Standard Deviation in Google Sheets (Sample & Population)

Authored by

Mohammed looti

November 2, 2025

Understanding Standard Deviation and Variability

The Standard Deviation (SD) is perhaps the most fundamental and widely utilized measure in the field of statistics. It serves as a critical indicator of the variability, or dispersion, present within a dataset. In essence, the standard deviation quantifies the average amount by which individual data points deviate from the central tendency, specifically the mean value. A clear comprehension of this metric is paramount for anyone involved in quantitative analysis, spanning diverse fields such as finance, quality control, and scientific research.

Interpreting the result is straightforward: a **low standard deviation** signifies that the data points are tightly clustered around the mean, suggesting consistency and predictability. Conversely, a **high standard deviation** implies that the data points are widely spread out across a broader range of values, indicating greater volatility or less reliable central tendency. This measure allows analysts to quickly gauge the risk or consistency associated with a set of observations.

When conducting data analysis using powerful, accessible tools like Google Sheets, calculating the standard deviation is highly efficient. However, successful application requires a careful distinction based on the scope of your data: are you analyzing an entire group (the population) or merely a representative portion of that group (the sample)? This distinction dictates which specific formula and function must be employed for accurate results.

The Critical Distinction: Population vs. Sample

While the goal of calculating standard deviation--measuring spread--remains constant, the underlying calculation differs based on whether your data represents the entire population or just a sample drawn from that population. Maintaining this statistical integrity is essential for drawing valid, unbiased conclusions from your analysis. Using the incorrect formula will inevitably lead to flawed inferences.

If your data set encompasses every single item or observation relevant to your study, you are working with the **Population Standard Deviation**. This calculation yields the true, definitive measure of dispersion for that specific group. However, collecting exhaustive population data is often impractical or outright impossible in real-world scenarios, such as measuring the lifetime of every light bulb produced or the income of every person in a country.

Consequently, the vast majority of statistical analyses rely on the **Sample Standard Deviation**. A

sample is a smaller, manageable subset chosen to represent the characteristics of the larger population. We use the properties observed in this sample to make reasonable estimates and inferences about the larger group. Since a sample inherently tends to exhibit less variability than the entire population from which it was drawn, a mathematical adjustment is required to prevent underestimation of the true spread.

This necessary modification is known as **Bessel's correction**. It involves adjusting the denominator in the sample formula from the sample size (n) to the sample size minus one ($n-1$). This adjustment ensures that the sample standard deviation serves as an unbiased estimator, providing a more statistically accurate representation of the population's true variability.

Calculating Population Standard Deviation ($sigma$)

The **Population Standard Deviation**, traditionally symbolized by the lowercase Greek letter sigma ($sigma$), is the definitive measure of dispersion used when the dataset comprises all members of the group under investigation. Calculating $sigma$ provides the precise measure of how much the individual values deviate from the population mean (mu).

The mathematical process involves calculating the squared difference between each data point (x_i) and the population mean (mu), summing these squared differences, dividing this total by the size of the population (N), and finally taking the square root of the result.

The formula used to calculate the population standard deviation ($sigma$) is presented below:

$sigma = sqrt{frac{sum(x_{i} - mu)^{2}}{N}}$

The components of the population formula represent specific statistical entities that must be correctly identified:

sum: This is the summation symbol, instructing you to total all the subsequent terms.
x_{i}: Represents the i^{th} individual value or observation within the complete dataset.
mu: The **population mean**, which is the arithmetic average of all values in the entire population.
N: The total **population size**, representing the total count of all data points.

Calculating Sample Standard Deviation (s)

The **Sample Standard Deviation**, denoted by the lowercase Latin letter s, is deployed when the

data has been gathered from a subset (the sample) of a much larger or potentially infinite population. Crucially, this measure provides an **estimate** of the true population standard deviation, not the exact value.

The key mathematical difference between the sample calculation and the population calculation lies in the denominator, which accounts for the degrees of freedom. Instead of dividing by the sample size (n), we divide by ($n - 1$). This adjustment, commonly referred to as the degrees of freedom adjustment or Bessel's correction, is vital. It corrects for the inherent tendency of a sample to display less variability than its parent population, thereby ensuring that s is an unbiased estimator of $sigma$.

The formula used to calculate the sample standard deviation (s) is provided below, reflecting the adjustment in the denominator:

$s = sqrt{frac{sum(x_{i} - bar{x})^{2}}{(n - 1)}}$

Understanding the variables specific to the sample calculation ensures correct application:

