

How to Perform a Two Sample t-Test in SPSS and Interpret the Results

Authored by
stats writer

March 14, 2026

RECOMMENDED CITATION

stats writer (2026). *How to Perform a Two Sample t-Test in SPSS and Interpret the Results*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=135736>

Theoretical Foundations of the Two-Sample T-Test

The **Two-Sample t-test**, frequently referred to as the **Independent Samples T-Test**, is a foundational inferential statistical procedure used to determine if there is a statistically significant difference between the **means** of two unrelated groups. In the realm of quantitative research, this test serves as a critical tool for investigators who seek to compare distinct populations based on a specific continuous variable. By analyzing the variance within each group and the difference between their averages, the test provides a mathematical basis for deciding whether observed differences are likely due to chance or represent a genuine phenomenon in the population.

To execute this analysis effectively, researchers often utilize specialized software such as **SPSS** (Statistical Package for the Social Sciences). This platform streamlines the complex calculations involving **standard deviations**, **standard errors**, and the resulting **t-statistic**. The primary objective is to evaluate the **null hypothesis**, which typically posits that no difference exists between the group means. If the resulting **p-value** falls below a predetermined threshold, the researcher can confidently reject this null hypothesis in favor of the **alternative hypothesis**.

Understanding the underlying mechanics of this test is essential for accurate data interpretation. The test assumes that the data follows a normal distribution and that the observations are independent of one another. Furthermore, it often considers the **homogeneity of variance**, which implies that the spread of data points is relatively equal across both groups. When these conditions are met, the **Two-Sample t-test** becomes a robust instrument for drawing evidence-based conclusions in fields ranging from psychology and medicine to engineering and business analytics.

Determining Suitability for the Two-Sample T-Test

Before initiating any analysis in **SPSS**, it is imperative to ensure that the **Two-Sample t-test** is the appropriate choice for your specific research design. This test is designed exclusively for comparing two independent groups, such as a treatment group versus a control group, or males versus females. If the study involves comparing the same group at two different time points, a paired-samples t-test would be required instead. The independent nature of the samples is a non-negotiable requirement for the validity of the results generated by this procedure.

The variable being measured, known as the dependent variable, must be continuous--measured on an interval or ratio scale. Examples include weight, height, temperature, or, in the case of our upcoming example, miles per gallon (mpg). The independent variable must be categorical and consist of exactly two levels or categories. This clear distinction allows **SPSS** to categorize the data points correctly and perform the necessary comparative arithmetic to determine **statistical significance**.

Beyond the structure of the variables, researchers must also consider the sample size and

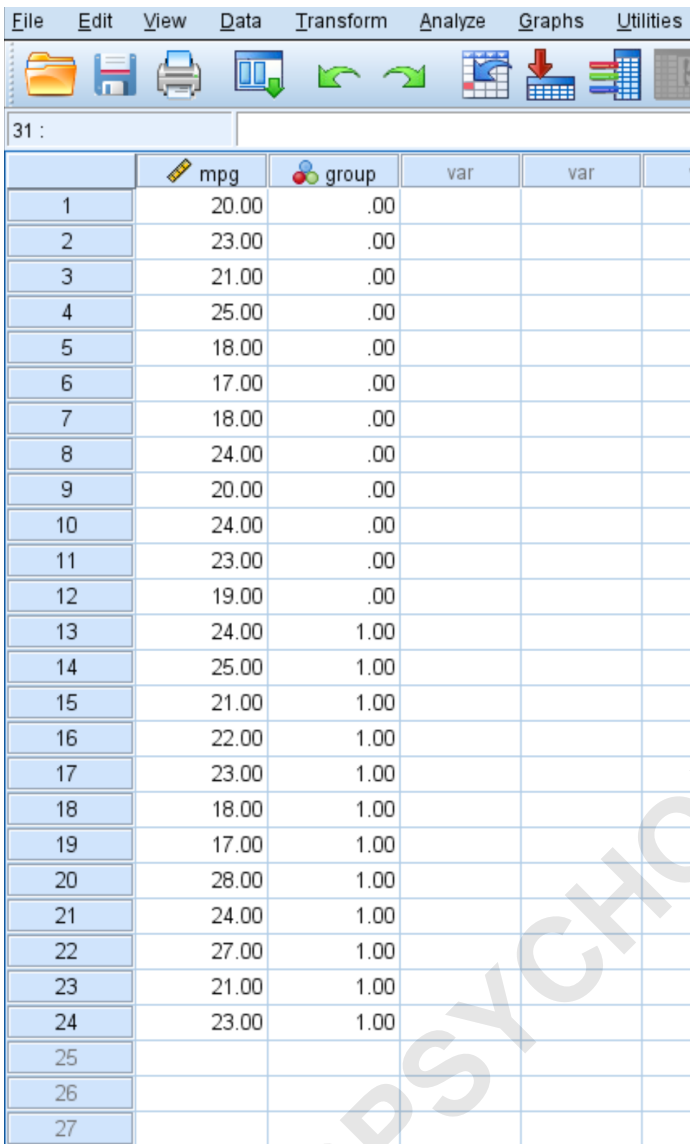
distribution. While the t-test is relatively robust to violations of normality in larger samples due to the Central Limit Theorem, smaller samples require more stringent adherence to distributional assumptions. Assessing these factors beforehand prevents the misinterpretation of data and ensures that the conclusions drawn from the **p-value** and **confidence intervals** are scientifically sound and reproducible.

