Coping with SPSS Syntax files on the DLI FTP and Web Sites

Chuck Humphrey, University of Alberta
Sharon Neary, University of Calgary

ACCOLEDS/DLI Training, December 2006

Outline

- What are the SPSS syntax files on the DLI FTP and Web sites?
- When and why would I or a patron use an SPSS syntax file?
- Where do I find them on the DLI FTP site?
- How do I use SPSS syntax files?
- What practices should you follow with SPSS syntax files?

File extensions on the DLI FTP site

	Frequency	Percent	Cumulative Percent
EXE	6151	26.9	26.9
ZIP	4628	20.3	47.2
PDF	3897	17.1	64.3
SPS	1364	6.0	70.3
TXT	1312	5.7	76.0
IVT	1292	5.7	81.7
DOC	995	4.4	86.0
WPD	804	3.5	89.6
SAS	674	3.0	92.5
XLS	501	2.2	94.7
TGZ	292	1.3	96.0
WP	159	.7	96.7
RTF	156	.7	97.4

File extensions association with …

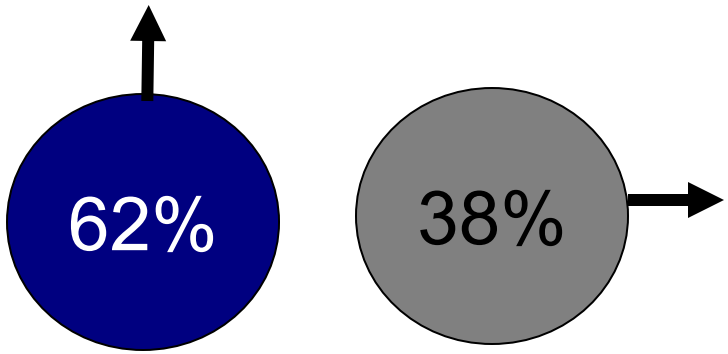
Data

EXE	6151
ZIP	4628
IVT	1292
XLS	501
TGZ	292

TXT	1312

Metadata

PDF	3897
SPS	1364
DOC	995
WPD	804
SAS	674
WP	159
RTF	156



What are DLI SPSS syntax files?

- SPSS syntax files contain code in the language used by SPSS to drive all of its operations.
- This language consists of a series of command names and a set of subcommands that specify the actions of the command.

```
FREQUENCIES  VARIABLES=JOBCAT
  GENDER   /PERCENTILES=25 50 75
   /BARCHART.
```

Command

What are DLI SPSS syntax files?

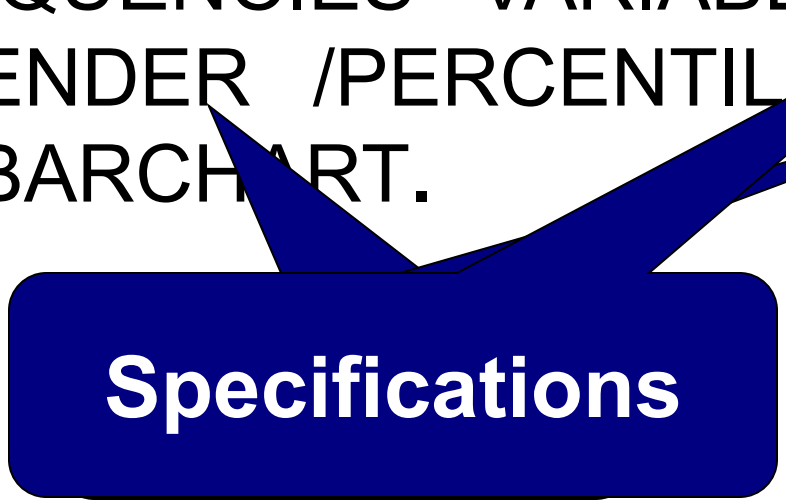
- SPSS syntax files contain code in the language used by SPSS to drive all of its operations.
- This language consists of a series of command names and a set of subcommands that specify the actions of the command.

```
FREQUENCIES   VARIABLES=JOBCAT
   GENDER   /PERCENTILES=25 50 75
    /BARCHART.
```

What are DLI SPSS syntax files?

- SPSS syntax files contain code in the language used by SPSS to drive all of its operations.
- This language consists of a series of command names and a set of subcommands that specify the actions of the command.

```
FREQUENCIES   VARIABLES=JOBCAT
   GENDER   /PERCENTILES=25 50 75
    /BARCHART.
```



What are DLI SPSS syntax files?

- SPSS syntax files contain code in the language used by SPSS to drive all of its operations.
- This language consists of a series of command names and a set of subcommands that specify the actions of the command.

```
FREQUENCIES   VARIABLES=JOBCAT
   GENDER   /PERCENTILES=25 50 75
    /BARCHART.
```

Ends Command

What are DLI SPSS syntax files?

- These commands can be grouped into three large, general sets:
  - **commands that define and read data,**
  - **commands that transform & manage data, and**
  - **commands that analyze data.**
- The syntax files on the DLI FTP site **define and read data files** (the exceptions are the few SPSS files containing code that makes use of boot strap weights.)

How to use SPSS syntax files

- There are typically five commands that define data for SPSS:
  - **File handle**
  - **Data list**
  - **Variable labels**
  - **Value labels**
  - **Missing values**
- SPSS syntax files are simple ASCII text files and can be edited by a word processor as well as the SPSS Syntax Editor.

Let's look at an SPSS syntax file

Go to the DLI Website and go the the list of files for the Adult Education and Training Survey, 2003.

http://www.statcan.ca/english/Dli/Data/Ftp/aets/aets2003.htm

Download and open the SPSS file for the Main file in the folder named by your instructor.

Now download the data file in the same folder.

When to use SPSS syntax files

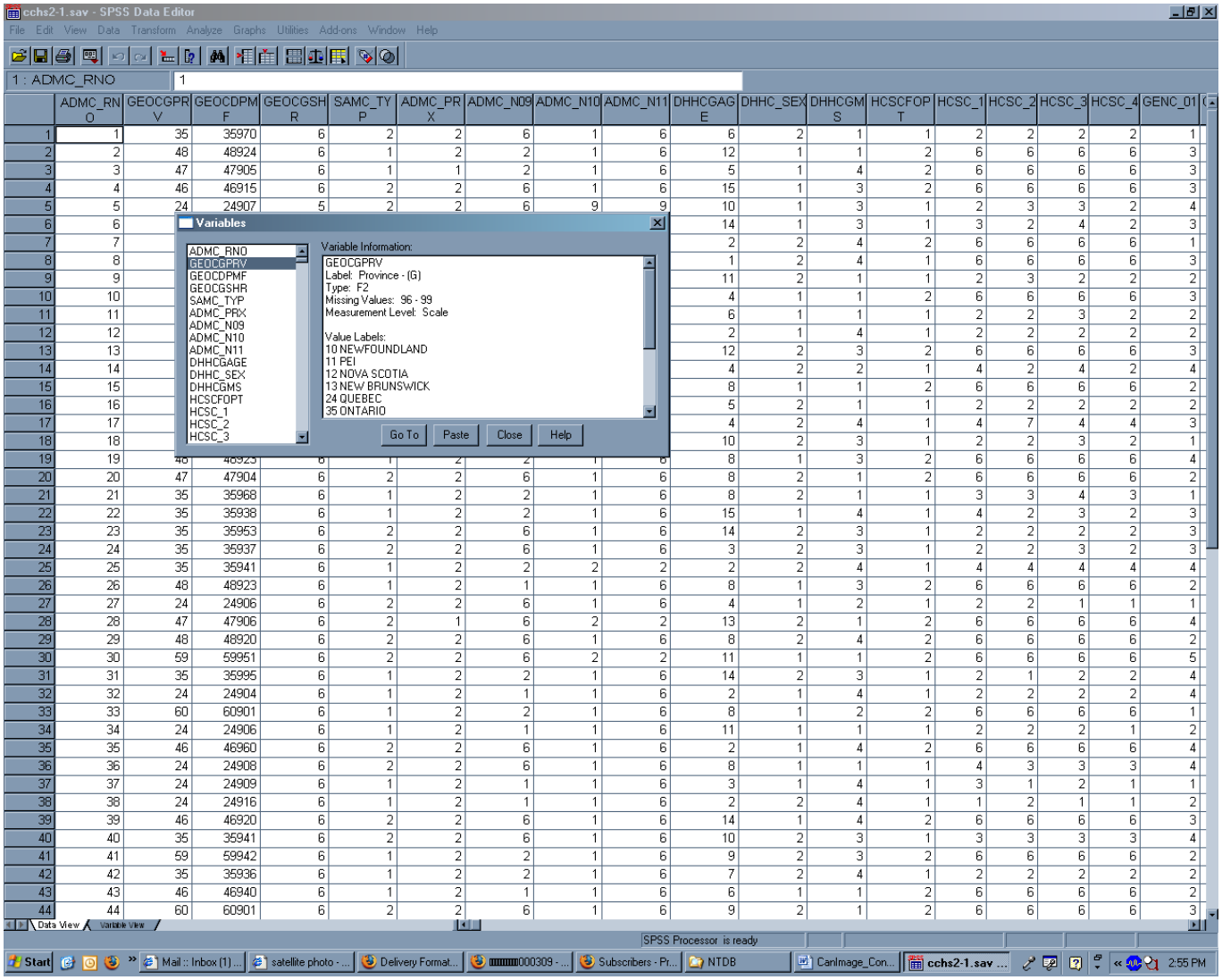
- The SPSS syntax files on the DLI FTP site are used when you or a patron needs to read an ASCII version of a microdata file or of an aggregate data file.
- To input a microdata or aggregate data file into SPSS, the physical location of the variables and their properties have to be described to the statistical system for it to read the file. This is the purpose of the Syntax files on the DLI FTP site.

CCHS 2.1 data file


```
0000013535943622616052412233322331322212269669666666666666666666611114212208102 9.73222
662262222229666966226662222222222222221666661296040203010500000000000101012266661212222
222222222212222122222220712666666662666666666666666666666666666662666666666666666666621
131123132311232666660301040101012962222222222266661122220000001111136666666666666626666
666666666666126666666666666666666666666666666666666666666622666666666666666266666666266696
666666666666661666666666662696666666669666666661101666666666666661112612244222222203 22
111122221222212126696969600 0.0000.3000.4000.0000.0001.4002.211122222222212222222222660
053001499669966996699669966996699669966996699660023996699669966996699669966996699669966
996699660101200.210033231069606050729696969696166666466666666662966666666269696969666666
666662696026666669669 9.996266666666629696966969.962666666962103396996996966669699666666
605699626669662666666266666666666266666666666666666662622112222666610202299699699699699
699699666666666666666996100000011214223296123296126296296229 69.9699.662666666699.66269
6969699.662669696969696969696969696969696969696969696969696969696969662669669696966966969
669696669669696612266966666666666666626666666666666666666669996666266666666666666666666
666996296969696129666666626666666666666696969626666666666666662969.9696961262996996996 9
969969969969969696261261144200002121112207303511626666661245504133320012 4.00
```

```
0000026060901612216081226666124221134326269669666666666666666666663111423260810 26.43212
662172222229666966226662222222221222111666661296010000000000000100000196962266662666666
666666666666666666666666011266666666266666666666666666666666666662666666666666666666622
662666666666666666660103040303012962222222222666631222101010022222666666666666666626666
666666666666666666666666666666666666666666666666666666666662666666666666666266666666266696
666666666666662666666666662696666666669666666662696666666666666666266666666666666669666
266666666666666266969696001 .0001.0000.3000.1000.4002.0004.911222222222222222222212120
903996699669966996699669966996699669966996699669966996699669966996699669966996699660104
010499660501203.110371112969696969629696969696266666666666666666296666666612099608062626
666661207326666669669 9.996266666666629696966969.962666666961601103020996966669699666666
601603126669662666666266666666666266666666666666666661166122212666610505100900900900900
900900966666666666666669961063009266666666966666966666966962969.9699.662666666699.66269
6969699.662669696969696969696969696969696969696969696969696969696969666669669696966966969
669696666966969661126696666666666666662666666666666666666669996666266666666666666666666
666996296969696269666666613454434554116304090126666666666666661000.00969626629969969969
969969969969969696261261244200003222134196304011626666661223203133320005 5.74
```

CCHS 2.1 SPSS data editor



CCHS 2.1 SPSS syntax file

```
TITLE"CCHS 2.1 (2003)"
LENGTH=NONE WIDTH=80.
FILE HANDLE cchs2003/NAME='drive:\path\HS.txt' LRECL=1381.
DATA LIST FILE=cchs2003/
 ADMC_RNO     1-   6
 GEOCGPRV     7-   8
 GEOCDPMF     9-  13
 GEOCGSHR    14-  14
 SAMC_TYP    15-  15
 ADMC_PRX    16-  16
 ADMC_N09    17-  17
 ADMC_N10    18-  18
 ADMC_N11    19-  19
 DHHCGAGE    20-  21
 DHHC_SEX    22-  22
 DHHCGMS     23-  23
 HCSCFOPT    24-  24
 HCSC_1      25-  25
 HCSC_2      26-  26
 HCSC_3      27-  27
 HCSC_4      28-  28
 GENC_01     29-  29
 GENC_02     30-  30
 GENC_02A    31-  31
 GENC_02B    32-  32
 GENC_07     33-  33
```

Locating DLI SPSS syntax files

- You will always need to match a data file with the SPSS syntax file prepared specifically for it; that is, always pair an ASCII data file with its SPSS syntax file.
- SPSS syntax files are often treated as part of the data documentation on the DLI FTP site.
- Consequently, syntax files are typically located in the folder named “docs” under a product’s folder.

Locating DLI SPSS syntax files

Let's take a look at the SPSS syntax files for the CCHS 3.1 on the DLI Website.

http://www.statcan.ca/english/Dli/Data/Ftp/cchs/cchs3-1.htm

Download the file identified as SAS_SPSS and uncompress.

Notice multiple data files and the need to match the correct syntax file with data file.

Characteristics of SPSS syntax files

- Create date on the DLI FTP site
  - **Early years, the syntax files may have come from other sources and may have been prepared for earlier versions of SPSS**
- Coding style
  - **Some author divisions prepare the SPSS syntax in one file, while others have placed the Data List command in one file and the variable and value label commands in separate files**
- Official language
  - **While not all SPSS files are in both official languages today, we will eventually have them in both languages**
- Location on the DLI FTP site
  - **As mentioned earlier, many are in the "doc" folder under a survey; some are at the root level of the folder; others are bundled in the zipped CD-image file**

Steps in Working with SPSS files

- Download the SPSS syntax file from the DLI FTP site or the DLI Web site into a folder specifically named for this survey.
- Download data file; unzip placing the ASCII version of the data in the same folder as the SPSS file.
- Edit the SPSS syntax file
  - **Scan the file for completeness of commands**
    - File Handle
    - Data List
    - Variable Labels
    - Value Labels
    - Missing Values

Completeness check

- If the syntax file does not have all five commands, check for additional files containing the missing commands.
- If there are no further commands, you need at a minimum the Data List command.
  - **You can read data into SPSS without the Variable Labels, Value Labels and Missing Values but be sure to let your patron know that this information is missing. Or you can see if someone on DLI list has a more complete version of the SPSS syntax or create it yourself.**

Fix the File Handle command

- The File Handle command will need to be edited to name the correct drive, folder and file name where the ASCII data are stored locally.
- The LRECL subcommand on the File Handle command declares the line or record length of the ASCII data file. This should be compared to the column specification of the last variable in the Data List command to ensure that the lengths match. The MAXLINE utility on the DLI FTP site can be used to check the line lengths. This information is also provided on the Web site.

Final checks

- Ensure that each command ends with a period.
- Browse to confirm that text delimiters are paired properly for the Variable and Value label commands.
  - **Common mistakes include the use of a single quote to delimit text and then including an apostrophe in the string; for example:**
    - ‘Respondent’s ID’
  - **Notice that unbalanced use of single quotes.**
  - **How to fix?**
    - Use double quotes as the delimiters for example: “Respondent’s ID”
    - Use consecutive single quotes to include the apostrophe ‘Respondent’’s ID’

Final checks

- Make the last command: Execute.
- Sometimes you will find a SAVE command in the syntax file. I recommend deleting this and using the File / Save option from the SPSS Data Editor menu.
- If there is a syntax error, SPSS will supply a message in the Output window. Some errors will result in the data not being read. Other errors just produce a warning message, which usually happens in conjunction with labeling variables or values. I wish that I could say that the error messages will always identify the problem for you. Sometimes you have to have to experiment to find the source of the problem.