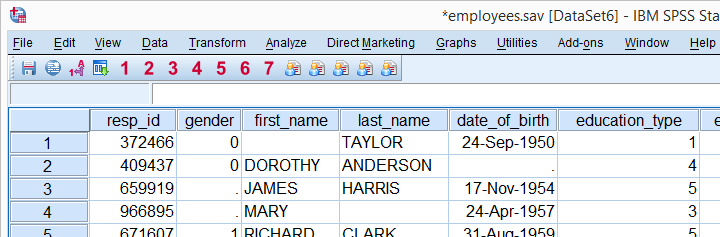
SPSS – What Is It?



**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”.

## SPSS - Quick Overview Main Features

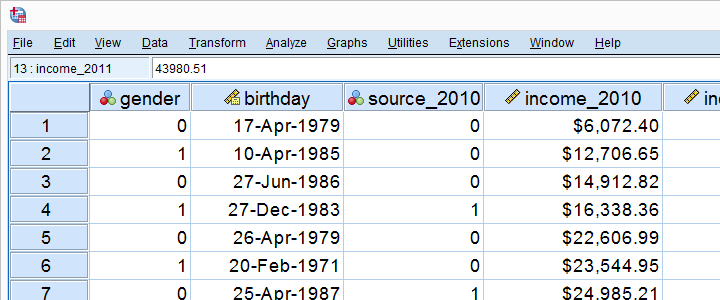
SPSS is software for **editing and analyzing all sorts of data**. These data may come from basically any source: scientific research, a customer database, Google Analytics or even the server log files of a website. SPSS can open all file formats that are commonly used for structured data such as

* spreadsheets from MS Excel or [OpenOffice](https://www.openoffice.org/): .xls (ms office <2007) , .xlsx (ms office >2005)
* plain text files (.txt – text file or .csv – comma separate version);
* relational (SQL) databases: MS SQL Server, Oracle , My SQL Server
* Stata and SAS.

Let's now have a quick look at what SPSS looks and feels like.

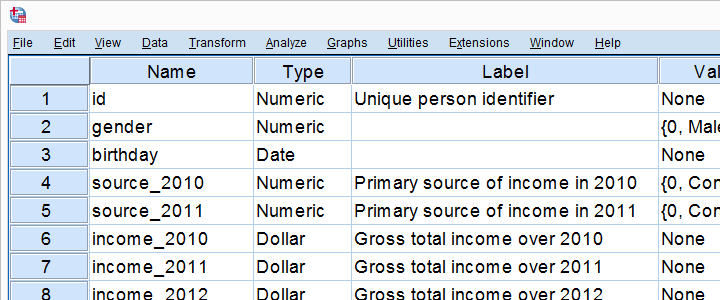
## SPSS Data View

After opening data, SPSS displays them in a **spreadsheet**-like fashion as shown in the screenshot below from [freelancers.sav](https://www.spss-tutorials.com/downloads/freelancers.sav).



This sheet -called [data view](https://www.spss-tutorials.com/spss-data-editor-window/#spss-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.

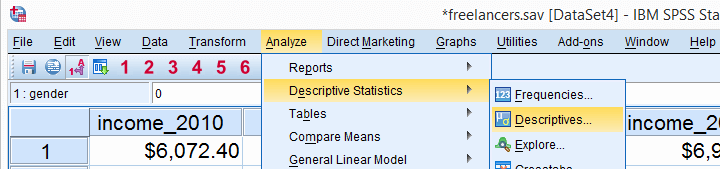
## SPSS Variable View



An SPSS data file always has a second sheet called [variable view](https://www.spss-tutorials.com/spss-data-editor-window/#spss-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](https://www.spss-tutorials.com/spss-dictionary/).

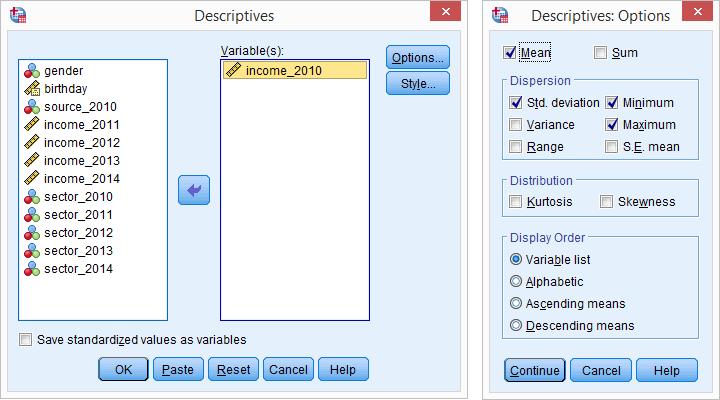
## Data Analysis

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.



