

# SPSS Data Management Syntax

- \* SPSS is **not case sensitive**, but most use UPPERCASE for **commands**.
- \* **Commands** must **end with a period** but can be on multiple lines.
- \* Typically, **spaces are optional** around operators or symbols.
- \* Only the **first 3 letters** are needed for most **commands** and **selectors**.

## In the Master File

Set Project Folder	<b>CD</b> "C:\My Docs\My Project".	stands for <b>C</b> hange <b>D</b> irectory
Open a SPSS file	<b>GET FILE =</b> "raw/myfile.sav".	location is relative to the project folder
Run a Syntax file	<b>INSERT FILE =</b> "syntax.sps".	as if the syntax was selected and run
Sort Cases	<b>SORT CASES BY</b> key .	
Merge Files:	<b>MATCH FILES</b>	<b>SORT</b> both files by <b>key</b> before <b>MATCH</b>
Add Variables	<b>/FILE =</b> 'one.sav' <b>/FILE =</b> 'two.sav' <b>/BY</b> key .	use <b>TABLE=</b> instead for the keyed table
Merge Files:	<b>ADD FILES</b>	variable types and widths must match
Add Cases	<b>/FILE =</b> 'one.sav' <b>/FILE =</b> 'two.sav' <b>/IN =</b> two .	put <b>IN</b> immediately after the file to indicate (value of 1 vs 0) in the <b>variable</b>
Delete Variables	<b>DELETE VARIABLES</b> v1 v2 v3.	immediate and permanent; it is often better to <b>SAVE</b> as a new file using <b>KEEP</b> ,
Delete Cases	<b>SELECT IF NOT MISSING</b> ( id ).	or <b>FILTER</b> then <b>UNSELECTED = DELETE</b>
Exclude Cases	<b>FILTER BY</b> valid .	can <b>DELETE</b> excluded cases using <b>SAVE</b>
Save a SPSS file	<b>SAVE OUTFILE =</b> "myfile.sav" <b>/UNSELECTED = DELETE</b> <b>/KEEP =</b> id v1 v2 v3 <b>/COMPRESSED</b> .	this is the only necessary line makes the <b>FILTER</b> permanent order matters, can also <b>DROP</b> use <b>ZCOMPRESSED</b> for a smaller file

## Comments

- \* Any Comments .
- \*\*\* Headings \*\*\*.
- /\*** Inside Commands **\*/**
- TITLE** "For Printouts".
- SUBTITLE** "Organize".

## Univariate Statistics

- FREQUENCIES** v1 v2.
- DESCRIPTIVES** v3 v4.
- EXAMINE** v1 **TO** v4.
- GRAPH HISTOGRAM** v3.

## Bivariate Statistics

- CROSSTABS** v1 **BY** v2  
**/CELL = COUNT ROW** .
- CORRELATIONS** v1 v2 v3.
- GRAPH SCATTER** v1 **WITH** v2.

## Analyze Sub-samples

- SPLIT FILE BY** group .  
do stuff for each subsample
- SPLIT FILE OFF**.
- FILTER BY** indicator .  
do stuff on a single subsample
- FILTER OFF** .

## Create and Modify Variables

Math & Logic	<b>COMPUTE</b> total = v1 + v2 . <b>COMPUTE</b> is_3 = ( v1 = 3 ) . /* TRUE → 1 */
Conditions	<b>IF</b> ( v1 > 20 <b>AND</b> v2 = 3 ) group = \$Sysmis.
Occurrences	<b>COUNT</b> n_of_3s = v1 v2 v3 ( 3 ). <b>COUNT</b> n_missing = v1 v2 v3 ( <b>MISSING</b> ).
Change Values	<b>RECODE</b> v1 ( 3, 4 = 0 ) ( <b>LO THRU</b> 2 = 1 ) ( 9 = <b>SYSMIS</b> ) ( <b>ELSE = COPY</b> ) <b>INTO</b> v1r.
Encode Strings	<b>AUTORECODE</b> string / <b>INTO</b> num . <i>/* if both are numeric, will clone */</i>
Create Strings	<b>STRING</b> new ( <b>A#</b> ) . <i>/* must run first */</i>
Update Data	<b>EXECUTE</b> . <i>/* run after changing data */</i>

## Variable Properties

Name	<b>RENAME VARIABLES</b> ( old = new ) .
Type	<b>RECODE</b> string ( <b>CONVERT</b> ) <b>INTO</b> num . <b>COMPUTE</b> string = <b>STRING</b> ( num, F5.0 ) .
Width	<b>ALTER TYPE</b> string ( <b>AMIN</b> ) .
Decimals	<b>FORMATS</b> num ( F5.0 ) . <b>FORMATS</b> date ( <b>ADATE</b> 10 ) .
Label	<b>VARIABLE LABELS</b> v1 'My Label' .
Values	<b>VALUE LABELS</b> v1 1 "One" 2 "Two" .
Missing	<b>MISSING VALUES</b> v1 v2 v3 ( 7, 9 ) .
Measure	<b>VARIABLE LEVEL</b> v1 v2 v3 ( <b>SCALE</b> ) .
Codebook	<b>DISPLAY DICTIONARY</b> .

### Operators

	<b>OR</b>
&	<b>AND</b>
~	<b>NOT</b>
=	<b>EQ</b> ual to
~=	<b>Not EQ</b> ual to
<	<b>Less Than</b>
<=	<b>Less than</b> or <b>EQ</b> ual to
>	<b>Greater Than</b>
>=	<b>Greater than</b> or <b>EQ</b> ual to

### Functions

<b>COMPUTE</b> new = <b>FUNCTION</b> ( ... ) .	or	<b>IF</b> <b>FUNCTION</b> ( ... ) new = 1 .
<b>One Variable:</b> <b>FUNCTION</b> ( v1 ) .		<b>Multiple Variables:</b> <b>FUNCTION</b> ( v1,v2,v3 ) .
Numeric <b>LG10</b> , <b>LN</b> , <b>RND</b> , <b>TRUNC</b> , <b>ABS</b>		Statistics <b>MEAN</b> , <b>SUM</b> , <b>MIN</b> , <b>MAX</b> , <b>SD</b>
String <b>RTRIM</b> , <b>UPCASE</b> , <b>LOWER</b> <b>SUBSTR</b> ( v1, pos, length ) <b>REPLACE</b> ( v1, "old", "new" )		Find <b>ANY</b> ( #, v1, v2, ...vk )
Missing <b>MISSING</b> , <b>NOT MISSING</b>		Missing <b>NVALID</b> , <b>NMISS</b> , <b>FUNCTION.#</b>
Extract #'s <b>NUMBER</b> ( <b>SUBSTR</b> ( string,1,1), F8.0 ).		<b>Functions</b> require only 1 valid input/value unless indicated ex. <b>MEAN.5</b> needs 5 valid values
Excel Date <b>DATE.MDY</b> (1,1,1900)+((v1-2)*24*3600).		<b>Operators</b> require all values to be valid or the result will be Missing