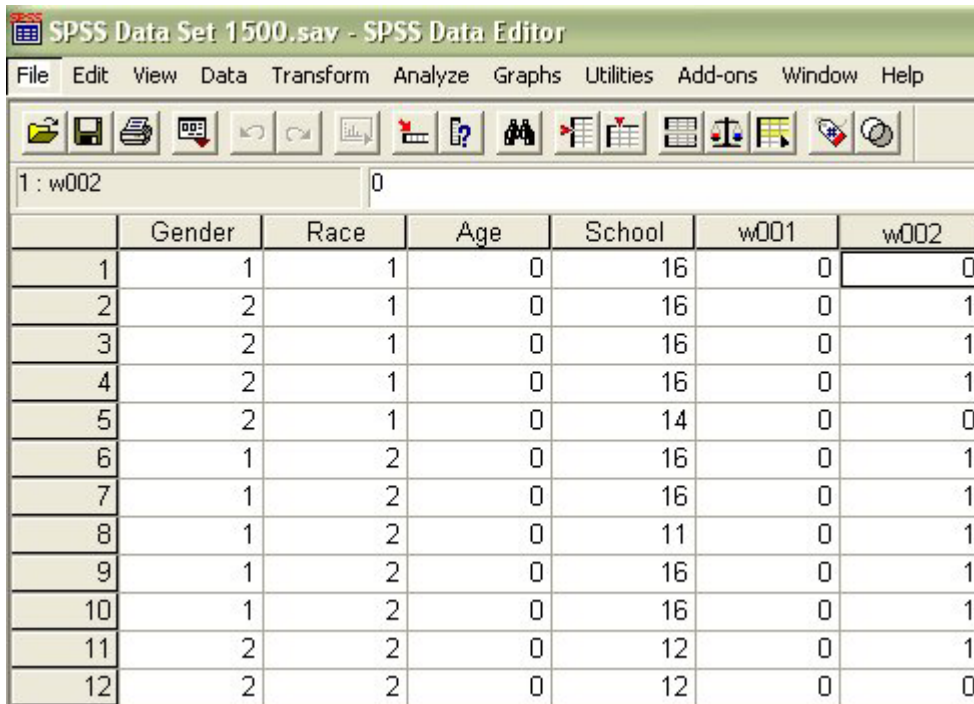


AN SPSS PRIMER

Prepared by Dennis C. Sweeney, Ph.D.

ENTERING DATA

SPSS Data Editor: (Data View)



The screenshot shows the SPSS Data Editor window for a file named 'SPSS Data Set 1500.sav'. The window title bar includes the file name and 'SPSS Data Editor'. Below the title bar is a menu bar with options: File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Add-ons, Window, and Help. Underneath the menu bar is a toolbar with various icons for file operations, editing, and analysis. The main area of the window displays a data table with the following columns: Gender, Race, Age, School, w001, and w002. The rows are numbered 1 through 12. The data values are as follows:

	Gender	Race	Age	School	w001	w002
1	1	1	0	16	0	0
2	2	1	0	16	0	1
3	2	1	0	16	0	1
4	2	1	0	16	0	1
5	2	1	0	14	0	0
6	1	2	0	16	0	1
7	1	2	0	16	0	1
8	1	2	0	11	0	1
9	1	2	0	16	0	1
10	1	2	0	16	0	1
11	2	2	0	12	0	1
12	2	2	0	12	0	0

This screen is similar to a spread sheet. It gives you access to all of the procedures for entering, transforming, printing, and analyzing your data.

OVERVIEW

- Each row contains the data for one person.
- Each column contains the data for one variable.
- Variable names are limited to eight characters
 - The names must start with a letter.
 - The names cannot contain a space.

FILE This menu contains all of the information for

- Saving the data
- Printing the data
- Opening a data file.
- Closing a data file.

DATA This menu lets you

- Create variables.
- Transform variables.
- Sort the data file.

ANALYZE This menu lets you do all of the statistical analyses.

GRAPH This menu has all of the graph functions.

FORMATTING VARIABLES

SPSS Data Editor: (Variable View)

	Name	Type	Width	Decimals	Label	Values	Missing
1	Gender	Numeric	1	0	Gender	{1, Male}...	0
2	Race	Numeric	1	0	Race	{1, Asian}...	0
3	Age	Numeric	3	0	Age	None	0
4	School	Numeric	6	0	Years of Scho	None	0
5	w001	Numeric	1	0	Ardent	None	-9999
6	w002	Numeric	1	0	Confident	None	-9999
7	w003	Numeric	1	0	Distant	None	-9999
8	w004	Numeric	1	0	Loyal	None	-9999

This screen lets you describe each variable found in the Data View. It is saved along with the data file. Each variable is represented by a row.

