

12.1 – Sampling, Frequency Distributions, and Graphs

Statistics is a branch of mathematics dealing with the collection, analysis, interpretation, and presentation of masses of numerical data. (www.m-w.com)

Statistics is divided into two main branches.

- **Descriptive statistics** is concerned with the collection, organization, and analysis of data.
- **Inferential statistics** is concerned with making generalizations or predictions from the data collected.

A statistician's interest lies in drawing conclusions about possible outcomes through observations of only a few particular events.

- The **population** consists of **all** items or people of interest.
- The **sample** includes *some* of the items in the population.

When a statistician draws a conclusion from a sample, there is always the possibility that the conclusion is incorrect. A sample should be representative of the population.

A **random sample** is a sample drawn in such a way that each item in a population has an equal chance of being drawn.

Example: (p.683: #2)

A **Frequency Distribution** or **Frequency Table** is a two column table that lists the data values or range of data values (called **classes**) on one column and the frequency in each on the other column.

Example: (p.683: #8)

The **Lower Class Limit** is the left most digit in a class of a frequency distribution. The **Upper Class Limit** is the right most digit in a class of a frequency distribution.

The **Class Width** is the difference between an Upper Class Limit and a Lower Class Limit. The class width is the same for each class.

A **Histogram** is a graph that uses bars that touch to illustrate the frequency or relative frequency of each class. Typically the horizontal axis is marked off by the data values or classes and the vertical axis has the frequency or relative frequency.

Example:

A **Frequency Polygon** is a line graph that connects line segments through the midpoints of the classes. Typically the horizontal axis has the midpoints marked off and the vertical axis has the frequency or relative frequency.

Example:

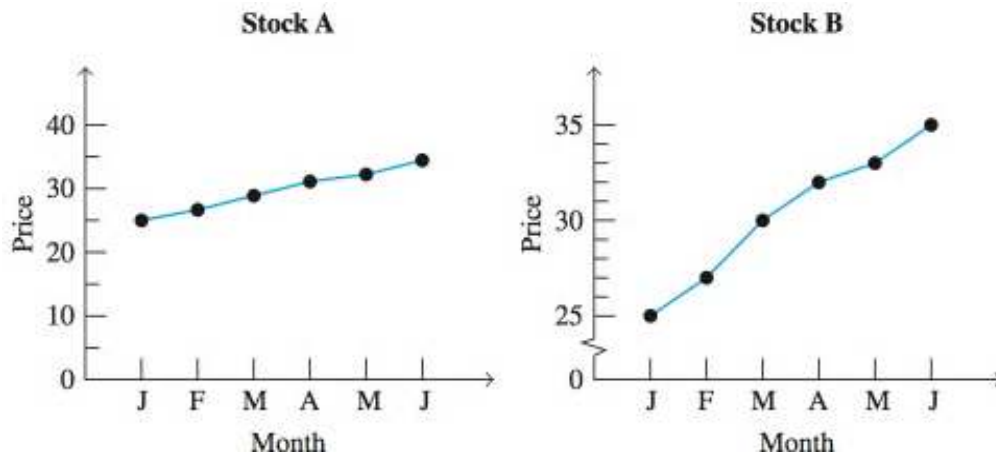
A **Stef-and-Leaf Plot** is a graph that splits the data values into **stems** (left most digits) and **leaves** (right most digit). The stems are shown in numerical order in a column. Each leaf for a data value is placed in the row corresponding to the appropriate stem. Typically the leaves will be rewritten in numerical value.

Example:

It is not uncommon to see statistics misused whether accidentally or on purpose.

Graphs can be misleading simply by altering the vertical axis.

While each graph presents identical information, the vertical scales have been altered.



Things to watch for in visual displays of data p.682 and Table 12.5, p. 682.