

# Understanding the Shapiro-Wilk Test for Normality Using SPSS: A Step-by-Step Guide

Authored by  
**Mohammed looti**

November 12, 2025

## RECOMMENDED CITATION

Mohammed looti (2025). *Understanding the Shapiro-Wilk Test for Normality Using SPSS: A Step-by-Step Guide*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=18143>

## Introduction to the Shapiro-Wilk Test and the Normality Imperative

The **Shapiro-Wilk test** is a fundamental statistical tool utilized by researchers to rigorously assess whether an observed dataset conforms to the foundational assumptions of a **normal distribution**. This evaluation of data distribution, commonly referred to as normality testing, is an absolutely mandatory precursor to the application of many powerful **parametric statistical tests**, including T-tests, ANOVA, and Linear Regression. These methods are mathematically predicated on the assumption that the underlying population residuals, or the data itself, follow the characteristic bell-shaped curve.

The choice of the normality test is critical, and the **Shapiro-Wilk test** is widely considered the gold standard, particularly when analyzing smaller sample sizes (N less than 50). It consistently demonstrates superior statistical power compared to alternatives like the Kolmogorov-Smirnov test in detecting even subtle deviations from theoretical normality. A deep understanding of your data's distribution is not merely a formality; it is the cornerstone of reliable statistical inference. Failure to confirm normality when required can lead to compromised results, making your conclusions inaccurate or misleading.

In the context of **SPSS** (Statistical Package for the Social Sciences), executing the Shapiro-Wilk test is streamlined and highly integrated within the Descriptive Statistics module. This guide will walk you through the precise procedural steps necessary to navigate the software, generate the required output, and accurately interpret the critical **p-value**. The process begins by navigating the primary menu structure: **Analyze > Descriptive Statistics > Explore**. This sequence opens the necessary dialogue box to initiate the distributional analysis.

## Formulating and Interpreting the Test Hypotheses

Every statistical test, including the **Shapiro-Wilk test**, is grounded in a defined framework of hypotheses that dictate the statistical decision-making process. Mastering the correct articulation and interpretation of these hypotheses is fundamentally important for drawing statistically valid conclusions from the **SPSS** output. The test always pits the observed data against the ideal scenario of perfect normality.

The formal hypotheses utilized in the Shapiro-Wilk assessment are structured as follows:

**H<sub>0</sub> (Null Hypothesis):** The data originate from a population that **is** normally distributed. This hypothesis posits that the observed sample distribution does not differ statistically significantly from a theoretical **normal distribution**.

**H<sub>A</sub> (Alternative Hypothesis):** The data originate from a population that **is not** normally distributed. This hypothesis asserts that the sample distribution significantly deviates from normality.

The ultimate interpretation of the test rests upon comparing the calculated **p-value** (often designated as 'Sig.' in the SPSS output) against a predetermined threshold known as the **significance level** ( $\alpha$ ). By convention in most social and behavioral sciences,  $\alpha$  is set at 0.05. If the resulting p-value is less than or equal to this threshold ( $p \leq 0.05$ ), we are compelled to **reject the null hypothesis** ( $H_0$ ). Rejection signifies strong statistical evidence that the distribution of the data is, in fact, non-normal. Conversely, if the p-value exceeds the significance level ( $p > 0.05$ ), we retain or **fail to reject the null hypothesis**, concluding that there is insufficient statistical evidence to claim that the data significantly violates the assumption of **normal distribution**.

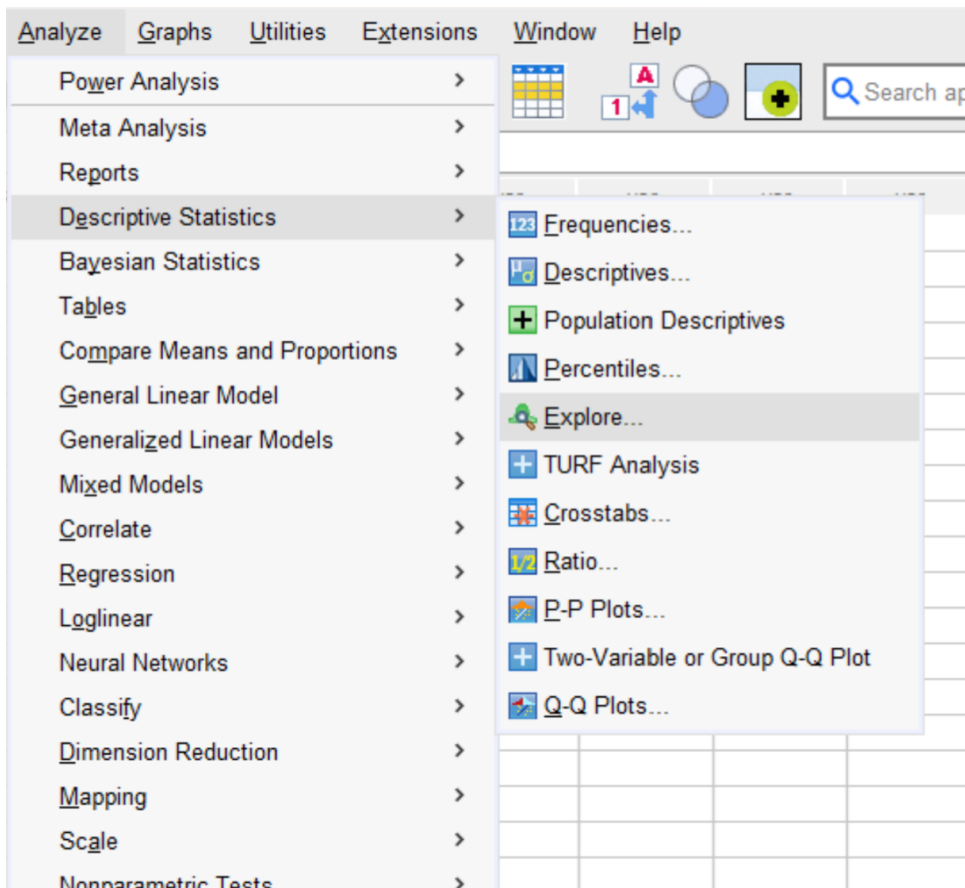
## Practical Example: Step-by-Step Execution in SPSS

To solidify this understanding, we will now execute a practical example. Consider a scenario where a researcher has collected the final examination scores from a sample of 25 university students. The primary analytical objective is to determine if the variable "Score" satisfies the crucial assumption of following a **normal distribution** before proceeding with any parametric analysis. We will utilize **SPSS** to perform the necessary computations.

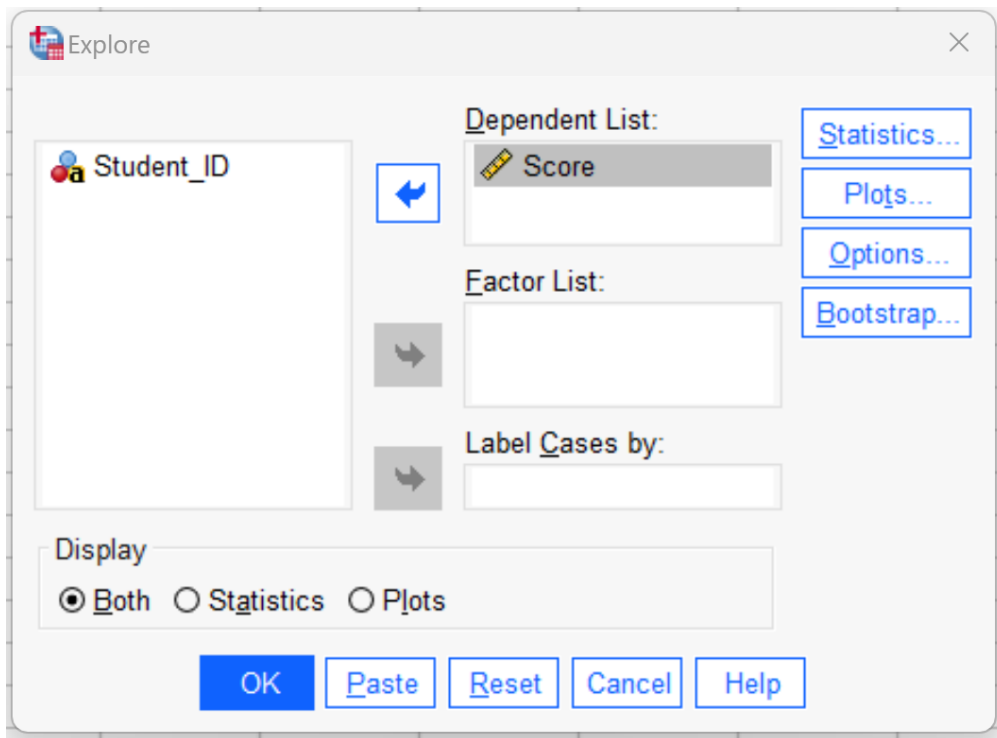
The initial dataset, which contains the scores for each individual student, must first be entered into the SPSS Data View editor. This view organizes the raw data ready for statistical processing.

