

### 3.9: Putting It All Together- SD and 3 M's

Consider the histogram of a frequency distribution in Figure 3.9.1.

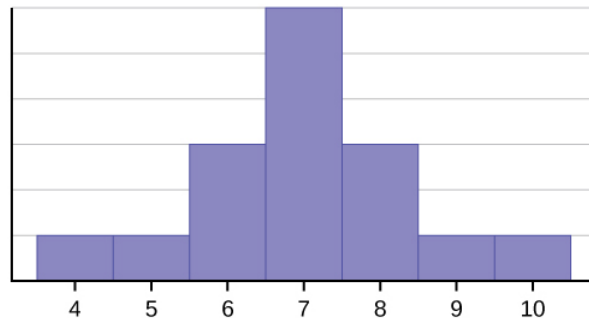


Figure 3.9.1. Example Symmetrical Histogram (Ciara Kidder of Marian University using work by [Barbara Illowsky & Susan Dean \(De Anza College\)](#) from [OpenStax](#))

The histogram displays a symmetrical distribution of data. A distribution is symmetrical if a vertical line can be drawn at some point in the histogram such that the shape to the left and the right of the vertical line are mirror images of each other. The mean, the median, and the mode are each seven for these data. In a perfectly symmetrical distribution, the mean and the median are the same. This example has one mode (unimodal), and the mode is the same as the mean and median. In a symmetrical distribution that has two modes (bimodal), the two modes would be different from the mean and median.

The histogram for the data in Figure 3.9.2 is not symmetrical. The right-hand side seems "chopped off" compared to the left side. A distribution of this type is called negatively skewed because the tail is pulled out to the left.

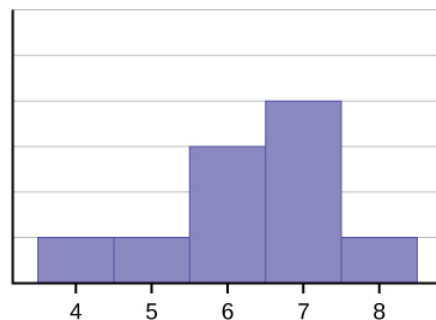


Figure 3.9.2- Example Negatively Skewed Histogram (Ciara Kidder of Marian University using work by [Barbara Illowsky & Susan Dean \(De Anza College\)](#) from [OpenStax](#))

#### ? Exercise 3.9.1

The data used to create Figure 3.9.2 is: 4, 5, 6, 6, 6, 7, 7, 7, 7, and 8. What is the mean? What is the median? What is the mode?

#### Answer

The mean is 6.3. The median is 6.5. The mode is 7.

(We cannot include units in these because this is example data; we don't know what was measured.)

Notice that the mean is less than the median, and they are both less than the mode. Distributions of data are negatively skewed when the mean is lower than the median. The extreme low scores bring the mean down, but their extremity doesn't affect the median (or mode).

The histogram in Figure 3.9.3 is also not symmetrical, but is positively skewed as the tail points to the right.

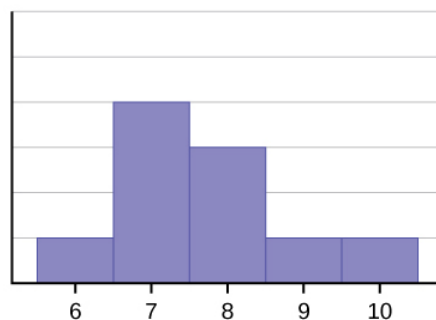


Figure 3.9.3- Example Positively Skewed Histogram (Ciara Kidder of Marian University using work by [Barbara Illowsky & Susan Dean](#) (De Anza College) from [OpenStax](#))

### ? Exercise 3.9.3

The data used to create Figure 3.9.3 is: 6, 7, 7, 7, 7, 8, 8, 8, 9, and 10. What is the mean? What is the median? What is the mode?

#### Answer

The mean is 7.7. The median is 7.5. The mode is 7.

(We cannot include units in these because this is example data; we don't know what was measured.)

In contrast to the negatively skewed distribution in Figure 3.9.2, the positively skewed distribution's mean is larger than the median. Looking at Figure 3.9.3, you can see how the few high scores pulled the mean up, but extreme scores do not affect what number is in the middle (the median).

*Generally, if the distribution of data is skewed to the left, the mean is less than the median, which is often less than the mode. If the distribution of data is skewed to the right, the mode is often less than the median, which is less than the mean.*

Skewness and symmetry become important when we discuss probability distributions in later chapters.

### 📌 Note

Do you think that the standard deviation is bigger in the symmetrical distribution shown in Figure 3.9.1 or the skewed distribution in Figure 3.9.2? Let's see!

First, let's work together to practice calculating a standard deviation using the data from Figure 3.9.1.

### ✓ Example 3.9.1

The data from Figure 3.9.1 is in Table 3.9.1. Use the table to find the Sum of Squares, then finish the standard deviation formula for a sample.

#### Solution

Table 3.9.1- Finding of the Sums of Squares

X	$X - \bar{X}$	Sums of Squares $((X - \bar{X})^2)$
4	$4 - \bar{X} = -3$	$(X - \bar{X})^2 = -3^2 = 9$
5	$5 - \bar{X} = -2$	$(X - \bar{X})^2 = -2^2 = 4$
6	$6 - \bar{X} = -1$	$(X - \bar{X})^2 = -1^2 = 1$
6	$6 - \bar{X} = -1$	$(X - \bar{X})^2 = -1^2 = 1$

X	$X - \bar{X}$	Sums of Squares $((X - \bar{X})^2)$
6	$6 - \bar{X} = -1$	$(X - \bar{X})^2 = -1^2 = 1$
7	$7 - \bar{X} = -0$	$(X - \bar{X})^2 = 0^2 = 0$
7	$7 - \bar{X} = -0$	$(X - \bar{X})^2 = 0^2 = 0$
7	$7 - \bar{X} = -0$	$(X - \bar{X})^2 = 0^2 = 0$
7	$7 - \bar{X} = -0$	$(X - \bar{X})^2 = 0^2 = 0$
7	$7 - \bar{X} = -0$	$(X - \bar{X})^2 = 0^2 = 0$
7	$7 - \bar{X} = -0$	$(X - \bar{X})^2 = 0^2 = 0$
8	$8 - \bar{X} = 1$	$(X - \bar{X})^2 = 1^2 = 1$
8	$8 - \bar{X} = 1$	$(X - \bar{X})^2 = 1^2 = 1$
8	$8 - \bar{X} = 1$	$(X - \bar{X})^2 = 1^2 = 1$
9	$9 - \bar{X} = 2$	$(X - \bar{X})^2 = 2^2 = 4$
10	$10 - \bar{X} = 3$	$(X - \bar{X})^2 = 3^2 = 9$
<b><math>\Sigma = 112</math></b>	<b><math>\Sigma = 0</math></b>	<b><math>\Sigma = 32</math></b>

The last column (the one on the right) is the squared deviations. So, when they are summed to 32, you have the Sum of Squares.

$$s = \sqrt{\frac{SS}{df}} = \sqrt{\frac{\sum(X - \bar{X})^2}{N - 1}}$$

As this suggests, you then divide the Sum of Squares by N-1 (N is the number of scores), so:

$$N - 1 = 16 - 1 = 15$$

Then,

$$s = \sqrt{\frac{SS}{df}} = \sqrt{\frac{\sum(X - \bar{X})^2}{N - 1}} = \frac{32}{15} = 2.1\bar{3}$$

And the final step of square rooting to get it back into whatever units we started with:

$$s = \sqrt{\frac{SS}{df}} = \sqrt{\frac{\sum(X - \bar{X})^2}{N - 1}} = \sqrt{2.1\bar{3}} = 1.460593$$

The standard deviation of the data from Figure 3.9.1 is 1.46!

Now, try it on your own!

### ? Exercise 3.9.2

Compute a standard deviation with the data from Figure 3.9.2 that is in Table 3.9.2.

Table 3.9.2- Your Turn to Find the Sums of Squares

X	$X - \bar{X}$	Sums of Squares $((X - \bar{X})^2)$
4		
5		

X	$X - \bar{X}$	Sums of Squares $((X - \bar{X})^2)$
6		
6		
6		
7		
7		
7		
7		
8		
$\Sigma = 63$	$\Sigma = 0$	$\Sigma = ?$

$$s = \sqrt{\frac{SS}{df}} = \sqrt{\frac{\sum (X - \bar{X})^2}{N - 1}}$$

### Answer

The standard deviation for the data from Figure 3.9.2 is 1.16 (rounded from 1.159502).

If you didn't get something close to this, check that you used the correct N (number of scores) for 10 in this sample. Another thing to check is that your Sum of Squares should be 12.1; often, simple adding or messing up the mean will wreck the rest of your calculations!

So what did we learn from Example 3.9.1 and Exercise 3.9.2? The skewed distribution had a smaller standard deviation than the symmetrical distribution. Why might that be? Maybe looking at the range will help? The range of the symmetrical distribution was 6 ( $10 - 4 = 6$ ), while the range of the skewed distribution was 4. Having a smaller range means that there was no extreme scores, so that could explain the lower standard deviation.

### Quiz Yourself!

You can use the following to test yourself to see if you understand the relationships between the measures of central tendency (mean, median, and mode) and the measures of variability (range and standard deviation) in relation to the shape of distributions of data.

#### Exercise 3.9.3

When the data are symmetrical, what is the typical relationship between the mean and median?

#### Answer

When the data are symmetrical, the mean and median are close or the same.

#### Exercise 3.9.4

Describe the shape of this distribution.


 This is a histogram which consists of 5 adjacent bars with the x-axis split into intervals of 1 from 3 to 7. The bar heights peak at the first bar and taper lower to the right.

Figure 3.9.4- Example Histogram (Ciara Kidder of Marian University using work by [Barbara Illowsky & Susan Dean](#) (De Anza College) from [OpenStax](#))

#### Answer

The distribution is skewed right because it looks pulled out to the right.

### Exercise 3.9.5

Describe the relationship between the mode and the median of this distribution.


 This is a histogram which consists of 5 adjacent bars with the x-axis split into intervals of 1 from 3 to 7. The bar heights peak in the middle and taper down to the right and left.

Figure 3.9.5- - Example Histogram (Ciara Kidder of Marian University using work by [Barbara Illowsky & Susan Dean \(De Anza College\)](#) from [OpenStax](#))

#### Answer

The mode and the median are the same. In this case, they are both five.

### Exercise 3.9.6

Describe the shape of this distribution.


 This is a histogram which consists of 5 adjacent bars over an x-axis split into intervals of 1 from 3 to 7. The bar heights from left to right are: 1, 1, 2, 4, 7.

Figure 3.9.6- Example Histogram (Ciara Kidder of Marian University using work by [Barbara Illowsky & Susan Dean \(De Anza College\)](#) from [OpenStax](#))

#### Answer

The distribution is skewed left because it looks pulled out to the left.

### Exercise 3.9.7

Which is the greatest, the mean, the mode, or the median of the data set?

11; 11; 12; 12; 12; 12; 13; 15; 17; 22; 22; 22

#### Answer

The mode is 12, the median is 12.5, and the mean is 15.1. The mean is the largest.

## Putting It All Together: SD & 3 M's

Let's look at everything that we know about the data for Figure 3.9.2.

First, we know the measures of central tendency. These tell us where the "center" of the distribution might be. For Figure 3.9.2, the sample mean was 6.3, the median was 6.5, and the mode was 7. This suggests that the center of the distribution of data was probably somewhere between 6.3 and 7. Even without seeing the histogram, I might think that the distribution was skewed because the mean was smaller than the median. It wasn't that much smaller (0.2), so I wouldn't bet my life on that one.

Second, we know some measures of variability. The range is 4 (the highest score of 8 minus the lowest score of 4), and the standard deviation was 1.16. Compared to a mean of 6.3, a standard deviation of 1.16 might be considered medium or slightly large. A medium standard deviation would suggest a bell-shaped curve, while a larger standard deviation would suggest a wide and flat distribution or a skewed distribution.

Finally, I can look at the actual Figure 3.9.2. That shows a negatively skewed distribution. I can find where the mode, median, and mean would be on Figure 3.9.2, and look to see what scores might fall within one standard deviation above and below the mean.

## Contributors and Attributions

- Barbara Illowsky and Susan Dean (De Anza College) with many other contributing authors. Content produced by OpenStax College is licensed under a Creative Commons Attribution License 4.0 license. Download for free at <http://cnx.org/contents/30189442-699...b91b9de@18.114>.
- Dr. MO (Taft College)

This page titled [3.9: Putting It All Together- SD and 3 M's](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Michelle Oja](#).

- **2.7: Skewness and the Mean, Median, and Mode** by OpenStax is licensed CC BY 4.0. Original source: <https://openstax.org/details/books/introductory-statistics>.