NAME This is the short variable name seen in the Variable View.

TYPE This identifies the variable as being one of the following

- Number
- Date
- Dollar
- String (Letter)

WIDTH This tells SPSS how many integers/letters wide to display the value. For numbers this includes the number of decimal places.

DECIMALS This tells SPSS how many decimals to display. It will save many more decimals in the Data File.

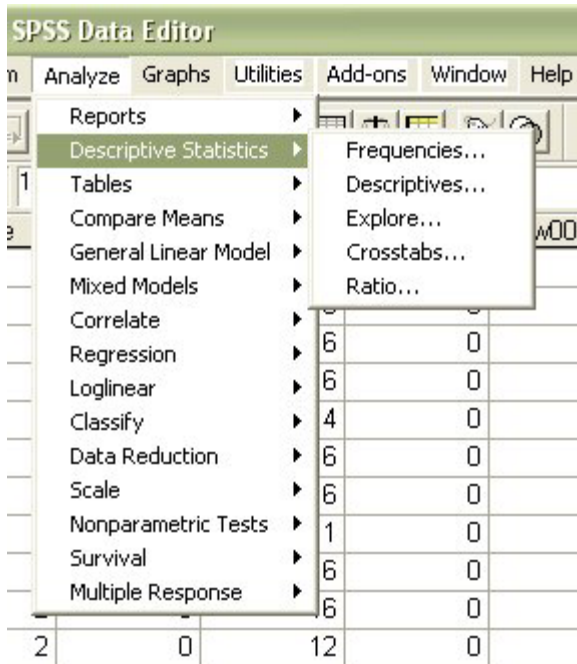
LABEL This is an extended label for the variable. This label can be almost any length.

VALUES This gives values for categorical each value. For example, Males are coded as 1 for Gender, Females are coded as 2.

MISSING Whenever a person has left something blank you should assign a missing value. Zero (0), unless you tell SPSS that this should be treated as missing, is treated as a number and will distort any statistical calculations.

DOING ANALYSES

SPSS Data Editor: Analyze> Descriptive Statistics



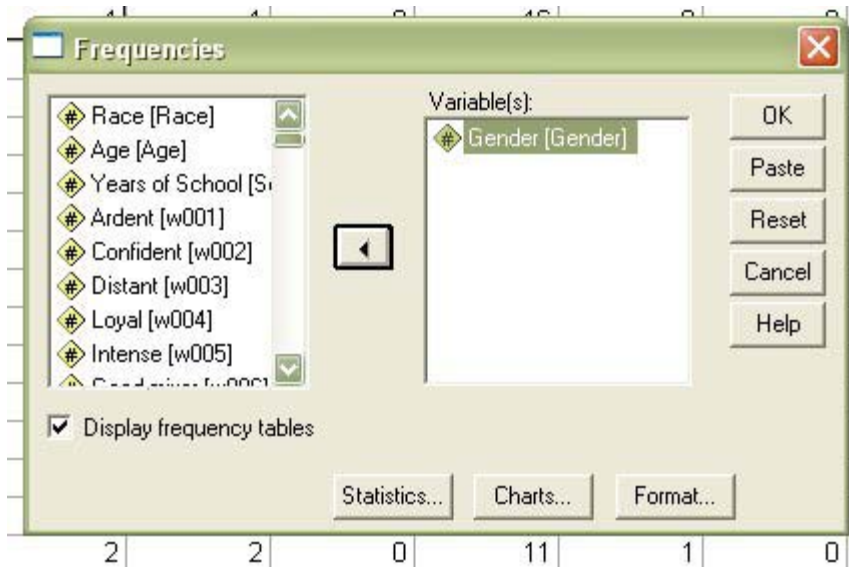
FREQUENCIES: This choice brings up a sub-menu for conducting frequency tables and graphs for the variables for which the frequencies are calculated. (See below)

DESCRIPTIVE This choice brings up a sub-menu for calculating descriptive statistics. It should be used for only interval and ratio data.

EXPLORE: This choice brings up a sub-menu for conducting a variety of simple inferential analyses including t-tests.

CROSSTABS: This choice brings up a sub-menu for creating crosstabulation tables and includes options for conducting chi-square analyses on the results.

