

Understanding Logarithmic Regression: A Step-by-Step Guide with Excel

Authored by
Mohammed looti

November 5, 2025

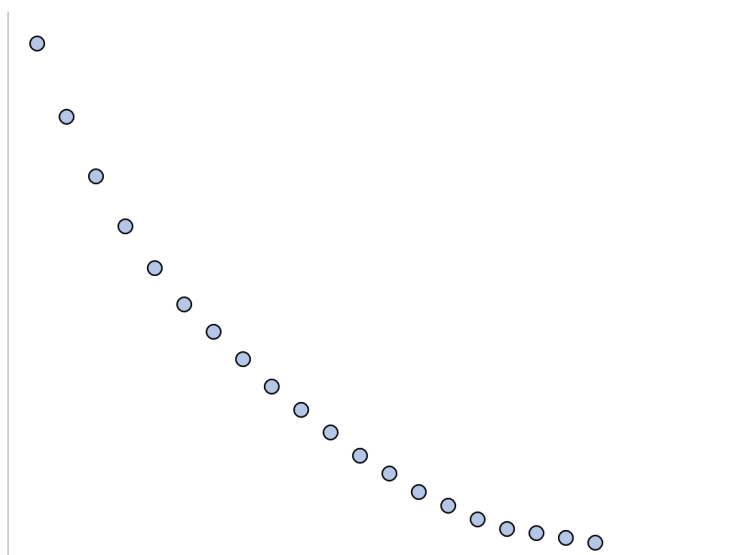
RECOMMENDED CITATION

Mohammed looti (2025). *Understanding Logarithmic Regression: A Step-by-Step Guide with Excel*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=10902>

Logarithmic regression is a sophisticated modeling technique within [regression analysis](#) designed to capture relationships characterized by a non-constant rate of change. This model is exceptionally valuable when analyzing phenomena where the initial impact is swift and dramatic, but the overall effect decelerates or stabilizes over subsequent observations, indicating a pattern of diminishing returns.

Standard linear models, which assume a constant slope, are inadequate for describing these curvilinear relationships. Logarithmic models, conversely, provide an accurate mathematical framework. Typical applications include analyzing biological growth that hits a plateau, modeling the decay rate of drug concentration in a system, or charting user adoption curves where market saturation eventually limits growth. Understanding how to perform this analysis in accessible software like Microsoft Excel is a critical skill for data professionals.

Consider the following visualization, which clearly illustrates a classic case of **logarithmic decay**. Notice how the steep initial decline gradually flattens out, approaching a steady state:



In scenarios such as this, forcing a straight line onto the data would lead to significant errors, resulting in a poor fit and heavily biased forecasts. Therefore, the relationship between the [predictor variable](#) (x) and the [response variable](#) (y) must be accurately defined using the dedicated logarithmic functional form.

The Core Logarithmic Regression Equation

The foundation of the logarithmic regression model lies in transforming the predictor variable. By taking the [natural logarithm](#) of the independent variable, the relationship is rendered linear, allowing us to estimate coefficients using standard linear methods like [Ordinary Least Squares](#)

(OLS).

The fundamental equation that defines this relationship is:

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

A clear understanding of the components of this equation is essential for both execution and interpretation:

y: Represents the [response variable](#) (dependent variable) whose values we are attempting to model and predict.

x: Denotes the [predictor variable](#) (independent variable), which must be strictly positive for the transformation to be valid.

ln(x): This is the [natural logarithm](#) of the predictor variable. This mathematical operation is the critical step that linearizes the underlying curvilinear pattern.

a, b: These are the estimated **regression coefficients**. 'a' represents the y-intercept (the predicted value of y when ln(x) equals zero), and 'b' is the slope that quantifies the change in y per unit change in ln(x).

The following comprehensive guide demonstrates precisely how to implement this model and interpret its results using Microsoft Excel's powerful built-in statistical functionalities.

Step 1: Prepare and Input the Data Set

The necessary first step for any regression analysis is the meticulous organization of data. You must structure your paired observations of the predictor variable (x) and the response variable (y) into two adjacent columns within your Excel spreadsheet. This clear structure is paramount for accurate range selection in later steps.

For this instructional demonstration, we will utilize a small, synthetic dataset specifically generated to illustrate a decaying logarithmic trend, where the output variable systematically decreases as the input variable increases, but at a continuously diminishing rate.

	A	B	C	D	E	F
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						
23						
24						

It is highly recommended to ensure that your data columns are correctly entered and clearly labeled (e.g., 'x' and 'y'). Proper labeling significantly simplifies the subsequent process of selecting the input ranges when configuring Excel's statistical tools.

Step 2: Calculate the Natural Logarithm (ln) of the Predictor Variable

Since standard regression procedures in Excel are optimized for linear models (using OLS), a transformation of the input data is mandatory to fit the logarithmic relationship. We must calculate the [natural logarithm](#) of every value in our original predictor variable, x . This step transforms the independent variable from x to $\ln(x)$.

To achieve this, create a new, distinct column in your spreadsheet, perhaps titled " $\ln(x)$ ". In this column, apply the Excel function **=LN()** to the corresponding values in the original predictor column. For example, if your initial x value is located in cell A2, you would enter the formula **=LN(A2)** into the adjacent cell (e.g., C2) and then drag this formula down to apply the calculation to all observations in the dataset.

The resulting spreadsheet should now include the new column, which serves as our transformed

independent variable, ready for the regression analysis:

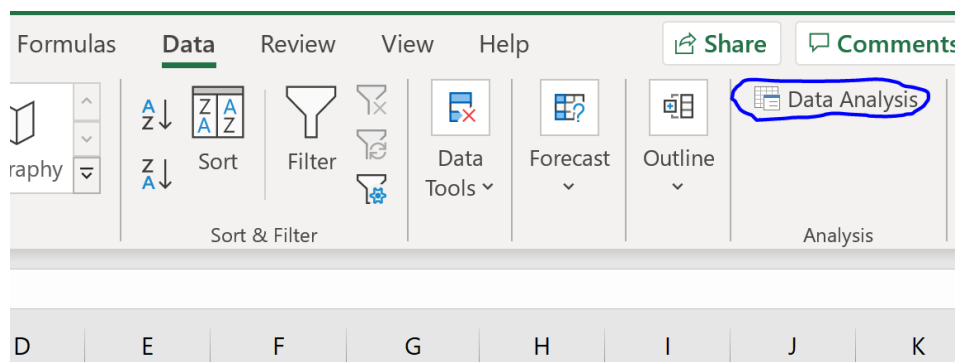
	A	B	C	D	E	F	G
1	x	y	ln(x)				
2	1	59	0	=LN(A2)			
3	2	50	0.693				
4	3	44	1.099				
5	4	38	1.386				
6	5	33	1.609				
7	6	28	1.792				
8	7	23	1.946				
9	8	20	2.079				
10	9	17	2.197				
11	10	15	2.303				
12	11	13	2.398				
13	12	12	2.485				
14	13	11	2.565				
15	14	10	2.639				
16	15	9.5	2.708				
17							
18							
19							
20							
21							
22							
23							
24							
25							

This critical transformation effectively converts the original curvilinear dataset into a structure that can be analyzed linearly. The "ln(x)" column, and not the original "x" column, will be used as the independent variable input in the next stage of the regression setup.

Step 3: Execute the Regression Analysis Using Excel's ToolPak

With the data successfully transformed, we can now leverage the **Regression** function found within Excel's indispensable [Data Analysis ToolPak](#) to estimate the model coefficients.

To initiate the process, navigate to the **Data** tab on the Excel ribbon. Locate the **Analysis** group on the far right and click the **Data Analysis** button.



If the **Data Analysis** option is not visible, you must first activate the [Data Analysis ToolPak](#) add-in via Excel options--a mandatory prerequisite for executing statistical modeling. Once the dialog box appears, select **Regression** and proceed by clicking **OK**.

The accurate input of the variable ranges is vital for model calculation:

For the **Input Y Range**, select the cells corresponding to the original [response variable](#) (y).

For the **Input X Range**, select the cells corresponding to the **transformed predictor variable (ln(x))**. It is essential to remember that the original x values are not used here.

If you included the header row in your selections (which is good practice), ensure the **Labels** box is checked. Specify a location for the output, such as a new worksheet, and click **OK** to generate the comprehensive statistical output.

	A	B	C	D	E	F	G	H	I	J
1	x	y	ln(x)							
2	1	59	0							
3	2	50	0.693							
4	3	44	1.099							
5	4	38	1.386							
6	5	33	1.609							
7	6	28	1.792							
8	7	23	1.946							
9	8	20	2.079							
10	9	17	2.197							
11	10	15	2.303							
12	11	13	2.398							
13	12	12	2.485							
14	13	11	2.565							
15	14	10	2.639							
16	15	9.5	2.708							
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										
27										
28										
29										

Regression

?

×

Input

Input Y Range:

Input X Range:

☒ Labels
 ☐ Constant is Zero

☐ Confidence Level: %

Output options

☒ Output Range:

☐ New Worksheet Ply:

☐ New Workbook

Residuals

☐ Residuals
 ☐ Residual Plots

☐ Standardized Residuals
 ☐ Line Fit Plots

Normal Probability

☐ Normal Probability Plots

OK

Cancel

Help

Step 4: Analyze the Regression Output and Model Significance

The Excel regression output is structured into three primary tables: Regression Statistics, ANOVA (Analysis of Variance), and Coefficients. Interpreting these results is crucial for validating the model's performance and statistical robustness.

E	F	G	H	I	J
SUMMARY OUTPUT					
<i>Regression Statistics</i>					
Multiple R	0.992242675				
R Square	0.984545526				
Adjusted R Square	0.98335672				
Standard Error	2.053603726				
Observations	15				
ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	3492.675253	3492.675	828.1803	3.70174E-13
Residual	13	54.82474741	4.217288		
Total	14	3547.5			
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>
Intercept	63.06859979	1.409031708	44.76024	1.25E-15	60.02457185
ln(x)	-20.19869943	0.701876896	-28.7781	3.7E-13	-21.71501227

Begin by assessing the overall utility of the model using the **ANOVA** table. Focus on the **F-statistic** (828.18 in the example) and the corresponding significance F, or **p-value** (3.70174E-13). Because this **p-value** is extremely small (far below the standard significance threshold of 0.05), we can confidently conclude that the model is statistically significant and highly effective for prediction purposes.

Next, evaluate the **goodness of fit** by reviewing the **R Square** value, located in the Regression Statistics section. This metric provides the proportion of the total variance in the **response variable** (y) that is successfully explained by the transformed predictor variable (ln(x)). A value close to 1 (as seen here) indicates that the logarithmic model provides an exceptionally strong and accurate fit to the observed data points.

Finally, attention must turn to the Coefficients table, which provides the actual numerical values required to construct the prediction equation.

Step 5: Define the Fitted Logarithmic Equation and Forecast Values

The ultimate goal of running the regression is to extract the **regression coefficients** and formalize the final **logarithmic regression** equation. These essential figures are found in the **Coefficients**

column of the output table.

The value labeled **Intercept** corresponds to coefficient 'a', and the value associated with the 'ln(x)' variable corresponds to coefficient 'b':

Intercept (a): 63.0686

ln(x) Coefficient (b): -20.1987

By substituting these derived constants into the general logarithmic model form ($y = a + b \cdot \ln(x)$), we obtain the specific fitted equation for our dataset:

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

This finalized equation is now a reliable tool for forecasting. It can be used to predict the expected value of the [response variable](#), y, given any new input value of the [predictor variable](#), x, provided x is positive and falls within the scope of the original data range.

For a practical forecasting exercise, suppose we need to predict the value of y when the original predictor variable $x = 12$. We must first calculate the [natural logarithm](#) of 12 ($\ln(12)$):

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

Since $\ln(12)$ is approximately 2.4849, the calculation proceeds as follows:

$$y = 63.0686 - 20.1987 \cdot (2.4849) \approx 63.0686 - 50.1986 \approx \mathbf{12.87}$$

Therefore, the model predicts that the value of y when x is 12 is 12.87. This step-by-step process demonstrates the complete methodology for developing and applying a logarithmic regression model in Excel.

Conclusion and Further Resources

[Logarithmic regression](#) offers a powerful and necessary alternative to linear modeling when analyzing non-linear trends characterized by diminishing returns or rapid decay. By mastering the required data transformation using the **=LN()** function and utilizing the robust [Data Analysis ToolPak](#), analysts can efficiently derive statistically sound models without the need for specialized software. This method provides reliable predictive equations essential for informed decision-making across various quantitative fields.

Bonus: For users seeking an automated approach to initial calculations, a variety of online calculators can instantly compute the logarithmic regression equation for a given predictor and response variable dataset.

Additional Resources for Regression Analysis

[How to Perform Simple Linear Regression in Excel](#)