

# **SPSS 12**

## **Data Analysis Basics**

**Linda E. Lucek, Ed.D.**

**[LindaL@niu.edu](mailto:LindaL@niu.edu)**

**815-753-9516**

**Technical Advisory Group  
Customer Support Services  
Northern Illinois University  
120 Swen Parson Hall  
DeKalb, IL 60115**

---

# SPSS Data Analysis Basics

<b>Introduction.....</b>	<b>2</b>
<b>The SPSS Interface .....</b>	<b>2</b>
<b>Startup.....</b>	<b>3</b>
<b>Finding Help .....</b>	<b>3</b>
<b>Creating a New Dataset in SPSS.....</b>	<b>4</b>
<i>Defining Dictionary Information from Variable View.....</i>	<i>5</i>
<i>Entering Data Values from Data View.....</i>	<i>7</i>
<b>Saving a Dataset.....</b>	<b>8</b>
<b>Running Procedures from the Menus .....</b>	<b>8</b>
<i>Frequencies Procedure .....</i>	<i>9</i>
<i>Scatter Plot Procedure.....</i>	<i>9</i>
<i>Means Procedure.....</i>	<i>9</i>
<i>Grouped T-Test Procedure.....</i>	<i>10</i>
<b>SPSS Output Files.....</b>	<b>11</b>
<b>Retrieving Existing Datasets .....</b>	<b>12</b>
<b>Data Transformations and Selection.....</b>	<b>12</b>
<i>Sort.....</i>	<i>12</i>
<i>Compute.....</i>	<i>12</i>
<i>Data Subset Selection.....</i>	<i>13</i>
<b>Printing .....</b>	<b>14</b>
<b>Exiting SPSS .....</b>	<b>14</b>
<b>Licensing SPSS .....</b>	<b>14</b>
<b>Windows System requirements for SPSS version 12: .....</b>	<b>15</b>
<b>Mac OS X System requirements for SPSS version 11.01:.....</b>	<b>15</b>
<b>Documentation .....</b>	<b>15</b>

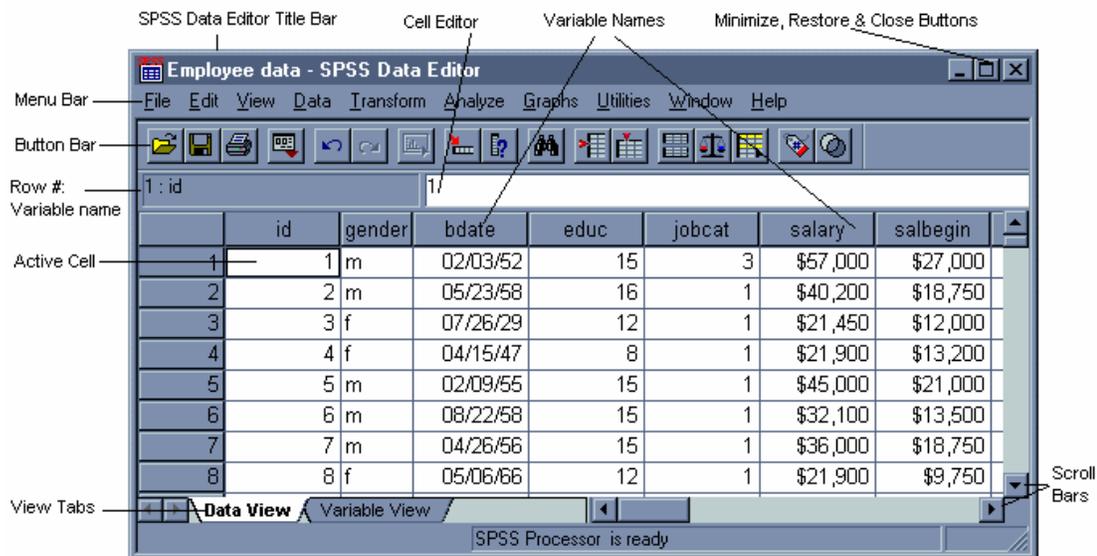
---

# SPSS Data Analysis Basics

## Introduction

Although SPSS stands for “Statistical Package for the Social Sciences,” this software is now widely used for data analysis and report generation across disciplines. SPSS for Windows or Mac OS features a spreadsheet-like data editor, integrated graphics, GUI interface, data management facilities (the number of variables and cases it can handle is limited only by the amount of RAM and disk space available), and comprehensive statistical procedures.

## The SPSS Interface



A spreadsheet-like data editor/viewer window opens up within the SPSS window frame. It is called the **Data Editor**. At this point you can either enter your own data into the *Untitled* data table or open an existing dataset.

SPSS typically stores data in its own proprietary format. A SPSS dataset saved from versions 10 or 11.x will have a file extension of *.sav*. Some other file types can also be read in -- such as SPSS portable files, various SAS formats, SYSTAT formats, dBase, MS Excel spreadsheets, Lotus spreadsheets, *.txt* and *.dat*.

## Hands-on Exercise

### Startup

To start the SPSS application, follow these steps.

1. From the Windows XP Desktop, select the following menu options:  
**Start > All Programs > SPSS for Windows > SPSS 12.0 for Windows**

SPSS starts up with an open Data Editor window and a dialog box with the query "What would you like to do?" From here your choices are to run a tutorial, type in new data, run an existing database query, create a new query, or open an existing dataset.

2. Click on the radio button next to *Type in data* and click on **OK**.

### Finding Help

The **FI** key can be used to get to context-sensitive Help Topics. Alternatively, on the top far right of the screen there is a Help menu which provides access to various help options including a tutorial (good for getting started), a link to the SPSS Home page on the Web, a Syntax Guide and Statistics Coach, and a version information screen, *About...* (this is where you can find the release # and site license #, which may be important to know if you are trying to get additional help).

1. Press **FI** key.
2. Click on the **Close** box in the upper right corner of the **Help** window when you are finished looking at the topics.

## Creating a New Dataset in SPSS

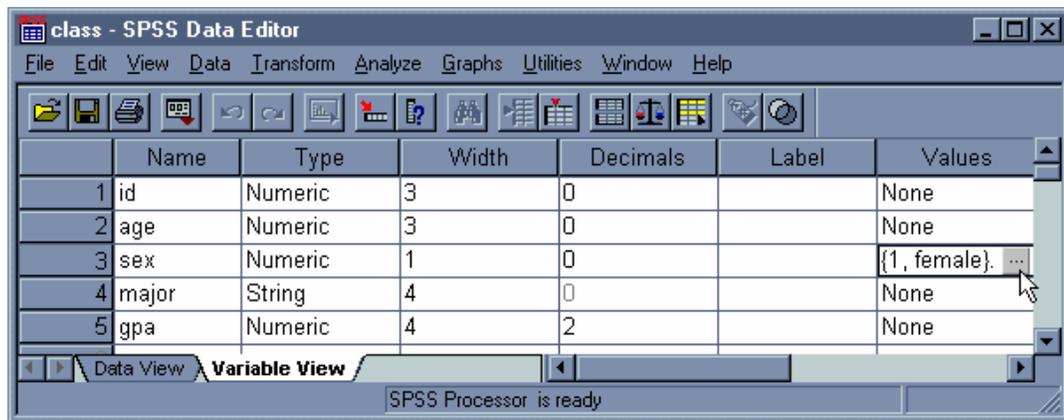
There are many ways to input the data you want to analyze. You can open existing SPSS data sets, open data from non-SPSS sources, or type in the values yourself. The data are displayed in a data table in which columns correspond to variables and rows correspond to cases. You can edit individual elements in the data table, and you can create new variables.

As an exercise, consider the following data to be the result of a small survey which you might wish to examine using SPSS. One individual's data represents a *case* in SPSS terms. The survey questions are considered *variables* and an individual's response to a particular question is the *data value*.

<u>id</u>	<u>age</u>	<u>sex</u>	<u>major</u>	<u>gpa</u>
1	19	1	psyc	3.55
2	19	2	csci	3.55
3	24	2	biol	4.0
5		1	biol	3.0
7	20	1	fcns	4.0
8	21	1		2.9
12	28	2	soci	3.7
13	20	2	phed	3.0

## Defining Dictionary Information from Variable View

1. To define the variables in the Data Editor, click on the Variable View tab at the lower left.
2. For the first variable, type in the variable name *id* and press the tab key.
3. The default type for a variable in SPSS is numeric with a width of 8 and 2 decimal places, but click in the cells and change these values as shown below. When you get to the variable named *sex*, you will find that SPSS will not allow you to change the Width to 1 until after you've changed the Decimals to 0.



We want to tell SPSS that the value "1" means "female" and the value "2" means "male" for this survey.

4. Click on the **Values** cell for the variable *sex* and then fill in the information for the first value as shown below.
5. Next, click on the **Add** button and enter a value of **2** with the value label *male*.
6. Click on **Add** once again and then on **OK**



The next variable, *major*, could also be coded numerically, but for the sake of practice, we'll treat it as a "string" or alphabetic variable and plan to enter the 4-character abbreviations used in class schedules.

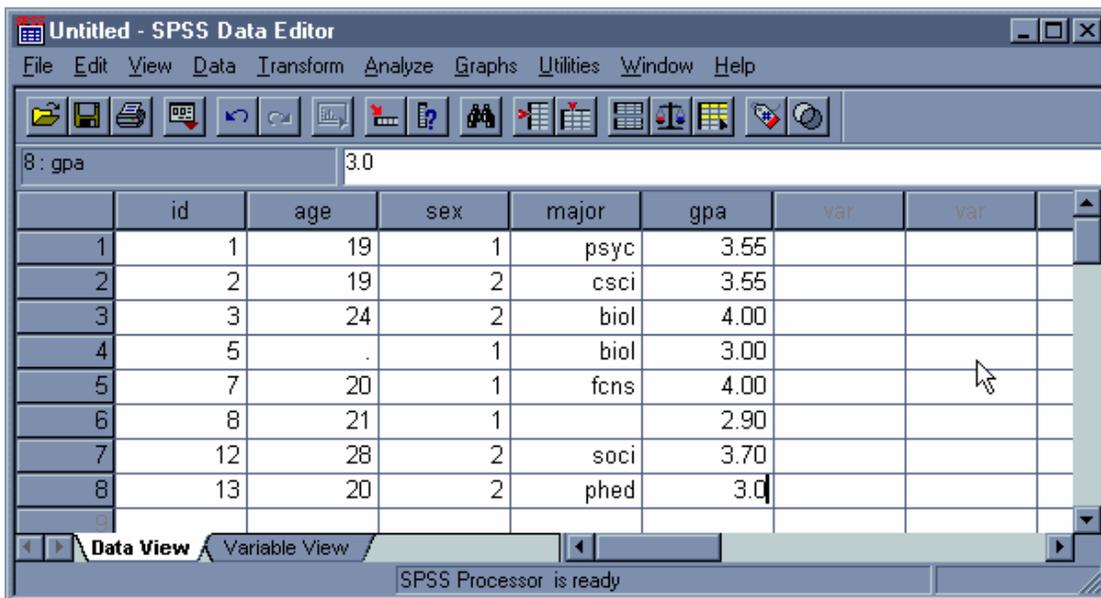
7. This time, after you type in the variable name *major*, change the variable Type by clicking on the dots in the *Type field* and selecting *String*.
8. Set the width to four characters and click *OK*.
9. Set up the next variable, *GPA* as a numeric variable with a width of four and two decimal places

This completes the data definition of all the variables in your dataset. To start entering the data values, click on the *Data View* tab.

## Entering Data Values from Data View

Generally, you start entering data in the first cell (upper left corner), typing in one data value at a time and pressing the **tab** key after each. This moves the next entry point across to the next variable (column) on the same case (row). This would be the way you would want to enter data if you were taking it right off the survey. However, if you have the numbers already laid out in columns like we do here, you may find it easier to enter all the cases for one variable at a time. You do that by pressing the **Enter** key instead of the **tab** key. This moves the insertion point down to the next row (case) instead of across to the next column (variable).

1. Click in the first table cell (upper left-hand corner, under *id* beside case # 1).
2. Enter the first id #.
3. Press the **tab** key to move the insertion point to next variable.
4. Continue entering new data values to match the table shown below.



The screenshot shows the SPSS Data Editor window titled "Untitled - SPSS Data Editor". The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Window, and Help. The toolbar contains various icons for file operations and data manipulation. The main window displays a data table with the following columns: id, age, sex, major, gpa, var, and var. The data is as follows:

	id	age	sex	major	gpa	var	var
1	1	19	1	psyc	3.55		
2	2	19	2	csci	3.55		
3	3	24	2	biol	4.00		
4	5	.	1	biol	3.00		
5	7	20	1	fcns	4.00		
6	8	21	1		2.90		
7	12	28	2	soci	3.70		
8	13	20	2	phed	3.0		

The status bar at the bottom indicates "SPSS Processor is ready".

When you enter values for the variable *major*, be consistent about whether you use upper or lower case. Either is acceptable, but you will need to use the same case later when you select for particular values. When you have finished entering all the survey values for one case (person), click the mouse in the first cell of the second row, to start entering data for the second case. When you reach points where there is missing data, just press the **tab** key to move on to the next cell. A dot (.) appears to indicate that the value for that variable for that particular case is missing (except for string variables such as *major*).

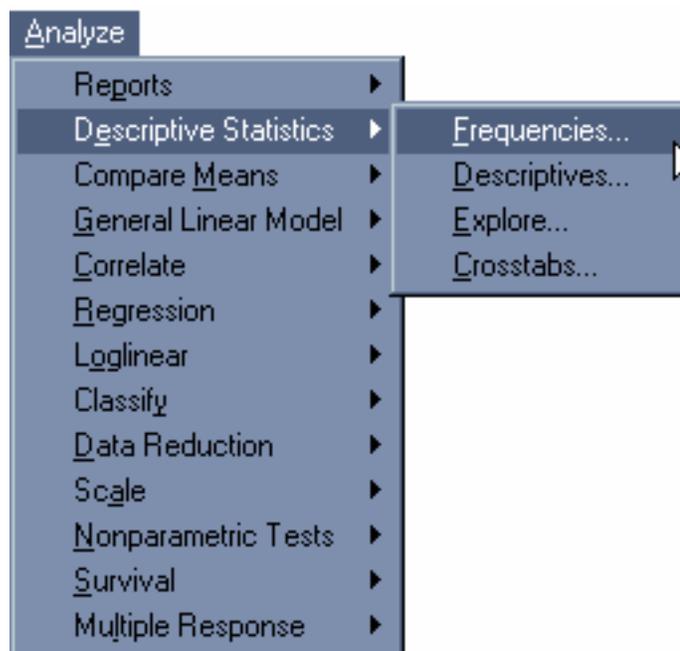
When you are finished entering the values, your dataset is *active* in memory and ready to use during the session, but in order to use it again at a later time, you will need to give it a file name and save it.

## Saving a Dataset

To save your data out as an SPSS system file, you will need to give the file a name. You should also check the location to see if you want to save it in a different directory or drive. To save to the floppy drive, make sure the diskette is in the drive.

1. From the SPSS menus, select **File > Save**
2. Next to File Name, type: *a:class*
3. Click on the *Save* button.

## Running Procedures from the Menus



Once the data to be analyzed are in the data table, you can access tasks and procedures from pull-down menus.

For example, to get counts and percentages for each value of a categorical variable, you run a frequencies procedure.

## Frequencies Procedure

1. From the SPSS menus, select **Analyze > Descriptive Statistics > Frequencies...**

This opens up a dialog box with your variables listed on the left.

2. Shift-click on the variable names *age*, *sex*, and *major*. Then click on the *right arrow* to move the selected names to the Variable List.
3. Click on **OK** to run the procedure.

The procedure output goes into a new window.

## Scatter Plot Procedure

Try submitting the following command, which provides a plot of *age* with *gpa*:

1. From the SPSS menus, select **Graphs > Scatter... > Simple**
2. Click the *Define* button
3. Select *gpa* for your X Axis and *age* for the Y Axis
4. Click on **OK** to run the procedure

## Means Procedure

Next try the **Means** procedure.

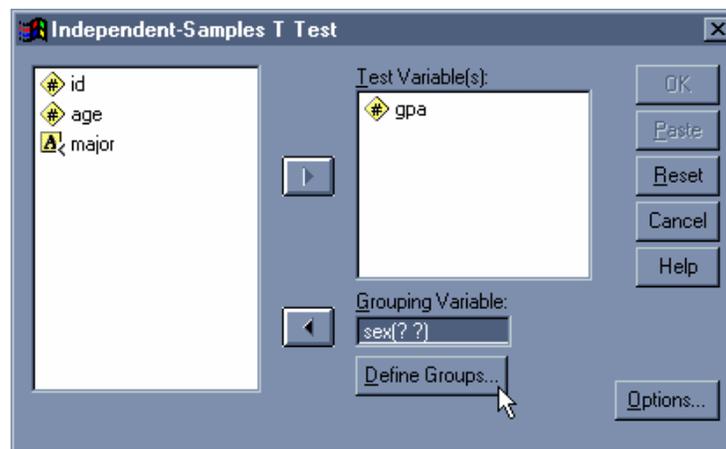
1. From the SPSS menus, select **Analyze > Descriptive Statistics > Descriptives...**
2. Select one or more appropriate variables (there is not much point in trying to get a mean for variables like *id* or *sex*)
3. Take a look at the options available through the *Options...* button
4. Click on **OK** to run the procedure.

Try running the same procedure a few times selecting different statistics.

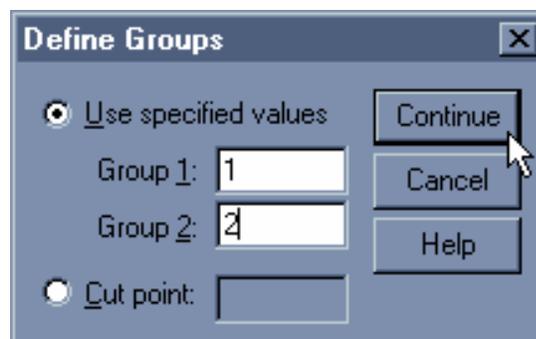
## Grouped T-Test Procedure

Next we'll try a grouped t-test, another commonly-used statistical procedure.

1. From the SPSS menus select **Analyze > Compare Means > Independent-Samples T Test**
2. Select a continuous variable (e.g., *gpa*) as your test variable and a categorical variable (e.g., *sex*) as your grouping variable, as shown in the following example.



3. You will also have to specify the values to be used for your grouping variable, by typing them between the parentheses or by clicking on the button labeled **Define Groups...** and filling in the boxes as shown below; then click on the **Continue** button.

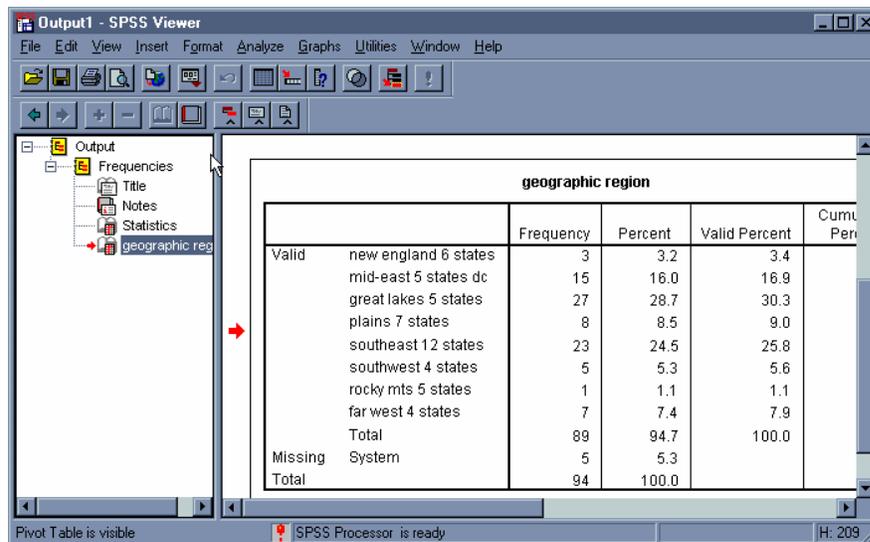


4. Click on **OK** to run the procedure.

## SPSS Output Files

SPSS has an Output Navigator, called the SPSS Viewer, in which procedural output is displayed as pivot tables. Pivot tables are OLE 2.0 objects which can be displayed and manipulated individually by SPSS with a great degree of flexibility.

In the Window shown here, pieces of the SPSS output are shown in outline form on the left. They can be individually or collectively selected by clicking on the icons. In this instance, the output table labeled "geographic region" has been selected.



		Frequency	Percent	Valid Percent	Cumu Per
Valid	new england 6 states	3	3.2	3.4	
	mid-east 5 states dc	15	16.0	16.9	
	great lakes 5 states	27	28.7	30.3	
	plains 7 states	8	8.5	9.0	
	southeast 12 states	23	24.5	25.8	
	southwest 4 states	5	5.3	5.6	
	rocky mts 5 states	1	1.1	1.1	
	far west 4 states	7	7.4	7.9	
	Total	89	94.7	100.0	
Missing	System	5	5.3		
Total		94	100.0		

To put this table into a Microsoft Word document just as it appears in SPSS, you can click anywhere inside the table, choose **Copy objects** from the Edit menu, move to your word processing document, and choose **Paste** from the Edit menu there. Essentially what gets pasted is a graphic clip, so the text in the table can no longer be edited easily. If you would like to change the titles or labels, you will be better off trying to do that from within SPSS. If you would rather just copy and paste the text from the table (and not include the borders), choose **Copy** instead of **Copy objects**.

You can, of course, print your original tables right from SPSS, by choosing **Print** from the File menu when the Output Navigator is active. You can also edit them within SPSS. A pivot table in the SPSS Output Navigator is displayed as a Windows Metafile. Changes can be made to it by first activating it (e.g., Right-click on the table in the Output Navigator and choose Edit from the SPSS Pivot Table Object submenu OR simply double-click on the table itself). After changes are made, select Exit from the Pivot Table Editor menu to return to the Output Navigator.

Output saved by choosing **Save** from the File menu is not in regular text format. It is saved as an **.spo** file, also sometimes called a Navigator file. That is, by default it will have an extension of **.spo** (presumably for "SPSS Output" file). Opening such a file from Windows Explorer will automatically start SPSS and open the Navigator, displaying the output file.

## Retrieving Existing Datasets

1. From the SPSS menus go to **File > Open > Data**
2. A window titled “Open File” appears which allows you to select file name and file type to open. Select the dataset named *school.sav* from your a: drive. (The extension probably will not be displayed.)

The default file type to read in is a *.sav* file (SPSS system file), but if you have another file type: *sys* (SPSS/PC+), *por* (SPSS portable format), *xls* (Excel), *slk* (SYLK), *dbf* (dBASE), *w\** (Lotus) or *dat* (tab-delimited), you can choose the correct one from the *Files of type* selector by pressing the down arrow to change the selection.

## Data Transformations and Selection

### Sort

This data is from a survey of 94 High School seniors. The cases were not entered in order by their id numbers. To sort them by *id*:

1. From the SPSS menus, select the following options: **Data > Sort Cases**
2. Double-click on the *id* variable.
3. Click on **OK**. The cases are sorted in ascending order by the *id* variable.

### Compute

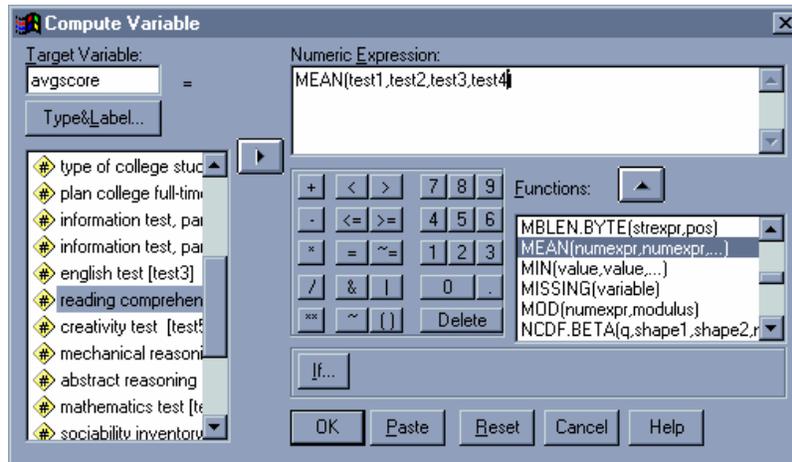
You can create new variables based on arithmetic and statistical functions. For example, working with the current dataset which includes various test scores, you might want to create a new variable which represents the mean across the test scores.

(Keep in mind that exercise is for practice only; in this case it would not really be appropriate to get a raw mean score across these tests since they are based on different scales).

1. From the SPSS menus, select the following options: **Transform > Compute**

We will create a new variable named *avgscore*. For each high school senior, *avgscore* will be the average of their first four test scores (test1, test2, test3, and test4).

