

Learning Skewness and Kurtosis: A Practical Guide with SPSS

Authored by
Mohammed loot

November 12, 2025

RECOMMENDED CITATION

Mohammed loot (2025). *Learning Skewness and Kurtosis: A Practical Guide with SPSS*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=18250>

The Foundation of Data Analysis: Understanding Distribution Shape

In the expansive field of [statistics](#), a critical prerequisite for any meaningful analysis is a deep understanding of the underlying shape of the data distribution. Before applying complex inferential tests or fitting predictive models, researchers must move beyond basic summary metrics like the mean and standard deviation to characterize precisely how data points are spread. This characterization is primarily achieved through two indispensable descriptive statistics: [skewness](#) and [kurtosis](#). These measures provide quantitative insight into the symmetry and the tail characteristics of a distribution, respectively, offering a necessary check against the fundamental assumptions of many parametric procedures, particularly the assumption of normality. Failing to evaluate these metrics can lead to misinterpretation of results, rendering subsequent statistical inferences unreliable.

While initial exploratory data analysis (EDA) often involves visual tools like histograms or box plots, these graphical representations offer only a qualitative assessment. **Skewness** and **kurtosis** convert these visual characteristics into precise, numerical values, allowing for objective comparison and formal assessment. Skewness quantifies the extent and direction of asymmetry--determining if the data is balanced or if it is stretched out by a long tail on one side. This is vital because a highly skewed dataset suggests that the central tendency (the mean) is being unduly influenced by extreme values, often necessitating data transformation or the selection of non-parametric alternatives that are robust to distributional irregularities.

Complementing the measure of asymmetry is **kurtosis**, which assesses the "tailedness" and peakedness of the distribution relative to a standard benchmark, typically the [normal distribution](#). Kurtosis essentially measures the concentration of data in the center and the propensity for extreme values, often referred to as [outliers](#), to appear in the tails. By analyzing both skewness and kurtosis together, analysts gain a comprehensive picture of the data's behavior, identifying potential issues such as heavy concentrations of values near the mean or, conversely, a surprisingly high frequency of extreme observations. These quantitative shape metrics are foundational tools that guide the entire process of statistical investigation.

Deconstructing Skewness: Quantifying Asymmetry

Skewness serves as the primary metric for assessing the degree to which a data distribution deviates from a state of perfect symmetry. A truly symmetrical distribution, where the left side is a mirror image of the right side, will theoretically yield a skewness value of zero. However, real-world data rarely achieves this ideal symmetry, making the sign and magnitude of the skewness value critical for proper interpretation. The numerical result of the skewness calculation directly indicates the direction in which the distribution's tail extends, providing immediate clues about the data's concentration relative to its center. The magnitude of this value helps determine if the deviation

from symmetry is statistically significant enough to warrant corrective action.

The interpretation of the skewness value is divided into three distinct categories, each carrying significant implications for how the three measures of central tendency--the mean, median, and mode--are positioned within the dataset. A noticeable skew, whether positive or negative, signals that the majority of observations are clustered toward one end of the scale, while a smaller subset of extreme values exerts a disproportionate gravitational pull on the mean. This phenomenon causes the mean to diverge from the median, which is a more robust indicator of the data's true center in the presence of asymmetry.

A **Negative Skew** (or Left-Skewed Distribution) is characterized by a tail that stretches toward the left side, encompassing the lower, or more negative, values. In such a distribution, the bulk of the observations are heavily concentrated on the right (higher values). Consequently, the mean is typically pulled toward the left and is less than the median. A classic example is a very easy exam, where most students score high marks, and only a few low scores drag the overall average down.

A **Positive Skew** (or Right-Skewed Distribution) occurs when the tail extends toward the right side, corresponding to higher, or more positive, values. Here, the majority of the data is clustered on the left (lower values). The presence of large, outlying observations on the right causes the mean to be greater than the median. Income distribution is a frequently cited instance of positive skew, where the vast majority of individuals earn lower incomes, but a small number of extremely high earners significantly inflate the overall mean income.

A value of **Zero Skewness** implies that the distribution is perfectly symmetrical. In this rare and ideal scenario, such as an exact [normal distribution](#), the mean, median, and mode coincide at the same central point, indicating an even balance of data around the center.

Exploring Kurtosis: Assessing Tail Heaviness and Risk

While skewness addresses horizontal asymmetry, **kurtosis** provides a measure of the vertical characteristics of the distribution, specifically focusing on the concentration of data in the peak and the thickness, or "heaviness," of the tails. It quantifies how much a distribution's tails differ from the standard thickness found in the [normal distribution](#). In modern statistical software like SPSS, kurtosis is typically reported as **excess kurtosis**, which is calculated relative to the normal distribution, assigned a reference value of 0. A high kurtosis value suggests a higher probability of observing extreme events or [outliers](#) compared to what the normal distribution would predict, which is a critical measure in fields relying on risk modeling.

The standard normal distribution is the benchmark, holding an excess kurtosis value of 0. Deviations from this zero baseline categorize distributions into three types, each revealing crucial information about the spread and propensity for extreme values. Understanding kurtosis is vital in

disciplines such as finance and quality control, where the likelihood and impact of extreme deviations (often called "fat tails") carry substantial operational or financial consequences. A distribution with significantly positive kurtosis indicates that, while most values may cluster very tightly around the mean (resulting in a high peak), the tails are thicker, meaning that the occurrence of extreme observations is more probable than expected under standard assumptions.

Mesokurtic Distributions: These distributions exhibit an excess kurtosis value of 0. They possess tail thickness and peakedness characteristics similar to the standard normal distribution, serving as the neutral baseline against which all other distributions are compared.

Platykurtic Distributions: If a distribution has an excess kurtosis less than 0 (i.e., negative), it is termed platykurtic. These distributions are characterized by lighter tails and are often flatter in the center. They tend to produce fewer and less extreme [outliers](#) than the normal distribution, indicating a more uniform and broader spread of data across the variable's range.

Leptokurtic Distributions: If a distribution has an excess kurtosis greater than 0 (i.e., positive), it is leptokurtic. This signifies a distribution with heavy, or "fat," tails and a high, sharp peak in the center. Leptokurtic distributions imply a higher frequency of [outliers](#) compared to the normal model. Although many values concentrate tightly, the probability of observing very extreme scores is elevated.

With these foundational definitions established, the next section will transition into a practical, step-by-step demonstration of how to accurately calculate and interpret both skewness and kurtosis for a typical dataset using the industry-standard statistical software package, [SPSS](#) (Statistical Package for the Social Sciences). This practical application will solidify the theoretical concepts discussed.

Step-by-Step Guide: Calculating Skewness and Kurtosis in SPSS

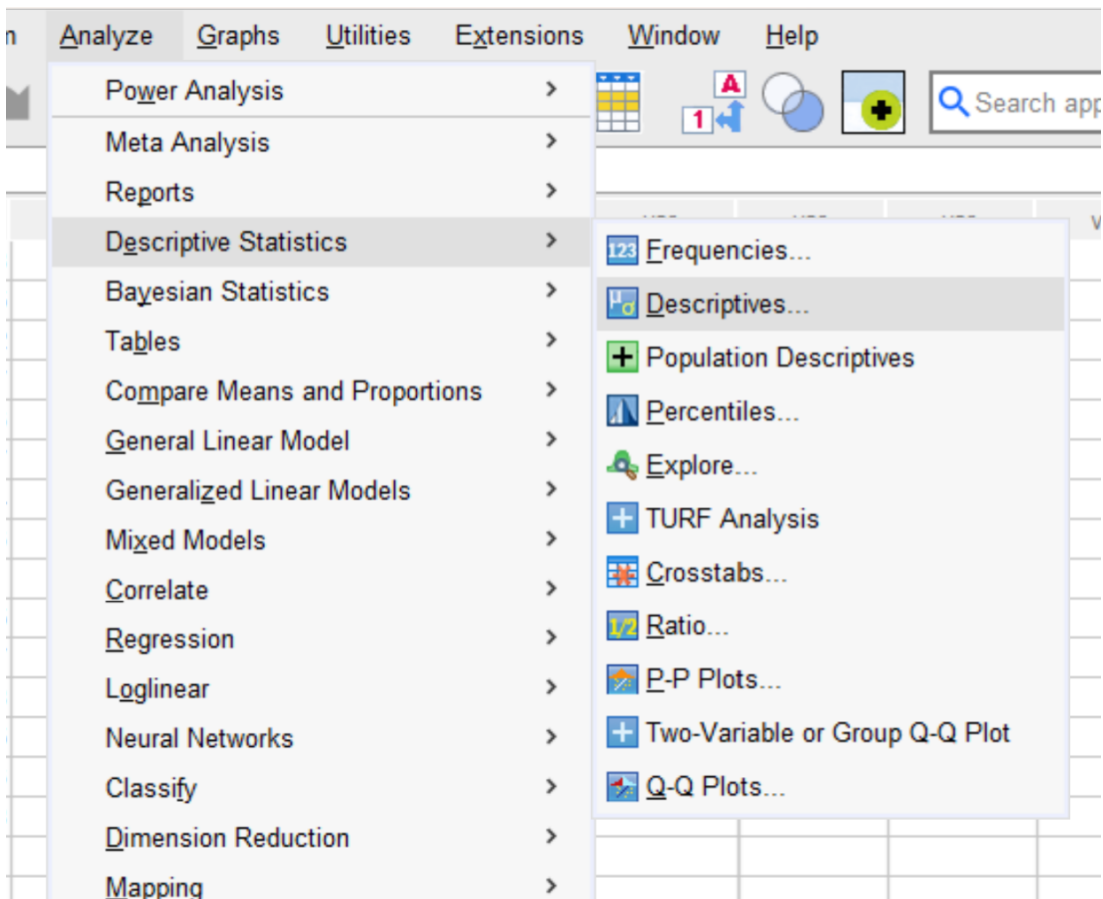
To demonstrate the calculation procedure effectively, we will utilize a hypothetical dataset within the [SPSS](#) environment that contains the final exam scores received by a group of students. This practical example, utilizing a common educational variable, allows us to observe firsthand how the software efficiently generates these essential descriptive statistics, which are necessary for completely characterizing the distribution of student performance.

For the purpose of this demonstration, assume we have the following data loaded into the SPSS Data View interface, with the variable clearly labeled as "Exam_Score." The goal is to determine the quantitative shape of the distribution of these scores, assessing how far they deviate from the ideal [probability distribution](#).

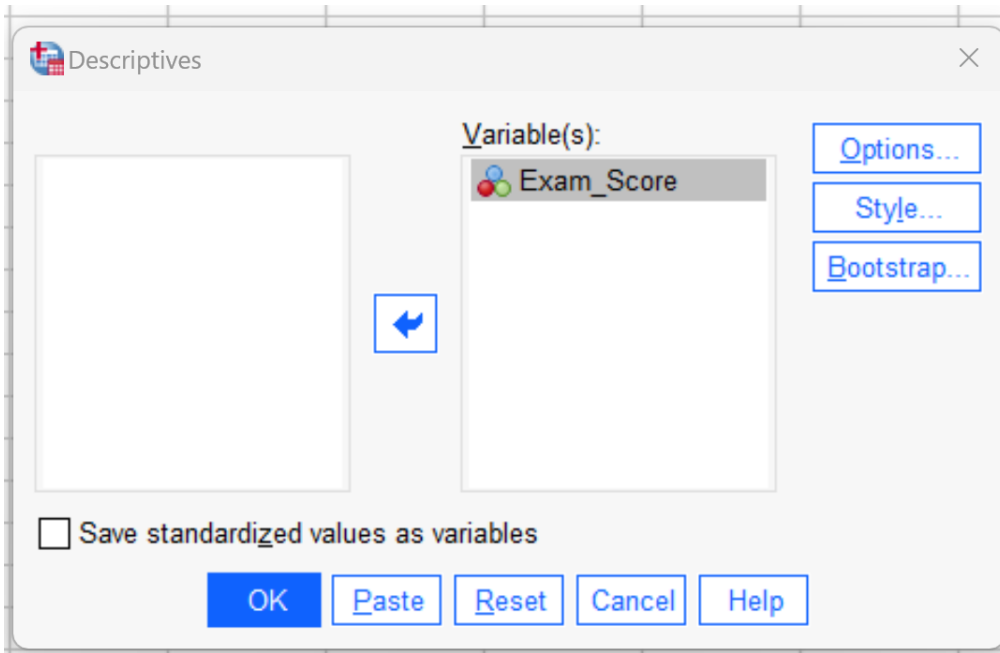
	Student_ID	Exam_Score	var	var	
1	1	88			
2	2	95			
3	3	92			
4	4	97			
5	5	96			
6	6	97			
7	7	94			
8	8	86			
9	9	91			
10	10	95			
11	11	97			
12	12	88			
13	13	85			
14	14	76			
15	15	68			
16					
17					
18					
19					

The procedure for calculating skewness and kurtosis for this variable is initiated through the standard SPSS menu bar, adhering to the conventional path for requesting descriptive analysis. This structured method ensures that the calculations are executed according to the rigorous statistical algorithms embedded within the software package, minimizing human error and providing reliable results.

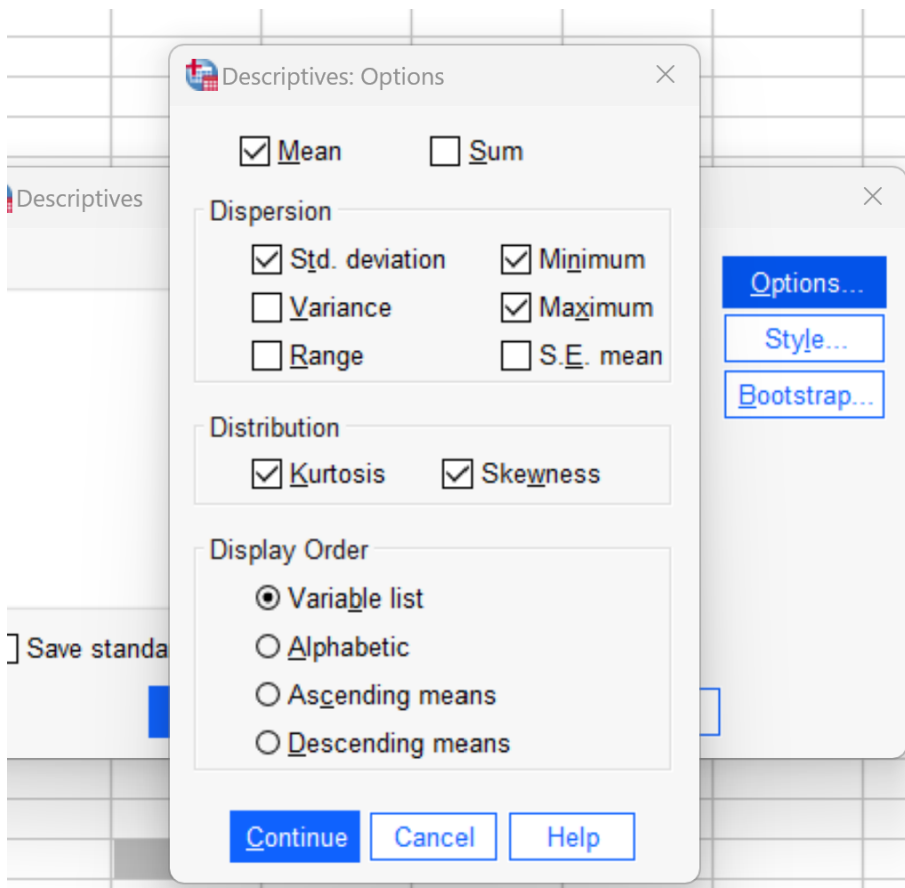
To begin the process, click the **Analyze** tab located in the top menu bar. Next, hover your cursor over the **Descriptive Statistics** option within the dropdown menu, and finally, click **Descriptives**. This crucial sequence opens the primary dialogue box, which is the control center where variable selection and the configuration of specific statistical options must be performed.



