

How to Fit a Quadratic Regression Curve in Excel

Authored by
stats writer

December 30, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Fit a Quadratic Regression Curve in Excel*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=109877>

Quadratic regression is a sophisticated type of regression analysis employed when seeking to model nonlinear relationships between variables. Unlike standard linear regression, which assumes a straight-line fit, quadratic regression fits the data to a second-order polynomial equation, resulting in a curve. While specialized statistical software often handles complex modeling, Microsoft Excel provides two straightforward methods for executing this analysis: using the charting Trendline feature for visualization, or utilizing the powerful Data Analysis ToolPak for detailed statistical output. The Data Analysis approach is preferred for rigorous analysis as it provides the necessary coefficients and diagnostic statistics for model evaluation and prediction.

To perform quadratic regression using the visual method in Excel, one must first generate a scatter plot of the observed data. Subsequently, the "Add Trendline" option is selected, and the user specifies a "Polynomial" trendline of the second degree. This generates the equation for the fitted curve directly on the chart, allowing for visual assessment and simple prediction. However, for a deep dive into the statistical significance and precise model fit, activating and utilizing the Data Analysis add-in is essential. This detailed guide focuses on the comprehensive method using the Data Analysis ToolPak, ensuring you obtain not only the curve but also the rigorous quantitative metrics required for expert-level interpretation.

The Necessity of Quadratic Regression

Regression is a fundamental statistical technique designed to quantify and explain the relationship between one or more predictor variables (also known as independent variables) and a single response variable (or dependent variable). In many real-world scenarios, the relationship between these variables is not a simple proportional increase or decrease. When the relationship is consistently proportional, standard linear regression, which adheres to the formula $Y = b_0 + b_1X$, is the appropriate choice. Linear models are useful for describing phenomena where an increase in effort leads to a consistent, proportional increase in outcome, such as predicting a student's exam score based on their study hours within a reasonable range.

However, statistical modeling frequently encounters diminishing returns, saturation points, or inverse effects, which manifest as nonlinear relationships. When plotted, these relationships do not follow a straight line but instead trace a curve. One of the most common and interpretable types of nonlinearity is the quadratic relationship. This relationship is characterized by a curve that resembles a 'U' shape or an inverted 'U' shape, meaning the effect of the predictor variable changes direction after reaching an apex or nadir.

The need for quadratic regression arises precisely when these turning points occur. For instance, consider the relationship between the number of hours spent working per week and an employee's self-reported job satisfaction. Initially, working more hours might lead to increased satisfaction due to greater fulfillment and achievement. Yet, once the working hours exceed a sustainable threshold

(say, 50 or 60 hours per week), additional work leads to exhaustion, stress, and subsequently, a sharp decrease in reported happiness. In this situation, a linear model would grossly misrepresent the data, providing inaccurate predictions for both low and high working hours. A quadratic model, which includes a squared term, elegantly captures this parabolic relationship, offering a far better fit for predicting the response variable across the entire range of predictor values.

Structuring Data for Quadratic Modeling

To illustrate the application of quadratic regression, let us use a practical example. Suppose we have collected data on the number of hours worked per week and the corresponding reported happiness level (measured on a validated scale from 0 to 100) for 16 different subjects. This data set is typical of what we would encounter in behavioral science or human resources research, where optimal performance or satisfaction levels often follow a non-linear curve.

The initial step in any regression analysis is organizing your data in Excel. Ensure your predictor variable (Hours Worked) is in one column (Column A) and your response variable (Happiness Level) is in the adjacent column (Column B). For our example, the data looks like this:

	A	B	C
1	Hours worked	Reported Happiness	
2	6	12	
3	9	18	
4	12	30	
5	12	42	
6	15	48	
7	21	78	
8	24	90	
9	24	96	
10	27	96	
11	30	90	
12	36	84	
13	39	78	
14	45	66	
15	48	54	
16	57	36	
17	60	24	
18			

Before proceeding with the regression, the critical difference between linear and quadratic analysis in preparation is the requirement for a squared term. The mathematical form of a quadratic regression equation is $Y = b_0 + b_1X + b_2X^2$. Notice that we have two predictor terms on the right

