

# How to Easily Compare Two Groups with the Mann-Whitney U Test in Excel

Authored by  
**stats writer**

December 28, 2025

## RECOMMENDED CITATION

stats writer (2025). *How to Easily Compare Two Groups with the Mann-Whitney U Test in Excel*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=109479>

The Mann-Whitney U Test is an indispensable tool in the field of nonparametric statistics. Also known as the Wilcoxon rank-sum test, this statistical procedure is specifically designed to compare the differences between two independent samples when the underlying assumptions for parametric tests, such as the two-sample t-test, are violated. Since the analysis relies on ranking the data rather than assuming a specific distribution, it provides a robust alternative for analyzing data that is not normally distributed or when working with small sample sizes.

While specialized statistical software is often used for these calculations, it is entirely possible to perform a complete Mann-Whitney U Test using Microsoft Excel. This approach involves calculating the ranks manually, determining the U statistic, and finally computing the necessary z test statistic and corresponding p-value to evaluate the null hypothesis. This guide provides a detailed, step-by-step tutorial on executing this essential nonparametric test within the familiar Excel environment.

## When to Use the Mann-Whitney U Test

The decision to employ the Mann-Whitney U Test hinges on the nature of your data and the fulfillment of parametric assumptions. This test is the appropriate choice when you are comparing two independent groups, but your data fails to meet the stringent requirements of a parametric test, specifically the assumption of normally distributed populations. This frequently occurs when dealing with ordinal data, heavily skewed interval data, or when the number of observations ( $n$ ) in the samples is small, typically less than 30 per group. The Mann-Whitney U Test serves as the most effective nonparametric statistical test equivalent to the independent samples t-test.

It is crucial to understand that the Mann-Whitney U Test does not compare the means of the two groups, but rather compares their distributions based on the ranks of the observations. Essentially, it tests whether a randomly selected observation from one population is likely to be greater or smaller than a randomly selected observation from the second population. Therefore, when the distribution shapes are similar, the test determines if the population medians are different. If the distributions are dissimilar, the test evaluates differences in overall population ranking, providing a highly generalized comparison between the two groups.

## Practical Example: Fuel Treatment Analysis

To illustrate the application of this test, consider a common research scenario in experimental design. Researchers are investigating the efficacy of a new fuel treatment and want to determine if its application leads to a statistically significant change in the average miles per gallon (mpg) achieved by automobiles. They set up a controlled experiment where they test two independent groups: the Treated group and the Untreated (Control) group.

In this specific study, the researchers collected data from 12 cars that received the fuel treatment

and 12 cars that did not. Given these relatively small sample sizes ( $n=12$  for each group) and the suspicion that the mpg data might not follow a standard Gaussian distribution, the team correctly chose the Mann-Whitney U Test as the robust method to compare the two independent samples. The goal is to determine if the location of the distributions (i.e., the median mpg) differs significantly between the cars that received the treatment and those that did not.

The null hypothesis ( $H_0$ ) posited for this experiment states that there is no difference in the distributions of mpg between the treated and untreated groups. Our task is to use Excel to calculate the test statistic necessary to evaluate this hypothesis against a significance level of 0.05.

## Step 1: Organizing the Data in Excel

The initial step in conducting the analysis involves structuring your dataset correctly within the Excel worksheet. For the Mann-Whitney U Test, the raw data for the two independent groups must be entered into two separate, adjacent columns. This segregation allows for easy manipulation and calculation of ranks in subsequent steps. For our fuel treatment example, we will label one column "Untreated" and the second column "Treated."

Ensure that all 12 observations for the Treated group and all 12 observations for the Untreated group are recorded accurately in their respective columns, maintaining clarity and integrity of the data input. This fundamental organization sets the stage for all complex calculations that follow, particularly the crucial ranking process.

	A	B	C	D	E
1	<b>Treated</b>	<b>Untreated</b>			
2	24	20			
3	25	23			
4	21	21			
5	22	25			
6	23	18			
7	18	17			
8	17	18			
9	28	24			
10	24	20			
11	27	24			
12	21	23			
13	23	19			
14					
15					
16					
17					
18					

## Step 2: Calculating the Ranks for Both Samples

The core of the Mann-Whitney U Test is the assignment of ranks. Since this is a nonparametric statistical test, we must combine all 24 observations from both groups into a single unified list and then assign numerical ranks. The smallest value receives the rank of 1, and the largest value receives the rank of N (24 in this case). When ties occur (identical values in the combined data), the average of the ranks they would have received is assigned to each tied observation.

