

Intro to SPSS

SPSS works like this:

You have a data set (either you create one or use an existing file such as the GSS).

You choose analysis techniques for the data using dialog boxes, then click PASTE.

Your analysis command is pasted into the Syntax window.

You go to the Syntax window and run the command.

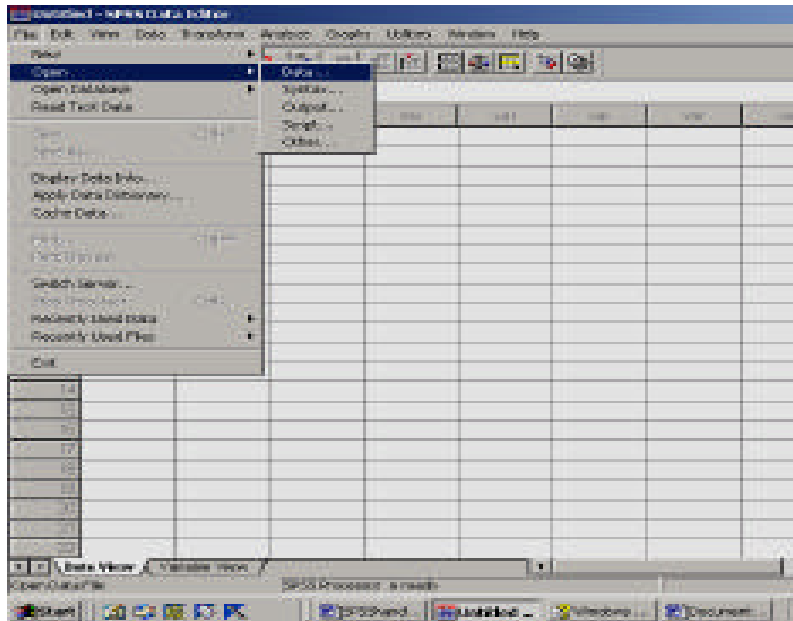
Your analysis results are shown in the Output window.

- Opening SPSS – SPSS is a windows based statistical package that is opened like any windows application- through the start menu or by double clicking the icon on the desktop/toolbar. SPSS is available on all of the SIPA lab computers.
- There are 3 main windows, and 3 main types of files, in SPSS: **Data, Syntax, and Output.**

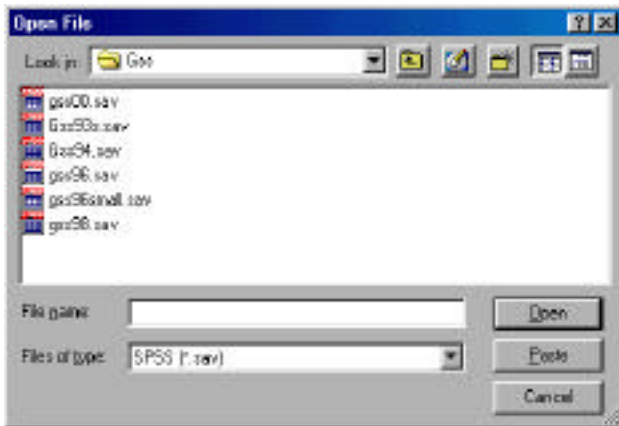
Data

- Existing data files (like the GSS data file) can be opened using the dialog boxes. Later in the year you'll be creating your own, new data sets by typing in data and setting up variables.
- SPSS only allows one data file to be open at a time.

OPENING A DATA FILE



- **Click on File, Open,**
- **then Data.** A list of data files will appear. The program defaults to the most recently used folder- which may or may not be the GSS folder on the SIPA server. Finding the GSS survey is explained below under **“Opening the 2000 GSS data set”**. Double-click on the one you want- In our case gss00.sav. It's in the O drive.



- Data files always have the extension “.sav”. The GSS is a (very big) data file. Notice the *data view* and *variable view* tabs at the bottom of the screen. In *Data View* the numbers in the farthest left column represent cases (i.e., survey respondents). The top row shows names of variables (i.e., survey questions). *Variable View* shows information about each variable, such as its name, values, number of decimal places, etc.

Syntax

Syntax files are records of your commands given to SPSS to manipulate/analyze the data. The Syntax window shows the commands you’ve given to SPSS.

When you analyze the data (by running descriptive statistics, recoding variables, etc.), always use **PASTE** instead of OK. This will literally paste your command into the Syntax window. You can then go to the syntax window, select the command you gave it, and **click the right arrow button** on the toolbar in the syntax window to **run the command**.

You can save syntax files if you anticipate re-running a particular type or series of analyses. SPSS allows for multiple syntax files to be open at the same time.

Saving your variables:

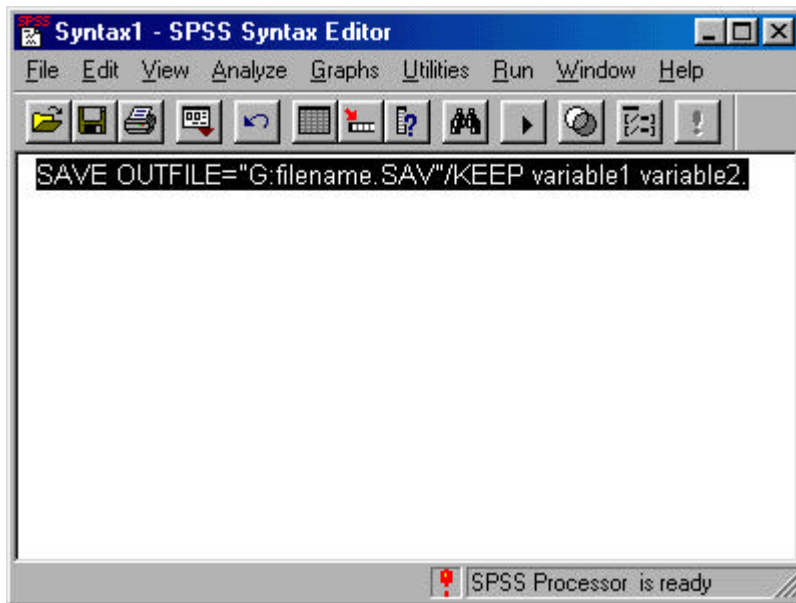
Because the GSS data set is so large it is useful to first save (or extract) only the variables you will be working with to your own SIPA drive (the G drive). First open a new syntax file:

Click on **File** on the menu bar at the top of the screen.

Go to **New**.

Click on **Syntax**.

A syntax window will appear.



Type into the Syntax window what you see in the box to the left. Instead of 'filename', type what you'd like this dataset to be called. Instead of 'variable1 variable2', type the mnemonic names of the variables (such as EDUC, INCOME, RACE, etc).

Select what you've typed and click on the arrow (underneath Run).

Now you've created a new data file containing only the variables you named. In order to work with this new file, you must open it. Go back to the data window, go to File, Open, Data, and then find your new file in your G drive.

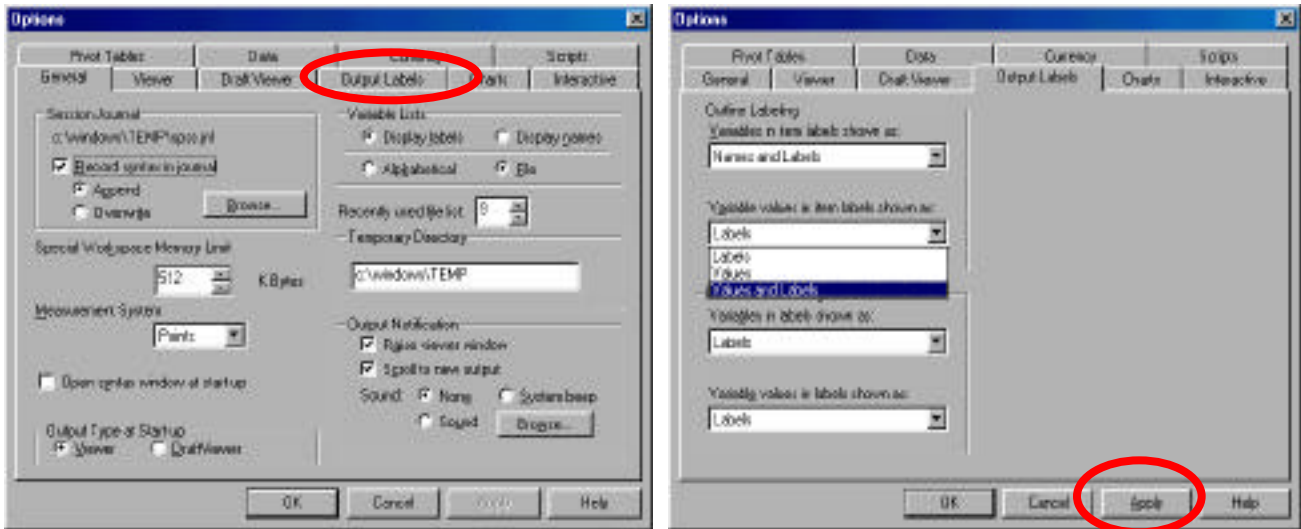
Output

SPSS automatically creates an output file when you run a command that generates output. The Output window shows your analysis - tables, graphs and charts.

In order to make reading your output easier, first change a couple of the default output settings.

Click on **Edit** on the menu bar at the top of the screen.
Go to **Options**.

| | wristat | hrs1 | hrs2 | awwa | |
|----|---------|------|------|------|----|
| 6 | 1 | 40 | -1 | | |
| 7 | 1 | 38 | -1 | | |
| 8 | 7 | -1 | -1 | | |
| 9 | 1 | 60 | -1 | | |
| 10 | 1 | 44 | -1 | | |
| 6 | 2000 | 6 | 1 | 40 | -1 |
| 7 | 2000 | 7 | 1 | 55 | -1 |
| 8 | 2000 | 8 | 1 | 42 | -1 |
| 9 | 2000 | 9 | 1 | 60 | -1 |
| 10 | 2000 | 10 | 1 | 60 | -1 |
| 11 | 2000 | 11 | 1 | 50 | -1 |



In the Options Dialogue Box click on Output Labels. SPSS sets the default to Labels, you will want to pull down the menu and **select Names and Labels for all four boxes**. Click apply button.

Tip: For presentation and printing, it is often easier to cut and paste the graphs and charts you want from the Output window on to a regular Word document. To do this, select the graphs and charts you want (you can use the outline on the left hand side of the screen), click on Copy Objects (Cntrl-k, or from the toolbar at the top of the screen under Edit), then go to a Word document. Click on Paste (Cntrl-v, or from the toolbar at the top of the screen under Edit).

To Switch between Data, Output, and Syntax windows just look on the bottom of the screen. All open windows are there.

Opening the 2000 GSS data set in SPSS

For this class we will be working extensively with the GSS (General Social Survey which is administered annually by the University of Chicago's National Opinion Research Center-NORC). The GSS 2000 data currently resides on the SIPA file server or O drive.

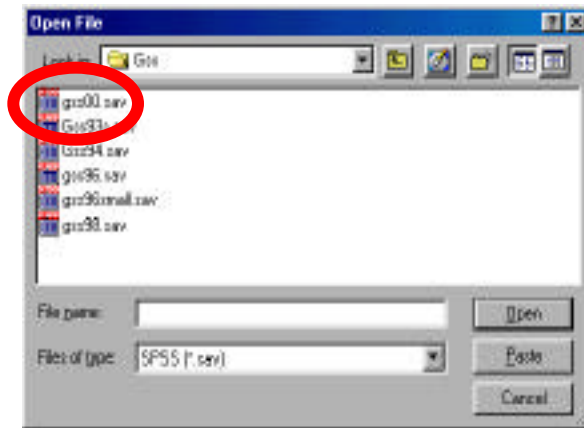
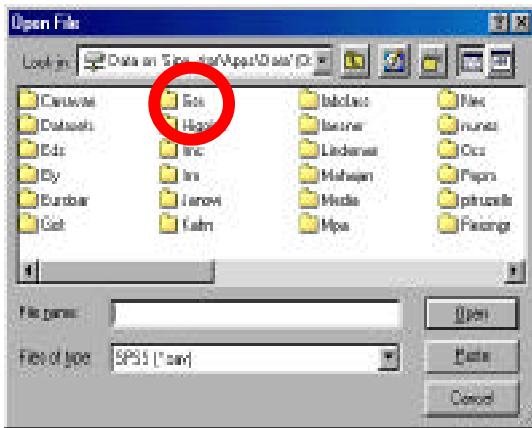
As above:

Click on **File** on the menu bar at the top of the screen.

Go to **Open**.

Click on **Data**.

In the OPEN FILE Dialogue box- click to the SIPA file server or O drive. Double-click on the GSS folder then double click on the GSS00.sav icon.



Running Frequencies

Once you have opened the GSS data set you can run any type of analysis you want. We will start with the most basic kind, frequencies. Frequencies show how many times each value appears for a variable in a dataset. For example, in a dataset there are 30 cases. 20 are men and 10 are women. If we ran frequencies on the variable GENDER, we would see that the value for “male” appears 20 times in the dataset, and the value for “female” appears 10 times in the dataset.

You can also get more descriptive statistics in your output window when you run frequencies, including mean, median, mode, range, and standard deviation, all relating to a specific variable.

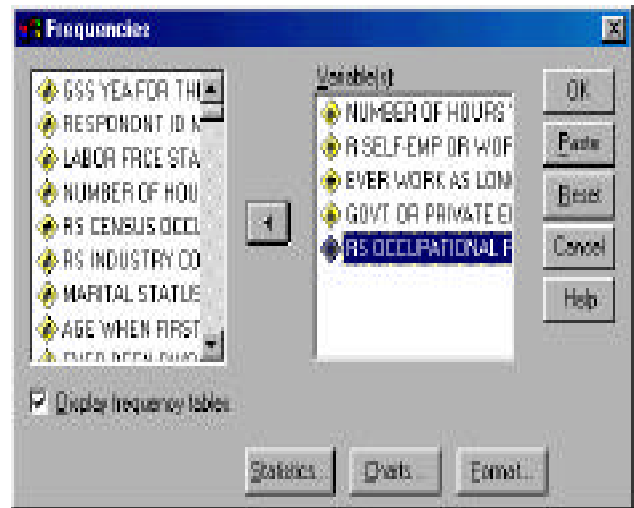
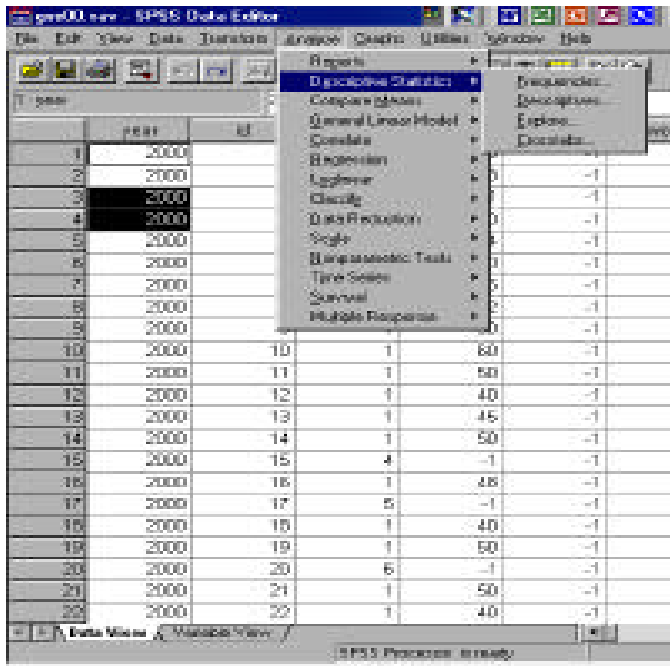
How to Run Frequencies

Make sure you're in the Data window (the main spreadsheet).

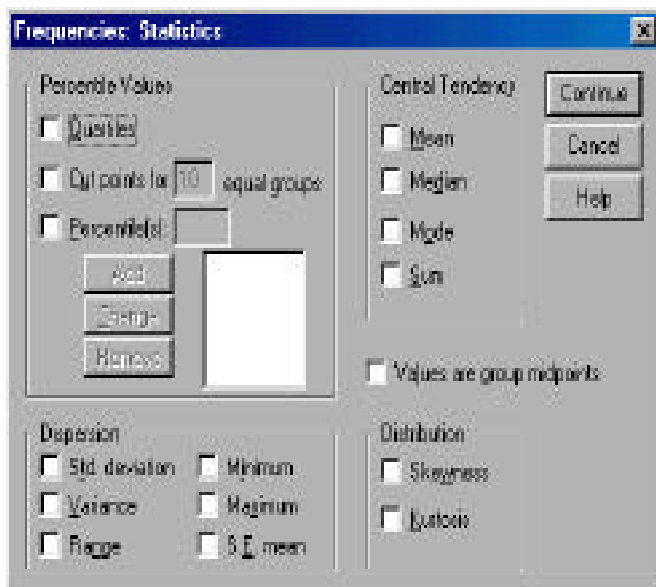
Click on **Analyze** from the top of the screen.

Go to **Descriptive Statistics**.

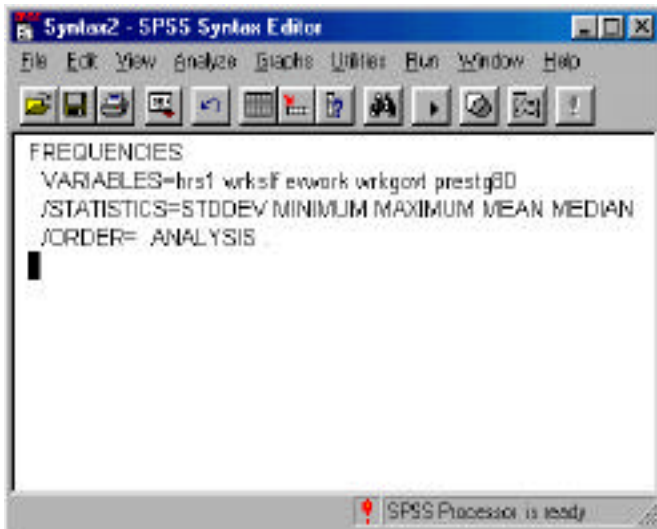
Click on **Frequencies**.



In the Frequency box, select the variable(s) you wish to analyze from the left hand list. You can select more than one at a time by holding down the Ctrl button. Press the arrow in the middle and the variable will move to the right hand box. **Double click on the Statistics button** and the dialogue box below appears.

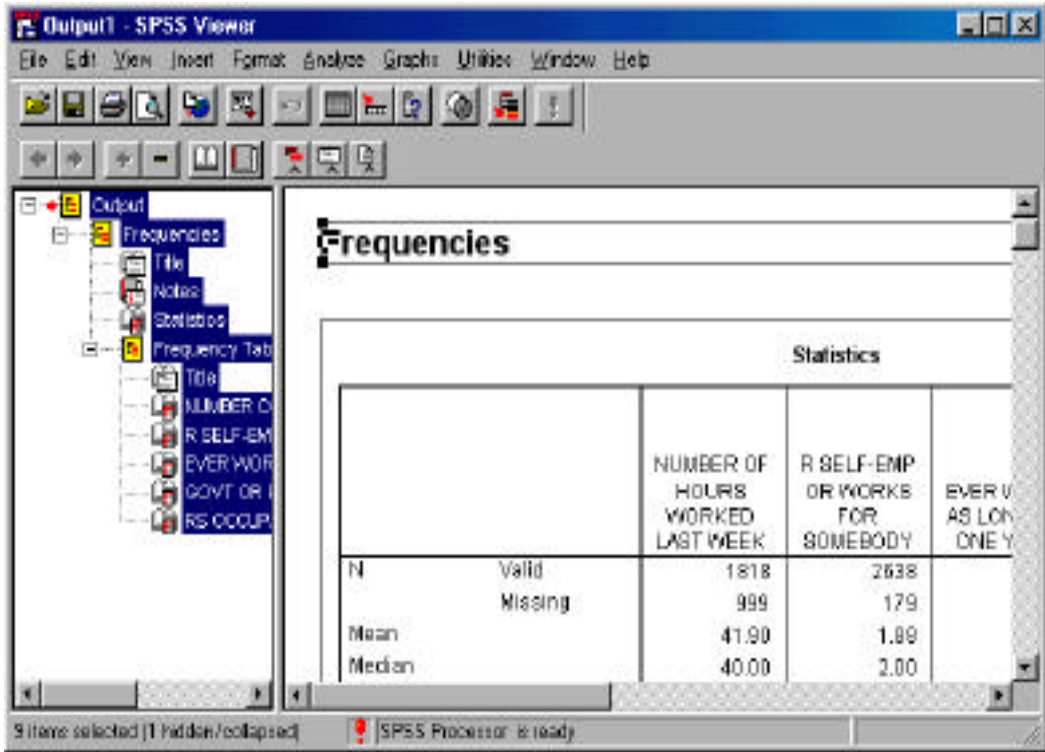


Here you can choose a number of statistics to be run on the variable(s) you have chosen. As you learn the basics it will become apparent which statistics are appropriate for which variables.



By pressing paste- SPSS will paste the Frequency command into a syntax file.

Highlight the whole command, and click the arrow underneath Run. Running this command produces the following output file.



We will go over how to interpret all of this output in detail. Below is an example of the statistics that we generated.

Statistics

| | | NUMBER OF HOURS WORKED LAST WEEK | R SELF-EMP OR WORKS FOR SOMEBODY | EVER WORK AS LONG AS ONE YEAR | GOVT OR PRIVATE EMPLOYEE | RS OCCUPA TIONAL PRESTIG E SCORE (1980) |
|----------------|---------|---|---|-------------------------------------|--------------------------------|--|
| N | Valid | 1818 | 2638 | 862 | 2678 | 2660 |
| | Missing | 999 | 179 | 1955 | 139 | 157 |
| Mean | | 41.90 | 1.88 | 1.15 | 1.85 | 44.05 |
| Median | | 40.00 | 2.00 | 1.00 | 2.00 | 43.00 |
| Std. Deviation | | 13.40 | .32 | .35 | .36 | 14.15 |
| Minimum | | 3 | 1 | 1 | 1 | 17 |
| Maximum | | 89 | 2 | 2 | 2 | 86 |