TYPES OF DATA

THREE TYPES OF DATA

- **Qualitative data** consist of words (Yes or No), letters (Y or N), or numerical codes (0 or 1) that represent a class or category.
- **Ranked data** consist of numbers (1st, 2nd, . . . 40th place) that represent relative standing within a group.
- **Quantitative data** consist of numbers (weights of 238, 170, . . . 185 lbs) that represent an amount or a count. To determine the type of data, focus on a single observation in any collection of observations

TYPES OF VARIABLES

A **variable** is a characteristic or property that can take on different values.

- The weights can be described not only as quantitative data but also as observations for a quantitative variable, since the various weights take on different numerical values.
- By the same token, the replies can be described as observations for a qualitative variable, since the replies to the Facebook profile question take on different values of either Yes or No.
- Given this perspective, any single observation can be described as a constant, since it takes on only one value.

Discrete and Continuous Variables

Quantitative variables can be further distinguished as discrete or continuous. A discrete variable consists of isolated numbers separated by gaps. Discrete variables can only assume specific values that you cannot subdivide. Typically, you count discrete values, and the results are integers.

Examples

- Counts- such as the number of children in a family. (1, 2, 3, etc., but never 1.5)
- These variables cannot have fractional or decimal values. You can have 20 or 21 cats, but not 20.5
- The number of heads in a sequence of coin tosses.
- The result of rolling a die.
- The number of patients in a hospital.
- The population of a country.

While discrete variables have no decimal places, the average of these values can be fractional. For example, families can have only a discrete number of children: 1, 2, 3, etc. However, the average number of children per family can be 2.2.

A **continuous variable** consists of numbers whose values, at least in theory, have no restrictions. Continuous variables can assume any numeric value and can be meaningfully split into smaller parts. Consequently, they have valid fractional and decimal values. In fact, continuous variables have an infinite number of potential values between any two points. Generally, you measure them using a scale.

Examples of continuous variables include weight, height, length, time, and temperature. Durations, such as the reaction times of grade school children to a fire alarm; and standardized test scores, such as those on the Scholastic Aptitude Test (SAT).

Independent and Dependent Variables

Independent Variable

In an experiment, an **independent variable** is the treatment manipulated by the investigator.

- Independent variables (IVs) are the ones that you include in the model to explain or predict changes in the dependent variable.
- Independent indicates that they stand alone and other variables in the model do not influence them.
- Independent variables are also known as predictors, factors, treatment variables, explanatory variables, input variables, x-variables, and right-hand variables—because they appear on the right side of the equals sign in a regression equation.
- It is a variable that stands alone and isn't changed by the other variables you are trying to measure.

For example, someone's age might be an independent variable. Other factors (such as what they eat, how much they go to school, how much television they watch).

The impartial creation of distinct groups, which differ only in terms of the independent variable, has a most desirable consequence. Once the data have been collected, any difference between the groups can be interpreted as being *caused* by the independent variable.

Dependent Variable

When a variable is believed to have been influenced by the independent variable, it is called a **dependent variable.** In an experimental setting, the dependent variable is measured, counted, or recorded by the investigator.

- The dependent variable (DV) is what you want to use the model to explain or predict. The values of this variable depend on other variables.
- It’s also known as the response variable, outcome variable, and left-hand variable. Graphs place dependent variables on the vertical, or Y, axis.
- A dependent variable is exactly what it sounds like. It is something that depends on other factors.

For example, the blood sugar test depends on what food you ate, at which time you ate etc. Unlike the independent variable, the dependent variable isn’t manipulated by the investigator. Instead, it represents an outcome: the data produced by the experiment.

Confounding Variable

An uncontrolled variable that compromises the interpretation of a study is known as a **confounding variable.** Sometimes a confounding variable occurs because it's impossible to assign subjects randomly to different conditions.