2. Enter the name of the Target variable as *avgscore* as shown below.



Although we could use arithmetic operators to derive the average, SPSS also has statistical functions available.

3. From the Functions listing on the right, select the *Mean* function by double-clicking on it. It appears in a formula box above (under *Numeric Expression:*).
4. Highlight the area between the parentheses and replace the text with: *test1,test2,test3,test4*
5. Click on *OK*.

The new variable, named *avgscore*, appears in the data editor to the far right. You will have to scroll to the right to see it.

## Data Subset Selection

1. From the SPSS menus, select the following options: **Data > Select Cases...**
2. Click on the radio button next to: *If condition is satisfied.*
3. Click on the *If* button.

4. From the variable picklist on the left, select the variable **Region**.
5. Next to **Region**, type: =2
6. Click on **Continue**.
7. Click on **OK**

The unselected cases are now displayed in the data editor with slash marks across the case numbers. If you were to select any new statistical procedures from the menus, they would only be performed on the active subset (where region equals 2).

To return to the complete dataset, turn the filter off .

8. From the SPSS menus, select the following options: **Data > Select Cases... > All Cases** (radio button)

## Printing

Make sure the window you want to print from is active (select it by clicking on it). Then choose Print from the File menu. You have the choice of printing only what is highlighted (selected) from the active window or printing the entire contents of the window.

## Exiting SPSS

You can switch to other Windows applications while SPSS is running by pressing Alt-Tab or clicking on another visible icon. To properly close an SPSS session, choose Exit from the File menu OR click on the close box of the Data Editor OR press Alt-F4 when the Data Editor window is active.

## Licensing SPSS

Information Technology Services at Northern Illinois University has a site license for distributing SPSS for Windows or Mac OS to NIU faculty and staff. The license includes the SPSS Base software, Regression Models, and Advanced Models products. Call Software Distribution at 753-7600 to obtain a subscription form. Students can purchase the SPSS Grad Pack at the Holmes Student Center Bookstore or rent the software from <http://www.e-academy.com/spss>

## Windows System requirements for SPSS version 12:

**Processor (CPU):** Pentium or Pentium-class processor running at 90MHz or faster.

**Memory (RAM):** 128MB minimum

**Display type:** SVGA (800 x 600 resolution) or higher

**Hard Disk Space:** Minimum of 220MB available disk space

**Media:** CD-ROM

**Mouse:** Required

**Platform:** Windows 98/Me/NT4 with SP6/2000/XP

## Mac OS X System requirements for SPSS version 11.01:

**Processor (CPU):** **G3 233 MHz processor or faster**, 350MHz recommended  
(optimized for G4 with Velocity Engine)

**Memory (RAM):** **192MB RAM** (random-access memory), 256MB or more recommended

**Display type:** Color monitor with 800x600 resolution. 1024x768 or higher recommended.

**Hard Disk Space:** 250MB free hard disk space for full installation  
(150MB for application only), including 3MB on System disk.

**Media:** CD-ROM drive (for installation only)

**Mouse:** Required

**Platform:** Mac OS X 10.1.2 or later

## Documentation

The SPSS 12 Documentation CD includes the following manuals in pdf format.

SPSS Advanced Models 12.0

SPSS Base User's Guide 12.0

SPSS Brief Guide 12.0

SPSS Categories 11.0

SPSS Complex Samples 12.0

SPSS Conjoint 8.0

SPSS Exact Tests 7.0

SPSS Interactive Graphics 10.0

SPSS Maps 10.0

SPSS Missing Value Analysis 7.5

SPSS Regression Models 12.0

SPSS Tables 11.5

SPSS Trends 10.0

Hardcopy SPSS Manuals can be purchased at NIU's Holmes Student Center Bookstore .  
Additional SPSS 12 Manuals on CD or hardcopy can be purchased from the on-line SPSS store  
at: <http://www.spss.com/estore/softwaremenu/index2.cfm>

Those interested in doing SPSS programming should note that the SPSS 12 Command Syntax Reference Guide is installed with SPSS and is available from the SPSS Help menu.

For those who need a basic introduction to data analysis methods and introductory statistics, the following Prentice Hall guide is well written and easy-to-follow, with examples of processing survey results using SPSS.

*SPSS Guide to Data Analysis* by Marija Norusis