

How to Perform and Interpret a Chi-Square Test in SPSS

Authored by
stats writer

January 24, 2026

RECOMMENDED CITATION

stats writer (2026). *How to Perform and Interpret a Chi-Square Test in SPSS*.
PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=127382>

The Chi-Square test (specifically the Pearson's Chi-Square Test of Independence) is a fundamental statistical method used extensively across social sciences, medicine, and business to determine if there is a statistically significant association between two or more categorical variables. This non-parametric test compares the observed frequencies in various categories to the frequencies that would be expected if the variables were entirely independent. The application of this test within SPSS (Statistical Package for the Social Sciences) automates complex calculations, providing researchers with a clear pathway to data analysis.

When running a Chi-Square analysis in SPSS, the output generates several key metrics crucial for interpretation. These metrics typically include the calculated **Chi-Square value** (which quantifies the difference between observed and expected frequencies), the **degrees of freedom** (df), and the critical **p-value** (Asymptotic Significance). These resulting values are essential for hypothesis testing, allowing the researcher to evaluate the evidence against the null hypothesis (H₀), which posits that there is no relationship or association between the variables under study.

The decision threshold hinges on the comparison between the computed **p-value** and a pre-determined level of significance, usually 0.05. If the p-value is less than the significance level, the result is deemed statistically significant, leading to the rejection of the null hypothesis and indicating a meaningful association between the categorical variables. Conversely, a non-significant p-value suggests that any observed difference is likely due to random chance, and we fail to reject the null hypothesis. Mastering the interpretation of the Chi-Square test output in SPSS is an indispensable skill for rigorous quantitative analysis.

Interpreting the Chi-Square Test of Independence Results in SPSS

Understanding the Purpose of the Chi-Square Test of Independence

The core objective of the Chi-Square Test of Independence is to assess whether the distribution of one categorical variable is dependent on the distribution of another categorical variable. Essentially, we are asking: are the classifications across the rows systematically related to the classifications across the columns? This test is particularly useful when dealing with nominal or ordinal data that has been categorized into distinct groups, such as gender, preference, or outcome (yes/no).

To perform this analysis correctly, the data must satisfy several underlying assumptions. A primary assumption is that the observations must be independent--meaning that the inclusion or classification of one subject should not influence the classification of another. Furthermore, the test requires the sample size to be sufficiently large, typically ensuring that the **expected count** in each

cell of the contingency table is greater than five for at least 80% of the cells, and that no cell has an expected count of zero. Violating these assumptions can lead to unreliable p-values and inaccurate conclusions regarding the association.

The following detailed example demonstrates the practical execution and interpretation of the Chi-Square Test of Independence using the SPSS statistical software environment. We will walk through the data entry, the selection of appropriate analysis options, and the detailed interpretation of the three main output tables generated by SPSS.

Case Study: Associating Gender and Political Party Preference

To illustrate the process, consider a research question focused on political science: Is there an association between a voter's **gender** and their **political party preference**? This inquiry involves two distinct categorical variables. Gender is categorized (e.g., Male, Female), and party preference is also categorized (e.g., Democrat, Republican, Independent). The objective is to determine if men and women distribute themselves differently across these political categories.

We begin by collecting data using a systematic sampling method, such as a simple random sample, ensuring the sample is representative of the target population of voters. For this example, let us assume we surveyed a sample size of 50 registered voters. The collected data is then meticulously entered into the SPSS Statistics data editor, with separate columns designated for 'Gender' and 'Party Preference'. Proper variable coding (e.g., 1=Male, 2=Female, 1=Democrat, 2=Republican, 3=Independent) is critical for accurate analysis.

The subsequent screenshot visually demonstrates how this raw data, reflecting the survey results from our 50 voters, is structured and input into the SPSS environment, setting the stage for the formal statistical test. Note the clarity and organization necessary in the data view for successful processing.

	VoterID	Gender	Party	var	var	
1	1	Male	Rep			
2	2	Male	Dem			
3	3	Female	Ind			
4	4	Male	Ind			
5	5	Female	Rep			
6	6	Male	Rep			
7	7	Male	Dem			
8	8	Male	Dem			
9	9	Female	Dem			
10	10	Male	Rep			
11	11	Male	Rep			
12	12	Female	Dem			
13	13	Female	Ind			
14	14	Male	Rep			
15	15	Male	Rep			
16	16	Female	Dem			
17	17	Male	Ind			
18	18	Female	Ind			
19	19	Female	Rep			
20	20	Male	Dem			
21	21	Male	Dem			
22	22	Female	Ind			

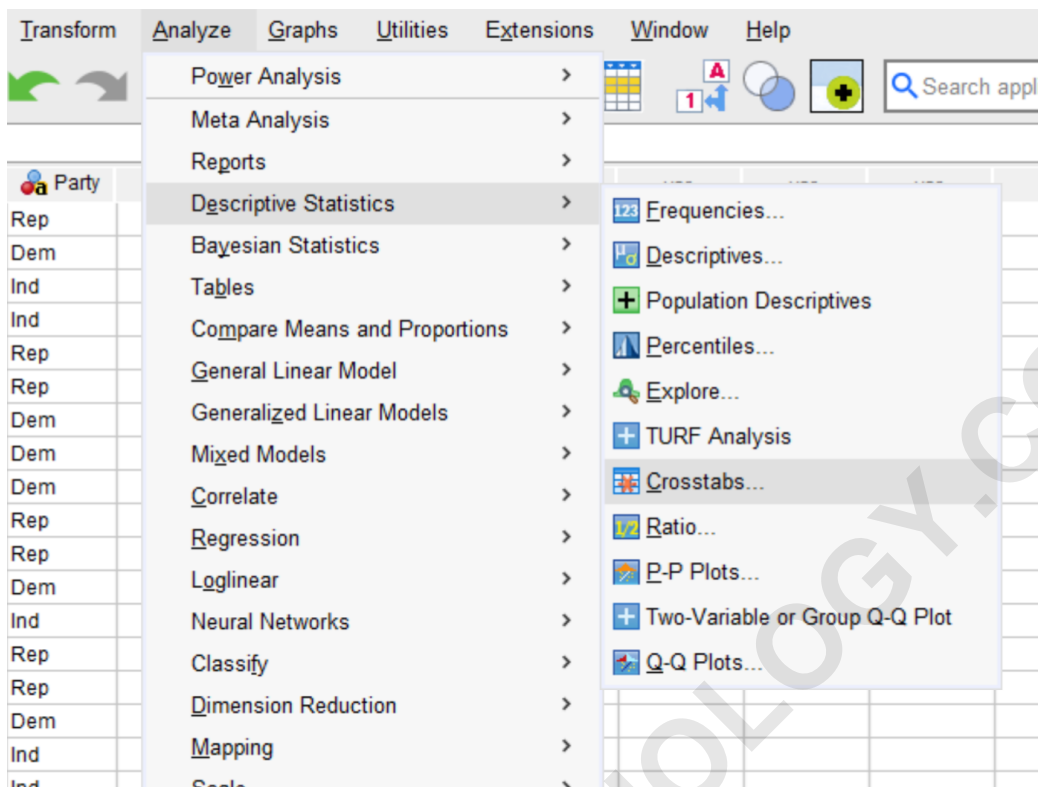
Setting Up the Analysis in SPSS: A Step-by-Step Guide

Executing the Chi-Square Test of Independence in SPSS is accomplished via the Crosstabs procedure, which efficiently constructs the contingency table and performs the necessary statistical calculations. The procedure begins by navigating the main menu bar. You must click the **Analyze** tab, hover over **Descriptive Statistics**, and finally select the **Crosstabs** option. This sequence initiates the primary dialogue box for setting up the analysis.

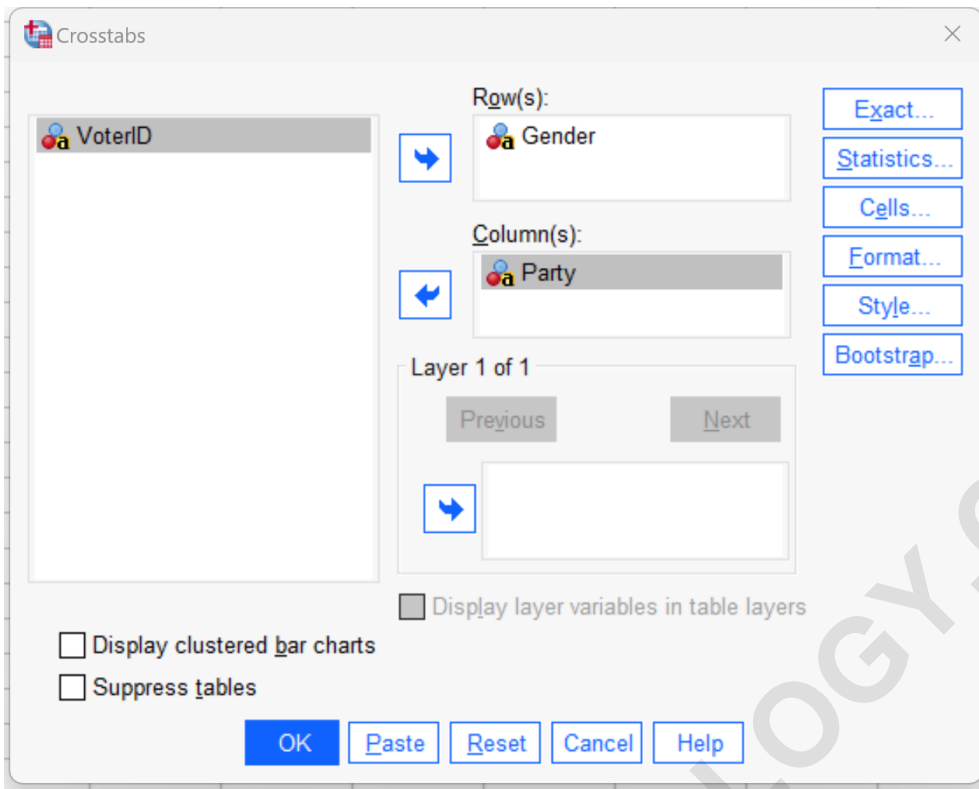
Within the Crosstabs dialog box, the researcher must carefully assign the variables to the appropriate axes of the resulting table. In line with conventional practice, the independent variable (in this case, **Gender**) is typically dragged into the **Rows** panel. Conversely, the dependent variable (**Party Preference**) is dragged into the **Columns** panel. This placement determines how the data will be aggregated and displayed in the resultant contingency table, facilitating clear interpretation of potential row-by-column associations.

The following image illustrates the precise navigational steps required to access the Crosstabs function within the SPSS interface, ensuring the user is correctly prepared to define the variables

for analysis.



The subsequent image confirms the correct placement of the Gender variable in the Rows and the Party variable in the Columns, a vital setup step before proceeding to request the specific output statistics.

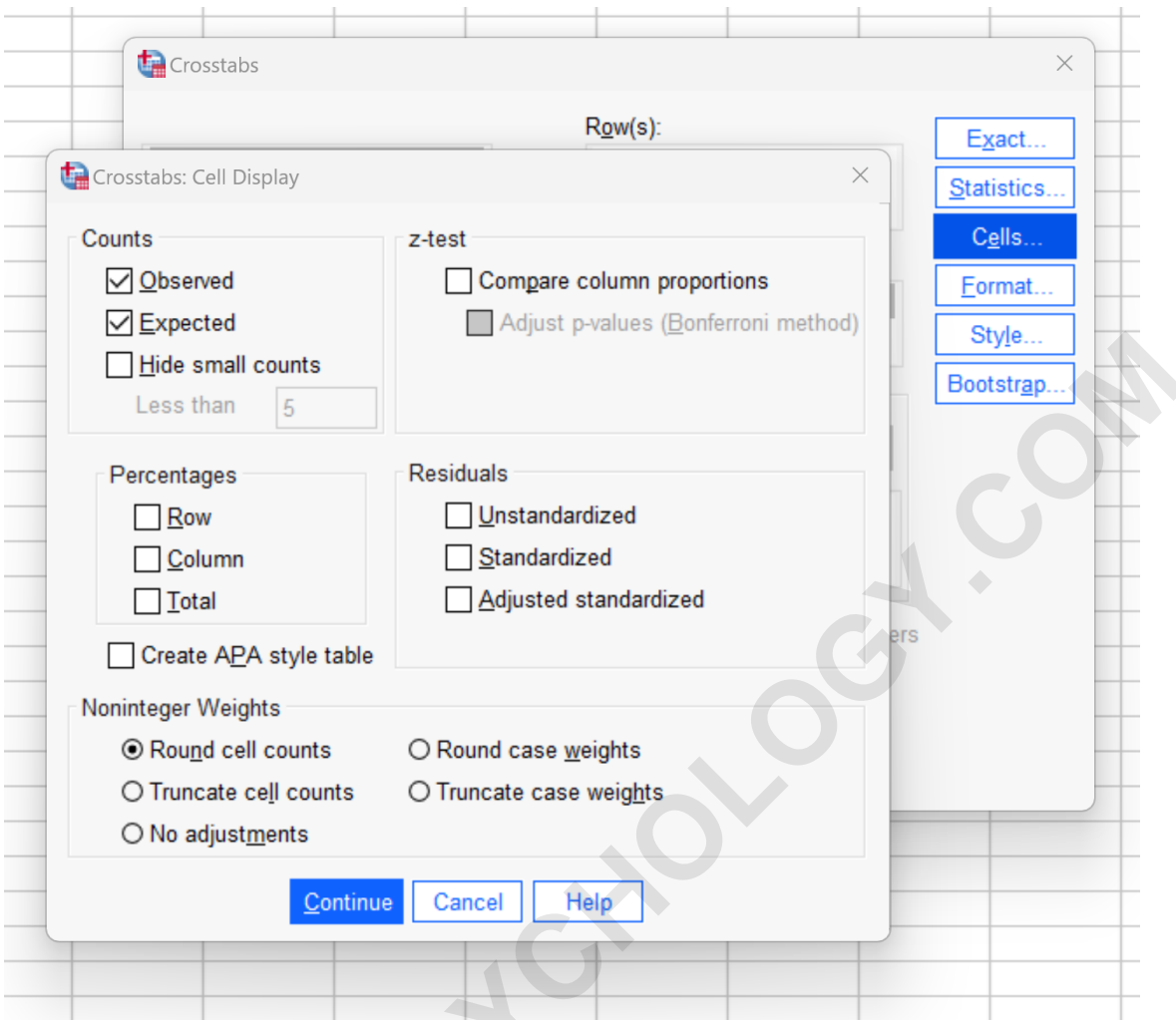


Requesting Essential Output: Observed and Expected Counts

Crucial to interpreting the Chi-Square test is the ability to compare the actual data distribution (the **Observed counts**) with the theoretical distribution (the **Expected counts**). To ensure SPSS provides this necessary comparison, the researcher must access the **Cells** sub-dialog box. This step is mandatory for checking the assumptions of the test and providing context to the final Chi-Square statistic.

Upon clicking the **Cells** button, a new window appears where you must check the boxes next to both **Observed** and **Expected** under the Counts section. The Observed counts represent the frequency we actually measured in our sample, while the Expected counts represent the frequencies we would anticipate seeing in each cell if the null hypothesis of independence were true. After selecting these options, click **Continue** to return to the main Crosstabs dialogue.

The accompanying visual guidance confirms the selection process for generating both observed and expected frequency tables, highlighting how essential this step is for a complete diagnostic analysis.

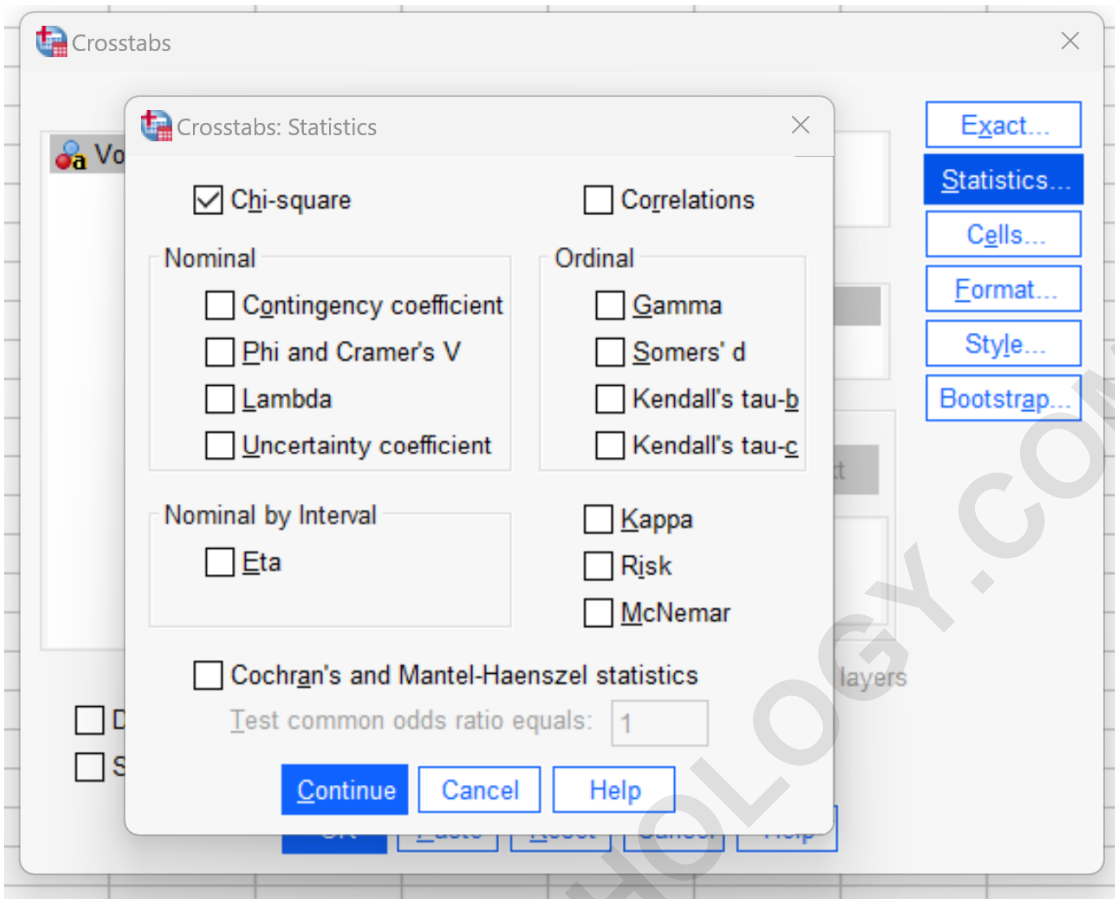


Generating the Chi-Square Statistic and Interpretation

The final step in setting up the analysis is instructing SPSS to compute the actual Chi-Square statistic. This is done by clicking the **Statistics** button within the Crosstabs dialogue. In the subsequent Statistics window, the researcher must check the box specifically labeled **Chi-square**. This selection tells SPSS to calculate Pearson's Chi-Square, the Likelihood Ratio, and other related tests.

Selecting the Chi-square option ensures that the primary measure of association based on the discrepancies between the observed and expected frequencies is calculated and presented in the output. After confirming this selection, click **Continue**, and then execute the analysis by clicking **OK** in the main Crosstabs window. The output viewer window will then populate with the results tables.

This image demonstrates the critical step of selecting the Chi-square test itself, confirming the final instruction needed before generating the results.



Once the analysis is executed, SPSS generates the comprehensive output tables, which include three major components: the Case Processing Summary, the Crosstabulation, and the Chi-Square Tests table. The collective presentation of these tables, as shown below, provides all the necessary information for hypothesis evaluation.

➔ **Crosstabs**

Case Processing Summary

	Valid		Cases Missing		Total	
	N	Percent	N	Percent	N	Percent
Gender * Party	50	100.0%	0	0.0%	50	100.0%

Gender * Party Crosstabulation

Gender			Party			Total
			Dem	Ind	Rep	
Female	Count		9	7	5	21
	Expected Count		8.0	6.3	6.7	21.0
Male	Count		10	8	11	29
	Expected Count		11.0	8.7	9.3	29.0
Total	Count		19	15	16	50
	Expected Count		19.0	15.0	16.0	50.0

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	1.118 ^a	2	.572
Likelihood Ratio	1.140	2	.566
N of Valid Cases	50		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 6.30.

Examining the Output: The Case Processing Summary and Crosstabulation

The first table generated is the **Case Processing Summary**. This table is vital for initial data quality assessment, as it displays the total number of observations analyzed and, crucially, how many cases were included (Valid) and how many were excluded due to missing data. This immediate check confirms that the analysis proceeded using the intended sample size. In our example, we observe **50** valid observations and **0** missing observations, verifying that all collected data points were included in the calculation.

The second key output is the **Crosstabulation** table (or contingency table). This is where the raw data comparison occurs. It systematically displays the intersection of the two categorical variables (Gender and Party Preference) and includes both the **Observed Count** and the **Expected Count** for every cell. The Observed Count tells us the actual frequency of voters falling into each specific

gender-party combination.

The **Expected Count** is a theoretical frequency--what we would statistically anticipate if Gender and Party Preference were completely independent of each other (i.e., if the null hypothesis were true). The calculation of expected counts involves multiplying the row total by the column total and dividing by the grand total sample size. The magnitude of the difference between the Observed and Expected counts directly influences the final Chi-Square test statistic. Large discrepancies suggest a strong association, while small discrepancies suggest independence.

Referencing external resources or statistical textbooks provides a deeper explanation of how to mechanically calculate these expected counts in a Chi-Square test, which is fundamental to understanding the mechanics of the test statistic.

Decoding the Core Results: The Chi-Square Tests Table

The most critical section for decision-making is the **Chi-Square Tests** table. This output contains the numerical results necessary to evaluate the null hypothesis. The focus is primarily on the row labeled **Pearson Chi-Square**, which provides three essential values: the Chi-Square statistic, the Degrees of Freedom (df), and the Asymptotic Significance (the p-value).

In our example, the calculated **Chi-Square test statistic** is **1.118**. This value is derived from summing the standardized, squared differences between the observed and expected frequencies across all cells. A larger Chi-Square value indicates greater discrepancy between what was observed and what was expected under independence, suggesting a stronger relationship between the variables.

Crucially, the corresponding two-sided **p-value** (labeled 'Asymp. Sig.') is calculated as **.572**. The degrees of freedom (df), which are calculated as $(\text{Number of Rows} - 1) * (\text{Number of Columns} - 1)$, reflect the number of values in the final calculation that are free to vary. For our 2x3 table (Gender x Party), $df = (2-1) * (3-1) = 2$. This p-value is the probability of observing a Chi-Square statistic as extreme as 1.118 if the null hypothesis of independence were truly correct.

Making the Final Decision: Rejecting or Failing to Reject the Null Hypothesis

Statistical inference requires comparing the calculated p-value against the predefined level of significance (α), conventionally set at 0.05. This comparison dictates the final conclusion regarding the relationship between the two categorical variables. The fundamental hypotheses for the Chi-Square Test of Independence are formally stated as:

H0 (Null Hypothesis): The two variables (Gender and Party Preference) are statistically independent; there is no association between them in the population.

HA (Alternative Hypothesis): The two variables are *not* independent; they are associated.

To make the decision, we apply the decision rule: If the **p-value ≤ 0.05** , we **reject H₀**; if the **p-value > 0.05** , we **fail to reject H₀**. In this specific scenario concerning gender and political party preference, the calculated p-value is **0.572**. Since 0.572 is significantly greater than the standard alpha level of 0.05, we must conclude that we **fail to reject the null hypothesis**.

The failure to reject the null hypothesis implies that we do not have sufficient statistical evidence, based on our sample data, to assert that there is a significant association between gender and political party preference in the population from which the sample was drawn. The observed differences in frequencies shown in the Crosstabulation are likely attributable to random sampling variation rather than a true relationship. Researchers should report the Chi-Square statistic ($\chi^2 = 1.118$), degrees of freedom ($df = 2$), and the p-value ($p = 0.572$) alongside their conclusion.

Conclusion and Practical Implications for Categorical Data Analysis

The Chi-Square Test of Independence remains one of the most widely used methods for analyzing frequency data and assessing relationships between categorical variables. While the result in our example indicated independence, a positive result ($p \leq 0.05$) would necessitate further exploration, often involving examining adjusted standardized residuals to pinpoint exactly which categories contribute most significantly to the overall association.

Understanding the full output generated by SPSS, from the initial Case Processing Summary to the crucial p-value in the Chi-Square Tests table, allows analysts to move beyond simple data description toward rigorous statistical inference. Careful attention to the calculation of expected counts is key, as violating the assumption of sufficient expected cell frequencies can render the Pearson Chi-Square result invalid, potentially requiring the use of alternatives like Fisher's Exact Test or the Likelihood Ratio Chi-Square.

For researchers seeking to expand their command of statistical software, the following tutorials explain how to perform other common tasks and tests within the SPSS environment, building upon the foundational knowledge gained from mastering the Chi-Square Test: