How do I calculate R-Squared in Google Sheets?

Authored by

stats writer

June 27, 2024

R-Squared, also known as the coefficient of determination, is a statistical measure used to determine how well a regression line fits a set of data points. In Google Sheets, the R-Squared value can be calculated by first creating a scatter plot of the data and then using the "TREND" function to generate the regression line equation. The R-Squared value is then obtained by squaring the correlation coefficient (r) from the regression line equation. This will provide a value between 0 and 1, with a higher value indicating a better fit of the regression line to the data points. This measure can be useful in analyzing the strength of a relationship between two variables in a data set.

Calculate R-Squared in Google Sheets

R-squared, often written as r2, is a measure of how well a fits a dataset.

In technical terms, it is the proportion of the variance in the response variable that can be explained by the predictor variable.

The value for r2 can range from 0 to 1:

A value of 0 indicates that the response variable cannot be explained by the predictor variable at all.A value of 1 indicates that the response variable can be perfectly explained without error by the predictor variable.

The following example shows how to calculate R-squared for two variables in Google Sheets.

Example: Calculating R-Squared in Google Sheets

Suppose we have the following data for the number of hours studied and the exam score received for 20 students:

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

Now suppose we want to fit a , using "hours" as the predictor variable and "score" as the response variable.

To find the R-squared for this model, we can use the

RSQ() function in Google Sheets, which uses the following syntax:

=RSQ(known_ys, known_xs)

where:

known_ys: the values for the response variableknown_xs: the values for the predictor variable

In our example, we can type the following formula into cell D2:

=RSQ(A2:A21, B2:B21)

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

D2 =RSQ(A2:A21, B2:B21)

	A	B	C	D
1	Hours	Score		R-squared
2	1	76		0.7272528139
3	2	78		
4	2	85		
5	4	88		
6	2	72		
7	1	69		
8	5	94		
9	4	94		
10	2	88		
11	4	92		
12	4	90		
13	3	75		
14	6	96		
15	5	90		
16	3	82		
17	4	85		
18	6	99		
19	2	83		
20	1	62		
21	2	76		
22				
23				
24				

This means that 72.73% of the variation in the exam scores can be explained by the number of hours studied.

Related:

The following tutorials explain how to perform other

common tasks in Google Sheets: