

GNU PSPP

A program for statistical analysis of sampled data.

John Darrington

¹School of Computer Science and Software Engineering
University of Western Australia
Perth, WA, Australia

The seventh Free and Open source Software Developers'
European Meeting, 2007



What is PSPP?

PSPP is:

- a GNU Project,



What is PSPP?

PSPP is:

- a GNU Project,
- a program for statistical analysis of sampled data,



What is PSPP?

PSPP is:

- a GNU Project,
- a program for statistical analysis of sampled data,
- a Free replacement for the proprietary program SPSS.



What is PSPP?

PSPP is:

- a GNU Project,
- a program for statistical analysis of sampled data,
- a Free replacement for the proprietary program SPSS.

Other Free statistical analysis projects include:

- GNU R (very powerful and complete. Not so easy to learn),



What is PSPP?

PSPP is:

- a GNU Project,
- a program for statistical analysis of sampled data,
- a Free replacement for the proprietary program SPSS.

Other Free statistical analysis projects include:

- GNU R (very powerful and complete. Not so easy to learn),
- **Goose (C++ library),**



What is PSPP?

PSPP is:

- a GNU Project,
- a program for statistical analysis of sampled data,
- a Free replacement for the proprietary program SPSS.

Other Free statistical analysis projects include:

- GNU R (very powerful and complete. Not so easy to learn),
- Goose (C++ library),
- GNU Octave (numerical analysis and linear algebra).



PSPP place in the Free Software community?

PSPP fills a gap in the suite of Free Software tools. Other Free projects are arguably more mature and flexible than PSPP. However, they are aimed at the hacker level, and are not so amenable to non-hackers. PSPP aims to be substantially similar to SPSS, — a *proprietary* program used by medical researchers, psychologists, market analysts and social scientists.



PSPP place in the Free Software community?

PSPP fills a gap in the suite of Free Software tools. Other Free projects are arguably more mature and flexible than PSPP. However, they are aimed at the hacker level, and are not so amenable to non-hackers. PSPP aims to be substantially similar to SPSS, — a *proprietary* program used by medical researchers, psychologists, market analysts and social scientists.

PSPP is to an SPSS user what Naltrexone is to an opium addict



What can PSPP do

PSPP currently features:



What can PSPP do

PSPP currently features:

- GNU GPL — no licensing fees, no timebombs.



What can PSPP do

PSPP currently features:

- GNU GPL — no licensing fees, no timebombs.
- Handles 2^{32} cases and 2^{32} variables.



What can PSPP do

PSPP currently features:

- GNU GPL — no licensing fees, no timebombs.
- Handles 2^{32} cases and 2^{32} variables.
- Very fast processing, (some reports indicate faster than SPSS), even when data exceeds memory capacity.



What can PSPP do

PSPP currently features:

- GNU GPL — no licensing fees, no timebombs.
- Handles 2^{32} cases and 2^{32} variables.
- Very fast processing, (some reports indicate faster than SPSS), even when data exceeds memory capacity.
- Complete set of mathematical functions, including random number distributions.



What can PSPP do

PSPP currently features:

- GNU GPL — no licensing fees, no timebombs.
- Handles 2^{32} cases and 2^{32} variables.
- Very fast processing, (some reports indicate faster than SPSS), even when data exceeds memory capacity.
- Complete set of mathematical functions, including random number distributions.
- **Identical command language to SPSS.**



What can PSPP do

PSPP currently features:

- GNU GPL — no licensing fees, no timebombs.
- Handles 2^{32} cases and 2^{32} variables.
- Very fast processing, (some reports indicate faster than SPSS), even when data exceeds memory capacity.
- Complete set of mathematical functions, including random number distributions.
- Identical command language to SPSS.
- **Data file format identical to SPSS.**



What can PSPP do

PSPP currently features:

- GNU GPL — no licensing fees, no timebombs.
- Handles 2^{32} cases and 2^{32} variables.
- Very fast processing, (some reports indicate faster than SPSS), even when data exceeds memory capacity.
- Complete set of mathematical functions, including random number distributions.
- Identical command language to SPSS.
- Data file format identical to SPSS.
- **GUI similar to SPSS.**



PSPP's GUI

The screenshot displays the PSPP Data Editor interface. The main window, titled 'Untitled0 --- PSPP Data Editor', shows a list of variables with the following columns: Name, Type, Width, Decimals, Label, Values, Missing, Columns, Align, and Measure. The variables listed are 'day' and 'month', both of type 'Numeric' with a width of 2 and 1 decimal place. A 'Syntax1 --- PSPP Syntax Editor' window is open, showing the command 'GET FILE=' followed by a file path and 'LIST.'. A 'Variable Type' dialog box is open, allowing the user to select a data type for a variable. The 'Date' option is selected, and the format 'dd-mmm-yyyy' is chosen from the list of date formats.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
0	day	Numeric	2	1		None	None	8	Right	Nominal
1	month	Numeric	2	1		None	None	8	Right	Nominal
2								8	Right	Nominal
3								8	Right	Nominal
4								8	Right	Scale
5								180	Left	Nominal
6								7	Left	Nominal
7										Scale
8										Ordinal
9										Nominal
10										Nominal
11										Nominal
12										Ordinal
13										Nominal
14										Nominal

Variable Type dialog box options:

- Numeric
- Comma
- Dot
- Scientific notation
- Date
- Dollar
- Custom currency
- String

Date format list:

- dd-mmm-yyyy
- dd-mmm-yy
- mm/dd/yyyy
- mm/dd/yy
- dd.mm.yyyy
- dd.mm.yy
- yyyy/mm/dd
- w/mm/dd

What PSPP can't do

PSPP is work in progress. Improvements are required:



What PSPP can't do

PSPP is work in progress. Improvements are required:

- More complete set of statistical procedures,



What PSPP can't do

PSPP is work in progress. Improvements are required:

- More complete set of statistical procedures,
- Improved output subsystem,



What PSPP can't do

PSPP is work in progress. Improvements are required:

- More complete set of statistical procedures,
- Improved output subsystem,
- Better integration with Gnome/OpenOffice.org,



What PSPP can't do

PSPP is work in progress. Improvements are required:

- More complete set of statistical procedures,
- Improved output subsystem,
- Better integration with Gnome/OpenOffice.org,
- **Tutorial type documentation.**



Acknowledgements

Thanks to:

- Ben Pfaff (creator and current maintainer),
- Jason Stover (statistician and contributor),
- Patrick Brunier (graphic designer and website maintainer).

<http://www.gnu.org/software/pspp/index.html>

