How to Perform a Mann-Whitney U Test in SPSS: A Step-by-Step Guide

Authored by

stats writer

March 14, 2026

RECOMMENDED CITATION

stats writer (2026). *How to Perform a Mann-Whitney U Test in SPSS: A Step-by-Step Guide*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=135745

Understanding Non-parametric Statistics and the Mann-Whitney U Test

In the realm of quantitative research, **non-parametric statistics** serve as a vital toolset for researchers who encounter data that does not conform to the stringent requirements of parametric tests. While parametric tests like the t-test assume that the underlying data follows a specific distribution, non-parametric methods are essentially distribution-free. The **Mann-Whitney U Test**, which is also frequently referred to as the **Wilcoxon rank-sum test**, is perhaps the most widely utilized non-parametric alternative for comparing two independent groups. This test is particularly advantageous when the **normal distribution** assumption is violated or when the sample sizes are too small to provide sufficient **statistical power** for parametric procedures.

The core logic behind the **Mann-Whitney U Test** involves ranking all observations from both groups combined, rather than comparing the raw means of the groups. By converting continuous or ordinal data into ranks, the test mitigates the influence of **outliers** and extreme skewness, which could otherwise distort the results of a standard t-test. Within the environment of **SPSS**, this procedure is streamlined, allowing users to efficiently determine whether there is a statistically significant difference between the distributions of two independent samples. Whether you are working in social sciences, clinical research, or business analytics, mastering this test ensures that your **hypothesis testing** remains robust even under imperfect data conditions.

Employing the **Mann-Whitney U Test** in **SPSS** requires a clear understanding of your data structure and the specific research question at hand. This test is designed specifically for **independent samples**, meaning that the subjects in one group must have no relationship to the subjects in the other group. If the data were paired or repeated measures, a different non-parametric test, such as the Wilcoxon signed-rank test, would be appropriate instead. By following a structured approach in **IBM SPSS Statistics**, researchers can generate comprehensive outputs that include rank sums, U statistics, and the critical **p-value** necessary for valid scientific inference.

Theoretical Foundations and Prerequisites for the Test

Before executing the analysis in **SPSS**, it is imperative to verify that your data meets the fundamental assumptions of the **Mann-Whitney U Test**. Although it is a non-parametric test, it still requires that the dependent variable be measured at least at the **ordinal** or continuous level. Furthermore, the independent variable must consist of two categorical, independent groups. One of the most common reasons researchers choose this test is the presence of a non-normal distribution, which can be identified through visual inspections like histograms or formal tests like the **Shapiro-Wilk test**.

Another nuance of the **Mann-Whitney U Test** is the interpretation of what exactly is being compared. If the distributions of the two groups have a similar shape, the test can be used to compare the **medians** of the groups. However, if the shapes of the distributions differ significantly, the test technically compares the mean ranks rather than the medians. This distinction is crucial for accurate **data analysis** and reporting, as it affects how you describe your findings in the final manuscript. Researchers must be diligent in examining the spread of their data to ensure they are making the correct claims about their populations.

