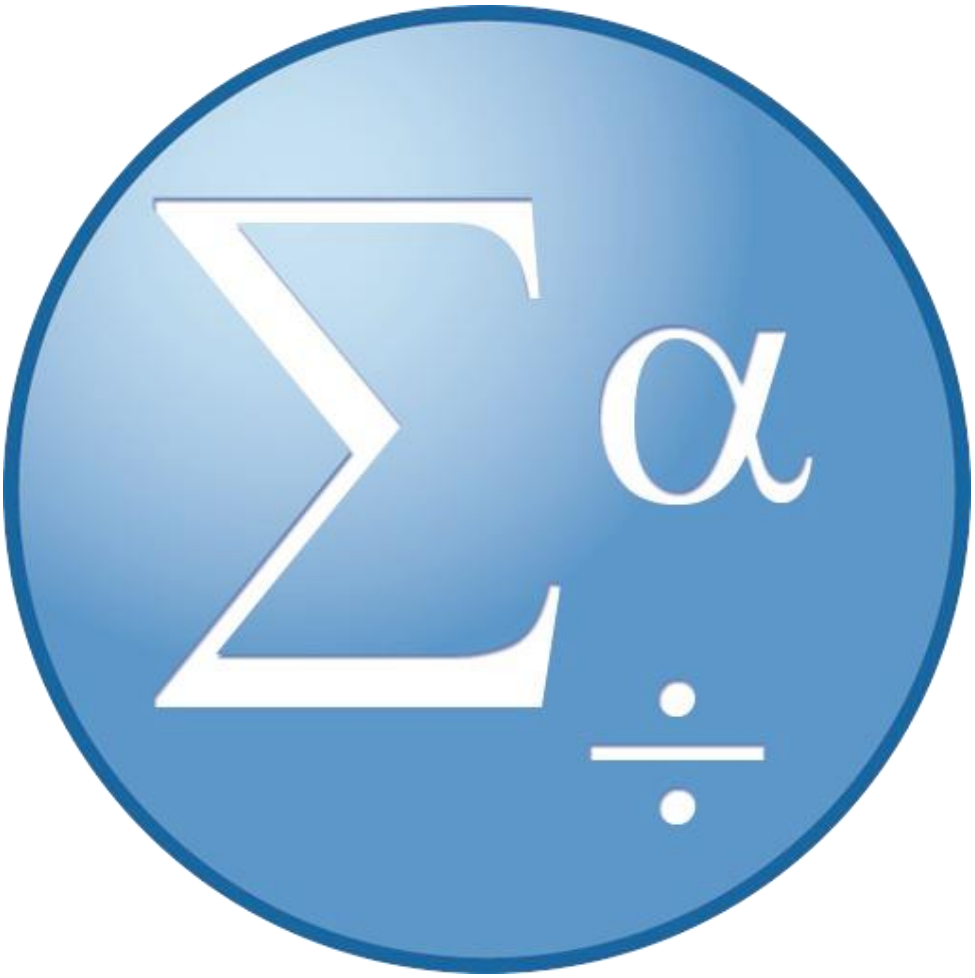
SPSS for Inferential Statistics



(Version 1:8/27/2014)

Background Information

IBM SPSS Statistics is a software package used for statistical analysis, data management, and data documentation. It allows even novice researchers to do their own statistical analysis with ease. This program is widely used by individuals with interests in social sciences, market research, health research, surveys, government, and education research.

This intermediate workshop is designed to further expand knowledge on statistical analysis and SPSS functions beyond the introductory level such as running descriptive or inferential analysis types, make charts with curves and change measurement types in chart builder, transform data, importing excel files, and using the variable window to manage data more effectively.

Required Skills

Before beginning, the following skills are required:

- Basic knowledge of Statistical terminology
- Basic knowledge of SPSS functions
- Experience with software navigation (keyboard and mouse)
- Basic Excel knowledge

Agenda

- Open an existing .sav or import from excel .xlsx
- How to sort data in Data view
- How to transform data (create new variable from old variables) in data view
- How to run Inferential analysis (Pearson correlation, t-test, regression, chi square ANOVA, MANOVA)
- Format charts

Opening your existing .sav file or importing a .xlsx file

Opening an existing .sav file from the welcome menu

1. In the opening welcome screen select *Opening an Existing Data Source>More Files > Okay > File > Open*

Opening an existing .sav file from the Data Editor window

1. In the data editor window go to the top menu
2. Select *File > Open > Data*

3. Alternatively you can use the *Open Data Document* button on the toolbar

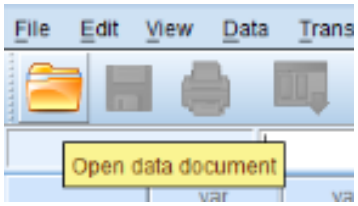


Figure 1 - Open data document button

Activity: Open the file Squirrel_Survey.sav

Importing an .xlsx file from Excel

1. Before you begin your import, make sure that the file is not open in Excel
2. In the opening welcome screen select *Opening an Existing Data Source*
3. In the *Open Data* window select the *Files of Type* drop down menu and select *Excel (*.xls, *.xlsx, *.xlsm)*
4. Select your Excel file to open in the Data Viewerand select *Okay*
5. An *Opening Excel Data Source* Window will appear
6. Check the *Read variable names from first row* if applicable
7. If your Excel file has multiple worksheets, you can select the worksheet you wish to import in the *Worksheet* dropdown menu
8. If you want to import only a portion of the spreadsheet, specify the range of cells to be imported in the *Range* text box
9. Click *Okay* to view the Excel file

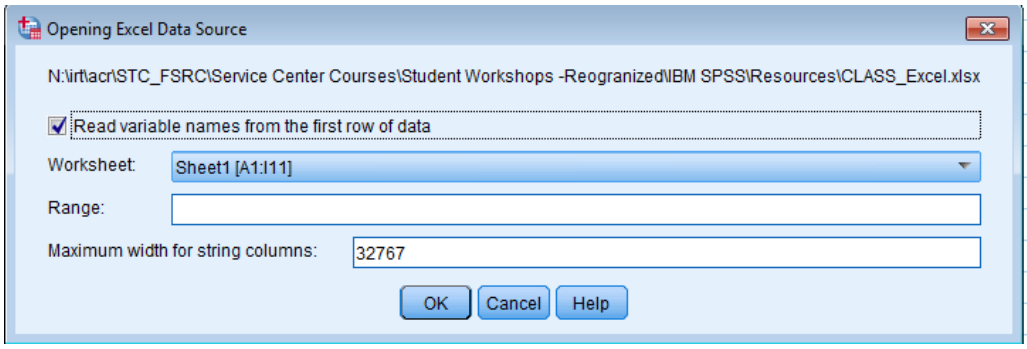


Figure 2 - Opening excel data source window (steps 5 and 6)

Activity: Import RawSquirrelData.xlsx

Saving your SPSS file

Saving your data in SPSS is similar to saving a file in any other program. The main difference is that you will have to save two files as your information from the *Output Viewer* and *Data Viewer* are separate files with separate file names.

To save a file in SPSS:

1. Go to the top menu and select *File > Save* and select a file location and name
   - Tip: name you output and data files similarly (ex: Squirrel_Survey_Data.sav and Squirrel_Survey_Output.spv)
2. Click *Save*

Sorting Your Data in Data View

Sorting your data visually organizes your data

1. Select *Data > Sort Cases*
2. In the *Sort Cases* window select your variable you wish to sort and move it to the *Sort By* window
3. Select your *Sort Order* and select *Okay*

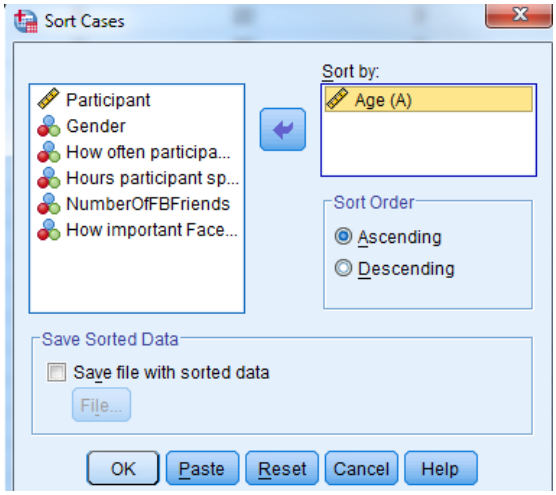


Figure 3 - Sort Cases window

Your data is now changed in the Data View tab

Activity: Sort the data by *Age* in ascending order

Transforming Data in Data View

You can create new variables from old variables with the transformation tool

1. In the Data Editor window go to the top menu
2. Select *Transform > Compute Variable*
3. Select your type of function in the *Function Group* box
4. Select your sub-function in the *Functions and Special Variables*
5. Follow the prompt in the lower center box and enter it in the *Numeric Expression* box (Ex: if MEAN is selected as the sub-function MEAN(x,y) will show in the lower prompt box)
6. Select your variables to transform and move them into your equation in the *Numeric Expression* Box
7. Label your new variable in the *Target Variable* box

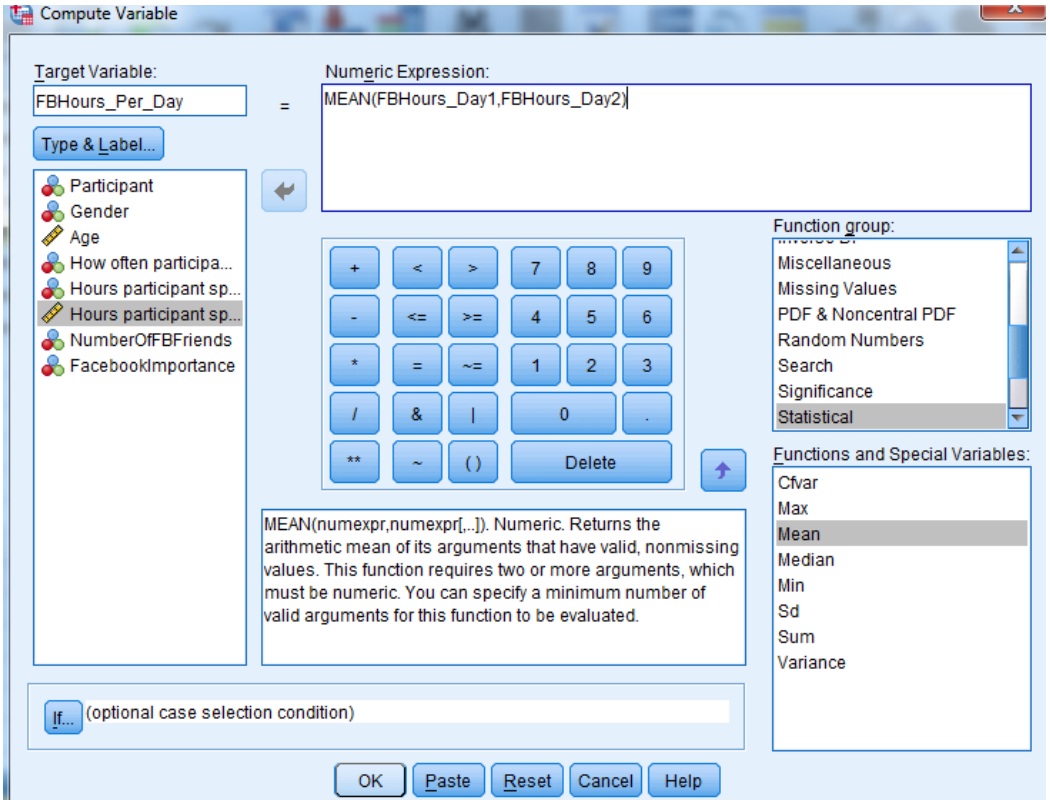


Figure 4 - Compute Variable window

8. Select the *Okay* button
9. Your new variable is now displayed in the Data View tab

Activity: Make a new variable using the *Statistics* function group, choose the function *Mean*, and select the variables: HrsPerDayPrsns1, HrsPerDayPrsns2, HrsPerDayPrsns3)

Running Inferential Analysis:

Inferential statistics is when we “make inferences”, do hypothesis testing, determine relationships, and make predictions.

Pearson’s Correlation

A correlation is a statistical device that measures the strength of a degree of a supposed linear association between two or more variables. One of the more common measures used is the Pearson Correlation, which estimates a relationship between two interval variables.

1. Go to the top menu
2. Select *Analyze > Correlate > Bivariate*. The Bivariate Dialog Box will open
3. Decide what variables you want to correlate (you can do many at once)
4. Move your variables over to the *Variable* box using the arrow button
5. Select the *Pearson Correlation Coefficient* check box (if not already checked)
6. Select the *two-tailed* or *one-tailed* test of significance bubble

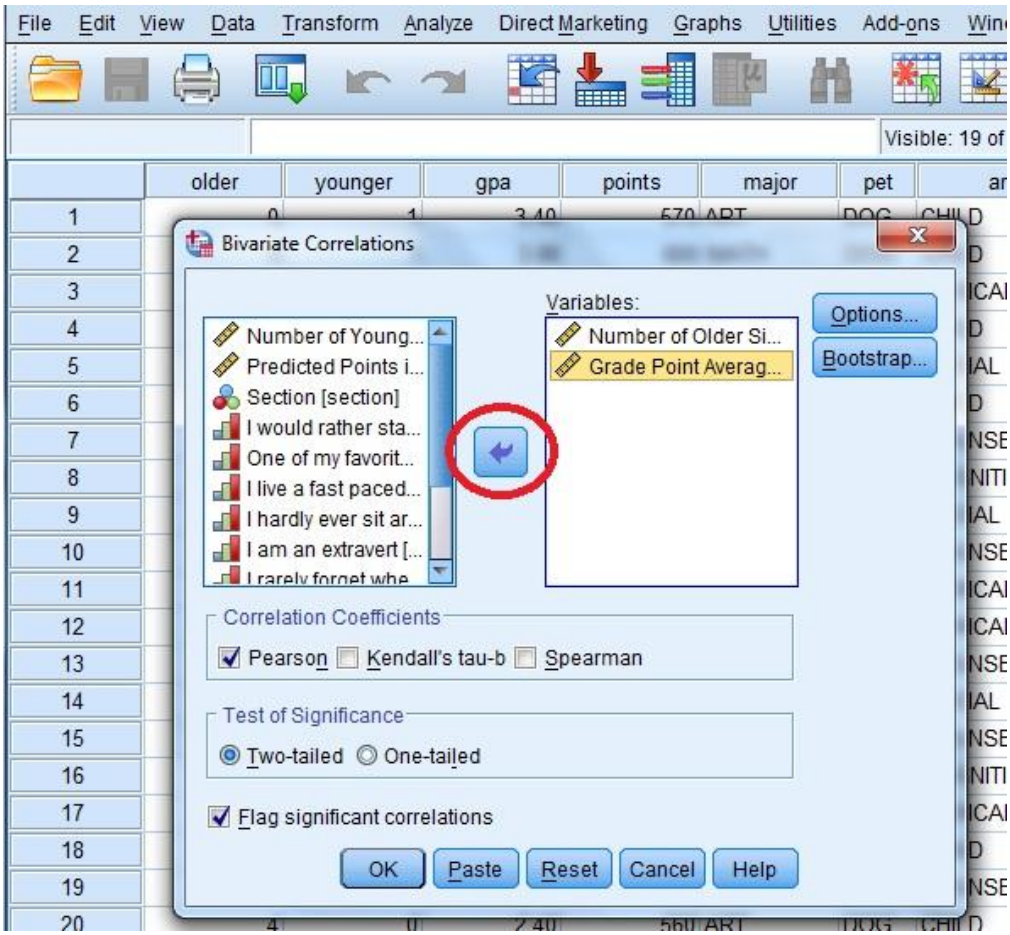


Figure 5- Bivariate Correlations window (step 4 shown)

7. Click *Okay* and the output viewer window will appear with the correlations table

Activity: Create a Pearson's Correlation using *CuteScale*,and*FeedingFrequency*

Linear Regression Analysis

Regression analysis is about predicting the future (the unknown) based on data collected from the past (the known). Such an analysis determines a mathematical equation that can be used to figure out what will happen within a certain range of probability. The analysis is performed on a single dependent variable and it takes into account which independent variables have more effect than others. One type of regression analysis linear regression, which is used when the projections are expected to be in a straight line with actual values.

1. Go to the top menu
2. Select *Analyze > Regression > Linear*
    - The Linear Regression window will appear
3. Select your dependent variable and move it to the *Dependent* box using the arrow button
4. Select your independent variable(s) and move them into the *Independent(s)* box using the arrow button

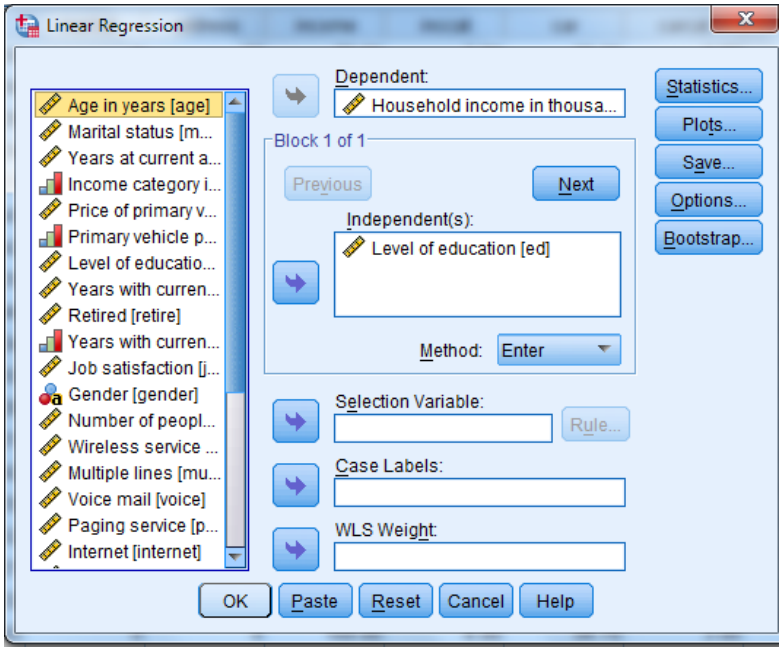


Figure 6 - Linear Regression window (steps 1-4 completed)

5. Click *Okay*

Several tables will appear in the output window, the *Coefficients*[a]table contains the coefficients required to make predictions.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	53.108	2.330		22.794	.000
	Level of education	6.327	.817	.096	7.742	.000

a. Dependent Variable: Household income in thousands

Figure 7 - Coefficients table

Activity: Independent T-test

An independent T-test is used to determine the likelihood that two independent data samples came from populations with identical means. If this were true, then the difference between the means should equal zero. In this case, the null hypothesis would indicate that two means are equal.

Two variables are required in the data set. One variable is the measured parameter. Examples include weight, height or frequency. The second variables divide the data set into two groups. The means of light and dark groups will be compared.

1. Go to the top menu
2. Select *Analyze > Compare Means > Independent Samples T Test.* The Independent-Samples T Test dialog box will open
3. Select continuous variables that you want to test from the list
4. Click on the arrow that will send them to the *Test Variables* box
5. Select the categorical variable from which you are going to extract the groups for comparison and send it to the *Grouping Variable* box by pressing the lower arrow

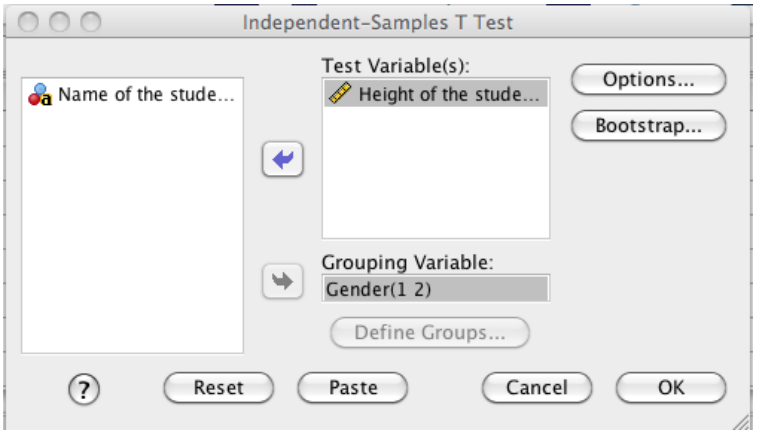


Figure 8 - Independent-Samples T Test window (steps 1-5 completed)

6. Click on the *Define Groups* button
7. The *Define Groups* window will pop up

8. Click *Continue* after specifying your group values

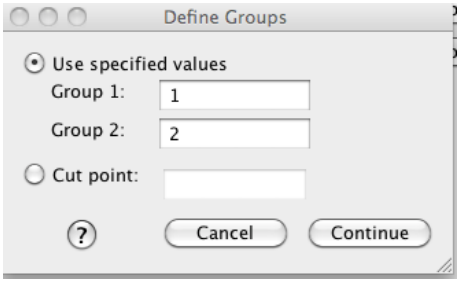


Figure 9 - Define Groups window

9. Click *Okay* to view your output

Paired Sample T-Test

This type of t-test is used if an observed difference between the two means of a paired samples set is statistically significant.

1. Select *Analyze > Compare Means > Paired Samples T-Test*
2. Highlight the two variables upon which you want to run your analysis. When you have the two highlighted, send them over to the right column with the arrow button. You can then define more variable pairs if you wish, but if that's all you want, then just click *OK*

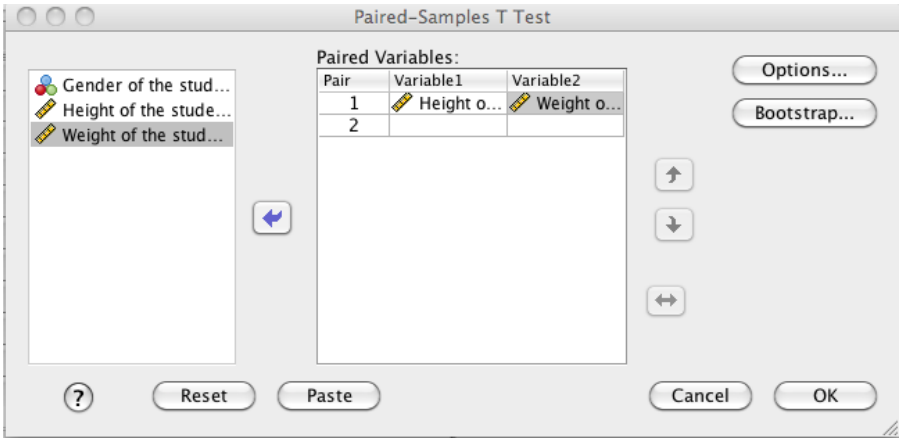


Figure 10 - Paired-Samples T Test window (step 2 shown)

Activity: Create a Paired Sample T-test using the variables *SquirrelAvoidance* and *ThrownItems*

Crosstabs and Chi-Square

1. Go to the top menu
2. Select *Analyze > Descriptive Statistics > Crosstabs*

3. A dialog box will appear, select your independent variable in the left box and move it over to *Row(s)*using the arrow button
4. Select your dependent variable and move it over to the *Column(s)* box
5. To find the Chi-Square:Click *Statistics>Chi-square>Continue*

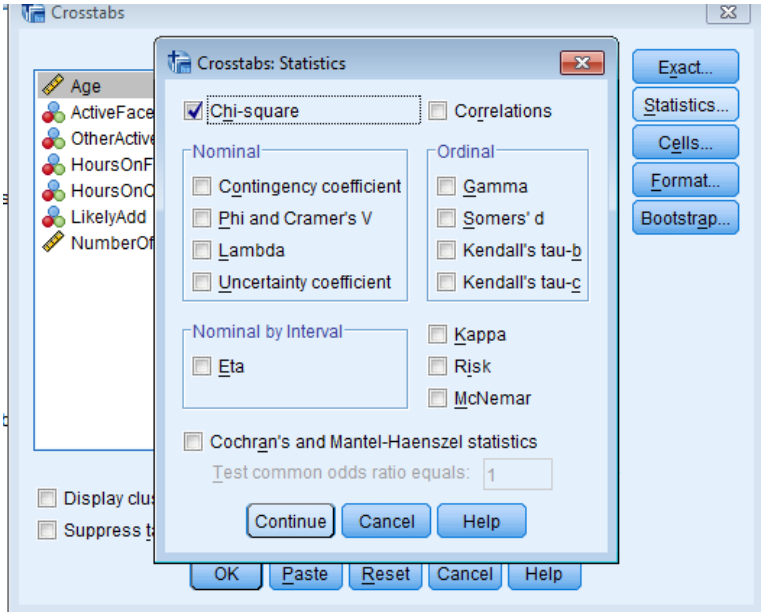


Figure 11 - Crosstabs: Statistics window in from of the Crosstabs window (step 5 shown)

6. Click the *Okay* button

Activity: Make a crosstab with a chi-square of *CuteScale* and *SquirrelOwnership*

ANOVA

One-Way ANOVA

One-way analysis of variance (One-Way ANOVA) is the analysis of the variance of values (of a dependent variable) by comparing them against another set of values (the independent variable). It is a hypothesis that the mean of the tested variable is equal to that of the factor.

1. In the Data View tab Go to the top menu
2. Select *Analyze > Compare Means > One-Way ANOVA*
3. Move your dependent variable to the *Dependent list* box
4. Move your independent variable to the *Factor* box
    a. Click *Options* if you want to add any extra statistics to your analysis

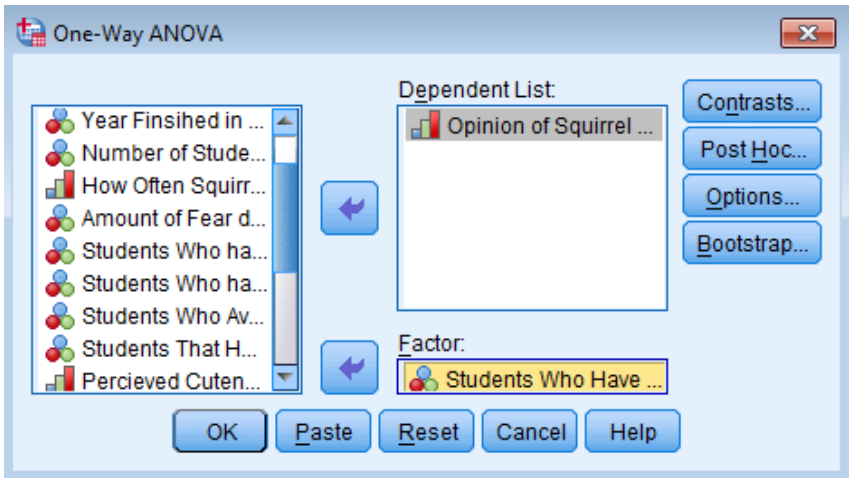


Figure 12 - One-Way ANOVA window (steps 3 and 4 shown)

5. Click *OK.* The output viewer window will open showing your One-Way ANOVA

Activity: Make a One-Way ANOVA with *SquirrelAcceptance* as the Dependent and *ThrownItems* as the factor

MANOVA (GLM Multivariate)

In SPSS MANOVA is referred to as GLM Multivariate.

MANOVA is a statistical test procedure for comparing multivariate (population) means of several groups. As a multivariate procedure, it is used when there are two or more dependent variables, although statistical reports provide individual p-values for each dependent variable in order to test for statistical significance.

1. In the Data view tab go to the top menu and select *Analyze >General Linear Model > Multivariate*
2. Move your dependent variables in the *Dependent Variables* box
3. Move your independent variable to the *Fixed Factor(s)* box
4. Click *OK*

Activity: Make a GLM multivariate with *SquirrelFearLvl*and *SquirrelAcceptance* as the DV and *JumpedOn*as the IV

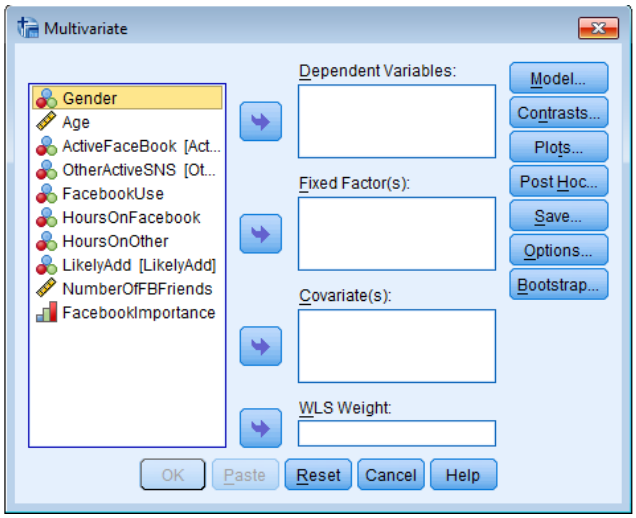


Figure 13 - Multivariate window

Formatting Charts

Selecting Chart Elements

To format your chart you must first select the elements of your chart that you wish to edit.

1. Open the Output viewer window
2. Double click your chart that you wish to format
3. The *Chart Editor* window will appear
4. Click on the exact element you wish to edit (it will be outlined in yellow when you do so)
    a. To select an individual bar press the *Ctrl* key and click and drag to select only one or several bars
    b. To directly edit the text double click on the area you wish to change

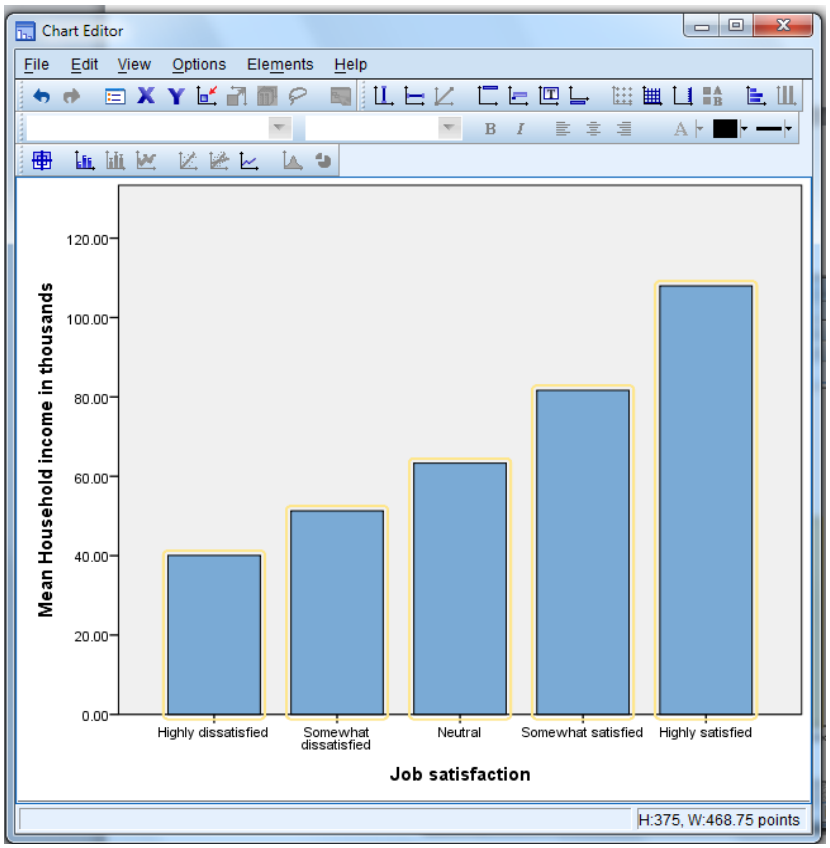


Figure 14 - Chart Editor window with all bars selected

Activity: Make a bar graph with *Opinion of Squirrels* on the x-axis and *HoursSpentArundSquirrels* in the y-axis

Changing Bar Colors

1. Open the *Chart Editor* window (See "Selecting Chart Elements" above)
2. Select the bars you wish to change
3. If the *Properties* dialog is not already open you can do so by going to the top menu in the *Chart Editor* window and select *Edit > Properties* or press *Ctrl + T* on your keyboard
4. To specify color attributes of graphic elements (excluding lines and markers) click on the *Fill & Border* tab
5. Click the swatch next to *Fill* to indicate that you want to change the fill color of the bars. The numbers below the swatch specify the red, green, and blue settings for the current color
6. Select the color you want to use from the palette on the right side of the window

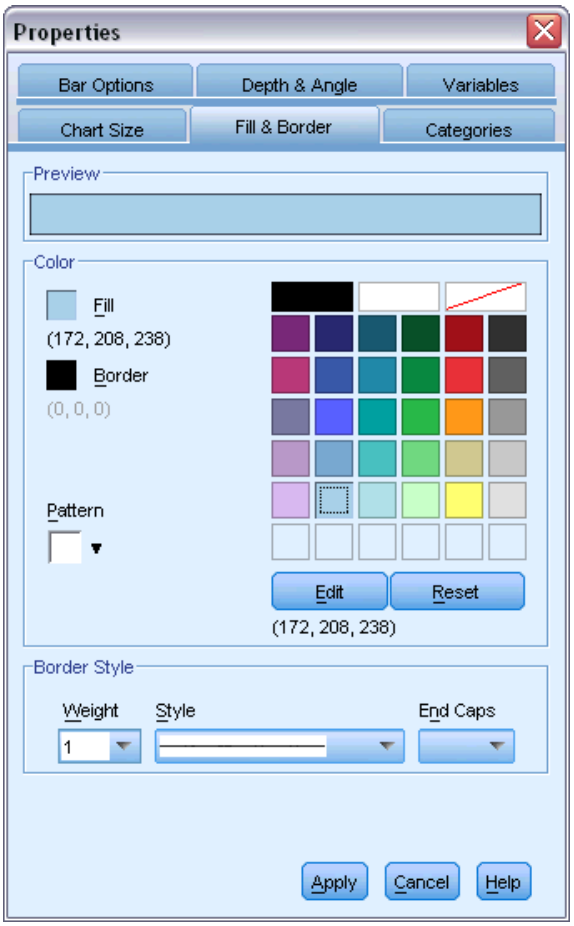


Figure 15 - Properties window with the Fill & Border tab open

7. Click *Apply*
8. Close the *Properties* and *Chart Editor* window if you are finished

Formatting Numbers in Tick Labels

Changing the numbers format in a tick label can make it easier to understand and more attractive to the reader.

1. Open the *Chart Editor* window (See "Selecting Chart Elements" above)
2. Select the y axis tick labels by clicking any one of them
3. If the *Properties* dialog is not already open you can do so by going to the top menu in the *Chart Editor* window and select *Edit > Properties* or press *Ctrl + T* on your keyboard
4. Click the *Number Format* tab
5. Type the number of decimal places you want to use in the *Decimal Places* box (if you do not want decimal places type: **0**)

6. Type your scale in the *Scaling Factor* box (The scaling factor is the number by which the Chart Editor divides the displayed number) Ex: Typing .001 changes 1 to 1000 Your change will be displayed in the *Sample* box before you apply it
7. Select the *Display Digit Grouping*box if you want to mark each thousandth place in the number

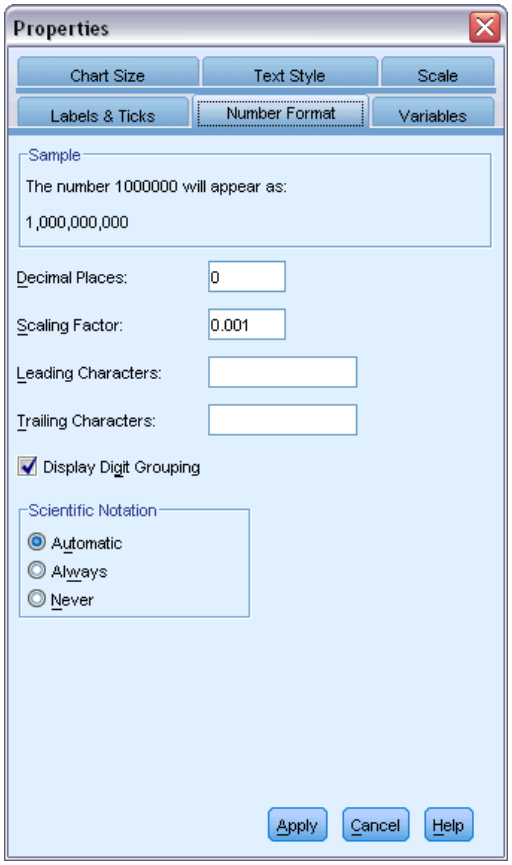


Figure 16 - Number Format tab in the Properties window (steps 1-7 completed)

8. Click *Apply*
9. Close the *Properties* and *Chart Editor* window if you are finished

Editing Text

1. Open the *Chart Editor* window (See "Selecting Chart Elements" above)
2. Click, pause, and Click again (A double click will open the *Properties* dialogue box) on the text that you wish to change (While in edit mode, the Chart Editor positions any rotated text horizontally. It also displays a flashing red bar cursor)
3. Type in the text you want
4. Press Enter to exit edit mode and update the title

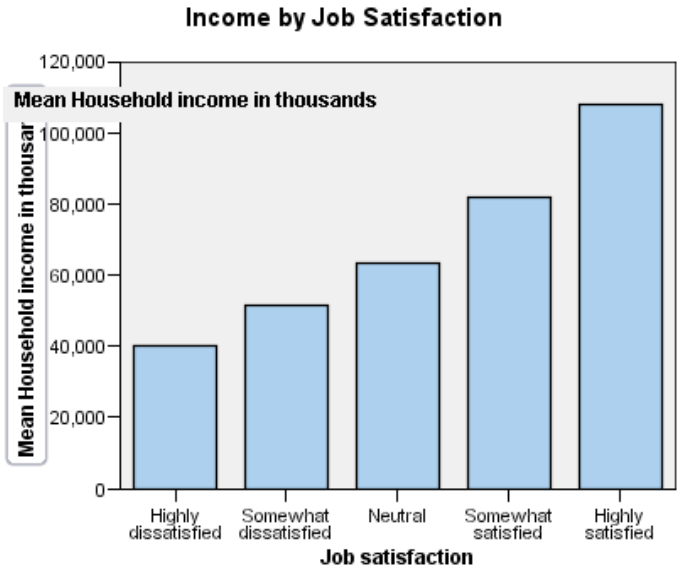


Figure 17 - Chart with edit mode open on the y axis

Formatting Text

1. Open the *Chart Editor* window (See "Selecting Chart Elements" above)
2. There are three ways to open the *Properties* dialogue box:
    a. The simplest way is to double click on the text you wish to format
    b. Single click on the text you want to format and press *Ctrl + T* on your keyboard
    c. Single click on the text you want to format and select *Edit > Properties* from the top menu
3. Select the tab *Text Style*
4. Change the text color in the *Color* box by selecting a color from the palette on the right side. Your selected color will appear in the *Text Color* box and in the *Preview in Preferred Size* box near the top of the window
5. You can change the font family, size and style with the labeled drop down menus in the *Font* box
6. Click the *Apply*
7. Close the *Properties* and *Chart Editor* window if you are finished

Displaying Data Value Labels

This is used to show the exact values associated with the graphic elements

1. Open the *Chart Editor* window (See "Selecting Chart Elements" above)
2. Select the graphic element you want to add data levels to
3. From the top menu in the *Chart Editor* window select *Elements > Show Data Labels*
4. Select what values you want to display and move them to the *Displayed* box

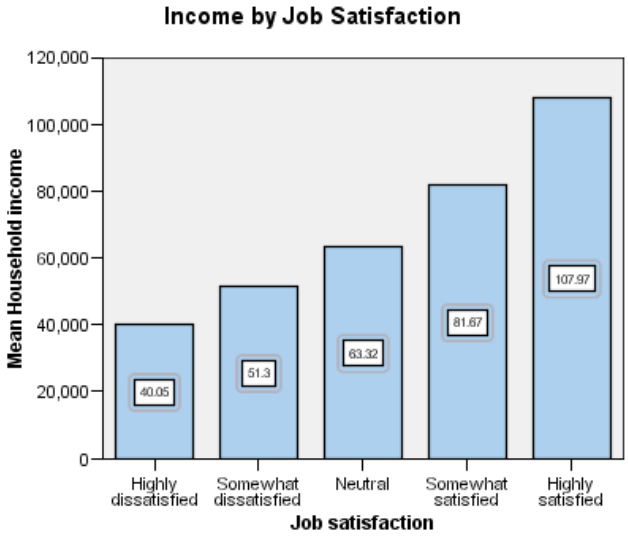


Figure 18 - Data Labels shown on chart

5. Select *Apply*
6. Close the *Chart Editor* window if you are finished

Wrapping Up:

- Opening your existing .sav file
    - open files from the welcome menu or from the Data Editor
    - Don't forget to save often
- Sorting data from SPSS organizes your data
- If you need to create a new variable from an old one use the Transformation tool
- Pearson's correlations are labeled as bivariate correlations in SPSS
- Linear Regression analysis
- Independent and paired sample T tests are found under *Compare Means* in the analyze drop down
- You can create a chi-square analysis through the crosstab feature
- While ANOVA is found under *Compare Means* MANOVA's are found under *General Linear Model*
- There are many chart formatting options in the *Chart Editor*, just double click on your chart in the Output Viewer to access this