1.1 Statistics: The Science and Art of Data

Statistics is the science and art of collecting, analyzing, and drawing conclusions from data.
To understand statistics, we have to outline the steps of the **statistical problem-solving process**.
1. Ask Questions: Clarify the research problem and ask one or more valid statistics questions.
2. Collect Data: Design and carry out an appropriate plan to collect the data.
3. Analyze Data: Use appropriate graphical and numerical methods to analyze the data.
4. Interpret Results: Draw conclusions based on the data analysis. Be sure to answer the research questions!

Most data tables follow a format where each row describes an individual and each column holds the values of a variable.
Sometimes the individuals in a data set are called *cases* or *observational units*.

An **individual** is a person, animal, or thing described in a set of data. A **variable** is any attribute that can take different values for different individuals.

Variables can be broken down into two different types:
A **categorical variable** assigns labels that place individuals into particular groups.
A **quantitative variable** takes number values for which it makes sense to find an average.
Not every variable that takes number values is quantitative

Example: Identify the individuals and variables in the data set. Classify each variable as categorical or quantitative.

	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-Am	3600
*	*	*	*	*	*	*
*	*	*	*	*	*	*
Patient 189	30	95	Yes	2	Asian	3147

Individuals: Patients

Variables: age @ delivery, weight, smoker, doctor visits, race, birth weight

Categorical: Smoker, race, age @ delivery

Quantitative: weight prior to pregnancy, doctor visits, birth weight