

Normalize Data in Google Sheets

Authored by
Mohammed looti

November 7, 2025

RECOMMENDED CITATION

Mohammed looti (2025). *Normalize Data in Google Sheets*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=12095>

The process of feature scaling, specifically known as [normalization](#) or Z-score transformation, is a cornerstone of modern [statistical analysis](#) and data preprocessing. This technique fundamentally involves rescaling a distribution of raw data points so that the resulting transformed dataset adheres to a standard distribution, possessing a central tendency or [mean](#) of 0 and a measure of spread known as the [standard deviation](#) of 1.

Normalization is not merely an optional step; it is indispensable, particularly in rigorous comparative statistics and, most critically, in the training of [machine learning](#) algorithms. Data standardization prevents variables that are measured on wildly different scales (e.g., income measured in thousands vs. age measured in years) from disproportionately influencing the final results. This comprehensive tutorial will guide you through the exact, step-by-step methodology required to achieve effective Z-score normalization using the accessible and widely utilized platform of [Google Sheets](#).

Preparing and Structuring the Raw Dataset

To properly demonstrate the mathematical mechanics of standardization, we must first organize a sample dataset within the spreadsheet environment. This initial setup is crucial for ensuring the subsequent calculations are accurate. For this illustration, assume you have a continuous array of raw numerical data points, such as student test scores, hourly temperature readings, or sales figures, meticulously arranged within a single column--for instance, Column A of your [Google Sheets](#) document.

The objective of this process is to convert these initial, raw scores into standardized scores (Z-scores). The resulting standardized scores will immediately and quantitatively reveal whether any specific data point deviates significantly above or below the dataset's overall average, thereby providing context that the raw data alone cannot offer.

Normalization is a three-part process that demands sequential execution: first, calculating the dataset's central tendency; second, measuring the data's dispersion; and third, applying the specific standardization function iteratively across all data points.

<i>fx</i>					
	A	B	C	D	
1	Data Values				
2	12				
3	14				
4	15				
5	15				
6	16				
7	17				
8	18				
9	20				
10	24				
11	25				
12	26				
13	29				
14	32				
15	34				
16	37				
17					
18					
19					
20					
21					

Calculating Essential Distribution Statistics

Before we can apply the Z-score transformation, we must determine the two fundamental statistical parameters that define the distribution of our raw data: the arithmetic mean and the standard deviation. We strongly recommend calculating and storing these resulting values in separate, clearly labeled cells within your spreadsheet (e.g., in cells C1 and C2). This practice improves transparency and simplifies the construction of the final standardization formula.

Step 1: Determine the Mean (Average). The first necessary calculation is finding the arithmetic [mean](#) (often denoted as $\bar{x}$ or μ). This value is the mathematical representation of the central point of your dataset. In Google Sheets, calculating this value is straightforward using the built-in **AVERAGE** function, which must be applied across the entire specified range of data values:

=AVERAGE(range of values)

Executing this formula on the sample data range (Column A in our example) yields the dataset's average score, providing the crucial first parameter needed for the normalization equation.

fx =AVERAGE(A2:A16)							
	A	B	C	D	E	F	G
1	Data Values						
2	12				Mean	22.26	=AVERAGE(A2:A16)
3	14						
4	15						
5	15						
6	16						
7	17						
8	18						
9	20						
10	24						
11	25						
12	26						
13	29						
14	32						
15	34						
16	37						
17							
18							
19							
20							
21							
22							

Step 2: Calculate the Standard Deviation. Next, we must quantify the degree to which individual data points are dispersed or spread out around the calculated mean. The [standard deviation](#) (s or σ) is the definitive metric used to measure this spread. Utilize the appropriate standard deviation function in Google Sheets:

=STDEV(range of values)

It is vital to select the correct function based on the nature of your data: if your data represents a statistical sample drawn from a larger population, use **STDEV.S**. If your data constitutes the entire defined population, the correct function is **STDEV.P**. The calculated standard deviation provides the 's' value, completing the necessary parameters for the [normalization](#) calculation.

fx =STDEV(A2:A16)							
	A	B	C	D	E	F	G
1	Data Values						
2	12				Mean	22.267	22.26666667
3	14				Std. Dev	7.96	=STDEV(A2:A16)
4	15						
5	15						
6	16						
7	17						
8	18						
9	20						
10	24						
11	25						
12	26						
13	29						
14	32						
15	34						
16	37						
17							
18							
19							
20							
21							

Applying the STANDARDIZE Function for Z-Scores

With the mean and [standard deviation](#) now accurately calculated and secured in designated cells, the final stage involves applying the standardization formula to every single raw data point. Fortunately, [Google Sheets](#) offers an extremely efficient, dedicated function for this operation: **STANDARDIZE**.

The **STANDARDIZE** function is designed to handle the Z-score calculation automatically. It requires three distinct positional arguments, which must be provided in the following order:

The individual data value being standardized (x).

The dataset's calculated mean.

The dataset's calculated standard deviation.

The syntax is structured precisely as follows:

=STANDARDIZE(x, mean, standard_dev)

This powerful function executes the transformation formula, efficiently calculating the Z-score for the specified value based on the established population or sample parameters.

CRITICAL NOTE ON CELL REFERENCES:

The **STANDARDIZE** function relies on the mathematical definition of the Z-score:

$$\text{Normalized value} = (x - \bar{x}) / s$$

When applying this formula to the first cell in your new column (e.g., cell B2), you will reference the individual raw data point (e.g., A2) and the fixed cells containing your statistical summaries (e.g., C1 and C2).

It is absolutely essential to use [absolute cell references](#) (denoted by dollar signs, e.g., \$C\$1 for the mean and \$C\$2 for the standard deviation). This mechanism locks the reference cells, preventing them from shifting when the formula is copied or autofilled down the entire column.

x = The individual data value (relative reference, e.g., A2).

$\bar{x}$ = The calculated [mean](#) of the entire dataset (absolute reference, e.g., \$C\$1).

s = The calculated [standard deviation](#) of the entire dataset (absolute reference, e.g., \$C\$2).

The image below illustrates the correct formula input using absolute references for standardizing the first value in our dataset:

fx	=STANDARDIZE(A2, \$F\$2, \$F\$3)					
	A	B	C	D	E	F
1	Data Values	Normalized Values				
2	12	-1.28842423	=STANDARDIZE(A2, \$F\$2, \$F\$3)		Mean	22.267
3	14				Std. Dev	7.968
4	15					
5	15					
6	16					
7	17					
8	18					
9	20					
10	24					
11	25					
12	26					
13	29					
14	32					
15	34					
16	37					
17						
18						
19						

Once the formula is correctly entered into the first standardization cell (B2), the remainder of the process is automated. Use the autofill handle feature in [Google Sheets](#): hover the cursor over the bottom-right corner of cell B2 until the cursor transforms into a small, bold plus sign. Double-click this handle, and the formula will instantaneously be copied down the column, completing the normalization process for the entire dataset. Every value in this new column is now standardized, forming a distribution centered at zero.

	A	B	C	D
1	Data Values	Normalized Values		
2	12	-1.288		
3	14	-1.037		
4	15	-0.912		
5	15	-0.912		
6	16	-0.786		
7	17	-0.661		
8	18	-0.535		
9	20	-0.284		
10	24	0.218		
11	25	0.343		
12	26	0.469		
13	29	0.845		
14	32	1.221		
15	34	1.472		
16	37	1.849		
17				
18				
19				
20				

Interpreting the Normalized Data (Z-Scores)

The resulting normalized values are formally recognized as [Z-scores](#) or standard scores. These scores provide the analyst with immediate, powerful, and context-dependent insight into where each original data point sits relative to the entire distribution. The primary utility of the Z-score is that it transforms raw data into a universal language of statistical distance.

The magnitude and algebraic sign of the Z-score clearly indicate how far an original data point deviates from the [mean](#), with the distance quantified in units of the [standard deviation](#).

A **positive Z-score**, such as +1.5, signifies that the data point is 1.5 standard deviations above the

mean.

A **negative Z-score**, such as -0.75, signifies that the data point is 0.75 standard deviations below the mean.

A Z-score approaching **0** indicates that the data point is very close to the mean of the distribution.

Consider the interpretation of the raw data point "12" from our sample dataset, which resulted in a normalized Z-score of -1.288:

	A	B	C	D
1	Data Values	Normalized Values		
2	12	-1.288		
3	14	-1.037		
4	15	-0.912		
5	15	-0.912		
6	16	-0.786		
7	17	-0.661		
8	18	-0.535		
9	20	-0.284		
10	24	0.218		
11	25	0.343		
12	26	0.469		
13	29	0.845		
14	32	1.221		
15	34	1.472		
16	37	1.849		
17				
18				
19				
20				

This score was derived using the following calculation:

$$\text{Normalized value} = (x - \text{Mean}) / \text{SD} = (12 - 22.267) / 7.968 = \mathbf{-1.288}$$

The interpretation is unambiguous: the value "12" is **1.288** standard deviations below the mean (22.267) of the original dataset. Conversely, if a score of 35 normalized to 1.598, it would be interpreted as 1.598 standard deviations above the average. This sophisticated yet simple technique permits analysts to compare variables, even if those variables were initially measured in entirely disparate units (e.g., liters vs. kilograms), by effectively placing them onto a single,

standardized, and comparable scale.

Further Resources for Data Transformation Techniques

While [Google Sheets](#) provides an expedient and accessible method for standardization, the core statistical principles of [normalization](#) are universally applicable across all data analysis platforms and programming languages. Understanding the Z-score mechanism described here allows for seamless transition to more complex statistical environments.

If your analytical workflow requires data transformation within alternative software environments that demand greater computational power or advanced scripting capabilities, the following related guides may prove beneficial for adapting these techniques:

[How to Normalize Data in Excel](#)

[How to Normalize Data in R](#)