

Understanding the Kolmogorov-Smirnov Test in SPSS: A Practical Guide

Authored by
Mohammed looti

November 12, 2025

RECOMMENDED CITATION

Mohammed looti (2025). *Understanding the Kolmogorov-Smirnov Test in SPSS: A Practical Guide*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=18149>

Introduction to the Kolmogorov-Smirnov Test for Normality

The **Kolmogorov-Smirnov (K-S) test** is a powerful and widely used [statistical test](#) employed to determine if a sample of data significantly deviates from a specified theoretical probability distribution. While the K-S test possesses broad utility across various distributions, its most frequent and critical application in quantitative research, particularly within the social sciences, is the assessment of whether a dataset follows a [normal distribution](#). This preliminary assessment of data normality is absolutely foundational because many core parametric statistical procedures, such as T-tests and Analysis of Variance (ANOVA), rely on the strong assumption that the underlying population data is normally distributed. Violating this distributional assumption can lead to severely biased results and fundamentally flawed inferential conclusions, highlighting why robust preliminary procedures like the **Kolmogorov-Smirnov test** are indispensable in modern statistical methodology.

The theoretical mechanism underpinning the K-S test involves comparing the observed empirical cumulative distribution function (ECDF) of the sample data against the expected theoretical cumulative distribution function (CDF) of the hypothesized distribution (e.g., the normal curve). The resulting test statistic, D , precisely quantifies the maximum vertical distance measured between these two distribution curves. Conceptually, a smaller distance signifies a better fit between the sample data and the theoretical distribution. Grasping this mathematical foundation is essential for accurately interpreting the complex numerical output generated by sophisticated statistical software packages, such as [IBM SPSS Statistics](#), which seamlessly automates these intricate calculations, allowing researchers to concentrate on meaningful data interpretation rather than tedious manual computation.

This comprehensive tutorial is specifically designed to guide users through the practical application of the one-sample **Kolmogorov-Smirnov test** for normality using [SPSS Statistics](#). This approach, which utilizes the intuitive menu path **Analyze > Descriptive Statistics > Explore**, is generally recognized as the most efficient and user-friendly method within the software. We will systematically dissect a practical example, meticulously detailing every necessary step, click sequence, and setting required. By the end of this guide, you will be fully equipped to confidently determine the crucial distributional characteristics of your own research data and correctly apply this vital preliminary statistical assessment.

Understanding the Hypotheses of Normality Testing

Prior to initiating any formal statistical procedure, it is mandatory to establish and clearly define the null and alternative hypotheses. The fundamental structure of hypothesis testing for the **Kolmogorov-Smirnov test**, especially when deployed to rigorously check for data normality, maintains a consistent structure across virtually all statistical software environments. These

hypotheses serve to frame the core research question: Does the distribution observed in our collected data significantly deviate from the perfect symmetry and characteristics of a theoretical normal distribution? The precise formulation of these hypotheses is critical, as the final statistical decision rests entirely on whether there is sufficient evidence to reject the initial, conservative assumption, which is invariably encapsulated within the [null hypothesis](#) (H0).

When assessing data normality, the K-S test employs the following specific and highly standardized hypotheses:

H0 (Null Hypothesis): The data **is** normally distributed. This hypothesis assumes the absence of an effect or difference, postulating that the sample data was accurately drawn from a population that inherently adheres to the theoretical [normal distribution](#).

HA (Alternative Hypothesis): The data **is not** normally distributed. This hypothesis is statistically accepted only if there is overwhelming statistical evidence that directly contradicts the null hypothesis, suggesting that the data exhibits a significant and non-random deviation from the expected bell-shaped curve.

The definitive judgment to reject or fail to reject the [null hypothesis](#) is based on the resulting [p-value](#), which is automatically calculated and provided by the statistical test. If this calculated p-value is numerically smaller than your pre-established [significance level](#) (typically set at $\alpha = 0.05$), then you are compelled to reject H0. A rejection of H0 signifies that the observed distributional difference is statistically significant, leading to the firm conclusion that the data is not normally distributed. Conversely, if the p-value exceeds the significance level, you retain the null hypothesis, concluding that there is insufficient statistical evidence to definitively claim non-normality.

Prerequisites: Preparing Your Data in SPSS

To vividly illustrate the precise, step-by-step procedure for conducting the K-S test, we will utilize a simulated, yet realistic, dataset prepared within the [SPSS](#) environment. For our purposes, imagine we have meticulously recorded the final exam scores attained by 25 distinct students enrolled in a specific university course. Before initiating any complex statistical analysis, the primary prerequisite is ensuring that this quantitative data is accurately entered, properly labeled, and correctly defined within the SPSS data editor. Achieving the proper data setup is recognized as the foundational and most critical initial step for ensuring the rigor and validity of any subsequent statistical analysis.

The dataset, which is visually represented in the image below, comprises a single, continuous quantitative variable explicitly named "Score." Our primary objective is to subject this variable to the rigorous scrutiny of the **Kolmogorov-Smirnov test** to determine definitively if this particular collection of scores adheres to the strict assumptions required for a [normal distribution](#), a prerequisite for future parametric statistical investigation.

	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					

Our analytical objective is unambiguous: to test the actual distribution of these 25 student scores against the idealized theoretical normal curve. If the test reveals that the distribution is significantly non-normal, this crucial finding will immediately guide our decisions toward the necessity of applying appropriate non-parametric statistical methods or implementing required data transformations. This initial, exploratory step, powered by the K-S test, functions as a vital quality control measure, ensuring the foundations of our inferential statistics are robust and trustworthy.

Executing the K-S Test Step-by-Step in SPSS

The most accessible, comprehensive, and reliable methodology for performing the **Kolmogorov-Smirnov test** within [SPSS Statistics](#) involves leveraging the dedicated Explore function. This

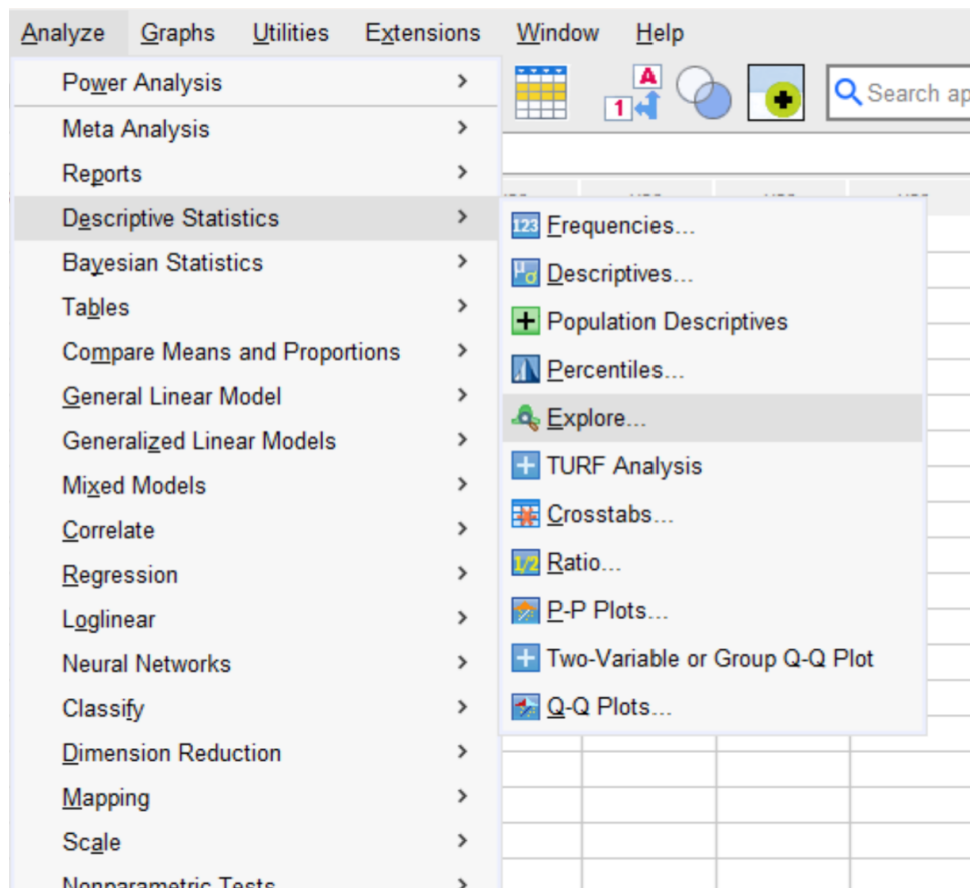
function is specifically engineered for in-depth distributional assessment, guaranteeing that researchers receive not only the definitive numerical test results but also invaluable complementary visual aids, such as detailed histograms and boxplots, which together offer a holistic and comprehensive view of the data's overall shape.

