

What are the different types of variables?

Authored by
stats writer

June 23, 2024

RECOMMENDED CITATION

stats writer (2024). *What are the different types of variables?*. PSYCHOLOGICAL SCALES.
Retrieved from <https://scales.arabpsychology.com/?p=148963>

Variables are an essential concept in mathematics, statistics, and computer programming. They are symbols or placeholders that represent quantities or values and can be manipulated or changed in a given context. There are several types of variables, including independent variables, dependent variables, discrete variables, and continuous variables. Independent variables are those that are controlled or manipulated in an experiment, while dependent variables are the ones that are affected by the changes in the independent variables. Discrete variables can only take on specific values, while continuous variables can take on any value within a range. Other types of variables include categorical variables, which represent different categories or groups, and quantitative variables, which represent numerical values. Understanding the different types of variables is crucial in interpreting and analyzing data accurately.

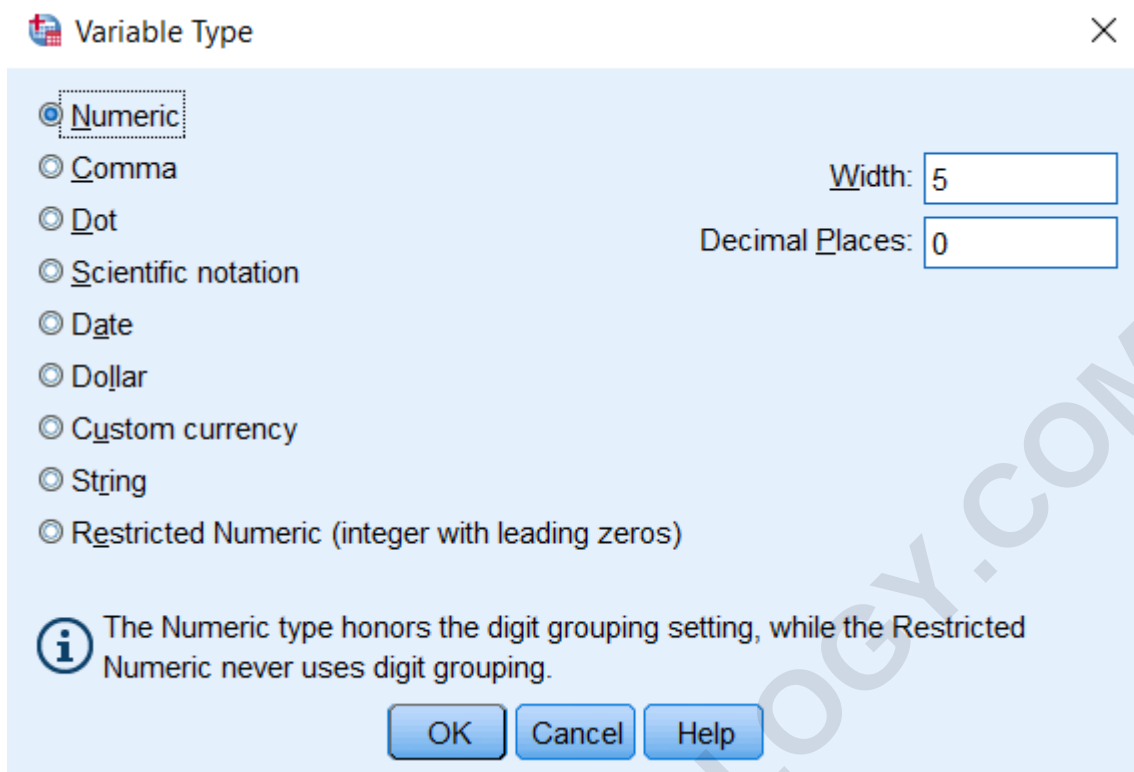
Variable Types

In order for your data analysis to be accurate, it is imperative that you correctly identify the *type* and *formatting* of each variable. SPSS has special restrictions in place so that statistical analyses can't be performed on inappropriate types of data: for example, you won't be able to use a continuous variable as a "grouping" variable when performing a *t*-test.

Information for the type of each variable is displayed in the **Variable View** tab. Under the "Type" column, simply click the cell associated with the variable of interest. A blue "..." button will appear.

Name	Type
ID	Numeric 

Click this and the Variable Type window will appear. You can use this dialog box to define the type for the selected variable, and any associated information (e.g., width, decimal places).



The image shows the 'Variable Type' dialog box in SPSS. It has a title bar with a red cross icon and the text 'Variable Type'. The dialog is light blue. On the left, there is a list of variable types with radio buttons: 'Numeric' (selected), 'Comma', 'Dot', 'Scientific notation', 'Date', 'Dollar', 'Custom currency', 'String', and 'Restricted Numeric (integer with leading zeros)'. On the right, there are two input fields: 'Width' with the value '5' and 'Decimal Places' with the value '0'. At the bottom, there is an information icon (i) followed by the text: 'The Numeric type honors the digit grouping setting, while the Restricted Numeric never uses digit grouping.' Below this text are three buttons: 'OK', 'Cancel', and 'Help'. A large, diagonal watermark 'ARABPSYCHOLOGY.COM' is visible across the dialog box.

The two common types of variables that you are likely to see are *numeric* and *string*.

Numeric

Numeric variables, as you might expect, have data values that are recognized as numbers. This means that they can be sorted numerically or entered into arithmetic calculations. When viewed in the Data View window, system-missing values for numeric variables will appear as a dot (i.e., "."). (Note that one should not type in a period character in a cell to specify a missing value. Simply leave the cell blank, and SPSS will recognize it as system-missing.)

Importantly, numeric variables in SPSS can also be used to denote nominal (unordered) or ordinal categorical variables. In those cases, it is almost always inappropriate to treat those variables as numbers, even though SPSS may not stop you from doing so. For example, it's extremely common to record demographic variables like sex using the number codes 1 and 2 instead of the words "male" and "female". Although these would be defined as numeric variables in your SPSS dataset, it would not be appropriate to use them in arithmetic operations, since the number codes are stand-ins for nominal categories (and nominal categories can't be used in arithmetic operations). So if you are examining a new dataset, you should not assume that all numeric variables represent interval or ratio variables.

All of the following are examples of variables that could be entered as numeric variables in an

SPSS dataset:

Example: Continuous variables that can take on any number in a range (e.g., height in centimeters and weight in kilograms) should be treated as numeric variables. The researcher can choose as many or as few decimal places as they feel are necessary. In this situation, the *Measure* setting should be defined as *Scale*; see the [Defining Variables](#) tutorial for more information on how to set measurement levels. This particular type of numeric variable is appropriate to use in arithmetic operations (adding, subtracting, multiplying, dividing).

Example: Counts (e.g., number of people living in a household) should be treated as numeric variables with zero decimal places. In this situation, the *Measure* setting should be defined as *Scale*. Certain mathematical calculations are valid when applied to count variables (e.g., mean and standard deviation), but some statistical procedures requiring continuous numeric variables may not be (e.g., the dependent variable in a linear regression), depending on the distribution of the variable.

Example: Nominal categorical variables that have been coded numerically (e.g., recording a subject's gender as 1 if male or 2 if female) should be treated as numeric variables with zero decimal places. In this situation, the *Measure* setting must be defined as *Nominal*. This type of numeric variable should *never* be used in mathematical calculations, nor used in any statistical procedure requiring continuous numeric variables (e.g. the dependent variable of a linear regression).

Example: Ordinal categorical variables that have been coded numerically (e.g., a questionnaire item with responses 1=Small, 2=Medium, 3=Large) should be treated as numeric variables with zero decimal places. In this situation, the number codes allow us to correctly convey that *Large* is "greater than" *Small* in a meaningful way; however, it is not safe to assume that the "distance" between *Large* and *Medium* is the same as the "distance" between *Medium* and *Small*. (This is because our choice of number codes is arbitrary and not tied to any physical meaning.) In this situation, the *Measure* setting must be defined as *Ordinal*. This type of numeric variable should never be used in any statistical procedure requiring continuous numeric variables (e.g. the dependent variable of a linear regression), and in most situations it is not appropriate to use ordinal variables in mathematical calculations, though there are some notable exceptions. (One such example is computing a composite score for a validated survey instrument by summing or averaging its constituent Likert items, though this is not without controversy.)

String

String variables -- which are also called *alphanumeric variables* or *character variables* -- have values that are treated as text. This means that the values of string variables may include numbers, letters, or symbols. In the Data View window, missing string values will appear as blank

cells. **However**, note that these blank cells are *not* recognized by SPSS as system-missing values (i.e., SPSS considers even blank strings to be non-missing)! This has important implications if you plan to use a string variable in an analysis, since it will affect your sample size.

Example: Zip codes and phone numbers, although composed of numbers, would typically be treated as string variables because their values cannot be used meaningfully in calculations.

Example: Any written text is considered a string variable, including free-response answers to survey questions.

The next few variable types are all technically numeric, but indicate special formatting. If your data has been recorded in one of these formats, you must set the variable type appropriately so that SPSS can interpret the variables correctly. (For example, SPSS can not correctly use dates in calculations unless the variables are specifically defined as date variables.)

Comma

Numeric variables that include commas that delimit every three places (to the left of the decimals) and use a period to delimit decimals. SPSS will recognize these values as numeric even if they contain commas or use scientific notation.

Example: Thirty-thousand and one half: 30,000.50

Example: One million, two hundred thirty-four thousand, five hundred sixty-seven and eighty-nine hundredths: 1,234,567.89

Dot

Numeric variables that include periods that delimit every three places and use a comma to delimit decimals. SPSS will recognize these values as numeric even if they contain periods or use scientific notation.

Example: Thirty-thousand and one half: 30.000,50

Example: One million, two hundred thirty-four thousand, five hundred sixty-seven and eighty-nine hundredths: 1.234.567,89

Note about comma versus dot notation: comma notation is standard in the United States. Oracle's [International Language Environments Guide](#) gives a list of countries and what form of notation is typically found in each.

Scientific notation

Numeric variables whose values are displayed with an E and power-of-ten exponent. Exponents can be preceded by either an E or a D, with or without a sign, or only with a sign (no E or D). SPSS will recognize these values as numeric, with or without an exponent.

Example: 1.23E2, 1.23D2, 1.23E+2, 1.23+2.

Date

Numeric variables that are displayed in any standard calendar date or clock-time formats. Standard formats may include commas, blank spaces, hyphens, periods, or slashes as space delimiters.

Example:Dates: 01/31/2013, 31.01.2013

Example:Time: 01:02:33.7

Dollar

Numeric variables that contain a dollar sign (i.e., \$) before numbers. Commas may be used to delimit every three places, and a period can be used to delimit decimals.

Example: Thirty-three thousand dollars and thirty-three cents: \$33,000.33

Example: One million dollars and twelve point three cents: \$1,000,000.123

Custom currency

Numeric variables that are displayed in a custom currency format. You must define the custom currency in the Variable Type window. Custom currency characters are displayed in the Data Editor but cannot be used during data entry.

Restricted number

Numeric variables whose values are restricted to non-negative integers (in standard format or scientific notation). The values are displayed with leading zeroes padded to the maximum width of the variable.

Example: 00000123456 (width 11)