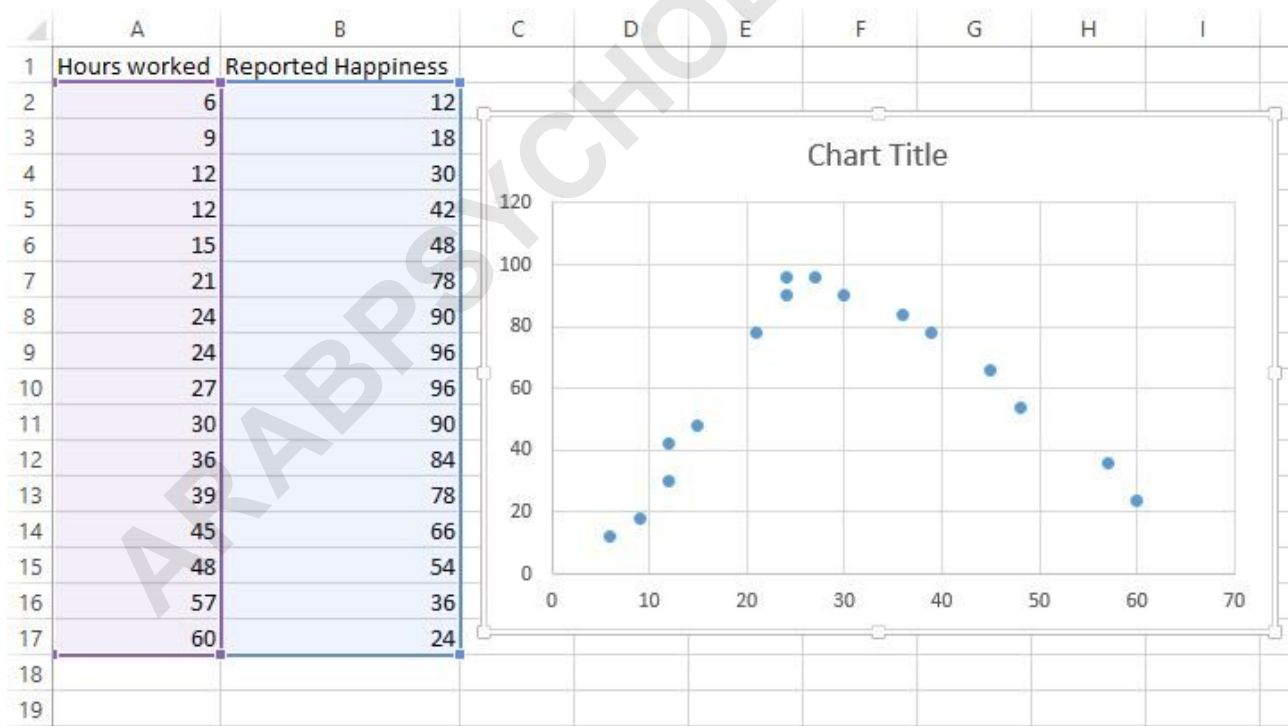
side: the original variable (X) and the squared variable (X^2). To run this model using Excel's Data Analysis ToolPak, we must explicitly create a new column containing the squared values of the predictor variable. This means transforming the original predictor variable into two distinct columns of independent variables: the linear term and the quadratic term.

Visual Assessment with a Scatter Plot

The subsequent step, essential for verifying the appropriate model choice, is creating a scatter plot. Visual inspection is the quickest way to confirm if a linear model is inadequate and if a quadratic model is potentially better suited. To generate the plot, highlight the data range containing both the predictor and response variables (cells **A2:B17** in our example).

Navigate to the INSERT tab on the Excel ribbon, and locate the Scatter chart option within the Charts group. Clicking this option will instantly produce a visual representation of the relationship between hours worked and reported happiness.

The resulting graph immediately reveals the underlying pattern:



As clearly demonstrated by the curvature, the relationship is highly non-linear; the data points trace an inverted 'U' shape. This observation confirms that a standard linear regression model would fail to capture the trend accurately. The data points show happiness peaking around 35 hours worked and then declining. This pattern makes the dataset an ideal candidate for quadratic regression,

which is specifically designed to model such curvilinear relationships.

Calculating the Squared Predictor Variable

As established, Excel's Data Analysis ToolPak requires all independent variables to be explicitly provided in separate columns. Since our quadratic model requires X and X^2 , we must calculate the X^2 term. First, for organizational clarity, it is helpful to shift the Happiness Level (Response Variable) data from Column B to Column C. This ensures the two required predictor variables (X and X^2) are adjacent to each other.

After moving the response variable, Column B is now available for the squared term. We label Column A as "Hours Worked (X)," Column B as "Hours Worked Squared (X^2)," and Column C as "Happiness Level (Y)."

	A	B	C	D
1	Hours worked		Reported Happiness	
2	6		12	
3	9		18	
4	12		30	
5	12		42	
6	15		48	
7	21		78	
8	24		90	
9	24		96	
10	27		96	
11	30		90	
12	36		84	
13	39		78	
14	45		66	
15	48		54	
16	57		36	
17	60		24	
18				

Next, we calculate the squared values. In cell **B2**, type the formula: `=A2^2`. This formula instructs Excel to square the value in cell A2 (6), yielding 36. After entering the formula, use the fill handle (the small square in the bottom-right corner of cell B2) and drag it down to apply the formula to all corresponding rows in Column B. This action populates the entire column with the squared values of the Hours Worked variable.

Once this step is completed, your data table will be structured correctly, with the response variable (Y) in the rightmost column and the two required predictor variables (X and X^2) grouped together

immediately to its left. This arrangement is crucial for correctly specifying the input range when running the Data Analysis tool.

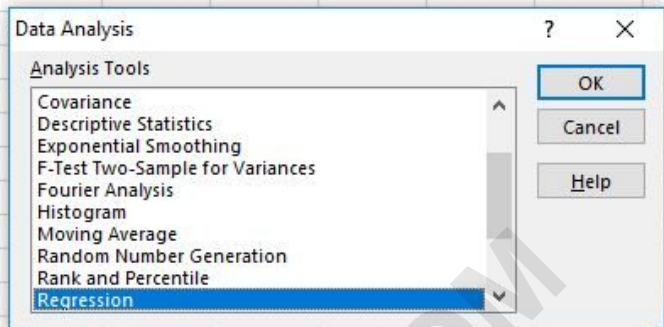
	A	B	C	D
1	Hours worked	Hours worked^2	Reported Happiness	
2	6	36	12	
3	9	81	18	
4	12	144	30	
5	12	144	42	
6	15	225	48	
7	21	441	78	
8	24	576	90	
9	24	576	96	
10	27	729	96	
11	30	900	90	
12	36	1296	84	
13	39	1521	78	
14	45	2025	66	
15	48	2304	54	
16	57	3249	36	
17	60	3600	24	
18				
19				

Executing the Regression using the Data Analysis ToolPak

With the data correctly prepared, we can now fit the quadratic regression model. Access the DATA tab on the ribbon and look for the Data Analysis option, typically located on the far right. If this option is missing, you must first enable the Analysis ToolPak Add-in through Excel Options.

Clicking Data Analysis opens a dialog box listing various statistical procedures. Select Regression from the list and click OK.

	A	B	C	D	E	F	G	H	I
1	Hours worked	Hours worked^2	Reported Happiness						
2	6	36	12						
3	9	81	18						
4	12	144	30						
5	12	144	42						
6	15	225	48						
7	21	441	78						
8	24	576	90						
9	24	576	96						
10	27	729	96						
11	30	900	90						
12	36	1296	84						
13	39	1521	78						
14	45	2025	66						
15	48	2304	54						
16	57	3249	36						
17	60	3600	24						
18									



The next dialog box requires careful input specification. For a quadratic model, the input ranges must be defined as follows:

Input Y Range: This is the range containing the response variable (Happiness Level). Select the data from Column C, including the header (e.g., C1:C17).

Input X Range: This is the most critical step for quadratic regression. Since we are fitting $Y = b_0 + b_1X + b_2X^2$, we must select **both** the linear term (X) and the squared term (X^2) columns simultaneously. Select the data range covering both Column A and Column B, including headers (e.g., A1:B17).

Labels: Check this box since you included the header rows in your ranges.

Output Options: Select a suitable location for the results, such as a New Worksheet Ply, which keeps the voluminous output separate from the raw data.

After configuring these settings, click OK to run the regression:

	A	B	C	D	E	F	G	H	I	J
1	Hours worked	Hours worked^2	Reported Happiness							
2	6	36	12							
3	9	81	18							
4	12	144	30							
5	12	144	42							
6	15	225	48							
7	21	441	78							
8	24	576	90							
9	24	576	96							
10	27	729	96							
11	30	900	90							
12	36	1296	84							
13	39	1521	78							
14	45	2025	66							
15	48	2304	54							
16	57	3249	36							
17	60	3600	24							
18										
19										
20										
21										

Regression [?] [X]

Input

Input Y Range: [icon]

Input X Range: [icon]

Labels Constant is Zero

Confidence Level: %

Output options

Output Range: [icon]

New Worksheet Ply:

New Workbook

Residuals

Residuals Residual Plots

Standardized Residuals Line Fit Plots

Normal Probability

Normal Probability Plots

[OK] [Cancel] [Help]

Interpreting the Regression Output

The Excel output provides a detailed summary of the quadratic model's performance, organized into three main tables: Regression Statistics, ANOVA (Analysis of Variance), and the Regression Coefficients table. Understanding these metrics is vital for determining the model's validity and predictive power.

