Chapter 2

Frequency Distributions and Graphs

Chapter 2 Overview

Introduction

- 2-1 Organizing Data
- 2-2 Histograms, Frequency Polygons, and Ogives
- 2-3 Other Types of Graphs

Chapter 2 Objectives

1. Organize data using frequency distributions.
2. Represent data in frequency distributions graphically using histograms, frequency polygons, and ogives.
3. Represent data using Pareto charts, time series graphs, and pie graphs.
4. Draw and interpret a stem and leaf plot.

2-1 Organizing Data

- Data collected in original form is called **raw data**.
- A **frequency distribution** is the organization of raw data in table form, using classes and frequencies.
- Nominal- or ordinal-level data that can be placed in categories is organized in **categorical frequency distributions**.

Chapter 2
Frequency Distributions and Graphs

Section 2-1

Example 2-1

Page #38

Categorical Frequency Distribution

Twenty-five army indicates were given a blood test to determine their blood type.

Raw Data: A,B,B,AB,O O,O,B,AB,B
B,B,O,A,O A,O,O,O,AB AB,A,O,B,A

Construct a frequency distribution for the data.

Categorical Frequency Distribution

Twenty-five army inductees were given a blood test to determine their blood type.

Raw Data: A,B,B,AB,O O,O,B,AB,B
B,B,O,A,O A,O,O,O,AB AB,A,O,B,A

Class	Tally	Frequency	Percent
A	𝍸	5	20%
B	𝍸 II	7	28%
O	𝍸 IIII	9	36%
AB	IIII	4	16%

Grouped Frequency Distribution

- **Grouped frequency distributions** are used when the range of the data is large.
- The smallest and largest possible data values in a class are the ***lower*** and ***upper class limits***. ***Class boundaries*** separate the classes.
- To find a class boundary, average the upper class limit of one class and the lower class limit of the next class.

Grouped Frequency Distribution

- The **class width** can be calculated by subtracting
  - successive lower class limits (or boundaries)
  - successive upper class limits (or boundaries)
  - upper and lower class boundaries
- The ***class midpoint*** X_m can be calculated by averaging
  - upper and lower class limits (or boundaries)

Rules for Classes in Grouped Frequency Distributions

1. There should be 5-20 classes.
2. The class width should be an odd number.
3. The classes must be mutually exclusive.
4. The classes must be continuous.
5. The classes must be exhaustive.
6. The classes must be equal in width (except in open-ended distributions).

Chapter 2
Frequency Distributions and Graphs

Section 2-1

Example 2-2

Page #41

Constructing a Grouped Frequency Distribution

The following data represent the record high temperatures for each of the 50 states. Construct a grouped frequency distribution for the data using 7 classes.

112	100	127	120	134	118	105	110	109	112
110	118	117	116	118	122	114	114	105	109
107	112	114	115	118	117	118	122	106	110
116	108	110	121	113	120	119	111	104	111
120	113	120	117	105	110	118	112	114	114

Constructing a Grouped Frequency Distribution

STEP 1 Determine the classes.

Find the class width by dividing the range by the number of classes 7.

$$\textit{Range} = \textit{High} - \textit{Low}$$

$$= 134 - 100 = 34$$

$$\textit{Width} = \textit{Range}/7 = 34/7 = 5$$

Rounding Rule: Always round up if a remainder.

Constructing a Grouped Frequency Distribution

- For convenience sake, we will choose the lowest data value, 100, for the first lower class limit.
- The subsequent lower class limits are found by adding the width to the previous lower class limits.

Class Limits
100 - 104
105 - 109
110 - 114
115 - 119
120 - 124
125 - 129
130 - 134

- The first upper class limit is one less than the next lower class limit.
- The subsequent upper class limits are found by adding the width to the previous upper class limits.

Constructing a Grouped Frequency Distribution

- The class boundary is midway between an upper class limit and a subsequent lower class limit. 104,104.5,105

Class Limits	Class Boundaries	Frequency	Cumulative Frequency
100 - 104	99.5 - 104.5		
105 - 109	104.5 - 109.5		
110 - 114	109.5 - 114.5		
115 - 119	114.5 - 119.5		
120 - 124	119.5 - 124.5		
125 - 129	124.5 - 129.5		
130 - 134	129.5 - 134.5		

Constructing a Grouped Frequency Distribution

STEP 2 Tally the data.

STEP 3 Find the frequencies.

Class Limits	Class Boundaries	Frequency	Cumulative Frequency
100 - 104	99.5 - 104.5	2	
105 - 109	104.5 - 109.5	8	
110 - 114	109.5 - 114.5	18	
115 - 119	114.5 - 119.5	13	
120 - 124	119.5 - 124.5	7	
125 - 129	124.5 - 129.5	1	
130 - 134	129.5 - 134.5	1	

Constructing a Grouped Frequency Distribution

STEP 4 Find the cumulative frequencies by keeping a running total of the frequencies.

Class Limits	Class Boundaries	Frequency	Cumulative Frequency
100 - 104	99.5 - 104.5	2	2
105 - 109	104.5 - 109.5	8	10
110 - 114	109.5 - 114.5	18	28
115 - 119	114.5 - 119.5	13	41
120 - 124	119.5 - 124.5	7	48
125 - 129	124.5 - 129.5	1	49
130 - 134	129.5 - 134.5	1	50

2-2 Histograms, Frequency Polygons, and Ogives

3 Most Common Graphs in Research

1. **Histogram**
2. **Frequency Polygon**
3. **Cumulative Frequency Polygon (Ogive)**

2-2 Histograms, Frequency Polygons, and Ogives

The ***histogram*** is a graph that displays the data by using vertical bars of various heights to represent the frequencies of the classes.

The class boundaries are represented on the horizontal axis.

Chapter 2
Frequency Distributions and Graphs

Section 2-2

Example 2-4

Page #51

Histograms

Construct a histogram to represent the data for the record high temperatures for each of the 50 states (see Example 2–2 for the data).

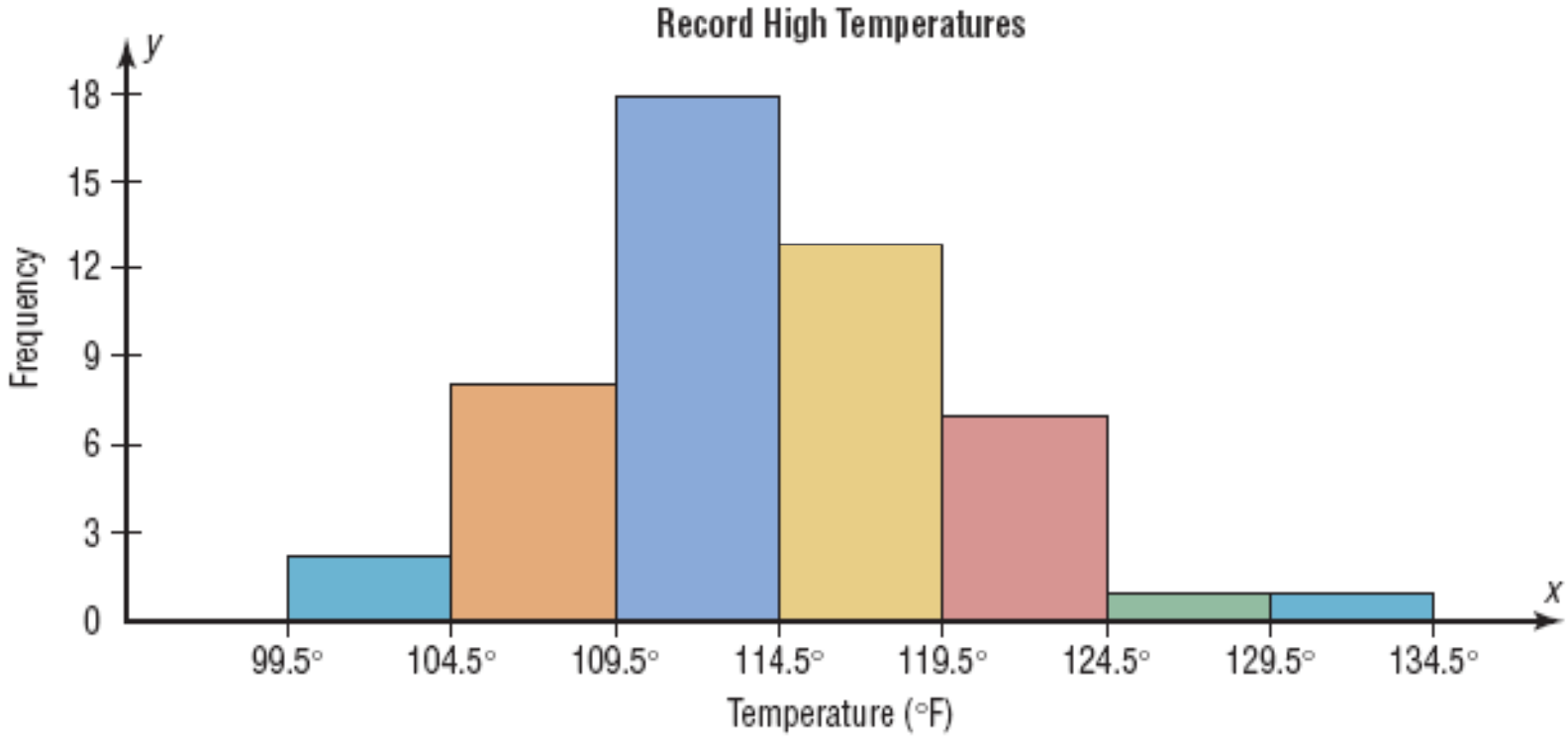
Histograms

Histograms use class boundaries and frequencies of the classes.

Class Limits	Class Boundaries	Frequency
100 - 104	99.5 - 104.5	2
105 - 109	104.5 - 109.5	8
110 - 114	109.5 - 114.5	18
115 - 119	114.5 - 119.5	13
120 - 124	119.5 - 124.5	7
125 - 129	124.5 - 129.5	1
130 - 134	129.5 - 134.5	1

Histograms

Histograms use class boundaries and frequencies of the classes.



2.2 Histograms, Frequency Polygons, and Ogives

- The ***frequency polygon*** is a graph that displays the data by using lines that connect points plotted for the frequencies at the class midpoints. The frequencies are represented by the heights of the points.
- The class midpoints are represented on the horizontal axis.

Chapter 2
Frequency Distributions and Graphs

Section 2-2

Example 2-5

Page #53

Frequency Polygons

Construct a frequency polygon to represent the data for the record high temperatures for each of the 50 states (see Example 2–2 for the data).

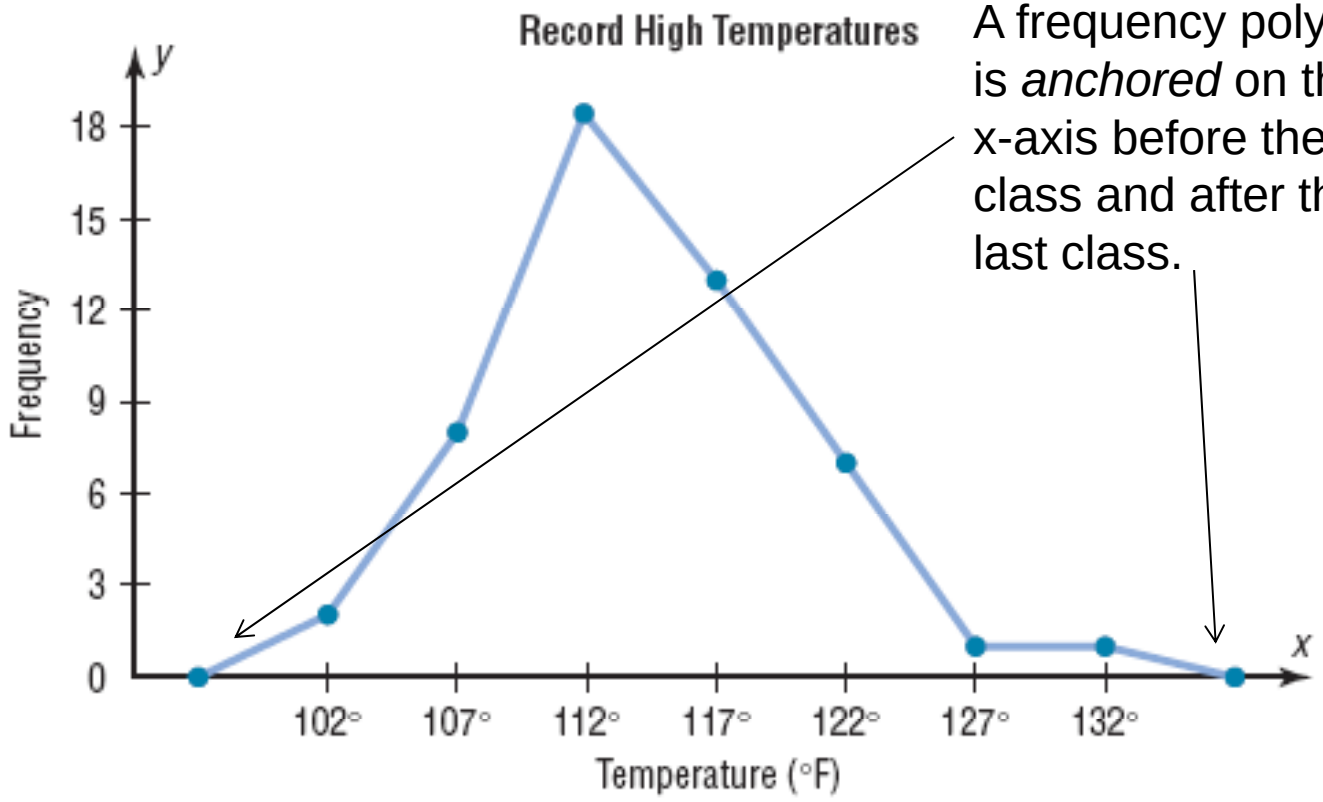
Frequency Polygons

Frequency polygons use class midpoints and frequencies of the classes.

Class Limits	Class Midpoints	Frequency
100 - 104	102	2
105 - 109	107	8
110 - 114	112	18
115 - 119	117	13
120 - 124	122	7
125 - 129	127	1
130 - 134	132	1

Frequency Polygons

Frequency polygons use class midpoints and frequencies of the classes.



A frequency polygon is *anchored* on the x-axis before the first class and after the last class.

2.2 Histograms, Frequency Polygons, and Ogives

- The ***ogive*** is a graph that represents the cumulative frequencies for the classes in a frequency distribution.

- The upper class boundaries are represented on the horizontal axis.

Chapter 2
Frequency Distributions and Graphs

Section 2-2

Example 2-6

Page #54

Ogives

Construct an ogive to represent the data for the record high temperatures for each of the 50 states (see Example 2–2 for the data).

Ogives

Ogives use upper class boundaries and cumulative frequencies of the classes.

Class Limits	Class Boundaries	Frequency	Cumulative Frequency
100 - 104	99.5 - 104.5	2	2
105 - 109	104.5 - 109.5	8	10
110 - 114	109.5 - 114.5	18	28
115 - 119	114.5 - 119.5	13	41
120 - 124	119.5 - 124.5	7	48
125 - 129	124.5 - 129.5	1	49
130 - 134	129.5 - 134.5	1	50

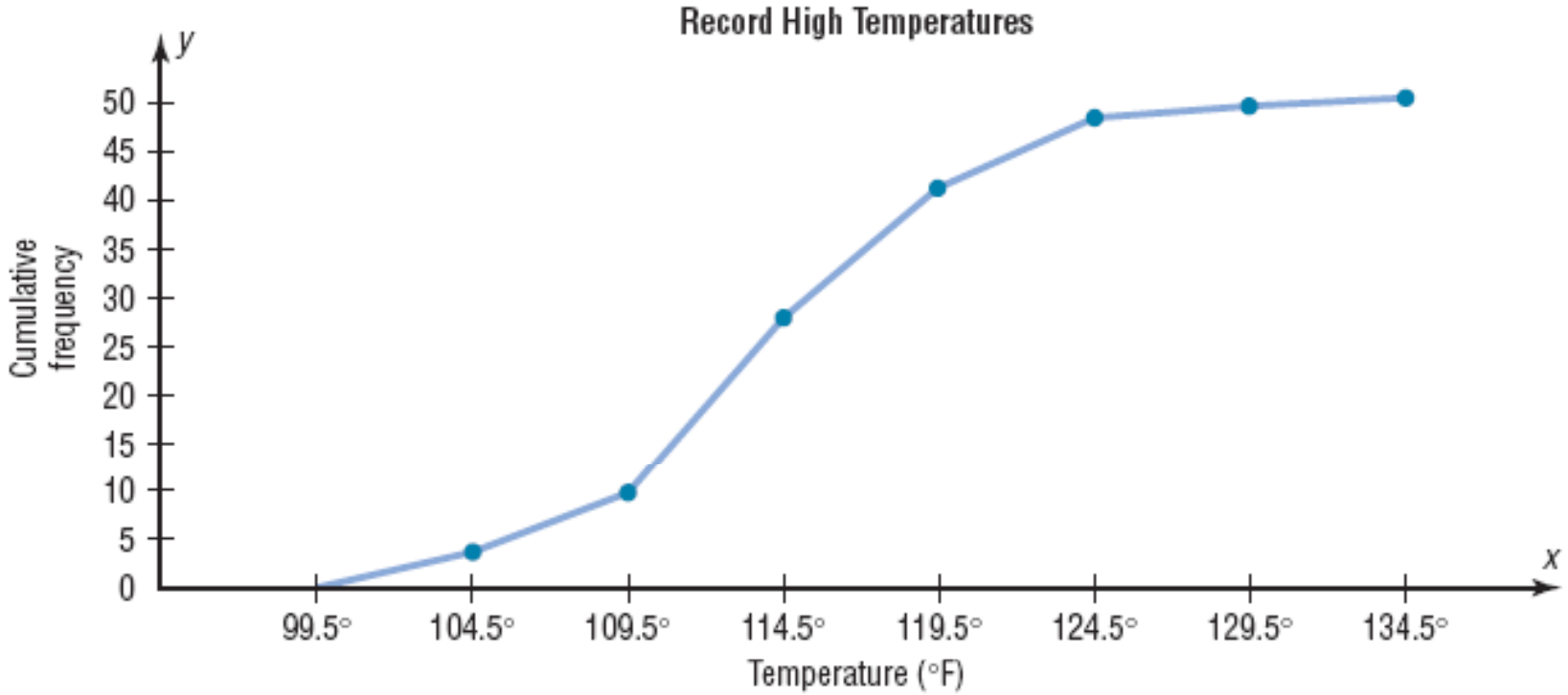
Ogives

Ogives use upper class boundaries and cumulative frequencies of the classes.

Class Boundaries	Cumulative Frequency
Less than 104.5	2
Less than 109.5	10
Less than 114.5	28
Less than 119.5	41
Less than 124.5	48
Less than 129.5	49
Less than 134.5	50

Ogives

Ogives use upper class boundaries and cumulative frequencies of the classes.



Procedure Table

Constructing Statistical Graphs

1: Draw and label the *x* and *y* axes.

2: Choose a suitable scale for the frequencies or cumulative frequencies, and label it on the *y* axis.

3: Represent the class boundaries for the histogram or ogive, or the midpoint for the frequency polygon, on the *x* axis.

4: Plot the points and then draw the bars or lines.

2.2 Histograms, Frequency Polygons, and Ogives

If proportions are used instead of frequencies, the graphs are called ***relative frequency graphs***.

Relative frequency graphs are used when the proportion of data values that fall into a given class is more important than the actual number of data values that fall into that class.

Chapter 2
Frequency Distributions and Graphs

Section 2-2

Example 2-7

Page #57

Construct a histogram, frequency polygon, and ogive using relative frequencies for the distribution (shown here) of the miles that 20 randomly selected runners ran during a given week.

Class Boundaries	Frequency
5.5 - 10.5	1
10.5 - 15.5	2
15.5 - 20.5	3
20.5 - 25.5	5
25.5 - 30.5	4
30.5 - 35.5	3
35.5 - 40.5	2

Histograms

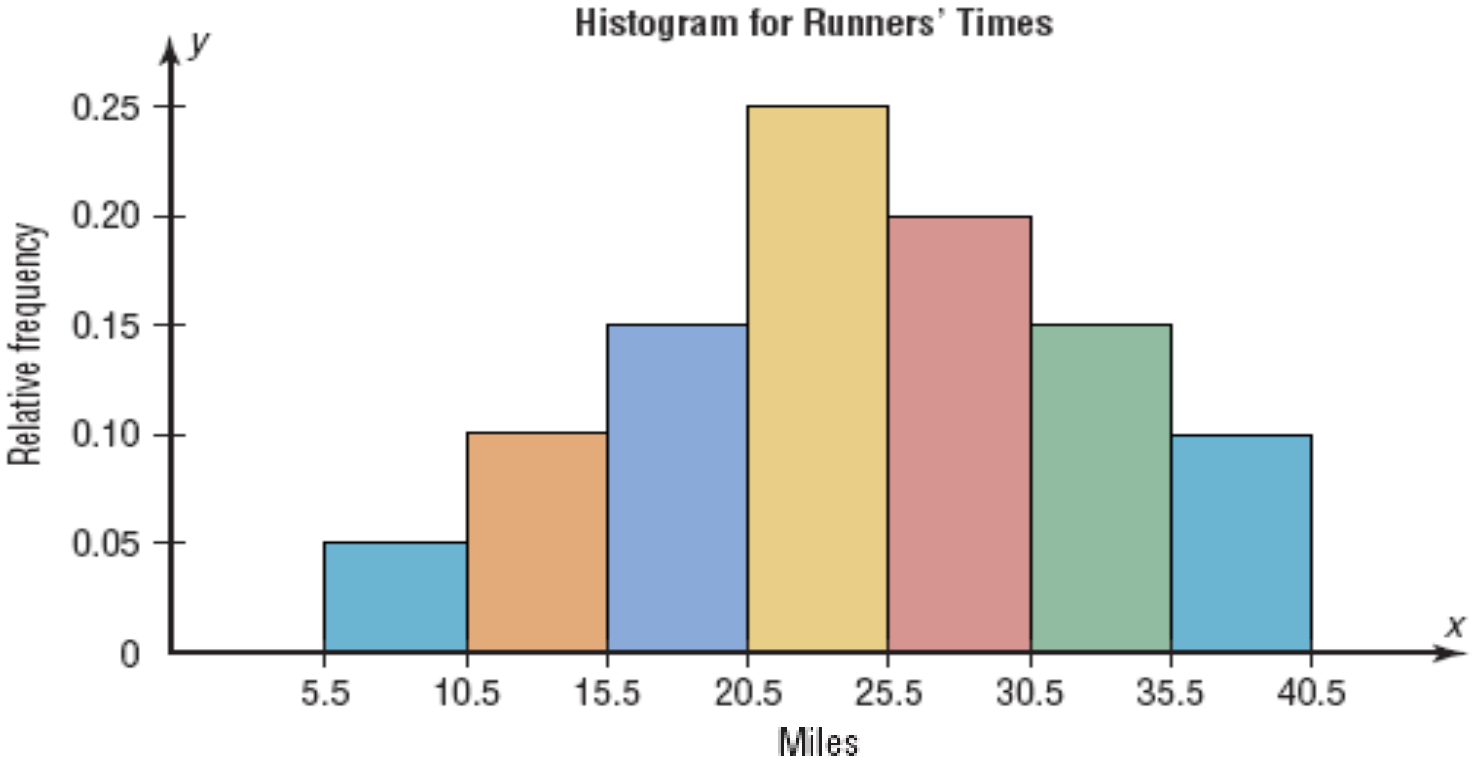
The following is a frequency distribution of miles run per week by 20 selected runners.

Class Boundaries	Frequency	Relative Frequency
5.5 - 10.5	1	1/20 = 0.05
10.5 - 15.5	2	2/20 = 0.10
15.5 - 20.5	3	3/20 = 0.15
20.5 - 25.5	5	5/20 = 0.25
25.5 - 30.5	4	4/20 = 0.20
30.5 - 35.5	3	3/20 = 0.15
35.5 - 40.5	2	2/20 = 0.10
	$\Sigma f = 20$	$\Sigma rf = 1.00$

Divide each frequency by the total frequency to get the relative frequency.

Histograms

Use the class boundaries and the relative frequencies of the classes.



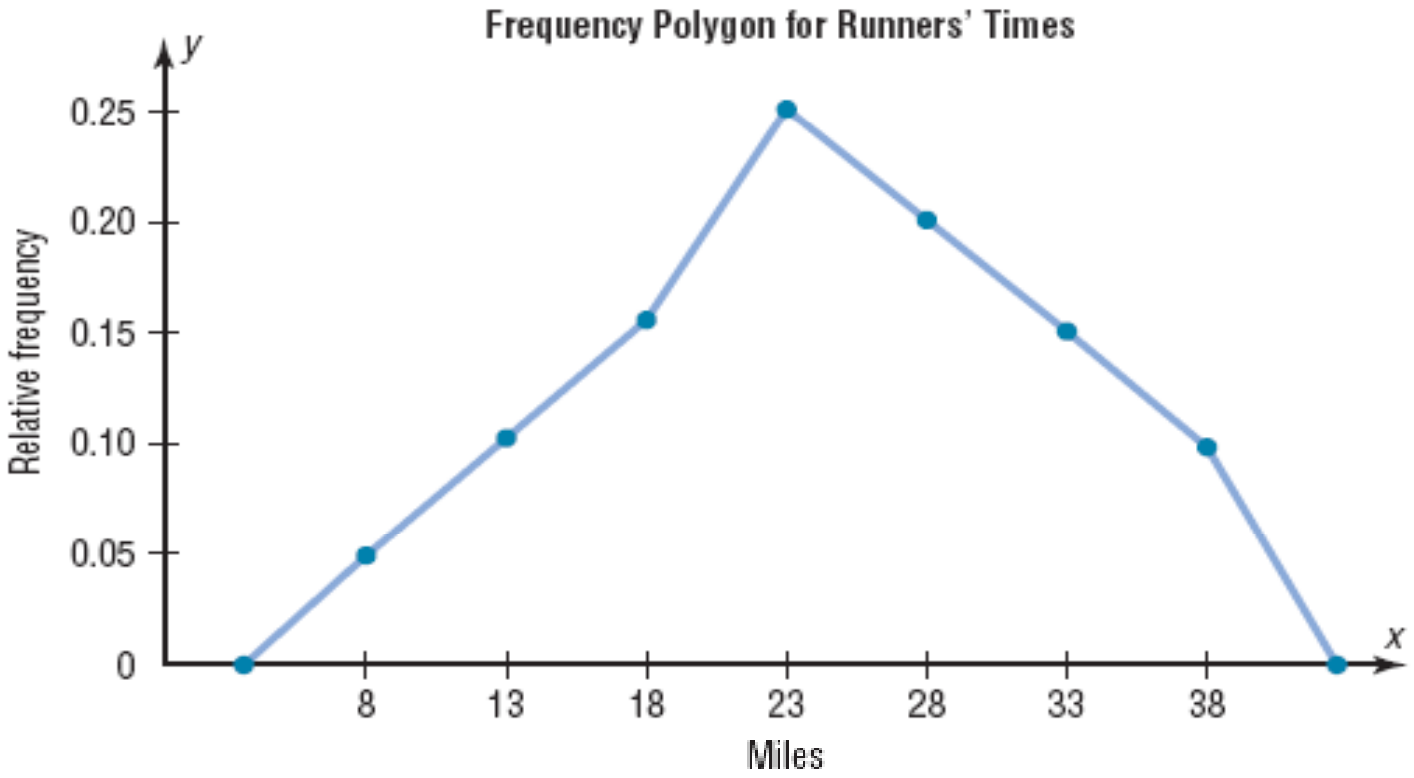
Frequency Polygons

The following is a frequency distribution of miles run per week by 20 selected runners.

Class Boundaries	Class Midpoints	Relative Frequency
5.5 - 10.5	8	0.05
10.5 - 15.5	13	0.10
15.5 - 20.5	18	0.15
20.5 - 25.5	23	0.25
25.5 - 30.5	28	0.20
30.5 - 35.5	33	0.15
35.5 - 40.5	38	0.10

Frequency Polygons

Use the class midpoints and the relative frequencies of the classes.



Ogives

The following is a frequency distribution of miles run per week by 20 selected runners.

Class Boundaries	Frequency	Cumulative Frequency	Cum. Rel. Frequency
5.5 - 10.5	1	1	1/20 = 0.05
10.5 - 15.5	2	3	3/20 = 0.15
15.5 - 20.5	3	6	6/20 = 0.30
20.5 - 25.5	5	11	11/20 = 0.55
25.5 - 30.5	4	15	15/20 = 0.75
30.5 - 35.5	3	18	18/20 = 0.90
35.5 - 40.5	2	20	20/20 = 1.00
	$\Sigma f = 20$		

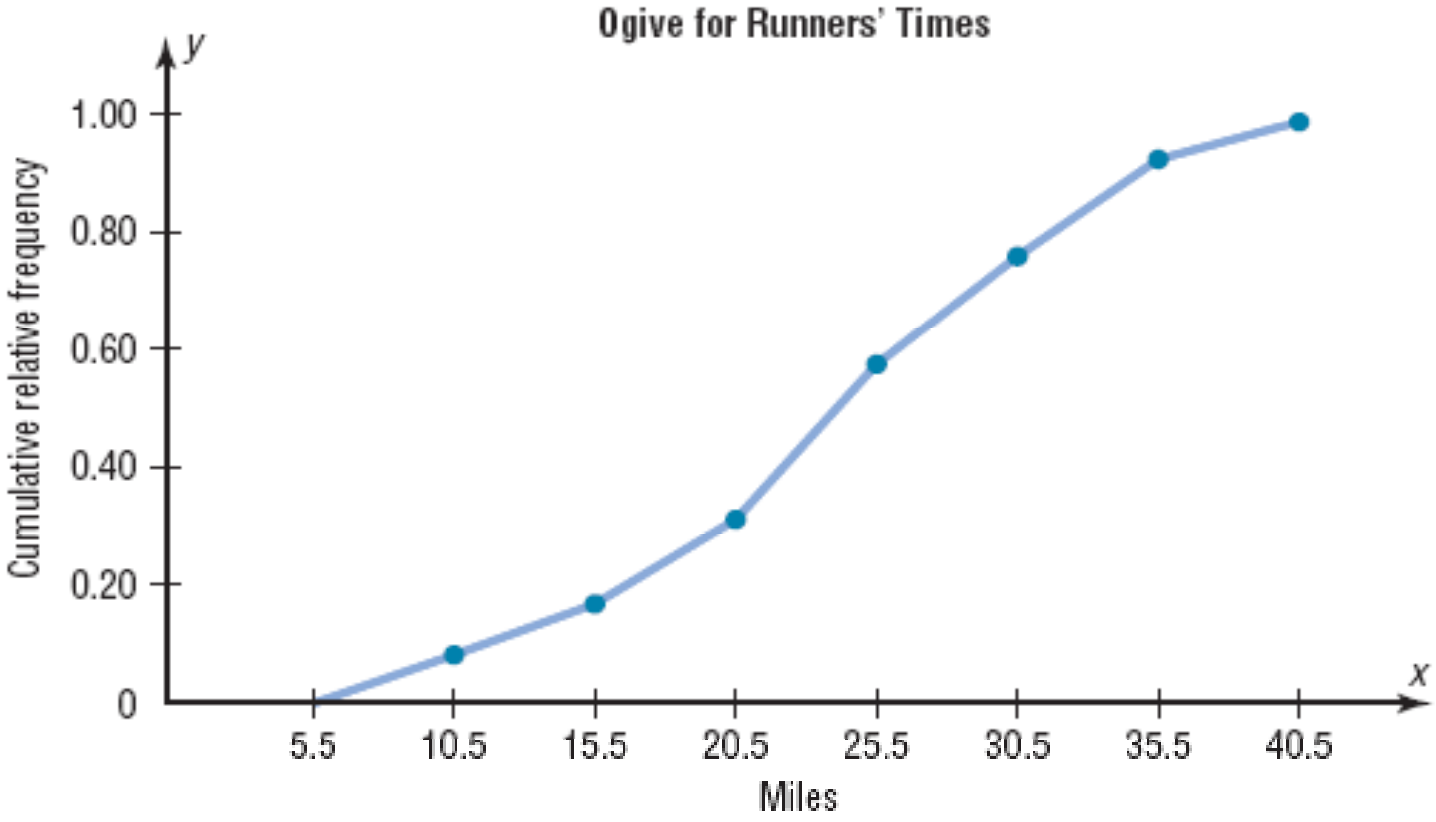
Ogives

Ogives use upper class boundaries and cumulative frequencies of the classes.

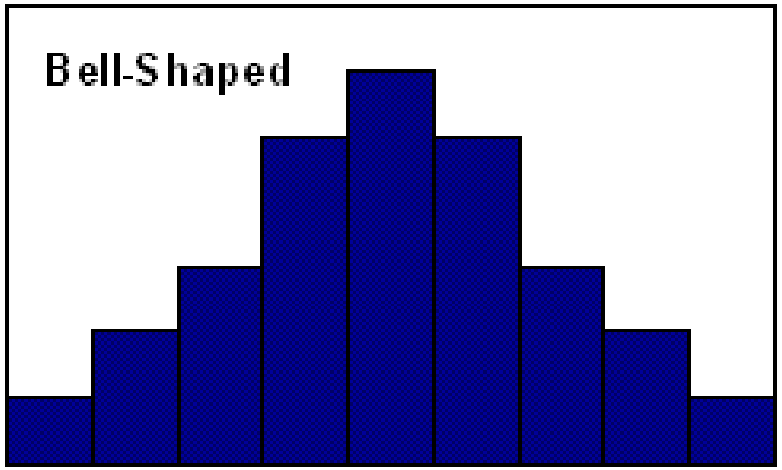
Class Boundaries	Cum. Rel. Frequency
Less than 10.5	0.05
Less than 15.5	0.15
Less than 20.5	0.30
Less than 25.5	0.55
Less than 30.5	0.75
Less than 35.5	0.90
Less than 40.5	1.00

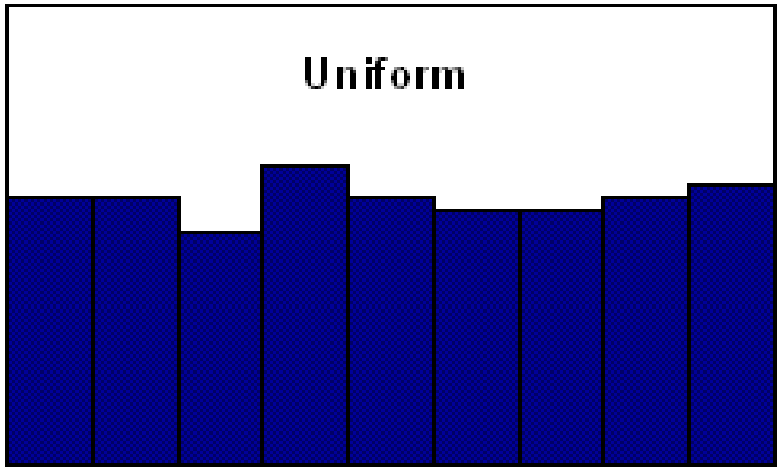
Ogives

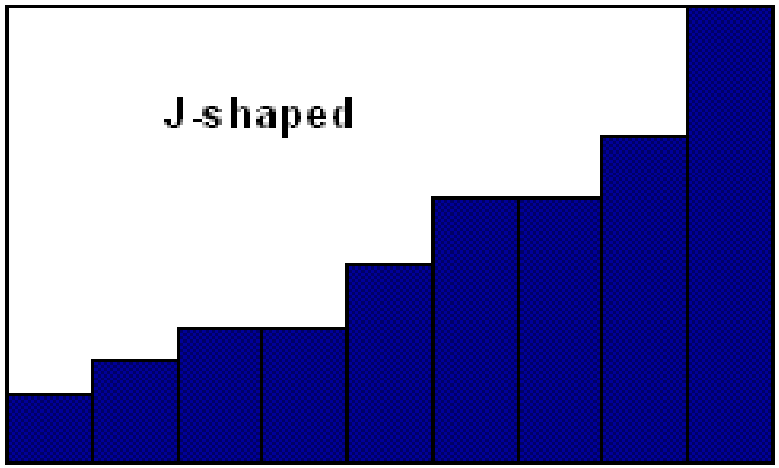
Use the upper class boundaries and the cumulative relative frequencies.

