

How to Run and Interpret Multiple Linear Regression in Excel

Authored by
stats writer

March 10, 2026

RECOMMENDED CITATION

stats writer (2026). *How to Run and Interpret Multiple Linear Regression in Excel*.
PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=134962>

Multiple linear regression in Excel is a statistical technique used to analyze the relationship between multiple independent variables and a single dependent variable. It is commonly used to predict the value of the dependent variable based on the values of the independent variables. To perform multiple linear regression in Excel, the user must first organize the data in a tabular format with the dependent variable in one column and the independent variables in separate columns. Then, the user can use the Data Analysis Toolpak to run the regression analysis, which will generate a regression equation and a summary of the model's performance. This allows for the identification of significant variables and the prediction of the dependent variable's value based on the values of the independent variables. Overall, using Excel for multiple linear regression provides a simple and efficient way to analyze and predict relationships between multiple variables.

Perform Multiple Linear Regression in Excel

Multiple linear regression is a method we can use to understand the relationship between two or more explanatory variables and a .

This tutorial explains how to perform multiple linear regression in Excel.

Note: If you only have one explanatory variable, you should instead perform .

Example: Multiple Linear Regression in Excel

Suppose we want to know if the number of hours spent studying and the number of prep exams taken affects the score that a student receives on a certain college entrance exam.

To explore this relationship, we can perform multiple linear regression using hours studied and prep exams taken as explanatory variables and exam score as a response variable.

Perform the following steps in Excel to conduct a multiple linear regression.

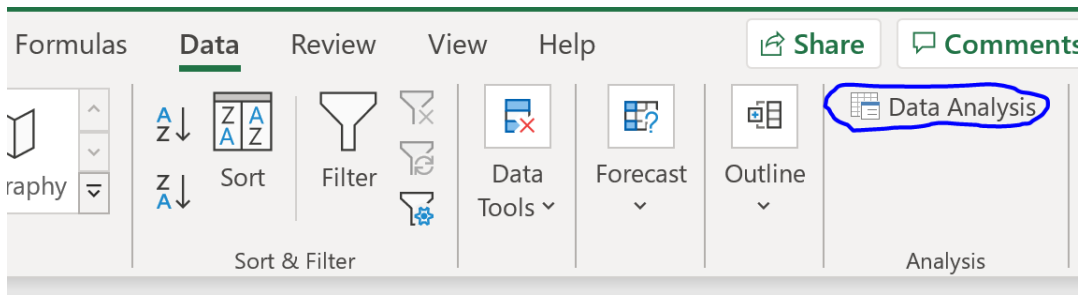
Step 1: Enter the data.

Enter the following data for the number of hours studied, prep exams taken, and exam score received for 20 students:

	A	B	C	D	E	F
1	hours	prep_exams	score			
2	1	1	76			
3	2	3	78			
4	2	3	85			
5	4	5	88			
6	2	2	72			
7	1	2	69			
8	5	1	94			
9	4	1	94			
10	2	0	88			
11	4	3	92			
12	4	4	90			
13	3	3	75			
14	6	2	96			
15	5	4	90			
16	3	4	82			
17	4	4	85			
18	6	5	99			
19	2	1	83			
20	1	0	62			
21	2	1	76			
22						
23						
24						

Step 2: Perform multiple linear regression.

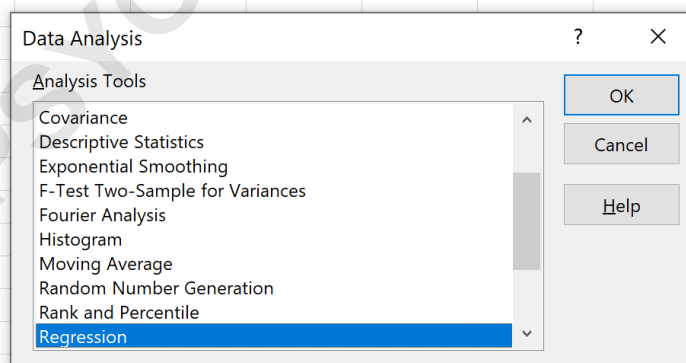
Along the top ribbon in Excel, go to the Data tab and click on Data Analysis. If you don't see this option, then you need to first .



D	E	F	G	H	I	J	K
---	---	---	---	---	---	---	---

Once you click on Data Analysis, a new window will pop up. Select Regression and click OK.

