SOLUTION

The z-score for the marketing test is $z = \frac{86 - 74}{10} = 1.20$

The z-score for the management test is $z = \frac{94 - 82}{11} \approx 1.09$

On the marketing test you scored 1.20 standard deviations above the mean, compared to only 1.09 standard deviations above the mean for the management test. Even though the raw score on the management test is larger than the raw score on the marketing test, relative to the means of the data sets, the performance on the marketing test was slightly better. Once again, changing the scale of the data has beneficial effects. It enables the comparison of two measurements that are drawn from different populations.

If a z-score is negative, the data value is less than the mean. Conversely, if the z-score is positive, the data value is greater than the mean. The z-score is also a unit-free measure. That is, regardless of the original units of measurement (whether the data are measured in centimeters, meters, or kilometers), an observation's z-score will be the same.

4.3 Exercises

Basic Concepts

1. What are two methods for describing relative position?
2. If a data value is calculated to be the 72nd percentile, what does this mean?
3. Describe how to find the percentile of a particular data value.
4. What are quartiles? Are they equivalent to percentiles? If so, how?
5. What is the interquartile range? What does it measure?
6. What are the advantages of using a box plot to display a data set?
7. What are the key calculations needed in order to construct a box plot?
8. What is an outlier? How can outliers be identified?
9. What is a z-score? Why is it useful?

Exercises

10. The following test scores were recorded for an economics final examination.

Test Scores														
60	81	100	44	90	56	71	42	64	100	69	80	90	87	94
41	78	100	50	96	77	61	38	41	68	50	69	85	47	86

  a. Calculate the 20th percentile.
  b. Calculate the 95th percentile.
  c. Interpret the meaning of each of these percentiles.
  d. Determine the percentile rank for the student who scored 56.
  e. Determine the percentile rank for the student who scored 80.

11. Copiers Etc. collects data on the number of copiers sold each day by each salesperson. The number of copiers sold for each salesperson for a small office on a randomly selected day is listed below.

Numbers of Copiers Sold											
1	5	2	3	7	6	1	0	0	3	4	5

a. Calculate the 25th percentile.

b. Calculate the 90th percentile.

c. Interpret the meaning of each of these percentiles.

d. Determine the percentile rank for the salesperson who sold 5 copiers.

e. Determine the percentile rank for the salesperson who sold 1 copier.

12. Subjects in a marketing study were shown a film and at the end of the film were given a test to measure their recall. The scores are listed in the following table.

Test Scores														
97	31	61	49	61	85	35	57	31	26	27	40	86	78	28
61	87	62	92	58	38	95	81	68	64	72	45	57	84	100

a. Calculate Q_1, the first quartile.

b. Calculate Q_2, the second quartile.

c. Calculate Q_3, the third quartile.

d. Explain the meaning of these quartiles in the context of the marketing study.

e. Calculate the interquartile range.

f. Construct a box plot for the test scores. Are there any outliers?

g. Compute the z-score for a test score of 81.

h. Compute the z-score for a test score of 62.

i. Explain what the z-scores in parts **g.** and **h.** are measuring.

13. A baseball recruiter is interested in 20 perspective players. He goes to several games and determines the batting average for each player. The batting averages are displayed in the following table.

Batting Averages									
.330	.260	.180	.150	.200	.400	.020	.190	.290	.200
.170	.150	.250	.270	.320	.280	.270	.220	.270	.300

a. Calculate Q_1, the first quartile.

b. Calculate Q_2, the second quartile.

c. Calculate Q_3, the third quartile.

d. Explain the meaning of these quartiles in the context of the batting averages.

e. Calculate the interquartile range.

f. Construct a box plot for the batting averages. Are there any outliers? (Guess which player is the pitcher.)

g. Compute the z-score for a batting average of .020.

h. Compute the z-score for a batting average of .330.

i. Explain what the z-scores in parts **g.** and **h.** are measuring.

j. Determine the percentile rank for a player who had a batting average of .270.

k. Determine the percentile rank for a player who had a batting average of .150.

14. Consider a set of data in which the sample mean is 64 and the sample standard deviation is 21. For the following specific values, calculate the *z*-score and interpret the results.

a. $x = 80$ **b.** $x = 64$ **c.** $x = 40$

15. A statistics student scored a 75 on the first exam of the semester and an 82 on the second exam of the semester. The average score and standard deviation of scores for the two exams are given in the following table. On which exam did the student perform relatively better?

Test Scores

	First Exam	Second Exam
μ	74	85
σ	10	7

16. A hospital measures babies' heights when they are born in both inches and centimeters. Eight baby girls are randomly selected and the following heights are recorded in both inches and centimeters.

Newborn Heights

Baby	1	2	3	4	5	6	7	8
Inches	17.75	18.50	19.25	19.75	20.25	20.50	20.50	20.75
Centimeters	45.09	46.99	48.90	50.17	51.44	52.07	52.07	52.71

a. Calculate the mean height in inches and centimeters for the baby girls.

b. Calculate the standard deviation of the heights of baby girls in both inches and centimeters.

c. Calculate the *z*-score for the height of Baby Girl 3 measured in inches.

d. For Baby Girl 3, calculate the *z*-score for the height measured in centimeters.

e. Consider the *z*-scores calculated in parts **c.** and **d.** Are the *z*-scores as you expected them to be? Explain.

4.4 Data Subsetting

Data subsetting is used to provide more clarity and structure to the data. Referring to the histogram in Example 4.2.6, we know that the data consists of tuition and fees of two-year and four-year institutions. The histogram below (Figure 4.4.1) depicts the same data in Example 4.2.6 but has more intervals to give us a better spread of the data. Due to the large number of observations between \$3,750 and \$5,000, the histogram appears to be right-skewed. Were the histogram bell-shaped, it would be easier to identify the center of the data. Thus, it is sometimes prudent to separate the tuition data (i.e., data subsetting) into two groups—tuition for two-year institutions and tuition for four-year institutions.