Within the Descriptive Statistics dialogue box that appears, you must carefully specify the variable targeted for analysis. Locate the variable named **Exam_Score** in the list displayed on the left and move it into the **Variables** panel on the right. This mandatory step informs **SPSS** precisely which column of raw data to process and summarize.



The next critical step involves instructing SPSS to include the specific measurements of distribution shape, as these are often excluded by default in the initial setup. Click the **Options** button situated on the right side of the dialogue box. A new sub-window will instantly pop up, presenting a list of available descriptive statistics. Within this window, ensure that the checkboxes corresponding to **Kurtosis** and **Skewness** are definitively checked. This configuration is essential for compelling SPSS to generate the required output table containing these two specific shape metrics.



After the options have been correctly configured, click **Continue** to close the Options sub-window and return to the main Descriptive Statistics box. Finally, click **OK** to execute the command. SPSS will rapidly process the request and display the resulting output table, containing the calculated descriptive statistics, in a separate, dedicated viewer window, ready for interpretation.

Interpreting SPSS Output and Visualizing the Distribution

Once the [SPSS](#) command is executed, the software immediately produces the output, which provides a concise summary of the descriptive metrics for the "Exam_Score" variable, including the newly calculated values for skewness and kurtosis. Analyzing this quantitative output is the moment of truth, allowing us to numerically characterize the distribution of student performance and objectively determine if the scores deviate significantly from the theoretical assumptions of a normal distribution.

The following table represents the structured output generated by SPSS, displaying the descriptive statistics for the variable:

→ **Descriptives**

Descriptive Statistics									
	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Skewness		Kurtosis	
						Statistic	Std. Error	Statistic	Std. Error
Exam_Score	15	68	97	89.67	8.372	-1.551	.580	2.230	1.121
Valid N (listwise)	15								

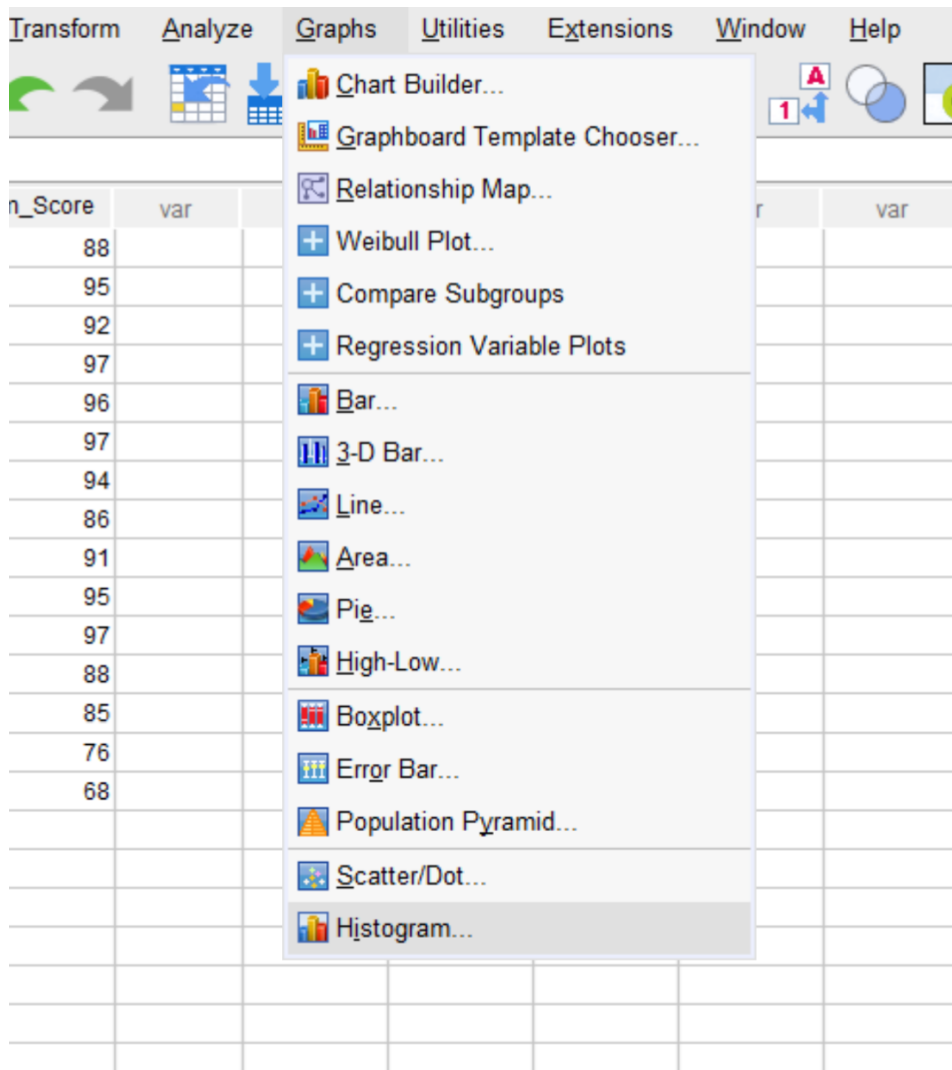
Based on the results presented in the table, we can extract and interpret the calculated values for the shape characteristics of the exam score distribution with high precision:

The calculated **Skewness** value is **-1.551**. Since this value is substantially negative (well below zero), it numerically confirms that the distribution is significantly **left-skewed** (negatively skewed). This interpretation means that the bulk of the scores are concentrated at the higher end, indicating that most students performed very well, while a comparatively long tail extends toward the lower scores, suggesting few students struggled.

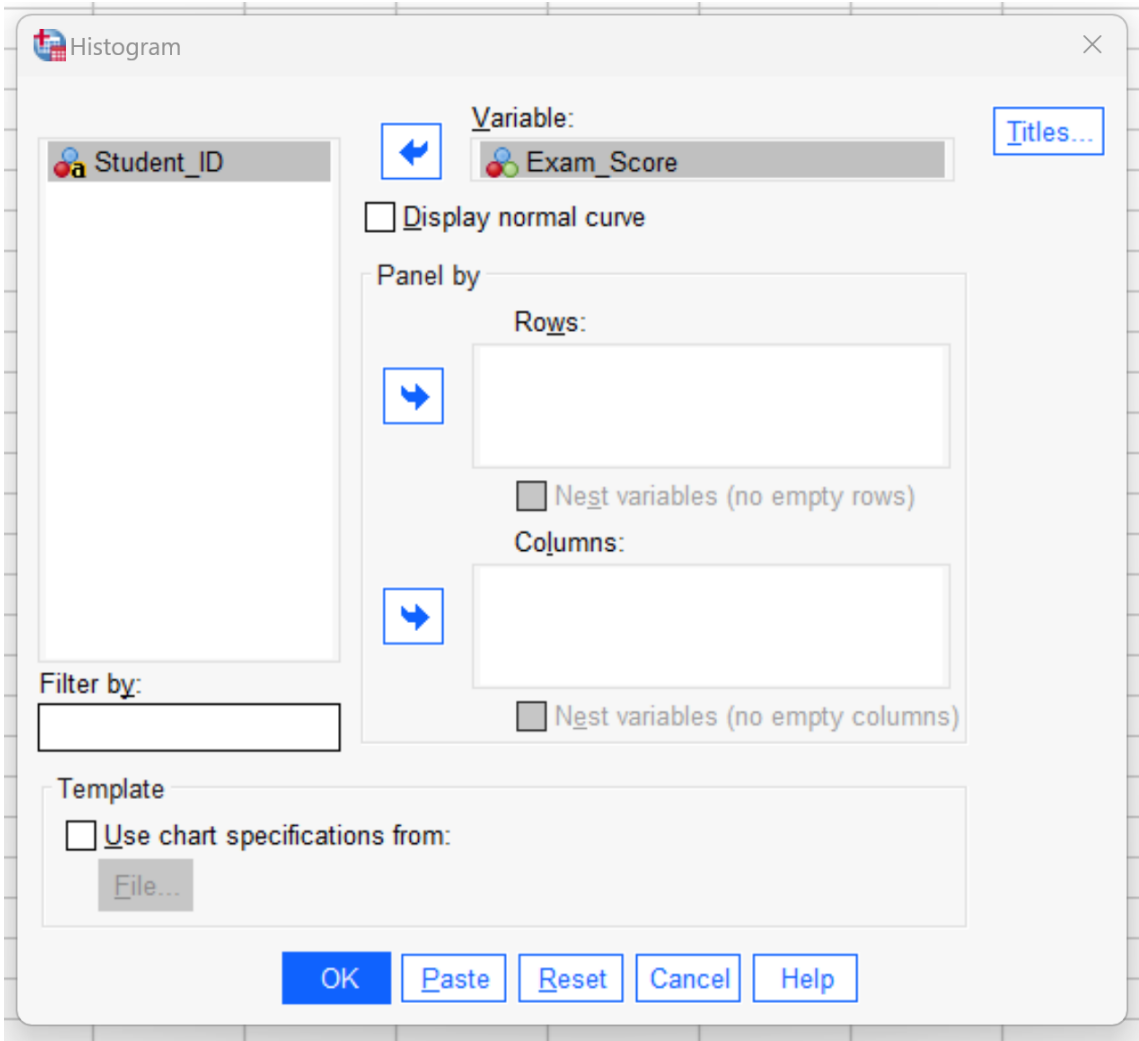
The calculated **Kurtosis** value is **2.230**. As this excess kurtosis value is positive and notably greater than zero, it signifies that the distribution is **leptokurtic**. A positive kurtosis implies that the distribution possesses heavier, or "fattened," tails compared to a normal distribution. This suggests that, relative to the main cluster of scores, there is a higher frequency of extremely high and extremely low scores (**outliers**), indicating high volatility or heterogeneity in performance.

While numerical metrics offer precision, they are often best understood and validated through visual inspection. To provide an intuitive confirmation of the calculated skewness and kurtosis, creating a **histogram** is highly recommended. A histogram effectively visualizes the frequency distribution of the data, allowing for an immediate, intuitive understanding of the data's density and asymmetry across different score ranges.