To successfully execute the normality test, follow these precise procedural steps within the SPSS interface:

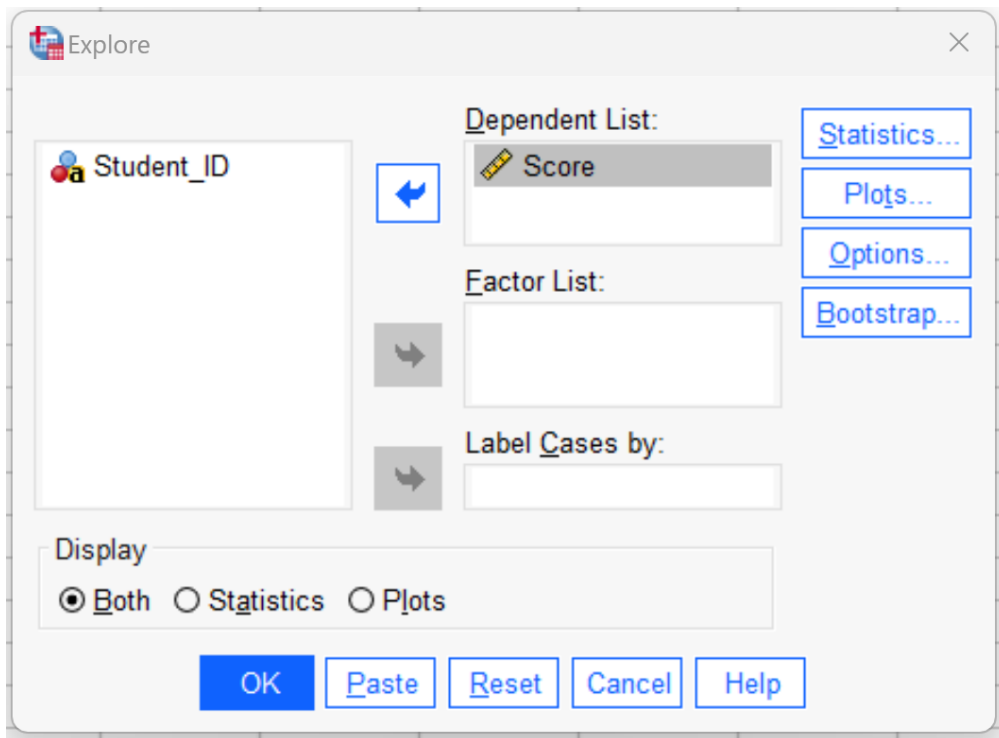
Navigate directly to the main menu bar located at the top of the SPSS window and click the **Analyze** tab.

Move your cursor over or click the **Descriptive Statistics** option.

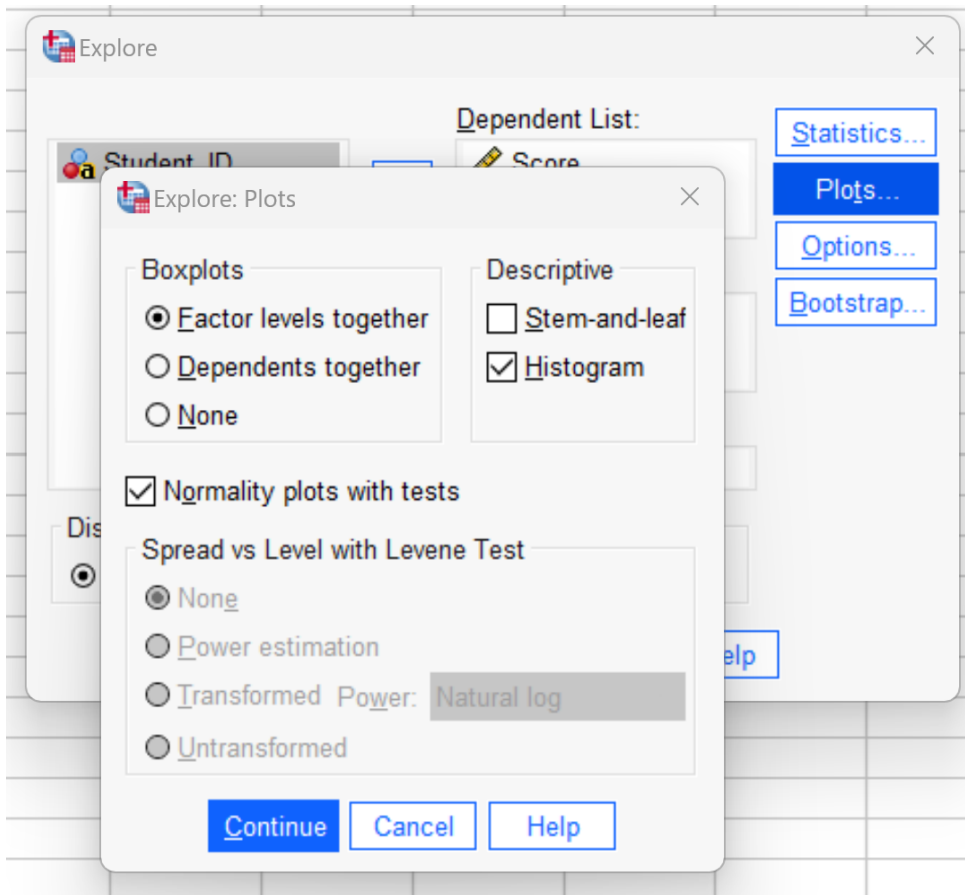
Select and click the **Explore** command:



Upon the appearance of the "Explore" dialog box, your immediate task is to designate the variable intended for analysis. Drag the variable containing the raw data you wish to test--in this case, the **Score** variable--from the source list on the left and accurately place it into the **Dependent List** panel on the right. This action formally designates the variable whose distributional characteristics are the focus of the normality assessment.



With the variable correctly positioned, the subsequent crucial step involves precisely specifying the required analytical outputs. Click the **Plots** button situated on the right side of the Explore dialog box. Within the resulting "Explore: Plots" configuration window, ensure you select the necessary options for a comprehensive analysis. Specifically, check the box adjacent to **Histogram** and, most importantly, check the box labeled **Normality plots with tests**. This final selection is the key command that instructs SPSS to calculate, generate, and display the full results for the **Kolmogorov-Smirnov test**, alongside the Shapiro-Wilk test.



After verifying that these selections are confirmed, click the **Continue** button to exit the Plots submenu. Finally, click **OK** in the main Explore window to successfully execute the command. SPSS will now process the request and swiftly generate the complete output in the Viewer window, providing all necessary descriptive statistics, graphical visualizations, and the definitive results of the normality test.

Interpreting the SPSS Output and Drawing Conclusions

The output generated by SPSS is typically extensive, but the most vital piece of information pertinent to the **Kolmogorov-Smirnov test** is consolidated within the "Tests of Normality" table. This table clearly presents the calculated K-S test statistic and, crucially, the associated [p-value](#), which is conveniently labeled as 'Sig.' (for Significance). It is this single p-value that dictates the final statistical decision regarding the [null hypothesis](#) (H_0).

