

How to Perform a Kolmogorov-Smirnov Test in SPSS to Check for Distribution Fit

Authored by
stats writer

January 24, 2026

RECOMMENDED CITATION

stats writer (2026). *How to Perform a Kolmogorov-Smirnov Test in SPSS to Check for Distribution Fit*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=127328>

The Kolmogorov-Smirnov test (often abbreviated as the **K-S test**) is a powerful, non-parametric statistical test designed to evaluate whether a sample of data plausibly originated from a specific theoretical probability distribution, or to compare two samples to determine if they originated from the same distribution. When analyzing data in fields ranging from social sciences to engineering, establishing the underlying distribution is often a critical prerequisite for selecting appropriate subsequent statistical methodologies. The K-S test operates by comparing the empirical cumulative distribution function (ECDF) of the sample data against the cumulative distribution function (CDF) of the hypothesized theoretical distribution.

The primary utility of the one-sample K-S test, which is the focus here, is its ability to test for goodness-of-fit. This means we are testing how well our observed data aligns with a theoretical model, such as the **Normal distribution**, the **Uniform distribution**, or the **Exponential distribution**. While other goodness-of-fit tests exist, the K-S test is widely used due to its simplicity and its sensitivity to deviations in both location and shape. Understanding the distributional properties of your data is fundamentally important because many advanced statistical models, such as t-tests and ANOVA, rely heavily on the assumption that the data are normally distributed.

Before proceeding with the practical execution in **SPSS** (Statistical Package for the Social Sciences), it is essential to appreciate the calculated statistic. The K-S test statistic (D) measures the maximum absolute difference between the observed ECDF and the expected CDF. A smaller value of D suggests a better fit between the sample data and the hypothesized distribution. The significance of this D statistic is then evaluated using a calculated p-value, which ultimately dictates whether we can accept or reject the premise that the data follows the specified distribution.

Assumptions and Applications of the K-S Test

Like all rigorous statistical procedures, the Kolmogorov-Smirnov test rests on specific assumptions regarding the nature of the data being analyzed. First and foremost, the test assumes that the sample observations are **independent**. This means that the measurement of one observation does not influence the measurement of any other observation in the sample. Secondly, the test requires that the data come from a **continuous distribution**. While the test can sometimes be adapted for discrete data, its power and validity are significantly reduced, making specialized tests like the **Chi-Square goodness-of-fit test** often more appropriate for categorical or highly discrete measurements. Failure to meet these fundamental assumptions can lead to incorrect conclusions regarding the distribution of the population.

The K-S test finds broad application across various analytical scenarios. For instance, in quality control, it can determine if product measurements conform to a specific standard distribution. In finance, analysts might use it to assess whether daily stock returns follow a Normal distribution, which is crucial for risk modeling. Furthermore, it serves a critical preparatory role in data analysis;

before applying parametric methods (those that assume a specific distribution), researchers often use the K-S test to confirm the normality assumption. If the assumption of normality is violated, the researcher must either employ non-parametric alternatives or apply transformations to the data.

It is important to acknowledge that the K-S test is generally less sensitive to deviations in the tails of the distribution compared to alternative tests, such as the **Shapiro-Wilk test**, particularly when testing for normality. Due to this characteristic, many statisticians recommend the Shapiro-Wilk test as the preferred method for assessing normality when the sample size is small to moderate (e.g., $n < 50$). However, for larger sample sizes, the K-S test remains a robust and acceptable option. Careful consideration of these nuances, often requiring consultation with a statistician or data analyst, ensures proper test selection and accurate interpretation of the results.

Hypotheses Formulation

Formal statistical testing always begins with the establishment of two opposing statements: the null hypothesis and the alternative hypothesis. When conducting a one-sample Kolmogorov-Smirnov test, particularly for assessing normality, these hypotheses are specifically framed around the concept of distribution adherence. The goal is to determine if there is enough statistical evidence to reject the baseline assumption that the sample data aligns with the hypothesized distribution.

The standard hypotheses for the K-S test are defined as follows:

H₀: The null hypothesis states that the data follow the specified theoretical distribution (e.g., "The data are normally distributed"). This is the position of no difference or no effect that the test attempts to disprove.

H_A: The alternative hypothesis (or H₁) states that the data do **not** follow the specified theoretical distribution (e.g., "The data are not normally distributed"). This hypothesis is accepted only if the evidence strongly suggests that H₀ should be rejected.

The decision to reject or retain the **null hypothesis** hinges entirely on the calculated p-value. The p-value represents the probability of observing our sample data (or data more extreme) if the null hypothesis were true. We compare this probability to a predetermined **significance level**, denoted as α (alpha). Typically, researchers set α at 0.05, meaning they are willing to accept a 5% risk of incorrectly rejecting a true null hypothesis (a Type I error). If the p-value resulting from the test is less than the chosen α (e.g., $p < 0.05$), then we have statistically significant evidence to reject H₀, concluding that the data does not conform to the hypothesized distribution.

Method 1: Direct Nonparametric Test in SPSS

While SPSS offers several pathways to conduct distribution tests, the most straightforward approach involves utilizing the dedicated Nonparametric Tests menu. This method is quick and

directly calculates the Kolmogorov-Smirnov test statistic and its corresponding significance value against a specified theoretical distribution. This procedure is particularly useful when the user only requires the raw test result without needing additional exploratory graphics or detailed descriptive statistics. Ensure your data file is open and the variable you wish to test is correctly defined in the data view before proceeding.

The execution steps for this direct method are precise:

Navigate to the top menu bar, click on **Analyze**, and hover over **Nonparametric Tests**.

From the cascading submenu, select **Legacy Dialogs**, and then click on **1-Sample K-S...** (for One-Sample Kolmogorov-Smirnov Test).

In the resulting dialog box, identify the variable containing your sample data (e.g., 'Exam_Scores') and move it into the **Test Variable List** panel using the arrow button.

In the **Test Distribution** section, you must specify the theoretical distribution against which your data will be tested. For example, check the box next to **Normal** if you are testing for normality. You may also specify other distributions like Uniform, Poisson, or Exponential, depending on your research question.

Ensure all settings are correct, and then click **OK** to run the analysis.

Upon execution, the results will appear instantly in the dedicated **SPSS Output Viewer** window. The output table, usually labeled "One-Sample Kolmogorov-Smirnov Test," provides the summary descriptive statistics, the calculated K-S test statistic (the maximum difference D), and crucially, the **Asymp. Sig. (2-tailed)**, which represents the p-value. Interpreting this p-value relative to your chosen significance level (e.g., 0.05) allows you to draw immediate conclusions about whether your sample data significantly deviates from the hypothesized distribution. If $p < 0.05$, reject the **null hypothesis** of distributional adherence.

Method 2: Preferred Approach via Descriptive Statistics and Exploration

While the direct Nonparametric Test method (Method 1) is functional, the method preferred by many data analysts, especially when assessing normality, involves using the **Explore** function under Descriptive Statistics. This pathway not only calculates the Kolmogorov-Smirnov test but also simultaneously generates essential exploratory graphics--such as **Histograms** and **Q-Q plots**--which provide visual confirmation of the distributional properties, enhancing the statistical decision-making process. Visual inspection is a vital complement to numerical tests, helping to identify skewness, outliers, and kurtosis.

To execute the K-S test using the Explore function in SPSS, follow these comprehensive steps:

Click the **Analyze** tab in the main menu.

Select **Descriptive Statistics**, and then click on **Explore...**

In the Explore dialog box, move the variable you are testing (the dependent variable, e.g., Score) into the **Dependent List** box.

Crucially, click on the **Plots** button located on the right side of the dialog box.

Within the Plots options, ensure the boxes next to **Histogram** and **Normality plots with tests** are checked. The 'Normality plots with tests' selection is what instructs SPSS to calculate and display both the K-S test and the Shapiro-Wilk test statistics in the output.

Click **Continue** to return to the main Explore dialog, and then click **OK** to generate the output.

This integrated approach is highly recommended for preliminary data analysis. The visual aids, specifically the histogram and the normal Q-Q plot, offer intuitive evidence that either supports or contradicts the numerical p-value result. For instance, a significantly skewed histogram or a Q-Q plot where the data points deviate sharply from the reference line visually reinforce a statistically significant K-S test result ($p < \alpha$), leading to greater confidence in rejecting the **null hypothesis** that the data is normally distributed.

Detailed Example: Assessing Exam Scores for Normality

To illustrate the practical application of the Kolmogorov-Smirnov test in SPSS, let us consider a hypothetical dataset containing the final exam scores of 25 students. Our objective is to determine whether the distribution of these scores adheres to the assumptions of a Normal distribution, a requirement for many parametric inferential statistics. Assume the data is already entered into the SPSS Data Editor, with the scores residing in a variable named 'Score'.

