**Population**: The collection of all individuals or items under consideration in a statistical study.

**Sample**: That part of the population from which information is obtained

**Inferential statistics** consists of methods for drawing and measuring the reliability of conclusions about a population based on information obtained from a sample of the population.

Descriptive statistics consists of methods for organizing and summarizing information.

Descriptive statistics includes the construction of graphs, charts, and tables and the calculation of various descriptive measures such as averages, measures of variation, and percentiles.

**Systematic Random Sampling**

Step 1 Divide the population size by the sample size and round the result down to the nearest whole number, m.

Step 2 Use a random-number table or a similar device to obtain a number, k, between 1 and m.

Step 3 Select for the sample those members of the population that are numbered k, k + m, k + 2m,....

**Cluster Sampling**

Step 1 Divide the population into groups (clusters).

Step 2 Obtain a simple random sample of the clusters.

Step 3 Use all the members of the clusters obtained in Step 2 as the sample.

**cluster sampling,** which is particularly useful when the members of the population are widely scattered geographically.

**In stratified sampling** the population is first divided into subpopulations, called strata, and then sampling is done from each stratum.

Step 1 Divide the population into subpopulations (strata).

Step 2 From each stratum, obtain a simple random sample of size proportional to the size of the stratum; that is, the sample size for a stratum equals the total sample size times the stratum size divided by the population size.

Step 3 Use all the members obtained in Step 2 as the sample.

**A frequency distribution** of qualitative data is a listing of the distinct values and their frequencies.

**Relative frequency**, which is the ratio of the frequency to the total number of observations:

Relative frequency = Frequency/Number of observations

**A pie chart** is a disk divided into wedge-shaped pieces proportional to the relative frequencies of the qualitative data.

**Standard deviation** measures variation by indicating how far, on average, the observations are from the mean. For a data set with a large amount of variation, the observations will, on average, be far from the mean; so the standard deviation will be large. For a data set with a small amount of variation, the observations will, on average, be close to the mean; so the standard deviation will be small.

**quartiles** divide a data set into quarters (four equal parts).

The first quartile is the median of the part of the entire data set that lies at or below the median of the entire data set.

The second quartile is the median of the entire data set.

The third quartile is the median of the part of the entire data set that lies at or above the median of the entire data set

**The interquartile range**, or IQR, is the difference between the first and third quartiles; that is,

IQR = Q3 − Q1.

**The five-number summary** of a data set is Min, Q1, Q2, Q3, Max