pie charts and bar graphs

pie charts show the distribution of a categorical variable as a “pie” whose slices are sized by the frequencies or relative frequencies of the categories.
bar graphs represent each category as a bar. The bar heights show the frequencies or relative frequencies of the categories.

relative frequency = frequency/sum of all frequencies

example 2.1, 2.3, 2.6.

Example: credit card used by last 50 customers in a store

Table 2.3 **Relative Frequency Distribution for Credit Cards**

Credit Card	Frequency	Relative Frequency
MasterCard	11	0.22
Visa	23	0.46
Am. Express	9	0.18
Discover	7	0.14

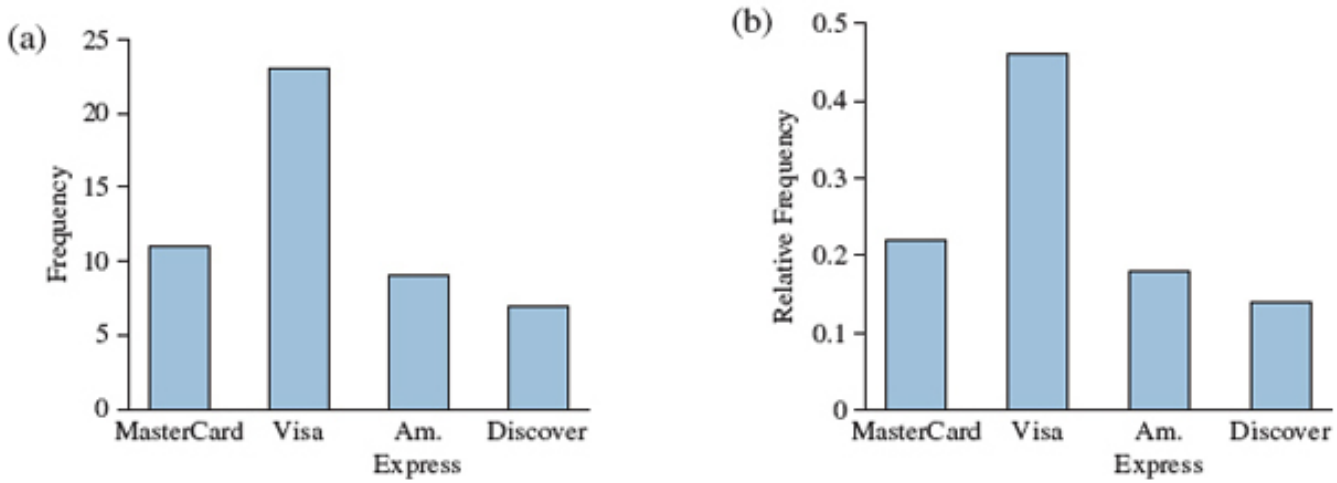


Figure 2.1 (a) Frequency bar graph. (b) Relative frequency bar graph.

Pie chart for the credit card data

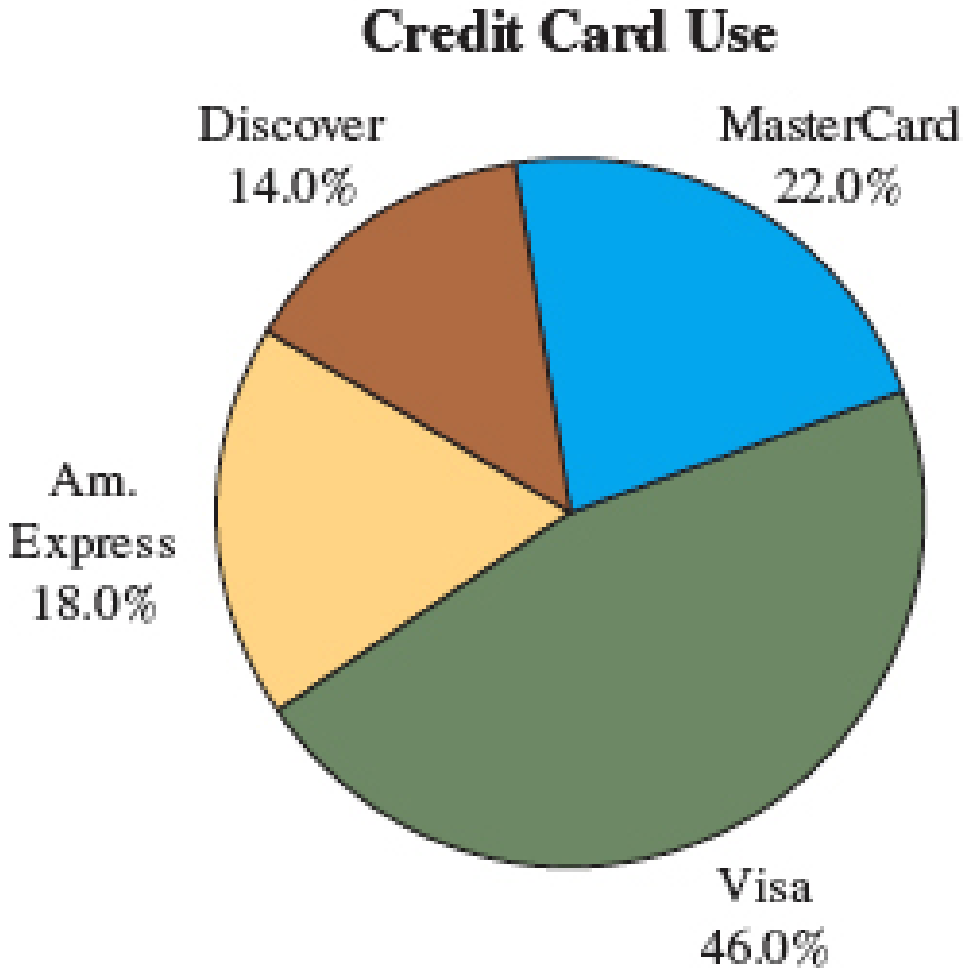


Figure 2.4 Pie chart for the credit card data in **Table 2.5**

histograms

For quantitative data:
Make a histogram:
1. choose the classes.
2. count the individuals in each class.
3. draw the histogram based on the frequency table.

example 2.7, 2.10, 2.12.

Example

Table 2.7 Particulate Emissions for 65 Vehicles

1.50	0.87	1.12	1.25	3.46	1.11	1.12	0.88	1.29	0.94	0.64	1.31	2.49
1.48	1.06	1.11	2.15	0.86	1.81	1.47	1.24	1.63	2.14	6.64	4.04	2.48
1.40	1.37	1.81	1.14	1.63	3.67	0.55	2.67	2.63	3.03	1.23	1.04	1.63
3.12	2.37	2.12	2.68	1.17	3.34	3.79	1.28	2.10	6.55	1.18	3.06	0.48
0.25	0.53	3.36	3.47	2.74	1.88	5.94	4.24	3.52	3.59	3.10	3.33	4.58

Table 2.9 Relative Frequency Distribution for Particulate Data

Class	Frequency	Relative Frequency
0.00–0.99	9	0.138
1.00–1.99	26	0.400
2.00–2.99	11	0.169
3.00–3.99	13	0.200
4.00–4.99	3	0.046
5.00–5.99	1	0.015
6.00–6.99	2	0.031

Histogram

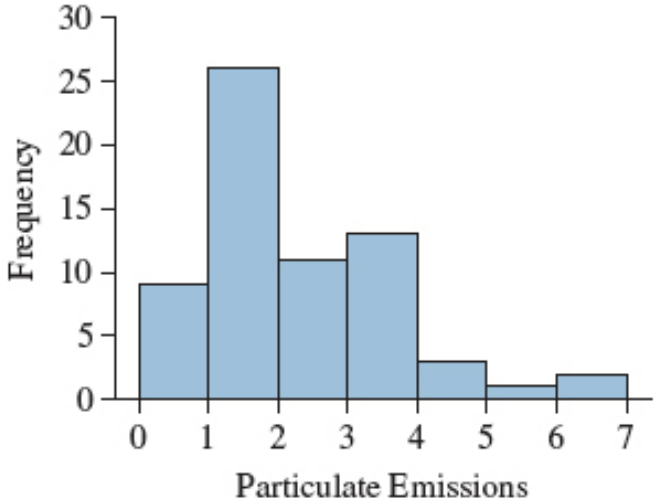


Figure 2.5 Frequency histogram for the frequency distribution in Table 2.10

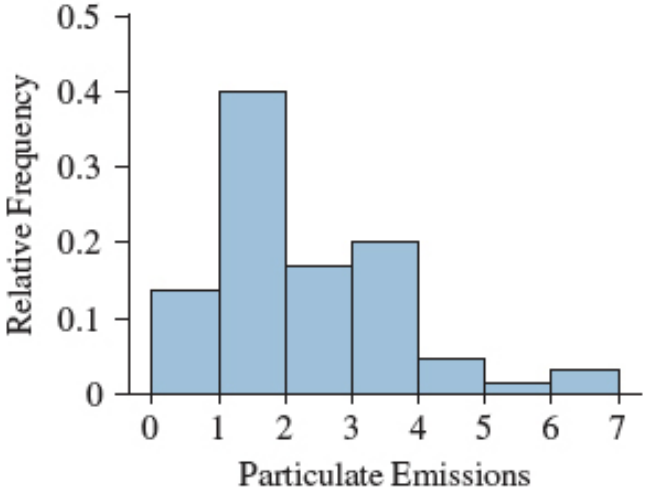


Figure 2.6 Relative frequency histogram for the relative frequency

Shapes of histogram

A distribution is **symmetric** if the right and left sides of the histogram are approximately mirror images of each other.

It is **skewed to the right** if the right side of the histogram extends much farther out than the left side.

It is **skewed to the left** if the left side extends much farther out.

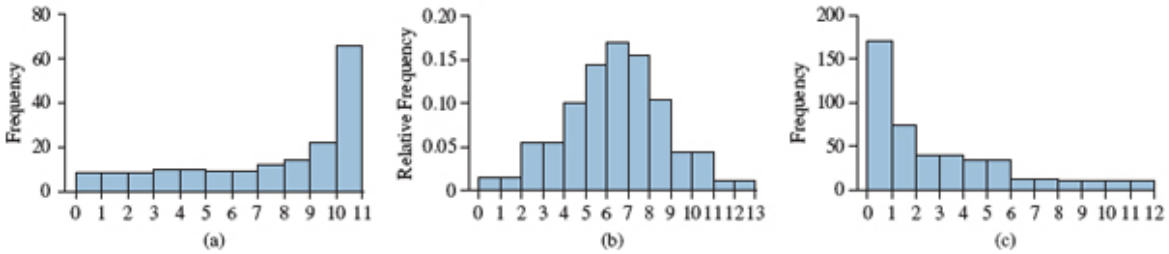
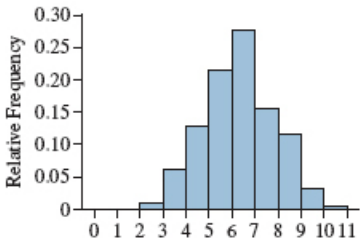


Figure 2.10 (a) A histogram skewed to the left. (b) An approximately symmetric histogram. (c) A histogram skewed to the right.

Shapes of histograms

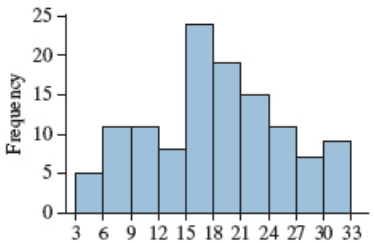
A peak of a histogram is called a **mode**.
Unimodal: has only one mode.
Bimodal: has two clearly distinct modes.
a and b are unimodal and c is bimodal below.

a.



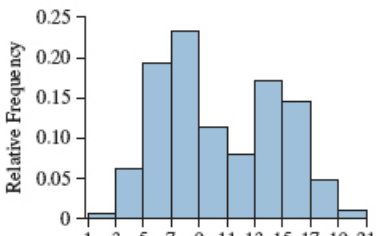
Answer

b.



Answer

c.



stem-and-leaf plot

For quantitative data.
Use a stem and a leaf to display an observation. The last digit is on the leaf and the remaining digits are on the stem.
Test scores: 12, 13, 21, 22, 22, 28, 29, 30, 32, 33, 33, 37, 42, 45, 49, 62

stem leaf
1 | 2 3
2 | 1 2 2 8 9
3 | 0 2 3 3 7
4 | 2 5 9
5 |
6 | 2

1 | 2 means 12.

Back-to-back stem-and-leaf plot

Back to back stem and leaf plot is useful for comparing two data sets.

High Altitude		Sea Level
	0	355669999
9851	1	011111122223333445556668 89
974110	2	11124556777
8863	3	0111334555678
744	4	026
9866633	5	9
5533	6	66
63	7	
97	8	
521	9	

Figure 2.16 Back-to-back stem-and-leaf plots comparing the emissions in vehicles driven at high altitude with emissions from vehicles driven at sea level

Dot plot

For quantitative data.
Useful when data set is not large and has some repeated values.

Table 2.18 Numbers of Children of U.S. Presidents and Their Wives

0	2	10	2	5	3	6	2	2	4	1
5	4	15	3	4	5	3	2	3	4	2
6	0	0	0	8	3	3	6	2	4	2
0	4	6	4	7	2	0	1	2	6	

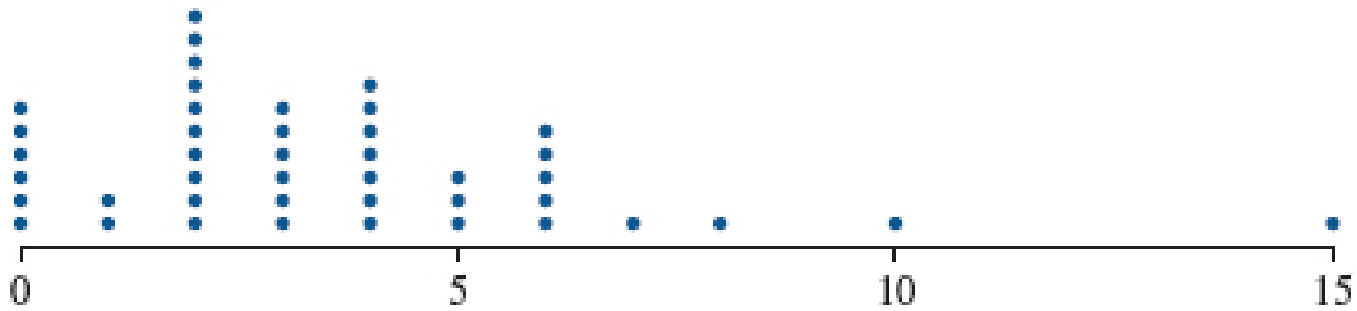


Figure 2.17 Dotplot for the data in Table 2.18

time-series plot

A **time-series plot** of a variable plots each observation against the time. Put time on the horizontal scale and the variable on the vertical scale. Connecting the data points by lines helps emphasize any change over time.

2. The National Institute on Drug Abuse surveyed U.S. high school seniors every two years to determine the percentage who said they had used marijuana one or more times. The following time-series plot presents the results.