Introduction to SPSS

This document will guide you through a general introduction to the SPSS interface as well as some of the basic functions and commands you would be likely to perform in SPSS.

BEFORE YOU BEGIN, PLEASE ENSURE YOU HAVE DOWNLOADED THE SAMPLE DATA FILE USED IN THIS GUIDE: SPSSsampledata.xls

The SPSS Interface

Once you have started SPSS you may be prompted to open existing data, run the tutorial, etc. For the purpose of this activity you will simply **click Cancel**.

At this point you will be viewing the Data Editor of SPSS in *Data View* mode. Your screen should look similar to the picture below.
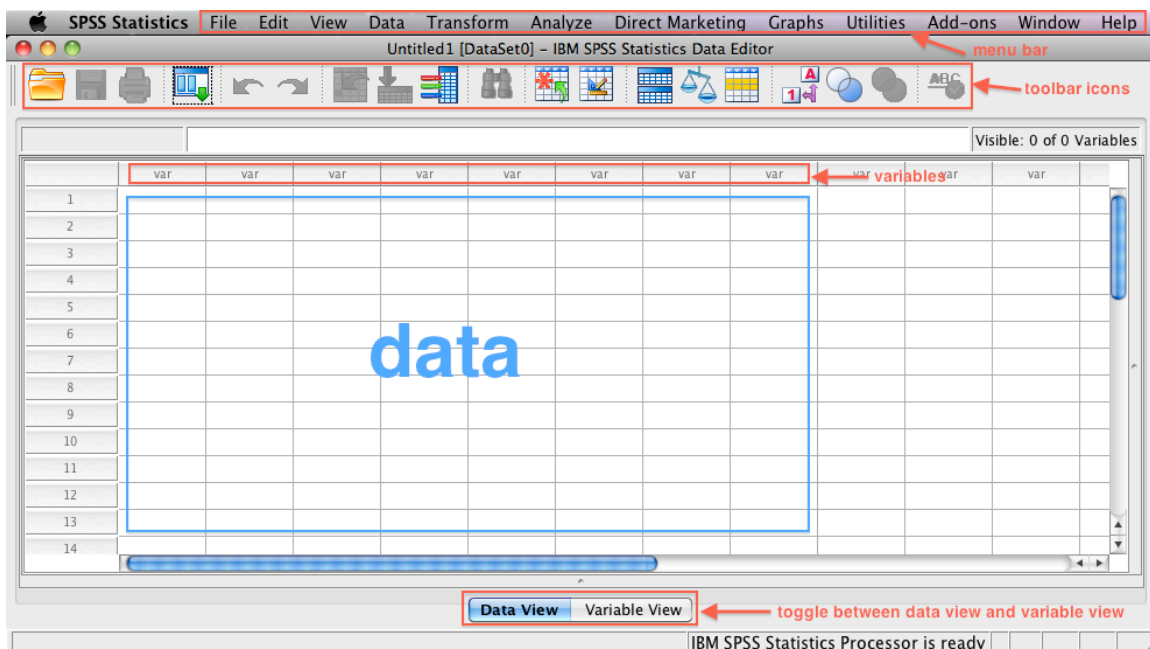
The Data Editor has two views that can be toggled near the bottom of the screen. **Data View** is where you can modify and see the contents of your data file. **Variable View** is where you can specify the format of your data.

The SPSS Menu Bar

1. **File**: similar to many other programs and includes options such as *Open, New, Save, Exit.*
2. **Edit:** includes the typical *Cut, Copy, Paste* as well as providing you a variety of options regarding how to display output generated with SPSS.
3. **Data:** allows you to select several options ranging from displaying data that is sorted by a specific variable to selecting certain cases for subsequent analysis.
4. **Transform:** includes several options to change variables (e.g., you can change continuous variables to categorical variables, change scores into rank scores, etc.)
5. **Graphs:** includes the commands to create various types of graphs

NOTE: The icons directly under the menu bar provide shortcuts to common commands found within the menu bar. If you hover your cursor over these items for a few seconds a description of the button command will appear.
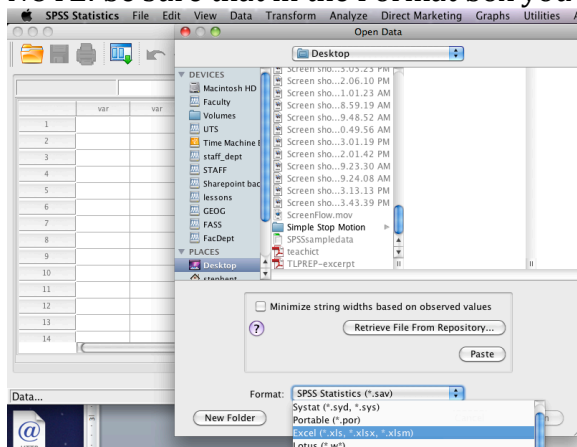
Opening the Sample Data File

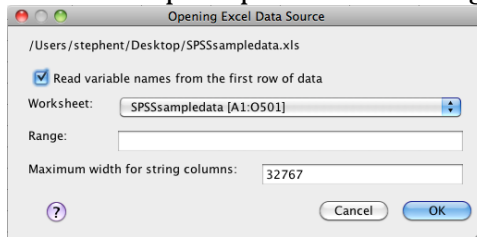1. **Click File → Open → Data ...**
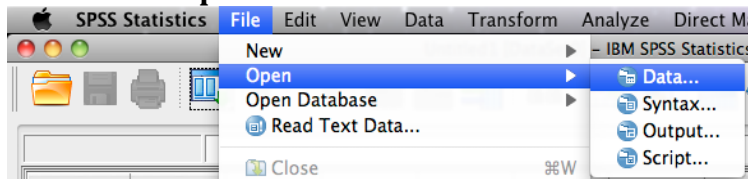
2. Browse to the location containing the SPSSsampledata.xls file that you downloaded earlier.
   NOTE: be sure that in the Format box you select Excel (*.xls)

3. **Click Open**

4. You will be prompted with a dialog similar to what you see below

Since this data has variable names in the first row of the file, there should be a tick in the box that states: Read variable names from the first row of data. **Click OK**
5. The data will import and you will be brought into to the data editor view with all of the data in the file displayed

Working in Variable View

The variable view contains information about the data set that is stored within the dataset

1. Toggle to the Variable View by **clicking the Variable View** button at the bottom of the screen.
2. Once in Variable View you should see something similar to the image below. The following is a description of the headings you see in Variable View

    a. **Name:** This is the variable name
        i. the first character must be alphabetic
        ii. variable names must be unique (e.g., no two rows the same)
        iii. variable name must be less than 64 characters
        iv. spaces are not allowed (you can use hyphens and underscores)
    b. **Type:** The most common types are
        i. Numeric: number
        ii. String: text
    c. **Width:** Allows you to define the number of characters SPSS will allow for the variable
    d. **Decimals:** The number of decimal places (must be between 0 and 16)
    e. **Label:** This is where you can specify the variable details (256 characters, including spaces)
    f. **Values:** This is used to suggest which numbers represent which categories when a variable represents a category (e.g., 0=Male, 1=Female)

Video walkthrough of the SPSS interface: http://vimeo.com/31788376

Exercise 1: Adding Labels and Value Labels

Often you will have raw data provided to you (e.g., from an online survey tool, from a larger database, etc.) that needs to be "tidied up" so it can be worked with more efficiently. In many cases you will first import your raw data into SPSS, then open the variable view to ensure that your data is labeled and configured in a manner that is easy for you to interpret your findings.

In this exercise you will
- add labels to variables
- create value labels

Background on the data file being used:

Variable	Label	Notes
ID		unique ID
VISUAL	Visualization Test	
MOSAIC	Patterns Test	
GRADES	Reported Grades	
MATHGR	Math Grades	
ALG	Algebra 1 Enroll	0=No, 1=Yes
ALG2	Algebra 2 Enroll	0=No, 1=Yes
GEO	Geometry Enroll	0=No, 1=Yes
TRIG	Trigonometry Enroll	0=No, 1=Yes
CALC	Calculus Enroll	0=No, 1=Yes
MATHACH	Math Achievement Test	
FAED	Father's Education	*see chart on right
MAED	Mother's Education	
SEX	Gender	0=Male, 1=Female
HEIGHT	Height in CM	

FAED/MAED Values:

Value	Category
2	Less than high school
3	High school graduate
4	Less than 2 years' vocational education
5	More than 2 years' vocational education
6	Less than 2 years' college/university education
7	More than 2 years' college/university education
8	College/University graduate
9	Master's degree
10	MD/PhD

Task:

Using the information provided, add the appropriate labels to variable and create value labels for SEX, FAED and MAED.