This crucial ranking process is efficiently handled in Excel using the RANK.AVG function. You must create two new columns, perhaps labeled "Rank (Untreated)" and "Rank (Treated)." The formula must reference the specific value being ranked against the entire combined range of data. The image below demonstrates the structure of the formula used to calculate the rank of the first value in the Treated group:

	A	B	C	D	E	F	G	H	I	J	K	L
1	<b>Treated</b>	<b>Untreated</b>		<b>Treated (ranks)</b>	<b>Untreated (ranks)</b>							
2	24	20		18.5								
3	25	23										
4	21	21										
5	22	25										
6	23	18										
7	18	17										
8	17	18										
9	28	24										
10	24	20										
11	27	24										
12	21	23										
13	23	19										
14												
15												
16												
17												
18												

Although the syntax for the ranking formula appears complex initially--especially ensuring the absolute references (\$B\$2:\$C\$13) are correct to fix the combined data range--this step only needs to be performed once per column. Once the initial rank is calculated, you can simply drag the formula down to automatically populate the ranks for all other observations in both the Treated and Untreated columns. This results in two new columns containing the individual ranks for each data point within its respective group, derived from the unified dataset. Note that the sum of these ranks across all 24 observations should equal  $N(N+1)/2$ , confirming the ranking accuracy.

	A	B	C	D	E	F
1	<b>Treated</b>	<b>Untreated</b>		<b>Treated (ranks)</b>	<b>Untreated (ranks)</b>	
2	24	20		18.5	7.5	
3	25	23		21.5	14.5	
4	21	21		10	10	
5	22	25		12	21.5	
6	23	18		14.5	4	
7	18	17		4	1.5	
8	17	18		1.5	4	
9	28	24		24	18.5	
10	24	20		18.5	7.5	
11	27	24		23	18.5	
12	21	23		10	14.5	
13	23	19		14.5	6	
14						
15						
16						
17						
18						

### Step 3: Determining Key Values for the U Statistic

With the ranks established, the next crucial phase is calculating the necessary components required to derive the U test statistic. This involves computing the sum of the ranks for each group (R1 and R2), identifying the sample sizes (n1 and n2), and subsequently calculating the two potential U statistics (\$U\_1\$ and \$U\_2\$). The final Mann-Whitney U Test statistic is defined as the minimum of \$U\_1\$ and \$U\_2\$.

The required calculations are performed using the following functions and formulas in designated cells:

**Sum of Ranks (R):** Use the `SUM` function (e.g., `=SUM(D2:D13)`) on the ranks calculated in Step 2 for each group.

**Sample Size (n):** Use the `COUNT` function to verify the number of observations in each sample (\$n\_1 = 12, n\_2 = 12\$).

**U Statistics:** The U statistics are calculated based on the rank sums (\$R\_1\$ and \$R\_2\$) and sample sizes (\$n\_1\$ and \$n\_2\$). The formulas are:

$$U_1 = n_1 n_2 + \frac{n_1 (n_1 + 1)}{2} - R_1$$

$$U_2 = n_1 n_2 + \frac{n_2 (n_2 + 1)}{2} - R_2$$

Once these components are determined, the overall U statistic for the test is simply the smaller value between \$U\_1\$ and \$U\_2\$. The figure below illustrates the resulting calculated values for the sum of ranks, sample sizes, and the subsequent U statistics based on our data. The minimal U value is the one we carry forward for the final calculation.

	A	B	C	D	E	F	G	H
1	<b>Treated</b>	<b>Untreated</b>		<b>Treated (ranks)</b>	<b>Untreated (ranks)</b>			
2	24	20		18.5	7.5			
3	25	23		21.5	14.5			
4	21	21		10	10			
5	22	25		12	21.5			
6	23	18		14.5	4			
7	18	17		4	1.5			
8	17	18		1.5	4			
9	28	24		24	18.5			
10	24	20		18.5	7.5			
11	27	24		23	18.5			
12	21	23		10	14.5			
13	23	19		14.5	6			
14								
15				<b>R1</b>	172	=SUM(D2:D13)		
16				<b>R2</b>	128	=SUM(E2:E13)		
17								
18				<b>N1</b>	12	=COUNT(A2:A13)		
19				<b>N2</b>	12	=COUNT(B2:B13)		
20								
21				<b>U1</b>	50	=E18*E19+E18*(E18+1)/2-E15		
22				<b>U2</b>	94	=E18*E19+E19*(E19+1)/2-E16		
23								
24				<b>U</b>	50	=MIN(E21, E22)		
25								
26								
27								