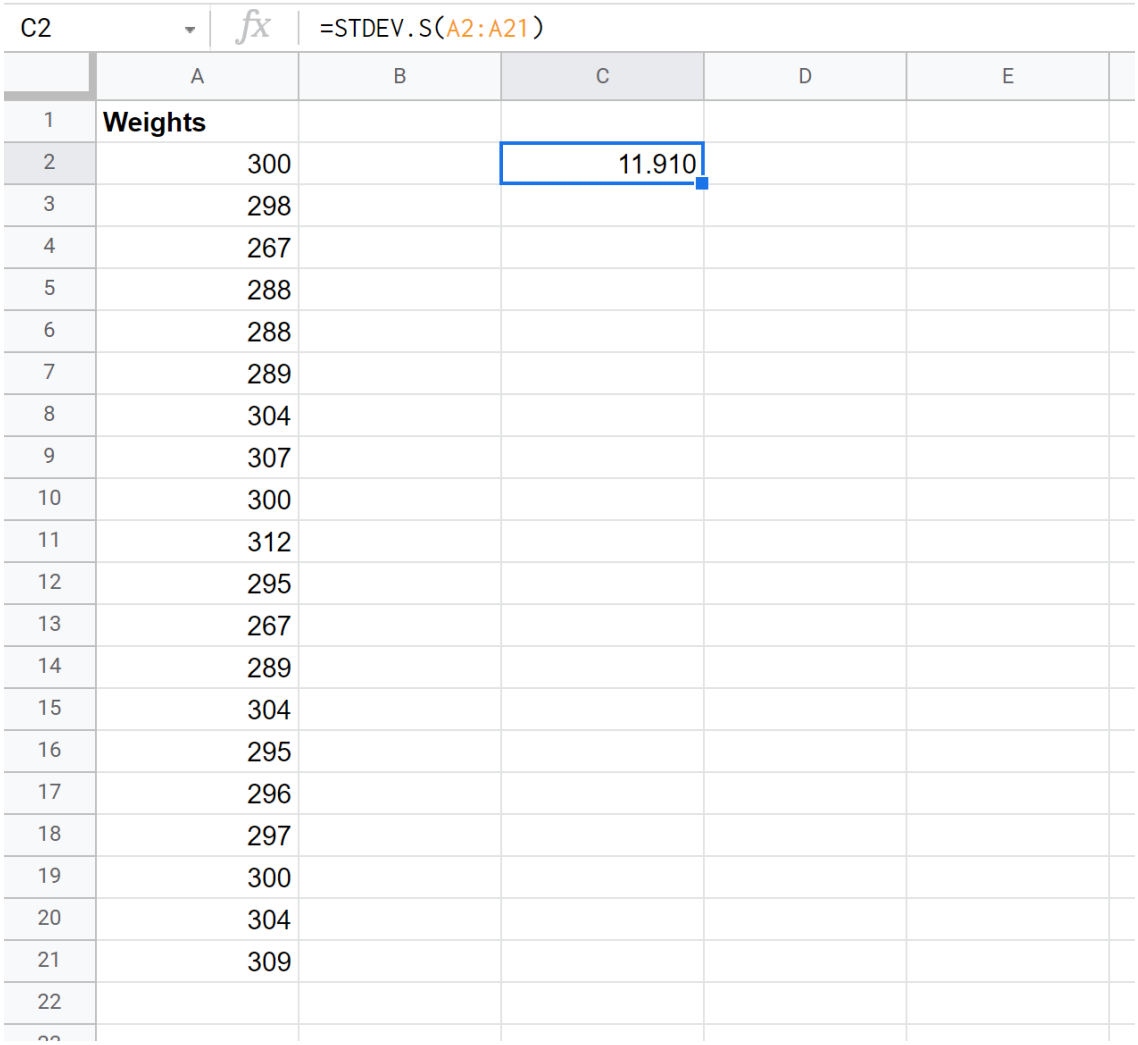
sum: The instruction to sum all the resulting squared differences between the individual values and the sample mean.
x_{i}: The i^{th} value observed in the sample taken from the population.
$bar{x}$: The **sample mean**, which is the average of the values within the specific sample dataset.
n: The **sample size**, which is the total number of observations contained in the sample.

Practical Application in Google Sheets: The Sample Function STDEV.S()

To solidify the practical application of the sample calculation, consider a marine biologist studying the dispersion of weights among a specific species of sea turtles. Given the impossibility of weighing every turtle of that species, she collects a representative sample of 20 individuals. Since this data represents only a subset of the total population, the sample standard deviation calculation is required to estimate the true population variability.

In Google Sheets, the function specifically designed for calculating the sample standard deviation is **STDEV.S()**. This function automatically performs the calculation using the degrees of freedom adjustment ($n-1$), thus providing an unbiased estimate of the population variability based on your sample data.

The procedure is straightforward: input your data into a designated range (e.g., cells A1 through A20) and then apply the function in an empty cell. The following screenshot illustrates the syntax and the resulting output when applying this function to the hypothetical turtle weight data:



Upon executing the formula `=STDEV.S(A1:A20)`, the calculated sample standard deviation is **11.91**. This result indicates that the weight of the sampled turtles deviates from the sample mean by approximately 11.91 units on average. While the older, non-suffixed function **STDEV()** is still available and performs the identical sample calculation, **STDEV.S()** is highly recommended for modern practice due to its superior clarity regarding the function's purpose.

