

How to Perform Quadratic Regression in Stata: A Step-by-Step Guide

Authored by
stats writer

March 8, 2026

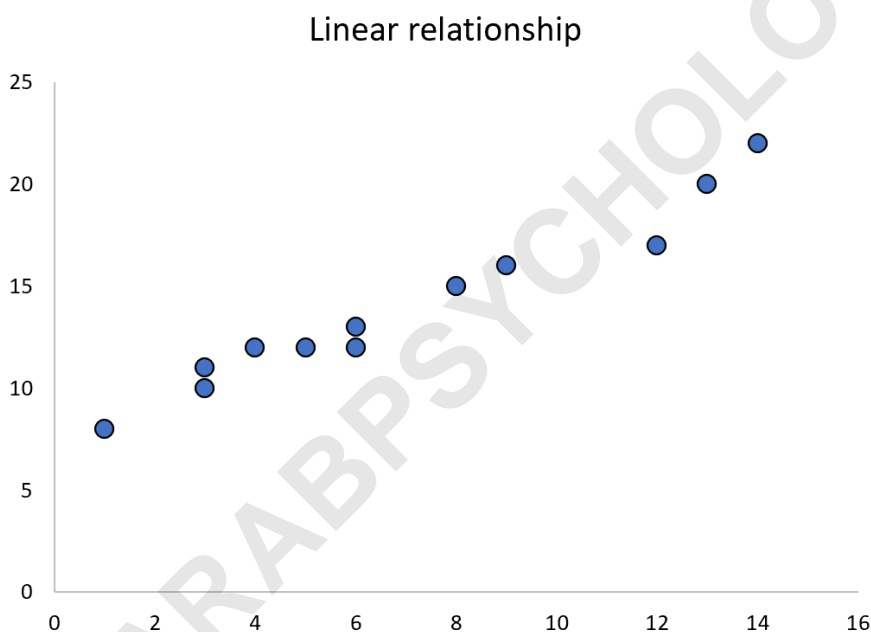
RECOMMENDED CITATION

stats writer (2026). *How to Perform Quadratic Regression in Stata: A Step-by-Step Guide*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=134579>

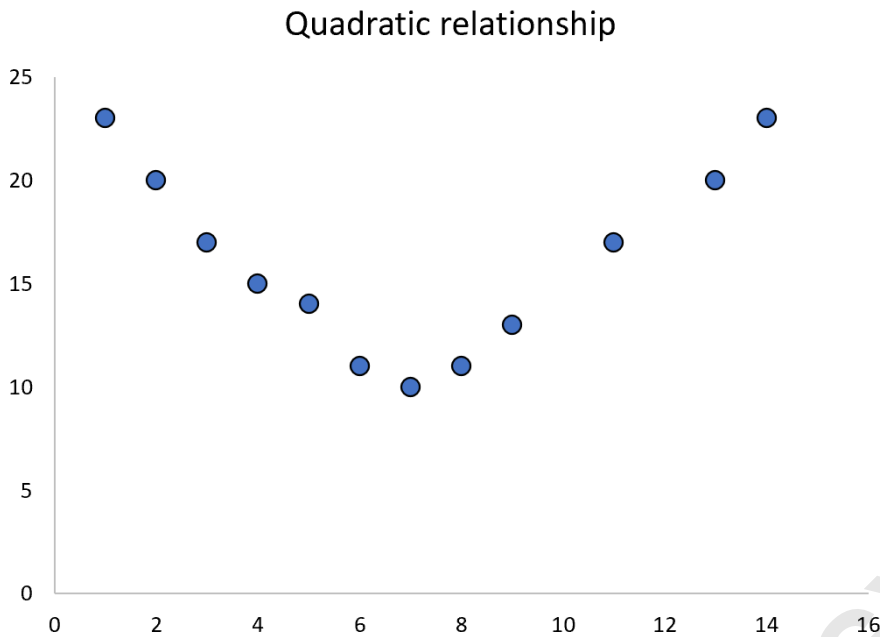
Quadratic regression is a statistical method used to model the relationship between a dependent variable and one or more independent variables, when the relationship is not linear but instead follows a quadratic curve. In Stata, this can be performed by using the "regress" command and specifying the quadratic term (x^2) in the model equation. Additionally, Stata allows for the use of other commands such as "nl" or "nlreg" for more advanced quadratic regression techniques. By following these steps, users can effectively perform quadratic regression in Stata and obtain valuable insights into the non-linear relationship between variables in their data.

Perform Quadratic Regression in Stata

When two variables have a linear relationship, you can often use to quantify their relationship.



However, when two variables have a quadratic relationship, you can instead use quadratic regression to quantify their relationship.



This tutorial explains how to perform quadratic regression in Stata.

Example: Quadratic Regression in Stata

Suppose we are interested in understanding the relationship between number of hours worked and happiness. We have the following data on the number of hours worked per week and the reported happiness level (on a scale of 0-100) for 16 different people:

	hours	happiness	
1	6	12	
2	9	18	
3	12	30	
4	12	42	
5	15	48	
6	21	78	
7	24	90	
8	24	96	
9	27	96	
10	30	90	
11	36	84	
12	39	78	
13	45	66	
14	48	54	
15	57	36	
16	60	24	

You can replicate this example by typing in this exact data into Stata using Data > Data Editor > Data Editor (Edit) along the top menu.

Use the following steps to perform a quadratic regression in Stata.

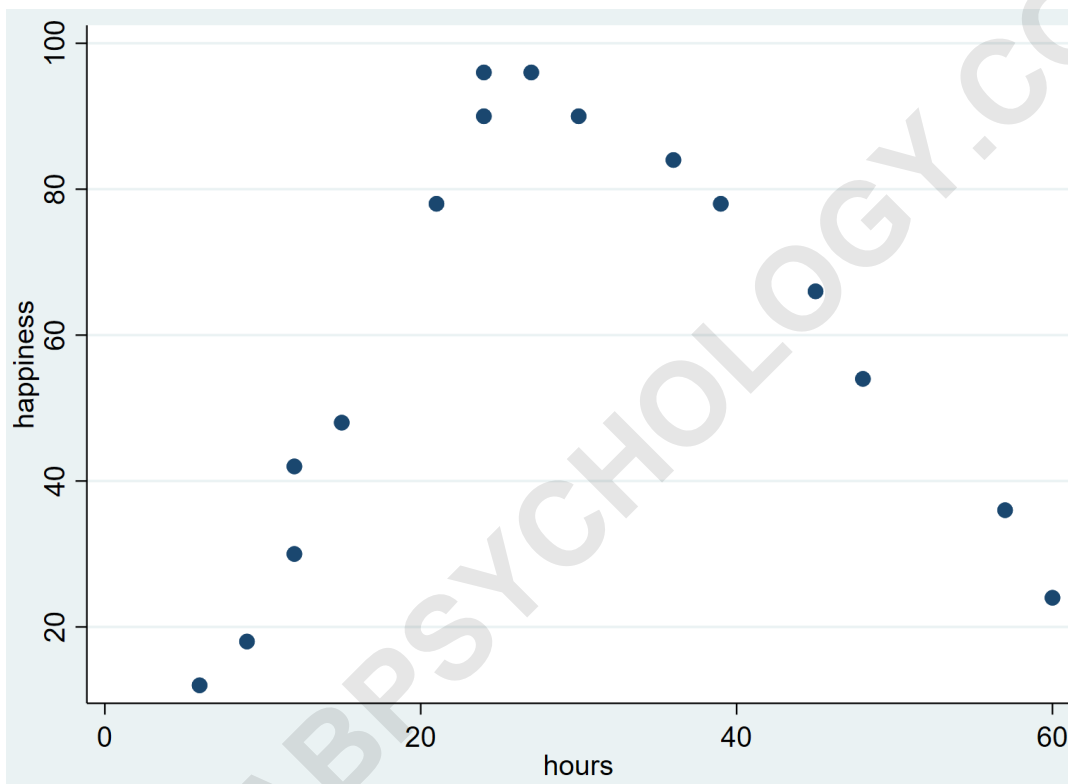
Step 1: Visualize the data.

Before we can use quadratic regression, we need to make sure that the relationship between the explanatory variable (hours) and response variable (happiness) is

actually quadratic. So, let's visualize the data using a scatterplot by typing the following into the Command box:

```
scatter happiness hours
```

This produces the following scatterplot:



We can see that happiness tends to increase as number of hours worked increases from zero up to a certain point, but then begins to drop lower as the number of hours worked exceeds about 30.

This upside down "U" shape in the scatterplot indicates that there is a quadratic relationship between hours

worked and happiness, which means we should use quadratic regression to quantify this relationship.

Before we fit the quadratic regression model to the data, we need to create a new variable for the squared values of our predictor variable *hours*. We can do so by typing the following into the Command box:

gen hours2 = hours*hours

We can view this new variable by going to *Data > Data Editor > Data Editor (Browse)* along the top menu.

	hours	happiness	hours2	
1	6	12	36	
2	9	18	81	
3	12	30	144	
4	12	42	144	
5	15	48	225	
6	21	78	441	
7	24	90	576	
8	24	96	576	
9	27	96	729	
10	30	90	900	
11	36	84	1296	
12	39	78	1521	
13	45	66	2025	
14	48	54	2304	
15	57	36	3249	
16	60	24	3600	

We can see that `hours2` is simply `hours` squared. Now we can perform quadratic regression using `hours` and `hours2` as our explanatory variables and `happiness` as our response variable. To perform quadratic regression, type the following into the Command box:

```
regress happiness hours hours2
```

```
. regress happiness hours hours2
```

Source	SS	df	MS	Number of obs	=	16
Model	11797.7001	2	5898.85003	F(2, 13)	=	65.09
Residual	1178.04994	13	90.6192262	Prob > F	=	0.0000
				R-squared	=	0.9092
				Adj R-squared	=	0.8952
Total	12975.75	15	865.05	Root MSE	=	9.5194

happiness	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
hours	7.173061	.6316306	11.36	0.000	5.808506 8.537616
hours2	-.1069887	.0094615	-11.31	0.000	-.127429 -.0865484
_cons	-30.25287	8.766065	-3.45	0.004	-49.1908 -11.31493

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

Prob > F: 0.000. This is the p-value for the overall regression. Since this value is less than 0.05, it means that the predictor variables *hours* and *hours2* combined have a statistically significant relationship with the response variable *happiness*.

R-squared: 0.9092. This is the proportion of the variance in the response variable that can be explained by the explanatory variable. In this example, 90.92% of the variation in happiness can be explained by *hours* and *hours2*.

Regression Equation: We can form a regression

equation using the coefficient values reported in the output table. In this case, the equation would be:

$$\text{predicted happiness} = -30.25287 + 7.173061(\text{hours}) - .1069887(\text{hours}^2)$$

We can use this equation to find the predicted happiness of an individual, given the number of hours they work per week.

For example, an individual that works 60 hours per week is predicted to have a happiness level of 14.97:

$$\text{predicted happiness} = -30.25287 + 7.173061(60) - .1069887(60^2) = 14.97.$$

Conversely, an individual that works 30 hours per week is predicted to have a happiness level of 88.65:

$$\text{predicted happiness} = -30.25287 + 7.173061(30) - .1069887(30^2) = 88.65.$$

Step 3: Report the results.

Lastly, we want to report the results of our quadratic regression. Here is an example of how to do so:

A quadratic regression was performed to quantify the relationship between the number of hours worked by an individual and their corresponding happiness level (measured from 0 to 100). A sample of 16 individuals was used in the analysis.

Results showed that there was a statistically significant relationship between the explanatory variables *hours* and *hours2* and the response variable *happiness* ($F(2, 13) = 65.09, p < 0.0001$).

Combined, these two explanatory variables accounted for 90.92% of explained variability in happiness.

The regression equation was found to be:

predicted happiness = $-30.25287 + 7.173061(\text{hours}) - .1069887(\text{hours}^2)$