

How to Easily Perform Logarithmic Regression in Google Sheets

Authored by
stats writer

November 28, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Easily Perform Logarithmic Regression in Google Sheets*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=101197>

Introduction to Logarithmic Regression

Regression analysis is a powerful statistical technique used across disciplines, from economics to biology, for modeling the relationships between variables. Specifically, logarithmic regression is a specialized form of regression essential for scenarios where the rate of change is not constant. Unlike linear regression, which assumes a straight-line relationship, logarithmic models are designed to capture phenomena characterized by rapid initial growth or decay that subsequently slows down significantly as the independent variable increases. This type of modeling is indispensable when analyzing natural or social processes that exhibit diminishing returns, providing a robust mathematical framework for analyzing non-linear data patterns.

Mastering the execution of logarithmic regression is vital for data analysts. While sophisticated statistical software like R or SPSS are often used for complex analyses, accessible tools like Google Sheets provide powerful, built-in functionalities to perform complex statistical calculations efficiently. This comprehensive guide will walk you through the precise steps required to implement logarithmic regression in Google Sheets, focusing on the rigorous mathematical approach using function formulas, followed by the simpler visual method using charts and trendline options.

Understanding the underlying structure of a logarithmic relationship is the first step toward effective modeling. This approach is frequently applied in fields such as pharmacology (analyzing drug concentration decay), finance (modeling compounding interest over time), and behavioral sciences (studying learning curves). The ability of the model to compress vast ranges of data into a manageable curve makes it a preferred tool for handling skewed distributions and heteroscedasticity often present in real-world datasets where marginal gains decrease as inputs increase.

Understanding the Logarithmic Regression Equation

The core mathematical foundation of logarithmic regression dictates a specific relationship between the predictor and the response. The model posits that the relationship between the variables can be linearized by applying the natural log function to the predictor variable. This transformation allows standard linear estimation techniques to be applied to the converted data, yielding highly accurate curve fits for the original non-linear relationship. This process effectively converts the complex relationship into a form compatible with standard least squares methodology.

The standard equation defining a logarithmic regression model takes the following generalized form, which is crucial for interpretation and manual calculation:

$$y = a + b \cdot \ln(x)$$

Where the components represent the fundamental parts of the statistical relationship:

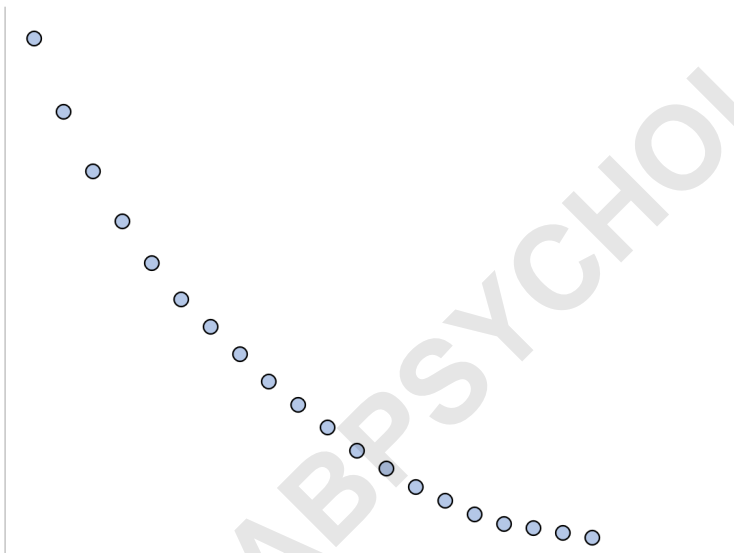
y: Represents the response variable, the outcome we are attempting to predict or model based on changes in the input.

x: Represents the predictor variable, the input or explanatory variable influencing the response.

ln(x): Refers to the natural log of the predictor variable, the mathematical transformation that linearizes the relationship between the two main variables.

a, b: These are the regression coefficients. The value 'a' is the intercept (the predicted value of y when ln(x) equals zero), and 'b' is the slope, describing the change in y for a unit change in ln(x).

