

[See this page in the course material.](#)

Learning OUTCOMES

- Distinguish between quantitative and categorical variables in context.

Data consist of **individuals** and **variables** that give us information about those individuals. An individual can be an object or a person. A variable is an attribute, such as a measurement or a label.

Example

Medical Records

This dataset is from a medical study. In this study, researchers wanted to identify variables connected to low birth weights.

	Age at delivery	Weight prior to pregnancy (pounds)	Smoker	Doctor visits during 1st trimester	Race	Birth Weight (grams)
Patient 1	29	140	Yes	2	Caucasian	2977
Patient 2	32	132	No	4	Caucasian	3080
Patient 3	36	175	No	0	African-American	3600
*	*	*	*	*	*	*
*	*	*	*	*	*	*
Patient 189	30	95	Yes	2	Asian	3147

In this example, the individuals are the patients (the mothers). There are six variables in this dataset:

- Mother's age at delivery (years)
- Mother's weight prior to pregnancy (pounds)
- Whether mother smoked during pregnancy (yes, no)
- Number of doctor visits during first trimester of pregnancy
- Mother's race (Caucasian, African American, Asian, etc.)
- Baby's birth weight (grams)

There are two types of variables: quantitative and categorical.

- **Categorical variables** take category or label values and place an individual into one of several groups. Each observation can be placed in only one category, and the categories are mutually exclusive. In our example of medical records, smoking is a categorical variable, with two groups, since each participant can be categorized only as either a nonsmoker or a smoker. Gender and race are the two other categorical variables in our medical records example.
- **Quantitative variables** take numerical values and represent some kind of measurement. In our medical example, age is an example of a quantitative variable because it can take on multiple numerical values. It also makes sense to think about it in numerical form; that is, a person can be 18 years old or 80 years old. Weight and height are also examples of quantitative variables.

Try It

We took a random sample from the 2000 US Census. Here is part of the dataset.

Sample of 2000 US Census Data

State	Zipcode	Family_Size	Annual_Income
Florida	32716	8	200
Alabama	35236	5	800
Florida	32116	6	13500
Florida	33679	5	21000
Alabama	36374	4	21000
California	94565	1	23000

Who are the individuals described in this data?

☐ People with families in the year 2000

[See this interactive in the course material.](#)

For the following questions, indicate whether the variable is categorical, quantitative or neither.

zip code [Select]

family size [Select]

annual income [Select]

Check Answer

[See this interactive in the course material.](#)

Try It

Consumer Reports analyzed a dataset of 77 breakfast cereals. Here is a part of the dataset.

(Note: *Consumer Reports* is an non-profit organization that rates products in an effort to help consumers make informed decisions.)

Sample of Consumer Reports Breakfast Cereal Data

Name	Manufactuer	Target	Shelf	Calories	Sodium	Fat
100% Bran	Nabisco	adult	top	70	130	1
100% Natural Bran	Quaker Oats	adult	top	120	15	5
All-Bran	Kelloggs	adult	top	70	260	1
All-Bran Extra Fiber	Kelloggs	adult	top	50	140	0
Almond Delight	Ralston Purnia	adult	top	110	200	2
Apple Cinnamon Cheerios	General Mills	child	bottom	110	180	2
Apple Jacks	Kelloggs	child	middle	110	125	0

Who are the individuals described in this data?

☐ Cereal manufacturers

[See this interactive in the course material.](#)

For the following questions, indicate whether the variable is categorical, quantitative or neither.

Cereal name [Select]

Cereal manufacturer [Select]

Target [Select]

Calories [Select]

Fat [Select]

[Check Answer](#)

[See this interactive in the course material.](#)

Licenses and Attributions

CC licensed content, Shared previously

- Concepts in Statistics. **Provided by:** Open Learning Initiative. **Located at:** <http://oli.cmu.edu>. **License:** [CC BY: Attribution](#)

</div