SPSS Data Editor: Analyze > Descriptive Statistics>Frequencies



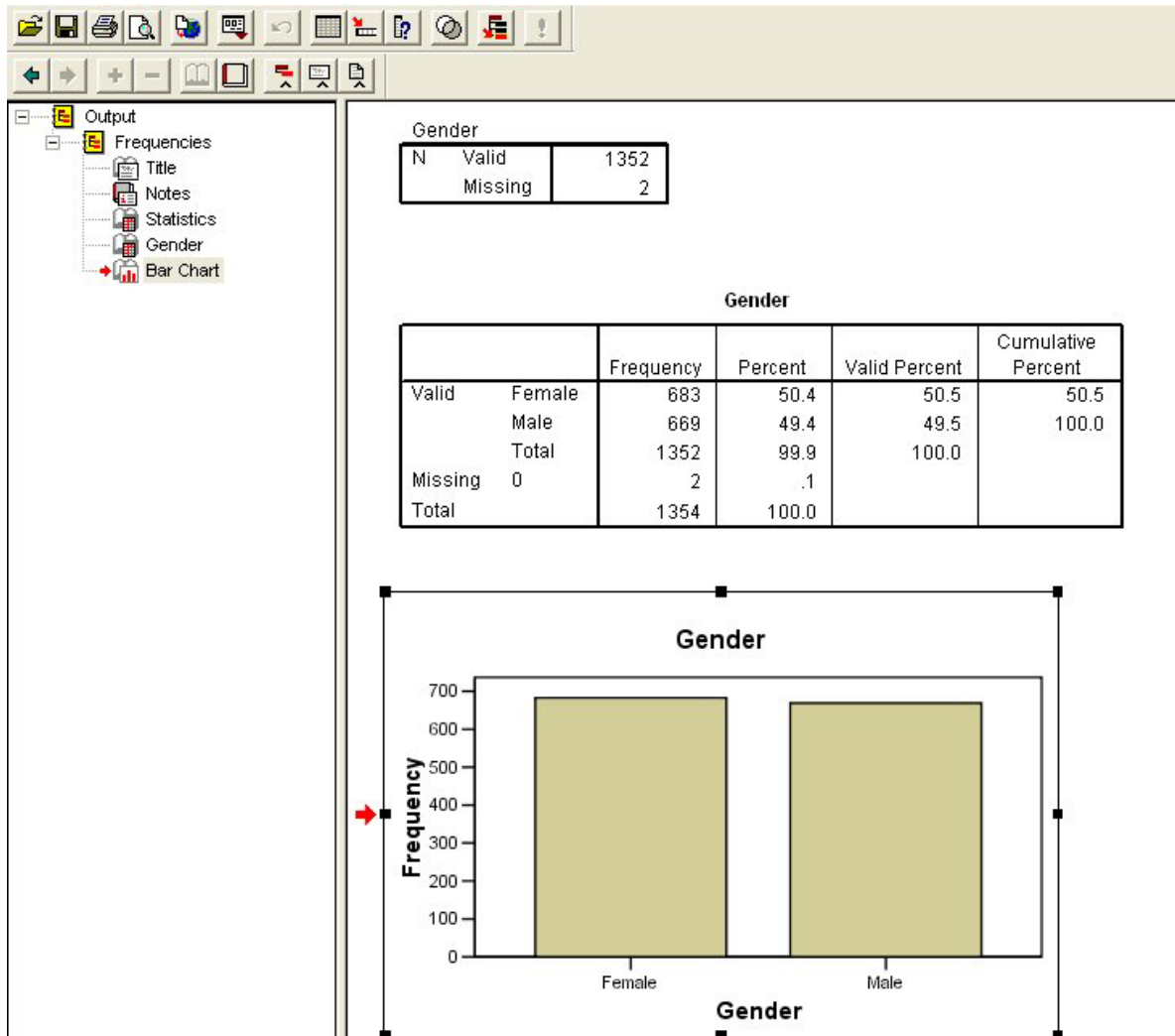
STATISTICS: This brings up a sub-menu for calculating descriptive statistics. Keep in mind that you would select this option only for interval and ratio data.

CHARTS: This option gives access to preparing a chart for the variables you for which you are computing frequency tables.

- Bar Charts should be selected for Nominal data.
- Histogram should be selected for Interval and Ratio data.

FORMAT: This option lets you tell SPSS how to order the data in the frequency table. You can select highest to lowest (preferred) or lowest to highest.

SPSS OUTPUT
SPSS Viewer



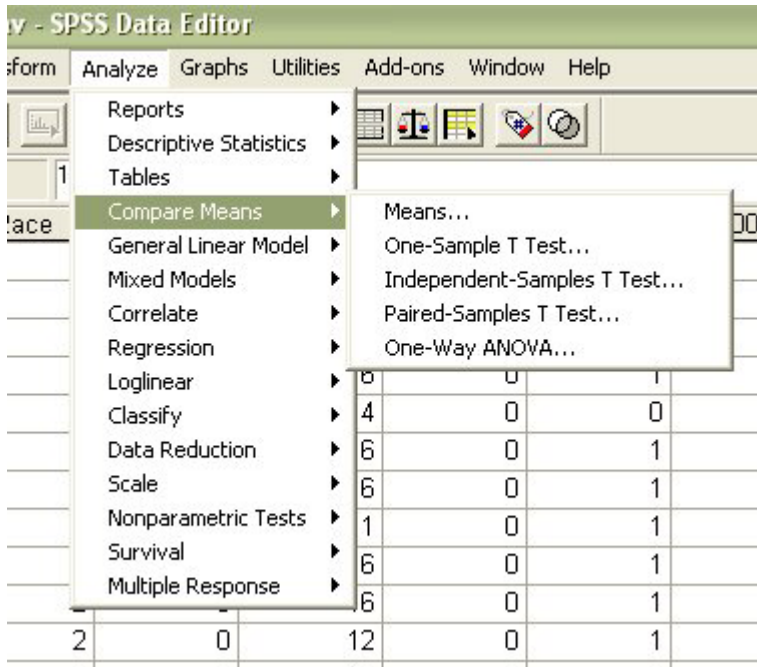
All SPSS output is contained in its own file that is separate from the data file. You can modify text in the output, copy tables and figures so you can place them in other documents, and print the whole file or parts of it.

To modify the label for tables and figures.

1. Click the table or figure to make it active. When a table or figure is active it is surrounded by a box with handles. The Gender bar graph is active.
2. Click on the title to make it active.
3. Type your text for the appropriate title.

t-test

SPSS Analyze ->Compare Means



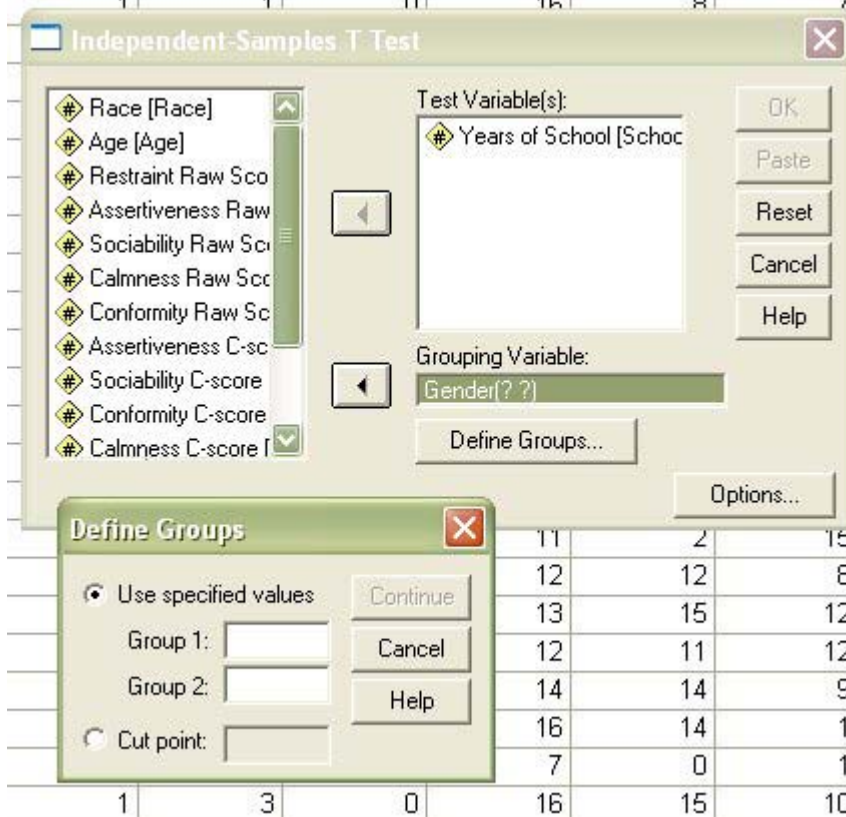
This screen allows you to select a variety of statistical tests that compare means.

MEANS compares the means for various broken down by another categorical variable (Gender, Race, etc.). It SPSS calculates the descriptive statistics but does not do a statistical test of the difference between the means.

ONE-SAMPLE T TEST compares the means for within group (repeated measures) studies.

INDEPENDENT-SAMPLES T TEST compares the means of two independent groups.

To do the t-test you need to do two things. First, identify the dependent variable. Second, identify the independent variable. The screen for doing this is shown below.



1. Test Variable(s): is the window for the dependent variable. You can enter one or more variables here.

2. Grouping Variable: This is the window for entering the independent variable. You need to tell SPSS what the values are for the grouping variable so, click ‘Define Groups...’ This will bring up the Define Groups screen. Then you need to enter the values for each group. For example, suppose the grouping variable is *Gender* and there are three values for *Gender* (1 for male, 2 for female, and 3 for undecided) and you want to compare males to females. Then enter 1 for Group 1 and 2 for Group 2 then click Continue.

3. Click OK to do the analysis.

t-test: Sample output.

SPSS computes two tables for the t-test. The **Group Statistics** table is a breakdown of the descriptive statistics by *Gender*.

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Restraint Raw Score	Male	669	10.58	4.485	.173
	Female	683	9.31	4.622	.177

The second table is the t-test summary. Levene's Test is an evaluation of the assumption that the variances of the two samples are equal (one of the t-test assumptions). If the differences are not statistically significant, shown in the **Sig.** column, then use the row labeled *Equal variances assumed*. Otherwise use the row labeled *Equal variances not assumed*.

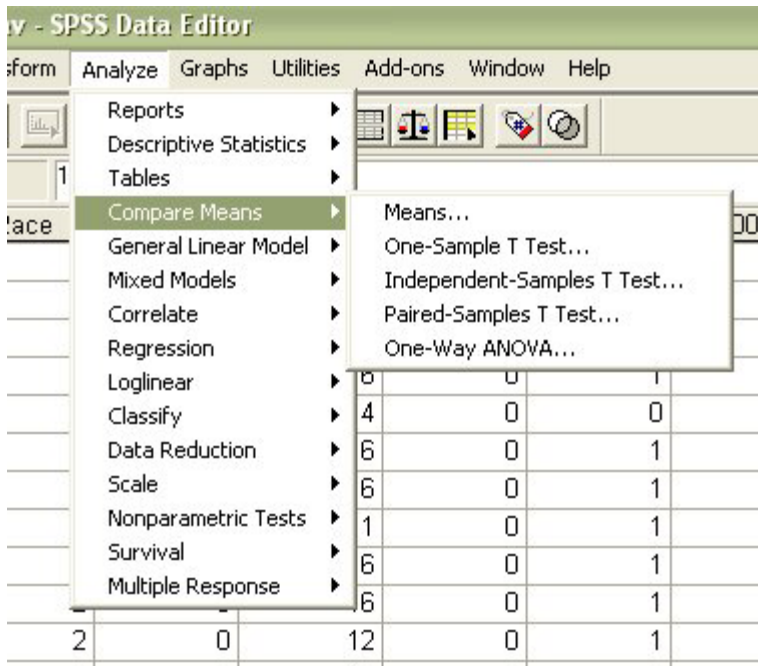
The *t* is shown in the *t* column, and the significance level is shown in the *Sig. (2-tailed)* column. SPSS calculates the actual significance level. In the example the significance level would be $p < .0001$. The *Mean Difference* column shows the difference between the Male and Female means.

Independent Samples Test

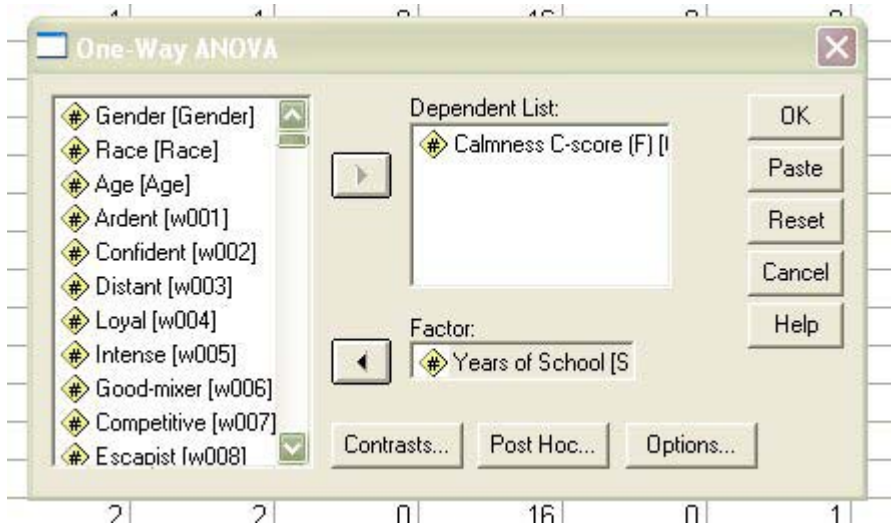
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Restraint Raw Score	Equal variances assumed	2.099	.148	5.125	1350	.000	1.270	.248	.784	1.756
	Equal variances not assumed			5.126	1349.878	.000	1.270	.248	.784	1.756