When the coefficient 'b' is negative, the model represents logarithmic decay, where the response variable decreases as the predictor variable increases, but at a continually slowing rate. Conversely, a positive 'b' signifies logarithmic growth, where the increase in the response variable decelerates over time. The image provided below demonstrates a classic example of logarithmic decay, where the curve flattens significantly as the X value increases, showcasing the non-linear relationship this model is designed to handle:



Why Use Logarithmic Regression?

Choosing the appropriate regression model is critical for ensuring the validity and accuracy of statistical inferences. Logarithmic regression is specifically chosen when visual inspection of the data, typically through a scatter plot, suggests a curvilinear relationship where the curve approaches an asymptote--meaning the changes in the dependent variable become increasingly smaller for constant increases in the predictor variable. This specific pattern strongly violates the assumption of strict linearity required for standard linear regression models.

In many real-world datasets, especially those involving biological or economic processes, the initial impact of a factor is strong, but the marginal returns diminish over time. For example, in agricultural

studies, the first unit of fertilizer dramatically improves crop yield, but subsequent units provide less proportional gain. Logarithmic transformation handles this phenomenon of diminishing returns gracefully, providing a mathematically sound fit where simpler linear models would significantly overestimate outcomes at higher predictor values and fail to accurately represent the data's true behavior.

Furthermore, using the natural log helps satisfy several crucial statistical assumptions. Specifically, the transformation can help stabilize the variance of the errors (promoting homoscedasticity) and can aid in normalizing the distribution of the response variable in certain contexts. By transforming the independent variable into its log form, we are effectively converting a non-linear estimation problem into a linear one, which allows us to leverage the stable and well-understood mathematical framework of ordinary least squares (OLS) estimation within environments like Google Sheets.

Step 1: Preparing Your Data in Google Sheets

The first essential step in performing logarithmic regression is the meticulous preparation and organization of the raw data. For this analysis, the data must be structured into two distinct columns: one for the predictor variable (x) and one for the response variable (y). It is paramount to ensure that all predictor values (x) are strictly positive, as the natural logarithm function is mathematically undefined for zero or negative numbers, which would halt the analysis.

To demonstrate the calculation process, we will utilize a representative dataset containing paired observations for 'x' and 'y'. Begin by entering your data into adjacent columns in a new Google Sheets workbook. For professional clarity and ease of reference, label the columns appropriately (e.g., 'X Value' in Column A and 'Y Value' in Column B), starting your data entry from row 2.

Here is the structure of the synthetic data we will employ for this detailed step-by-step example:

	A	B	C	D
1	x	y		
2		1	59	
3		2	50	
4		3	44	
5		4	38	
6		5	33	
7		6	28	
8		7	23	
9		8	20	
10		9	17	
11		10	15	
12		11	13	
13		12	12	
14		13	11	
15		14	10	
16		15	9.5	
17				
18				
19				
20				
21				
22				

This initial layout is critical because the mathematical calculation functions in Sheets, such as the powerful LINEST array function, rely on correctly defined and contiguous ranges for both the dependent and independent variables. A well-organized dataset facilitates accurate processing and minimizes potential errors in subsequent calculations, especially when dealing with increasingly large volumes of statistical observations.

Step 2: Transforming the Predictor Variable using the Natural Log

The core mathematical requirement for fitting a logarithmic regression is the transformation of the predictor variable, x , into its natural log equivalent, $\ln(x)$. This essential transformation successfully linearizes the relationship between the variables, making it possible to apply standard linear regression techniques. This intermediate step is mandatory when utilizing the formula-based approach in Google Sheets, as the platform's standard regression calculation functions expect a linear input.

To execute this step, create a new, dedicated column immediately adjacent to the original 'x'

column (e.g., Column C), and label it clearly, perhaps 'LN(X)'. In the first data cell of this new column (e.g., cell C2), you must input the natural log formula targeting the corresponding x value. The specific function used in Google Sheets for calculating the natural logarithm (base \$e\$) is `LN()`.