	Student_ID	Score	var	var	v
1	1	88			
2	2	95			
3	3	92			
4	4	97			
5	5	96			
6	6	97			
7	7	94			
8	8	86			
9	9	91			
10	10	95			
11	11	97			
12	12	88			
13	13	85			
14	14	76			
15	15	68			
16	16	65			
17	17	87			
18	18	82			
19	19	77			
20	20	70			
21	21	90			
22	22	59			
23	23	68			
24	24	90			
25	25	92			
26					
27					
28					

To initiate the test, the user must meticulously follow the software's menu hierarchy. Begin by clicking the **Analyze** tab in the main toolbar. Next, hover the cursor over **Descriptive Statistics**, and finally, select the **Explore** option. This action is the gateway to generating comprehensive descriptive statistics, including essential tests for distributional assumptions.



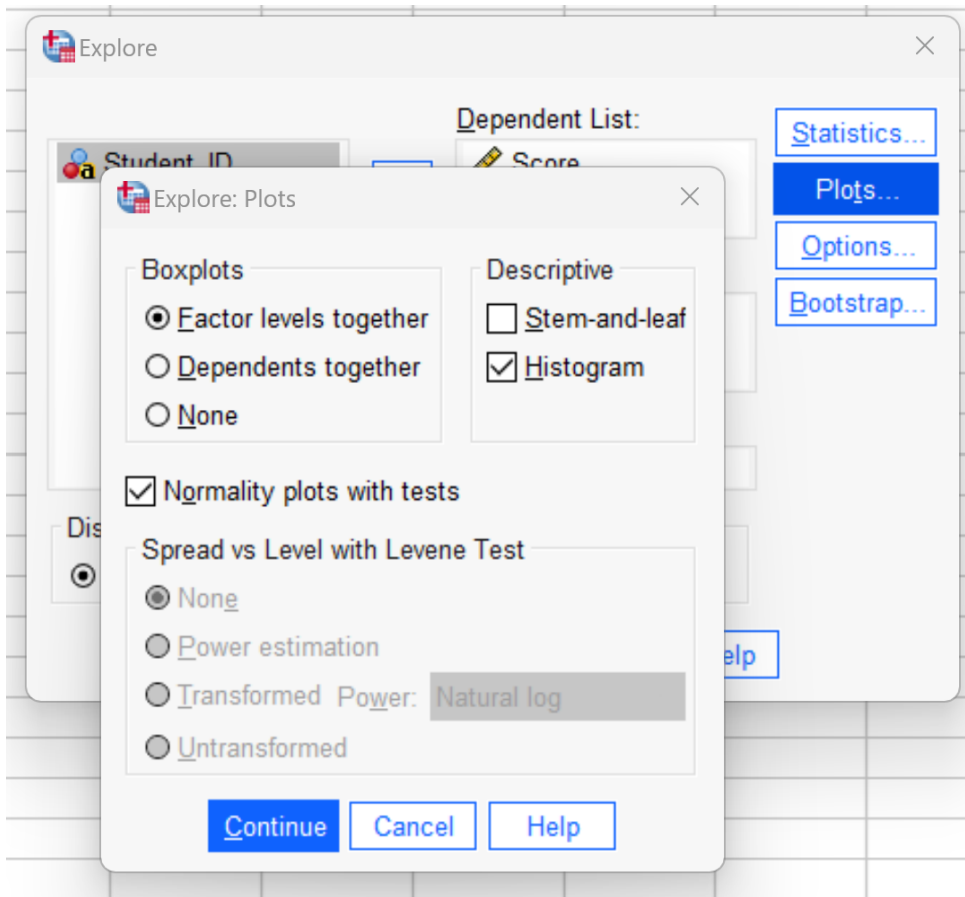
Once the **Explore** dialogue box appears, the variable under examination--in our case, the **Score** variable--must be transferred into the **Dependent List** pane. This panel is specifically designated for variables whose distributional properties and descriptive summaries are required for investigation. Ensure only the target variable is placed here before proceeding to configure the specific test options.