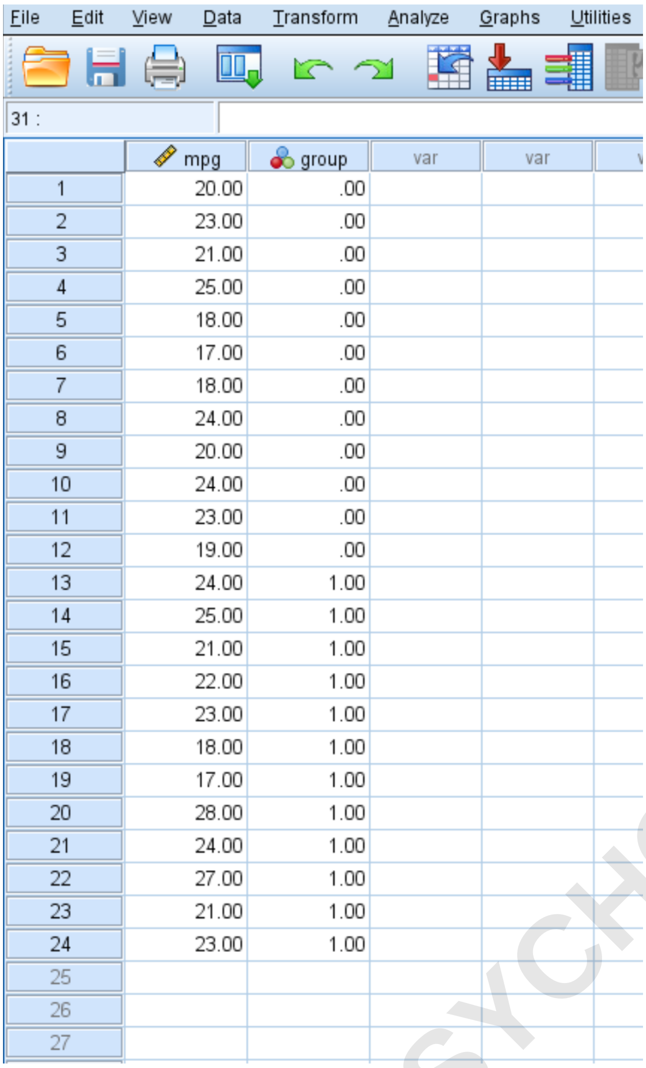
The calculation of the U statistic itself is based on the number of times a value from one group precedes a value from the other group when all data points are ranked in ascending order. If the two groups are similar, the U statistic will be relatively large, and the resulting **p-value** will exceed the standard **significance level** of 0.05. Conversely, a small U statistic indicates that the values of one group are consistently higher or lower than the other, suggesting a significant difference. **SPSS** automates these complex ranking calculations, providing both the U statistic and a standardized **Z-score** for larger sample sizes.

Practical Application: The Fuel Treatment Research Scenario

To illustrate the practical application of this test, consider a study where researchers are investigating the efficacy of a new fuel treatment. The primary objective is to determine if the treatment leads to a change in the average miles per gallon (mpg) of a vehicle. In this experiment, 24 cars are utilized, with 12 cars receiving the treatment and 12 cars serving as the **control group**. This small sample size (n=12 per group) is a classic indicator that a non-parametric approach like the **Mann-Whitney U Test** may be more appropriate than a parametric t-test, particularly if the mpg data is skewed.

The dataset for this analysis must be organized carefully within **SPSS**. Each car represents a single case (row), and there should be two primary columns: one for the dependent variable (mpg) and one for the independent grouping variable (treatment status). In the grouping variable, numeric codes are typically assigned--for instance, 0 represents the control group (no treatment) and 1 represents the experimental group (fuel treatment). This binary coding is essential for **SPSS** to distinguish between the two samples during the calculation process.

The following screenshot displays the raw data as it would appear in the **SPSS Data View**. Each row corresponds to an individual car's performance and its respective group assignment:



	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				

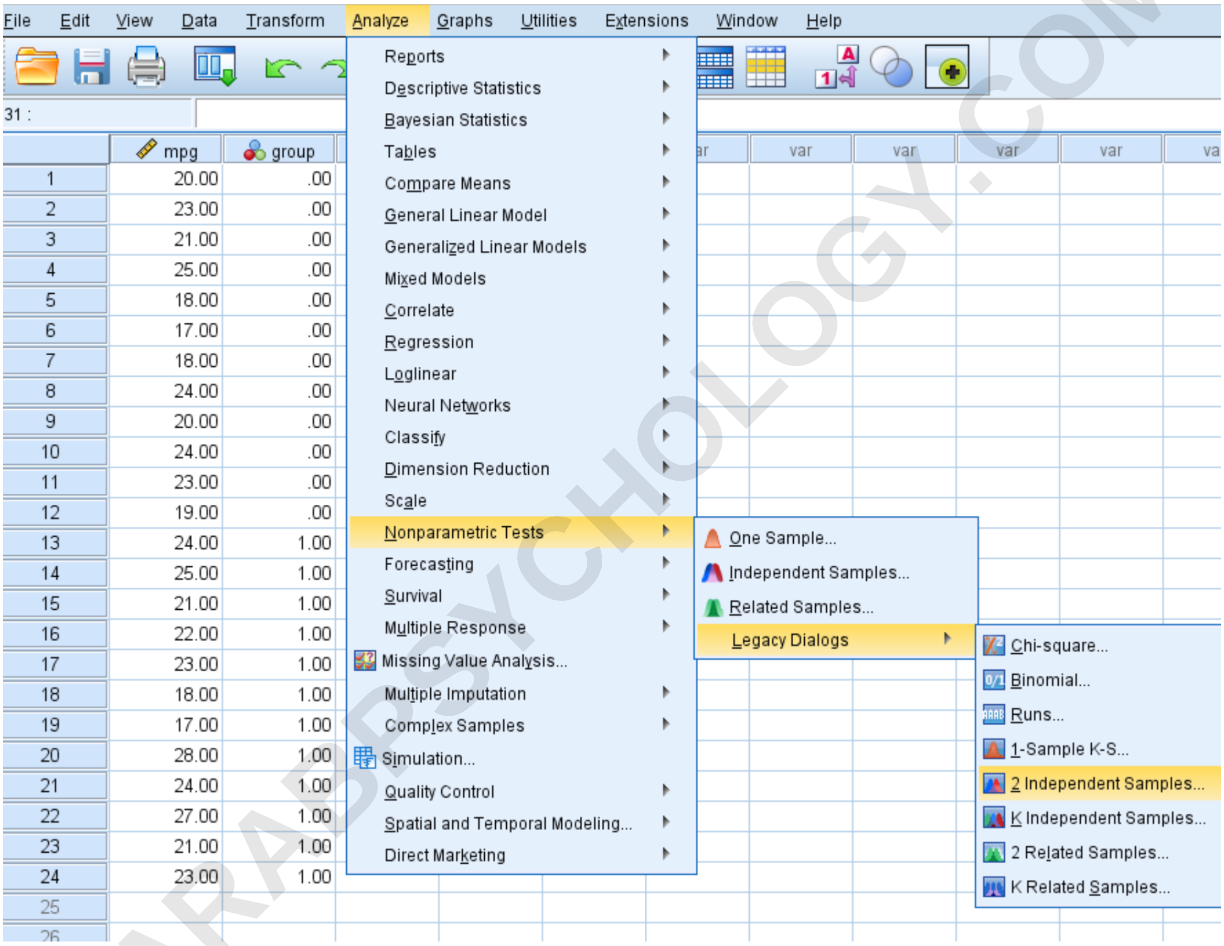
Because the researchers suspect that the underlying distribution of mpg might not be normal--perhaps due to the small sample size or potential outliers in engine performance--they opt for the **Mann-Whitney U Test**. This choice ensures that the resulting **statistical inference** is not compromised by the unmet assumptions of parametric testing. By using **SPSS**, the researchers can conduct this comparison with high precision and reliability.

Step 1: Initiating the Mann-Whitney U Procedure in SPSS

The first step in performing the **Mann-Whitney U Test** is navigating through the **SPSS** menu system. While **IBM SPSS Statistics** offers several ways to access non-parametric tests, the **Legacy Dialogs** path remains a favorite among researchers for its straightforward interface and clear configuration options. To begin, click on the **Analyze** menu located at the top of the screen. From the dropdown menu, hover over **Nonparametric Tests**, then select **Legacy Dialogs**, and finally click on **2 Independent Samples**.

This path is designed for scenarios where you are comparing exactly two groups that are independent of one another. Selecting this option will open a new dialog box where you will define the variables for your analysis. It is important to ensure that your data is correctly formatted before reaching this step; specifically, your grouping variable should be numeric, as **SPSS** requires numeric values to define group boundaries in the **Legacy Dialogs**.

The following image illustrates the exact menu path you should follow within the **SPSS** interface to initiate the test:



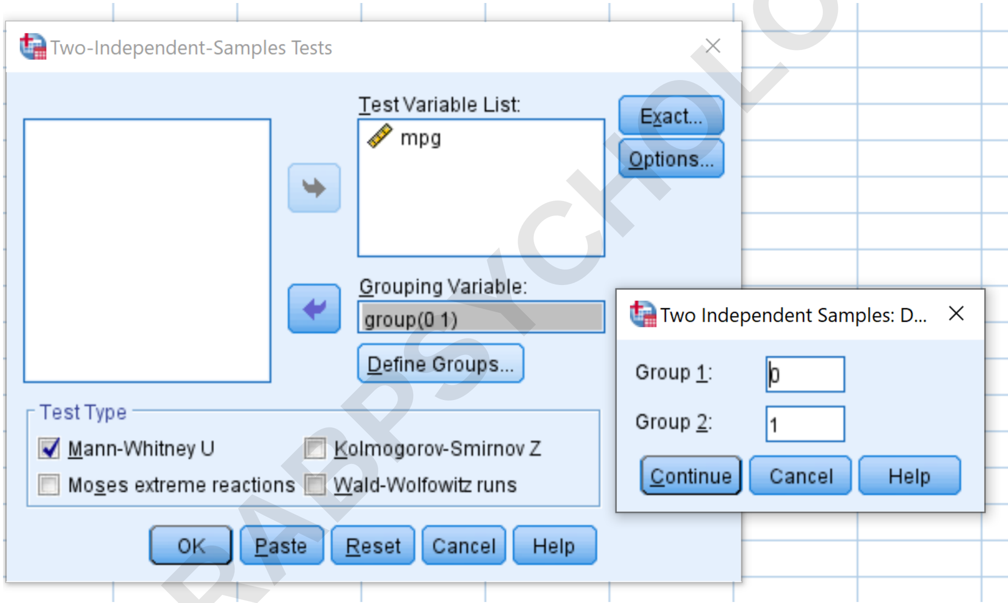
Once you have selected **2 Independent Samples**, the software is ready to receive the specific instructions regarding which variables to test and how the groups are categorized. This stage is critical because misidentifying the test variable or the grouping variable will lead to errors in the execution of the **algorithm**. Always double-check that your active dataset is the one containing the fuel treatment data before proceeding further.

Step 2: Configuring Variables and Defining Groups

After the **Two-Independent-Samples Tests** dialog box appears, you must move your variables into the appropriate fields. Locate the **mpg** variable in the list on the left and move it into the box labeled **Test Variable List**. This represents the continuous dependent variable that you wish to compare. Next, locate the **group** variable and move it into the **Grouping Variable** box. This variable tells **SPSS** which cases belong to which sample.

Once the grouping variable is in place, you will notice that the **Define Groups** button becomes active. You must click this button to specify which numeric codes correspond to your two samples. In our car experiment, you would enter 0 for Group 1 and 1 for Group 2. This step is vital because **SPSS** needs to know exactly which values in the dataset represent the two independent groups being compared. After entering these values, click **Continue** to return to the main dialog box.

Before finalizing the test, ensure that the **Mann-Whitney U** checkbox is selected under the **Test Type** section. This is usually the default setting, but it is always worth verifying to avoid running a different test like the Moses extreme reactions or Kolmogorov-Smirnov test by mistake. The configuration should look like the following screenshot:



With all parameters set, clicking **OK** will trigger the **SPSS** processor to perform the ranking and calculate the U statistic. The software will then generate the output in a new window, providing the statistical evidence needed to address your **null hypothesis**.

