

## SPSS Help

Sorry, no pretty pictures here. You can find those in the Help Command of SPSS, books, and other internet sites. Included here are the basic steps in a very abbreviated style.

- **Help Command in SPSS**
  - I prefer the Contents option (under Topics). There is also a Search option (also under Topics). And they provide a Tutorial.
- **SPSS Book**
  - If you purchased the Student Version of SPSS, the book that came with the software is an excellent resource.
- **SPSS Help on the Internet**
  - Use any search engine to find web site (many at Universities) that offer SPSS Help. For example: <http://www.ats.ucla.edu/stat/spss/>, <http://my.ilstu.edu/~mshesso/SPSS/tutorial.html>.
- **Open Excel File in SPSS**
  - File—Open—Data. Change Files/Type to “Excel.” Find and select the Excel file. If the first row includes variable names, select “read variable names from the first row...” If the first row does not include variable names, de-select “read variable names from the first row...” Finally, save in SPSS as an SPSS file {SPSS (\*.sav)}.
- **Dichotomize a continuous variable**
  - You will be dividing a continuous variable (example: GDP in million \$) into just two categories: low (units below the median) and high (units above the median).
  - Find the median value in SPSS. Analyze, Descriptive Statistics, Frequencies—Turn off “Display frequency tables”—Select the variable—Statistics, Median, Continue, OK.
  - Recode the variable in SPSS. Transform, Recode, Into Different Variables. Select variable, Output Variable Name (“New Name”), Label (“New Variable Label”), Old and New Values, Range—Lowest through *put in median value*, Value = 1, Add; AND Range—*put in median value + increment* through Highest, Value – 2, Add. Continue. Change. OK.
- **Recode a variable**
  - Example: You want to record a discrete variable with 7 categories into just 2 categories with equal units in each category.
  - Generate the frequencies for the variable: Analyze, Descriptive Statistics, Select Variable, OK.

- Determine which categories need to be collapsed. Example: say category I had 2 values, category II had 3 values, category III had 2 values, category IV had 4 values, and category V had 3 values. You would want to collapse categories I, II, and III (to produce a single “low” category with 7 values) and categories IV and V (to produce a single “high” category also with 7 values).
  - Transform, Recode, Into Different Variable. Select this variable, Output Variable Name (“New Name”), Label (“New Variable Label”), Old and New Values, Range—1 through 3, Value = 1, Add; AND Range 4 through 5, Value = 2, Add. Continue. Change. OK.
- **Missing Value**
    - A missing value is a value you want to “ignore” in your analysis. Example: no response, no data, refuse to answer, do not know, etc.
    - Variable View, Column “Missing” and Row “this variable.” Select the small, shaded box in this cell. Discrete Missing Values—put each missing value in one of the three boxes. OK.
- **Variable Label**
    - Variable Labels allow you to add a lengthier variable name to an abbreviated variable name. Example: variable name is MEPCGNP1989 and you add a variable label that is “Military Expenditures as a Percent of GNP in 1989.:
    - Variable View, Column “Label” and Row “this variable.” Simply type the variable label in this cell.
- **Value Label**
    - Value Labels provide the English-language labels for the numeric values. For example, the variable Gender will have values of 1 and 2. Value Labels will provide the information about which value is male and which is female.
    - Variable View, Column “Values” and Row “this variable.” Select the small, shaded box in this cell. Value = *whichever value* and Value Label = *the value label* (example, Value = 1 and Value Label = male). Add. Then repeat for all Value Labels. OK.
- **Frequencies**
    - Analyze, Descriptive Statistics, Frequencies
    - Select variable and move it to the “variable” box using the right arrow.
    - For statistics (mean, median, mode), select statistics.
- **Cross-Tabs and Chi-Square**
    - Chi-Square is a test of significance for cross-tabs. Use smallest value under “Asymp. Sig. (2-sided).” The criteria is 0.05 (5 percent probability that the null hypothesis is true).
    - Analyze, Descriptive Statistics, Crosstabs.
    - Select your independent variable and move it to the “column” box.
    - Select your dependent variable and move it to the “row” box.

- If you want just Cross-Tabs (or Cross-Tabs and Chi-Square) do not select “suppress tables.” If you only want Chi-Square, do select “suppress tables.”
  - For the Chi-Square statistic: select Statistic—select Chi-Square—Continue.
  - For Cross-Tabs: select Cells—turn-off “Observed” and Select “Column Percentages”—Continue.
- OK.
- **Lambda**
  - Lambda is a measure of association for nominal level data (only categories). The range is 0 to 1. 0 would be a perfect “non-relationship” while 1 would be a perfect relationship. It must be used in comparison.
  - Analyze—Descriptive Statistics—Crosstabs.
  - Move the X variable to the “column” box and the Y variable to the “row” box.
  - Suppress table
  - Statistics—Lambda (under Nominal).
  - Continue.
  - OK.
  - Output—Use Lambda, Symmetric. The “Value” column shows the value for Lambda.
- **Gamma**
  - Gamma is a measure of association for ordinal level data (rank-ordered; > or <). The range is -1 to +1. The sign only indicates inverse versus directly proportional relationships. 0 would be a perfect “non-relationship” while an absolute value of 1 would be a perfect relationship. It must be used in comparison.
  - Very simply, identical to Lambda above; except after Statistics, select Gamma (under Ordinal).
- **Correlation**
  - The correlation coefficient is a measure of association for interval level data (know the intervals; + or -). The sign only indicates inverse versus directly proportional relationships. 0 would be a perfect “non-relationship” while an absolute value of 1 would be a perfect relationship. It must be used in comparison.
  - Analyze—Correlate—Bivariate.
  - Select your two variables and move them to the “Variables” box.
  - Leave all default selections “as-is” (Pearson, Two-tailed, and Flag significant correlations should be “turned-on”).
  - OK
  - Your output will be a 2x2 table, but only one cell is relevant. Use either cell that shows the relation between your two variables (there are two such identical cells). The first value in that cell is the correlation coefficient.
- **Partial Correlation**

- Essentially a “multi-variate” correlation coefficient—or a correlation between two variables holding a third variable constant.
  - Analyze—Correlate—Partial.
  - Two “original” variables in “Variables” box.
  - Third variable in “Controlling For” box.
- **Regression**
    - Produces the linear equation:  $Y = a + bX$ ; or  $Y = mX + b$
    - Analyze—Regression—Linear.
    - Put Y variable in the “Dependent” box.
    - Put X variable in the “Independent” box.
    - OK.
    - Output contains 4 tables.
    - The 4<sup>th</sup> (last) table contains the coefficients for the regression equation ( $Y=a+bX$ ).
    - The coefficients are in the column “Unstandardized Coefficients—B.”
    - The (Constant) row contains “a” and the independent variable row contains “b.”
- **Multiple Regression**
    - Essentially a multi-variable regression analysis. One dependent variable with two (or more) independent variables.
    - Analyze—Regression—Linear.
    - Put Y variable in the “Dependent” box.
    - Put two X variables in the “Independent(s) box.”
- **Split File**
    - Suppose you had 170 nations labeled by 7 different geographic areas. And this was the variable Area. And you wanted the mean value of a GNP variable for each of these areas. You would use the Split File command first.
    - Data, Split File, Organize Output by Groups
    - Groups based on: put the variable (in this case, “Area,” that contains the 7 groups or areas.
    - Select your analysis. For example, Mean. Analyze, Descriptive Statistics, Frequencies, turn off “Display frequency tables,” Statistics, Mean, OK.
    - If you want to group by another variable, repeat the process with the other variable in the “Groups based on:” box.
    - When complete, turn off the Split File. Data, Split File, select “Analyze all cases, do not select groups.”