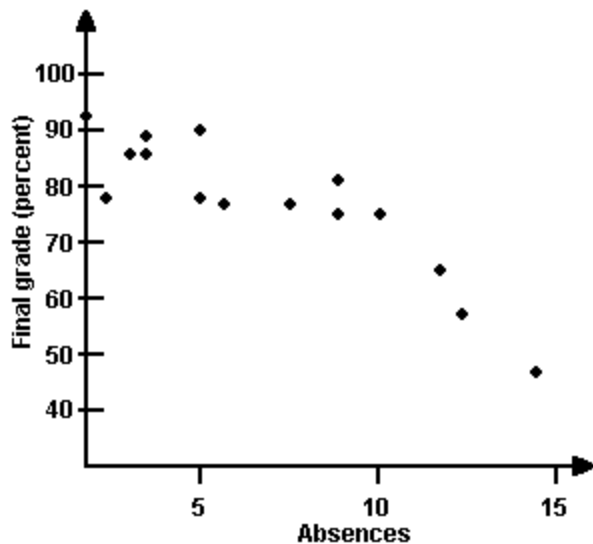
- 1) A
- 2) a. The lower quartile is 73, the upper quartile is 89, and the median is 81.
- b. The interquartile range is  $89 - 73 = 16$ . The score of 39 is a potential outlier since it is less than  $73 - 1.5(16) = 49$ .
- c. No scores fall outside the outer fences, 25 and 137. Only the score of 39 lies between the inner and outer fences.
- 3) The horizontal axis extends from 10 to 20, with each tick mark representing one unit.



## 2.9 Graphing Bivariate Relationships (Optional)

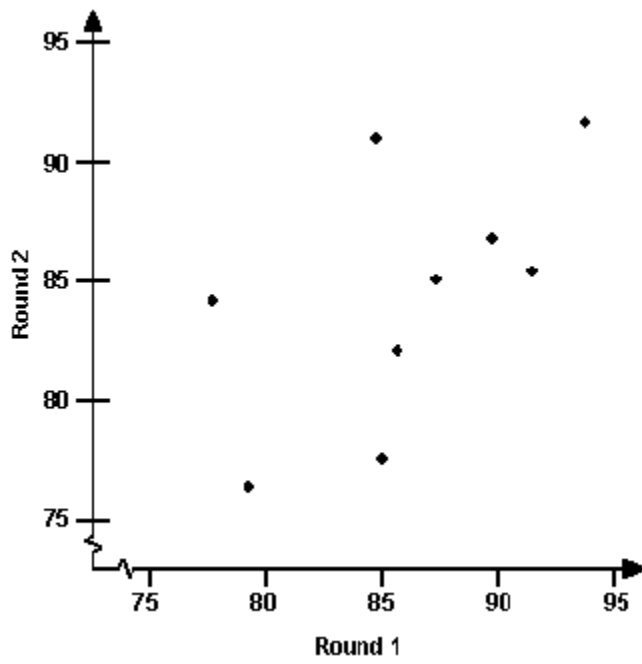
### 1 Construct, Interpret Scatterplot

1)



There appears to be a trend in the data. As the number of absences increases, the final grade decreases.

2)



3) B

4) A

## 2.10 Distorting the Truth with Descriptive Techniques

### 1 Understand Misleading Statistics

- 1) Stretching the vertical axis may overemphasize the differences in the heights of the bars making the taller bars look much taller than the shorter bars.
- 2) Using a scale break on the vertical axis may make the shorter bars look disproportionately shorter than the taller bars.
- 3) The reader may think that the area of the bar represents the quantity rather than the height of the bar, giving a disproportionate emphasis on the taller bars.
- 4) When comparing means from two different distributions, the difference between them may be insignificant if the variability in one or both of the distributions is large.