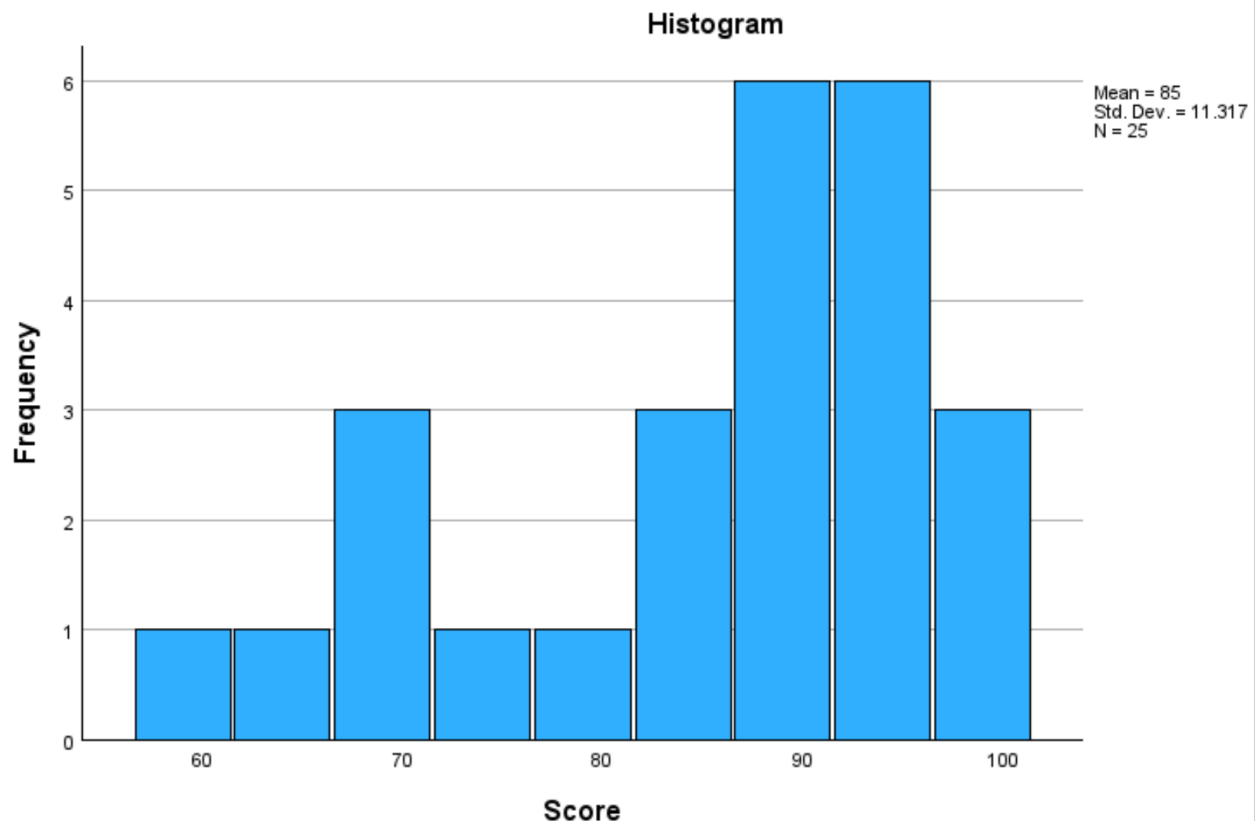
The following output table summarizes the numerical results derived from the normality test performed on our sample of student scores, provided alongside the corresponding histogram for essential visual context:

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Score	.180	25	.036	.877	25	.006

a. Lilliefors Significance Correction

Score



Focusing intently on the "Tests of Normality" table, specifically the row dedicated to the **Kolmogorov-Smirnov test**, we observe that the p-value (Sig.) is calculated to be **.036**. To interpret this result correctly, we must recall the established hypotheses:

H₀: The data **is** normally distributed.

H_A: The data **is not** normally distributed.

Since the calculated [p-value](#) of \$0.036\$ is numerically less than the conventional critical [significance level](#) of \$0.05\$, we possess sufficient statistical evidence to confidently reject the null hypothesis. The unavoidable conclusion is therefore that the distribution of our observed exam scores **is not** normally distributed, as the deviation from the theoretical curve is statistically significant.

Final Assessment and Practical Implications

The definitive numerical conclusion derived from the **Kolmogorov-Smirnov test** is powerfully corroborated by the visual evidence presented in the output. The histogram, which was generated automatically by [SPSS](#), clearly illustrates that the distribution of exam scores fails to conform to the expected symmetrical, bell-shaped pattern characteristic of a [normal distribution](#). Instead, the data exhibits noticeable negative skewness, visually confirming the statistical rejection of the null hypothesis.

This critical finding--the determination that the data is not normally distributed--carries significant implications for all subsequent statistical modeling efforts. If the researcher had planned to utilize parametric tests, which explicitly assume normality, they must now critically re-evaluate their analytical strategy. The viable primary options include attempting data transformations (such as logarithmic or square root transformations) in an effort to achieve approximate normality, or, more frequently, transitioning to the adoption of appropriate non-parametric statistical alternatives. Non-parametric tests are specifically designed to be valid and robust when dealing with data that fundamentally violates strict distributional assumptions, thereby ensuring the integrity and trustworthiness of the research findings when normality cannot be assumed or successfully achieved.

Mastering the efficient application and accurate interpretation of the **Kolmogorov-Smirnov test** in SPSS is an essential and non-negotiable skill for any serious data analyst or researcher, guaranteeing that the foundational distributional assumptions required for rigorous statistical inference are correctly addressed before proceeding to more complex and demanding hypothesis testing.

Additional Resources for SPSS Operations

The following supplementary tutorials provide detailed explanations on how to perform other common and necessary operations in SPSS, effectively building upon your foundational knowledge of data distribution analysis: