

Standardization (Z-score Normalization):

1. **Scale:** Standardization scales data to have a mean of 0 and a standard deviation of 1. It centers the data around zero and measures the data's spread in terms of standard deviations.
2. **Formula:** The formula for Standardization is:

$$\text{Standardized Value} = \frac{\text{Value} - \text{Mean}}{\text{Standard Deviation}}$$

where 'Value' is the original data point, 'Mean' is the mean (average) value of the dataset, and 'Standard Deviation' is the standard deviation of the dataset.

3. Use Cases:

- Standardization is used when data does not have a well-defined range or when you want to remove the effect of outliers.
- It's commonly applied to features when their units or scales are not meaningful or when you want to make data suitable for algorithms that assume a normal distribution of features.
- It's particularly useful for linear regression, K-means clustering, and principal component analysis (PCA).

4. Advantages:

- Centers data around zero, which can be important for some algorithms.
- Makes data suitable for algorithms that assume a normal distribution of features.
- Reduces the impact of outliers on the dataset.

Summary:

- Use Normalization when you have a specific range in mind and want to maintain the relationships between data points.
- Use Standardization when you want to center data around zero and have unit variance or when dealing with data that lacks a meaningful range.

The choice between these techniques depends on your data and the requirements of the machine learning algorithm you plan to use. It's essential to consider the characteristics of your data and the assumptions of the algorithm to make an informed decision about which scaling method to apply.