#### Step 4: Calculating the Z Test Statistic and P-Value

For sample sizes where both  $n_1$  and  $n_2$  are greater than 10 (as in our case, where  $n_1 = n_2 = 12$ ), the distribution of the U statistic can be accurately approximated by the standard normal distribution. Therefore, we convert the calculated U statistic into a standardized z test statistic. This conversion requires calculating the mean ( $\mu_U$ ) and standard deviation ( $\sigma_U$ ) of the U distribution under the null hypothesis.

The formulas used in Excel to calculate the mean and standard deviation of U, and subsequently the Z score, are implemented as follows:

**Mean ( $\mu_U$ ):** Calculated as  $\frac{n_1 n_2}{2}$ .

**Standard Deviation ( $\sigma_U$ ):** Calculated as  $\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$ .

**Z Statistic (Z):** Calculated as  $\frac{U - \mu_U}{\sigma_U}$ . (Note: Often a continuity correction of -0.5 is applied to U in the numerator for enhanced accuracy in the Z approximation.)

Once the Z score is obtained, the final step is calculating the two-tailed p-value associated with this Z score. This is typically done using the `NORM.S.DIST` function in Excel, which provides the cumulative probability associated with the Z score. By subtracting this cumulative probability from 1 and multiplying by 2 (for a two-tailed test), we obtain the final p-value that determines statistical significance.

	A	B	C	D	E	F	G	H	I	J
1	<b>Treated</b>	<b>Untreated</b>		<b>Treated (ranks)</b>	<b>Untreated (ranks)</b>					
2	24	20		18.5	7.5					
3	25	23		21.5	14.5					
4	21	21		10	10					
5	22	25		12	21.5					
6	23	18		14.5	4					
7	18	17		4	1.5					
8	17	18		1.5	4					
9	28	24		24	18.5					
10	24	20		18.5	7.5					
11	27	24		23	18.5					
12	21	23		10	14.5					
13	23	19		14.5	6					
14										
15				<b>R1</b>	172	=SUM(D2:D13)				
16				<b>R2</b>	128	=SUM(E2:E13)				
17										
18				<b>N1</b>	12	=COUNT(A2:A13)				
19				<b>N2</b>	12	=COUNT(B2:B13)				
20										
21				<b>U1</b>	50	=E18*E19+E18*(E18+1)/2-E15				
22				<b>U2</b>	94	=E18*E19+E19*(E19+1)/2-E16				
23										
24				<b>U</b>	50	=MIN(E21, E22)				
25										
26				<b>z</b>	-1.270170592	=(E24-E18*E19/2)/SQRT(E18*E19*(E18+E19+1)/12)				
27				<b>p</b>	0.20402387	=NORM.DIST(E26, 0, 1, TRUE)*2				
28										

## Interpreting the Results

The final stage of any statistical test is interpreting the resulting p-value against a predetermined significance level ( $\alpha$ ), which is conventionally set at 0.05. The null hypothesis ( $H_0$ ) for this test states that the two groups have the same distribution (i.e., the median mpg is the same for treated and untreated cars).

In this specific example, the calculated p-value is approximately **0.20402387**. Since this value (0.2040) is substantially larger than our chosen significance level of  $\alpha = 0.05$ , we fail to reject the null hypothesis. The observed difference in the ranking of mpg between the treated and untreated vehicles can reasonably be attributed to random sampling variation rather than a genuine effect of the treatment.

Therefore, based on the statistical evidence derived from the Mann-Whitney U Test conducted in Excel, we must conclude that we do not have sufficient evidence to claim that the true average mpg is statistically significantly different between the group of cars that received the fuel treatment and the control group. The fuel treatment, based on this experiment, does not appear to have a measurable impact on the central tendency of the car's fuel efficiency.

ARABPSYCHOLOGY.COM