19.5 teacher notes

Objectives

- Calculate deviation from the mean.
- Calculate mean absolute deviation.
- Determine how many data points lie within one or two mean absolute deviations.

Vocabulary

Statistics (Lesson 19-1)
Mean (Lesson 19-1)
Deviation from the mean
Mean absolute deviation
Absolute value (Lesson 1-2)
Dispersion

Prerequisites

Finding the mean of a data set

Graphing points on a number line

Finding the distance between points on a number line

Calculating the absolute value

Get Started

- Set up a hypothetical scenario concerning students and part-time employment. Assign the students the letters A, B, C, D, E or use names of students in the class or use names offered by the class. Student A has a salary of \$2 per hour. Students B, C, D, and E have salaries of \$4, \$5, \$7, and \$7 per hour, respectively. Make a chart on the board listing the names in a vertical column and the salary in a second vertical column next to the names, so that a name and the corresponding amount are on the same line. Labels may be put at the top of the columns. Label the first column “students” and the second column “salary.”
- Ask students to find the mean of the salaries. Remind students the mean is determined by adding the data values and dividing the sum by the number of values. The correct response is the mean of the salaries is \$5.
- Ask the students how each salary compares to the mean of \$5. Compare each salary to the mean and record the difference in a third column next to the salary column. Student A’s salary is \$3 less than the mean, so record negative three in the third column next to the \$2

salary. Do the same for each of the salaries. The third column will contain differences of –3, –1, 0, 2, and 2. Notice that the third student has a salary equal to the mean. Explain to students the recorded numbers represent how much each salary differed, or deviated, from the mean. Label the third column "deviation from the mean." Tell students this lesson further explains deviation.

- Ask students to add the deviations. The sum will be zero.
- Repeat the exercise using six students with salaries of $1, $2, $2, $5, $6, and $8. The mean is $4; the deviations from the mean are –3, –2, –2, 1, 2, and 4. Notice that there is no salary equal to the mean. The sum of the deviations will again be zero.
- Through discussion and leading questions, students should conclude the sum of the deviations will always be equal to zero.

Section 1

Expand Their Horizons

In this lesson, students will examine deviation from the mean and will learn to compute mean absolute deviation. Recall that mean, median, and mode are measures of central tendency in statistics. Range, interquartile range, and mean absolute deviation are measures of dispersion or variability. They describe how data is spread out in the set.

In real world applications, measures of dispersion are important. Consider a company that makes batteries. The company president might be pleased to hear the mean life of one of his company's batteries is 15 hours. However, he might not be pleased to hear that half the batteries last 29 hours and half the batteries last only one hour, even though the mean life is 15 hours. He may very well expect a more consistent performance by ***all*** the batteries. The president would need more information than just the mean to have an accurate picture of the performance of the batteries as a whole. He would need information provided by measures of dispersion. One measure of dispersion is the mean absolute deviation.

A deviation from the mean for a data value is found by subtracting the mean from that data value. Every data value below the mean has a negative deviation, and every data value above the mean has a positive deviation. Because the mean is the average of the data values, the sum of all the deviations will always be zero. In fact, adding the deviations to get zero is a good way for students to check their calculations. Most textbooks on statistics use the symbol $\bar{x}$ (said "x bar") to denote the mean of the x-values in a data set. Thus, a deviation from the mean is indicated by $(x_i - \bar{x})$, where x_i is a value in the data set.

The mean absolute deviation of a data set describes the dispersion or spread of the data. It is the average of the absolute values of the distances of all the data values from the mean. If the data values are clustered close together, the distances from the mean are small, and the mean absolute deviation is small. If the data values are spread out in a wide dispersion, some of the distances from the mean are large, and the mean absolute deviation is large. To find the mean absolute deviation, the absolute value of each deviation is found; the absolute values are added; and the sum is divided by the number of data values. In effect, the mean absolute deviation is the average of the distances from the mean, where all of the distances are made nonnegative by using absolute values.

Consider the two following sets of test scores. For convenience, the scores are placed in order of least to greatest.

Class A of 10 students: {46, 54, 66, 66, 72, 85, 89, 94, 98, 100}

Class B of 10 students: {46, 73, 76, 77, 78, 79, 79, 80, 82, 100}

Both classes have a mean of 77, a maximum of 100, a minimum of 46, and a range of 54 (found by subtracting the lowest score, 46, from the highest score, 100). However, the data in the sets are spread out differently. By examining the tables below, it is apparent that the minimum and maximum scores in Class B are outliers. (An outlier is significantly less or greater than most of the other data values).

