# Read and Interpret a Regression Table

Authored by

**Mohammed looti**

November 9, 2025

In the field of statistics, regression analysis is a fundamental technique employed to rigorously analyze and quantify the relationship between one or more potential influencing factors, known as predictor variables, and a specific outcome, termed the response variable.

When statistical software packages--such as R, SAS, or SPSS--are utilized to execute a regression analysis, the primary output is a comprehensive regression table. This table summarizes the complex results of the analysis in a structured format. Developing the ability to accurately read and interpret this output is **essential** for deriving meaningful conclusions from any statistical modeling exercise, ensuring that the insights gained are both valid and actionable.

This detailed tutorial is designed to demystify the process of interpreting regression output. We will navigate through a practical example of a multiple linear regression analysis, providing an in-depth, step-by-step explanation of every critical element found within a standard regression table.

## A Practical Regression Example

To illustrate the interpretation process, consider a hypothetical dataset compiled from 12 distinct students. This dataset captures three key metrics for each student: the total number of hours dedicated to studying, the total count of practice or prep exams taken, and the final score achieved on the main examination.

The raw data collected for this scenario is presented below:

| | Study Hours | Prep Exams | Final Exam Score |
|---|---|---|---|
| Student 1 | 3 | 2 | 76 |
| Student 2 | 7 | 6 | 88 |
| Student 3 | 16 | 5 | 96 |
| Student 4 | 14 | 2 | 90 |
| Student 5 | 12 | 7 | 98 |
| Student 6 | 7 | 4 | 80 |
| Student 7 | 4 | 4 | 86 |
| Student 8 | 19 | 2 | 89 |
| Student 9 | 4 | 8 | 68 |
| Student 10 | 8 | 4 | 75 |
| Student 11 | 8 | 1 | 72 |
| Student 12 | 3 | 3 | 76 |

Our objective is to assess how the amount of time spent studying and the frequency of taking

practice exams collectively influence the final exam score. To achieve this, we employ a multiple linear regression model, designating Study Hours and Prep Exams as our primary **predictor variables**, with the final exam score serving as the crucial **response variable**.

Upon executing the regression analysis using statistical software, the following output table is generated, which we will now proceed to break down section by section:

| *Regression Statistics* | |
|---|---|
| Multiple R | 0.728550902 |
| R Square | 0.530786416 |
| Adjusted R Square | 0.426516731 |
| Standard Error | 7.326766656 |
| Observations | 12 |

ANOVA

| | *df* | *SS* | *MS* | *F* | *Significance F* |
|---|---|---|---|---|---|
| Regression | 2 | 546.53308 | 273.2665 | 5.090515 | 0.033202256 |
| Residual | 9 | 483.1335867 | 53.68151 | | |
| Total | 11 | 1029.666667 | | | |

| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* |
|---|---|---|---|---|---|---|
| Intercept | 66.99010508 | 6.211445265 | 10.78495 | 1.9E-06 | 52.93883968 | 81.04137 |
| Study Hours | 1.299900324 | 0.417012868 | 3.117171 | 0.012375 | 0.356551677 | 2.243249 |
| Prep Exams | 1.117275106 | 1.025145307 | 1.08987 | 0.30409 | -1.201764693 | 3.4363149 |

## Evaluating the Overall Fit and Strength of the Model

The initial segment of the regression table is dedicated to providing statistics that measure the quality of the model's fit. These metrics indicate how effectively the established regression model is able to explain or "fit" the variability observed within the underlying dataset.

The relevant statistics for model fit are typically grouped together, as shown in the image below:

| *Regression Statistics* | |
|---|---|
| Multiple R | 0.728550902 |
| R Square | 0.530786416 |
| Adjusted R Square | 0.426516731 |
| Standard Error | 7.326766656 |
| Observations | 12 |

Understanding each of these summary statistics is paramount to judging the model's explanatory power. We detail the interpretation of each number in this section subsequently:

## Multiple R

The Multiple R represents the correlation coefficient. It is a measure that quantifies the strength of the linear relationship existing between the combined set of predictor variables and the singular response variable. A value of 1 signifies a perfect positive linear relationship, whereas a value of 0 indicates the complete absence of any linear association. Mathematically, the Multiple R is derived by calculating the square root of the R-squared value.

In our current example, the calculated value for the Multiple R is **0.72855**. This figure suggests a fairly **strong linear correlation** between the combination of predictors--study hours and prep exams--and the outcome, which is the final exam score.

## R-Squared (Coefficient of Determination)

The R-squared value, or the coefficient of determination, is a critical metric whose range spans from 0 to 1. An R-squared of 0 implies that the predictor variables offer no explanatory power for the variation in the response variable. Conversely, an R-squared of 1 signifies that the response variable can be predicted perfectly, without any error, solely by the predictor variable.

For this specific model, the R-squared is **0.5307**. Interpreting this, we can state that **53.07% of the total variance** observed in the final exam scores can be statistically accounted for or explained by the combined effect of the number of hours studied and the number of prep exams taken. (Related: What is a Good R-squared Value?)

## Adjusted R-Squared

The Adjusted R-squared is a modification of the standard R-squared statistic. It incorporates an

adjustment based on the number of predictors included in the model. Because it penalizes the inclusion of superfluous variables, the Adjusted R-squared is invariably lower than the unadjusted R-squared. This metric is particularly useful when comparing the relative **goodness-of-fit** between different regression models that utilize varying numbers of predictors.

In the context of our current analysis, the Adjusted R-squared is calculated as **0.4265**.

## Standard Error of the Regression

The Standard Error of the Regression, sometimes referred to as the standard error of the estimate, provides a measure of the **average distance** that the observed data points fall from the estimated regression line. Essentially, it quantifies the typical prediction error of the model in the units of the response variable. A smaller standard error generally indicates a more precise model fit.

In this analysis, the Standard Error of the Regression is **7.3267**. This means, on average, the actual observed values fall approximately 7.33 units from the regression line. (Related: Understanding the Standard Error of the Regression)

## Observations

This entry is the most straightforward, indicating the sample size used in the analysis. It is simply the total count of data points, or observations, that were included in the dataset used to build the model.

For our student performance example, the total number of observations is **12**.

## Assessing the Global Significance of the Regression Model (ANOVA Table)

The subsequent section of the regression output is often presented in an Analysis of Variance (ANOVA) format. This segment details the metrics required to test the overall significance of the entire regression model, determining if the predictors, taken together, contribute significantly to predicting the response variable.

This section includes the degrees of freedom, the sum of squares, mean squares, the F statistic, and the overall significance level (P-value).

ANOVA

| | *df* | *SS* | *MS* | *F* | *Significance F* |
|---|---|---|---|---|---|
| Regression | 2 | 546.53308 | 273.2665 | 5.090515 | 0.033202256 |
| Residual | 9 | 483.1335867 | 53.68151 | | |
| Total | 11 | 1029.666667 | | | |

A clear understanding of these values is essential for hypothesis testing concerning the model's utility:

## Regression Degrees of Freedom (df)

The regression degrees of freedom (df) is calculated as the number of regression coefficients minus one. In this multiple regression example, we have an intercept term and two predictor variables, totaling three coefficients. Therefore, the regression degrees of freedom is calculated as 3 - 1 = **2**.

## Total Degrees of Freedom (df)

The total degrees of freedom is simply the total number of observations (N) minus one. Since our dataset comprises 12 observations, the total degrees of freedom is calculated as 12 - 1 = **11**.

## Residual Degrees of Freedom (df)

The residual degrees of freedom represents the remaining degrees of freedom after accounting for the model's predictors. It is calculated by subtracting the regression df from the total df. In this example, the residual degrees of freedom is 11 - 2 = **9**.

## Mean Squares (MS)

Mean Squares are obtained by dividing the Sum of Squares (SS) by their corresponding degrees of freedom. These values are crucial stepping stones for calculating the F statistic.

The regression Mean Squares (MS) is calculated as Regression SS / Regression df: 546.53308 / 2 = **273.2665**.
The residual Mean Squares (MS) is calculated as Residual SS / Residual df: 483.1335 / 9 = **53.68151**.

## F Statistic

The F statistic is the ratio of the regression Mean Squares to the residual Mean Squares (MS Regression / MS Residual). This statistic indicates whether the regression model provides a significantly better fit to the data than a model that contains no independent variables. It tests the overall utility of the model.

In this example, the F statistic is calculated as 273.2665 / 53.68151 = **5.09**.

## Significance of F (P-value)

The last value in the table is the p-value associated with the F statistic. To evaluate the significance of the overall regression model, this p-value is compared to a predetermined significance level (e.g., 0.05).

If the p-value is less than the significance level, we conclude that the model fits the data significantly better than a model with no predictors. In this example, the p-value is **0.033**. Since 0.033 is less than the common significance level of 0.05, we conclude that the overall regression model is statistically **significant**.

## Analyzing Individual Predictor Significance and Coefficients

The final section of the regression table details the specific results for each term included in the model, allowing for individual inference about the predictors. This output includes the coefficient estimates, their standard errors, the t-statistics, individual p-values, and the associated confidence intervals for each term.

| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* |
|---|---|---|---|---|---|---|
| Intercept | 66.99010508 | 6.211445265 | 10.78495 | 1.9E-06 | 52.93883968 | 81.04137 |
| Study Hours | 1.299900324 | 0.417012868 | 3.117171 | 0.012375 | 0.356551677 | 2.243249 |
| Prep Exams | 1.117275106 | 1.025145307 | 1.08987 | 0.30409 | -1.201764693 | 3.4363149 |

## Coefficients

The coefficient values are the numerical constants required to formulate the estimated regression equation. For a multiple linear regression with two predictors, the equation takes the general form:

**yhat = b0 + b1x1 + b2x2**

Using the coefficients derived from our example, the specific estimated equation for predicting the

final exam score is:

**final exam score = 66.99 + 1.299(Study Hours) + 1.117(Prep Exams)**

Each individual coefficient is interpreted as the average increase in the response variable for each one-unit increase in a given predictor variable, assuming that all other predictor variables are held constant. For example, for each additional hour studied, the average expected increase in final exam score is **1.299 points**, provided that the number of prep exams taken remains unchanged.

The intercept (66.99) is interpreted as the expected average final exam score for a student who studies for zero hours and takes zero prep exams. It is important to be careful when interpreting the intercept, especially if the value is negative or if zero falls outside the observed data range, as it may lack a practical real-world meaning.

## Standard Error, T-statistics, and P-values

The Standard Error is a measure of the uncertainty around the estimate of the coefficient for each variable. It reflects the precision of the estimate.

The T-statistic is calculated simply by dividing the coefficient by its standard error. For example, the T-statistic for Study Hours is 1.299 / 0.417 = **3.117**.

The final column shows the p-value associated with the T-statistic. This value determines the statistical significance of that specific predictor within the model. We observe that the p-value for Study Hours is **0.012**, indicating it is a significant predictor. The p-value for Prep Exams is **0.304**, suggesting that Prep Exams is **not** a statistically significant predictor when controlling for Study Hours.

## Confidence Interval for Coefficient Estimates

The last two columns provide the lower and upper bounds for a 95% confidence interval for the coefficient estimates. Since our coefficient estimate is derived from a sample, the confidence interval gives a range of likely values for the true population coefficient.

For Study Hours, the 95% confidence interval is (0.356, 2.24). Crucially, this interval does not contain the number 0. This reinforces that the true value for the coefficient of Study Hours is non-zero, confirming its statistical significance.

Conversely, the 95% confidence interval for Prep Exams is (-1.201, 3.436). This interval **does contain the number 0**, meaning that zero is a plausible value for the true population coefficient. This supports the conclusion that Prep Exams is not a statistically significant predictor of final exam scores.

## Summary and Further Reading

Mastering the interpretation of a regression table is a foundational skill in statistics. By systematically examining the model fit statistics, the overall F-test results, and the individual coefficient estimates, analysts can confidently draw conclusions regarding the relationships between variables and the predictive power of their models.

Below are additional resources for deepening your understanding of specific regression concepts:

What is a Good R-squared Value?

Understanding the Standard Error of the Regression