How do I perform a Mann-Whitney U test in SPSS?

Authored by

stats writer

December 26, 2025

Introduction to the Mann-Whitney U Test

The Mann-Whitney U test is a fundamental statistical tool in the field of non-parametric analysis. It is specifically designed for comparing two independent groups when the assumptions required for standard parametric tests, such as the independent samples t-test, are violated. This robust method determines whether two populations are stochastically different, meaning it assesses if values from one population tend to be larger or smaller than values from the other population. It is a critical choice when dealing with ordinal data or when interval data distributions are significantly skewed or contain outliers, which is common in many real-world research settings.

Often referred to interchangeably as the Wilcoxon rank-sum test, the U test operates by converting observed data values into ranks and then comparing the sum of ranks between the two groups. This procedure allows the test to bypass the need for normality assumptions that underpin tests like the independent samples t-test. In the context of the powerful statistical software SPSS (Statistical Package for the Social Sciences), performing this analysis is straightforward, provided the correct steps and dialogue settings are followed precisely.

Why Use the Non-Parametric Approach?

Choosing between parametric and non-parametric tests is one of the most important decisions in statistical analysis. The primary reason for utilizing the Mann-Whitney U test is when the data fail to meet the strict prerequisites of a parametric test. Specifically, if the sample distributions are non-normally distributed, or if the sample sizes are particularly small (often cited as less than 30 observations per group), relying on the U test becomes essential for valid conclusions and reliable statistical inference.

While the independent samples t-test compares the means of two groups, the Mann-Whitney U test compares the medians or, more accurately, the underlying distributions of the two independent groups. Because it relies on ranks rather than the actual raw scores, it is far less sensitive to extreme outliers, offering a more reliable assessment when data characteristics are questionable. It is universally regarded as the nonparametric equivalent to the independent samples t-test, offering robust comparative power without demanding stringent distributional assumptions.

Setting Up the Example Scenario in SPSS

To illustrate the application of this test, consider a common research question in applied statistics. Suppose researchers are interested in evaluating the effectiveness of a new fuel treatment designed to improve vehicle efficiency. They hypothesize that this treatment will lead to a significant change in the average miles per gallon (mpg) achieved by a car, comparing treated vehicles against a control group.

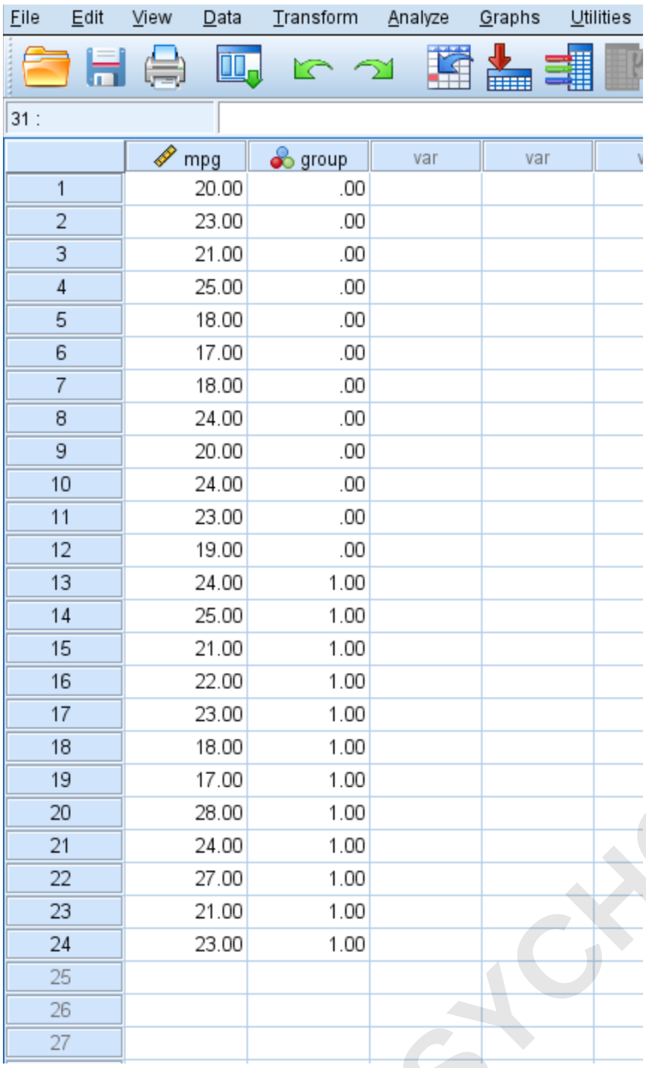
To investigate this, they set up a controlled experiment. They collect data from two independent groups of cars: one group receives the experimental fuel treatment (Group 1), and the other group serves as the control and does not receive the treatment (Group 0). Due to practical constraints, they utilize a relatively small sample size, measuring the mpg for 12 cars with the treatment and 12 cars without it, resulting in a total sample size of 24.

Given the small sample size (n=12 per group) and an initial suspicion that the resulting mpg data might not follow a perfect normal distribution, the researchers decide that the Mann-Whitney U test is the most appropriate statistical method to determine if there is a statistically significant difference in mpg between the two defined groups.

Data Entry and Verification

Before conducting the test, the data must be accurately entered into the SPSS Data View and Variable View editors. In this example, two key variables are required: the quantitative dependent variable, named 'mpg', representing the miles per gallon achieved; and the independent grouping variable, named 'group', which is coded numerically. Typically, the grouping variable is defined using binary values: 0 representing the 'no fuel treatment' group (Control) and 1 representing the 'fuel treatment' group (Experimental).

The following visual representation confirms the structure of the data setup in SPSS, showing the achieved mpg for each car along with its corresponding group assignment. This preparation ensures that SPSS can correctly distinguish between the two independent samples we intend to compare during the analysis phase.



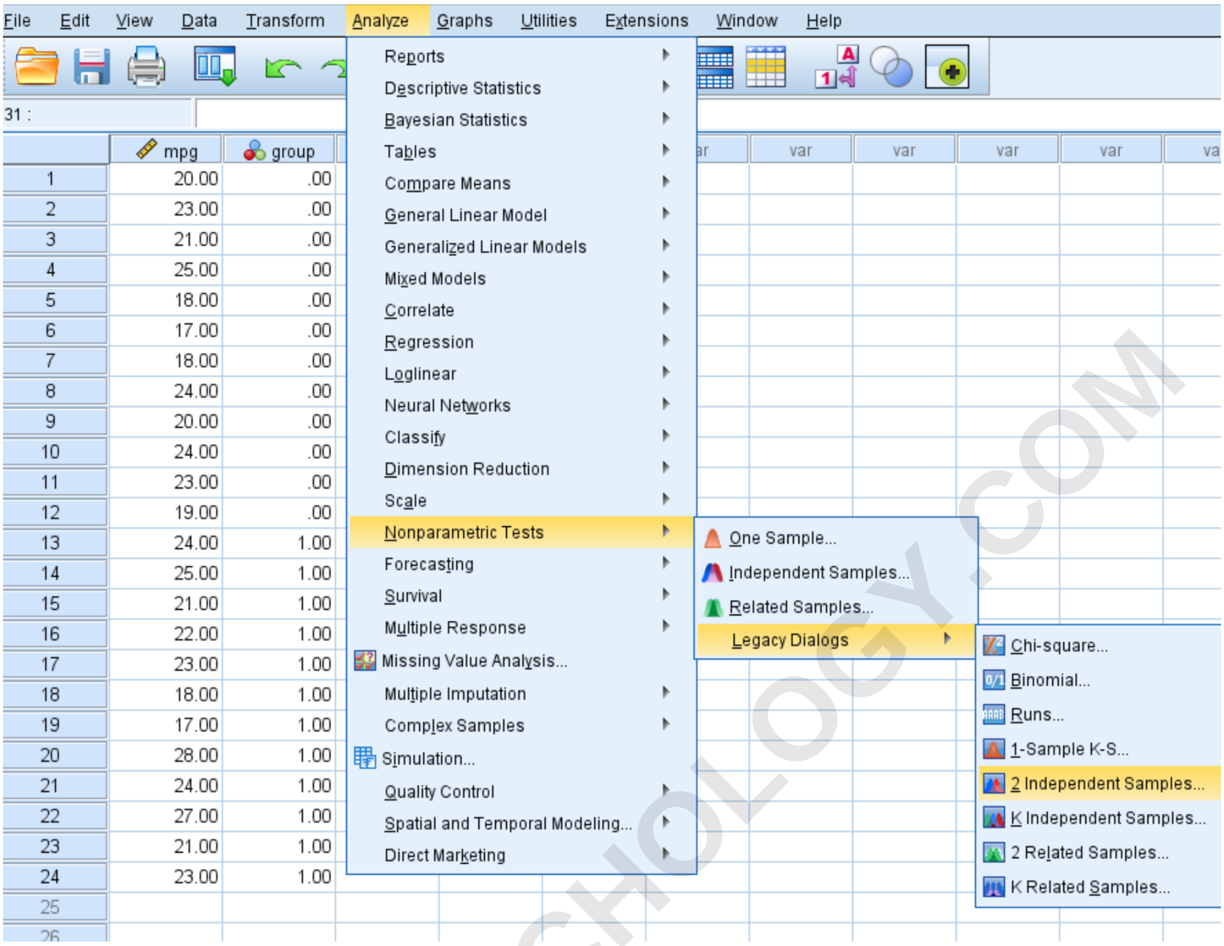
	mpg	group	var	var
1	20.00	.00		
2	23.00	.00		
3	21.00	.00		
4	25.00	.00		
5	18.00	.00		
6	17.00	.00		
7	18.00	.00		
8	24.00	.00		
9	20.00	.00		
10	24.00	.00		
11	23.00	.00		
12	19.00	.00		
13	24.00	1.00		
14	25.00	1.00		
15	21.00	1.00		
16	22.00	1.00		
17	23.00	1.00		
18	18.00	1.00		
19	17.00	1.00		
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				

Step-by-Step Guide: Conducting the Test (Legacy Dialogs)

The procedure for executing the Mann-Whitney U test in SPSS is straightforward and typically involves using the Legacy Dialogs pathway, which provides a reliable interface for this specific analysis. We begin by accessing the analysis menu structure from the top ribbon:

Step 1: Select the Mann-Whitney U Test Option.

Click the **Analyze** tab on the menu bar. Then, hover over **Nonparametric Tests**. Within the subsequent sub-menu that appears, select **Legacy Dialogs**. Finally, choose the option labeled **2 Independent Samples**. This precise sequence opens the dedicated dialogue box necessary for configuring and running the non-parametric comparison test.



Configuring Variables and Group Definitions

Once the Two Independent Samples dialogue box is open, the variables must be correctly assigned to their respective roles in the analysis. This step is crucial, as incorrect assignment will lead to erroneous results. Therefore, careful attention is required during the selection and definition process.

Step 2: Fill in the Necessary Values to Perform the Test.

Move the quantitative dependent variable, **mpg**, into the box labelled **Test Variable List**. This is the variable whose distribution differences we are investigating between the groups.

Next, move the categorical grouping variable, **group**, into the box labelled **Grouping Variable**. This variable instructs SPSS how to separate the observations into the two comparison samples.

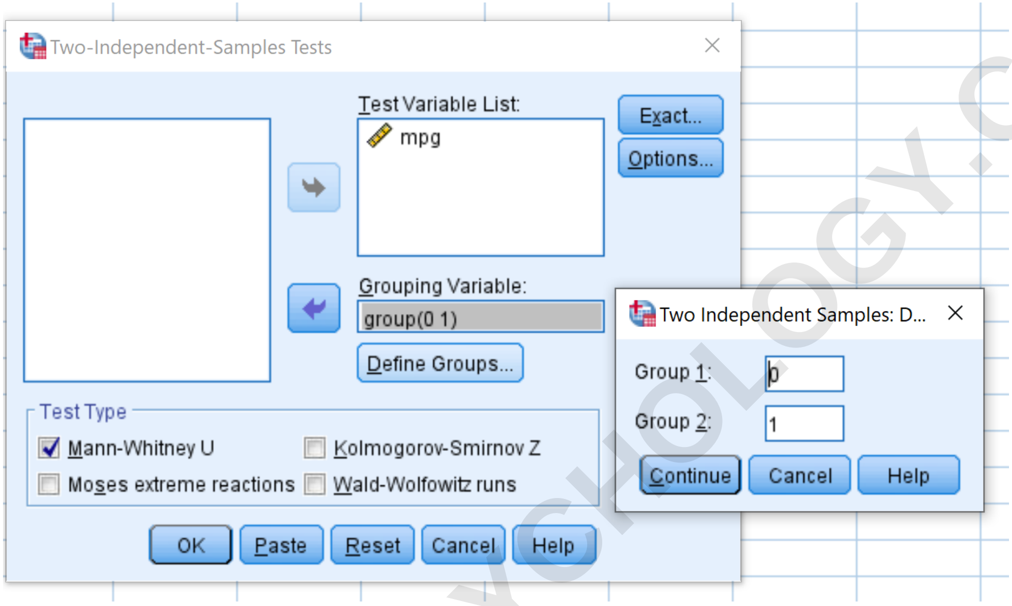
After assigning the grouping variable, click on the **Define Groups** button. In the subsequent definition window, explicitly input the numerical values used for the two groups: define Group 1 as the rows with value 0 (no treatment) and define Group 2 as the rows with value 1 (fuel treatment).

Click **Continue**.

Ensure that the checkbox next to **Mann-Whitney U** is selected under the Test Type section, confirming that we are running the desired non-parametric test.

Finally, click **OK** to execute the analysis and generate the output tables in the viewer.

The dialogue box, once configured with the Test Variable and the defined Grouping Variable, should visually match the screenshot provided below, confirming the settings prior to execution.



Interpreting the Mann-Whitney U Test Output

Upon clicking **OK**, SPSS generates the output viewer displaying the results of the Mann-Whitney U test. These results typically include both a Rank Table (summarizing the mean ranks for comparison) and a Test Statistics Table. The Test Statistics Table is the most important for drawing statistical inferences, as it provides the calculated U statistic, the Z test statistic, and the significance level (Asymp. Sig. 2-tailed).

Step 3: Interpret the Results.

The generated output image below isolates the key statistical findings needed to assess the null hypothesis regarding the fuel treatment effect.

NPar Tests

Mann-Whitney Test

Ranks

	group	N	Mean Rank	Sum of Ranks
mpg	.00	12	10.67	128.00
	1.00	12	14.33	172.00
	Total	24		

Test Statistics[a]

	mpg
Mann-Whitney U	50.000
Wilcoxon W	128.000
Z	-1.279
Asymp. Sig. (2-tailed)	.201
Exact Sig. [2*(1-tailed Sig.)]	.219[b]

a. Grouping Variable: group

b. Not corrected for ties.

Focusing specifically on the values in the Test Statistics table, the critical figures required for our conclusion are:

Z Test Statistic: -1.279

Asymptotic Significance (2-tailed p-value): .201

Drawing Conclusions Based on the P-Value

The core of statistical decision-making rests on comparing the calculated p-value against a predetermined level of significance, typically denoted as alpha ($alpha$). In most behavioral and applied science research, the conventional alpha level is set at 0.05. This threshold represents the maximum acceptable probability of rejecting the null hypothesis when it is actually true (a Type I error).

Our null hypothesis (H_0) states that there is no difference in the distribution of mpg between the two fuel treatment groups. The alternative hypothesis (H_A) states that a difference does exist. Since the calculated two-tailed p-value (.201) is substantially greater than the established significance level of 0.05, we must consequently fail to reject the null hypothesis.

This statistical conclusion implies that, based on the collected sample data, there is insufficient statistical evidence to assert that the true median miles per gallon achieved is significantly different between the cars that received the new fuel treatment and those that did not. In practical terms, the fuel treatment did not produce a statistically distinguishable effect on vehicle efficiency in this specific experimental sample.

Reporting the Findings in APA Format

The final and crucial step in any statistical analysis is reporting the results clearly, accurately, and in a standardized format, such as the guidelines set forth by the American Psychological Association (APA). This ensures that the findings are transparent, easily understood, and reproducible by other researchers.

Step 4: Report the Results.

When reporting the outcomes of the Mann-Whitney U test, it is necessary to include the sample sizes for each group, the calculated Z statistic, and the exact p-value. Below is an example of how this specific study's findings should be formally documented in a research paper or report:

A Mann-Whitney U test was conducted on 24 cars to determine if a new fuel treatment led to a significant difference in miles per gallon (mpg). The sample consisted of two independent groups: a treatment group (n=12) and a control group (n=12).

Results indicated that the distribution of mpg was not statistically significantly different between the treatment and control groups ($Z = -1.279$, $p = .201$). This result suggests that the observed differences were likely due to chance rather than the effect of the treatment, as the p-value exceeded the conventional significance threshold of $alpha = 0.05$.

Based on these findings, we conclude that the new fuel treatment, as evaluated in this study, does not have a significant measurable impact on the miles per gallon achieved by the cars.