CLASS A

Score x	Deviation $x - \bar{x}$	Absolute Deviation $\lvert x - \bar{x} \rvert$
46	–31	31
54	–23	23
66	–11	11
66	–11	11
72	–5	5
85	8	8
89	12	12
94	17	17
98	21	21
100	23	23
Sum: 770	0	162

Mean = 77

Mean Absolute Deviation = 16.2

CLASS B

Score x	Deviation $x - \bar{x}$	Absolute Deviation $\lvert x - \bar{x} \rvert$
46	–31	31
73	–4	4
76	–1	1
77	0	0
78	1	1
79	2	2
79	2	2
80	3	3
82	5	5
100	23	23
Sum: 770	0	72

Mean = 77

Mean Absolute Deviation = 7.2

In each table, the first column contains the scores, denoted as x. In both classes, the sum of the scores is 770. When that sum is divided by the 10 (the number of scores) a mean of 77 is obtained. The second column in each table gives the deviation from the mean, denoted as $(x - \bar{x})$; the mean of 77 is subtracted from each score to get these numbers. As previously mentioned, the sum of the deviations is always zero. Note that many of the deviations in Class A are much greater than those in Class B. To the teacher of these classes, this information might reveal that all but one member of Class B had a fairly solid understanding of the material being tested, but that several members of Class A did not understand the material.

The third column in each table gives the absolute value of each deviation, denoted as $|x - \bar{x}|$. The mean of the numbers in this column is the mean absolute deviation. In Class A, the mean absolute deviation is 16.2. In Class B, the mean absolute deviation is 7.2. The average distance of the Class A scores from the mean is 16.2, and the average distance of the Class B scores from the mean is 7.2. Many of the scores in Class A are more widely dispersed than the scores in Class B. This dispersion would not be obvious by examining only the measures of central tendency or by the range.

Number of points scored in a basketball game by players who played for 30 minutes or more: 24, 15, 10, 8, 3

1 **Find the deviations from the mean.** First, the mean of the data set must be found. To find the mean, add the values and divide by the number of values in the set. The sum of the values is 60, and there are five values in the set, so the mean is $60 \div 5 = 12$. The deviations from the mean are found by subtracting the mean from each of the values. The deviations are $24 - 12 = 12$, $15 - 12 = 3$, $10 - 12 = -2$, $8 - 12 = -4$, and $3 - 12 = -9$. To check the deviations from the mean, be sure their sum is zero: $12 + 3 - 2 - 4 - 9 = 0$. (This does not ***prove*** that all the deviations are correct, but it does give some assurance).

2 **Find the mean absolute deviation.** The deviations from the mean are 12, 3, –2, –4, and –9. The mean absolute deviation is found by adding the absolute values of the deviations and dividing that sum by the number of data values. The absolute values of the deviations are 12, 3, 2, 4, and 9. The sum of these is 30, and the number of data values is five; since $30 \div 5 = 6$, the mean absolute deviation is six.

3 **Determine the number of players who scored within one mean absolute deviation of the mean.** The players that scored within one mean absolute deviation of the mean scored within six points of the mean. The mean is 12, so to find the scores that are within six points of 12, subtract six from 12 and add six to twelve: $12 - 6 = 6$ and $12 + 6 = 18$. The scores that are within six points of 12 are the scores between six and 18. The scores between six and 18 are 15, 10, and 8. The absolute values of the deviations of 15, 10, and 8 from the mean are three, two, and four, respectively. Therefore, there are three players who scored within one mean absolute deviation of the mean.

Connections

Measures of dispersion are commonly used in industry for quality control purposes. For example, a tire manufacturer may want to know the average life span of a tire. With this information, the manufacturer can determine the mean life expectancy. A tire with a small mean absolute deviation means the tire is consistent, and the average life expectancy is close to the mean.

Look Beyond

Among statisticians, variance and standard deviation are commonly used measures of dispersion. The variance is the mean of the squares of the deviations (or in some applications, an approximate mean). The standard deviation is the square root of the variance.

Additional Examples

1. The monthly electric bills (rounded to the nearest dollar) for a family last year are \$97, \$88, \$96, \$89, \$95, \$150, \$169, \$221, \$204, \$199, \$139, and \$97. Find the mean dollar amount for a monthly bill and find the deviations from the mean.

The mean is found by adding the values and dividing by the number of values. The sum of the bills is \$1,644 which, when divided by 12 months, yields a mean dollar amount of \$137. The mean is then subtracted from each score to find each deviation. See the last column in the table below for the deviations. Note that the sum of the deviations is zero.

Month	Bill x	Deviation $x - \bar{x}$
Jan	\$97	–\$40
Feb	\$88	–\$49
Mar	\$96	–\$41
Apr	\$89	–\$48
May	\$95	–\$42
Jun	\$150	\$13
Jul	\$169	\$32
Aug	\$221	\$84
Sep	\$204	\$67
Oct	\$199	\$62
Nov	\$139	\$2
Dec	\$97	–\$40
	\$1,644	\$0

2. Using the data set from Additional Examples Question 1, find the mean absolute deviation. What does it reveal?

The absolute value of each deviation must be used. The absolute values are shown in the table below.

Month	Absolute Deviation $\lvert x - \bar{x} \rvert$
Jan	\$40
Feb	\$49
Mar	\$41
Apr	\$48
May	\$42
Jun	\$13
Jul	\$32
Aug	\$84
Sep	\$67
Oct	\$62
Nov	\$2
Dec	\$40
	\$520

Find the average of the absolute values to find the mean absolute deviation. The sum of the absolute deviations is \$520, and $520 \div 12 \approx 43.33$. The mean absolute deviation, rounded to the nearest penny, is \$43.33. This reveals a fairly wide dispersion of the bills, possibly explained by the use of air conditioning in the summer months.