Section 1.1
Types of Data

Statistics is the science of the collection, organization and interpretation of data. When data is analyzed, it can help people make decisions when faced with uncertainty. We will need to know what questions need to be answered, how much data will need to be collected and how we should go about collecting it, how can we summarize the data and what decisions can be made based on our question and the data collected.

Types of Data

Population data is everything (or everyone) we are studying. It is a set of data that consists of all possible values pertaining to a certain set of observations or an investigation.

Sample data represents a subset of the population. It is just a small section of the population taken for the purpose of investigation

Example 1: Identify the population and the sample for each of the following:
a. A university in interested in how many of their students commute to school. They randomly choose 100 students at the student center to interview.

Population data: All students

Sample data: 100 students

b. An elementary school is creating a new lunch menu. They send questionnaires to students with last names that begin with the letters A through L.

Population data: All students

Sample data: Students with last names beginning with A – L

A **variable** is a characteristic of an individual that can assume more than one value. Variables can be classified as **categorical** (qualitative) or **quantitative** (numeric).

Categorical (or qualitative) variables describe qualities or attributes of an individual.
For example: blood type, brand of potato chips

Quantitative variables include measurements or observations that take on numeric values.
For example: number of dvds a person owns, height of men over 18

Quantitative variables can be classified as **discrete** or **continuous**.

A **discrete quantitative** variable is a countable set of values.
For example: number of dvds a person owns

Continuous quantitative variables are data that can take on any values within some interval.
For example: height of men over 18

When trying to distinguish between categorical and quantitative data, it is often helpful to think if it makes sense to average the values. However, beware of categorical data that is represented numerically. For example, "on a scale of 1 to 10, rate your interest in this class where 1=poor and 10=excellent". When this is true, the term **ordinal data** is sometimes used to describe that data.

Example 2: Classify the following variables as categorical or quantitative. If quantitative, state whether the variable is discrete or continuous.

a. Number of siblings.
Categorical or Quantitative
If quantitative, discrete or continuous.

b. Time it takes to be on hold when calling the IRS at tax time.
Categorical or Quantitative
If quantitative, discrete or continuous.

c. The number of quizzes a student has missed.
Categorical or Quantitative
If quantitative, discrete or continuous.

d. Blood type.
Categorical or Quantitative
If quantitative, discrete or continuous.