## Configuring Output Options and Running the Test

After correctly assigning the variable to the Dependent List, the subsequent critical step involves instructing **SPSS** to calculate and display the normality tests. Locate and click the **Plots** button situated on the right side of the main Explore dialogue box. This summons a sub-dialogue where both graphical representations and statistical test options can be specified.

Within the Plots window, it is imperative to check the box labeled **Normality plots with tests**. This specific command triggers **SPSS** to compute the required statistics, including the coefficients for the [Shapiro-Wilk test](#) and, for comparison, the Kolmogorov-Smirnov test. Furthermore, for robust qualitative analysis, we strongly recommend checking the box for the **Histogram**. The histogram offers an indispensable visual aid, allowing researchers to quickly gauge the data's shape, detect potential skewness, and assess the kurtosis visually.



Once all necessary options are activated, click **Continue** to exit the Plots configuration window. Finally, execute the analysis by clicking **OK** in the primary Explore dialogue box. SPSS will immediately process the request, generating the comprehensive output. This output typically includes summary descriptive statistics, the requested visual plots, and, most importantly, the dedicated "Tests of Normality" table containing the vital test results.

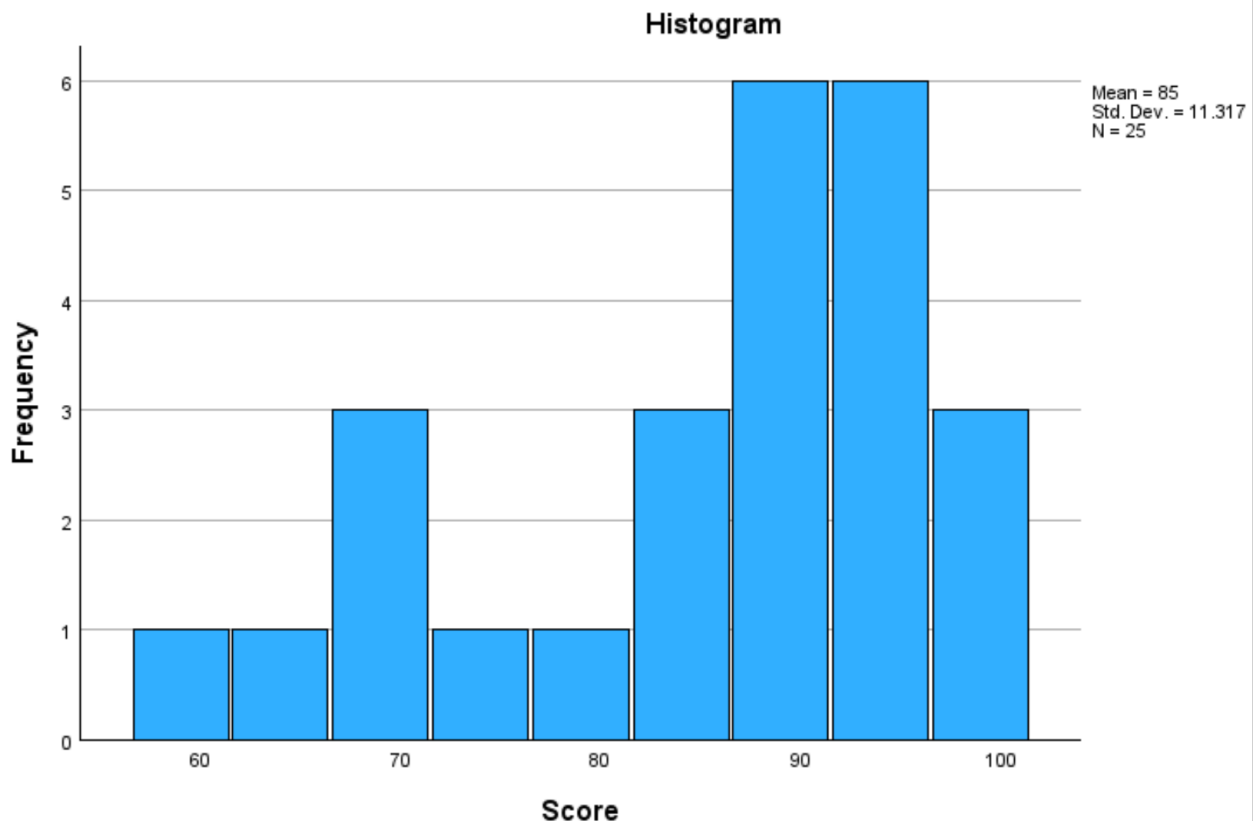
## Interpreting the Statistical and Visual Output