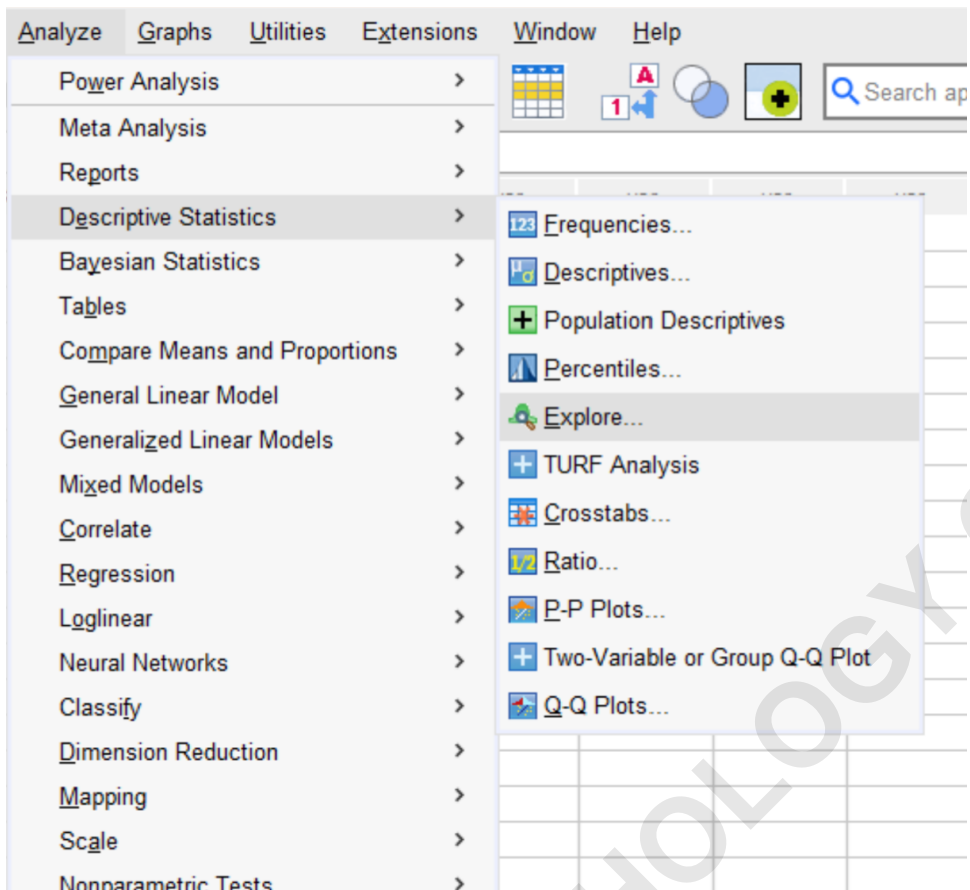
The raw dataset, as visualized in the SPSS Data Editor, appears as follows:

	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					

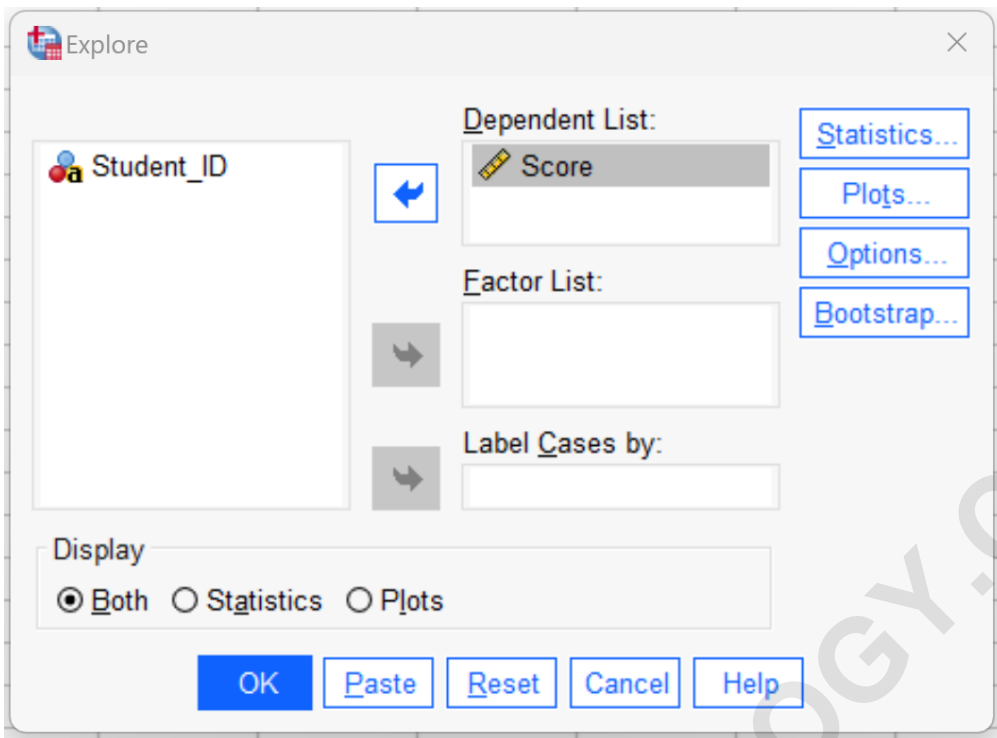
Following Method 2, we initiate the analysis by selecting **Analyze > Descriptive Statistics > Explore**. We place the 'Score' variable into the Dependent List. The crucial step is accessing the Plots menu and ensuring **Normality plots with tests** is checked, alongside the **Histogram** option. This setup ensures that both the numerical results and the graphical evidence are generated simultaneously, providing a holistic view of the score distribution.

The subsequent steps involve navigating the dialog boxes:

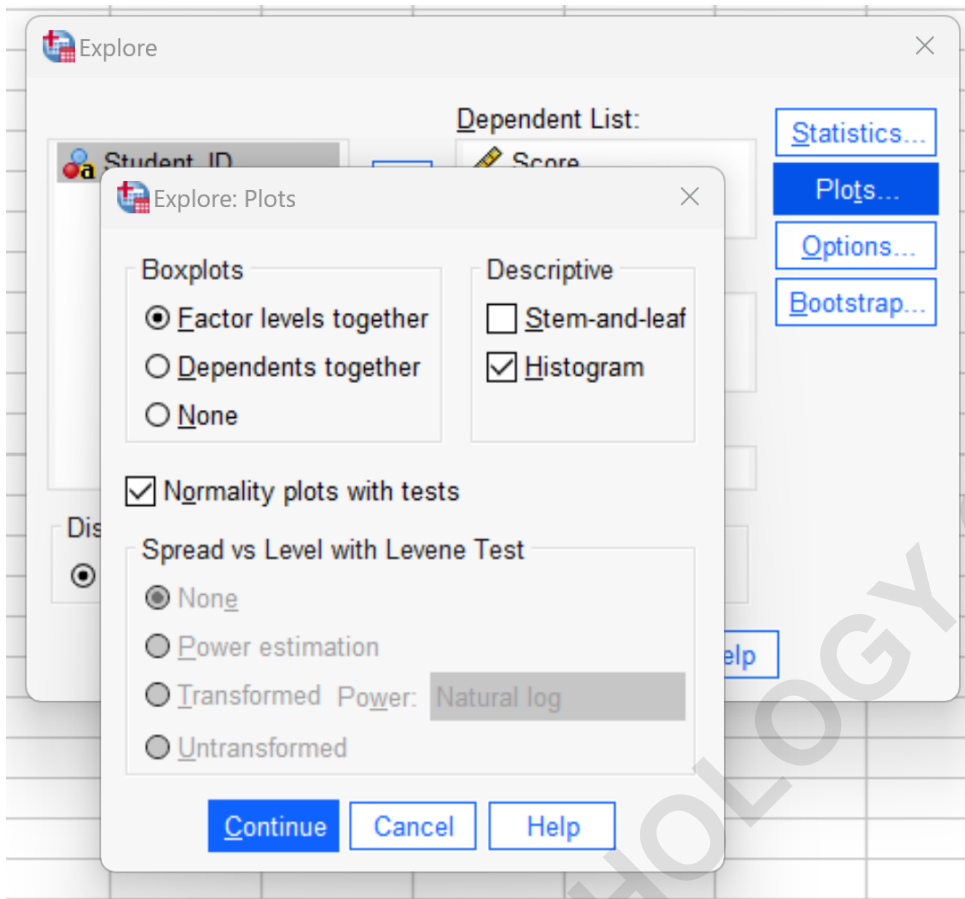
First, initiating the Explore function:



Second, defining the dependent variable:



Third, configuring the plot options to include the necessary tests:



Once all settings are confirmed and the analysis is run, the SPSS Output Viewer provides a series of tables and charts, including the critical **Tests of Normality** table, which summarizes the K-S results for the exam scores variable. This output forms the basis for our statistical conclusion.

Interpreting the SPSS Output and Drawing Conclusions

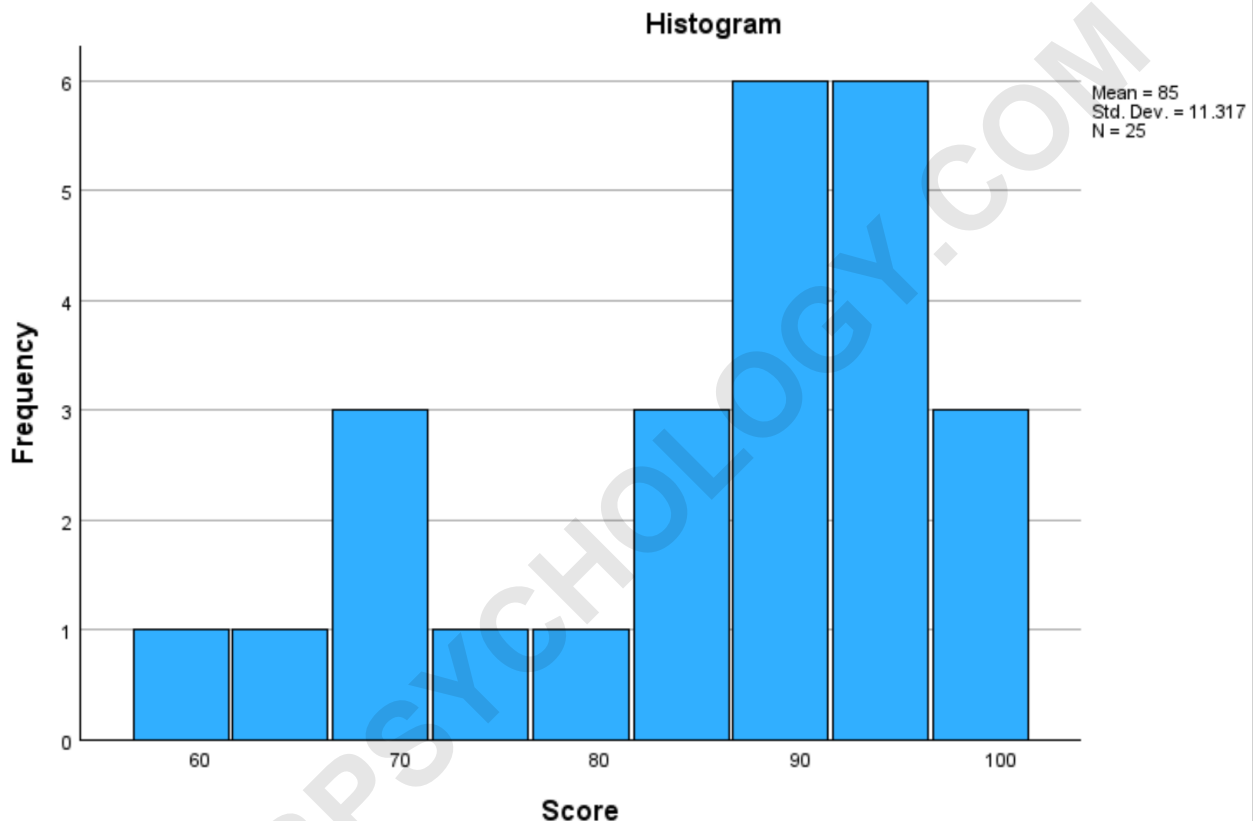
The final output generated by SPSS is comprehensive, but the focus for the Kolmogorov-Smirnov test lies primarily within the "Tests of Normality" table. This table presents the calculated K-S statistic and the associated significance level (p-value). For our example concerning the exam scores, the relevant portion of the output looks like this, typically accompanied by the calculated statistic D and degrees of freedom (df):

Tests of Normality

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

a. Lilliefors Significance Correction

Score



Examining the output table reveals the asymptotic significance value (p-value) for the Kolmogorov-Smirnov test. In this specific case, we observe that the p-value is **0.036**. To interpret this result, we must compare it to our chosen significance level, $\alpha = 0.05$. Since $0.036 < 0.05$ ($p < \alpha$), we adhere to the standard decision rule in hypothesis testing: we reject the null hypothesis (H_0).

The rejection of the null hypothesis leads to the conclusion that there is statistically significant evidence to state that the distribution of exam scores **is not** a Normal distribution. Had the p-value been greater than 0.05 (e.g., 0.150), we would have failed to reject H_0 , suggesting that the data did not significantly deviate from a normal distribution. Furthermore, the accompanying histogram, which is also generated by the Explore function, visually confirms this finding. The distribution of scores does not display the characteristic symmetrical "bell curve" shape typically associated with a perfectly Normal distribution, instead showing some clear asymmetry or

skewness, reinforcing the statistical conclusion.

Important Considerations and Next Steps

While the Kolmogorov-Smirnov test is a useful tool, practitioners must remain aware of its limitations and the implications of its results. Specifically, when the sample size is very large (e.g., $N > 1,000$), the K-S test can become overly sensitive, often detecting minor, practically irrelevant deviations from the theoretical distribution and leading to the rejection of the **null hypothesis** even when the data is reasonably close to normal. Conversely, with very small samples, the test may lack the necessary power to detect meaningful non-normality.

If the test results, whether K-S or Shapiro-Wilk, indicate a significant departure from normality (as in our example), the researcher must adjust their methodology. Potential next steps include employing non-parametric statistical tests, which do not rely on distributional assumptions (e.g., the **Mann-Whitney U test** instead of the independent samples t-test). Alternatively, data transformation techniques (such as log transformation or square root transformation) can sometimes normalize the data, allowing the researcher to proceed with parametric testing. However, transformations must be used judiciously, as they can sometimes complicate the interpretation of results.

Finally, irrespective of the software used or the specific test performed, the ultimate interpretation of complex statistical outputs requires careful consideration. It is always highly recommended to seek consultation from a qualified **statistician** or experienced **data analyst**, especially when dealing with critical research questions or datasets where distributional assumptions are foundational. Accurate statistical practice ensures the validity and reliability of research findings.

Perform a Kolmogorov-Smirnov Test in SPSS

The **Kolmogorov-Smirnov (K-S) test** is a statistical procedure utilized to determine whether or not a sample of observations originates from a specified theoretical distribution, often used as a goodness-of-fit metric.

In most practical applications, the K-S test is employed to check if a given data distribution follows a Normal distribution, which is a prerequisite for parametric tests.

The K-S test uses the following hypotheses framework:

H₀: The data **is** normally distributed (The sample data conforms to the theoretical distribution).

H_A: The data **is not** normally distributed (The sample data significantly deviates from the theoretical distribution).

If the p-value that results from the test is less than your chosen significance level (e.g. 0.05) then you can reject the null hypothesis and conclude that the data is not normally distributed.

The easiest and most informative way to perform a Kolmogorov-Smirnov test in SPSS is by navigating through **Analyze > Descriptive Statistics > Explore**, which generates graphical support along with the test statistics.

The following example demonstrates how to perform this integrated test in practice using sample data.

Example: How to Perform a Kolmogorov-Smirnov Test in SPSS

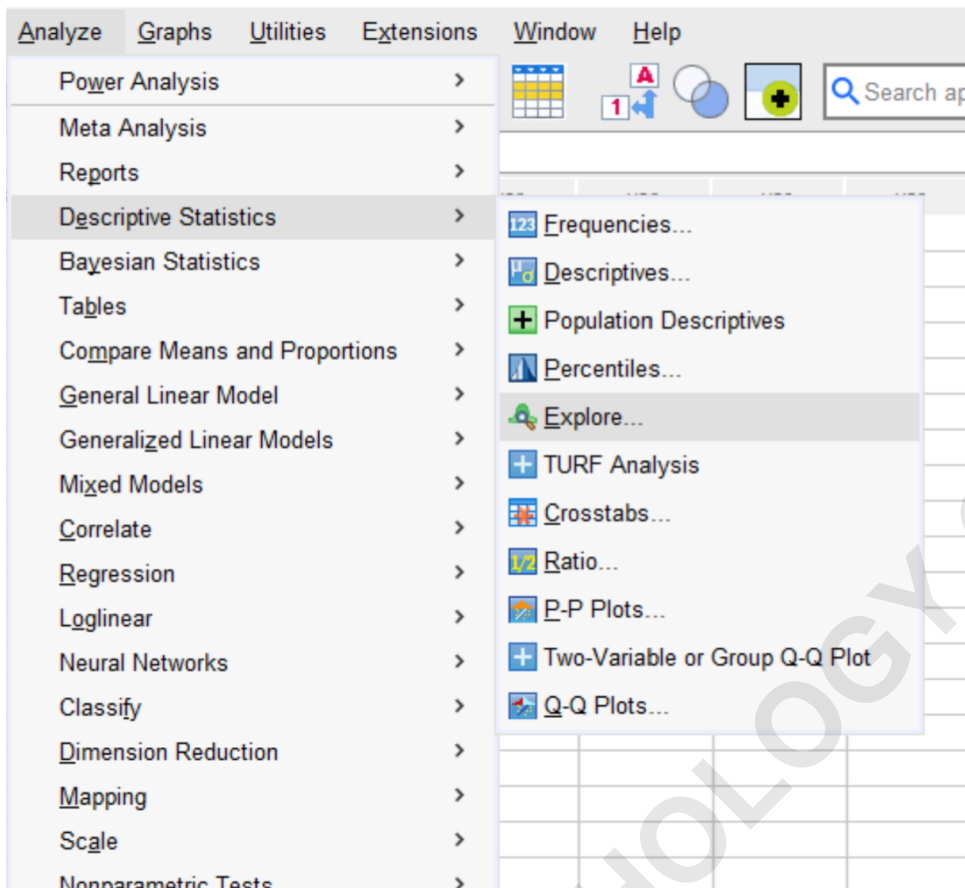
Suppose we have the following dataset in SPSS that contains information about the final exam scores received by 25 different students in some class:

ARABPSYCHOLOGY.COM

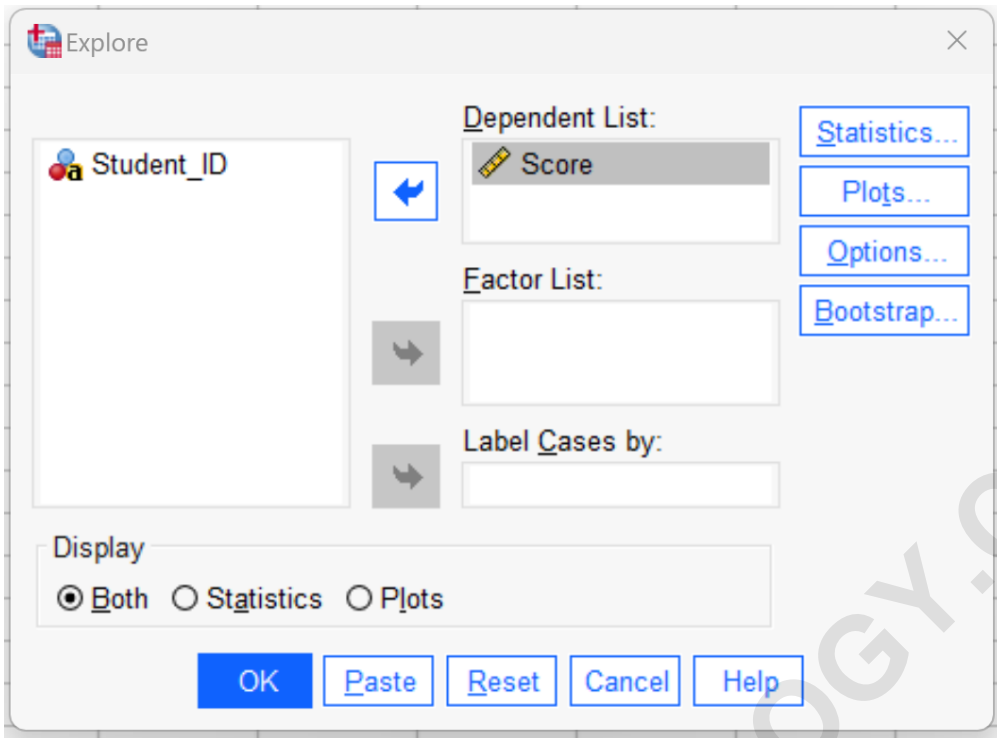
	Student_ID	Score	var	var	v
1	1	88			
2	2	95			
3	3	92			
4	4	97			
5	5	96			
6	6	97			
7	7	94			
8	8	86			
9	9	91			
10	10	95			
11	11	97			
12	12	88			
13	13	85			
14	14	76			
15	15	68			
16	16	65			
17	17	87			
18	18	82			
19	19	77			
20	20	70			
21	21	90			
22	22	59			
23	23	68			
24	24	90			
25	25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

Suppose that we would like to perform a Kolmogorov-Smirnov test to determine if the exam scores are normally distributed, thereby confirming the suitability for parametric testing.

To do so, click the **Analyze** tab, then click **Descriptive Statistics**, then click **Explore**:

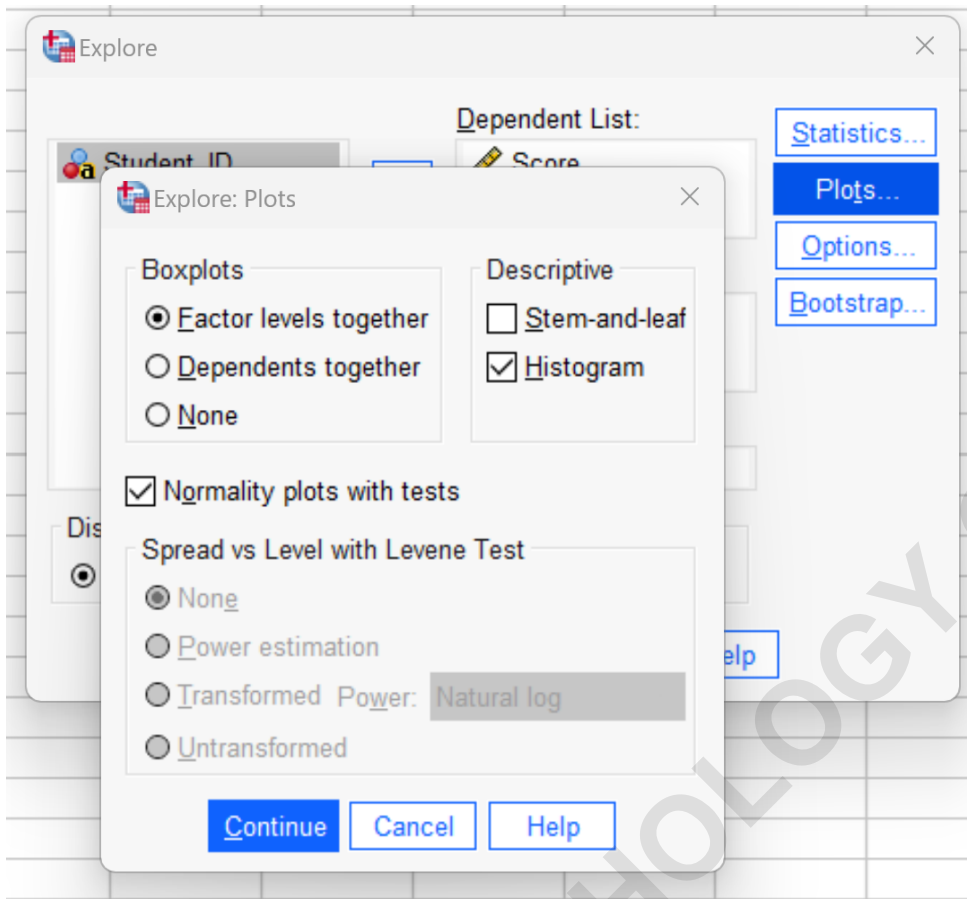


In the new window that appears, drag the **Score** variable into the **Dependent List** panel:



Then click the **Plots** button to customize the output visuals and tests.

