

Why is Standard Deviation Important?

Authored by
stats writer

May 4, 2024

RECOMMENDED CITATION

stats writer (2024). *Why is Standard Deviation Important?*. PSYCHOLOGICAL SCALES.
Retrieved from <https://scales.arabpsychology.com/?p=142789>

Why is Standard Deviation Important? (Explanation + Examples)

The standard deviation is used to measure the spread of values in a sample.

We can use the following formula to calculate the standard deviation of a given sample:

$$\sqrt{\sum (x_i - \bar{x})^2 / (n-1)}$$

where:

Σ : A symbol that means "sum"

x_i : The i th value in the sample

$\bar{x}$: The mean of the sample

n : The sample size

The higher the value for the standard deviation, the more spread out the values are in a sample. Conversely, the lower the value for the standard deviation, the more tightly packed together the values.

One question students often have is: *Why is the standard deviation important?*

The answer: Standard deviation is important because it tells us how spread out the values are in a given dataset.

Whenever we analyze a dataset, we're interested in finding the following metrics:

The center of the dataset. The most common way to measure the "center" is with the mean and the median.

The spread of values in the dataset. The most common way to measure spread is with the standard deviation.

By knowing where the center is located and how spread out the values are, we can gain a good understanding of the distribution of values in any dataset.

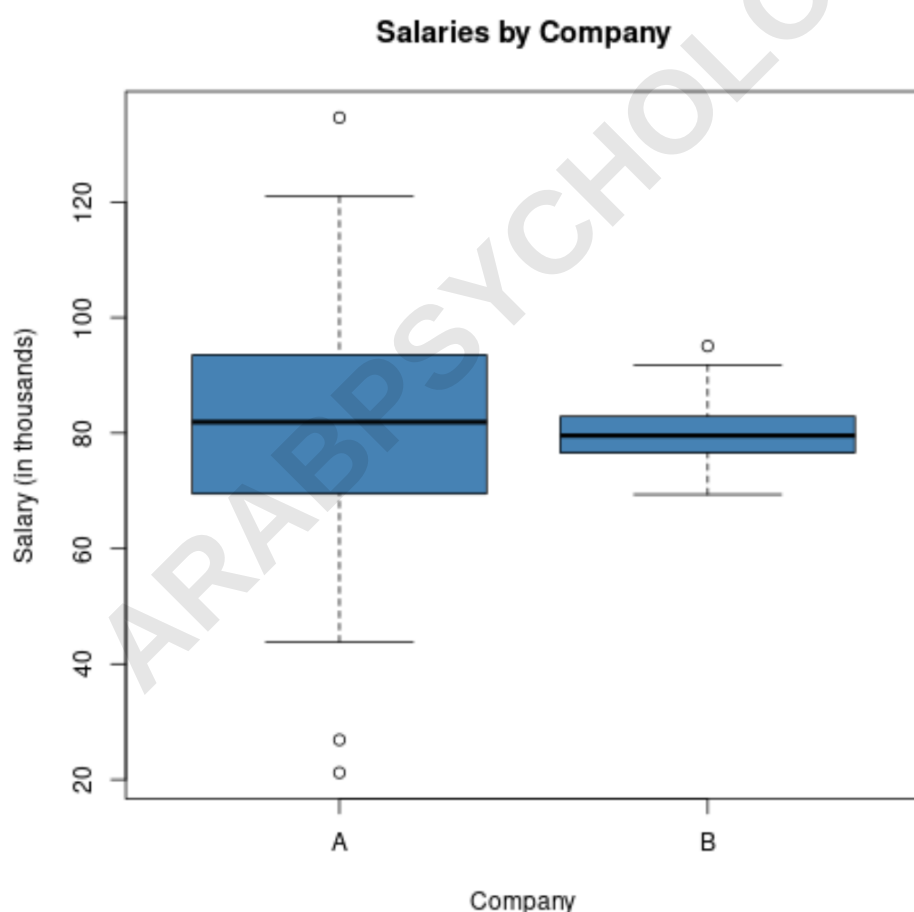
The following examples illustrate the importance of the standard deviation in practice.

Example 1: Distribution of Salaries

Suppose the mean salary at company A is \$80,000 and the standard deviation is \$20,000. Since the standard deviation is so large, there's no guarantee that you will get paid close to \$80,000 per year if you work at this company since there's such a variation in salaries.

Conversely, suppose the mean salary at company B is also \$80,000 but the standard deviation is only \$4,000. Since this standard deviation is so small, you can be sure that you'll get paid close to \$80,000 because there's very little variation in salaries.

If we created a boxplot to visualize the distribution of salaries at these two companies, it might look something like this:



Both companies have the same mean salary, but the

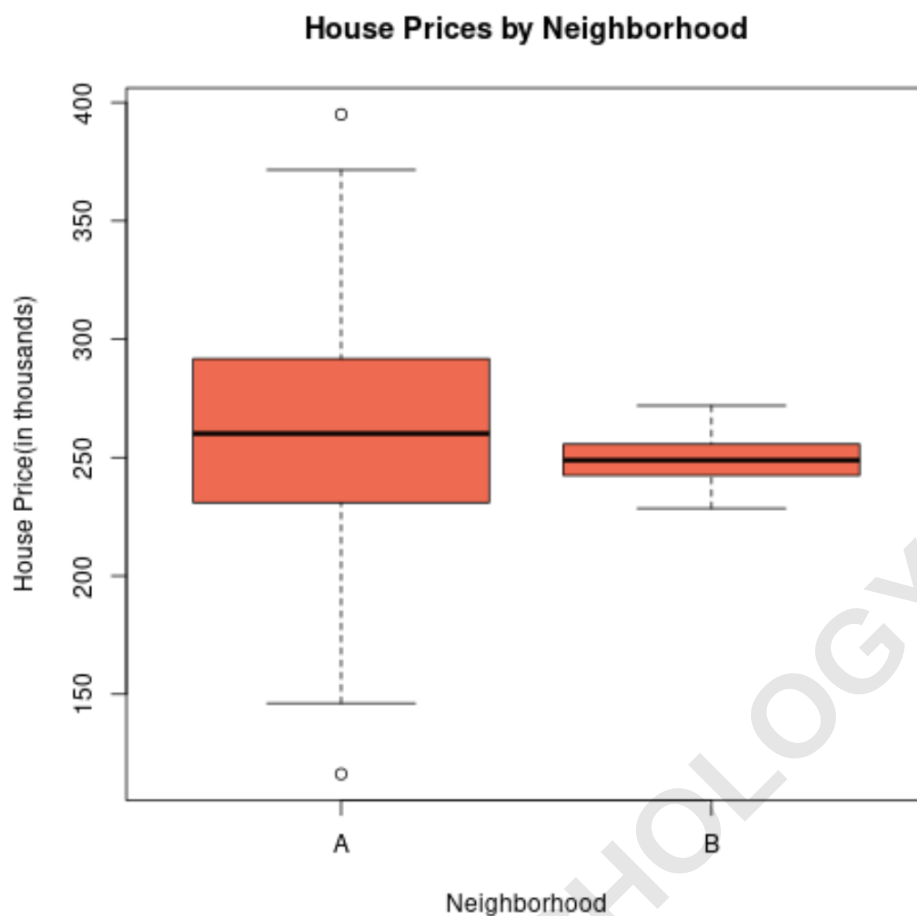
spread of salaries is much higher at company A.

Example 2: Distribution of House Prices

Suppose the mean house price in neighborhood A is \$250,000 and the standard deviation is \$50,000. Since the standard deviation is quite large, this means that some of the house prices will be far greater than \$250,000 and some will be far less. If you look at a given house in this neighborhood, there's no guarantee that the price will be close to the mean.

Conversely, suppose the mean house price in neighborhood B is also \$250,000 but the standard deviation is only \$10,000. Since this standard deviation is fairly small, you can be sure that any given house you look at in the neighborhood is likely to be close to this price.

If we created a boxplot to visualize the distribution of house prices in these two neighborhoods, it might look something like this:



The length of the boxplot for neighborhood A is so much greater because the standard deviation of house prices is so much higher.

In fact, house prices range from lower than \$150k to higher than \$400k for neighborhood A, while prices only range from about \$230k to \$270k for neighborhood B.

By simply knowing the standard deviation of house prices in each neighborhood, we can know how much variation to expect in prices in each neighborhood.