

# Defining Data Variability

by Sophia



## WHAT'S COVERED

This lesson discusses measures of data spread, or variability. You will be able to recognize what changes in these measures indicate about your data set. Specifically, this lesson covers:

1. [Variability](#)
2. [Standard Deviation](#)

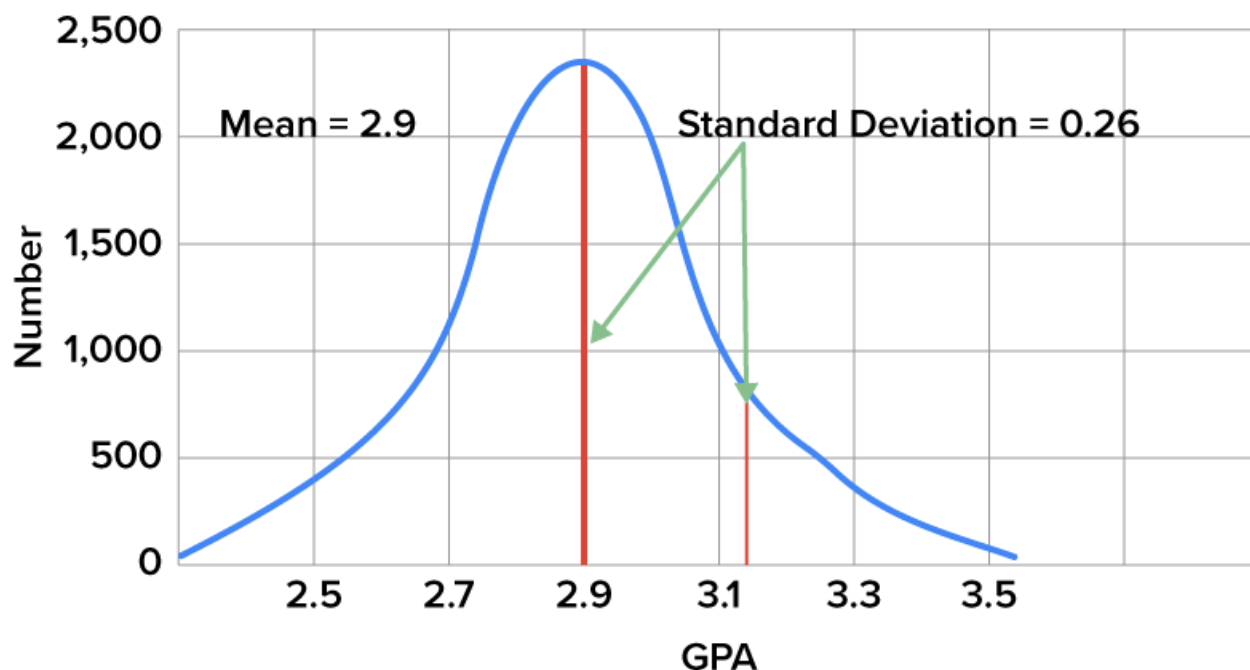
## 1. Variability

In addition to finding the center of a data set, we may also be interested in finding a number that tells us how far the data is spread out from the mean. Like finding range and interquartile, measuring how spread out data is from the mean can determine a different measure of variability.

If the variability in the data is small, the mean will generally serve as a good estimate of a typical value in the data set. If the variability in the data is large, the mean is generally not a good estimate of a typical value in the data set.

Two measures of variability that can be used are **variance** and **standard deviation**. Take a look at a sample of high school grade point averages drawn from a group of recent graduates. By knowing the standard deviation (0.26) and the sample mean (2.9) for this specific sample, you can get a good sense of how the data is distributed.

## High School GPA



### TERMS TO KNOW

#### Variance

Takes into account how widespread the data may be.

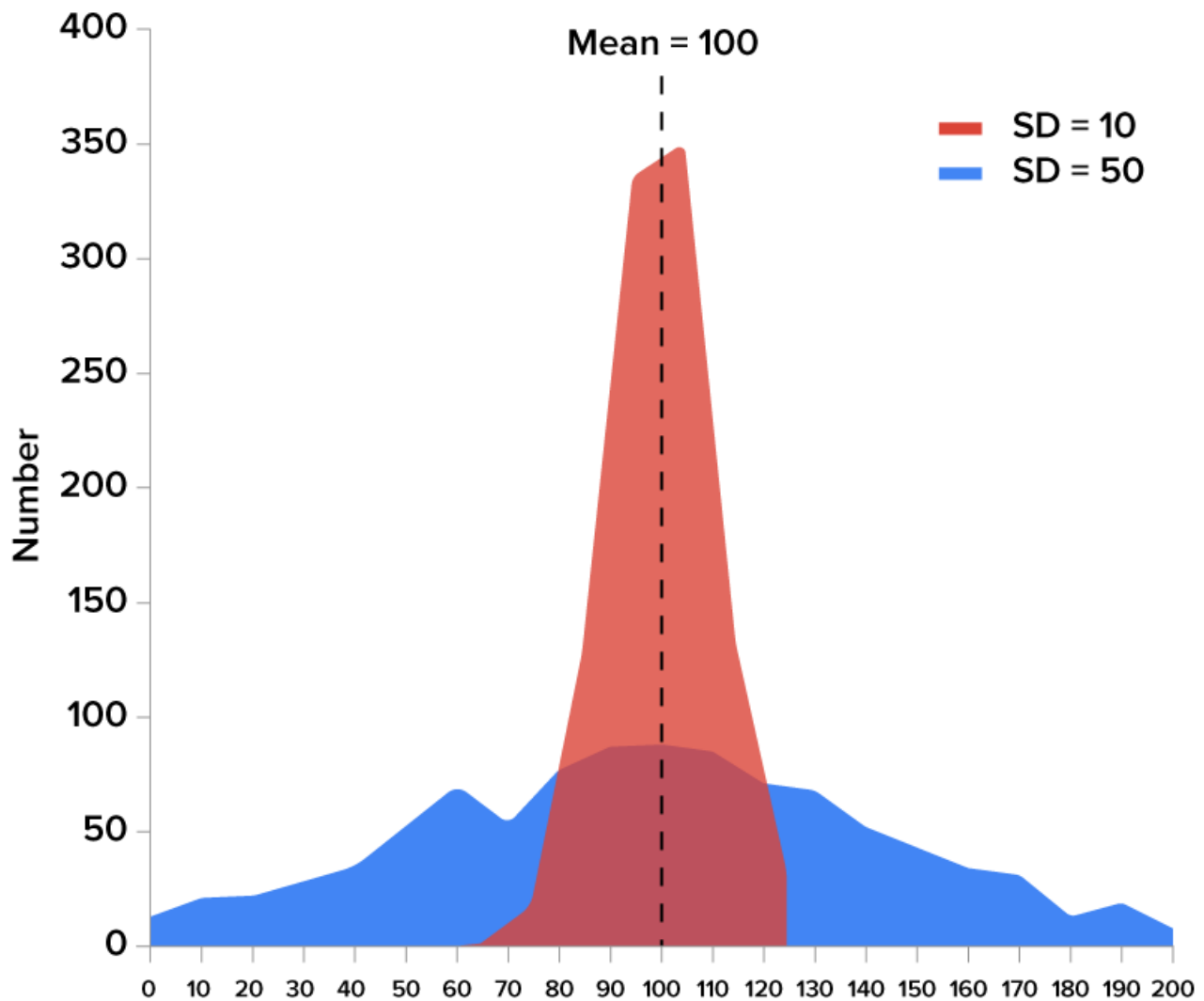
#### Standard Deviation

Indicates how closely the data is distributed around the mean.

## 2. Standard Deviation

By knowing the variability of a data set in terms of variance and standard deviation, you can determine what percentage of data falls within a certain range of values. You will learn more about this in later lessons. Knowing measures of variability allows you to gain important insights about your data.

In the example below, notice that both distributions have the same means (average) and number of values, but very different standard deviations (SD). The red distribution has a standard deviation of 10, and the blue distribution has a standard deviation of 50. The standard deviation provides you with an indication of how closely the data is distributed from the mean.



#### THINK ABOUT IT

Suppose you had delivery time data from all your local pizza restaurants.

Would you prefer to order pizza from a restaurant with a small or a large standard deviation in delivery time?

+

A restaurant with a smaller standard deviation would deliver your pizza within a more predictable time, and a restaurant with a larger standard deviation would be more variable. Keep in mind standard deviation tells you nothing about the mean (average) delivery time.



#### SUMMARY

In this lesson, you learned that **standard deviation** indicates how spread out data is from the mean. Knowing just the mean and standard deviation gives you a good idea of the **variability**, or spread, in the

data set.

Source: THIS TUTORIAL WAS AUTHORED BY DAN LAUB FOR SOPHIA LEARNING. PLEASE SEE OUR [TERMS OF USE](#).



## ATTRIBUTIONS

- [Comparison Standard Deviation](#) | Author: JRBrown | License: Public Domain



## TERMS TO KNOW

### Standard Deviation

Indicates how closely the data is distributed around the mean.

### Variance

Takes into account how widespread the data may be.