

how to Interpret ANOVA Results in SAS?

Authored by
stats writer

November 19, 2025

RECOMMENDED CITATION

stats writer (2025). *how to Interpret ANOVA Results in SAS?*. PSYCHOLOGICAL SCALES.
Retrieved from <https://scales.arabpsychology.com/?p=96959>

Welcome to this comprehensive guide dedicated to mastering the interpretation of Analysis of Variance (ANOVA) results within the powerful statistical software environment of SAS. ANOVA is a foundational statistical method used extensively across scientific disciplines to test for differences between group means. Understanding the output generated by SAS is crucial for accurately drawing conclusions from your research data.

This tutorial focuses specifically on the one-way ANOVA, providing a detailed, step-by-step breakdown of how to read the complex tables produced by the **PROC ANOVA** procedure. We will cover everything from setting up the data to dissecting the final F-test, ensuring you can confidently determine whether statistically significant differences exist among your groups.

The Purpose of the One-Way ANOVA

The one-way ANOVA is a fundamental inferential statistical test. Its primary function is to determine if there is a statistically significant difference among the means of three or more independent groups concerning a single continuous dependent variable. It achieves this by partitioning the total variability observed in the dependent variable into two components: the variance explained by the differences between the groups (the model) and the variance unexplained (the error or residual).

When conducting a one-way ANOVA, we are essentially testing the null hypothesis that all group population means are equal against the alternative hypothesis that at least one group mean is different. This distinction is critical because if the overall test is not significant, further pairwise comparisons are generally unwarranted. The strength of the ANOVA lies in its ability to control the Type I error rate (the risk of false positives) compared to performing multiple individual t-tests.

To illustrate the application and interpretation of this test, we will proceed with a detailed case study implemented entirely in SAS. This practical approach will highlight how different components of the SAS output relate directly back to the underlying statistical theory, providing a robust understanding of the results.

Case Study: Comparing Study Methods

Imagine a research scenario where an educational psychologist wishes to compare the effectiveness of three distinct studying methods (Method A, Method B, and Method C) on exam performance. The researcher recruits 30 students and randomly assigns them to one of the three methods. After a period of preparation, all students take the same standardized exam, and their scores are recorded.

The resulting dataset, crucial for our analysis, consists of the studying method (the categorical independent variable) and the exam score (the continuous dependent variable). This structure

perfectly aligns with the requirements for a one-way ANOVA. Since the groups are independent and we have more than two groups, the ANOVA test is the most appropriate statistical tool to determine if the means of the exam scores across the three methods are statistically distinct.

The raw data collected for the 30 participants, showing their assigned method and corresponding score, is presented below. This raw data forms the input for our statistical analysis in SAS, allowing us to generate the necessary output tables for interpretation.

Method A	Method B	Method C
78	81	84
81	83	88
82	83	88
82	85	89
85	86	90
88	88	93
88	90	95
90	91	98

Data Preparation and PROC ANOVA Implementation in SAS

Before executing the analysis, the raw data must be structured into a valid SAS dataset. We use the **DATALINES** statement to input the scores and the corresponding method identifier directly into the environment. This step ensures that SAS correctly recognizes the variables and their types--specifically, that `Method` is a character or categorical variable, and `Score` is a numeric variable.

The following code snippet demonstrates the creation of the dataset named `my_data`. Notice the use of the dollar sign (\$) after `Method` in the `input` statement, which designates it as a character variable, crucial for its role as the grouping factor in the ANOVA model:

```
/*create dataset*/  
data my_data;  
input Method $ Score;  
datalines;  
A 78  
A 81  
A 82  
A 82  
A 85
```

```
A 88  
A 88  
A 90  
B 81  
B 83  
B 83  
B 85  
B 86  
B 88  
B 90  
B 91  
C 84  
C 88  
C 88  
C 89  
C 90  
C 93  
C 95  
C 98  
;  
run;
```

Once the data is prepared, we initiate the ANOVA procedure using **PROC ANOVA**. This procedure is specifically designed for balanced designs (where the number of observations in each group is equal, as is the case here with 8 students per method). The CLASS statement identifies the grouping variable (Method), while the MODEL statement specifies the relationship we are testing, showing that Score is modeled by Method.

Crucially, we include the MEANS statement along with the TUKEY and CLDIFF options. The inclusion of the Tukey post-hoc test is a preparatory measure; it instructs SAS to automatically calculate pairwise mean comparisons only if the overall p-value from the main ANOVA test indicates a statistically significant difference. The CLDIFF option specifically requests confidence intervals for those mean differences, adding valuable context to the results.

```
/*perform one-way ANOVA*/  
proc ANOVA data=my_data;  
class Method;  
model Score = Method;  
means Method / tukey cldiff;
```

run;

Deconstructing the ANOVA Summary Table

The first and most critical piece of output generated by **PROC ANOVA** is the ANOVA summary table itself. This table encapsulates the partition of variance and provides the necessary statistics to test the null hypothesis. Understanding each column is paramount to correct interpretation.

The ANOVA Procedure
Dependent Variable: Score

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	175.5833333	87.7916667	5.26	0.0140
Error	21	350.2500000	16.6785714		
Corrected Total	23	525.8333333			

R-Square	Coeff Var	Root MSE	Score Mean
0.333914	4.698685	4.083941	86.91667

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Method	2	175.5833333	87.7916667	5.26	0.0140

We must meticulously examine the contents of this table, paying close attention to the sources of variation (Model, Error, and Corrected Total) and their associated numerical values. The calculations for these values are based on the total number of observations ($N=24$) and the number of groups ($k=3$). This section walks through the calculation and meaning of the core statistics.

DF (Degrees of Freedom)

The degrees of freedom (DF) represent the number of independent values that can vary within a statistical calculation. In the ANOVA table, DF is calculated for the model (Between Groups) and the error (Within Groups):

DF Model: This represents the degrees of freedom associated with the differences between the study methods. It is calculated as the number of groups minus one ($k - 1$). With three methods, $DF_{Model} = 3 - 1 = 2$.

DF Error: This represents the degrees of freedom associated with the residuals or unexplained

variation within the groups. It is calculated as the total number of observations minus the number of groups ($N - k$). With 24 observations and 3 groups, $DF \text{ Error} = 24 - 3 = 21$.

Corrected Total: This is the total degrees of freedom, calculated as the sum of $DF \text{ Model}$ and $DF \text{ Error}$ ($2 + 21 = 23$), or equivalently, $N - 1$ ($24 - 1 = 23$).

Sum of Squares (SS)

The Sum of Squares measures the variability in the data. SS is partitioned into $SS \text{ Model}$ (variation due to the factor) and $SS \text{ Error}$ (random variation):

Sum of Squares Error (SS Error): The sum of squares associated with the residuals, quantifying the variability within each studying method group. In our output, this value is **350.25**.

Sum of Squares Corrected Total (SS Total): The total variability in the exam scores. This is the sum of $SS \text{ Model}$ and $SS \text{ Error}$. In the example, the total variability is **525.833**.

Mean Square (MS) and the F-Statistic

The Mean Square (MS) is the average variability, calculated by dividing the Sum of Squares by its corresponding degrees of freedom. These values are crucial because they form the ratio used to calculate the F-statistic.

Mean Square Model: This measures the average variability explained by the differences in the study methods. It is calculated as $SS \text{ Model} / DF \text{ Model}$ ($175.583 / 2 = 87.79$).

Mean Square Error: Often referred to as the pooled variance estimate, this measures the unexplained or residual average variability. It is calculated as $SS \text{ Error} / DF \text{ Error}$ ($350.25 / 21 = 16.68$).

F Value: The overall F-statistic for the ANOVA model. This is the test statistic used to assess the null hypothesis. It represents the ratio of the variance explained by the model to the variance unexplained (random error). $F \text{ Value} = \text{Mean Square Model} / \text{Mean Square Error}$ ($87.79 / 16.68 = 5.26$).

Interpreting the Global F-Test and P-Value

The most important value derived from the ANOVA summary table is the p-value, listed in the **Pr > F** column. This value determines whether the calculated F-statistic (5.26) is sufficiently large to reject the hypothesis that all group means are equal. The p-value quantifies the probability of observing our data (or more extreme data) if the null hypothesis were true.

Recall the hypotheses we are testing in a one-way ANOVA:

H0 (Null hypothesis): All group means are statistically equal ($\mu_A = \mu_B = \mu_C$).

HA (Alternative hypothesis): At least one group mean is different from the others.

In our SAS output, the calculated p-value ($Pr > F$) is **0.0140**. We typically compare this value against a predetermined significance level, α , usually set at 0.05. Since 0.0140 is less than 0.05, we meet the criterion for statistical significance. Therefore, we confidently reject the null hypothesis.

The rejection of the null hypothesis leads to the conclusion that there is a statistically significant difference in the mean exam scores among the three studying methods. However, the global F-test does not tell us which specific pairs of groups differ; it only confirms that the groups are not all the same. To pinpoint the exact location of these differences, we must turn to the results of the post-hoc analysis.

Post-Hoc Analysis: Tukey's Multiple Comparison Test

Because our global ANOVA F-test was significant ($p < 0.05$), we proceed to examine the output of the multiple comparison procedure. We requested the Tukey post-hoc test, which is highly recommended for conducting all pairwise comparisons while controlling the family-wise error rate.

The SAS output for the Tukey test, often presented as a matrix or table comparing every pair of means, allows us to directly identify the specific pairs responsible for the overall significant finding. This table provides the mean difference, confidence intervals, and an associated significance marker (usually stars or letters).

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Score

Note: This test controls the Type I experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	21
Error Mean Square	16.67857
Critical Value of Studentized Range	3.56462
Minimum Significant Difference	5.1469

Comparisons significant at the 0.05 level are indicated by ***.

Method Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
C - B	4.750	-0.397	9.897	
C - A	6.375	1.228	11.522	***
B - C	-4.750	-9.897	0.397	
B - A	1.625	-3.522	6.772	
A - C	-6.375	-11.522	-1.228	***
A - B	-1.625	-6.772	3.522	

To determine statistical significance between any two methods, we must look for the significance indicators next to the pairwise comparisons. In this specific SAS output format, significant differences are marked by three asterisks (***), indicating that the difference between those two means is unlikely to be zero.

Upon reviewing the Tukey post-hoc test results, we observe that the comparison between Group A and Group C carries the significance markers. This is the only comparison that reached statistical significance. No statistically significant differences were found between Method A and Method B, nor between Method B and Method C.

Focusing on the significant comparison (C vs. A), SAS provides precise estimates. The mean difference in exam scores between Group C and Group A is **6.375** points. Furthermore, the 95% confidence interval for this mean difference is . Since this interval does not contain zero, it confirms the significant difference detected by the Tukey test, suggesting that Method C leads to significantly higher exam scores than Method A.

Conclusion and Next Steps

The interpretation process for ANOVA results in SAS follows a clear, logical sequence. First, we establish the existence of a statistical difference using the global F-test and its associated p-value. Upon finding significance ($p = 0.0140$), we reject the null hypothesis.

Second, we use the post-hoc analysis, in this case, the Tukey test, to perform granular comparisons between specific pairs of groups. Our analysis revealed that Method C resulted in significantly higher exam scores than Method A (mean difference = 6.375), while all other pairwise comparisons were non-significant.

Researchers should always complement these statistical findings by calculating and reporting descriptive statistics (such as group means and standard deviations) and generating visualizations (like box plots) to fully contextualize the differences observed. For those working with more complex designs, such as factorial or repeated measures models, additional tutorials offer guidance on adapting these interpretation principles.

For further reading and advanced applications of multiple comparison tests in ANOVA:

[A Guide to Using Post Hoc Tests with ANOVA](#)