**Types of variables**

A variable is defined as a feature which can be used to store data of unknown numbers or alphabets. Variables can be things like height, age, income, province of birth, school grades, and kind of dwelling. Categorical and quantitative variables are the two basic categories into which variables can be divided. Then, each category is divided into two subcategories: discrete or continuous for quantitative variables, and nominal or ordinal for categorical variables.

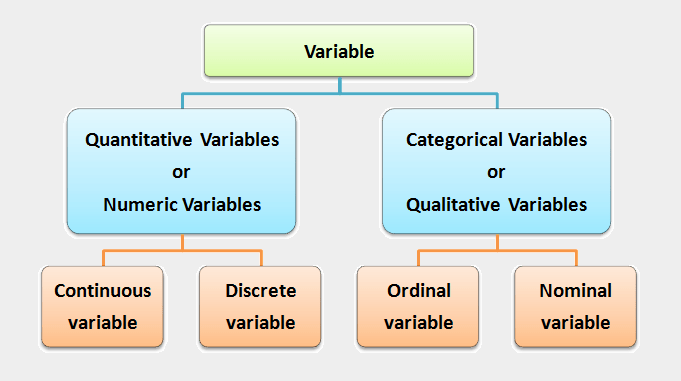


Fig-1: Types of Variables

**Quantitative variables**

When you collect quantitative data, the numbers you record represent real amounts that can be added, subtracted, divided, etc. There are two types of quantitative variables: discrete and continuous.

**Table 1. Discrete vs continuous variables**

|  |  |  |
| --- | --- | --- |
| **Type of variable** | **What does the data represent?** | **Examples** |
| Discrete variables (aka integer variables) | Counts of individual items or values. | 1. Number of students in a class 2. Number of different tree species in a forest |
| Continuous variables (aka ratio variables) | Measurements of continuous or non-finite values. | 1. Distance 2. Volume 3. Age |

**Categorical variables**

Categorical variables represent groupings of some kind. They are sometimes recorded as numbers, but the numbers represent categories rather than actual amounts of things.

There are three types of categorical variables: binary, nominal, and ordinal variables.

**Table 2. Binary vs nominal vs ordinal variables**

|  |  |  |
| --- | --- | --- |
| **Type of variable** | **What does the data represent?** | **Examples** |
| Binary variables (aka dichotomous variables) | Yes or no outcomes. | 1. Heads/tails in a coin flip 2. Win/lose in a football game |
| Nominal variables | Groups with no rank or order between them. | 1. Species names 2. Colors 3. Brands |
| Ordinal variables | Groups that are ranked in a specific order. | 1. Finishing place in a race 2. Rating scale responses in a survey, such as Likert scales\* |

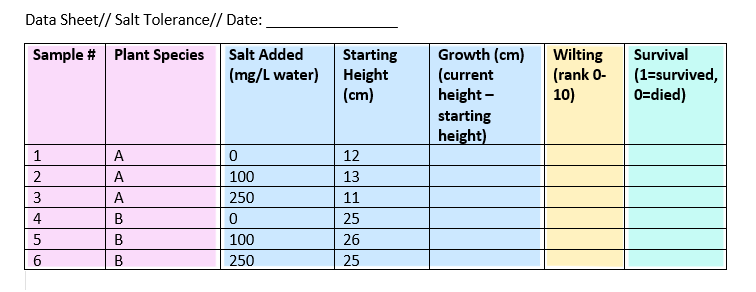
\*Note that sometimes a variable can work as more than one type! An ordinal variable can also be used as a quantitative variable if the scale is numeric and doesn’t need to be kept as discrete integers. For example, star ratings on product reviews are ordinal (1 to 5 stars), but the average star rating is quantitative.

Reference Video: <https://www.youtube.com/watch?v=FqB5Es1HXI4>

**Example data sheet**

To keep track of your salt-tolerance experiment, you make a data sheet where you record information about the variables in the experiment, like salt addition and plant health.

To gather information about plant responses over time, you can fill out the same data sheet every few days until the end of the experiment. This example sheet is color-coded according to the type of variable: nominal, continuous, ordinal, and binary.



Nominal Data Continuous, Ordinal, Binary

**Independent vs Dependent variables Vs Control Variable:**

Experiments are usually designed to find out what effect one variable has on another – in our example, the effect users review on products.

You manipulate the independent variable (the one you think might be the cause) and then measure the dependent variable (the one you think might be the effect) to find out what this effect might be.

A control variable is anything that is held constant or limited in a research study. It’s a variable that is not of interest to the study’s objectives, but is controlled because it could influence the outcomes.

**Table 3. Independent vs Dependent variables Vs Control Variable**

|  |  |  |
| --- | --- | --- |
| **Type of variable** | **What does the data represent?** | **Examples** |
| Independent variables | Variables you manipulate in order to affect the outcome of an experiment. | 1. No. of brand name 2. No. of Product 3. User Name |
| Dependent variables | Variables that represent the outcome of the experiment. | 1. Star Rating 2. Reviews |
| Control variables | Variables that are held constant throughout the experiment. | 1. No. of hours of study, survey. 2. Age of respondents. |

**Composite Variable:**

Two or more variables combined to make a more complex variable. A composite variable is a type of variable that is formed by combining two or more other variables. It is also sometimes referred to as a derived variable.

**For example:** In google search rank task, We have a Text component having some variables: Product Name, product Features, Weight, Quantity, Brand Name etc.

This all things create a combined variable / component called Text Component.

**Confounding Variable:**

A confounding variable is a third variable that influences both the independent and dependent variables. It’s important to consider potential confounding variables and account for them in your research design to ensure your results are valid. Left unchecked, confoudning variables can introduce many research biases to your work, causing you to misinterpret your results.

**For example:** Weight and Packaging and Star Rating of Products.

Sometime, less weight products gets more buyers and more star rating. Colourful packaging get more buyers. But that doesn’t mean these products are good in quality.