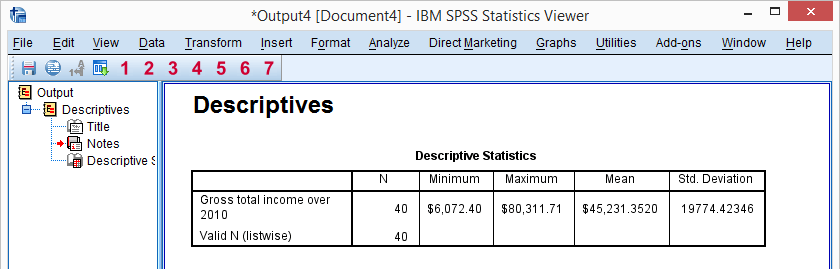
Doing so opens a dialog box in which we select one or many variables and one or several statistics we'd like to inspect.

Variable/column/field



## SPSS Output Window

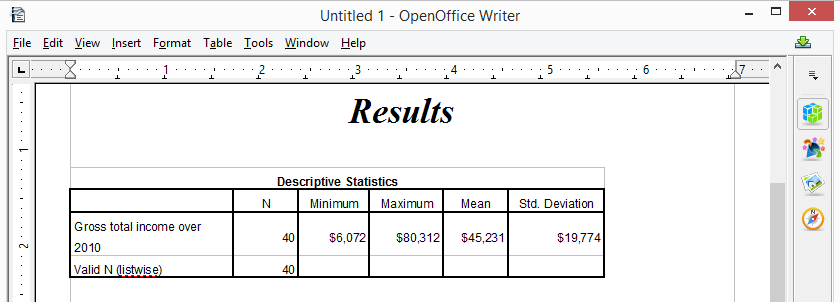
After clicking Ok, a new window opens up: SPSS’ [output viewer window](https://www.spss-tutorials.com/spss-output/). It holds a nice table with all statistics on all variables we chose. The screenshot below shows what it looks like.



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.

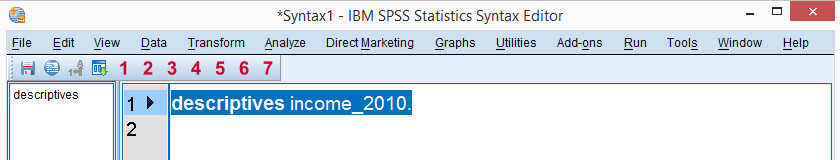
## SPSS Reporting

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](https://www.spss-tutorials.com/spss-syntax/). 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

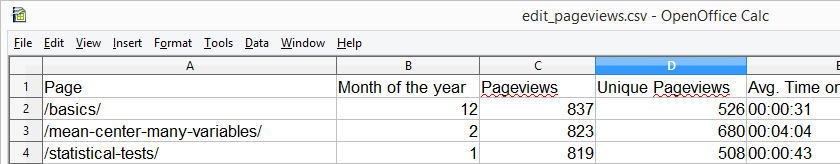
Now that we have a basic idea of how SPSS works, let's take a look at what it can do. Following a typical project workflow, SPSS is great for

* [Opening data files](https://www.spss-tutorials.com/spss-what-is-it/#opening-data), either in SPSS’ own file format or many others;
* [editing data](https://www.spss-tutorials.com/spss-what-is-it/#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](https://www.spss-tutorials.com/spss-what-is-it/#tables-charts) containing frequency counts or summary statistics over (groups of) cases and variables.
* running [inferential statistics](https://www.spss-tutorials.com/spss-what-is-it/#inferential-statistics) such as ANOVA, regression and factor analysis.
* [saving data and output](https://www.spss-tutorials.com/spss-what-is-it/#saving-data) 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.

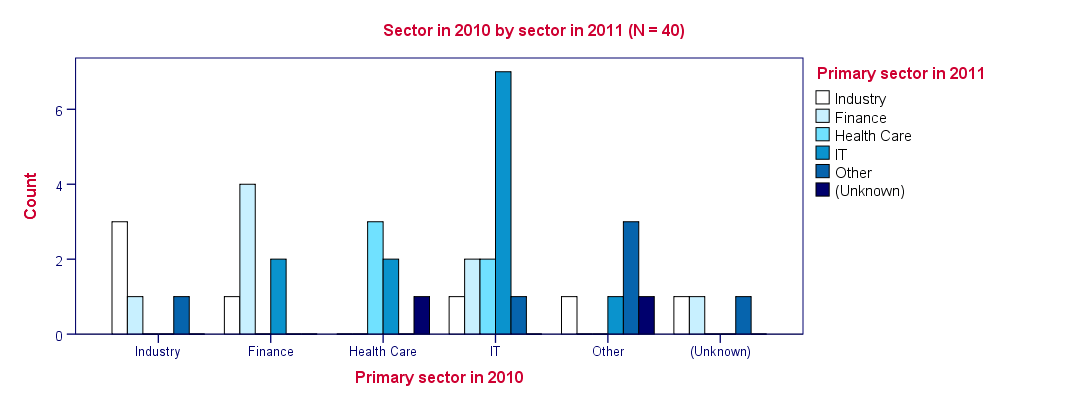
*Web analytics data -often downloaded as Excel files- can easily be opened and further analyzed in SPSS*

## 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](https://www.spss-tutorials.com/spss-main-numeric-functions/), [string functions](https://www.spss-tutorials.com/spss-string-variables-tutorial/), [date functions](https://www.spss-tutorials.com/spss-date-variables-tutorial/#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](https://www.spss-tutorials.com/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](https://www.spss-tutorials.com/spss-chart-templates/). Doing so, however, requires a fair amount of effort and expertise.

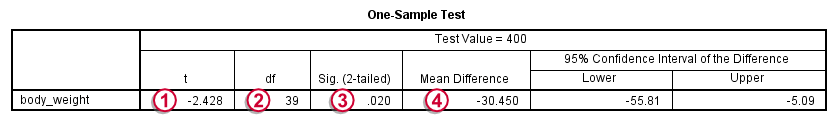
*SPSS clustered bar chart with chart template applied*

## Inferential Statistics

SPSS contains all basic statistical tests and multivariate analyses such as

* [t-tests](https://www.spss-tutorials.com/t-test/);
* [chi-square tests](https://www.spss-tutorials.com/chi-square-independence-test/);
* [ANOVA](https://www.spss-tutorials.com/anova-what-is-it/);
* [correlations](https://www.spss-tutorials.com/pearson-correlation-coefficient/) and other association measures;
* [regression](https://www.spss-tutorials.com/regression/);
* [nonparametric tests](https://www.spss-tutorials.com/nonparametric-tests/);
* [factor analysis](https://www.spss-tutorials.com/spss-factor-analysis-tutorial/);
* cluster analysis.

Some analyses are available only after purchasing additional SPSS options on top of the main program. An overview of all commands and the options to which they belong is presented in [Overview All SPSS Commands](https://www.spss-tutorials.com/overview-all-spss-commands/).

*SPSS One Sample T-Test Output Example*

## Saving Data and Output

SPSS **data** can be saved as a variety of file formats, including

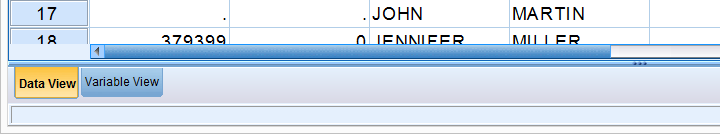
* MS Excel;
* plain text (.txt or .csv);
* Stata;
* SAS.

The options for **output** are even more elaborate: charts are often copy-pasted as images in .png format. For tables, rich text format is often used because it retains the tables’ layout, fonts and borders.  
Besides copy-pasting individual output items, all output items can be **exported in one go**to .pdf, HTML, MS Word and many other file formats. A terrific strategy for writing a report is creating an SPSS output file with nicely styled tables and chart. Then export the entire document to Word and insert explanatory text and titles between the output items.

## SPSS Data View & Variable View

An SPSS data file always has two tabs in the left bottom corner:

* **Data View** is where we inspect our actual data and
* **Variable View** is where we see additional information about our data.

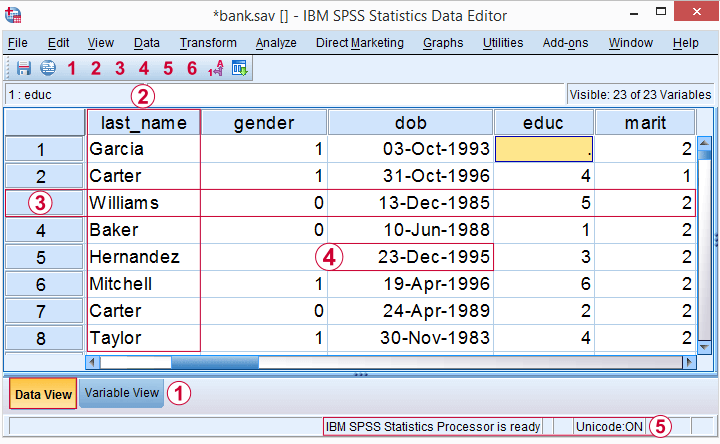


You can **switch** between Data View and Variable View by

* clicking the tabs in the left bottom corner;
* using the Ctrl + t [shortkey](https://www.spss-tutorials.com/spss-keyboard-shortcuts/);
* double-clicking a variable name in Data View;
* double-clicking an outline number in Variable View.

Let's first take a close look at the main parts of the Data View tab. We'll then proceed with [variable view](https://www.spss-tutorials.com/spss-data-editor-window/#spss-variable-view).

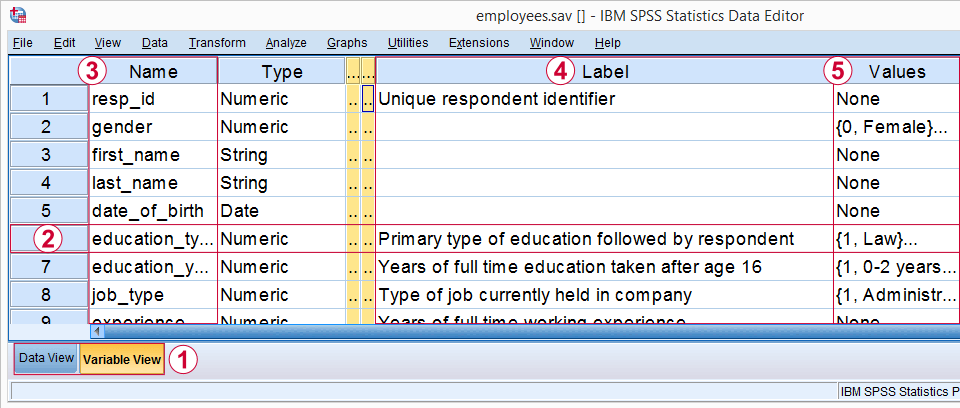
## SPSS Data View



https://spss-tutorials.com/img/b1.png The data editor has **tabs** for switching between Data View and Variable View. For now, make sure you're in Data View.  
https://spss-tutorials.com/img/b2.png Columns of cells are called **variables**. Each variable has a unique name (“gender”) which is shown in the column header.  
https://spss-tutorials.com/img/b3.png Rows of cells are called **cases**. Oftentimes, each respondent in a study is represented as a single case.  
https://spss-tutorials.com/img/b4.png In SPSS, **values** refer to cell contents.  
https://spss-tutorials.com/img/b5.png The **status bar** may give useful information on the data, for instance whether a [WEIGHT](https://www.spss-tutorials.com/spss-weight-command/), [FILTER](https://www.spss-tutorials.com/spss-filter-command/), [SPLIT FILE](https://www.spss-tutorials.com/spss-split-file-command/) or Unicode mode is in effect.

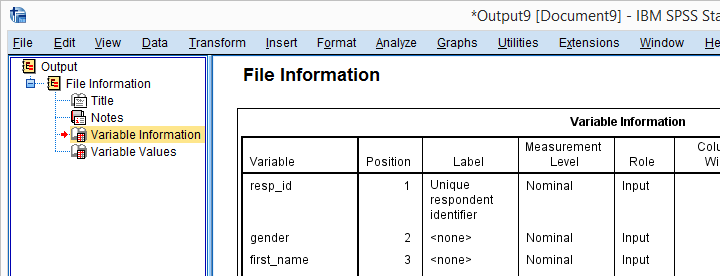
These are the main elements in Data View. Let's now switch to Variable View.

## SPSS Variable View



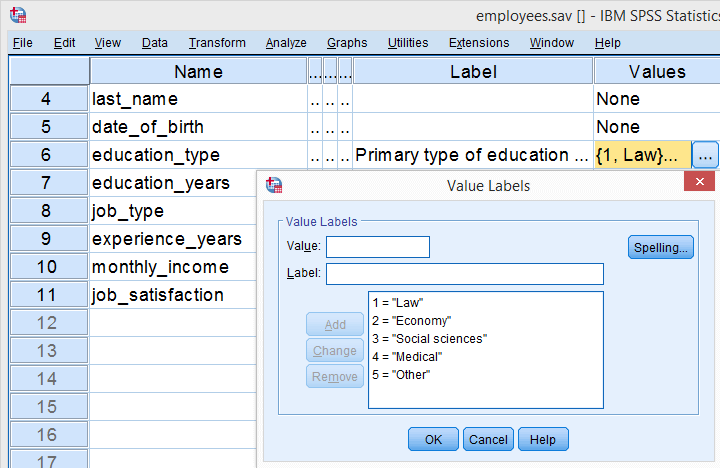
https://spss-tutorials.com/img/b1.png In the left bottom corner we find **tabs** for switching between Variable View and Data View. For now, select Variable View.  
https://spss-tutorials.com/img/b2.png In Variable View, **variables** are shown as rows of cells.  
https://spss-tutorials.com/img/b3.png The first column shows the **variable name** for each variable.  
https://spss-tutorials.com/img/b4.png The fifth column may or may not contain a **variable label**. This describes the exact meaning of each variable.  
https://spss-tutorials.com/img/b5.png The sixth column shows **value labels**: descriptions of the meaning of one, many or all values that a variable may contain.

In short, Variable View does not show the data itself but, rather, information about the data. This is sometimes called “metadata” or “the codebook”. In SPSS, however, it's called the [dictionary](https://www.spss-tutorials.com/spss-dictionary/).  
This is important to know because you may find commands like DISPLAY DICTIONARY or APPLY DICTIONARY in manuals. If you're familiar with [syntax](https://www.spss-tutorials.com/spss-syntax-editor-window/), runningDISPLAY DICTIONARY.creates the output shown below: dictionary information as seen in variable view.

*Dictionary information reported by running DISPLAY DICTIONARY.*

## Variable View - Value Labels

For some variables, it's immediately clear what their values mean: a value of € 2500,- in a variable “gross monthly income” represents a gross monthly income of € 2500,-.  
This is not always the case, however: answer categories for [categorical variables](https://www.spss-tutorials.com/measurement-levels/#categorical) are often represented by numbers -usually 1 through x. What these values represent is then stored in their **value labels**. Clicking the open value labels icon https://spss-tutorials.com/img/spss-open-value-labels-icon.png for education\_type displays all value labels for this variable.

*Value labels for education\_type.*

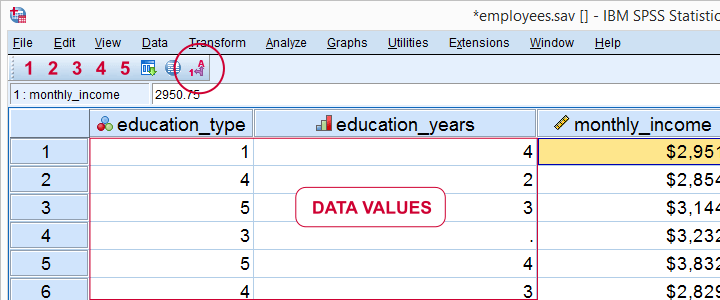
These value labels tell us that a person with a value of 1 on education\_type indicates somebody who studied “Law”. In a similar vein, “Economy” is represented by a value of 2, and so on.

## Dictionary Information in Data View

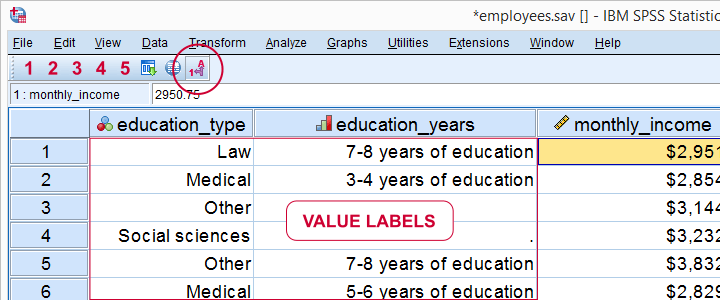
Thus far, we explained that SPSS’ Data Editor always has 2 tabs:

* Data View in which we inspect our actual **data values** and
* Variable View in which we find information about our data -**dictionary information**.

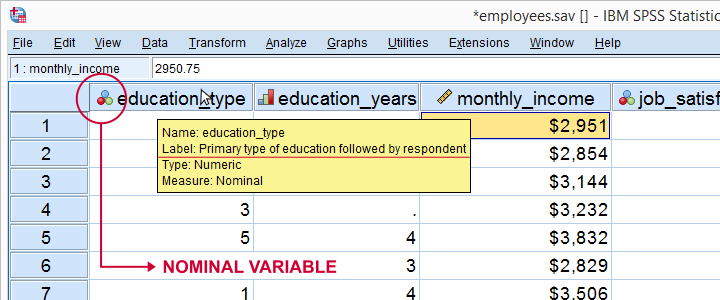
Little known by many SPSS users is that we can see some dictionary information in Data View too. Let's start off with value labels. Initially, we just see data values in Data View as shown below.



Now, if we click the **value labels icon** https://spss-tutorials.com/img/spss-show-value-labels-icon.png we'll see value labels instead of data values in data view.



So this allows you to look up what your data mean without having to switch between Data View and Variable View. Perhaps even more useful: place your **mouse pointer on a variable name** in Data View without clicking it. Now a yellow box with a lot of dictionary information pops up for a few seconds.



Starting from SPSS version 22, icons next to variable names tell us something about our [variable types](https://www.spss-tutorials.com/spss-variable-types-and-formats/), formats and measurement levels -if correctly set, that is.

## Final Notes

So basically, “data” consist of 2 components:

* **data values** which we see in Data View and
* **dictionary information** about our data in Variable View.

## SPSS Variable Types

SPSS has **2 variable types**:

* **Numeric variables** contain only numbers and are suitable for numeric calculations such as addition and multiplication.
* **String variables** may contain letters, numbers and other characters. You can't do calculations on string variables -even if they contain only numbers.

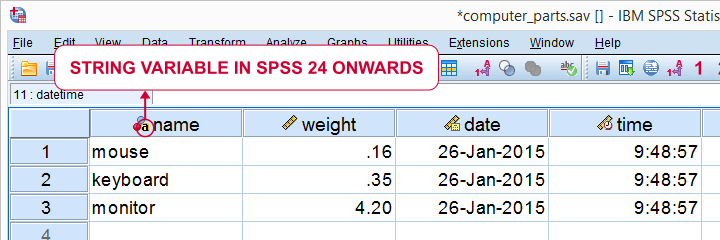
There are **no other variable types** in SPSS than string and numeric. However, numeric variables have several different formats that are often confused with variable types. We'll see in a minute how [variable view](https://www.spss-tutorials.com/spss-data-editor-window/#spss-variable-view) puts users on the wrong track here.  
The only way to change a string variable to numeric or reversely is [ALTER TYPE](https://www.spss-tutorials.com/spss-alter-type-command/). However, there's several ways to make a numeric copy of a string variable or reversely. We'll get to those in a minute.

## So What's Better: String or Numeric?

The simplest rule of thumb is that**only nominal variables with many categories  
should be string variables in SPSS.**Examples are names of people, email addresses, passport numbers and so on. Although such variables can be useful, we don't usually analyze them.  
We do sometimes analyze [nominal variables](https://www.spss-tutorials.com/measurement-levels/#nominal-variable) with few categories -such as nationality, blood group or profession. If these are string variables, they may or may not cause trouble. For example, the independent variable for [ANOVA](https://www.spss-tutorials.com/anova-what-is-it/) may or may not be a [string variable](https://www.spss-tutorials.com/spss-string-variables-basics/) depending on the exact command you use for it.\*  
You may get away by leaving such variables as strings. However, copying them into numeric variables makes sure you'll avoid all trouble. A decent way to do so is [AUTORECODE](https://www.spss-tutorials.com/spss-autorecode-command/). For converting metric string variables -holding just numbers- into numeric variables, see [SPSS Convert String to Numeric Variable](https://www.spss-tutorials.com/spss-convert-string-to-numeric-variable/).

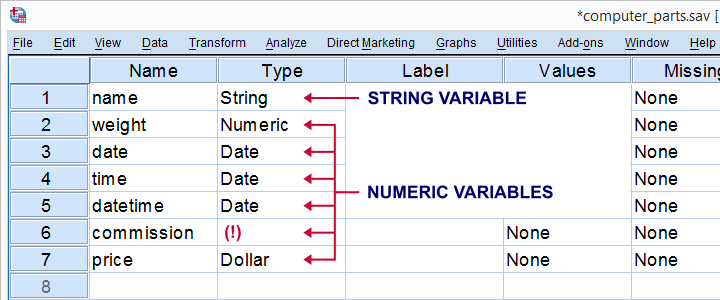
## Determining SPSS Variable Types

So how do we know if a variable is string or numeric? In SPSS versions 24 and higher, tiny icons in front of variable names tell us the variable type, format and even measurement level. The icon for “nominal” may contain a tiny **“a”** which indicates it's a string variable.



For SPSS versions 23 and earlier, we'll inspect our [variable view](https://www.spss-tutorials.com/spss-data-editor-window/#spss-variable-view) and use the following rule:

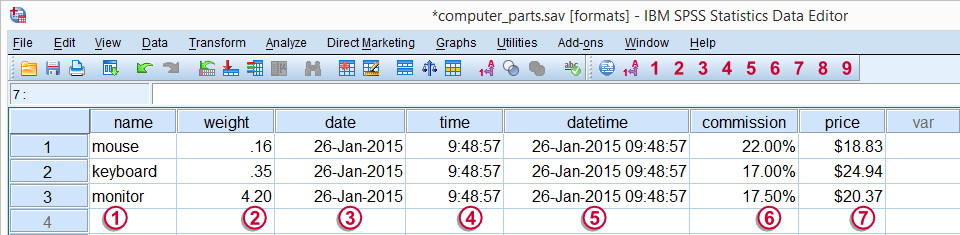
* if Type says “String”, you're dealing with a string variable;
* if Type does not say “String”, you're dealing with a numeric variable.



SPSS suggests that “Date” and “Dollar” are variable types as well. However, these are formats, **not types**. The way they are shown here among the actual variable types (string and numeric) is one of SPSS’ most confusing features.

## SPSS Variable Formats - Introduction

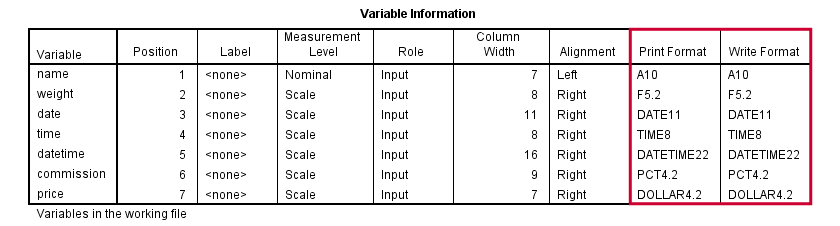
Let's now have a look at the data in data view as shown the screenshot below. We'll briefly describe the kinds of variables we see.



Regarding these data, we stated earlier that  
  
https://spss-tutorials.com/img/b1.png is a string variable and  
https://spss-tutorials.com/img/b2.png through https://spss-tutorials.com/img/b7.png are numeric variables and contain **only numbers**.  
  
However, values such as “26-jan-2015” sure don't look like numbers, do they? This is because SPSS can display numbers in very different ways. These ways of displaying data values are referred to as **variable formats**.

## Determining SPSS Variable Formats

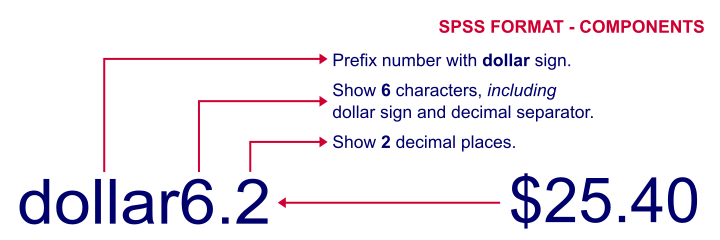
As we saw earlier, “Type” under variable view shows a confusing mixture of variable types and formats. We'll see the actual formats by runningdisplay dictionary.Part of the result is shown by the screenshot below.



SPSS distinguishes print and write formats but we don't bother about this distinction. SPSS variable formats consist of two parts. One or more letters indicate the **format family**. Most of them speak to themselves, except for the first two variables:

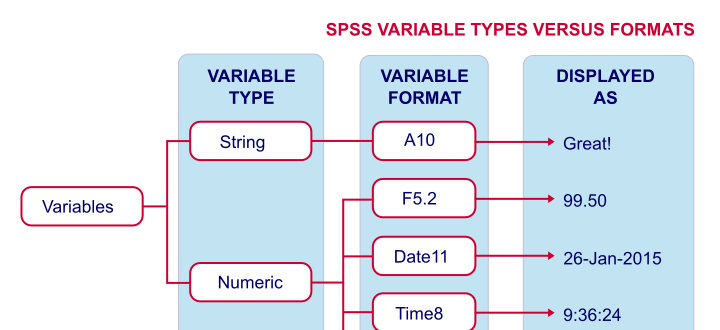
* [**A**](https://www.spss-tutorials.com/spss-string-variables-basics/#a-format) (“Alphanumeric”) is the usual format for **string variables**;
* **F**, (“Fortran”) indicates a **standard numeric variable**.

Formats end with numbers, indicating the number of characters to be shown. If a period is present, the number after the period indicates the number of decimal places to be displayed. The figure below illustrates these points.



## SPSS Common Variable Formats

The figure below now summarizes some common variable types and formats we'll encounter in SPSS.



## Setting Variable Formats in SPSS

You can set variable formats for numeric variables with the [FORMATS](https://www.spss-tutorials.com/spss-formats-command/) command. For example,formats weight (f4.3).shows weight with 3 decimal places. Doing so affects the [output](https://www.spss-tutorials.com/spss-output-viewer-window/) you create: most tables will add an extra decimal place for weight as well. If you'd like to see this for yourself, run the syntax below and compare the 2 resulting tables.

**\*Show 2 decimal places for weight and run descriptives.**  
formats weight(f3.2). : float 123.12  
  
descriptives weight.  
  
**\*Show 3 decimal places for weight and run descriptives.**  
formats weight(f4.3).  
  
descriptives weight.  
  
**\*Note that second output table shows more decimal places.**

Keep in mind that**changing variable formats does not change your data**in any way. The actual values are still the exact same numbers. They are merely displayeddifferently.

## Variable Types and Formats - Why Bother?

Basically, “what you see is not what you get” in data view. For example, we see $20.37 but the actual value is just 20.37. So we can identify products costing $20,- or more by [running the syntax](https://www.spss-tutorials.com/spss-syntax/#paste-spss-syntax) below:compute expensive = (price >= 20).We **don't include the dollar sign** in our syntax. Although SPSS shows a dollar sign in data view, the actual values are just numbers and these are what the syntax acts upon.  
  
Or let's say we'd like to add 30 days to our date variable. We could do so by runningcompute newdate = datesum(date,30,'days').The resulting values are 13644236937.72. These are the correct numbers but they'll display as readable dates only after running something likeformats newdate (date11).Another reason for bothering about variable formats is setting **decimals places for output** tables. For SPSS version 22 onwards, [OUTPUT MODIFY](https://www.spss-tutorials.com/spss-output-modify-command/) does the trick as shown below.

**\*Create basic descriptives table.**  
descriptives weight.  
  
**\*Set 2 decimal places (format = f3.2) for mean and SD (columns 4 and 5).**  
output modify  
/select tables   
/tablecells select = [position(4) position(5)] selectdimension = columns format = 'f3.2'.

In a similar vein, CTABLES allows choosing different formats for different statistics in your output.

**\*Create descriptives table with different decimal places for different statistics.**  
ctables  
/table commission [count 'N' f3 Minimum pct3 Maximum pct3 mean 'Mean

## Overview Most Useful SPSS Shortkeys

|  |  |  |
| --- | --- | --- |
| **WHERE** | **SHORTKEY** | **USEFUL FOR** |
| Data Editor | CTRL + t | Switch between data view and variable view. |
| Data Editor | CTRL + keyboard arrow / keyboard arrow | Go to first/last variable/case (depending on view). Allows for very quick case and variable count when combined with CTRL + t. |
| Syntax Window | Shift + keyboard arrow / keyboard arrow | Select lines of [syntax](https://www.spss-tutorials.com/spss-syntax/) above/below cursor position. |
| Syntax Window | F2 | Select Entire Command in which cursor is located. |
| Syntax Window | CTRL + r | Run selected syntax. |
| Any Window | CTRL + Home / End | Go to beginning/end of window contents. |
| Syntax Window | CTRL + c | Copy selected syntax. |
| Syntax Window | CTRL + x | Cut selected syntax. |
| Syntax Window | CTRL + v | Paste selected syntax. |
| Syntax Window | CTRL + a | Select all syntax in window. |
| Syntax Window | CTRL + f | Find expression within syntax window. |
| Syntax Window | CTRL + h | Replace expression within syntax window. |
| Syntax Window | CTRL + z | Undo last edit(s). Note: doesn't always seem to work properly in recent SPSS versions. |
| Syntax Window | Home / End | Move to start/end of line. |
| Any Window | Alt + tab | Switch between Syntax Window, Data Editor and Viewer Windows. Actually an MS Windows shortkey for switching between applications. |
| Any Window | Alt + F4 | Close any window. You'll usually get a pop-up window asking whether you'd like to save it if you haven't done so yet. |
| Syntax Window | F1 | When your cursor is located anywhere in a command, help for this command can be opened by pressing F1. |
| Syntax Window | Shift + Home / End | Select from cursor through start/end of line. |
| Any Window | CTRL + Shift + Home / End | Select from cursor through beginning/end of window contents. |
| Syntax Window | CTRL + y | Redo last edit(s). |
| Any Window | Alt + F6 | Back to previous window. Handy for switching between Syntax Window and Data Editor. |