Variables

A **variable** is a value or characteristic that can be different from individual to individual.

Variables

A **variable** is a value or characteristic that can be different from individual to individual.

Examples:

- favorite color
- age
- population (the "individuals" might be countries or cities, for example)

Quantitative variables

Quantitative variables have **numerical values**, often with **units of measurement**.

Quantitative variables

Quantitative variables have **numerical values**, often with **units of measurement**.
They arise as answers to questions like, "How much?" or, "How many?"

Quantitative variables

Quantitative variables have **numerical values**, often with **units of measurement**.
They arise as answers to questions like, "How much?" or, "How many?"

Examples:

- age (possible units: years)
- annual income (possible units: thousands of dollars)
- number of children (this is a "count" and has no other units)

Continuous and discrete quantitative variables

Sometimes we are willing to believe that the values of a quantitative variable can be arbitrarily close to each other. Such a variable is called **continuous**.

Continuous and discrete quantitative variables

Sometimes we are willing to believe that the values of a quantitative variable can be arbitrarily close to each other. Such a variable is called **continuous**.

Height, weight, and age are examples of variables often considered continuous, though in practice we can only measure them up to a certain degree of accuracy.

Continuous and discrete quantitative variables

Sometimes we are willing to believe that the values of a quantitative variable can be arbitrarily close to each other. Such a variable is called **continuous**.

Height, weight, and age are examples of variables often considered continuous, though in practice we can only measure them up to a certain degree of accuracy.

But quantitative variables like "number of children" take on values that are separated from each other by fixed amounts. For example, the number of children in a household can be 0, 1, 2, and so on, consecutive values being separated by 1. But it can't be 1.25. Such variables are called **discrete**.

Continuous and discrete quantitative variables

Sometimes we are willing to believe that the values of a quantitative variable can be arbitrarily close to each other. Such a variable is called **continuous**.

Height, weight, and age are examples of variables often considered continuous, though in practice we can only measure them up to a certain degree of accuracy.

But quantitative variables like "number of children" take on values that are separated from each other by fixed amounts. For example, the number of children in a household can be 0, 1, 2, and so on, consecutive values being separated by 1. But it can't be 1.25. Such variables are called **discrete**.

The distinction between continuous and discrete isn't absolute.

Continuous and discrete quantitative variables

Sometimes we are willing to believe that the values of a quantitative variable can be arbitrarily close to each other. Such a variable is called **continuous**.

Height, weight, and age are examples of variables often considered continuous, though in practice we can only measure them up to a certain degree of accuracy.

But quantitative variables like "number of children" take on values that are separated from each other by fixed amounts. For example, the number of children in a household can be 0, 1, 2, and so on, consecutive values being separated by 1. But it can't be 1.25. Such variables are called **discrete**.

The distinction between continuous and discrete isn't absolute. Sometimes we will "discretize" a continuous variable, for example by measuring age in "completed years."

Continuous and discrete quantitative variables

Sometimes we are willing to believe that the values of a quantitative variable can be arbitrarily close to each other. Such a variable is called **continuous**.

Height, weight, and age are examples of variables often considered continuous, though in practice we can only measure them up to a certain degree of accuracy.

But quantitative variables like "number of children" take on values that are separated from each other by fixed amounts. For example, the number of children in a household can be 0, 1, 2, and so on, consecutive values being separated by 1. But it can't be 1.25. Such variables are called **discrete**.

The distinction between continuous and discrete isn't absolute.
Sometimes we will "discretize" a continuous variable, for example by measuring age in "completed years."
And sometimes, especially when we're doing approximations, we will consider discrete variables like test scores to be continuous.

Categorical and qualitative variables

Some variables don't have numerical values; rather, the data fall into categories that have no particular order or ranking in relation to each other. Such variables are called **categorical** or **qualitiative**.

Categorical and qualitative variables

Some variables don't have numerical values; rather, the data fall into categories that have no particular order or ranking in relation to each other. Such variables are called **categorical** or **qualitiative**.

Examples:

- favorite color
- gender
- nationality

Categorical and qualitative variables

Some variables don't have numerical values; rather, the data fall into categories that have no particular order or ranking in relation to each other. Such variables are called **categorical** or **qualitiative**.

Examples:

- favorite color
- gender
- nationality

Some variables aren't numerical but their values have a natural order or ranking; such variables are sometimes called **qualitative ordinal**.

Categorical and qualitative variables

Some variables don't have numerical values; rather, the data fall into categories that have no particular order or ranking in relation to each other. Such variables are called **categorical** or **qualitiative**.

Examples:

- favorite color
- gender
- nationality

Some variables aren't numerical but their values have a natural order or ranking; such variables are sometimes called **qualitative ordinal**.

Examples:

- temperature: low/medium/high
- job performance: poor/satisfactory/excellent

Categorical and qualitative variables

Some variables don't have numerical values; rather, the data fall into categories that have no particular order or ranking in relation to each other. Such variables are called **categorical** or **qualitiative**.

Examples:

- favorite color
- gender
- nationality

Some variables aren't numerical but their values have a natural order or ranking; such variables are sometimes called **qualitative ordinal**.

Examples:

- temperature: low/medium/high
- job performance: poor/satisfactory/excellent

Again, these distinctions aren't absolute, and we won't spend a lot of time worrying about them.

A word of caution

Just because a variable has numerical values, that doesn't mean it is quantitative.

A word of caution

Just because a variable has numerical values, that doesn't mean it is quantitative.

For example, if your variable is "favorite color," you can assign a numerical code to each color and say that "red" is 1, "blue" is 2, and so on. That doesn't change the variable from categorical to quantitative.

A word of caution

Just because a variable has numerical values, that doesn't mean it is quantitative.

For example, if your variable is "favorite color," you can assign a numerical code to each color and say that "red" is 1, "blue" is 2, and so on. That doesn't change the variable from categorical to quantitative.

You can't do any arithmetic on those numbers. For example, it doesn't make sense to take their average. The assigned numerical values are just labels, as were the original words "red," "blue," and so on.

A word of caution

Just because a variable has numerical values, that doesn't mean it is quantitative.

For example, if your variable is "favorite color," you can assign a numerical code to each color and say that "red" is 1, "blue" is 2, and so on. That doesn't change the variable from categorical to quantitative.

You can't do any arithmetic on those numbers. For example, it doesn't make sense to take their average. The assigned numerical values are just labels, as were the original words "red," "blue," and so on.

But if you code yes/no answers as 1/0, then the average of all your 1's and 0's does indeed make sense and turns out to be a very interesting quantity. We'll see this later in the course; stay tuned!