

May 2005

Guide to ICPSR Direct Web Downloads for Windows

The Inter-University Consortium for Political and Social Research (ICPSR) at the University of Michigan serves member colleges and universities in the U.S. and abroad. The ICPSR provides access to the world's largest archive of computerized social science data.

Most ICPSR data are restricted to Northwestern students, faculty or staff members. This document assists you with downloading and using data from the ICPSR. You may download ICPSR data from any computer connected to the campus network. From off-campus, access is enabled via proxy service or VPN. For help, see <http://www.library.northwestern.edu/help/proxy/>.

Files downloadable from the ICPSR can be classified as documentation and data files.

- Documentation files include a codebook file, a questionnaire file, a data map or data dictionary file, and a data definition file (SPSS and SAS statements).
- Data files contain raw data that need further processing to be used in a statistical package program. Portable or transport files can be used without additional processing.

File types can easily be identified by filenames as shown in the following file list from the study, "SETUPS: Voting Behavior, The 1996 Election" – ICPSR study number 2000.

```
da2000.txt [raw data file]
sa2000.sas [SAS data definition file]
sp2000.sps [SPSS data definition file]
sp2000.por [SPSS portable file] this file is only an example
```

The following steps show you how to create an SPSS portable/system file from a raw data file and an SPSS data definition file on a PC, using the above study as an example. For more detailed help, see <http://webapp.icpsr.umich.edu/cocoon/ICPSR-FAQ/0062.xml>.

STEP 1

Download da2000.gz and sp2000.gz from the ICPSR website by following website instructions. Save the files in a local directory (C:\WORK\ in this example). Decompress both files using a zip program such as WINZIP. Do not save the data definition file (sp2000) with the ".txt" extension because SPSS for Windows will try to read it as a data file: *you must change it to a ".sps" extension*. (If downloading a portable file, some computers will show it with a ".txt" extension which *must be changed to ".por"* .)

STEP 2

Open sp2000.sps in SPSS for Windows [*File* → *Open* → *Syntax*]. Most ICPSR data definition files start with a header that describes the contents of the file, followed by SPSS command lines (See *Example*). Note that command lines with an SPSS comment marker (*) are not run. If the

MISSING VALUES section is written with a comment marker, these values will not be run. Delete the comment marker to retain missing values as originally assigned in the data.

STEP 3

After opening sp2000.sps, scroll to the DATA LIST command. Modify it with the full path and filename of the data file you downloaded from the website.

```
DATA LIST FILE = 'C:\WORK\da2000' records= 1
Note: card-image data requires that you specify the number of records per
case. Example: 'C:\WORK\da7732' records= 6
```

To subset data for only the variables you want, select and edit your chosen variables in each command section. (*Example below selects 4 variables.*) Finally, scroll down to the end of the data definition file and add the following command to save the resulting SPSS portable or system file in a local directory:

```
EXPORT OUTFILE= 'C:\WORK\mydata.por' . [Or]
SAVE OUTFILE='C:\WORK\mydata.sav' .
```

STEP 4

Run sp2000.sps by clicking on [*Run*→*All*]. When SPSS has completed reading the commands, the status bar at the bottom of the screen will show *SPSS for Windows Processor is ready*. When the processor is finished, you can see the data by opening the data window [*Window*→*SPSS Data Editor*]. Any error message will be printed in a log file in the output window [*Window*→*SPSS Output Window*].

Edited Example: sp2000.sps

```
/* SETUPS VOTING BEHAVIOR: THE 1996 ELECTION */
* SPSS file definition statements
* The data list statement assumes the raw data file is on the C drive.
data list file='c:\work\da2000' records=1
/caseid 1-4 weight 5-10(4) v001 to v161 11-171
v201 to v215 172-216 v216 to v218 217-219 v219 220-221
v220 222-225.
variable labels
v001 "Voted in Election"
v002 "Presidential Vote"
v003 "Presidential Vote in 1992"
v004 "House Vote"
(...)
value labels v001 1 "1. Yes" 2 "2. No"/
v002 1 "1. Clinton" 2 "2. Dole" 3 "3. Perot" 9 "9. NA"/
v003 1 "1. Bush" 2 "2. Clinton" 3 "3. Perot" 9 "9. NA"/
v004 v006 1 "1. Democrat" 2 "2. Republican" 9 "9. NA"/
(...)
missing values v001 to v161 v216 to v218(9).
(...)
export outfile='c:\work\mydata.por' .
```