ONE-WAY ANALYSIS OF VARIANCE

SPSS Analyze->Compare Means ->One-way ANOVA



One-Way ANOVA. This is used when there is one independent variable with more than two levels and one dependent variable.



This screen is used to select the dependent variables '**Dependent Variable List**' and the independent variables: '**Factor**'

Post Hoc tests can be selected using the **Post Hoc** button. Select those that are appropriate for the analysis you are using: Either Tukey or Scheffe.

A sample of the One-way ANOVA output is shown on the next page.

An SPSS Primer

ONE-WAY ANALYSIS OF VARIANCE

SPSS Output

SPSS provides several tables summarizing the ANOVA analysis. The first is the ANOVA summary table shown below. As with the t-test SPSS calculates the actual significance level in the *Sig.* column.

Conformity Raw Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	467.299	3	155.766	9.323	.000
Within Groups	22556.536	1350	16.709		
Total	23023.835	1353			

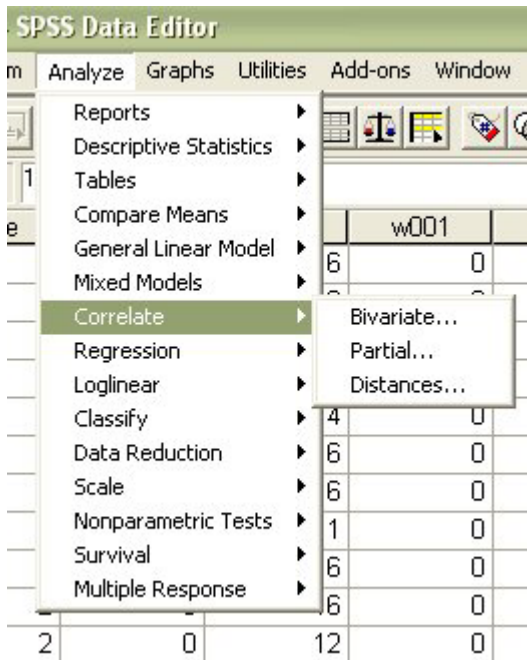
Second major table is the summary of the pairwise comparison of the means for the independent variable. This table has a lot of redundancy. For example, in the second row, the Asian mean is compared to Black, White and Hispanic means. In the third major row the Black mean is compared to the Asian, White, and Hispanic means. SPSS flags significant differences with an *.

Dependent Variable: Conformity Raw Score

Tukey HSD

(I) Race	(J) Race	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Asian	Black	-.046	.612	1.000	-1.62	1.53
	White	.414	.518	.855	-.92	1.75
	Hispanic	-2.115(*)	.688	.012	-3.88	-.35
Black	Asian	.046	.612	1.000	-1.53	1.62
	White	.460	.371	.601	-.49	1.41
	Hispanic	-2.068(*)	.585	.002	-3.57	-.56
White	Asian	-.414	.518	.855	-1.75	.92
	Black	-.460	.371	.601	-1.41	.49
	Hispanic	-2.529(*)	.485	.000	-3.78	-1.28
Hispanic	Asian	2.115(*)	.688	.012	.35	3.88
	Black	2.068(*)	.585	.002	.56	3.57
	White	2.529(*)	.485	.000	1.28	3.78

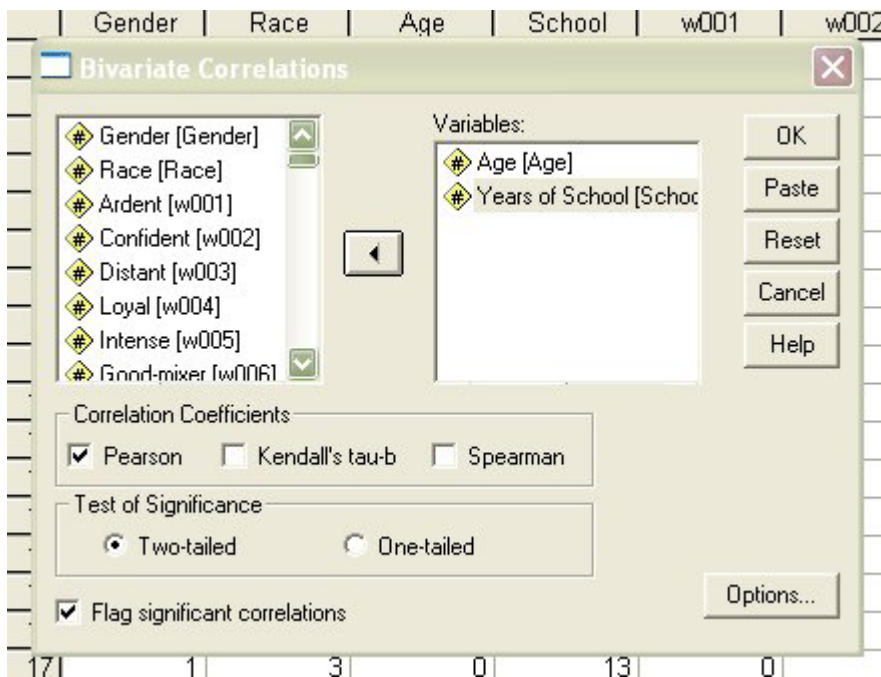
An SPSS Primer
 CORRELATIONAL ANALYSES.
 SPSS Analyze -> Correlate



This screen lets you select three types of correlational analyses.

BIVARIATE- This is the correlation between two variables. For regular correlations select 'Bivariate'.

PARTIAL -This is the analysis of the correlation between two variables removing the influence of a third variable.



Use this screen to select the variables to analyze. You can pick as many variables as you want.

Note: As shown here,

-The normal Pearson correlation will be calculated.

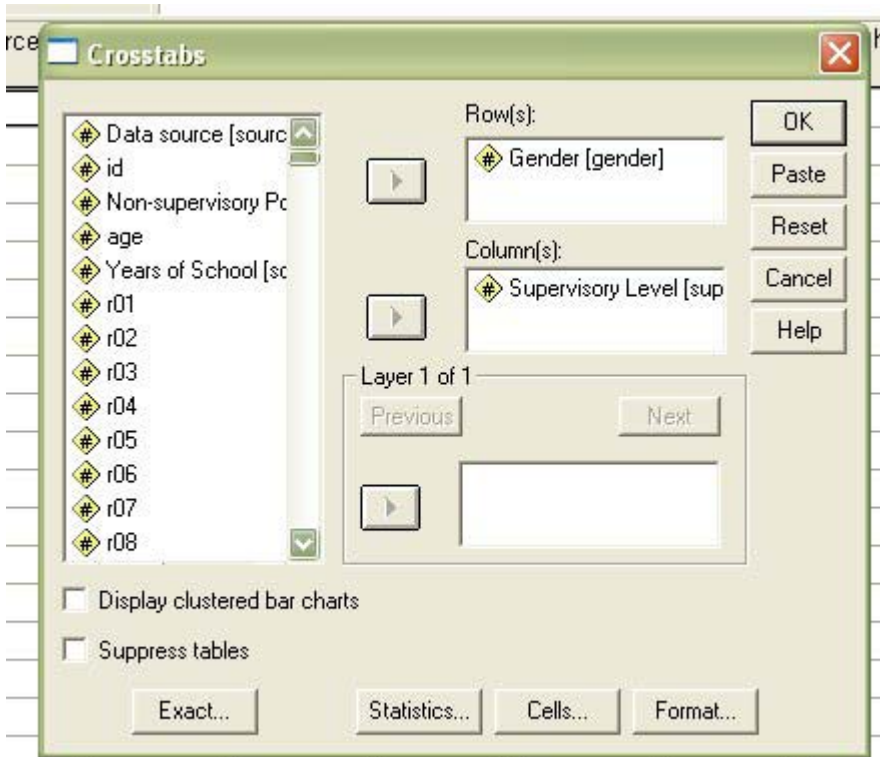
- SPSS will calculate a two-tailed significance level.

- Correlations that are statistically significant will be flagged with an asterisk.

- The **Options** button permits you to select various options such as

calculation of the descriptive statistics.

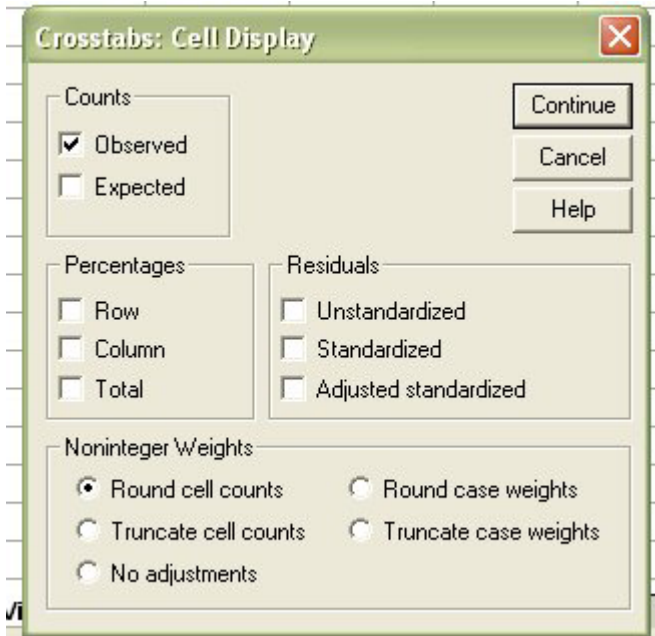
CROSSTABS - This selection gives a breakdown of two or more nominal variables.



Rows. This selection identifies the variable that will be listed in each row.

Column(s) This selection identifies the variable that will be listed in the table columns.

Statistics... Use this button to access the Chi-square option that is located in the upper left-hand corner of this option box.



Cells: This option gives various options for displaying the cross tab table. One useful option is the *Percentages* check box. This will tell SPSS to put the percent of responses for each Row or Column or the Total percentages. This is not always necessary but can be helpful for interpreting the data.

A sample of the chi-square output is shown on the next page.

NOTE: You can also compute Chi-square by going to the Analyze -> Nonparametric Tests selection. If you chose this option you do not get the Crosstabs table.

An SPSS Primer
 CHI-SQUARE ANALYSIS
 SPSS Output

This is a sample of a Crosstab table. It shows the Row Percentages (% within Gender), Column percentage (% within Supervisory Level) and Total percent (% of Total). Sometimes these percent figures can be confusing, so they should be used only when needed.

Gender * Supervisory Level Crosstabulation

			Supervisory Level					Total
			Supervisor.	First -Level Manager	Mid-Level Manager	Executive or V.P.	President or CEO	
Gender	Male	Count	279	262	365	212	44	1162
		% within Gender	24.0%	22.5%	31.4%	18.2%	3.8%	100.0%
		% within Supervisory Level	46.5%	57.6%	80.6%	86.9%	95.7%	64.6%
		% of Total	15.5%	14.6%	20.3%	11.8%	2.4%	64.6%
		Count	321	193	88	32	2	636
Female	Female	% within Gender	50.5%	30.3%	13.8%	5.0%	.3%	100.0%
		% within Supervisory Level	53.5%	42.4%	19.4%	13.1%	4.3%	35.4%
		% of Total	17.9%	10.7%	4.9%	1.8%	.1%	35.4%
		Count	600	455	453	244	46	1798
		% within Gender	33.4%	25.3%	25.2%	13.6%	2.6%	100.0%
Total	Total	% within Supervisory Level	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
		% of Total	33.4%	25.3%	25.2%	13.6%	2.6%	100.0%

This table summarizes the Chi-Square analysis. The Pearson Chi-Square is the one used most frequently.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	218.761(a)	4	.000
Likelihood Ratio	235.176	4	.000
Linear-by-Linear Association	208.672	1	.000
N of Valid Cases	1798		

a 0 cells (.0%) have expected count less than 5. The minimum expected count is 16.27.

THE SPSS Data Set

These data are subset of a larger set of data collected to restandardize a personality test called the AVA. They were collected over a two-year period in 1992 and 1993 and represent a stratified sample of working adults around the United States.

The AVA measures five constructs Assertiveness (V-1), Sociability (V-2), Calmness (V-3), Conformity (V-4) and Conscious Restraint (V-5). People taking the test are given a set of self-descriptive adjectives and are asked to check those that they feel describe them.

My goal in this project was to develop a new, updated version of the AVA. The data set consists of the following variables.

GENDER	Male or Female
RACE	White, Black, Hispanic, or Asian
AGE	Age at the most recent birthday
SCHOOL	Number of years of school the person attended.
Raw scores	These are the raw scores reflecting the count of the number of words checked for each of the scales.
C-Scores	These are standardized scores for each scale based on a mean of 50 and a standard deviation of 10.