

# Reading Raw Data into SPSS

---

This document shows how to read raw data into SPSS. SPSS is a comprehensive, easy to learn statistical analysis and data management software program. This document uses SPSS for Windows for illustration, but because SPSS for Windows and for Mac are very similar in terms of their interface, users of SPSS for Mac also benefit from this document with ease. No prior knowledge of SPSS is required to follow the steps in this document; however, we recommend reading our document “Using SPSS for Windows and Macintosh” (see “For Further Assistance” below) if you plan to use SPSS for data analyses later on.

Currently, SPSS is publicly available in the Meyer Library cluster, Tressider cluster, Student dorm clusters and the Social Sciences Resource Center (SSRC) Reading Room cluster in the Bing Wing of Green Library. To use any of the public computers on campus, you must have a SUNet ID (Stanford University Network Identifier). If you do not have a SUNet ID, visit the Web at <http://www.stanford.edu/services/sunetid> for information on obtaining one.

## Table of Contents

---

Who Should Read This Document? .....	1
What is Raw Data? .....	2
Types of Electronic Data Files .....	2
Scenario 1: SPSS System or Portable File Available .....	3
Scenario 2: ASCII Data File and SPSS Set-up File Available.....	4
Scenario 3: ASCII Data File Available but No SPSS Set-up File .....	8
For Further Assistance .....	10
SPSS Help .....	10
SSDS Software Services at Stanford.....	10

## Who Should Read This Document?

---

This document is intended for users who have acquired electronic data and intend to use them in SPSS but do not know where to start. Specifically, this document illustrates step-by-step how to read *fixed format raw data* into SPSS using example data downloaded from ICPSR (Inter-university Consortium for Political and Social Research) data archive (<http://www.icpsr.umich.edu>). Scenarios covered here include: 1) when the data are available as an SPSS system or portable file; 2) when the ASCII data file and SPSS setup file are available; 3) when the ASCII data file and a codebook file are available.

For other data reading needs, we also have the following documents available both as hard copies and online at <https://www.stanford.edu/group/ssds/cgi-bin/drupal/content/getting-started-guides-documents>.

If your data come in a format readable by other statistical software such as Stata or SAS, but you would like to use them in SPSS, you can use Stat/Transfer to convert the data directly into SPSS format. Please read the documents under *Data Conversion Software* for instructions. **CAN'T ONE JUST READ THEM DIRECTLY INTO SPSS?**



**Find & Analyze Data**

Find ICPSR Data

Bibliography of Data-Related Literature

Variables Database

Analyze Data Online

Thematic Collections

Restricted Data

Publication-Related Archive

**Download--Study No. 9043**

**Title:** Policy Research on Aging and Mental Health Services (PRAMHS)Project

This study is maintained and distributed by the [Substance Abuse & Mental Health Data Archive \(SAMHDA\)](#). SAMHDA is supported by the [Office of Applied Studies, Substance Abuse & Mental Health Services Administration](#).

System File(s)

ASCII File(s)

Dataset	SAS	SPSS	Stata	Other Data	All Files	File Size
All Datasets	↓	↓	↓	↓	↓	4,928 KB
Community Mental Health Centers	↓	↓	↓	↓	↓	1,752 KB
Area Agencies on Aging	↓	↓	↓	↓	↓	1,544 KB
State Units on Aging	↓	↓	↓	↓	↓	840 KB
State Mental Health Agencies	↓	↓	↓	↓	↓	792 KB

Different downloading options give you different sets of files (you do not need all of them):

Download Option	Files Extracted	File Content and Use
SPSS	xxxx-Data.por	SPSS portable file; opens data directly in SPSS
	xxxx-Data.sav	SPSS system file; opens data directly in SPSS
ASCII data file + SPSS Setup Files	xxxx-Data.txt	Raw/ASCII data file; contains data in text format
	xxxx-SPSS_setup.sps	SPSS syntax file; contains SPSS commands to read raw data in
Codebook file	xxxx-Codebook.txt (or .pdf, .doc)	Codebook file; comes with all options if available; contains information about variable coding, etc.; needed to generate SPSS syntax file if one is not available

## Scenario 1: SPSS System or Portable File Available

SPSS system (.sav) and portable (.por) files can open directly in SPSS. Therefore, whenever an SPSS system or portable file is available, choose that data format. After you have downloaded the .sav or .por file, go to the computer's program menu and launch IBM SPSS 19, then from the **File** menu, go to **Open** → **Data**, and select the appropriate file type (\*.sav or \*.por). Click on "open."

The data file will open in the PASW Statistics Data Editor like this:

09043-0004-Data.sav [DataSet1] - PASW Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

1: CASEID 1 Visible: 60 of 60 Variables

	CASEID	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16
1	1	1	4	1	112582	999	99	55	0	4	2	2	
2	2	4	4	4	33562	999	99	25	999	2	3	1	
3	3	5	4	5	140	2	1	99	99	4	3	1	
4	4	6	4	6	78139	19	16	84	80	4	2	1	
5	5	8	4	8	21826	999	99	999	0	2	4	1	
6	6	10	4	9	2877	20	30	75	5	1	4	2	
7	7	11	4	10	250752	37	89	87	99	2	3	1	
8	8	12	4	11	999999	999	99	999	10	2	4	3	
9	9	13	4	12	228	4	2	75	0	2	3	2	
10	10	17	4	16	999999	999	99	999	999	2	2	1	
11	11	18	4	17	67796	999	99	25	24	1	2	1	
12	12	22	4	21	14832	2	0	99	99	3	3	2	
13	13	23	4	22	490000	999	99	999	999	4	4	3	
14	14	24	4	25	3000	11	1	88	0	1	2	9	

Data View Variable View

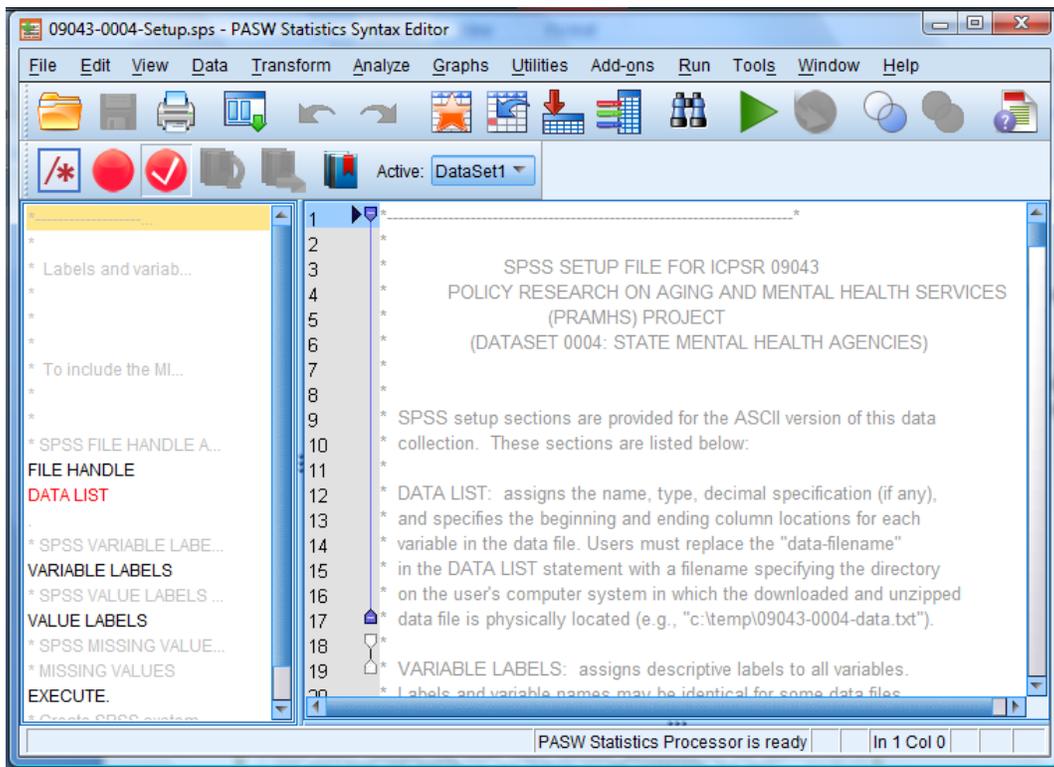
PASW Statistics Processor is ready

If you used the portable file, be sure to **save the data file** as an SPSS system file (\*.sav). You can do so by selecting **Save** from the **File** menu.

## Scenario 2: ASCII Data File and SPSS Set-up File Available

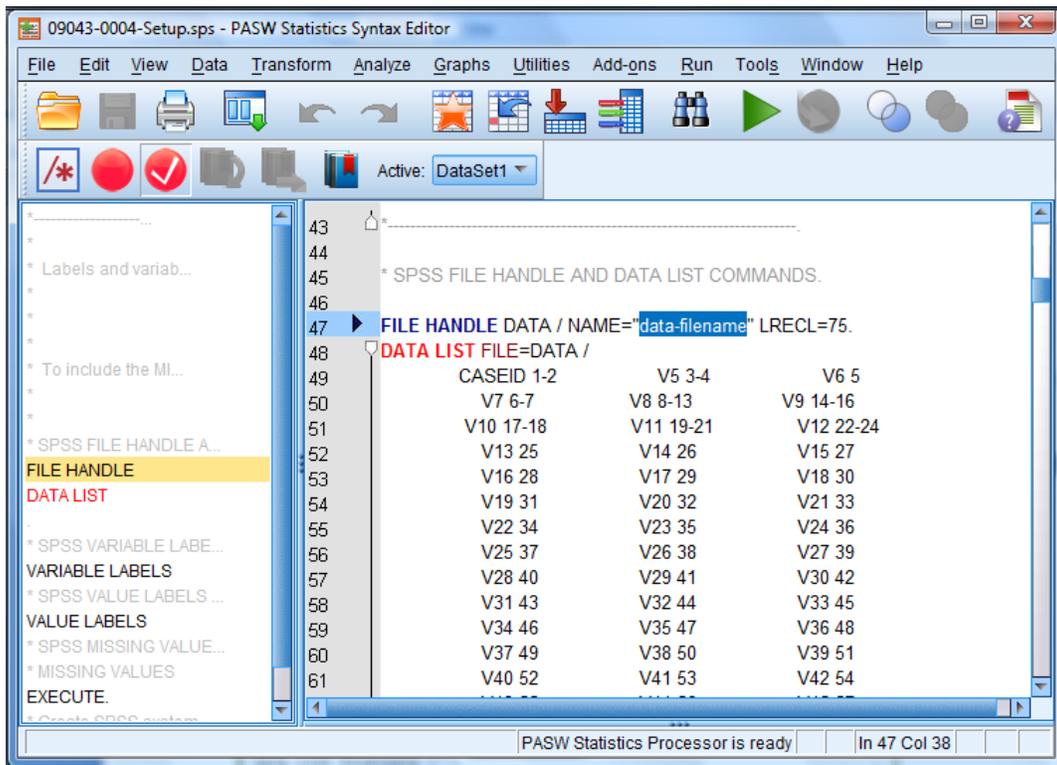
Not all datasets have SPSS system or portable files available; data digitized earlier (e.g., before mid 1990s) often only have the ASCII data file and the codebook or setup files. In this case, you will need to use a combination of the raw data file and the SPSS set-up file. Raw data files usually end with the extension **.txt** or **.dat**. SPSS set-up files are SPSS syntax files (or programs) that end with the extension **.sps**. You do not need to open the raw data file after downloading it to your computer, but will need to open and modify the SPSS set-up file.

Launch IBM SPSS Statistics 19. From the **File** menu, select **Open → Syntax**, then the \*.sps file in your data directory. This file will open in SPSS Syntax Editor (a new window) like the screenshot below. The beginning of the file, contents in the boxed area, contains information about the syntax and instructions on how to modify the syntax. It is important to read this section carefully first.



There are a few places in this syntax file that you will need to modify. The first executable commands in the file (after the instructions, which are grayed out as comments to indicate non-commands) are usually `FILE HANDLE` and `DATA LIST`, which together tell SPSS where the raw data is located, and how the variables are delimited.

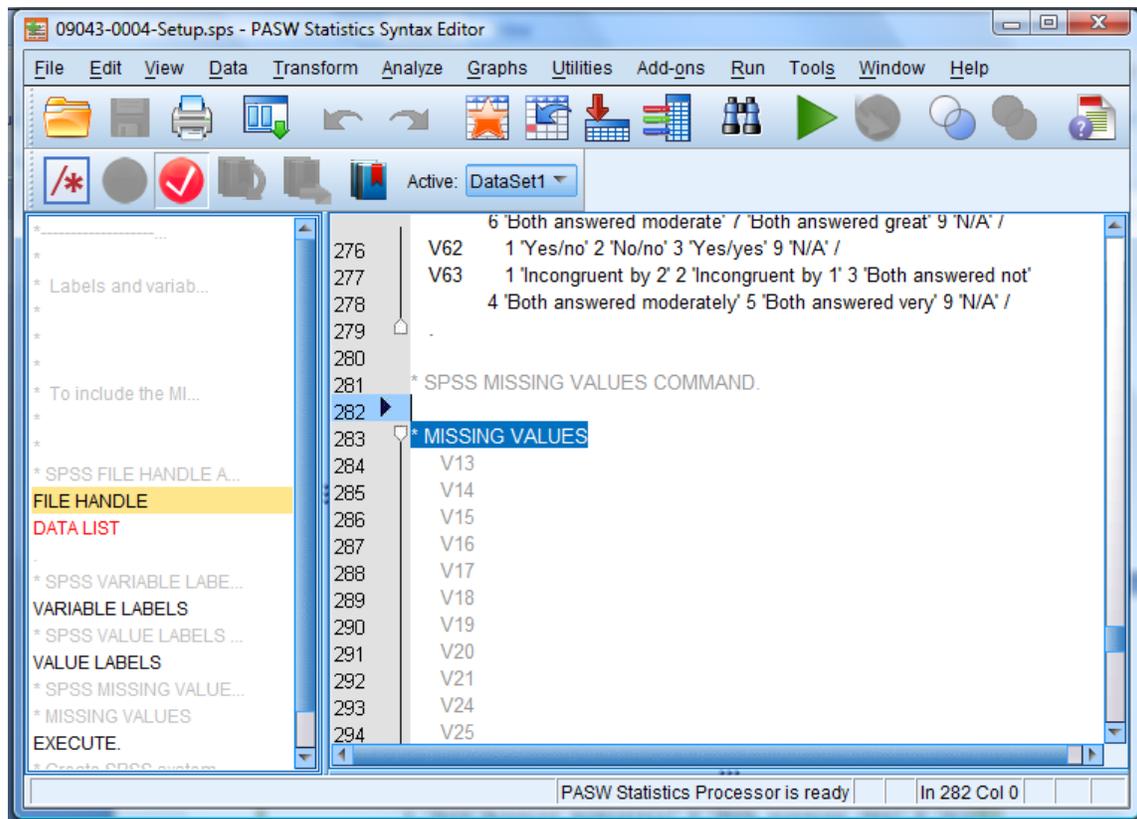
You will need to change the highlighted text below into the directory and name of the raw data file, e.g., `"C:\SSDS\ICPSR_09043\SMHA\09043-data.txt"`.



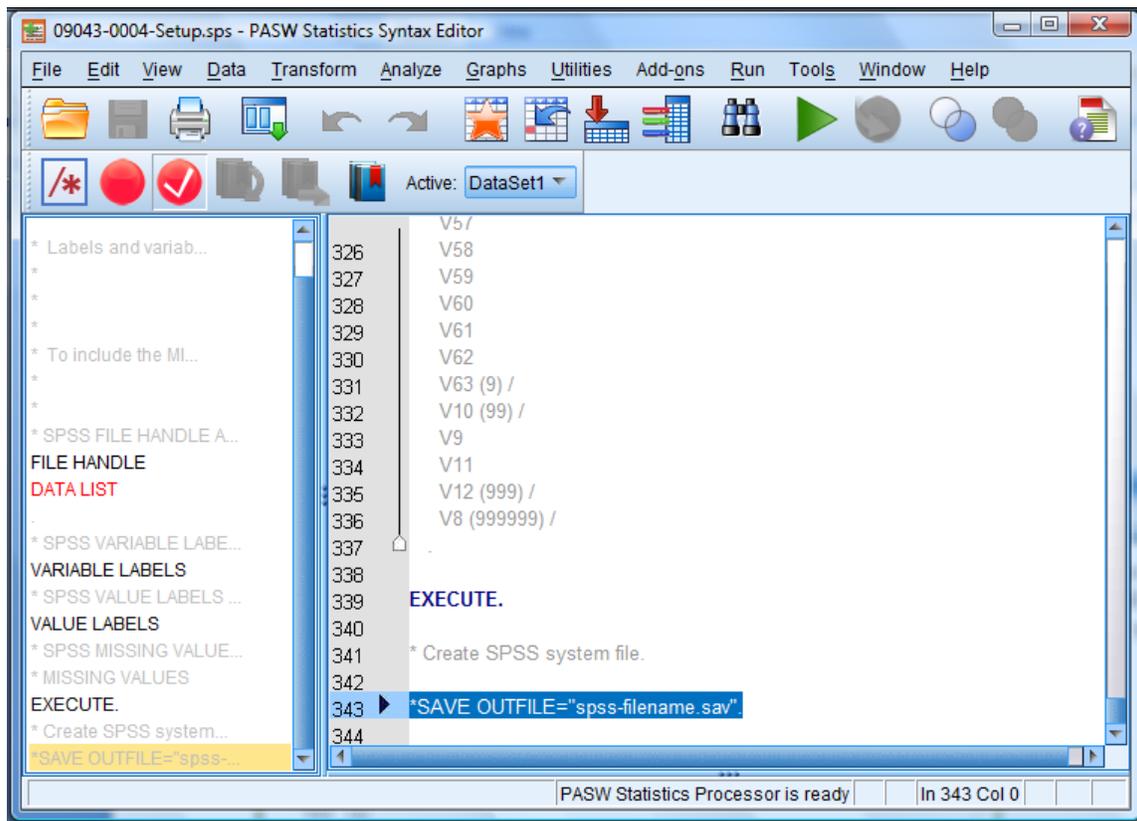
The DATA LIST command usually takes multiple lines. In this example, it is telling SPSS that “CASEID” (identifier for each case) takes its value from columns 1 to 2 in the raw data file, “V5” (name for variable 5) takes its value from columns 3 to 4, “V6” (variable 6) takes its value from column 5, etc. You do not need to modify this part, but you may delete the variables (and their column numbers) that you do not need in the final data. This will be useful when you only need a few variables from a large dataset.

After the DATA LIST command, there is usually a VARIABLE LABELS command and a VALUE LABELS command, which attach labels to the variable names and the values in each variable. You do not need to modify anything in these two commands.

Sometimes, the setup file will also include a MISSING VALUES command as the next screenshot shows. This is to specify which values represents missing data. If you would like to run this command, remove the “\*” at the beginning of this command. (In SPSS language, a “\*” at the beginning of a line makes that line non-executable; it is also used to insert a comment or note in the program).



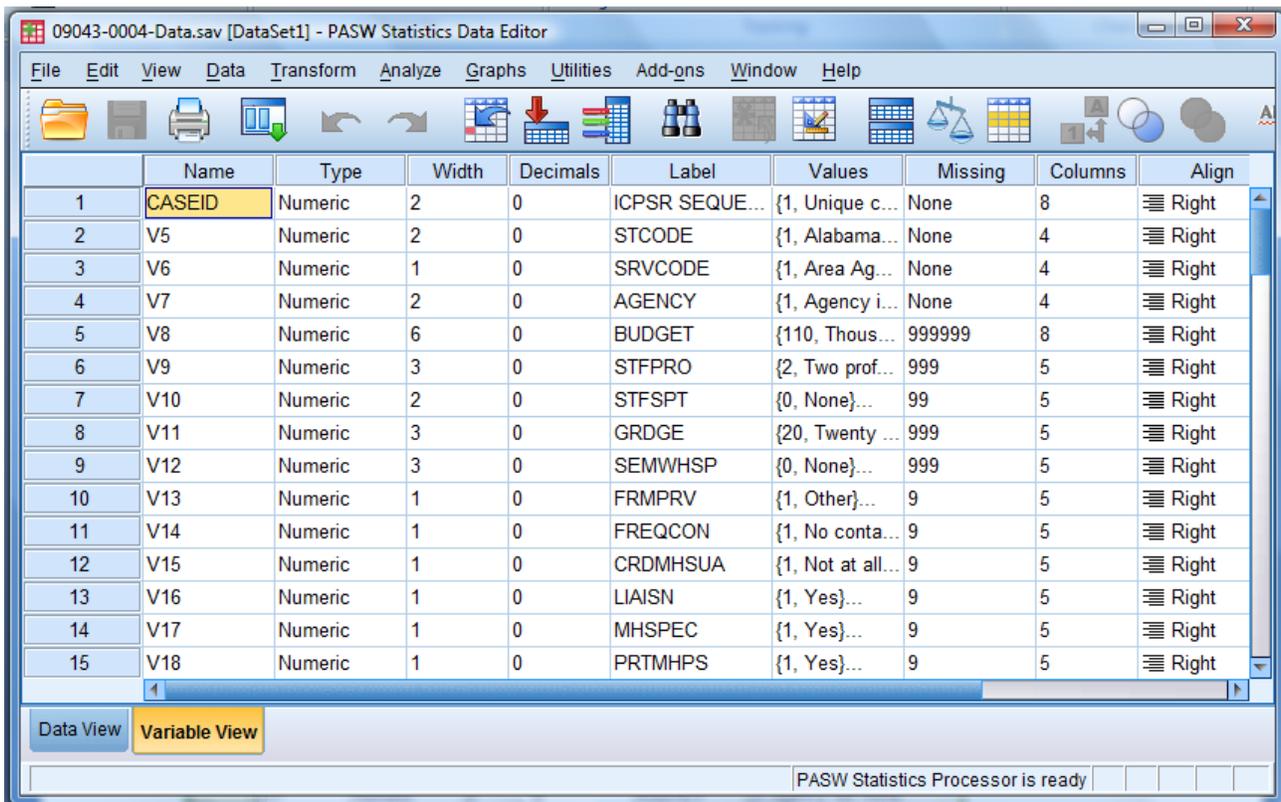
The last thing to modify in the setup file is the SAVE OUTFILE= command at the end of the file. In the following screenshot, you will need to 1) remove the “\*” at the beginning of the highlighted line; and 2) change the file name in quotation marks into the directory and file name you would like your data to be stored and called, e.g. “c:\DATA\health\_survey.sav.”



Now you may go to the **Run** menu and click on **All**. This will run all the commands in the setup file, and will read the raw data into a spreadsheet-looking window (IBM SPSS Statistics Data Editor), where each column is a variable and each row is an observation:

	CASEID	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16
1	1	1	4	1	112582	999	99	55	0	4	2	2	
2	2	4	4	4	33562	999	99	25	999	2	3	1	
3	3	5	4	5	140	2	1	99	99	4	3	1	
4	4	6	4	6	78139	19	16	84	80	4	2	1	
5	5	8	4	8	21826	999	99	999	0	2	4	1	
6	6	10	4	9	2877	20	30	75	5	1	4	2	
7	7	11	4	10	250752	37	89	87	99	2	3	1	
8	8	12	4	11	999999	999	99	999	10	2	4	3	
9	9	13	4	12	228	4	2	75	0	2	3	2	
10	10	17	4	16	999999	999	99	999	999	2	2	1	
11	11	18	4	17	67796	999	99	25	24	1	2	1	
12	12	22	4	21	14832	2	0	99	99	3	3	2	
13	13	23	4	22	490000	999	99	999	999	4	4	3	
14	14	24	4	25	3000	11	1	88	0	1	2	9	

If you click on the “**Variable View**” tab at the bottom of the window, you will get a summary view of your data. Here, the “label” column tells you what each variable means, and the “values” column tells you what each number in a variable means.

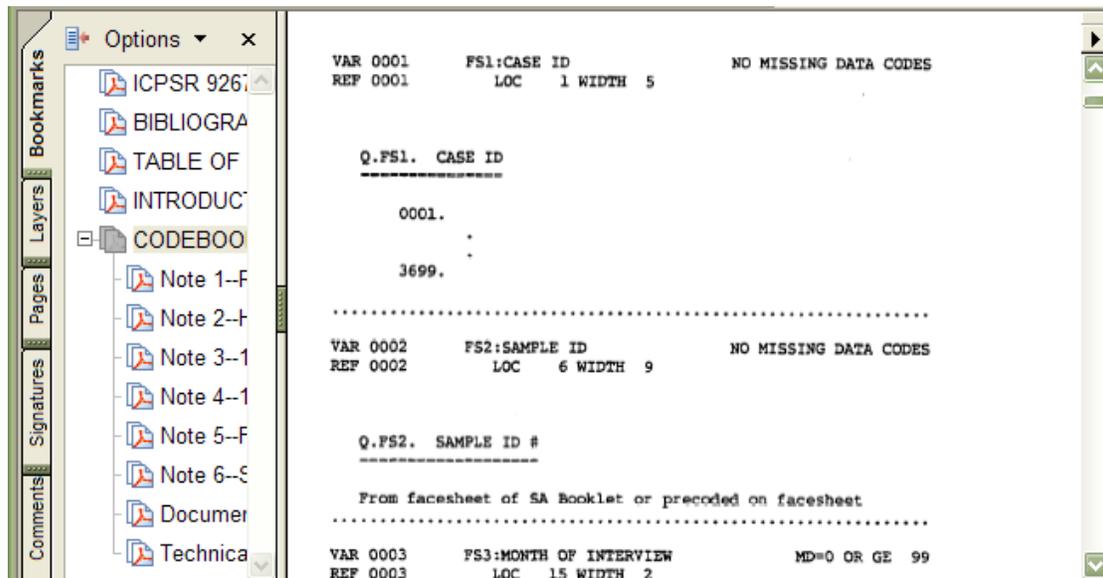


If you did not use the “SAVE OUTFILE” command in the setup file, remember to save the data file at this point by using the **File** → **Save** menu option.

### Scenario 3: ASCII Data File Available but No SPSS Set-up File

It may happen that all the electronic files available for a dataset are the raw data file and the codebook. In this situation, in order to read the raw data file into SPSS, you will need to write an SPSS syntax file (such as the one used in Scenario 2) using information from the codebook.

For example, suppose the study “Americans’ Changing Lives: Wave 1, 1986” comes in two files: **Codebook.pdf** and **Data.txt**. Open the codebook file and scroll to the page where you see variable information like this screenshot:



This codebook indicates that variable 0001 (VAR 0001) is case id (FS1:CASE ID); it starts at column 1 (LOC 1) and is 5 columns long (WIDTH 5).

To write your own SPSS setup file, launch SPSS. From the **File** menu, select **New → Syntax**. A blank syntax window will open. This is where you will write your own SPSS setup file (see Scenario 2 for an example of SPSS setup files). As in Scenario 2, your SPSS setup file will consist of 5 parts.

Part 1. Define and Select the Variables. Type in the DATA LIST command as follows, substituting the shaded text with your data directory and file name:

```
DATA LIST FILE="c:\your_data_directory\Data.txt" /
      CASEID 1-5      SAMPLEID 6-14
      VAR03 15-16     VAR04 17-18
      [add the rest of the variables you need]
.
```

Indented lines below the DATA LIST command defines the variables. For example, using information from the codebook above, we can call the first variable “CASEID” and tell SPSS that it takes columns 1-5, variable “SAMPLEID” takes columns 6-14, and so on. Note that if a variable is stored as letters (“string”) rather than numbers, you will need to add “(A)” after you specify the columns for that variable. For example: NAME 20-40 (A). You may select only the variables you will use to include in this command. The period at the end is necessary – it tells SPSS that this command is completed.

Part 2. Label the Variables. This step is optional but usually useful. The VARIABLES LABEL command can attach labels to all variables at the same time. Each indented line below is one variable, and variables are separated by “/”.

```
VARIABLE LABELS
  CASEID      'case id' /
  SAMPLEID    'sample id' /
  VAR03       'month of interview' /
  VAR04       'date of interview' /
  [add the rest of the variables and labels you need]
.
```

Part 3. Label the Values. This step is necessary to know what the numbers mean for each variable. Each indented line below is one variable. For each variable, type the number and the corresponding label in single quotes; repeat this for all values in the variable. Variables are separated by “/”.

```
VALUE LABELS
  VAR03      1 'Jan' 12 'Dec' /
  REGION     1 'Northeast' 2 'Midwest' 3 'South' 4 'West' /
  [add other variables you need]
.
```

Part 4. Missing Values. Missing values in a variable are usually indicated by a special number. The codebook will tell you which number means that the value for that variable is missing. It is recommended that you tell SPSS what missing values are so that your data analyses will give accurate results. For example:

```
MISSING VALUES
  V8 (999999)
  V9 (9)
  V10 (99)
.
```

Part 5. Save Data File. You may write a `SAVE` command as below to save the data in SPSS format, or use the Save option from the File menu after running the SPSS setup file.

```
SAVE OUTFILE = "C:\your_data_directory\your_data_name.SAV".
```

## For Further Assistance

---

### SPSS Help

SPSS has an extensive hypertext help system that provides information about using SPSS and understanding the results. To access the SPSS Help system, go to the **Help** menu on the far right side of the screen and select an option.

### SSDS Software Services at Stanford

Software Services provides technical support for SPSS users at Stanford. As mentioned earlier, this document focuses on *fixed* format raw data. If you need help with other formats, please contact the software consultants. Users can ask questions or make appointments with the consultants via our website. For more information or to contact us, see the web at:

<http://ssds.stanford.edu/>

*Note: This document is based on IBM SPSS 19 for Windows.*

---

Copyright © 2011, by The Board of Trustees of the Leland Stanford Junior University. Permission granted to copy for non-commercial purposes, provided we receive acknowledgment and a copy of the document in which our material appears. No right is granted to quote from or use any material in this document for purposes of promoting any product or service.

---

*Social Science Data and Software  
Document updated: 8/2/2011*