## Introduction to SPSS

### SPSS – What Is It?

### SPSS Data Editor Window

### SPSS Syntax Introduction

### SPSS Dictionary

### SPSS Output – Basics, Tips & Tricks

## Must-Know Statistics

### Pearson Correlations – Quick Introduction

### Standard Deviation – What Is It?

### What is a Frequency Distribution?

### SPSS – What Is It?

**SPSS** means “**S**tatistical **P**ackage for the **S**ocial **S**ciences” and was first launched in 1968. Since SPSS was acquired by IBM in 2009, it's officially known as IBM SPSS Statistics but most users still just refer to it as “SPSS”.

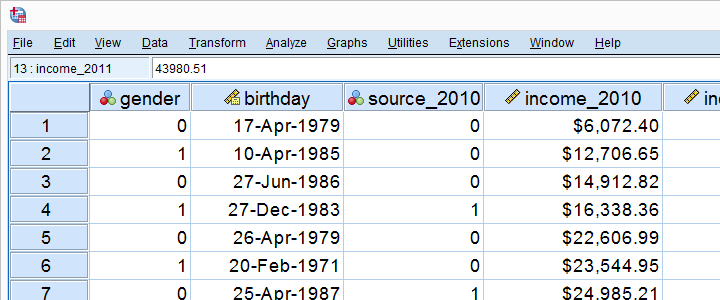
## Quick Overview Main Features

SPSS is software for **editing and analyzing all sorts of data**. These data may come from basically any source: scientific research, a customer database, Google Analytics or even the server log files of a website. SPSS can open all file formats that are commonly used for structured data such as

* spreadsheets from MS Excel or OpenOffice;
* plain text files (.txt or .csv);
* relational (SQL) databases;
* Stata and SAS.

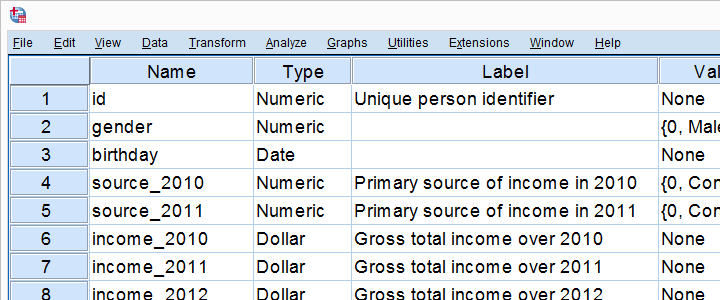
## SPSS Data View

After opening data, SPSS displays them in a **spreadsheet**-like fashion as shown in the screenshot below from freelancers.sav.



This sheet -called data view- always displays our **data values**. For instance, our first record seems to contain a male respondent from 1979 and so on. A more detailed explanation on the exact meaning of our variables and data values is found in a second sheet shown below.

## SPSS Variable View



An SPSS data file always has a second sheet called variable view. It shows the metadata associated with the data. Metadata is information about the **meaning of variables** and data values. This is generally known as the “codebook” but in SPSS it's called the dictionary.

SPSS Dictionary - Complete Overview

Note: we tried to sort these properties from most to least important. "Optional" indicates whether a property can (technically) be absent.

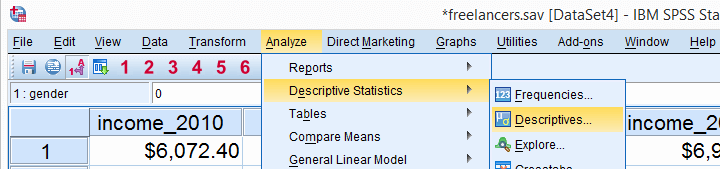
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **NAME** | **APPLIES TO** | **FUNCTION** | **IMPORTANCE** | **OPTIONAL** |
| variable name | Variable | Variable identifier. | High | No |
| Label | Variable | Normal language description of the meaning of variables. | High | Yes |
| Value Labels | Value | Normal language description of the meaning of values. | High | Yes |
| User Missing Values | Value | Tells SPSS which values to ignore in calculations. | High | Yes |
| Type | Variable | Tells SPSS how to store values internally. | High | No |
| Format | Variable | Tells SPSS how to display numeric values. | High | No |
| Document | Data file | Long data file description. | Low | Yes |
| Width | Variable | Maximum number of characters that values may consist of. | Low | No |
| Columns | Variable | Variable's column width (as displayed on screen). | Low | No |
| Variable Attribute | Variable | Descriptive tags for variables. | Low | Yes |
| Datafile attribute | Data file | Data file description using arrays. | Low | Yes |
| File label | Data file | Short data file description. | Low | Yes |
| Align | Variable | Alignment of data values (on screen). | Low | No |
| Measure | Variable | Measurement level nominal, ordinal or scale (= metric). | Low | No |
| Role | Variable | The variable's supposed relation to other variables. | Low | No |

Previous tutorial: SPSS Data Preparation 6 – Inspect Cases

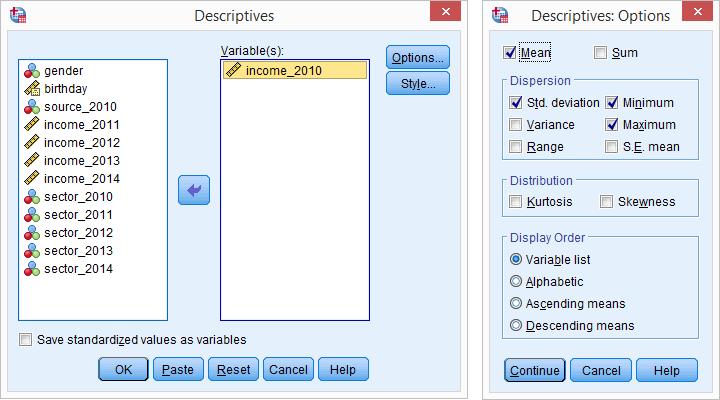
Next tutorial: Changing Variable Properties 1 – Introduction

Data Analysis

Right, so SPSS can open all sorts of data and display them -and their metadata- in two sheets in its Data Editor window. So **how to analyze your data** in SPSS? Well, one option is using SPSS’ elaborate menu options.  
For instance, if our data contain a variable holding respondents’ incomes over 2010, we can compute the average income by navigating to Descriptive Statistics as shown below.

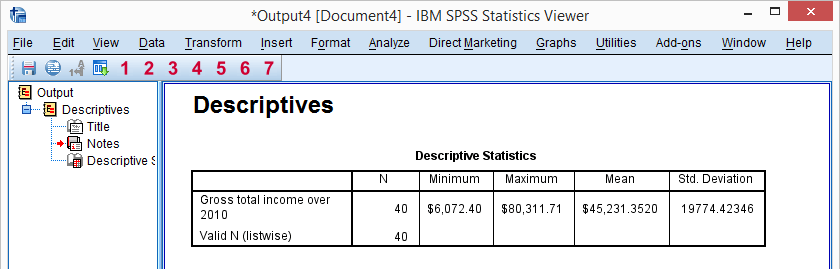


Doing so opens a dialog box in which we select one or many variables and one or several statistics we'd like to inspect.



SPSS Output Window

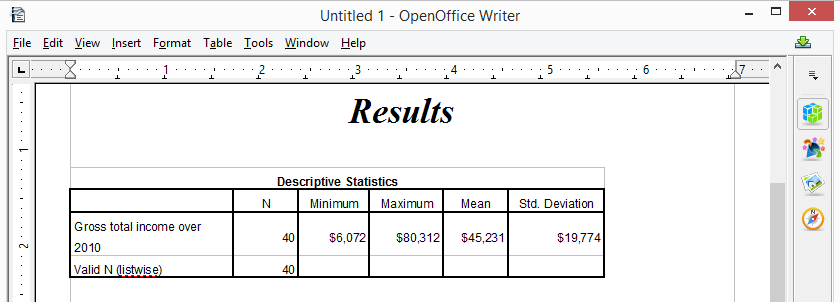
After clicking Ok, a new window opens up: SPSS’ output viewer window. It holds a nice table with all statistics on all variables we chose. The screenshot below shows what it looks like.



As we see, the **Output Viewer window** has a different layout and structure than the Data Editor window we saw earlier. Creating output in SPSS does not change our data in any way; unlike Excel, SPSS uses different windows for data and research outcomes based on those data.  
  
For non SPSS users, the look and feel of SPSS’ Output Viewer window probably comes closest to a Powerpoint slide holding items such as blocks of text, tables and charts.

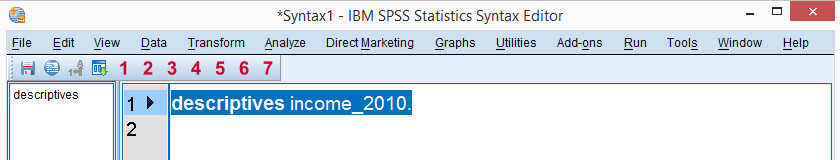
SPSS Reporting

SPSS Output items, typically tables and charts, are easily **copy-pasted** into other programs. For instance, many SPSS users use a word processor such as MS Word, OpenOffice or GoogleDocs for reporting. Tables are usually copied in rich text format, which means they'll retain their styling such as fonts and borders. The screenshot below illustrates the result.



SPSS Syntax Editor Window

The output table we showed was created by running Descriptive Statistics from SPSS’ menu. Now, SPSS has a second option for running this (or any other) command: we can open a third window, known as the syntax editor window. Here we can type and run SPSS code known as **SPSS syntax**. For instance, runningdescriptives income\_2010.has the exact same result as running this command from SPSS’ menu like we did earlier.



Besides typing commands into the Syntax Editor window, most of them can also be pasted into it by clicking through SPSS’ menu options. Like so, SPSS users unfamiliar with syntax can still use it. But **why use syntax** if SPSS has such a nice menu?  
The basic point is that **syntax can be saved**, corrected, rerun and shared between projects or users. Your syntax makes your SPSS work **replicable**. If anybody raises any doubts regarding your outcomes, you can show exactly what you did and -if needed- correct and rerun it in seconds.  
  
For non SPSS users, the look and feel of SPSS’ Syntax Editor window probably come closest to Notepad: a single window basically just containing plain text.

SPSS - Overview Main Features

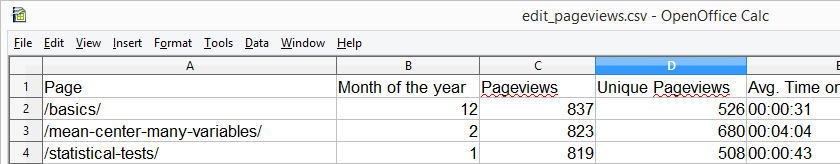
Now that we have a basic idea of how SPSS works, let's take a look at what it can do. Following a typical project workflow, SPSS is great for

* Opening data files, either in SPSS’ own file format or many others;
* editing data such as computing sums and means over columns or rows of data. SPSS has outstanding options for more complex operations as well.
* creating tables and charts containing frequency counts or summary statistics over (groups of) cases and variables.
* running inferential statistics such as ANOVA, regression and factor analysis.
* saving data and output in a wide variety of file formats.

We'll now take a closer look at each one of these features.

Opening Data Files

SPSS has its own data file format. Other file formats it easily deals with include MS Excel, plain text files, SQL, Stata and SAS.

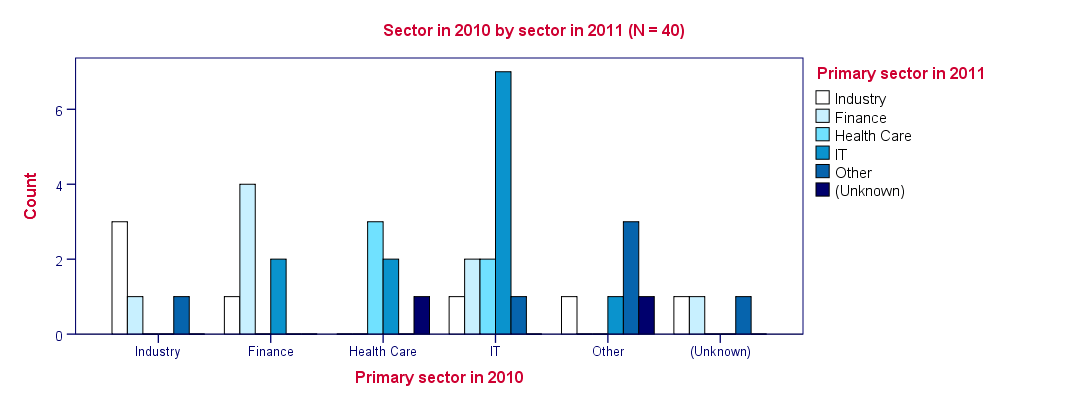
*Web analytics data -often downloaded as Excel files- can easily be opened and further analyzed in SPSS*

Editing Data

In real world research, raw **data usually need some editing** before they can be properly analyzed. Typical examples are creating means or sums as new variables, restructuring data or detecting and removing unlikely observations. SPSS performs such tasks -and more complex ones- with amazing efficiency.  
For getting things done fast, SPSS contains many numeric functions, string functions, date functions and other handy routines.

Tables and Charts

All basic tables and charts can be created easily and fast in SPSS. Typical examples are demonstrated under Data Analysis. A real weakness of SPSS is that its **charts tend to be ugly** and often have a clumsy layout. A great way to overcome this problem is developing and applying SPSS chart templates. Doing so, however, requires a fair amount of effort and expertise.

*SPSS clustered bar chart with chart template applied*

Inferential Statistics

SPSS contains all basic statistical tests and multivariate analyses such as

* t-tests;
* chi-square tests;
* ANOVA;
* correlations and other association measures;
* regression;
* nonparametric tests;
* factor analysis;
* cluster analysis.