

## Terminology

For us to have an understanding of what the subject of **statistics** is all about, we need to introduce some terminology. First we will explain what we mean by the subject of statistics.

**Definition 1: Statistics:** Statistics is the science of collecting, organizing, summarizing, analyzing, and making inferences from data. The subject of statistics is divided into two broad areas that incorporate collecting, organizing, summarizing, analyzing, and making inferences from data. These categories are *descriptive statistics* and *inferential statistics*.

In order to obtain information, **data** are collected from variables used to describe an event. Data are the values or measurements that variables describing an event can assume.

**Definition 2: Variables:** Variables whose values are determined by chance are called *random variables*. There are two types of variables:

- 1) Qualitative variables
- 2) Quantitative variables.

Qualitative variables are nonnumeric in nature.

Quantitative variables can assume numeric values and can be classified into two groups: **discrete variables** and **continuous variables**. A collection of values is called a *data set*, and each value is called a *data value*.

**Definition 3: Quantitative data:** Quantitative data are data values that are numeric.

For example, the heights of female basketball players are quantitative data values.

**Definition 4: Qualitative data:** Qualitative data are data values that can be placed into distinct categories, according to some characteristic or attribute. For example, the eye color of female basketball players is classified as qualitative data.

**Definition 5: Discrete variables:** Discrete variables are variables that assume values that can be counted. For example, the number of days it rained in your neighborhood for the month of March.

**Definition 6: Continuous variables:** Continuous variables are variables that can assume all values between any two given values.

**Definition 7: Population:** A population consists of all elements that are being studied.

For example, we might be interested in studying the distribution of math scores of freshmen at a college campus. In this case, the population will be the math scores of all the freshmen on that particular campus.

**Definition 8: Sample:** A sample is a subset of the population.

For example, we might be interested in studying the distribution of math scores of freshmen at a college campus. In this case, we might select the math scores of every 10<sup>th</sup> freshman from an alphabetical list of the students' last names.

**Definition 9: Census:** A census is a sample of the entire population.

For example, we might be interested in studying the distribution of math scores of freshmen at a college campus. In this case, we may list the math scores for all freshmen on that particular campus.

**Definition 10: Parameter:** A parameter is a characteristic of or a fact about a population.

For example, we might be interested in studying the distribution of math scores of freshmen at a college campus. In this case, the average math score for all freshmen on this particular campus may be 25.

Since parameters are descriptions of the population, a population can have many parameters. When selecting a sample, statisticians would like to select values in such a way that there is no inherent bias. One way of doing this is by selecting a **random sample**.

**Definition 11: Random sample:** A random sample of a particular size is a sample selected in such a way that each group of the same size has an equal chance of being selected.

For example, in a lottery game in which six numbers are selected, this will be a random sample of size six, since each group of size six will have an equal chance of being selected.