For our example dataset, starting in cell C2, the precise formula to be entered is: `=LN(A2)`. This formula instructs Sheets to calculate the natural log of the value in cell A2.

After successfully entering the formula in the first cell, utilize the small square drag-handle feature at the bottom-right corner of the cell to quickly apply this calculation to all remaining rows in the dataset. This action calculates the natural log for every observation of the original predictor variable. This newly populated column, LN(X), will now function as the primary predictor variable for the subsequent regression calculation using the LINEST function.

The resulting data structure, incorporating the newly calculated natural log values, should visually align with the following image, confirming the successful completion of the essential transformation stage:

	A	B	C	D
1	x	y	ln(x)	
2	1	59	0	
3	2	50	0.6931471806	
4	3	44	1.098612289	
5	4	38	1.386294361	
6	5	33	1.609437912	
7	6	28	1.791759469	
8	7	23	1.945910149	
9	8	20	2.079441542	
10	9	17	2.197224577	
11	10	15	2.302585093	
12	11	13	2.397895273	
13	12	12	2.48490665	
14	13	11	2.564949357	
15	14	10	2.63905733	
16	15	9.5	2.708050201	
17				
18				
19				
20				
21				

Step 3: Calculating Regression Coefficients using the LINEST Function

With the data successfully transformed, the calculation of the regression coefficients can be performed using Google Sheets' specialized array function, `LINEST`. The `LINEST` function is specifically designed to calculate the statistics for the straight line that provides the best fit for your data using the highly reliable method of least squares. Since we ensured a linear relationship between y and $\ln(x)$ in Step 2, `LINEST` is the perfectly appropriate and robust tool for this estimation.

The general syntax for `LINEST` is `LINEST(known_data_y, known_data_x, ,)`. For our logarithmic regression, our 'known_data_y' will be the original response variable range (Column B), and our 'known_data_x' will be the newly calculated LN(X) range (Column C). We will perform the calculation without requesting extra statistics for simplicity.

To initiate the calculation, select two adjacent blank cells where the two primary coefficients will be displayed (the first cell must contain the slope 'b', and the second cell will contain the intercept 'a'). Because `LINEST` is an array function, you must input the formula into the first cell and then finalize the entry by pressing **Ctrl + Shift + Enter** (or **Command + Shift + Enter** on Mac) to correctly output the results across multiple cells simultaneously.

For our example data, where Y values are located in B2:B16 and LN(X) values are located in C2:C16, the formula is entered as follows:

=LINEST(B2:B16, C2:C16)

Upon successful execution using the array key combination, the output will immediately display the slope coefficient (b) followed by the intercept coefficient (a). These values represent the statistical parameters of your fitted logarithmic model and are necessary for the precise formulation of the regression equation.

Interpreting the Regression Output and Equation

Once the `LINEST` function has been successfully executed, the regression coefficients of the logarithmic regression model are presented in the selected output cells. Interpreting these numerical values correctly is essential for drawing statistically sound and meaningful conclusions from the analysis of the data relationship.

As depicted in the output image below, the results from the calculation provide the estimated values for both the slope (b) and the intercept (a):

E2 fx =LINEST(B2:B16, C2:C16)						
	A	B	C	D	E	F
1	x	y	ln(x)			
2		1	59	0	-20.19869943	63.06859979
3		2	50	0.6931471806		
4		3	44	1.098612289		
5		4	38	1.386294361		
6		5	33	1.609437912		
7		6	28	1.791759469		
8		7	23	1.945910149		
9		8	20	2.079441542		
10		9	17	2.197224577		
11		10	15	2.302585093		
12		11	13	2.397895273		
13		12	12	2.48490665		
14		13	11	2.564949357		
15		14	10	2.63905733		
16		15	9.5	2.708050201		
17						
18						
19						
20						
21						
22						
23						
24						

From this precise output, we extract the following critical regression coefficients:

Intercept (a): 63.0686

Slope (b): -20.1987

By substituting these specific estimates back into the general logarithmic equation form, $y = a + b \cdot \ln(x)$, we can derive the exact fitted model specific to our dataset:

$$y = 63.0686 - 20.1987 \cdot \ln(x)$$

The derived slope coefficient of -20.1987 indicates a clear inverse relationship: for every one-unit increase in the natural logarithm of x , the predicted value of the response variable y decreases by approximately 20.2 units. This quantitatively confirms the pattern of logarithmic decay that was visually observed in the data structure, suggesting that the response variable rapidly decreases as the predictor variable begins to increase, but the rate of decrease diminishes

significantly over the range of observations. A thorough understanding of these coefficient interpretations ensures that the model provides robust and actionable insights, moving beyond mere raw numerical output.

Applying the Model for Prediction

The ultimate objective of fitting any regression model is often its practical predictive capability. Once the logarithmic regression equation is successfully established, it can be utilized to reliably forecast the value of the response variable, y , for any new given value of the predictor variable, x , provided the new x value falls reasonably within the scope of the original observed data (interpolation). Predictive accuracy decreases significantly when extrapolating far outside the observed data range.

To perform a specific prediction, simply substitute the chosen new value of x into the fitted equation: $y = 63.0686 - 20.1987 * \ln(x)$.

For example, let us determine the predicted value of y when x is equal to 12. We must first calculate the natural logarithm of 12 using the LN function, which yields $\ln(12)$ approx 2.4849. We then meticulously substitute this value into the derived equation:

$$y = 63.0686 - 20.1987 * \ln(12)$$

$$y = 63.0686 - 20.1987 * (2.4849)$$

$$y = 63.0686 - 50.1986$$

$$y = 12.87$$

Therefore, based on the rigorous statistical model derived from our sample data, if the predictor variable x equals 12, the corresponding predicted response variable y is calculated to be approximately **12.87**. This ability to predict outcomes based on established quantitative relationships highlights the immense practical utility and precision of logarithmic regression in data-driven research and strategic decision-making. You may also use dedicated online statistical tools to verify these manual predictions, ensuring consistency and accuracy across complex analytical projects.

Alternative Method: Visualizing the Trendline

While the `LINEST` function provides statistically precise coefficients and enables subsequent calculation of metrics like the R-squared value, [Google Sheets](#) also offers a significantly quicker, though often less rigorous, method for fitting and visually presenting logarithmic regression: utilizing the native chart and trendline features. This method is exceptionally useful for rapid visual

confirmation of the relationship and for generating the equation directly onto the graph itself.

The process begins by plotting the original, untransformed data (X and Y columns). First, select the entire relevant data range (e.g., A2:B16) and initiate the chart creation process by navigating to the Insert menu and selecting Chart. It is essential to ensure that the initial chart type selected is a **Scatter chart**, as this is the universally accepted visualization method for displaying bivariate relationships in regression analysis.

Once the scatter chart is displayed on the sheet, access the Customize tab within the Chart Editor pane located on the right side of the screen. Expand the Series configuration section and locate the 'Trendline' option checkbox. Check this box to prompt the visualization of a fitted line. Crucially, under the subsequent 'Type' dropdown menu, select **Logarithmic**. Selecting this option instructs Google Sheets to apply the necessary internal log transformation to the predictor variable and fit the appropriate curved model to your data points.

To retrieve the exact equation derived by the sheet's internal calculations, remain within the Series section and locate the 'Label' dropdown menu. Change the selection from 'None' to **Use Equation**. Google Sheets will instantaneously display the calculated logarithmic equation directly on the chart canvas. This visual approach is highly efficient for verification purposes and for simple, fast coefficient extraction, making it an invaluable preliminary step in data exploration.

Bonus: Feel free to use this online tool to automatically compute the logarithmic regression equation for a given predictor and response variable, allowing for immediate comparison and verification against the coefficients calculated manually using the `LINEST` function in Google Sheets.