

How can I use the LINEST function in Excel to perform multiple linear regression?

Authored by
stats writer

June 27, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I use the LINEST function in Excel to perform multiple linear regression?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=154623>

The LINEST function in Excel is a powerful tool that allows users to perform multiple linear regression analysis. This function is used to determine the relationship between two or more independent variables and a dependent variable. To use the LINEST function, the user must first enter the data for the independent and dependent variables into a table. Then, by selecting the appropriate cells and entering the function, Excel will generate a regression analysis, including the coefficients, intercept, and standard error values. This function also allows users to create a scatter plot and visually analyze the relationship between the variables. Overall, the LINEST function in Excel is a valuable tool for conducting multiple linear regression analysis and gaining insights into the relationships between variables.

Excel: Use LINEST to Perform Multiple Linear Regression

You can use the LINEST function in Excel to fit a to a dataset.

This function uses the following basic syntax:

```
=LINEST(known_y's, , , )
```

where:

known_y's: An array of known y-values
known_x's: An array of known x-values
const: Optional argument. If TRUE, the constant b is treated normally. If FALSE, the constant b is set to 1.
stats: Optional argument. If TRUE, additional regression statistics are returned. If FALSE, additional regression statistics are not returned.

The following step-by-step example shows how to use this function in practice.

Step 1: Enter the Data

First, let's enter the following dataset in Excel:

	A	B	C	D	E	F
1	x1	x2	x3	y		
2	1	8	7	8		
3	2	7	5	12		
4	2	7	5	14		
5	2	7	4	15		
6	3	6	6	15		
7	4	6	5	14		
8	6	4	6	19		
9	5	5	8	29		
10	6	6	10	24		
11	5	5	14	26		
12	7	4	7	22		
13	9	2	6	30		
14	8	1	4	34		
15						
16						
17						
18						
19						
20						
21						

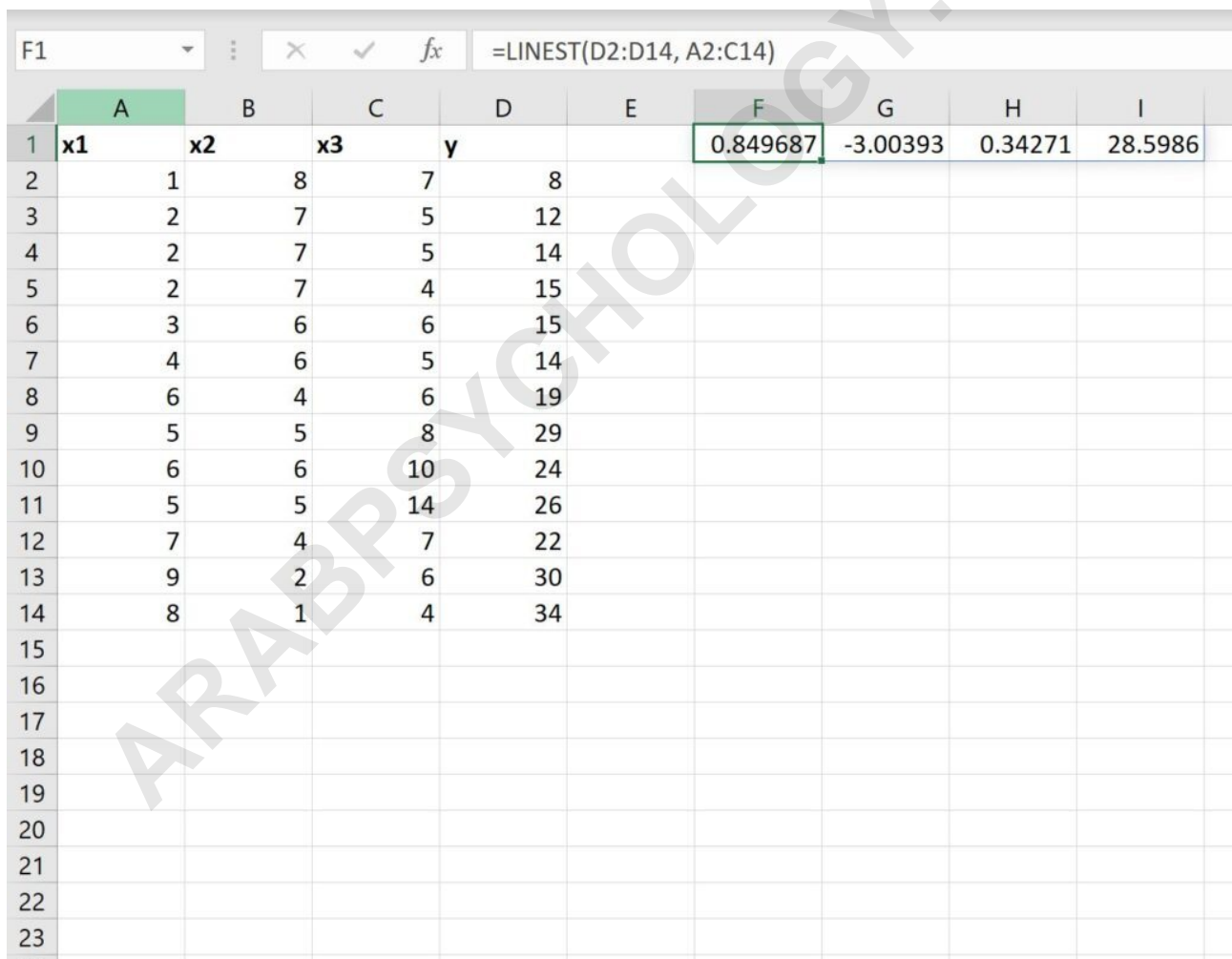
Step 2: Use LINEST to Fit Multiple Linear Regression Model

Suppose we would like to fit a multiple linear regression model using x1, x2, and x3 as predictor variables and y as the response variable.

To do so, we can type the following formula into any cell to fit this multiple linear regression model

=LINEST(D2:D14, A2:C14)

The following screenshot shows how to use this formula in practice:



	A	B	C	D	E	F	G	H	I
1	x1	x2	x3	y		0.849687	-3.00393	0.34271	28.5986
2		1	8	7	8				
3		2	7	5	12				
4		2	7	5	14				
5		2	7	4	15				
6		3	6	6	15				
7		4	6	5	14				
8		6	4	6	19				
9		5	5	8	29				
10		6	6	10	24				
11		5	5	14	26				
12		7	4	7	22				
13		9	2	6	30				
14		8	1	4	34				
15									
16									
17									
18									
19									
20									
21									
22									
23									

Here's how to interpret the output:

The coefficient for the intercept is 28.5986. The coefficient for x1 is 0.34271. The coefficient for x2 is -3.00393. The coefficient for x3 is 0.849687.

$$y = 28.5986 + 0.34271(x1) - 3.00393(x2) + 0.849687(x3)$$

Step 3 (Optional): Display Additional Regression Statistics

We can also set the value for the stats argument in the LINEST function equal to TRUE to display additional regression statistics for the fitted regression equation:

=LINEST(D2:D14, A2:C14, TRUE, TRUE)									
	A	B	C	D	E	F	G	H	I
1	x1	x2	x3	y		0.849687	-3.00393	0.34271	28.5986
2		1	8	7	8	0.453295	1.626423	1.327566	13.20088
3		2	7	5	12	0.838007	3.707539	#N/A	#N/A
4		2	7	5	14	15.51935	9	#N/A	#N/A
5		2	7	4	15	639.9797	123.7126	#N/A	#N/A
6		3	6	6	15				
7		4	6	5	14				
8		6	4	6	19				
9		5	5	8	29				
10		6	6	10	24				
11		5	5	14	26				
12		7	4	7	22				
13		9	2	6	30				
14		8	1	4	34				
15									
16									
17									
18									
19									
20									

The fitted regression equation is still the same:

$$y = 28.5986 + 0.34271(x_1) - 3.00393(x_2) + 0.849687(x_3)$$

Here's how to interpret the other values in the output:

The standard error for x3 is 0.453295. The standard error for x2 is 1.626423. The standard error for x1 is 1.327566. The standard error for the intercept is 13.20088. The R2 for the model is .838007. The residual standard error for y is 3.707539. The overall F-statistic is 15.51925. The degrees of freedom is 9. The regression sum of squares is 639.9797. The residual sum of squares is 123.7126.

In general, the most interesting metric in these additional statistics is the R2 value, which represents the proportion of the variance in the response variable that can be explained the predictor variable.

The value for R2 can range from 0 to 1.

Since the R2 for this particular model is .838, it tells us that the predictor variables do a good job of predicting the value of the response variable y.

Related:

The following tutorials explain how to perform other common operations in Excel:

ARABPSYCHOLOGY.COM