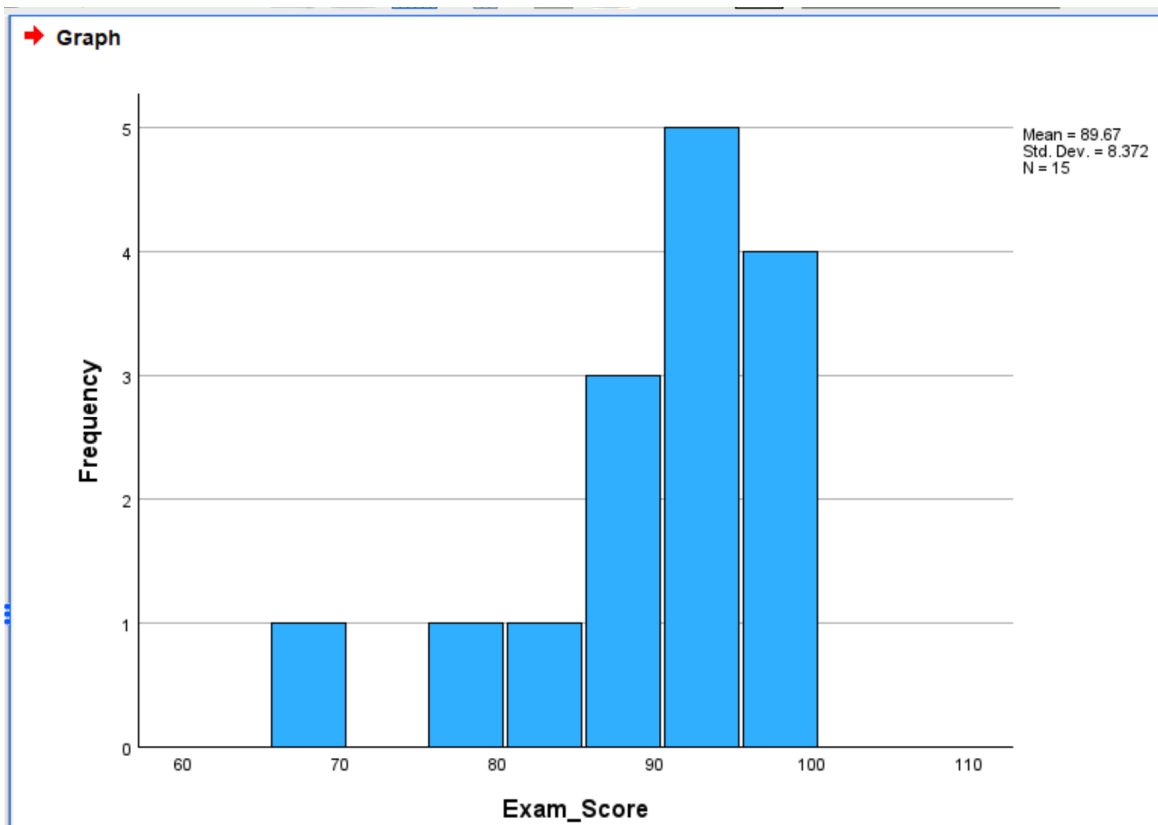
To generate this necessary visual aid in SPSS, navigate back to the menu bar: click the **Graphs** tab, then hover over **Legacy Dialogs** (if using an older version of SPSS), and then select **Histogram**. This sequence activates the plotting function specifically tailored for frequency distributions.



In the resulting Histogram dialogue box, drag the **Exam_Score** variable into the **Variable** panel, mirroring the step performed during the descriptives calculation. Although optional, checking the box to display the normal curve overlay can be helpful for visually comparing the data distribution against the theoretical ideal.



Finally, click **OK**, and the histogram will be promptly generated in the output viewer, providing a clear graphical depiction of the frequency distribution of the exam scores across various ranges. This visual confirmation constitutes the final, essential step in the exploratory data analysis process for robustly assessing distribution shape.



The generated visualization unequivocally confirms the numerical finding: the distribution is distinctly **left-skewed**, as the majority of the data mass is concentrated at the right side (higher scores), and the tail extends visibly toward the lower scores on the left. By systematically calculating the skewness and kurtosis metrics and rigorously complementing them with a histogram visualization, we achieve a robust, comprehensive, and validated understanding of the distribution of exam scores within this dataset--an understanding that forms the essential bedrock for any subsequent statistical modeling or inferential procedures. Further statistical tests, such as the [Shapiro-Wilk test](#), could then be employed to formally test the null hypothesis of normality, using the calculated shape metrics as initial evidence of non-normality.

Additional Resources for Advanced SPSS Analysis

Building upon this foundational knowledge of descriptive statistics and distribution shape, the following tutorials explain how to perform other common and critical tasks in [SPSS](#), enabling more complex data manipulation and inferential testing:

Understanding the calculation and significance of the [Standard Deviation](#) and variance.

Procedures for testing assumptions of normality using formal statistical criteria.

Implementing data transformations, such as logarithmic or square-root transformations, to address

significant skewness or kurtosis.

By mastering the calculation and interpretation of skewness and kurtosis along with creating a histogram, analysts are well-equipped to assess data quality and select the most appropriate inferential statistics for their research questions.