E	F	G	H	I	J	K	L	M
SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.953526							
R Square	0.909211							
Adjusted R Square	0.895244							
Standard Error	9.519413							
Observations	16							
ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	2	11797.7	5898.85	65.09491	1.69E-07			
Residual	13	1178.05	90.61923					
Total	15	12975.75						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	-30.2529	8.766065	-3.45113	0.004299	-49.1908	-11.3149	-49.1908	-11.3149
Hours worked	7.173061	0.631631	11.35642	4.04E-08	5.808506	8.537616	5.808506	8.537616
Hours worked^2	-0.10699	0.009461	-11.3078	4.25E-08	-0.12743	-0.08655	-0.12743	-0.08655

Key Statistics from the Output Summary:

R Square (Coefficient of Determination): This value represents the proportion of the variance in the response variable (Happiness) that is explained by the predictor variables (Hours Worked and Hours Worked²). In our example, the R Square is **0.9092**. This is an exceptionally strong result, indicating that 90.92% of the variation in reported happiness levels can be accounted for by the quadratic model relating to the number of hours worked. Models with high R-square values demonstrate a strong fit to the observed data.

Standard Error: The Standard error of the regression (also known as the root mean square error) quantifies the average distance that the observed data points fall from the fitted regression line. A smaller standard error signifies a tighter fit. Here, the standard error is **9.519 units**, meaning the model's predictions are, on average, off by about 9.5 units on the happiness scale.

Evaluating Model Significance and Coefficients

The ANOVA table allows us to evaluate the overall statistical significance of the regression model.

F Statistic: Calculated as the ratio of the Mean Square (MS) Regression to the MS Residual, the F Statistic tests the null hypothesis that all regression coefficients (excluding the intercept) are zero.

Essentially, it determines if the entire model is useful or if it fits the data no better than a simple mean. Our F statistic is **65.09**, with a corresponding p-value (Significance F) less than 0.0001. Since this p-value is far smaller than the standard significance level ($\alpha = 0.05$), we strongly reject the null hypothesis, confirming that the quadratic regression model as a whole is statistically significant and provides valuable explanatory power.

Regression Coefficients: The final table provides the intercept (b_0) and the estimated coefficients for each predictor variable (b_1 for X and b_2 for X^2). These values are the constants needed to formulate the estimated regression equation, which is the core output of the analysis.

The general form of the quadratic regression equation is:

$$\hat{y} = b_0 + b_1x_1 + b_2x_1^2$$

By extracting the values from the Coefficients column in the Excel output, we formulate our specific estimated equation:

Intercept (b_0): -30.252

Hours Worked (b_1): 7.173

Hours Worked Squared (b_2): -0.106

Therefore, the estimated regression equation for our data is:

$$\text{Reported Happiness Level} = -30.252 + 7.173(\text{Hours worked}) - 0.106(\text{Hours worked})^2$$

Using the Quadratic Model for Prediction

The final step in the quadratic regression process is using the derived equation for predicting the response variable based on new values of the predictor variable. The model allows us to forecast outcomes within the observed range of the independent variable, though caution should be exercised when extrapolating beyond that range.

For instance, suppose we want to predict the expected happiness level of an individual who works exactly 30 hours per week. We substitute $X = 30$ into our derived equation:

$$\text{Substitute } X = 30: \text{Happiness} = -30.252 + 7.173(30) - 0.106(30)^2$$

$$\text{Calculate the linear term: } 7.173 * 30 = 215.19$$

$$\text{Calculate the quadratic term: } 0.106 * (900) = 95.4$$

$$\text{Solve the equation: Happiness} = -30.252 + 215.19 - 95.4$$

The calculation yields an expected reported happiness level of **89.538**. This demonstrates the power of the quadratic model: it provides a precise, statistically grounded prediction that accounts for the curvilinear nature of the relationship, suggesting that 30 hours per week is near the optimal point for maximizing reported happiness based on this data set. Mastery of this technique in Excel enables robust analysis of non-linear trends in diverse fields, from economics to engineering.

ARABPSYCHOLOGY.COM