Step 3: Interpreting the Statistical Output

Upon clicking **OK**, **SPSS** generates an **Output Viewer** window containing several tables. The most important tables are the **Ranks** table and the **Test Statistics** table. The Ranks table provides the mean rank for each group, which gives you an initial visual indication of which group performed better. For instance, if the fuel treatment group has a significantly higher mean rank, it suggests higher mpg values in that group. However, the **Test Statistics** table is where you find the formal confirmation of **statistical significance**.

In the **Test Statistics** table, you will find the **Mann-Whitney U** value, the **Wilcoxon W**, and the **Z-statistic**. For small samples, the exact significance is often reported, while for larger samples, **SPSS** provides an **Asymptotic Significance (2-tailed)**. This **p-value** is the most critical number for your conclusion. If the p-value is less than or equal to your alpha level (typically 0.05), you reject the **null hypothesis** and conclude that there is a significant difference between the groups.

Referring to our specific fuel treatment example, the results of the **Mann-Whitney U Test** are as follows:

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.

Z test statistic: -1.279
p-value (Asymp. Sig. 2-tailed): .201

In this case, the **p-value** of .201 is substantially higher than the 0.05 threshold. Consequently, we fail to reject the **null hypothesis**. This indicates that the observed differences in mpg between the cars with the fuel treatment and those without it are likely due to random **sampling error** rather than the treatment itself. The evidence is insufficient to claim that the fuel treatment has a significant impact on vehicle efficiency.

Step 4: Academic Reporting and Final Conclusions

The final stage of the process is communicating your findings in a clear and professional manner, often following **APA style** or other academic formatting guidelines. When reporting a **Mann-Whitney U Test**, you should include the name of the test, the sample sizes for each group, the U or Z statistic, and the **p-value**. It is also helpful to mention the medians or mean ranks to provide context for the direction of the difference, even if that difference is not statistically significant.

For the car fuel treatment study, a formal report might be written as follows:

A Mann-Whitney U test was conducted on a sample of 24 vehicles to evaluate whether a new fuel treatment influenced the mean miles per gallon. The study compared two independent groups, each consisting of 12 cars. The results indicated that the fuel treatment did not lead to a statistically significant difference in mpg (z = -1.279, p = .201) at the 0.05 significance level. Consequently, the data does not support the claim that the treatment improves fuel efficiency.

By providing this level of detail, you ensure that your research is transparent and reproducible. The **Mann-Whitney U Test** in **SPSS** serves as a powerful tool in your analytical arsenal, allowing you to handle non-normal data with confidence. Whether your results are significant or not, following these rigorous steps ensures the integrity of your **statistical** conclusions and contributes to the overall quality of your scientific work.

Conclusion and Best Practices for Future Research

The **Mann-Whitney U Test** is an essential component of **non-parametric statistics**, providing a reliable way to compare groups when the assumptions of parametric tests are not met. Throughout this guide, we have explored the theoretical underpinnings of the test, the step-by-step execution within **SPSS**, and the proper interpretation of the resulting data. It is important to remember that while **SPSS** simplifies the computation, the researcher's role in verifying assumptions and understanding the data remains paramount.

For future research, always consider performing a **power analysis** before your study to ensure your sample size is adequate to detect a meaningful effect. Additionally, always report **effect size** measures, such as r (calculated from the Z score), to provide a more complete picture of the practical significance of your findings. Even when a **p-value** is non-significant, the effect size can offer valuable insights into the magnitude of the differences observed.

By adhering to these best practices and utilizing the robust features of **IBM SPSS Statistics**, you can navigate complex data landscapes with ease. The **Mann-Whitney U Test** ensures that your comparative analyses are both scientifically sound and resistant to the distortions of non-normal distributions, making it a cornerstone of modern **quantitative research** methodology.