The output generated by **SPSS** will feature a table titled "Tests of Normality." This table is the focal point of our analysis. We must direct our attention exclusively to the row dedicated to the **Shapiro-Wilk test**. The output for our sample dataset of exam scores is presented below for careful interpretation:

Score	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Score	.180	25	.036	.877	25	.006

a. Lilliefors Significance Correction

### Score



To derive the statistical conclusion, we examine the column labeled "Sig." (Significance), which furnishes the exact **p-value**. In this specific example, the p-value reported is **.006**. We then apply our established decision rule by comparing this value against the standard **significance level** ( $\alpha = 0.05$ ). Since  $0.006$  is substantially less than  $0.05$ , we are led to the unequivocal conclusion that there is sufficient statistical evidence to **reject the null hypothesis** ( $H_0$ ).

The definitive statistical conclusion derived from this test is that the distribution of final exam scores **is not** normally distributed. This finding carries significant implications for subsequent analysis. If the researcher had planned to employ a parametric test that strictly requires the assumption of **normal distribution**, this violation necessitates the immediate consideration of methodological alternatives. These alternatives typically include applying data transformation techniques to normalize the scores, or, more commonly, adopting appropriate **non-parametric tests**, which do not rely on distributional assumptions, thereby ensuring the validity and trustworthiness of the final

statistical inference.

## Visual Reinforcement: Assessing Skewness and Kurtosis

While the p-value offers a crucial binary (reject/fail to reject) decision, the histogram generated alongside the test provides invaluable qualitative context regarding the nature of the deviation from normality. For data that perfectly adhered to a normal distribution, the histogram would closely mirror a perfectly symmetrical bell curve.

Upon analyzing the histogram produced in the **SPSS** output for the exam scores, a visual inspection immediately reveals a lack of symmetry. The bulk of the scores are clearly concentrated toward the higher end of the scale, resulting in a long, pronounced tail extending towards the left (lower scores). This specific asymmetrical pattern is the defining characteristic of a **negatively skewed** (or left-skewed) distribution.

The visual confirmation of this severe skewness serves to strongly reinforce the quantitative result obtained from the [Shapiro-Wilk test](#) ( $p = 0.006$ ). This powerful convergence of evidence--the statistical rejection of the null hypothesis coupled with clear graphic proof of skewness--solidifies the conclusion: the assumption of normality is convincingly violated for this particular sample dataset. Researchers must therefore proceed with caution and utilize robust statistical methods appropriate for skewed data.

## Next Steps in Statistical Analysis and Further Resources

Successfully mastering the execution and interpretation of the **Shapiro-Wilk test** represents a vital foundational skill in quantitative research using **SPSS**. Since normality is such a pervasive assumption, researchers must be proficient in subsequent procedures that either rely on this assumption or serve as alternatives when it is violated.

To continually build expertise and ensure methodological rigor, researchers should explore advanced tutorials and resources covering related statistical practices. Key areas for development include:

Detailed instructions on how to compute and accurately interpret fundamental measures of **central tendency** (mean, median, mode) and **variability** (standard deviation, variance).

Guidance on executing core [parametric statistical tests](#), such as the Independent Samples T-Test, when the normality assumption holds true.

Advanced techniques for identifying and correcting for common statistical issues like **heteroscedasticity**.

Comprehensive guidance on the systematic selection and application of appropriate [non-parametric tests](#) whenever the critical assumption of [normal distribution](#) cannot be maintained

or is definitively violated.

Adhering to these foundational steps ensures that statistical inferences are robust, ethical, and defensible in academic or professional contexts.