

What is the difference between interpolation and extrapolation?

Authored by
stats writer

May 12, 2024

RECOMMENDED CITATION

stats writer (2024). *What is the difference between interpolation and extrapolation?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=143765>

Interpolation and extrapolation are two methods used in data analysis and prediction to estimate values between or beyond existing data points. Interpolation is the process of estimating data points within the range of known data, while extrapolation is the process of estimating data points outside the range of known data. In simpler terms, interpolation is used to fill in missing values within a set of data, while extrapolation is used to make predictions beyond the known data. While both methods involve estimation, they differ in terms of the reliability and accuracy of the results. Interpolation is generally considered more accurate as it is based on existing data, whereas extrapolation can be less reliable as it involves making assumptions beyond the known data points. It is important to carefully consider the limitations and potential errors of both methods when using them in data analysis and prediction.

Interpolation vs. Extrapolation: What's the Difference?

Two terms that students often confuse in statistics are interpolation and extrapolation.

Here's the difference:

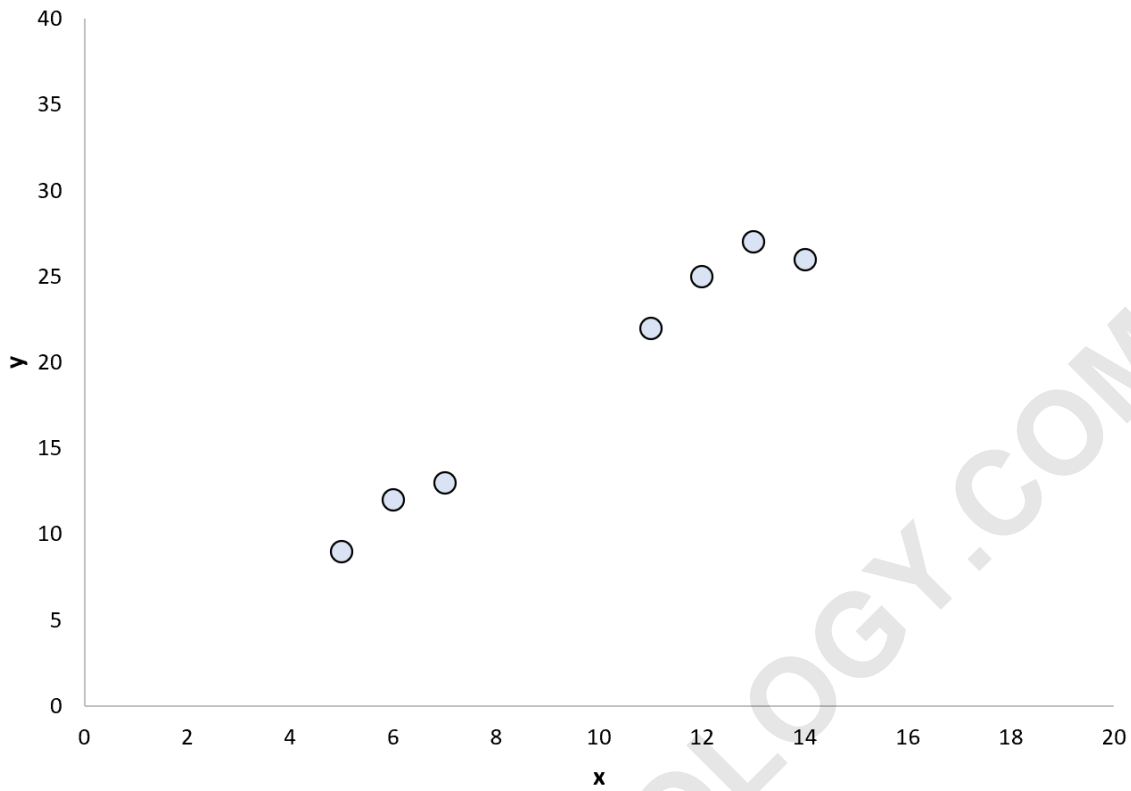
Interpolation refers to predicting values that are *inside* of a range of data points.

Extrapolation refers to predicting values that are *outside* of a range of data points.

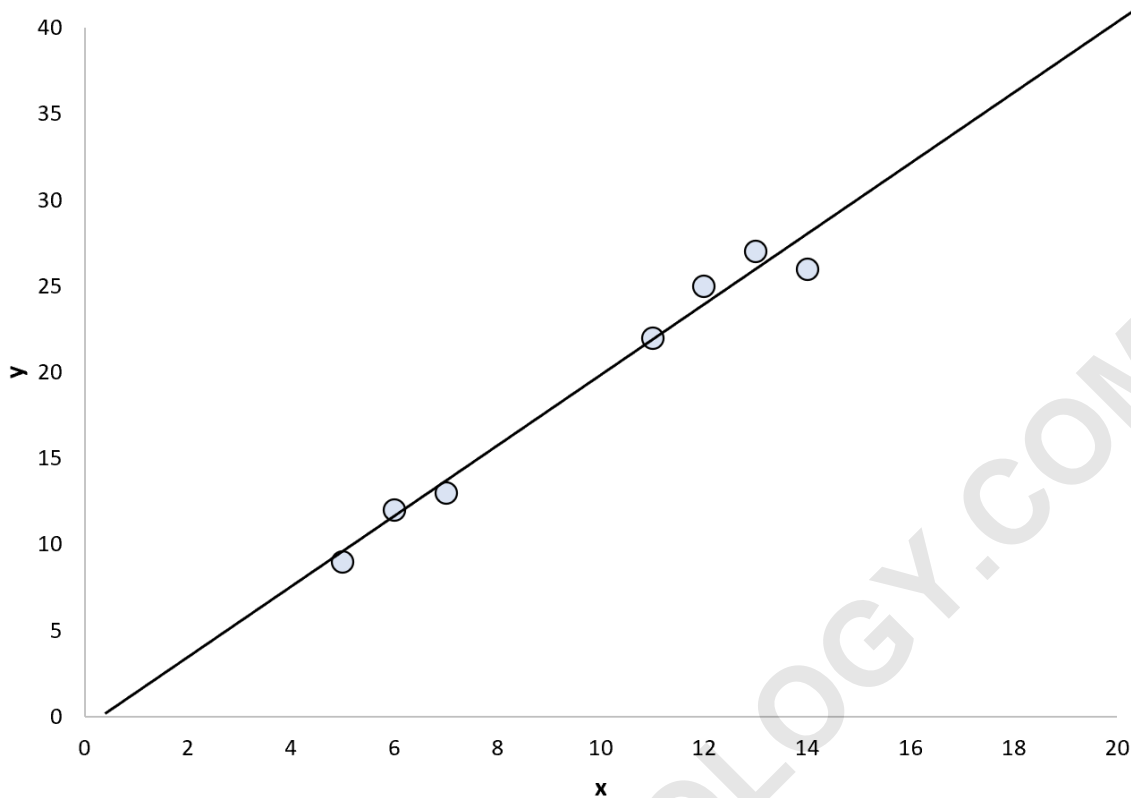
The following example illustrates the difference between the two terms.

Example: Interpolation vs. Extrapolation

Suppose we have the following dataset:



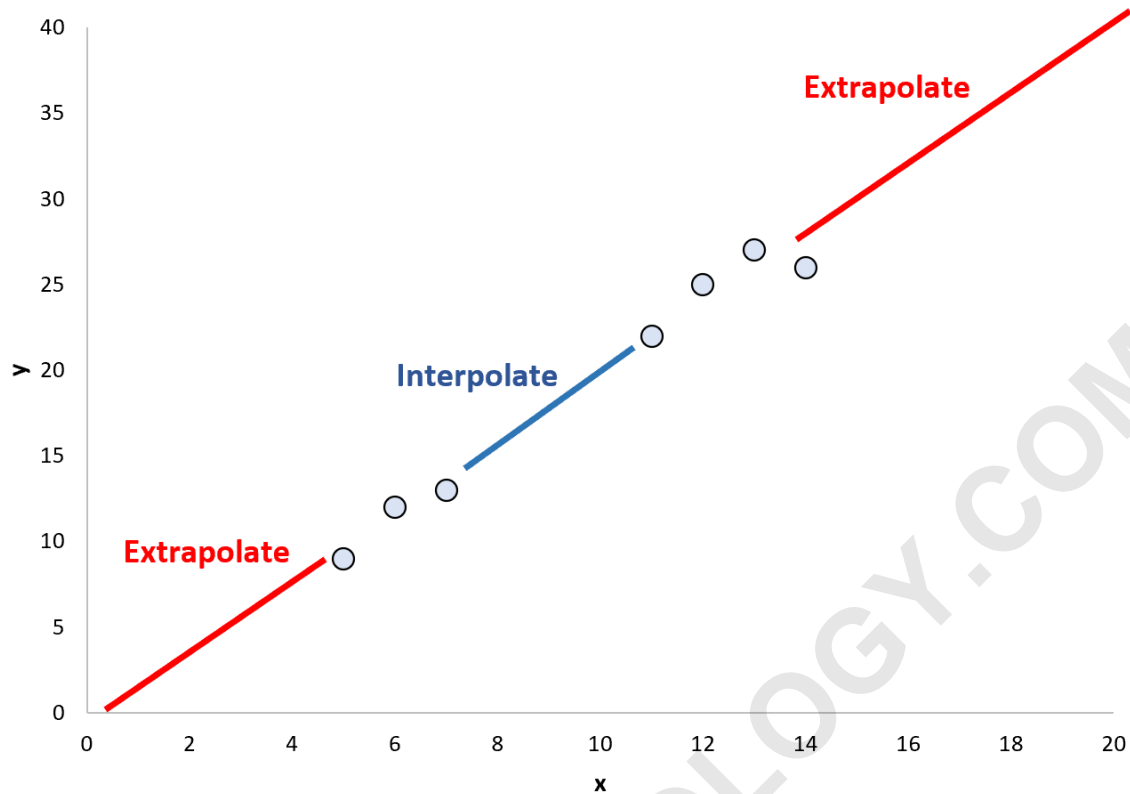
We may decide to fit a to these points:



We could then use the fitted regression model to predict the values of points both *inside* and *outside* of the range of data points.

When we use the fitted regression model to predict the values of points inside the existing range of data points it is known as interpolation.

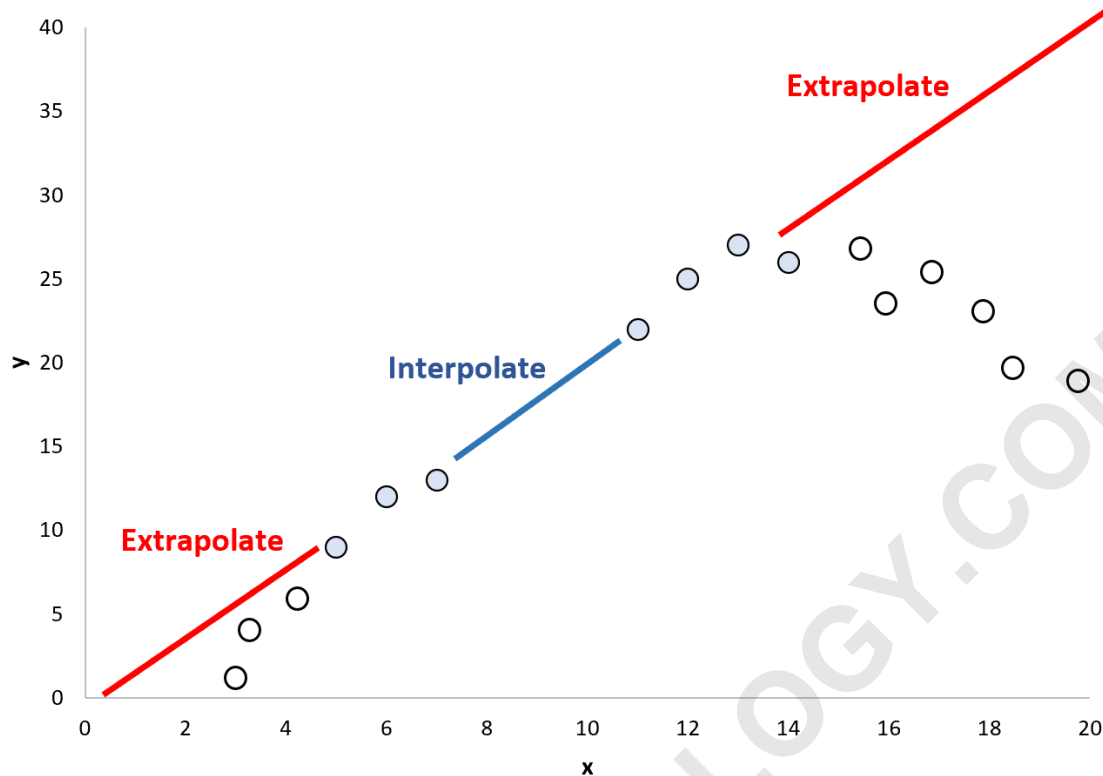
Conversely, when we use the fitted regression model to predict the values of points outside the existing range it is known as extrapolation:



The Potential Danger of Extrapolation

When we perform extrapolation, we assume that the same pattern that exists inside the current range of data points also exists outside of the range as well.

However, this can be a dangerous assumption because it's possible that the pattern that exists outside the current range of data points is quite different:



For this reason, it can be dangerous to use extrapolation to predict the values of data points that fall outside of the range of values that was used to build the regression model.

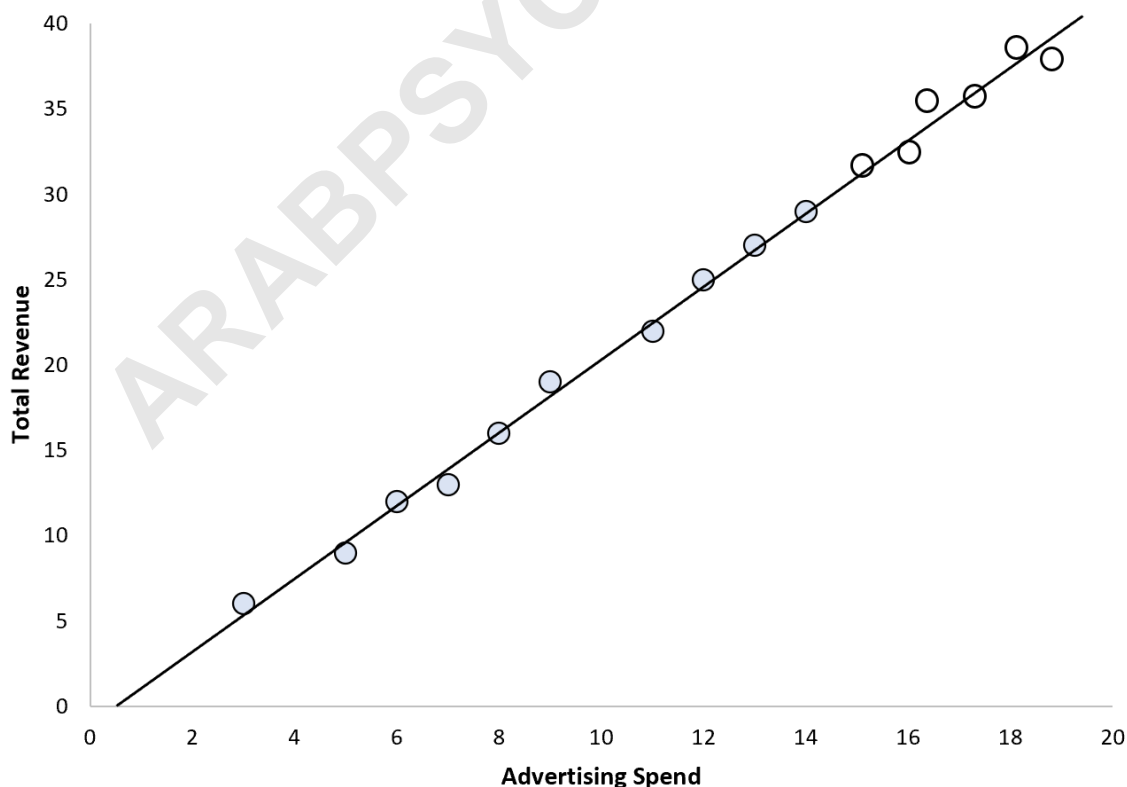
In practice, it's often fine to use extrapolation to predict the values of points that fall just slightly outside of the range of existing values but the further outside the range the higher the likelihood that the difference between the predicted value and the actual value will be large.

When to Use Extrapolation

Often it requires domain-specific expertise to determine if extrapolation is a reasonable idea or not.

For example, suppose a marketing department at a business fits a simple linear regression model using advertising spend as the predictor variable and total revenue as the response variable.

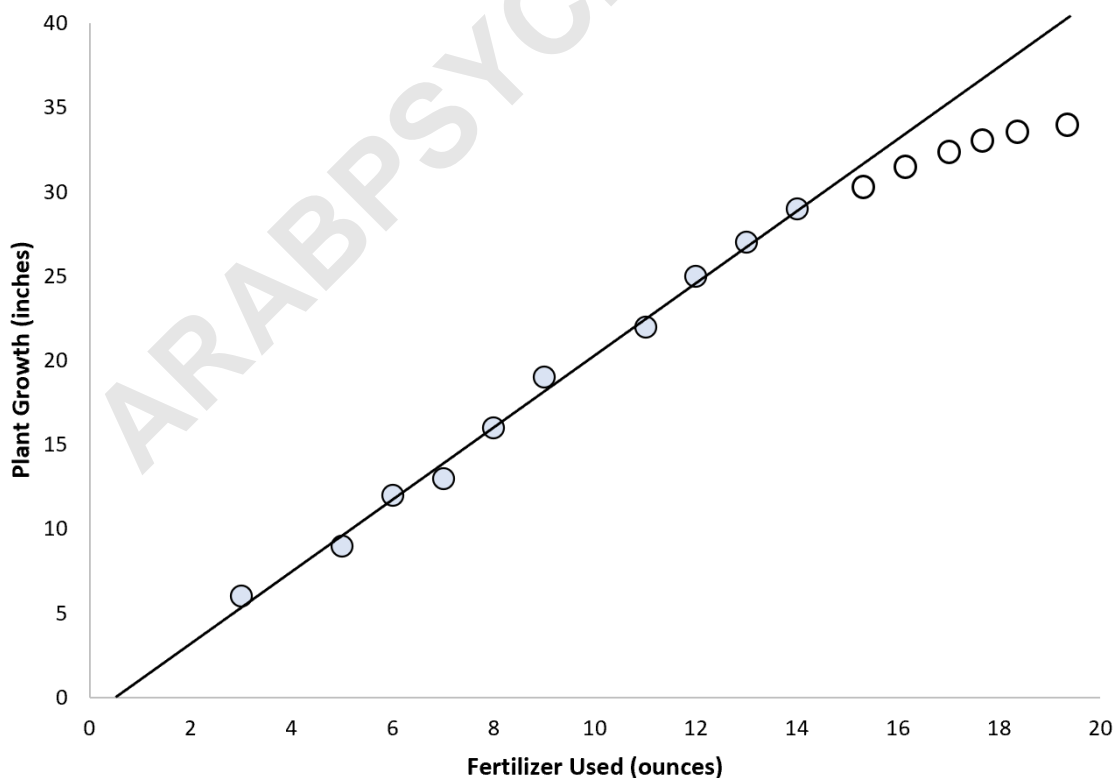
In this scenario, it may be reasonable to assume that a steady increase in advertising spend will lead to a predictable increase in total revenue:



In this scenario, we may be quite confident in our ability to extrapolate values.

However, consider a scenario where a biologist wants to use total fertilizer to predict plant growth.

She may decide to fit a simple linear regression model to the data points, but since there is an upper limit on how tall plants can grow, it probably doesn't make sense to use extrapolation to predict the values of points outside of the range of values used to fit the model:



In this scenario, we may be considerably less confident in our ability to extrapolate values.

The Takeaway: Extrapolation can make sense in some fields more than others, but there is always a potential danger that the pattern that exists within the range of values used to fit the model does not exist outside of the range.

ARABPSYCHOLOGY.COM