

Learning Quadratic Regression Analysis Using SPSS: A Step-by-Step Guide

Authored by
Mohammed Iooti

November 8, 2025

RECOMMENDED CITATION

Mohammed Iooti (2025). *Learning Quadratic Regression Analysis Using SPSS: A Step-by-Step Guide*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=12901>

When analyzing the relationship between two variables, researchers often begin by fitting a simple [linear regression](#) model to quantify the association. This approach is highly effective when the data exhibits a straight-line pattern.

However, real-world data frequently presents complex relationships that are inherently [non-linear](#). When a simple straight line fails to adequately capture the curvature of the data, attempting to force a linear model results in poor predictive accuracy and misleading interpretations. In such scenarios, particularly when the relationship resembles a U-shape or an inverted U-shape, employing [quadratic regression](#) becomes necessary.

This comprehensive tutorial is designed to guide you through the process of performing and interpreting a quadratic regression analysis within the statistical software [SPSS](#) (Statistical Package for the Social Sciences).

The Necessity of Quadratic Modeling

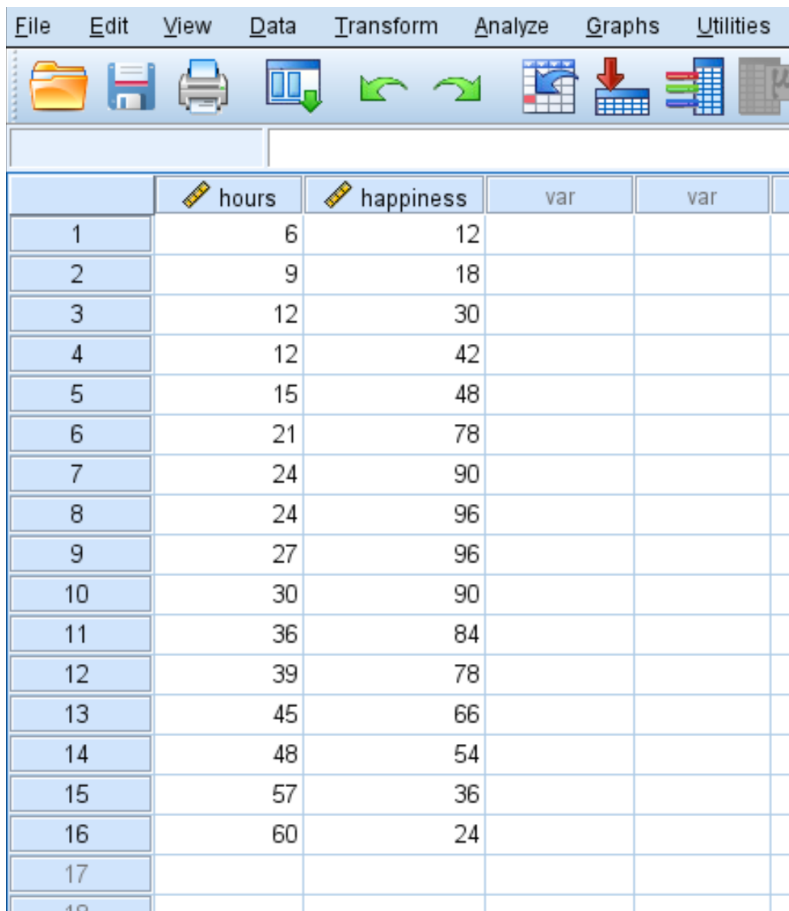
Linear models assume that for every unit increase in the independent variable (X), the dependent variable (Y) changes by a constant amount. This assumption often breaks down in psychological, economic, or behavioral research where variables interact in diminishing or accelerating ways. For instance, increasing study time may lead to a large increase in test scores initially, but the benefit eventually plateaus or even declines due to fatigue.

Quadratic regression addresses this limitation by introducing a squared term of the independent variable (X^2) into the model. Mathematically, the quadratic model takes the form $Y = \beta_0 + \beta_1X + \beta_2X^2 + \epsilon$. By including this squared component, the model is able to fit a parabolic curve, allowing for a turning point where the relationship between X and Y shifts direction.

This technique is crucial for accurately modeling phenomena where intermediate levels of a predictor variable are optimal, while both low and high levels yield poorer outcomes. Understanding how to correctly implement this in a powerful statistical tool like **SPSS** is essential for advanced data analysis.

Example Scenario and Data Visualization (Step 1)

To illustrate the procedure, consider a classic research question regarding the relationship between productivity or effort and well-being. We are interested in understanding the association between the number of hours worked per week and an individual's reported happiness level. Happiness is measured on a continuous scale from 0 to 100. We have collected data from 16 participants:

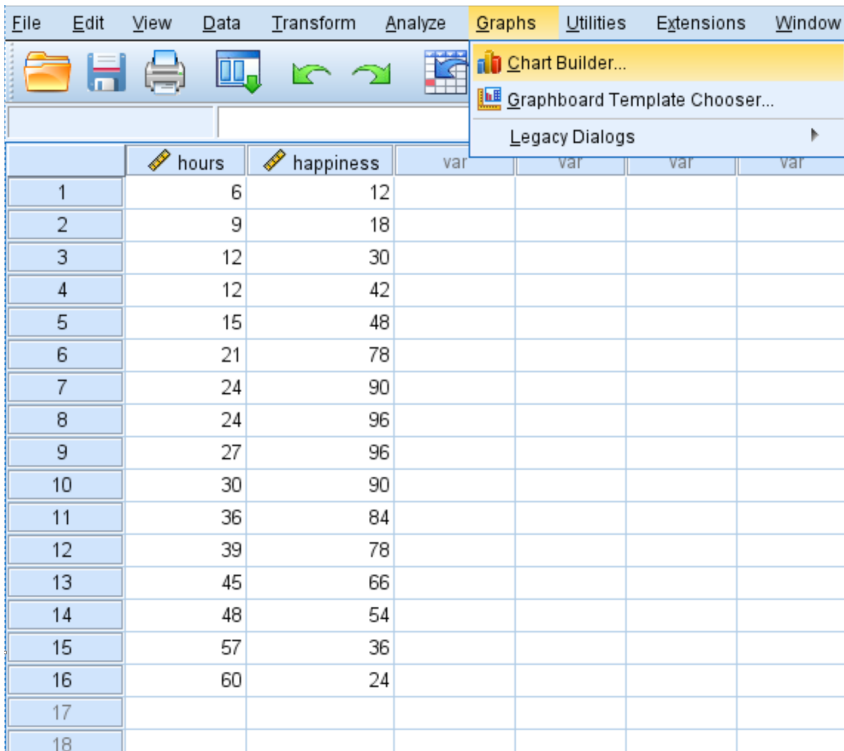


The screenshot shows the SPSS software interface with a menu bar (File, Edit, View, Data, Transform, Analyze, Graphs, Utilities) and a toolbar. Below the toolbar is a data table with the following data:

	hours	happiness	var	var
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		
17				
18				

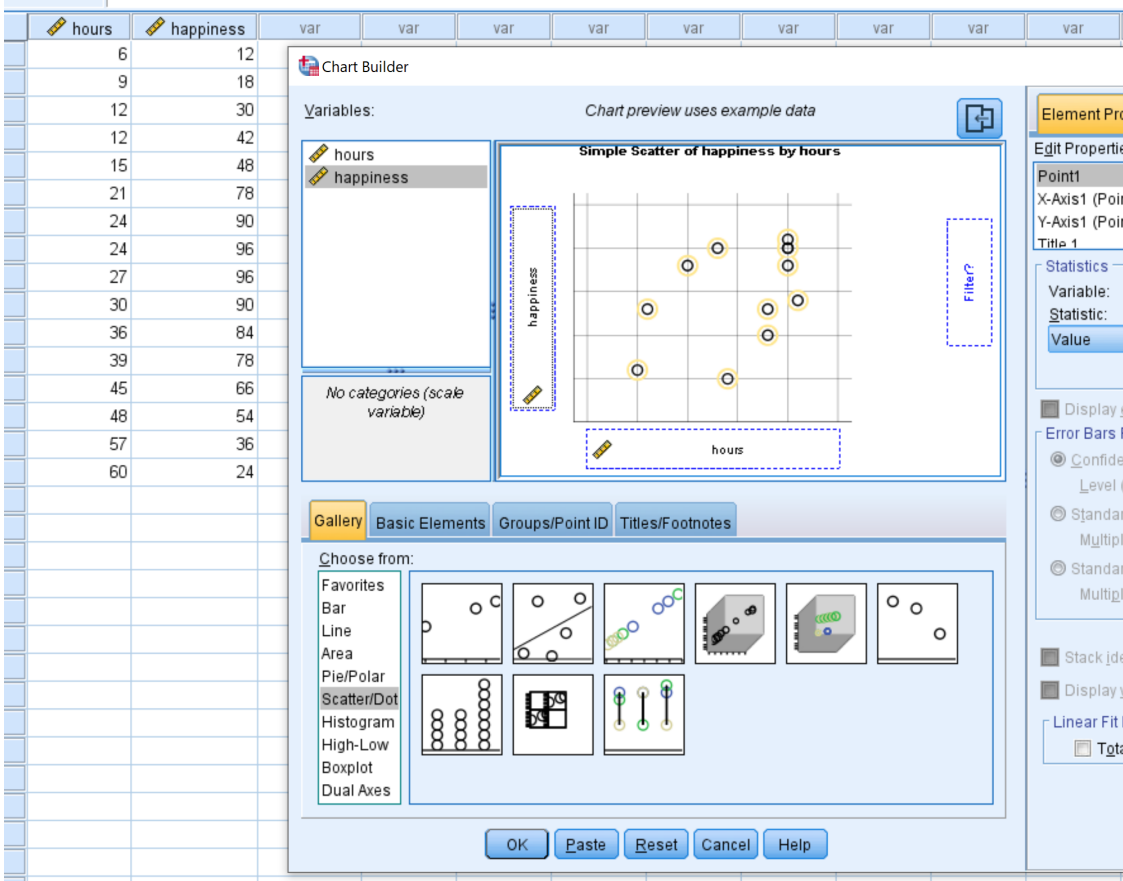
The initial and most critical step in any regression analysis is visualizing the raw data. This step helps confirm the underlying structure of the relationship and validates the choice of a quadratic model over a simpler linear alternative.

To generate a [scatterplot](#) in SPSS, navigate to the menu: Click the **Graphs** tab, then select **Chart Builder**.



	hours	happiness	var	var	var	var
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				
17						
18						

In the Chart Builder interface, select **Scatter/Dot** from the **Choose from** list at the bottom. Drag the **Simple Scatter** chart type into the main editing window. Position the independent variable, **hours**, onto the x-axis, and the dependent variable, **happiness**, onto the y-axis. Once configured, click **OK** to generate the plot.



The resulting scatterplot visually confirms the hypothesized relationship, revealing a clear curvature:

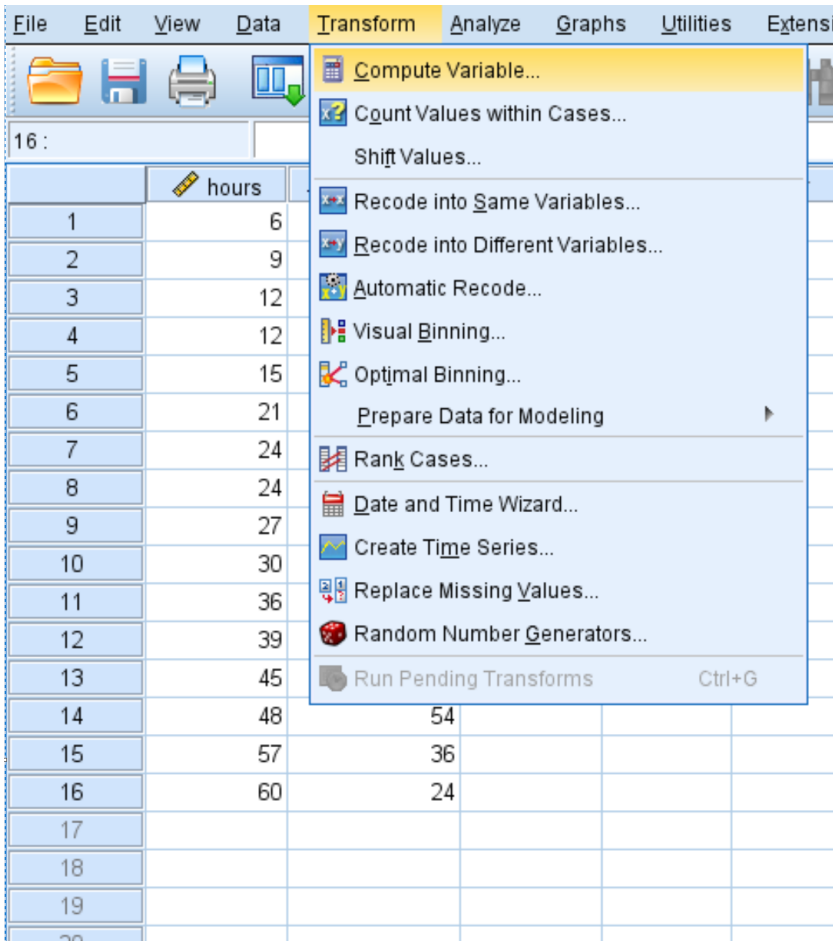


The scatterplot demonstrates that happiness increases with hours worked up to a certain point (around 30-40 hours), after which happiness begins to decline as hours worked continue to rise. This inverted U-shape is the definitive sign that **quadratic regression** is the appropriate modeling technique for this dataset, as it can capture this turning point far more effectively than a linear model.

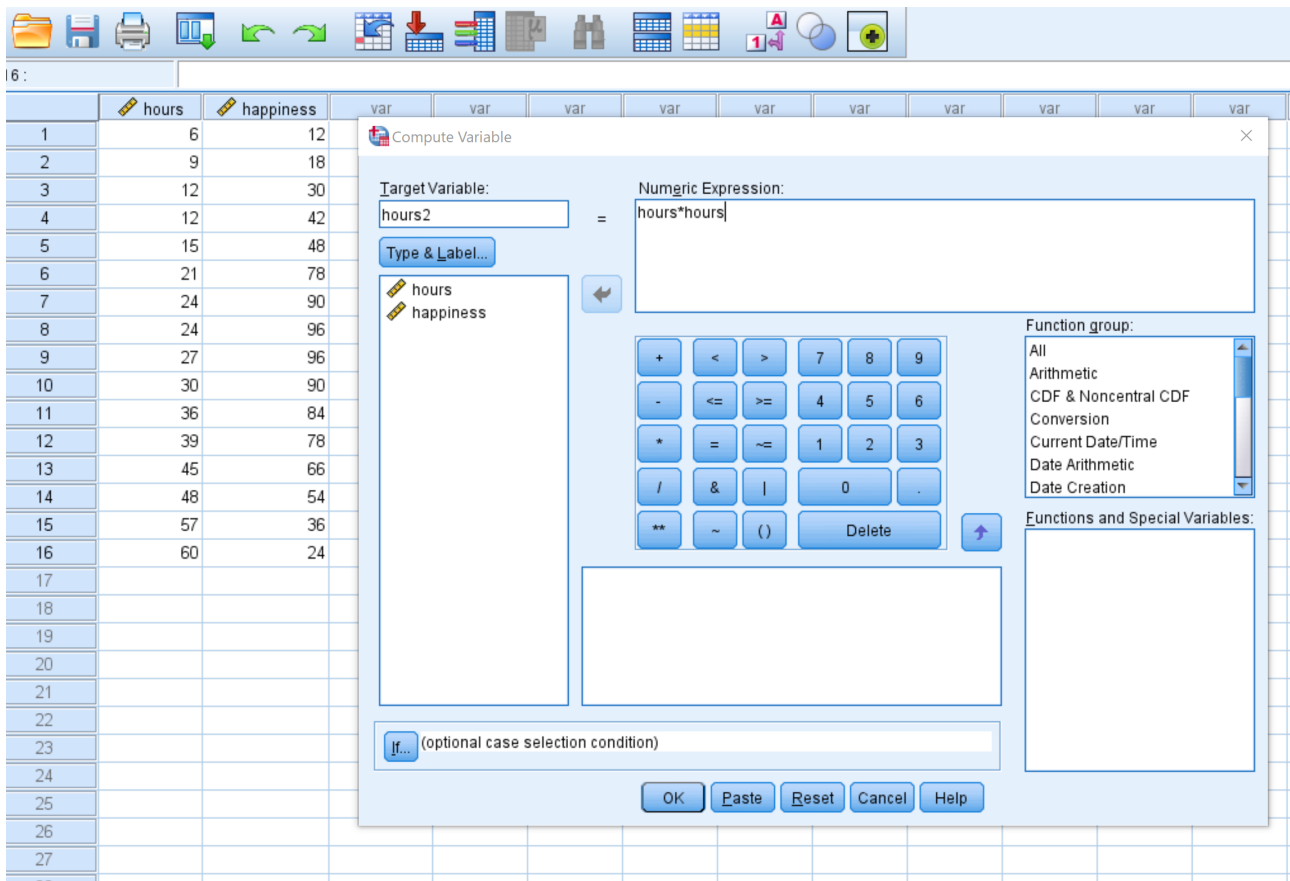
Preparing the Data for Quadratic Regression (Step 2)

Although we are modeling a quadratic relationship, the process in SPSS utilizes the standard linear regression procedure. To achieve this, we must first create a new variable that represents the squared term (X^2) of our independent variable (hours).

To create this new squared variable, navigate to the menu: Click the **Transform** tab, then select **Compute variable**.



In the Compute Variable dialog box, define the new squared term as follows: Name the **Target Variable** as **hours2**. In the **Numeric Expression** box, input the expression **hours*hours**. This creates the necessary quadratic term.



The screenshot shows the SPSS 'Compute Variable' dialog box overlaid on a data view. The data view has columns for 'hours' and 'happiness'. The dialog box is configured to create a new variable named 'hours2' with the numeric expression 'hours*hours'. The 'Function group' is set to 'All'. The 'If...' field is empty, indicating no case selection conditions. The 'OK' button is highlighted.

	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
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		

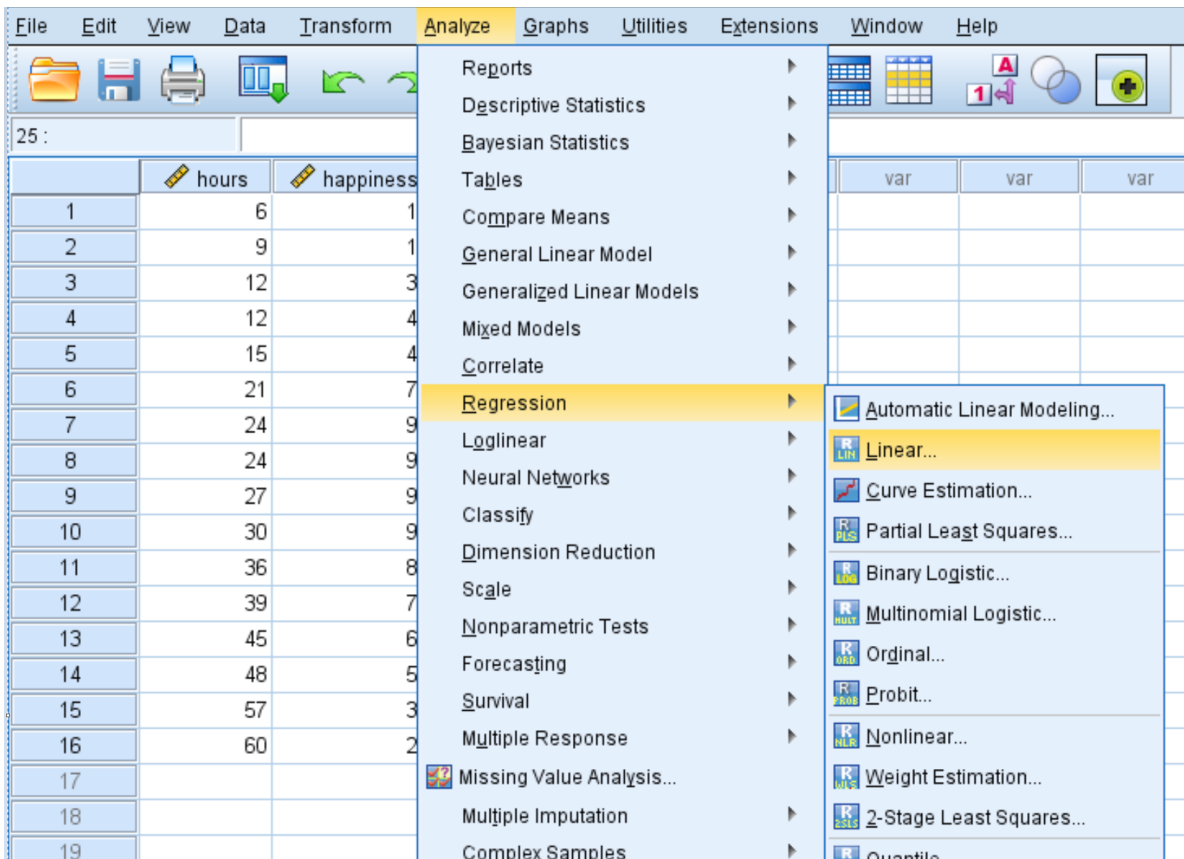
After clicking **OK**, a new column labeled **hours2** will appear in your Data View. This column contains the squared values of the original hours variable, preparing the dataset for the multiple linear regression procedure that models the quadratic curve.

	hours	happiness	hours2	var
1	6	12	36.00	
2	9	18	81.00	
3	12	30	144.00	
4	12	42	144.00	
5	15	48	225.00	
6	21	78	441.00	
7	24	90	576.00	
8	24	96	576.00	
9	27	96	729.00	
10	30	90	900.00	
11	36	84	1296.00	
12	39	78	1521.00	
13	45	66	2025.00	
14	48	54	2304.00	
15	57	36	3249.00	
16	60	24	3600.00	
17				
18				
19				

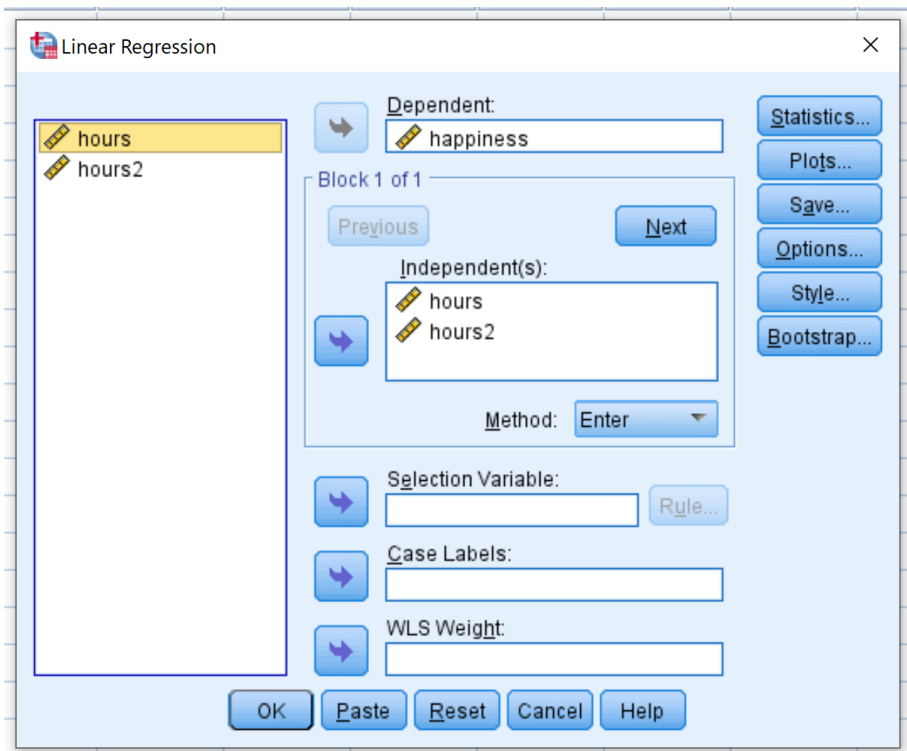
Executing the Quadratic Regression in SPSS (Step 3)

With the squared term now computed, we can proceed to run the regression. Since the quadratic model $Y = \beta_0 + \beta_1X + \beta_2X^2$ is statistically equivalent to a multiple linear regression where X and X^2 are treated as separate predictors, we use the standard linear regression dialogue in SPSS.

Navigate to the menu: Click on the **Analyze** tab, then **Regression**, and finally select **Linear**.



In the Linear Regression dialogue box, correctly assign your variables: Drag the dependent variable, **happiness**, into the box labeled **Dependent**. Then, drag **both** the original independent variable, **hours**, and the newly created squared variable, **hours²**, into the box labeled **Independent(s)**. Ensure that the default method, "Enter," is selected, and click **OK** to execute the analysis.



Detailed Interpretation of Regression Output (Step 4)

Upon execution, SPSS generates several output tables crucial for evaluating the model fit and determining the regression equation. We must examine three key tables in succession: Model Summary, ANOVA, and Coefficients.

Interpreting the Model Summary

The **Model Summary** table provides an overview of how well the combination of predictor variables explains the variation in the dependent variable.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.954 ^a	.909	.895	9.519

a. Predictors: (Constant), hours2, hours

Key metrics in this table include:

R: The multiple correlation coefficient, representing the linear relationship between the predictor

variables and the dependent variable.

R Square: This value, also known as the Coefficient of Determination, indicates the proportion of the variance in the response variable that is explained by the explanatory variables. In this example, the **R Square** value is **0.909**. This means that **90.9%** of the total variation observed in happiness levels can be accounted for by the combined effect of the variables **hours** and **hours²**. This high value suggests an excellent model fit.

Adjusted R Square: This is a modified version of R Square that adjusts for the number of predictors in the model. It is particularly useful when comparing models with different numbers of predictors.

Std. Error of the Estimate: This measure represents the average distance that the observed values fall from the estimated regression line (or curve). It serves as an indicator of the model's precision. Here, the observed happiness levels fall an average of **9.519** units from the fitted quadratic curve.

Interpreting the ANOVA Table

The **ANOVA** (Analysis of Variance) table assesses the overall statistical significance of the regression model. It determines whether the predictors, as a group, significantly improve the prediction of the dependent variable compared to a model with no predictors.

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	11797.700	2	5898.850	65.095	.000 ^b
	Residual	1178.050	13	90.619		
	Total	12975.750	15			

a. Dependent Variable: happiness

b. Predictors: (Constant), hours2, hours

The critical components of the ANOVA table are:

F Statistic: This is the overall F ratio for the regression model, calculated by dividing the Mean Square Regression by the Mean Square Residual. A larger F value indicates that the model explains more variance than is unexplained.

Sig. (p-value): This is the probability associated with the overall F statistic. It tests the null hypothesis that all regression coefficients are zero. In this case, the **p-value** is **0.000** (which SPSS reports when the value is less than 0.0005). Since 0.000 is far below the conventional significance level of 0.05, we conclude that the regression model, including **hours** and **hours²**, is statistically significant in predicting happiness.

Interpreting the Coefficients Table

The **Coefficients** table is the most important output for constructing the final regression equation and understanding the unique contribution of each predictor.

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-30.253	8.766		-3.451	.004
	hours	7.173	.632	4.145	11.356	.000
	hours2	-.107	.009	-4.127	-11.308	.000

a. Dependent Variable: happiness

We utilize the values found in the **Unstandardized B** column to formulate the estimated quadratic regression equation for this dataset:

The constant (Intercept) is -30.253. The coefficient for **hours** is 7.173. The coefficient for **hours²** is -0.107.

Estimated happiness level = **-30.253 + 7.173*(hours) - 0.107*(hours²)**

The negative coefficient on the squared term (-0.107) confirms the inverted U-shape observed in the scatterplot, indicating that the curve reaches a maximum point. We can now use this robust equation to predict happiness based on working hours. For example:

For an individual who works **60 hours** per week, the estimated happiness level is calculated as:

Estimated happiness level = $-30.253 + 7.173*(60) - 0.107*(60^2) = 14.97$.

Conversely, an individual working the optimal number of **30 hours** per week is predicted to have a much higher happiness level:

Estimated happiness level = $-30.253 + 7.173*(30) - 0.107*(30^2) = 88.65$.

Reporting the Findings (Step 5)

The final step in the analysis is to formally report the results of the quadratic regression, summarizing the statistical findings and the resulting predictive equation. Statistical reports should be clear, concise, and include the key metrics derived from the SPSS output (F-statistic, degrees of freedom, p-value, and R²).

A standard way to report these results might look like this:

A quadratic regression analysis was conducted to quantify the non-linear relationship between the number of hours worked per week and the corresponding self-reported happiness level (measured from 0 to 100), based on a sample of 16 individuals.

The overall model was found to be statistically significant, indicating that the explanatory variables (hours and hours²) reliably predict happiness ($F(2, 13) = 65.095, p < 0.001$).

Combined, the linear and quadratic terms accounted for a substantial 90.9% of the variability in happiness ($R^2 = 0.909$). The statistical significance of the quadratic term (hours²) confirms the inverted U-shaped relationship, where optimal happiness is achieved at intermediate working hours.

The resulting estimated regression equation is:

$$\text{Estimated happiness level} = -30.253 + 7.173(\text{hours}) - 0.107(\text{hours}^2)$$