Contextualizing the Research Case Study

To illustrate the practical application of the **Two-Sample t-test**, let us consider a practical scenario involving automotive engineering and fuel efficiency. Researchers are interested in determining whether a newly developed fuel treatment significantly impacts the average fuel economy of a specific vehicle model. This investigation is crucial for verifying manufacturer claims and providing consumers with accurate data regarding the efficacy of the additive. To maintain scientific rigor, the experiment must be controlled and the data must be analyzed using objective statistical methods.

In this experiment, the researchers utilize a sample of 24 identical cars. These vehicles are divided into two distinct groups: 12 cars receive the new fuel treatment (the experimental group), while the remaining 12 cars do not receive any treatment (the control group). By maintaining a consistent environment and driving conditions for all vehicles, the researchers aim to isolate the effect of the fuel treatment on the dependent variable, which is the **mean** miles per gallon (mpg) achieved by each group.

The following dataset highlights the collected mpg values alongside a categorical grouping variable. In **SPSS**, categorical data is often represented numerically for ease of processing. In this instance, "0" represents the group without fuel treatment, and "1" represents the group that received the treatment. This structured approach allows for a seamless transition into the software's analytical environment.



The screenshot shows the SPSS Data Editor window with a dataset containing 27 rows and 5 columns. The columns are labeled 'mpg', 'group', 'var', 'var', and 'v'. The 'mpg' column contains values ranging from 17.00 to 28.00. The 'group' column contains values of .00 and 1.00. The 'var' and 'v' columns are empty.

	mpg	group	var	var	v
1	20.00	.00			
2	23.00	.00			
3	21.00	.00			
4	25.00	.00			
5	18.00	.00			
6	17.00	.00			
7	18.00	.00			
8	24.00	.00			
9	20.00	.00			
10	24.00	.00			
11	23.00	.00			
12	19.00	.00			
13	24.00	1.00			
14	25.00	1.00			
15	21.00	1.00			
16	22.00	1.00			
17	23.00	1.00			
18	18.00	1.00			
19	17.00	1.00			
20	28.00	1.00			
21	24.00	1.00			
22	27.00	1.00			
23	21.00	1.00			
24	23.00	1.00			
25					
26					
27					

Formulating Statistical Hypotheses for Fuel Efficiency

Before running the analysis in **SPSS**, it is vital to formally state the **null hypothesis** (H_0) and the **alternative hypothesis** (H_1). These statements guide the interpretation of the results and define the parameters of the statistical test. For our fuel treatment study, the hypotheses are defined as follows:

H_0 : $\mu_1 = \mu_2$ -- This represents the **null hypothesis**, asserting that the population **mean** mpg for cars with the treatment is equal to the population mean mpg for cars without the treatment. Essentially, it suggests the treatment has no effect.

H_1 : $\mu_1 \neq \mu_2$ -- This is the **alternative hypothesis**, which claims that the population means are not equal. This suggests that the fuel treatment does indeed cause a change in fuel efficiency, though it does not specify the direction of that change.

In addition to the hypotheses, we must establish a significance level, denoted as α (alpha). The most common threshold in academic research is $\alpha = 0.05$. This value represents the probability of rejecting the **null hypothesis** when it is actually true (a Type I error). By setting this level, we establish a rigorous standard for determining whether the difference in mpg between the two groups is **statistically significant** or merely a result of sampling variability.

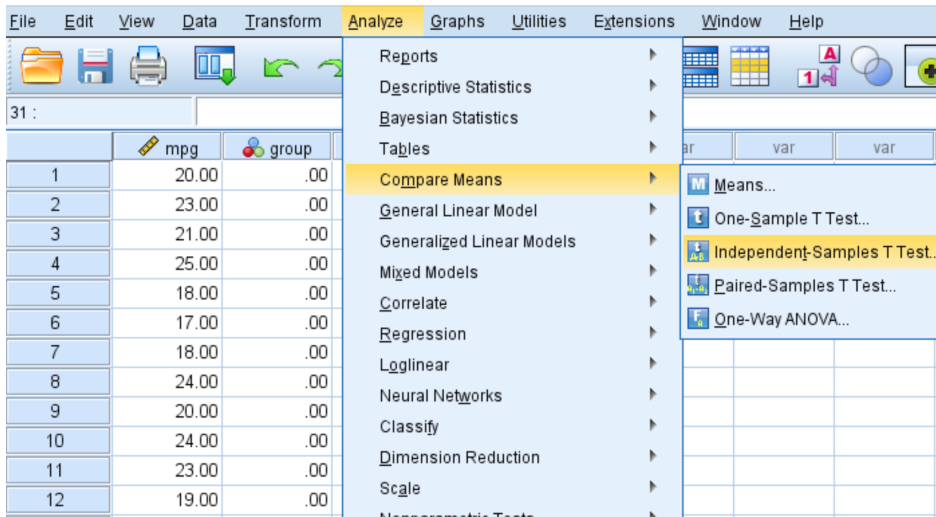
By defining these parameters upfront, the researcher creates a clear framework for decision-making. Once the **p-value** is generated by **SPSS**, it will be compared directly against the $\alpha = 0.05$ threshold. If the p-value is less than or equal to 0.05, the results are deemed significant, leading to the rejection of the null hypothesis. If the p-value is greater than 0.05, we fail to reject the null hypothesis, concluding that there is insufficient evidence to support a difference between the groups.

Execution of the Independent Samples Procedure in SPSS

To begin the actual analysis within the **SPSS** environment, you must first ensure your data is correctly formatted in the Data View tab, with one column for the dependent variable (mpg) and another for the grouping variable (treatment). Once the data is ready, navigate to the top menu bar. This interface provides access to a wide array of statistical tools, but for this specific test, we focus on the comparison of group averages.

Step 1: Choose the Independent Samples T Test option.

Navigate through the menu by clicking the **Analyze** tab. From the resulting dropdown menu, hover over **Compare Means** to reveal further options, and then select **Independent-Samples T Test**. This specific path is designed for comparing the means of two independent populations, making it the correct choice for our fuel efficiency study. The following visual guide demonstrates the exact location of these commands within the **SPSS** interface:



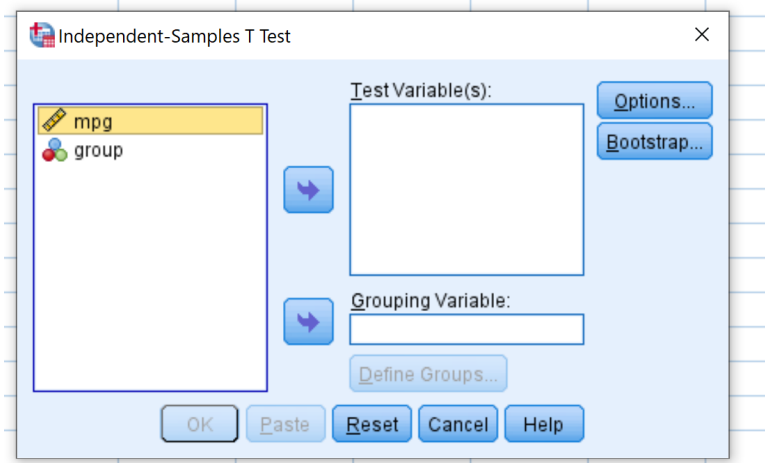
Selecting this option will open a new dialogue box specifically tailored for the **Two-Sample t-test**. It is within this window that you will define which variables **SPSS** should use for the calculation and how the software should distinguish between the two groups in your dataset. Accuracy at this stage is paramount, as incorrect variable placement will result in errors or misleading output.

Configuring Test Variables and Grouping Parameters

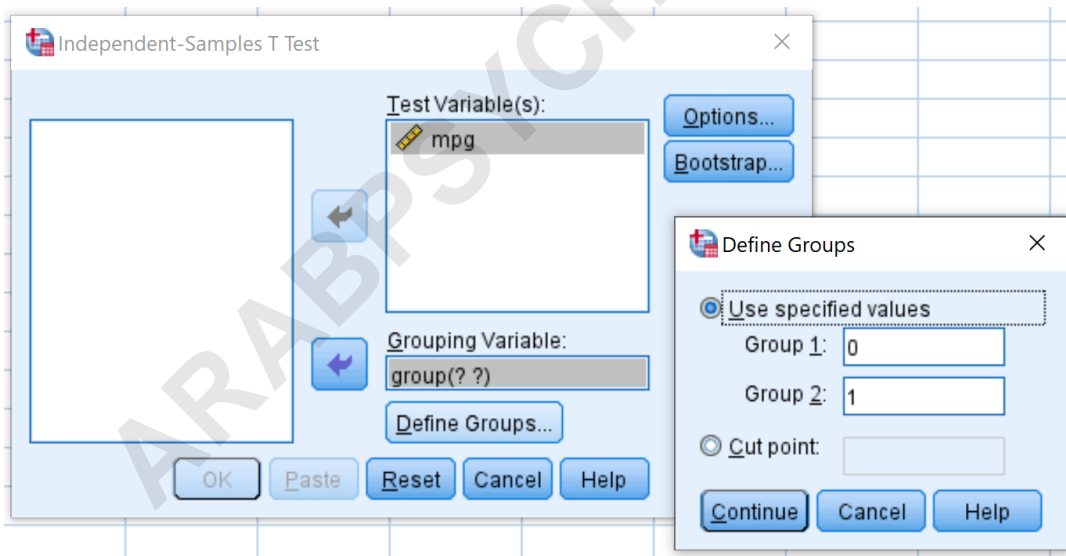
Once the **Independent-Samples T Test** dialogue box appears, you are required to assign your variables to their respective roles. This step informs **SPSS** which metric is being compared and how the participants are categorized. The user-friendly interface allows for a simple drag-and-drop or selection process to move variables from the list on the left into the appropriate functional boxes on the right.

Step 2: Fill in the necessary values to perform the two sample t-test.

As illustrated in the screenshot below, you should drag the **mpg** variable into the area labeled **Test Variable(s)**. This identifies fuel efficiency as the continuous dependent variable for the analysis. Next, drag the **group** variable into the **Grouping Variable** box. This informs the software that the "group" column contains the information necessary to split the cars into two distinct sets for comparison.



However, simply adding the grouping variable is not enough; you must also explicitly **Define Groups**. Click the button labeled **Define Groups** and enter the numerical values that represent your categories. In our case, define Group 1 with the value "0" and Group 2 with the value "1". This ensures **SPSS** knows exactly which rows of data belong to the control group and which belong to the treatment group. After these parameters are set, click **Continue** and then **OK** to run the procedure.



Analyzing Descriptors in the Group Statistics Output

Upon clicking **OK**, **SPSS** will generate an Output Viewer window containing two primary tables. The first table, titled **Group Statistics**, provides a high-level descriptive overview of the data for both the control and treatment groups. Examining this table is a critical first step in interpretation, as it allows the researcher to see the raw differences in **means** and the spread of the data before

diving into the inferential results.

T-Test

Group Statistics					
	group	N	Mean	Std. Deviation	Std. Error Mean
mpg	.00	12	21.0000	2.73030	.78817
	1.00	12	22.7500	3.25087	.93845

Independent Samples Test										
		Levene's Test for Equality of Variances			t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
mpg	Equal variances assumed	.034	.855	-1.428	22	.167	-1.75000	1.22552	-4.29157	.79157
	Equal variances not assumed			-1.428	21.362	.168	-1.75000	1.22552	-4.29597	.79597

The **Group Statistics** table includes several vital metrics for each group:

N: This represents the sample size for each group. In our example, both groups have an N of 12, indicating a balanced design.

Mean: This is the average mpg for each group. By looking at these values, you can see the absolute difference between the control and treatment groups.

Std. Deviation: The **standard deviation** indicates how much the mpg values vary or deviate from the mean within each group.

Std. Error Mean: This value represents the **standard error** of the mean, calculated as the standard deviation divided by the square root of the sample size (s/\sqrt{n}). It provides an estimate of how much the sample mean might differ from the true population mean.

While the descriptive statistics might show a slight difference in the averages--for instance, one group might have a mean of 20 mpg while the other has 22 mpg--they do not tell us if this difference is **statistically significant**. To determine whether the gap between the two **means** is meaningful or just the result of random variation in this specific sample of 24 cars, we must proceed to the second table in the output.

Navigating the Complexities of the T-Test Results Table

The second table, labeled **Independent Samples Test**, contains the core inferential statistics required to test our hypotheses. This table is divided into two main sections: **Levene's Test for Equality of Variances** and the **t-test for Equality of Means**. It provides two rows of results: "Equal variances assumed" and "Equal variances not assumed." The choice between these two rows depends on the outcome of the initial Levene's test.

In most standard analyses, if the **variances** are similar, we use the first row. The key components of this row include:

t: The calculated test statistic. A higher absolute **t-value** indicates a larger difference between the two groups relative to the variability in the data. In this study, the value is -1.428.

df: The **degrees of freedom**, which for an independent samples t-test is typically calculated as $(n_1 + n_2 - 2)$. Here, $12 + 12 - 2$ equals 22.

Sig. (2-tailed): This is the **p-value**. It tells us the probability of observing a t-statistic as extreme as ours if the **null hypothesis** were true. Our result is .167.

Mean Difference: The exact numerical difference between the two sample means.

Std. Error Difference: The standard error associated with the difference between the means.

95% C.I. of the Difference: The **confidence interval** provides a range of values within which the true difference between population means is likely to fall. If this range includes zero, it is a strong indicator that the result is not significant.

By carefully analyzing these values, a researcher can determine the strength of the evidence against the **null hypothesis**. The **p-value** is the most scrutinized number in this table, as it directly informs the final conclusion of the study. However, it should always be interpreted alongside the **confidence interval** to understand the precision of the estimate.

Evaluating the Assumption of Homogeneity of Variance

Before finalizing the conclusion, one must look at **Levene's Test for Equality of Variances**, which is located in the first two columns of the results table. This test checks the **homogeneity of variance** assumption. The null hypothesis for Levene's test is that the variances of the two groups are equal. If the "Sig." value for Levene's test is greater than 0.05, the assumption is met, and we read from the "Equal variances assumed" row.

If the "Sig." value for Levene's test is less than or equal to 0.05, it indicates that the **variances** are significantly different. In such a case, the standard t-test might provide inaccurate results. To correct for this, **SPSS** provides the "Equal variances not assumed" row, which utilizes **Welch's t-test**. This version of the test does not require equal variances and adjusts the **degrees of freedom** accordingly to maintain the integrity of the **p-value**.

In our specific fuel treatment example, the two rows yield very similar results, which suggests that the difference in variances is not extreme enough to drastically alter the outcome. However, adhering to the protocol of checking Levene's test first is a hallmark of a rigorous statistical analysis. It ensures that the mathematical foundations of the **Two-Sample t-test** are respected, leading to more reliable and defensible research findings.

Final Synthesis and Implications of Research Results

The final step in the process is to synthesize the data into a clear conclusion based on the **p-value** and the established significance level ($\alpha = 0.05$). In our **SPSS** output, the **Sig. (2-tailed)** value is reported as .167. Because .167 is significantly greater than our threshold of 0.05, we **fail to reject the null hypothesis**. This means that the observed difference in mpg between the cars with the fuel treatment and those without it is not large enough to be considered **statistically significant**.

From a practical standpoint, this suggests that the new fuel treatment did not produce a meaningful change in the average fuel economy of the cars in this study. While there may have been a slight difference in the sample **means**, the statistical evidence indicates that this difference could easily have occurred due to random chance. Consequently, the researchers cannot claim with 95% confidence that the treatment has a real effect on the population of this car model.

When reporting these results in a formal paper or report, it is standard practice to include the t-statistic, the **degrees of freedom**, and the **p-value** in a format such as: $t(22) = -1.428, p = .167$. Providing the **confidence interval** also adds depth to the report, showing the range of the mean difference. By following this structured approach in **SPSS**, you ensure that your research is grounded in mathematical reality and that your conclusions are supported by a transparent and standardized analytical process.