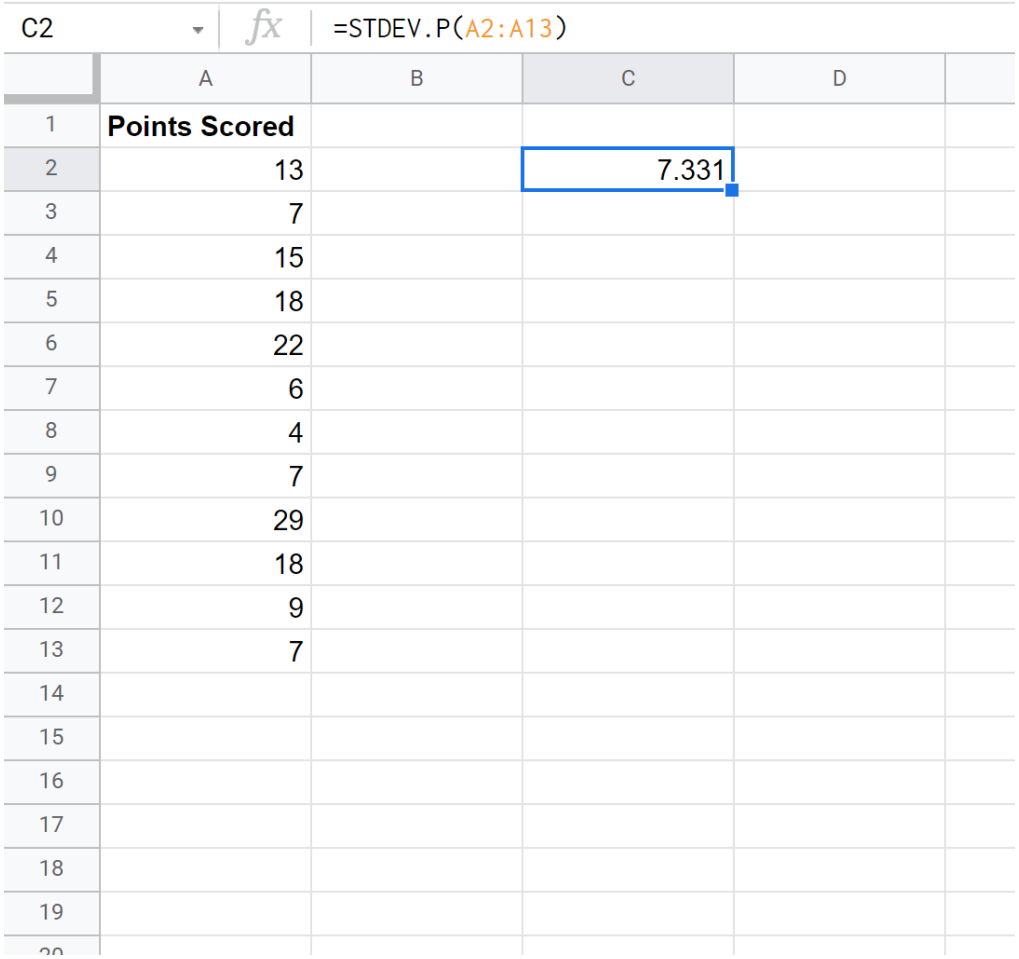
Practical Application in Google Sheets: The Population Function STDEV.P()

Conversely, we must use the population standard deviation when the data set is complete. Imagine

a basketball coach who wishes to analyze the scoring consistency solely among the twelve players currently on his team. Since his interest is confined exclusively to the variability within these specific twelve individuals--and not the entire league or hypothetical future players--his dataset constitutes the entire population of interest for this particular study.

For data sets that represent the complete group, the appropriate function in Google Sheets is **STDEV.P()**. This function calculates the true standard deviation of the provided data, performing the calculation by dividing the sum of squared differences by the total number of observations (N), without applying Bessel's correction.

To calculate the population standard deviation for the players' scores, the data is entered into a range, and the specific population function is applied. The screenshot below clearly demonstrates the use of the **STDEV.P()** function on the coach's scoring data:



The execution of the formula `=STDEV.P(B2:B13)` yields a population standard deviation of **7.331**. This figure provides the exact measure of how much the players' individual scores deviate from the team's precise average mean score. In this case, the relatively low value suggests a high degree of scoring consistency across the twelve players, which is valuable information for the coach's analysis.

Deepening Your Quantitative Analysis

Mastering the calculation and interpretation of standard deviation is a crucial foundational step in developing strong quantitative analysis skills. Understanding how data dispersion is measured is often the gateway to more complex statistical concepts. For users seeking to deepen their knowledge of statistical spread and its practical application within Google Sheets and other statistical environments, exploring related measures of variability is highly recommended.

Measures closely related to standard deviation, such as variance (which is simply the standard deviation squared) and range, provide additional context regarding data distribution and are often prerequisites for advanced statistical tests. The resources below offer further context on these related measures of spread and central tendency:

The following tutorials offer additional information about standard deviation, variance, and related concepts in statistical theory:

Understanding the fundamental difference between variance and standard deviation.
How to construct and interpret confidence intervals based on calculated standard deviation.
Exploring the concept of Z-scores and their direct relationship to standard deviation and normalization.

The following resources explain how to calculate other essential measures of spread and central tendency directly within Google Sheets functions:

Calculating the Range using MAX() and MIN() functions in Google Sheets.
Calculating the Interquartile Range (IQR) using the QUARTILE.EXC or QUARTILE.INC functions.