

A brief introduction to SPSS

The aim of this handout is to introduce you to the basics of the statistical application *SPSS for Windows*. *SPSS for Windows* is a simple but powerful statistical analysis and data management system. SPSS has a very exhaustive and clear help facility accessed by clicking the *Help* button from the *Applications Window* (see below). I suggest that you become familiar with this resource.

Starting SPSS

SPSS for Windows is activated by double-clicking on the SPSS icon. On PCs located at the Health Sciences campus, SPSS is activated from the *Start* menu. Click the *Start* button in the lower left hand corner of the screen, then choose *Programs, Applications, SPSS*.

The SPSS For Windows Environment

When *SPSS for Windows* is activated a dialog box will appear asking "*What would you like to do?*". For this class, we will usually select "*Open an existing file*" to load an already created SPSS file from a floppy disk or your hard disk (usually a data set, but perhaps an output or syntax file), or the *Cancel* button. If you choose "*Open an existing file*" from a PC used by many others (like those in the library or the Learning Center) you will most likely then need to select "*More Files...*" to select your file (this opens the "*Open File*" dialog box from which you can find and select your file). *The Data Editor* window, which initially appears in the background behind the dialog box, will become active after you select a file.

The Data Editor Window

At the top of the screen is the title bar with *SPSS for Windows Data Editor* written on it (prefaced by the name of an active data file or prefaced by "Untitled" if a file has not been loaded into active memory). Underneath the title bar is the menu bar with the headings *File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Window* and *Help*. The *File, Edit, View, Window* and *Help* menus are similar to those found in other Windows packages. The *Data* to *Utilities* menus provide the options to perform operations on the data and produce varying output.

To reveal the drop down menu for one of the options in the menu bar, you can either press the **Alt** key and the underlined letter in the heading or you can click on the menu heading using the mouse. Once the menu has been selected, the menu options are displayed. To select a menu option, click on the item using your mouse. The bottom line of the *Data Editor* window, the status bar, presents messages indicating what the package is doing. The initial message reads *SPSS Processor is ready*.

The *Data Editor* window has the appearance of a spreadsheet, with columns headed by the variable names of the active data set (or *var* if the window is empty) and rows numbered sequentially. Each column of the grid represents a single variable, and each row represents a single case. The spreadsheet in SPSS is more limited than a standard spreadsheet, as the cells can only contain values, not formula or functions. There are scroll bars along the bottom and down the right side of the window to allow access to the rest of the spreadsheet. You can only have one *Data* window open at any one time. At the bottom left of the window are two tabs: the *Data View* and the *Variable View* tabs.

Clicking these tabs allows you to toggle between viewing the spreadsheet (*Data View*), and viewing information about the variables contained in the spreadsheet (*Variable View*)

Other windows in SPSS include the *Output* and *Syntax* window. The first time any procedure is invoked during your session an *Output* will be created and the output of your procedure will be directed there. You can also open a new or existing *Output* file through the *Files* menu (you may have a number of output windows on the screen simultaneously). For example, you may have two *Output* windows open during a session. I would advise you to open a "*Draft Output*" window to collect your output. Your output may not look as nice, but you may run into some initial difficulties (like not being able to see all of your output) if you use the more highly formatted default *Output* window.) To Create Draft Output: from the menus choose: *File, New, Draft Output*. To make draft output the default output type, from the menus choose: *Edit, Options...*, click the *General* tab, select *Draft Viewer* under *Output Type at Start-up*. Draft output is saved as an rtf file.

Note: New output is always displayed in the designated Viewer window. If you have both a Viewer and Draft Viewer window open, the designated window is the one opened most recently or the one designated with the Designate Window tool (the exclamation point) on the toolbar.

You can designate which *Output* window will receive the output from the procedures that you are performing. This is done by selecting the desired window from the *Window* menu. It is identified by having a ✓ before its title, e.g. ✓ *2Output – SPSS Viewer*. SPSS for Windows also allows you to open files saved in previous sessions into the *Output* window for continued work.

The *Syntax* window allows for collecting a series of commands together into a file. We generally will not need to do this.

Any window (*Data Editor, Output, etc.*) can be made active, i.e. made to appear on the screen, by selecting the window from the *Window* menu.

Entering data

Data can be entered from the keyboard, from a text file, from an applications system file, or from an SPSS systems file. For now, I will make available to you SPSS system files that may be read in to the *Data Editor* from the initial dialog box appearing when SPSS is invoked or by selecting *File/Open* and then choosing the desired file from the appropriate folder.

Once your data are located within the *Data Editor* window you use SPSS to analyze the data. Most of the analyses are obtained by selecting the procedure from the menu in the *Application* window.

Clicking on the *Application* window *Analyze* menu reveals a drop-down menu with some or all of these options:

- Reports*
- Descriptive Statistics*
- Compare Means*
- General Linear Model*
- Correlate*
- Regression*
- Loglinear*
- Classify*
- Data Reduction*
- Scale*
- Nonparametric Tests*
- Survival*
- Multiple Response*

If an option has an arrowhead to its right, then there is a subsidiary menu. For example, if you selected *Descriptive Statistics*, a subsidiary menu containing these options will appear:

- Frequencies ...*
- Descriptives ...*
- Explore ...*
- Crosstabs ...*

Most statistical procedures in SPSS for Windows involve similar methods of selecting variables, i.e. by highlighting them and clicking on the selection (arrowhead) button. Most procedures also have associated with them a variety of options and statistics. We will discuss these as they become relevant.

As an example consider the “Hospital” data referenced in HW 1. These data are stored in a file named *hospital.sav* (*sav* is the extension reserved for SPSS system files). To perform the analyses necessary for problems 1-3 you could use the following procedures.

Analyze/Descriptive Statistics/Statistics (and then choosing the variable 'duration')—this gives descriptive statistics including the sample size, minimum, maximum, mean and standard deviation (other statistics are available by clicking the *Options* button in the dialog box for this procedure).

Analyze/Descriptive Statistics/Explore—produces a wide range of descriptive statistics and plots. To produce both descriptive statistics and plots select *Both* from the *Display* window. Select the statistics you want by clicking on the *Statistics* tab and then selecting the desired statistics (selecting *Descriptives* will provide the mean, standard deviation, variance, median, interquartile range and a few other statistics). Plots (notably histograms, stem and leaf and Box plots) may be selected from the *Plots* tab. Place the variables for analysis into the *Dependent List* window. You may obtain statistics and plots by levels of another variable by placing that variable in the *Factor List* window. For example to obtain descriptive statistics and plots for 'duration' by 'sex' place 'duration' in the *Dependent List* window and 'sex' in the *Factor List* window. To produce Box plots of 'duration' for each sex side-by-side, select *Factor levels together* from the plots tab.

Analyze/Descriptive Statistics/Frequencies—produces frequency distributions, descriptive statistics, percentiles, and plots (bar charts, pie charts, and histograms.)

Histograms may also be obtained from *Graphs*. Additional procedures will be described as necessary.

Leaving SPSS for Windows

You can leave by selecting *File /Exit* from any window. You will be prompted to save files.

The contents of any window may be saved through the *File* menu. For example to save the contents of the *Data Editor* choose *File/Save* to save the data as an SPSS system file, or choose *File/Save as* to save the data in another format (e.g. ASCII). In order to save the contents of a window that window must be current.