Statistics involves
1. Asking questions
2. Collecting data to address those questions
3. Analyzing the data
4. Drawing conclusions from the data
5. Communicating the results

- The **observational units** in a statistical study are the objects described by a set of data (people, animals, things).
- A **variable** is any characteristic of an observational unit, which can take different values (i.e., can vary) for different observational units.
  - Some variables are **quantitative**, taking numerical values on which ordinary arithmetic operations make sense.
  - Other variables are **categorical**, taking category designations.
    - A categorical variable with only two categories is called **binary**.

**Example 1-1: Variables on You**
Consider the students in this class as the observational units in a statistical study.

a) For each of the following variables, indicate whether it is quantitative or categorical:
   - How many *Harry Potter* books have you read?
   - With which hand do you write?
   - How many hours have you slept in the past 24 hours?
   - Have you slept for at least 7 hours in the past 24 hours?
   - On what day of the week were you born?
   - Do you have a Facebook account?
   - How many Facebook friends do you have?

b) Explain why the following questions are *not* variables:
   - What is the average number of *Harry Potter* books read by a student in this class?
   - What percentage of students in this class are left-handed?

c) What would the observational units have to be in order for these questions in b) to be legitimate variables?

d) Explain why the following questions are *not* variables:
   - Have women at Cal Poly read more *Harry Potter* books, on average, than men?
   - Is the proportion of undergraduate Cal Poly students with a Facebook account smaller than the proportion of Cal Poly graduate students with a Facebook account?
Summaries are not variables.
Research questions are not variables.

Example 1-2: Statistical Studies
For each of the following questions that led to statistical studies, identify the observational units and variables. Also classify each variable as quantitative or categorical (also binary?).

a) How much did an average American consumer spend on Christmas presents in 2016?

b) Is the price of a house related to its size?

c) Is the residence situation of a college student (on-campus, off-campus with parents, off-campus without parents) related to how much alcohol the student consumes?

d) Can you predict how far a cat can jump based on factors such as its length and weight?

e) Among heterosexual couples, who is more likely to say “I love you” first: the man or the woman?

f) An article in a 2006 issue of Journal of Behavioral Decision Making reports on a study involving 47 undergraduate students at Harvard. All of the participants were given $50, but some (chosen at random) were told that this was a “tuition rebate,” while the others were told that this was “bonus income.” After one week, the students were contacted again and asked how much of the $50 they had spent and how much they had saved. Those in the “rebate” group had spent an average of $22.04, while those in the “bonus” group had spent an average of $9.55. Is this difference statistically significant?

g) Statistical evidence was used in the murder trial of Kristen Gilbert, a nurse who was accused of killing patients. More than one thousand 8-hour hospital shifts were analyzed. Was the proportion of shifts with a death significantly higher for the shifts that Gilbert worked?

Much of statistics involves making inferences from a sample to a population.

- **Population**: the *entire* group of observational units (people or objects) about which information is desired
- **Sample**: a (typically small) *part* of the population from which data are gathered
  - **Sample size**: number of observational units in the sample