Step 1 – Adding Variable Labels
1. Toggle into variable view
2. Click on the appropriate cell in the *label column* and type in the label
3. Repeat this process for all the labels in the chart (above left)

Step 2 – Creating Value Labels
1. Toggle into variable view
2. Click on the appropriate cell in the *values column* and then click the … button to bring up the value labels dialogue
3. For the SEX Value Labels
   a. Enter **0** for the Value, and **Male** for the Value Label then **click Add**
   b. Enter **1** for the Value, and **Female** for the Value Label then **click Add**
   c. **Click OK**
4. Repeat this process for the FAED and MAED Values and Value Labels by referencing the chart (above right)

Video of the completion of this exercise: http://vimeo.com/31788964

Exercise 2: Recoding Data

There are many instances in which you will need to recode variables, such as changing values from letters to numbers, increasing or decreasing the number of possible groups, imposing cut-off scores, etc. SPSS allows you to recode variables and then use these recoded variables in statistical analyses.

Task:
Create 4 groupings (instead of 9) for the variables FAED (Father's Education) and MAED (Mother's Education) as listed in the table below (right).

Old FAED/MAED Values:

Value	Category
2	Less than high school
3	High school graduate
4	Less than 2 years' vocational education
5	More than 2 years' vocational education
6	Less than 2 years' college/university education
7	More than 2 years' college/university education
8	College/University graduate
9	Master's degree
10	MD/PhD

New FAED/MAED Values:

Value	Category
1	Less than high school
2	High school graduate
3	Some post-secondary education
4	College/University graduate & beyond

Step 1 – Recoding the Data into a New Variable
1. Click **Transform → Recode → Into different variables…**
2. In the Recode into different variables dialogue box you will see the list of variables on the left. **Click FAED**, and then **click the ➡ button**. You will notice that FAED appears in the right box.
3. In the Name field type **FAED_GROUP**
4. In the Label field type **Father's Education**
5. Next, **Click the Old and New Values button**
6. The **Old and New Values** dialogue box will appear
   a. In the old value box type 2, and type 1 in the new value box and then **click Add** (this will recode all the old value 2 into a new value 1)
   b. In the old value box type 3, and type 2 in the new value box and then **click Add**
   c. In the old value box **select Range,** and then enter 4 **through** 7, and type 3 in the new value box and then **click Add** (this will recode any old value of 4 through 7 into a new value 3)
   d. In the old value box **select Range,** and then enter 8 **through** 10, and type 4 in the new value box and then **click Add**
   e. **Click Continue**
   f. **Click Change**
   g. **Click OK**
7. Repeat steps 1-6 for the MAED variable
Step 2 – Adding Value Labels to the FAED_GROUP and MAED_GROUP variables
1. Revisit exercise 1 and add the appropriate value labels to each variable

Video of the completion of this exercise: http://vimeo.com/31789653

Exercise 3: Computing a Variable

In some instances you may need to use mathematical formulas to create new variables from your data (e.g., you might need to convert units, find the logarithm, etc.). SPSS allows you to write simple or complex algorithms to compute new variables.

Task:

Given the variable HEIGHT (cm), compute a new variable called HEIGHT_INCHES

Step 1 – Computing the Variable HEIGHT_INCHES

1. Click **Transform → Compute Variable → Into different variables...**
2. In the Compute variable dialogue box type in the name of the new variable: **HEIGHT_INCHES**

   NOTE: Since 2.54 cm = 1 inch, the formula for conversion from centimeters to inches involves the division by 2.54

3. From the variable box select HEIGHT and then **click the → button** to add this into the Numeric Expression box. Next click the **/** symbol and then select (or type) 2.54. The numeric expression should read **HEIGHT / 2.54**
4. **Click OK**

Video of the completion of this exercise: http://vimeo.com/31789942

Exercise 4: Frequencies and Descriptive Statistics

Often times you simply need to find out basic information about your data (e.g., frequencies, measures of central tendency, measures of dispersion, etc.). SPSS allows you to find this data easily using its built in descriptive statistics operations.

Task A:

Find the frequencies of occurrence for the following variables: GENDER, FAED_GROUP, and MAED_GROUP

Step 1 – Running the Frequencies report
1. Click **Analyze → Descriptive Statistics → Frequencies...**
2. In the frequencies dialogue box, select the variables you wish to analyze (i.e., GENDER, FAED_GROUP, MAED_GROUP) and then **click the ➡ button**
3. **Click OK**

Task B:

Find the mean, standard deviation, variance, minimum and maximum for the following variables: VISUAL, MOSAIC, and MATHACH

Step 1 – Running the Descriptives report
1. Click **Analyze → Descriptive Statistics → Descriptives...**
2. In the descriptives dialogue box, select the variables you wish to analyze (VISUAL, MOSAIC, MATHACH) and then **click the ➡ button**
3. **Click the Options... button**
4. In the descriptives options dialogue box, select the measures of central tendency and dispersion you wish to obtain (i.e., Mean, Std. deviation, Variance, Minimum, Maximum)
5. **Click Continue**
6. **Click OK**

Video of the completion of this exercise: http://vimeo.com/31790326

Pearson's correlation is used to find a correlation between at least two continuous variables. The value for a Pearson's can fall between 0.00 (no correlation) and 1.00 (perfect correlation). Other factors such as group size will determine if the correlation is significant. Generally, correlations above 0.80 are considered pretty high.

Task:

Determine if a significant correlation exists between VISUAL (visualization test), MATHACH (math achievement test), and GRADES (reported grades)

Step 1 – Perform Correlation Analysis
1. Click **Analyze → Correlate → Bivariate…**
2. In the bivariate correlations dialogue box, select the variables you wish to analyze (i.e., VISUAL, MATHACH, GRADES) and then **click the ➥ button**
3. **Click the Options… button**
4. **Check Means and standard deviations** (this will provide some descriptive statistics in the output)
5. **Click Continue**
6. **Click OK**
7. Consult a statistics book in order to classify the strength of the correlation

Video of the completion of this exercise: http://vimeo.com/31790734

Exercise 6: Crosstabs

In some situations you may need to find the descriptive statistics of two or more variables as a group (e.g., how many males over 55 inches took calculus). Crosstabs is an SPSS procedure that cross-tabulates two (or more) variables, thus displaying their relationship in tabular form. In contrast to Frequencies, which summarizes information about one variable, Crosstabs generates information about bivariate relationships.

Crosstabs creates a table that contains a cell for every combination of categories between the variables. Inside each cell is the number of cases that fit that particular combination of responses. SPSS can also report the row, column, and total percentages for each cell of the table.

Task:
Create a crosstab between GENDER and MAED_GROUP including the mean and standard deviation.

Step 1 – Perform crosstabs
1. Click **Analyze → Descriptive Statistics → Crosstabs...**
2. In the crosstabs dialogue box, select the variables you wish to analyze (i.e., GENDER, MAED_GROUP)
   a. Select GENDER from the list and **click the ➡ button** to place it in the **ROW** field
   b. Select MAED_GROUP from the list and **click the ➡ button** to place it in the **COLUMN** field
3. **Click the Cells... button** (this will allow you to include within group ROW and COLUMN data if required)
   a. **Select, Row, Column and Total percentages**
4. **Click Continue**
5. **Click OK**

Video of the completion of this exercise: http://vimeo.com/31791263

Further Information

Variable Types

In the Variable View you will noticed that there are three types of variables. Here is a description of each variable type and their appropriate use:

Nominal: A variable can be treated as nominal when its values represent categories with no intrinsic ranking (for example, the department of the company in which an employee works). Examples of nominal variables include region, zip code, and religious affiliation.

Ordinal: A variable can be treated as ordinal when its values represent categories with some intrinsic ranking (for example, levels of service satisfaction from highly dissatisfied to highly satisfied). Examples of ordinal variables include attitude scores representing degree of satisfaction or confidence and preference rating scores.

Scale: A variable can be treated as scale (continuous) when its values represent ordered categories with a meaningful metric, so that distance comparisons between values are appropriate. Examples of scale variables include age in years and income in thousands of dollars.

Notes