Chapter 3: Plotting Distributions

Dotplots

Stem-and-Leaf displays

Histograms

- frequency histograms
- relative frequency histograms
- density histograms

Types of Data (Review!)

Continuous: Can potentially take any value on an interval of numbers.

Categorical: indicates the group or category to which an item belongs.

Longitudinal data study the results from a process over time.

Data collected from a group of people or things at one point in time are called **cross-sectional** data.

When the time dimension of longitudinal data is ignored, we will also say that we are looking at cross-sectional aspects of the data.

Dotplots (page 52)

A **dotplot** consists of a horizontal scale (a number line) on which dots are placed to show the numerical values of the data.

If a data value repeats, the dots are piled up at that location, one dot for each repetition.

We say that the dotplot displays the **distribution** of the data.

Example:

Stem-and-Leaf Displays

(page 56)

The **stem-and-leaf display** (or stemplot) is a combination table and graphic display.

It retains the numerical values of the individual data points while at the same time producing a profile that graphically displays the frequencies of the grouped data values.

We let the digits in the data values do the work of grouping the data while also displaying the frequencies.

Example:

Histograms (page 62)

frequency histograms

relative frequency histograms

density histograms

All require:

- grouping the data into classes, the class intervals
- counting the number of observations in each class, the class frequencies
- displaying in a plot, the histogram

Class intervals (with midpoints and end points) cover the full range of the observations.

Frequency histograms plot the frequency of data above the corresponding class interval.

Relative frequency histograms plot the relative frequency of data above the corresponding class interval.

Density histograms plot the density of data above the corresponding class interval where density is given as

$$\text{Density} = \frac{\text{Relative Frequency}}{\text{Length of Class Interval}}$$

Key: Area of density histogram ‘bar’ should be proportional to frequency of corresponding class interval.