

How do I use the LINEST function in Google Sheets?

Authored by
stats writer

July 1, 2024

RECOMMENDED CITATION

stats writer (2024). *How do I use the LINEST function in Google Sheets?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=164348>

The LINEST function in Google Sheets is a statistical function that can be used to calculate the least squares regression line for a set of data points. This function allows users to find the best-fit line for their data, which can be helpful in analyzing trends and making predictions. To use the LINEST function, users must input the range of data points and the range of corresponding y-values, and the function will return an array of values, including the slope, y-intercept, and other relevant statistics. This function is particularly useful for analyzing large sets of data and can provide valuable insights into the relationship between variables.

LINEST

Given partial data about a linear trend, calculates various parameters about the ideal linear trend using the least-squares method.

Sample Usage

```
LINEST(B2:B10, A2:A10)
```

```
LINEST(B2:B10, A2:A10, FALSE, TRUE)
```

Syntax

```
LINEST(known_data_y, , , )
```

known_data_y - The array or range containing dependent (y) values that are already known, used to curve fit an ideal linear trend.

If **known_data_y** is a two-dimensional array or range, **known_data_x** must have the same dimensions or be omitted.

If **known_data_y** is a one-dimensional array or range, **known_data_x** may represent multiple independent variables in a two-dimensional array or range. I.e. if **known_data_y** is a single row, each row in **known_data_x** is interpreted as a separated independent value, and analogously if **known_data_y** is a single column.

known_data_x - - The values of the independent variable(s) corresponding with **known_data_y**.

If **known_data_y** is a one-dimensional array or range, **known_data_x** may represent multiple independent variables in a two-dimensional array or range. I.e. if **known_data_y** is a single row, each row in **known_data_x** is interpreted as a separated independent value, and analogously if **known_data_y** is a single column.

Note: For multiple independent variables, the order of the output parameters are corresponding to the input variables in reverse.

`calculate_b` - - Given a linear form of $y = m \cdot x + b$, calculates the y-intercept (b) if `TRUE`. Otherwise, forces b to be 0 and only calculates the m values if `FALSE`, i.e. forces the curve fit to pass through the origin.

`verbose` - - A flag specifying whether to return additional regression statistics or only the linear coefficients and the y-intercept (default).

If `verbose` is `TRUE`, in addition to the set of linear coefficients for each independent variable and the y-intercept, `LINEST` returns

The standard error for each coefficient and the intercept,

The coefficient of determination (between 0 and 1, where 1 indicates perfect correlation),

Standard error for the dependent variable values,

The F statistic, or F-observed value indicating whether the observed relationship between dependent and independent variables is random rather than linear,

The degrees of freedom, useful in looking up F statistic values in a reference table to estimate a confidence level,

The regression sum of squares, and

The residual sum of squares.

See Also

`TREND`: Given partial data about a linear trend, fits an ideal linear trend using the least squares method and/or predicts further values.

`LOGEST`: Given partial data about an exponential growth curve, calculates various parameters about the best fit ideal exponential growth curve.

`GROWTH`: Given partial data about an exponential growth trend, fits an ideal exponential growth trend and/or predicts further values.

Examples