### How to Find if the Data Is normal?

By using a displot we can see if our data is distributed normally or not so a displot plot for you which will help you identify how your data is looking, how it is appearing. It will show you how your data is distributed and how skewed it is. Secondly, we can use the pd.skew() function. If it shows 0, the data is fine, but if it shows a negative value, the data is skewed. Another plot we have is the qq plot. The qq plot is a reliable way for us to determine whether the data is normally distributed or not

# Column Transformer

In our dataset we have faced different problems on different columns. Means in one column you have to impute missing values or maybe in another column you have to scale it etc. So fixing these columns separately is kind of challenging for us. So that's why we use a column transformer in which we will transform our columns using a single class.

# Function Transformer

A FunctionTransformer in scikit-learn is a class that allows you to apply a custom function to each element or row of your dataset. It's particularly useful when you need to apply a transformation that isn't readily available as a built-in preprocessing step in scikit-learn.

Here's a breakdown of FunctionTransformer and its usage:

Custom Transformation: FunctionTransformer allows you to define a custom function that operates on your data. This function can be as simple or complex as needed, depending on the transformation you want to apply.
Flexible Usage: You can use FunctionTransformer to apply transformations to individual elements, entire rows, or even entire columns of your dataset. This flexibility makes it suitable for a wide range of preprocessing tasks.
Integration with Pipelines: FunctionTransformer seamlessly integrates with scikit-learn pipelines, allowing you to incorporate custom transformations into your overall preprocessing workflow.
Example Use Cases:

- Scaling or normalizing data using a custom scaling function.

- Applying mathematical transformations such as logarithm or exponential functions.
- Handling missing values by imputing them with a custom strategy.
- Feature engineering tasks such as creating new features based on existing ones.

When to Use:

- You might consider using FunctionTransformer when:
- Your data preprocessing requires custom transformations that aren't available in scikit-learn's built-in preprocessing modules.

# Log Transform

A log transform is a mathematical transformation applied to data by taking the logarithm of each observation. The most commonly used base for the logarithm is the natural logarithm (base e), but other bases such as 10 or 2 can also be used depending on the context.

Log transforms are commonly used on datasets with positive-valued variables that are right-skewed, such as:

- Income data
- Population data
- Sales data
- Stock price data

If we have right skewed data then we will apply log transform and we can not use this on negative values. And if we have left skewed data we will use square transform.

# Power Transformer Class

### Box-Cox Transform

We can convert any given distribution into a normal distribution, and this is a general transformation. In fact, log transform, sqrt transform, and sq transform are special cases of this.

Data Types: The Box-Cox transform is suitable for continuous, positive-valued data. It's often applied to data that exhibit positive skewness or unequal variances across groups. Common examples include:

- Financial data such as stock prices and incomes.
- Count data such as the number of customer transactions or defects.
- Non-negative data such as measurements of height or weight.

Benefits:

- Normalization: The Box-Cox transform can normalize the data, making it more suitable for statistical techniques that assume normality.
- Homogenization of Variance: It can stabilize the variance, making it more consistent across different levels of the data.
- Improved Interpretability: After transformation, relationships in the data may become more interpretable.

Note: Box-Cox Transformation is only applicable on those numbers greater than zero.

## Yeo - Johnson Transform

What Yeo - Johnson Transform do, basically it has all of the Box-Cox Transformer characteristics but there is only one change which is also applicable on those numbers which are zero or negative numbers.