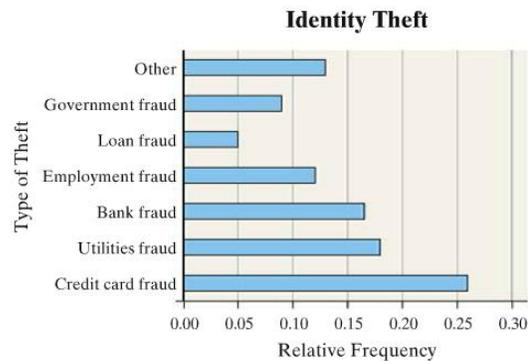


Chapter 2

Ch2.1 Organizing Qualitative Data

Example 1 :

Identity Theft Identity fraud occurs someone else's personal information is used to open credit card accounts, apply for a job, receive benefits, and so on. The following relative frequency bar graph represents the various types of identity theft based on a study conducted by the Federal of Trade Commission.



Source: Federal Trade Commission

- (a) Approximate what percentage of identity theft was loan fraud (such as applying for a loan in someone else's name)?
- (b) If there were 10 million cases of identity fraud in 2008, how many were credit card fraud (someone uses someone else's credit card to make a purchase)?

B1. Frequency / Relative Frequency Distribution

- A _____ lists each category of data and the frequency which is the number of occurrences for each category data.
- A _____ lists each category of data and the relative frequency which is the proportion of observation within a category.

Relative frequency =

Example 1: In a national survey conducted by the Centers for Disease Control to determine health-risk behaviors among college students, college students were asked, “How often do you wear a seat belt when driving a car?” The frequencies were as follows:

Response	Frequency
I do not drive a car	249
Never	118
Rarely	249
Sometimes	345
Most of the time	716
Always	3093

(a) Construct a relative frequency distribution.

Response	Frequency	Relative Frequency
I do not drive a car	249	
Never	118	
Rarely	249	
Sometimes	345	
Most of the time	716	
Always	3093	

(b) What percentage of respondents answered “Always”?

(c) What percentage of respondents answered “Never” or “Rarely”?

(d) Suppose that a representative from the Centers for Disease Control says, “2.5% of the college students in this survey responded that they never wear a seat belt.” Is this a descriptive or inferential statement?

(d) Construct a relative frequency bar graph.

(e) Construct a Pareto chart.

(f) Construct a pie chart.

affiliation	relative freq	degree of pie

Chapter 2.2 Organizing Quantitative Data: The Popular Displays

Objective A : Histogram

- A _____ is constructed by drawing rectangles for each class of data. If the discrete data set is small, each number is a class. If the discrete data set is large or the data are continuous, the classes must be created using interval of numbers. The _____ of each rectangle is the frequency or relative frequency of the class. The _____ of each rectangle is the same and the rectangles _____ each other.

Construct Frequency Distribution and Histogram for Discrete Data

Example 1: The following data represent the number of customers waiting for a table at 6:00 P.M. for 40 consecutive Saturdays at Bobak's Restaurant:

11 5 11 3 6 8 6 7 4 5 13 9 6 4 14 11 13 10 9 6 8
10 9 5 10 8 7 3 8 8 7 8 7 9 10 4 8 6 11 8

(a) Are these data discrete or continuous? Explain.

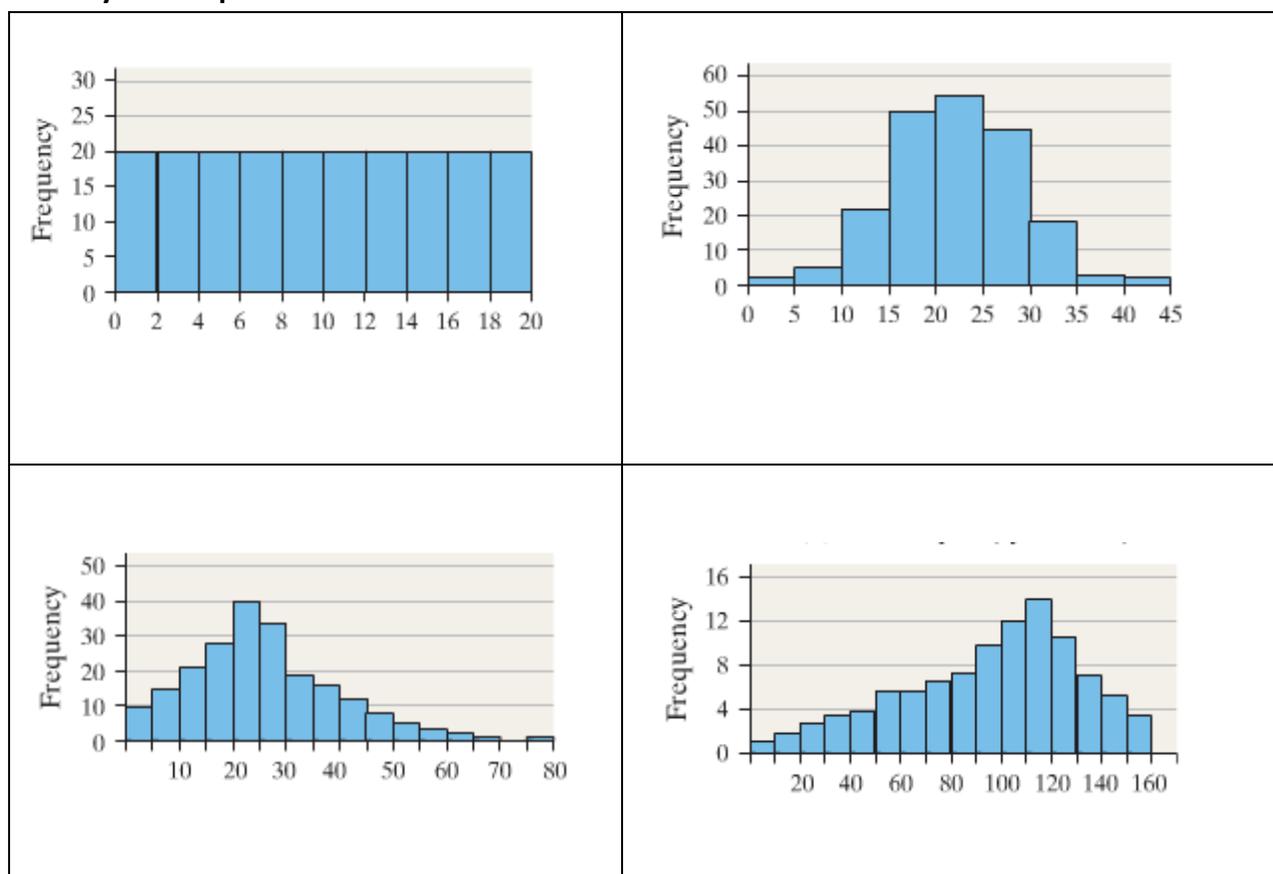
(b) Construct a frequency distribution of the data.

(c) Construct a relative frequency distribution of the data.

(d) What percentage of the Saturdays had 10 or more customers waiting for a table at 6:00 p.m.?

(e) Construct a frequency histogram of the data.

Identify the shape of each distribution.



Objective B : Constructing a Stem-and-Leaf Plot

The _____ of a data value will consist of the digits to the left of the rightmost digit.

The _____ of a data value will be the rightmost digit.

Example 1: The following data represent the number of miles per gallon achieved on the highway for small cars for the model year 2008.

27 31 28 30 52 25 33 33 29 23 27 37 30 45 24 32
34 35 31 44 42 26 43 35 36 36 54 33 32 35 34 37

(a) Construct a stem-and-leaf plot.

(b) Describe the shape of the distribution.

Objective C : Construct Frequency Distributions and Histogram for Continuous Data

- _____ are categories into which data are grouped.
- The _____ class limit is the smallest value within a class.
- The _____ class limit is the largest value within a class.
- The _____ is the difference between consecutive lower class limits.

- The class width is computed by the following formula.

$$\text{Class width} \approx \frac{\text{largest data value} - \text{smallest data value}}{\text{number of classes}}$$

-----> Round this value _____ to the same decimal place as the raw data.

Example 1: The following data represent the fall 2006 student headcount enrollments for all public community colleges in the state of Illinois.

Number of Students Enrolled	Number of Community Colleges ^a
0–4,999	15
5,000–9,999	16
10,000–14,999	9
15,000–19,999	4
20,000–24,999	0
25,000–29,999	1

Source: Illinois Board of Higher Education, Report to the Governor and General Assembly, December 2007

^aTreats the City Colleges of Chicago as seven distinct institutions, but all other multicampus districts as single institutions.

- Find the number of classes.
- Find the class limits.
- Find the class width.

Example 2 : Uninsured Rates The following data represent the percentage of people without health insurance for the 50 states and the District of Columbia in 2009. (Ch 2.2 Q36 p. 94)

4.2	10.6	12.6	14.8	17.8	18.9	21.4
8.6	10.6	13.0	15.5	18.1	19.4	22.2
9.2	10.9	13.3	15.9	18.3	19.6	25.0
9.6	10.9	13.4	15.9	18.3	19.7	
9.6	11.3	13.9	16.1	18.4	20.6	
9.7	11.4	14.0	16.1	18.4	21.1	
10.2	11.6	14.3	16.1	18.6	21.2	
10.5	12.3	14.7	16.2	18.7	21.3	

Source: Gallup

With the first class having a lower class limit of 4 and a class width of 2:

(a) Construct a frequency distribution.

Put the data in ascending order first:

4.2 8.6 9.2 9.6 9.6 9.7 10.2 10.5 10.6 10.6 10.9 10.9 11.3 11.4 11.6
12.3 12.6 13.0 13.3 13.4 13.9 14.0 14.3 14.7 14.8 15.5 15.9 15.9 16.1 16.1
16.1 16.2 17.8 18.1 18.3 18.3 18.4 18.4 18.6 18.7 18.9 19.4 19.6 19.7 20.6
21.1 21.2 21.3 21.4 22.2 25.0

(b) Construct a relative frequency distribution.

(c) Construct a frequency histogram of the data.

(d) Construct a relative frequency histogram of the data.

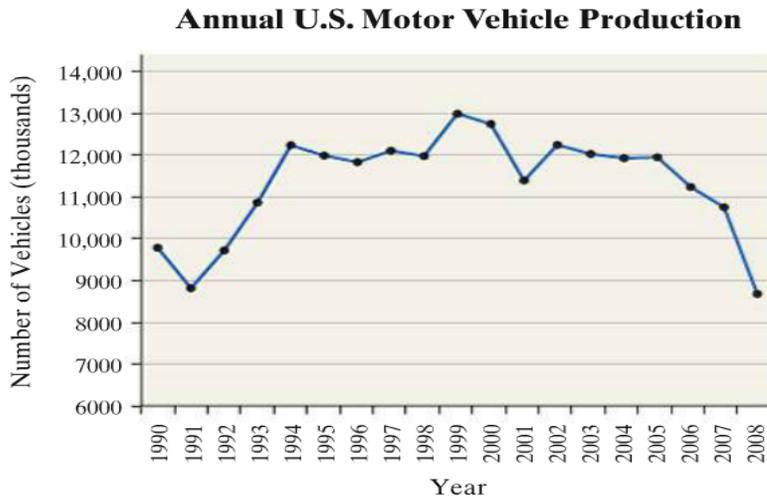
(e) Describe the shape of the distribution.

Example : The largest value of a data set is 125 and the smallest value of the data set is 27.
If six classes are to be formed, calculate an appropriate class width.

Objective D : Time Series Graphs

- A _____ represents the values of a variable that have been collected over a specified period of time. The horizontal axis is the _____ and the vertical axis is the value of the _____. Line segments are drawn by connective consecutive points of time and corresponding value of the variable.

Example 1: The following time-series graph shows the annual U.S. motor vehicle production from 1990 through 2008.



Source: Bureau of Transportation Statistics

- Estimate the number of motor vehicles produced in the United States in 1991.
- Estimate the number of motor vehicles produced in the United States in 1999.
- Use the results from (a) and (b) to estimate the percent increase in the number of motor vehicles produced from 1991 to 1999.
- Estimate the percent decrease in the number of motor vehicles produced from 1999 to 2008.