

SPSS 6 – Reading non-SPSS data

SPSS data formats

- Binary

- .sav

- GET FILE

- FL2@(#) SPSS DATA FILE MS Windows

(The exact text varies)

- ýýýçýýŒ ®GázŦ^V@Ë ; E Ŧ óMF@=

- ASCII

- .por

- IMPORT FILE

- ?????@????@????@????@????@????@????@????@????@????@ASCII SPSS PORT FILE

- 5C/5K/1E/2L.CI/19.D70/D.2C/23/50/5Q/18/2L.Q3/20.1JF/8.IR/1I/5K/66/1E/2D.OR/1K.G6

Data conversion utilities

- Stat/Transfer
- DBMS/Copy

Other binary formats

- SAS

```
get sas data= 'Y:\spss\data\fitness.sas7bdat'
```

- Stata

```
get stata file= 'Y:\spss\data\fitness.dta'
```

- Excel

```
get data
```

```
  /type=xls
```

```
  /file='Y:\spss\data\fitness.xls'
```

```
  /sheet=name 'fitness'.
```

- See also: `get translate file=`

Data Sources

- In-line
- In a file

Text Preliminaries

- Look at your data in a text editor first, *not* SPSS
- Does the file have anything but data values? E.g. variable names? Other documentation?
- What marks the beginning/end of an observation? E.g. a new line of text? A specific number of data values?

Preliminaries (cont.)

- What marks each data value? E.g. a special character between data values, specific column position?
 - Can the special character also be a data value?
- How are missing data denoted? E.g. a special character? Blanks in a column? Double delimiters?

Text data

- Delimited
- Fixed
- Details
 - Records/case and cases/record
 - Missing values
 - Text qualifiers

Two text import commands

- DATA LIST
 - Fewer options, easier to write
- GET DATA /TYPE=TXT
 - Many options, has some quirky details
- Crucial details
 - Column numbering and input formats differ
 - DATA LIST is a transformation command

Methods of Delimiting Data Values

- “Fixed” or “column” format

```
Gregor 10 2 40 0
```

```
Molly 15 5 10 1000
```

More Methods of Delimiting Data Values

- “Delimited” or “list” or “free” format

- Space delimited

```
Gregor 10 2 40 0  
Molly 15 5 10 1000
```

- Comma delimited

```
Gregor,10,2,40,0  
Molly,15,5,10,1000
```

- Tab delimited (often hard to distinguish visually)

```
Gregor      10      2      40      0  
Molly15     5       10     1000
```

Delimiters and Qualifiers

- When a delimiter is also valid as part of a data value

```
"Gregor Mendel" 10 2 40 0
```

```
"Molly Ivens" 15 5 10 1000
```

Records per Case

Cases per Record

Missing Values

- Multiple consecutive delimiters *MAY* indicate a missing value

```
Gregor,10,2,40,0
```

```
Molly,15,,10,1000
```

Or they *MAY NOT*

```
Gregor 10 2 40 0
```

```
Molly 15 5 10 1000
```