How to Easily Report Your Regression Results

Authored by
stats writer

December 5, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Easily Report Your Regression Results*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=105562

When presenting the findings of any statistical analysis, particularly those derived from regression results, clarity, precision, and completeness are paramount. An effective report serves not only to communicate the observed relationships but also to establish the scientific validity and reliability of the model. It is essential to move beyond simply listing numerical outputs; instead, a strong report weaves those numbers into a coherent narrative that addresses the initial research question.

A comprehensive summary must include several key elements: the specific research question, the model structure, the estimated parameters or coefficients, diagnostic statistics, and the overall goodness-of-fit measures. Furthermore, a meaningful report requires a detailed discussion of the practical implications of the findings, consideration of potential limitations, and, most critically, the inclusion of sufficient methodological detail to allow independent verification and replication of the analysis. Transparency ensures that the analysis stands up to scrutiny and contributes reliably to the body of knowledge.

The Fundamentals of Reporting Regression Results

In the realm of quantitative analysis, linear regression models are fundamental tools used by researchers across various disciplines, from economics to psychology, to quantify and establish the relationship between one or more independent variables (often called predictor variables) and a single dependent variable (the response variable). Properly reporting these results requires a systematic approach that bridges complex statistical calculations with accessible interpretation. The goal is to provide a narrative that is both statistically rigorous and contextually relevant.

The first step involves clearly defining the model used. Whether employing simple linear regression (one predictor) or multiple linear regression (two or more predictors), the report must explicitly state which variables were included and why. This context is crucial for readers to assess the suitability of the model for the research question. Subsequent reporting must then focus on the model's performance, detailing how well the predictors collectively explain the variability in the response variable. This is where statistics like the F-test and the coefficient of determination come into play.

Finally, the report must address the individual contribution of each predictor. This means detailing the regression coefficients (often denoted as $beta$), their standard errors, and their corresponding statistical significance as determined by the t-statistic and the p-value. When interpreting these coefficients, it is vital to relate them back to the real-world units of measurement, explaining, for example, how a one-unit increase in the predictor variable affects the mean change in the response variable, while controlling for other factors in the model.

Standard Template for Simple Linear Regression

For a rigorous yet clear presentation of a simple linear regression analysis--which investigates the relationship between one predictor and one response variable--we can utilize a standardized, fourpart format. This structured approach ensures all necessary statistical components are included and presented logically. The first section introduces the hypothesis, the second presents the specific equation derived from the data, the third confirms the overall model fit, and the final section details the predictive power of the single variable.

We can use the following general format to report the results of a simple linear regression:

Phase 1: Hypothesis and Scope. Simple linear regression was employed to test the hypothesis that significantly predicted .

Phase 2: Model Specification. The fitted regression model derived from the analysis was explicitly stated as: .

Phase 3: Overall Model Significance and Fit. The overall regression equation was found to be statistically significant (R2 = , F(df regression, df residual) = , p =).

Phase 4: Individual Coefficient Interpretation. Specifically, it was determined that significantly predicted (β = , p =).

Reporting Guidelines for Multiple Linear Regression

Multiple linear regression expands upon the simple model by incorporating two or more predictors, allowing for more nuanced modeling of complex relationships and controlling for confounding variables. Reporting these results requires differentiating between the overall model performance and the unique contribution of each predictor, as their collinearity might affect individual significance. The template provided below ensures all these dimensions are covered comprehensively.

We can use the following format to report the results of a multiple linear regression:

Phase 1: Hypothesis and Scope. Multiple linear regression was conducted to determine if the set of predictors, including , , and others, significantly predicted .

Phase 2: Model Specification. The complete fitted regression model established by the data was: .

Phase 3: Overall Model Significance and Fit. The collective predictive power of the model was statistically significant, indicating that the predictors together account for a meaningful proportion of the variance in the response variable (R2 = , F(df regression, df residual) = , p =).

Phase 4: Individual Coefficient Interpretation. It was found that significantly predicted (β = , p =).

Phase 5: Non-Significant Predictors. It was found that did not significantly predict (β = , p =), suggesting its unique contribution is negligible in the presence of the other predictors.

Understanding Key Statistical Components

A high-quality regression report depends entirely on the correct interpretation and presentation of the core statistical outputs. These statistics function as evidence to support or reject the hypothesized relationships. Three measures are central to evaluating the overall model fit and significance: the Coefficient of Determination (R2), the F-statistic, and the model's overall p-value. Misunderstanding these components can lead to flawed conclusions regarding the research hypothesis.

The R2 value, often expressed as a percentage, quantifies the proportion of the variance in the dependent variable that is predictable from the independent variable(s). For example, an R2 of 0.73 means 73% of the variability in the response variable is explained by the model. While a higher R2 generally indicates a better fit, it must be reported alongside the adjusted R2, particularly in multiple regression, to account for the increasing complexity of the model and prevent overestimation of fit due to adding unnecessary variables.

The F-statistic (and its accompanying degrees of freedom, df) tests the null hypothesis that all regression coefficients are zero simultaneously. If the calculated F-statistic is large enough, leading to a small overall p-value (typically $p < 0.05$), we reject the null hypothesis, concluding that the model, as a whole, performs significantly better than a model with no predictors. This is the primary indicator of the collective relevance of the predictor variables.

Detailed Example: Simple Linear Regression Results

To illustrate the application of the simple linear regression template, consider a pedagogical scenario where a university professor wishes to analyze the relationship between the number of hours a student studied and their subsequent exam score. Data is collected from 20 students, and a simple linear regression model is fitted, treating hours studied as the single predictor variable.

The following screenshot shows the output of the regression model, providing the raw statistical data for formal reporting:

D	E	F	G	H	I	J	K	L
SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.8528							
R Square	0.7273							
Adjusted R Square	0.7121							
Standard Error	5.2805							
Observations	20							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	1338.2906	1338.2906	47.9952	0.0000			
Residual	18	501.9094	27.8839					
Total	19	1840.2000						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	67.1617	2.6633	25.2178	0.0000	61.5664	72.7570	61.5664	72.7570
hours	5.2503	0.7578	6.9279	0.0000	3.6581	6.8424	3.6581	6.8424

Here is how to report the results of this simple linear regression model using the standardized format:

Simple linear regression was used to test if hours studied significantly predicted exam score.

The fitted regression model was: **Exam score = 67.1617 + 5.2503 * (hours studied).**

The overall regression was statistically significant (R2 = .73, F(1, 18) = 47.99, p < .000). This indicates that hours studied explains 73% of the variance in exam scores.

It was found that hours studied significantly predicted exam score (β = 5.2503, p < .000). The positive coefficient suggests that for every additional hour studied, the predicted exam score increases by approximately 5.25 points.

Detailed Example: Multiple Linear Regression Results

Next, suppose the professor incorporates a second variable, the number of prep exams taken, to see if it adds unique predictive power to the model beyond hours studied. He collects data for the same 20 students and fits a multiple linear regression model. This allows for a more complex assessment of how multiple factors contribute to the final exam score.

The following screenshot shows the output of the multiple regression model, which now contains separate coefficients and p-values for both predictor variables:

D	E	F	G	H	I	J	K
SUMMARY OUTPUT							
Regression Statistics							
Multiple R	0.857						
R Square	0.734						
Adjusted R Square	0.703						
Standard Error	5.366						
Observations	20						
ANOVA							
	df	*SS*	*MS*	*F*	*Significance F*		
Regression	2	1350.76	675.38	23.46	0.00		
Residual	17	489.44	28.79				
Total	19	1840.20					
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	
Intercept	67.67	2.82	24.03	0.00	61.73	73.61	
hours	5.56	0.90	6.18	0.00	3.66	7.45	
prep_exams	-0.60	0.91	-0.66	0.52	-2.53	1.33	

The formal report must now address the unique contribution of each variable:

Multiple linear regression was used to test if hours studied and prep exams taken significantly predicted exam score.

The fitted regression model was: **Exam Score = 67.67 + 5.56*(hours studied) - 0.60*(prep exams taken).**

The overall regression was statistically significant (R2 = 0.73, F(2, 17) = 23.46, p < .000). The model collectively accounts for 73% of the variance in exam scores.

It was found that hours studied significantly predicted exam score (β = 5.56, p < .000).

It was found that prep exams taken did not significantly predict exam score (β = -0.60, p-value = 0.52).

Interpreting and Contextualizing the Findings

Beyond the presentation of raw statistical outputs, a superior report must dedicate significant space to interpreting the implications of the derived model. Interpretation means explaining what the statistically significant and non-significant coefficients mean in the context of the original research question. For instance, in the multiple regression example, while hours studied is a strong positive predictor, the non-significance of prep exams taken suggests that simply counting the number of prep exams offers little unique predictive value once study time is accounted for. This nuanced interpretation is crucial for deriving practical conclusions.

Furthermore, every report must address the limitations inherent in the analysis. This includes discussing potential issues such as sample size restrictions, the presence of outliers, or violations of key regression assumptions (e.g., linearity, homoscedasticity, and independence of errors). By acknowledging these limitations, the researcher strengthens the credibility of the findings and guides future research toward addressing the identified weaknesses. Failure to mention critical assumptions diminishes the trustworthiness of the reported regression results.

Finally, the discussion section should synthesize the findings, linking them back to relevant existing literature or theoretical frameworks. This step elevates the report from a mere calculation summary to a piece of scholarly work. It explains why the results matter--whether they confirm existing theories, contradict previous findings, or suggest entirely new directions for exploration. This contextualization provides the reader with a full understanding of the analytical contribution.

Best Practices for Ensuring Reproducibility

A fundamental pillar of modern scientific reporting is the commitment to reproducibility. Readers must be given sufficient information to replicate the analysis independently, ensuring the transparency and verifiability of the findings. This often requires documenting aspects that go beyond the final coefficient table, ensuring that the process is as clear as the outcome.

To achieve high standards of reproducibility, the report should include a detailed methodology section covering:

Data Description: Precise details on how the data was collected, the sample size, and how any missing data or outliers were handled.
Software and Packages: Specification of the statistical software package used (e.g., R version 4.3.1, SPSS version 29, or Excel Analysis ToolPak), including any specific libraries or custom functions utilized.
Model Selection Rationale: Justification for choosing the specific form of the linear regression model, especially if transformations or variable interactions were included.

Beyond textual description, best practices increasingly involve sharing the actual syntax, code, or data files, whenever ethically permissible. If direct sharing is not possible, providing the summarized data inputs and the exact function calls used in the statistical software allows another researcher to meticulously check the calculation process and confirm the reported goodness-of-fit measures and coefficients. This robust documentation is key to maintaining scientific integrity.