Check the box next to **Histogram** and **Normality plots with tests**:



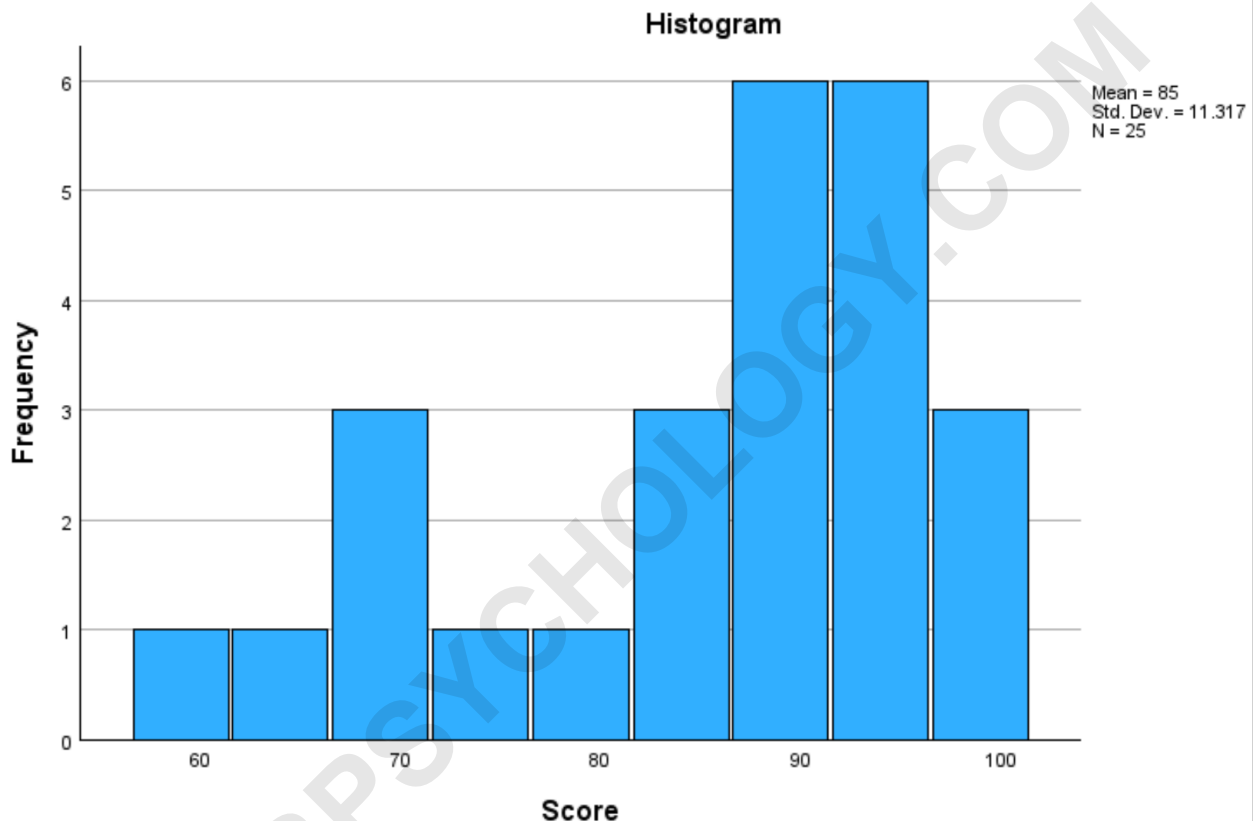
Then click **Continue** to save the plot settings. Then click **OK** in the main Explore dialog to execute the command.

The following output will be generated, displaying the numerical results of the Kolmogorov-Smirnov test along with a histogram and Q-Q plots to help us visualize the distribution of exam scores:

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

a. Lilliefors Significance Correction

Score



From the output in the "Tests of Normality" table, we can see that the p-value of the K-S test (Asymp. Sig. column) is **.036**.

Recall that the K-S test uses the following hypotheses:

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

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

Since the p-value of the test (0.036) is less than the typical significance level of .05, we have sufficient statistical evidence to reject the null hypothesis.

We would therefore conclude that the distribution of exam scores **is not** normally distributed, and thus non-parametric testing or data transformation should be considered.

The histogram generated by the Explore procedure visually agrees with this conclusion.

The distribution of exam scores in the histogram do not follow the expected symmetrical "bell curve" that is typical of a normal distribution, showing distinct deviation from normality.

The following tutorials explain how to perform other common operations in SPSS:

ARABPSYCHOLOGY.COM