## Introduction

*Statistics* is the art of learning from data.
It is concerned with the collection of data, its subsequent description, and its analysis, which often leads to the drawing of conclusions.

## Data Collection and Descriptive Statistics

Sometimes a statistical analysis begins with a given set of data.
Statistics can be used to describe, summarize, and analyze these data.

In other situations, data are not yet available; in such cases statistical theory can be used to design an appropriate experiment to generate data.
The experiment chosen should depend on the use that one wants to make of the data.

At the end of the experiment, the data should be described. In addition, summary measures such as the averge should be presented.
This part of statistics, concerned with the description and summarization of data, is called *descriptive statistics*.

## Inferential Statistics and Probability Models

After the preceding experiment is completed and the data are described and summarized, we hope to be able to draw a conclusion about which teaching method is superior.
This part of statistics, concerned with the drawing of conclusions, is called *inferential statistics*.

To be able to draw logical conclusions from data, we usually make some assumptions about the chances (or *probabilities*) of obtaining the different data values. The totality of these assumptions is referred to as a *probability model* for the data.

## Populations and Samples

In statistics, we are interested in obtaining information about a total collection of elements, which we will refer to as the *population*.
The population is often too large for us to examine each of its members.

In such cases, we try to learn about the population by choosing and then examining a subgroup of its elements. This subgroup of a population is called a *sample*.

If the sampel is to be informative about the total population, it must be, in some sense, representative of that population.

In certain situations, we are presented with a sample and must then decide whether this sample is reasonably representative of the entire population. In practice, a given sample generally cannot be assumed to be representative of a population unless that sample has been chosen in a random manner. This is because any specific nonrandom rule for selecting a sample often results in one that is inherently biased toward some data values as opposed to others.

Thus, although it may seem paradoxical, we are most likely to obtain a representative sample by choosing its members in a totally random fashion without any prior considerations of the elements that will be chosen.

## A Brief History of Statistics