2.2: Organizing Quantitative Data

Recall: Types of Data

Nominal (categorical): Values are names or categories. Values do not measure the amount of any characteristic, and do not have a meaningful order.

Ordinal: Values have a meaningful order.

Interval: Values have a meaningful order, and also a meaningful difference (interval) between values. Values do not have a meaningful zero. (unless it is ratio also)

Ratio: Values have a meaningful order, a meaningful difference (interval) between values, and also have a meaningful zero. The zero value indicates the absence of the characteristic being measured.

Quantitative data can be classified as *discrete* or *continuous*.

Discrete: A *discrete* variable takes on either a finite number of values, or a countably infinite number of values. *Countable* means the values can be counted (1, 2, 3, …); countably infinite means that you can't ever finish counting the values. A discrete variable cannot take on every value in an interval.

Continuous: A *continuous* variable takes on every value in an interval. It does not make sense to count (assign the numbers 1, 2, 3, ….) to the values.

Examples of discrete variables:

integers

# of students

# of participants

power value in contact lens prescription

Examples of continuous variables:

finishing times for runners in a race

(19:57.3, 39:54.9, ...)

real numbers

height, weight, distance, time

radius of eyeball

Histograms:

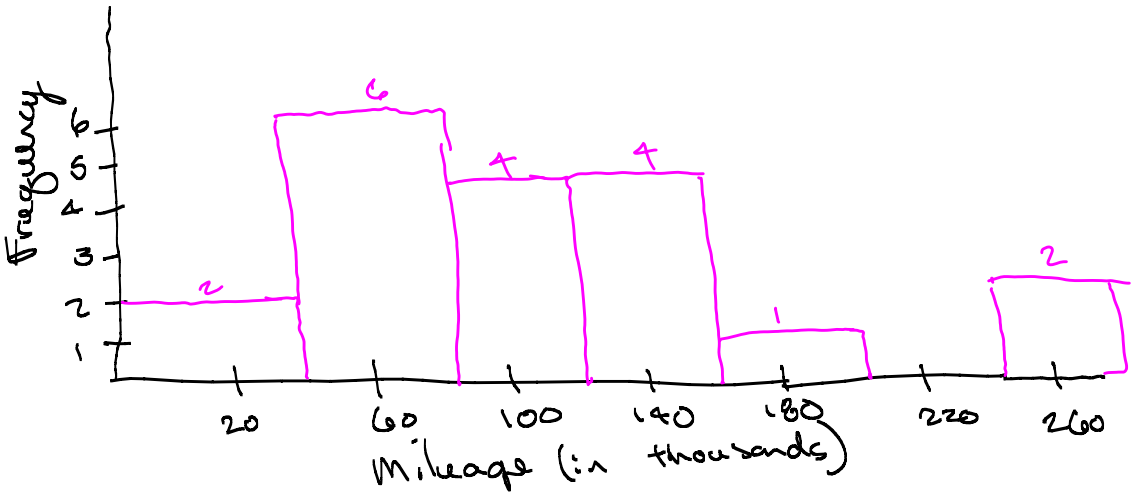
A histogram is a specific type of bar graph used to represent quantitative data. The values for the variable are on the horizontal axis, and the frequencies or relative frequencies are on the vertical axis. All values within the range of the variable are represented on the *x*-axis.

Example 1:

250 000
160 000
80
108
109
70
50
120
70

17
31
130
130
50
50
44
114
130
260

Mileage Category	Frequency
0 – 39 999	11
40 000 – 79 999	6
80 000 – 119 999	4
120 000 – 159 999	4
160 000 – 199 999	1
200 000 – 239 999	0
240 000 – 279 999	2



Stem-and-leaf plots:

A *stem-and-leaf plot* provides a similar summary display of relative frequencies as a bar chart or histogram, but preserves the individual data points. In a stem-and-leaf plot, each value is divided into a "stem" and a "leaf." For example, the number 29 could be represented as a stem of 20 and a leaf of 9. The number 15.7 could be represented as a leaf of 15 and a stem of 0.7.

Example 2:

40	20	10
25	20	36
30	30	47
50	15	18
10	37	15
10	7	44

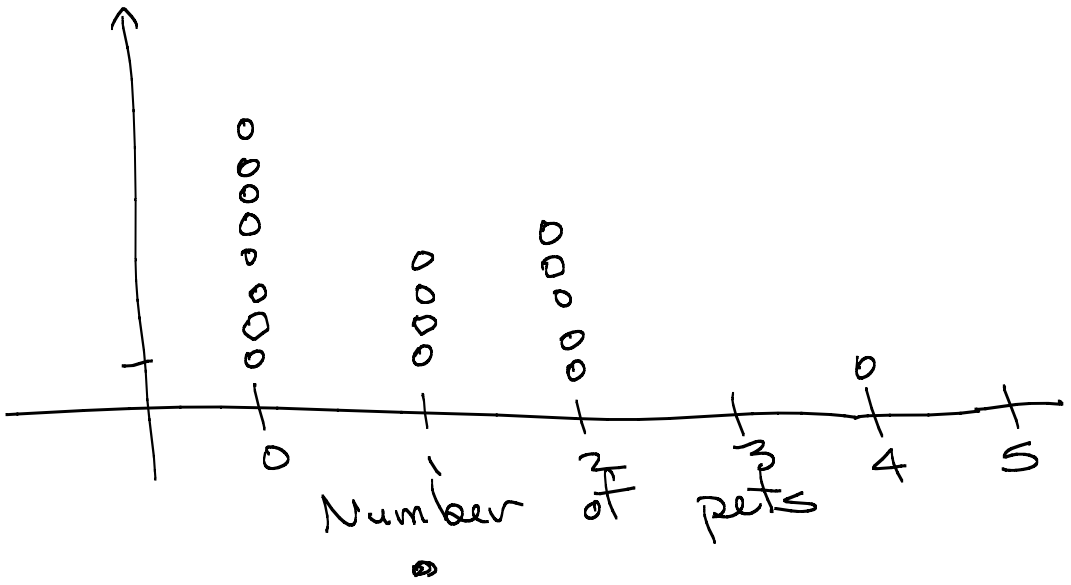
The 10s digit will be the "stem" and the 1s digit will be the "leaf"

0	7
1	0 0 5 0 8 5
2	5 0 0
3	0 0 7 6
4	0 7 4
5	0
6	

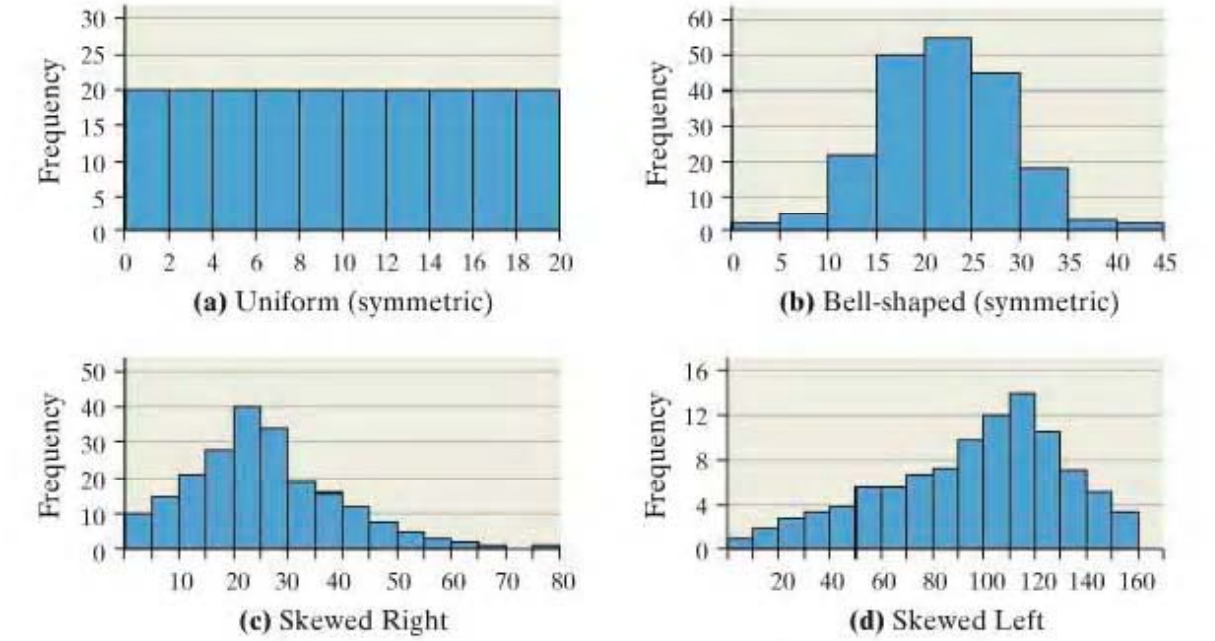
Dot plots:

A dot plot is a simple graph in which each observation is represented by a dot. The dots are place in a column above the value they represent. A dot plot is used for discrete variables (not for continuous variables).

Example 1:



Identifying the shape of a frequency distribution:



Time-series plots:

A *time-series plot* is used to analyze trends in data over time. The horizontal axis represents time; the vertical axis represents the value of the variable.

Example 2: