How can SPSS help me document my data?

Authored by

stats writer

June 30, 2024

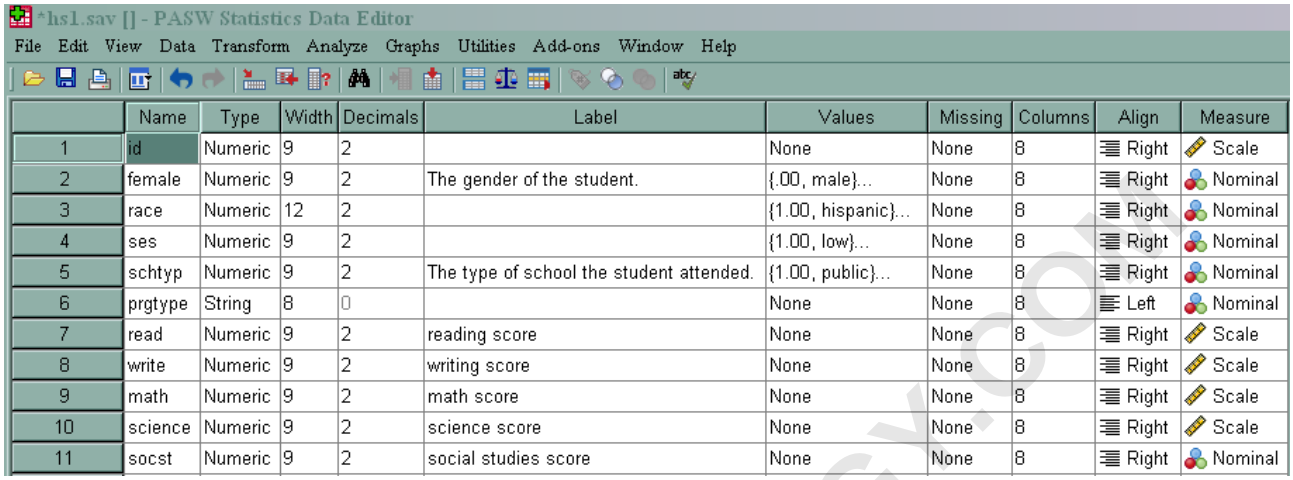
SPSS (Statistical Package for the Social Sciences) is a software tool that can greatly aid in documenting and analyzing data. With its user-friendly interface and powerful data management capabilities, SPSS can assist in organizing, storing, and documenting data efficiently and accurately. Its wide range of statistical and graphical tools can also help in visualizing and summarizing complex data sets, making it easier to identify patterns and trends. Additionally, SPSS offers customizable reporting options that allow users to document their data in a clear and concise manner, making it an ideal tool for creating professional and comprehensive research reports. Overall, SPSS can greatly assist individuals in efficiently and effectively documenting their data, saving time and effort in the research process.

How can SPSS help me document my data? | SPSS FAQ

The codebook command was introduced in SPSS version 17. It
provides information about the variables in a dataset, such as the type,
variable labels, value labels, as well as the number of cases in each level of
categorical variables and means and standard deviations of continuous variables.
This information can be as important as the data themselves, because it helps to
give meaning to the data. Also, this information can help you distinguish
between two similar datasets.

The examples below will use the hs1.sav dataset.
Let's start by looking at the Variable View.

get file "D:datahsb1.sav".


*hs1.sav [] - PASW Statistics Data Editor

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

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	id	Numeric	9	2		None	None	8	Right	Scale
2	female	Numeric	9	2	The gender of the student.	{.00, male}...	None	8	Right	Nominal
3	race	Numeric	12	2		{1.00, hispanic}...	None	8	Right	Nominal
4	ses	Numeric	9	2		{1.00, low}...	None	8	Right	Nominal
5	schtyp	Numeric	9	2	The type of school the student attended.	{1.00, public}...	None	8	Right	Nominal
6	prgtype	String	8	0		None	None	8	Left	Nominal
7	read	Numeric	9	2	reading score	None	None	8	Right	Scale
8	write	Numeric	9	2	writing score	None	None	8	Right	Scale
9	math	Numeric	9	2	math score	None	None	8	Right	Scale
10	science	Numeric	9	2	science score	None	None	8	Right	Scale
11	socst	Numeric	9	2	social studies score	None	None	8	Right	Nominal

You can access the codebook command via the point-and-click interface by
clicking on Analyze -> Reports -> Codebook.

Let's consider the syntax below. Although it may look complicated, only
the command itself is necessary. If you issue the codebook command
by itself, you will get the variable information for all of the variables in the
dataset; counts and percents for all categories of nominal and ordinal
variables; and means, standard deviations and quartiles for scale variables.

This may be more output than you want, so you may prefer to select which
variables and what information about them you would like to see. In the
example below, we have selected six variables from our dataset. In square
brackets ( ) after each variable name, we have indicated the measurement
level. Scale variables (AKA continuous variables) are indicates with an s, ordinal variables
(AKA categorical variables)
with an o, and nominal variables with an n. The measurement
level specified in the command may or may not match that shown in the Variable
View. For example, as we can see above, the variable socst has a
nominal measurement; however, in the codebook command below, we have
specified it as a scale variable. The type of measurement determines what will
be provided in the output for the variable: counts and percents for all
categories of nominal and ordinal variables; means,

**standard deviations and
quartiles for scale variables.**

**On the varinfo subcommand, we request some of the information that we see in
the Variable View. On the fileinfo subcommand, we request information on
the data file itself, such as the name of the data file, its location, the file
label, any documents attached to the data file and a count of the number of
cases in the dataset. On the statistics subcommand, we request the count
and percent, which gives the number of cases and percent of cases in each level
of nominal and ordinal variables. We also request the mean and standard
deviation of scale variables.**

```
codebook ses prgtype write science socst
/varinfo position label type format measure valuelabels missing
/fileinfo name location label documents casecount
/statistics percent mean stddev.
```

Codebook

D:\data\hs1.sav

File Information

File Name		hs1.sav
Location		D:\data
Label		
Number of Cases	Unweighted	200
	Weighted	200

ses

		Value	Count	Percent
Standard Attributes	Position	4		
	Label	<none>		
	Type	Numeric		
	Format	F9.2		
	Measurement	Nominal		
Valid Values	1.00	low	47	23.5%
	2.00	middle	95	47.5%
	3.00	high	58	29.0%

prgtype

		Value	Count	Percent
Standard Attributes	Position	6		
	Label	<none>		
	Type	String		
	Format	A8		
	Measurement	Nominal		
Valid Values	academic		105	52.5%
	general		45	22.5%
	vocati		50	25.0%

write

		Value
Standard Attributes	Position	8
	Label	writing score
	Type	Numeric
	Format	F9.2
	Measurement	Scale
N	Valid	200
	Missing	0
Central Tendency and Dispersion	Mean	52.7750
	Standard Deviation	9.47859

science

		Value
Standard Attributes	Position	10
	Label	science score
	Type	Numeric
	Format	F9.2
	Measurement	Scale
N	Valid	195
	Missing	5
Central Tendency and Dispersion	Mean	51.6615
	Standard Deviation	9.86603

socst

		Value
Standard Attributes	Position	11
	Label	social studies score
	Type	Numeric
	Format	F9.2
	Measurement	Nominal
N	Valid	200
	Missing	0
Central Tendency and Dispersion	Mean	52.4050
	Standard Deviation	10.73579

In the example below, we show how to get minimal output (by using the keyword

none on the statistics subcommand), and ordering the output in
alphabetical order (by using specifying varorder = alpha on the
options subcommand).

```
codebook ses prgtype science socst
/varinfo label type valuelabels
/options varorder = alpha
/statistics none.
```

Codebook

[DataSet1] D:\Data\hs1.sav

prgtype

		Value
Standard Attributes	Label	<none>
	Type	String

science

		Value
Standard Attributes	Label	science score
	Type	Numeric

ses

		Value
Standard Attributes	Label	<none>
	Type	Numeric
Valid Values	1.00	low
	2.00	middle
	3.00	high

socst

		Value
Standard Attributes	Label	social studies score
	Type	Numeric

For more information about documenting data in SPSS, please visit

SPSS Learning Modules: Labeling and documenting data and

Statistical Consulting Seminars: Introduction to SPSS Syntax, Part 1 (section 13).