	A	B	C	D	E	F	G	H	I	J
1	hours	prep_exams	score							
2	1	1	76							
3	2	3	78							
4	2	3	85							
5	4	5	88							
6	2	2	72							
7	1	2	69							
8	5	1	94							
9	4	1	94							
10	2	0	88							
11	4	3	92							
12	4	4	90							
13	3	3	75							
14	6	2	96							
15	5	4	90							
16	3	4	82							
17	4	4	85							
18	6	5	99							
19	2	1	83							
20	1	0	62							
21	2	1	76							
22										
23										
24										



For Input Y Range, fill in the array of values for the response variable. For Input X Range, fill in the array of

values for the two explanatory variables. Check the box next to Labels so Excel knows that we included the variable names in the input ranges. For Output Range, select a cell where you would like the output of the regression to appear. Then click OK.

	A	B	C	D	E	F	G	H	I	J
1	hours	prep_exams	score							
2	1	1	76							
3	2	3	78							
4	2	3	85							
5	4	5	88							
6	2	2	72							
7	1	2	69							
8	5	1	94							
9	4	1	94							
10	2	0	88							
11	4	3	92							
12	4	4	90							
13	3	3	75							
14	6	2	96							
15	5	4	90							
16	3	4	82							
17	4	4	85							
18	6	5	99							
19	2	1	83							
20	1	0	62							
21	2	1	76							
22										
23										
24										
25										
26										

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

D	E	F	G	H	I	J	K
SUMMARY OUTPUT							
<i>Regression Statistics</i>							
Multiple R	0.857						
R Square	0.734						
Adjusted R Square	0.703						
Standard Error	5.366						
Observations	20						
ANOVA							
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>		
Regression	2	1350.76	675.38	23.46	0.00		
Residual	17	489.44	28.79				
Total	19	1840.20					
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	
Intercept	67.67	2.82	24.03	0.00	61.73	73.61	
hours	5.56	0.90	6.18	0.00	3.66	7.45	
prep_exams	-0.60	0.91	-0.66	0.52	-2.53	1.33	

Step 3: Interpret the output.

Here is how to interpret the most relevant numbers in the output:

R Square: 0.734. This is known as the coefficient of determination. It is the proportion of the variance in the response variable that can be explained by the explanatory variables. In this example, 73.4% of the variation in the exam scores can be explained by the number of hours studied and the number of prep exams taken.

Standard error: 5.366. This is the average distance that the observed values fall from the regression line. In this example, the observed values fall an average of 5.366 units from the regression line.

F: 23.46. This is the overall F statistic for the regression model, calculated as regression MS / residual MS.

Significance F: 0.0000. This is the p-value associated with the overall F statistic. It tells us whether or not the regression model as a whole is statistically significant. In other words, it tells us if the two explanatory variables combined have a statistically significant association with the response variable. In this case the p-value is less than 0.05, which indicates that the explanatory variables hours studied and prep exams taken combined have a statistically significant association with exam score.

P-values. The individual p-values tell us whether or not each explanatory variable is statistically significant. We can see that hours studied is statistically significant ($p = 0.00$) while prep exams taken ($p = 0.52$) is not statistically significant at $\alpha = 0.05$. Since prep exams taken is not statistically significant, we may end up

deciding to remove it from the model.

Coefficients: The coefficients for each explanatory variable tell us the average expected change in the response variable, assuming the other explanatory variable remains constant. For example, for each additional hour spent studying, the average exam score is expected to increase by 5.56, assuming that prep exams taken remains constant.

Here's another way to think about this: If student A and student B both take the same amount of prep exams but student A studies for one hour more, then student A is expected to earn a score that is 5.56 points higher than student B.

We interpret the coefficient for the intercept to mean that the expected exam score for a student who studies zero hours and takes zero prep exams is 67.67.

Estimated regression equation: We can use the coefficients from the output of the model to create the following estimated regression equation:

$$\text{exam score} = 67.67 + 5.56 * (\text{hours}) - 0.60 * (\text{prep exams})$$

We can use this estimated regression equation to calculate the expected exam score for a student, based on the number of hours they study and the number of prep exams they take. For example, a student who studies for three hours and takes one prep exam is expected to receive a score of 83.75:

$$\text{exam score} = 67.67 + 5.56*(3) - 0.60*(1) = 83.75$$

Keep in mind that because prep exams taken was not statistically significant ($p = 0.52$), we may decide to remove it because it doesn't add any improvement to the overall model. In this case, we could perform simple linear regression using only hours studied as the explanatory variable.

The results of this simple linear regression analysis can be found .

Once you perform multiple linear regression, there are several assumptions you may want to check including:

1. .

2. .

3. .

ARABPSYCHOLOGY.COM