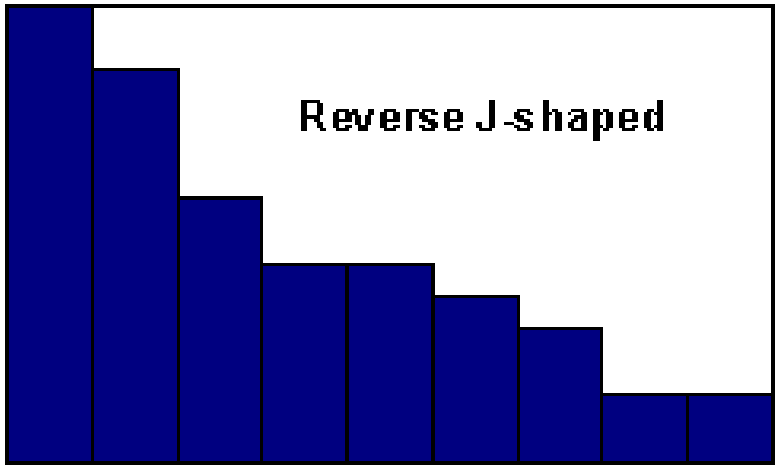


Shapes of Distributions

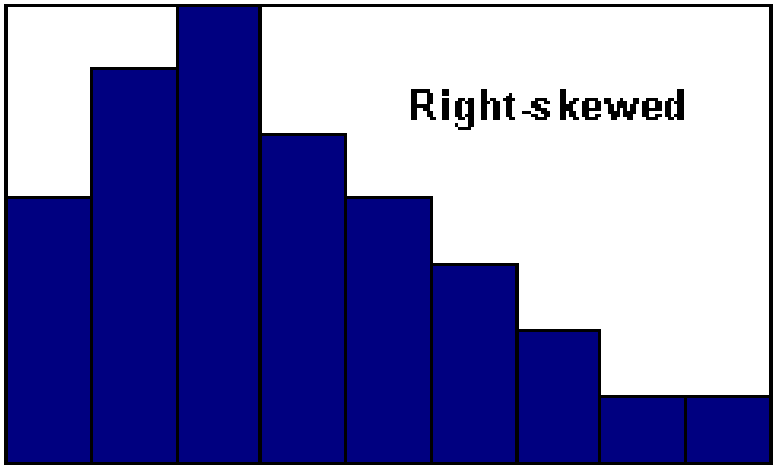


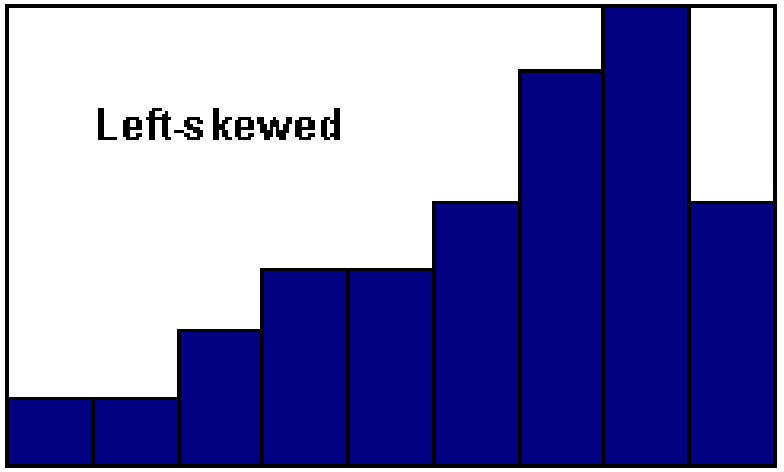


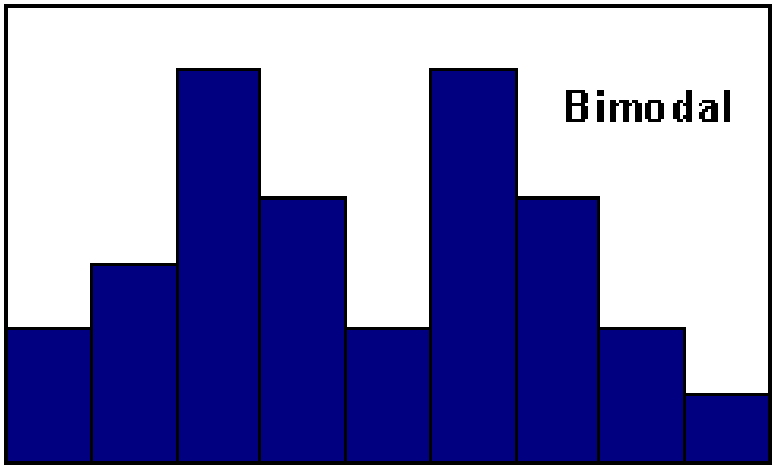


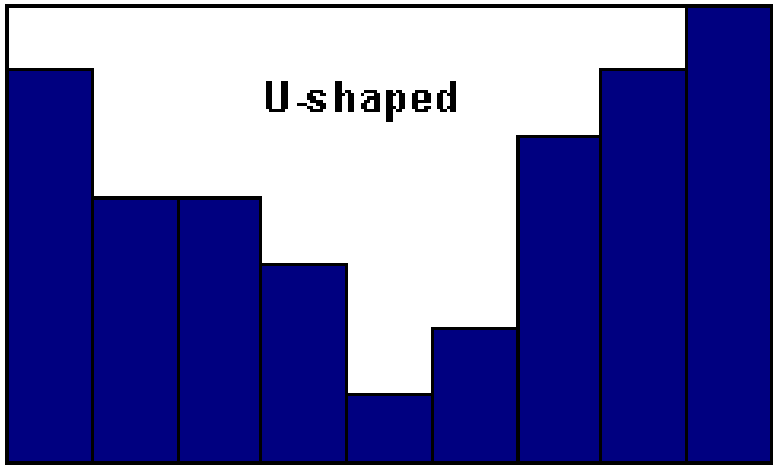


Shapes of Distributions

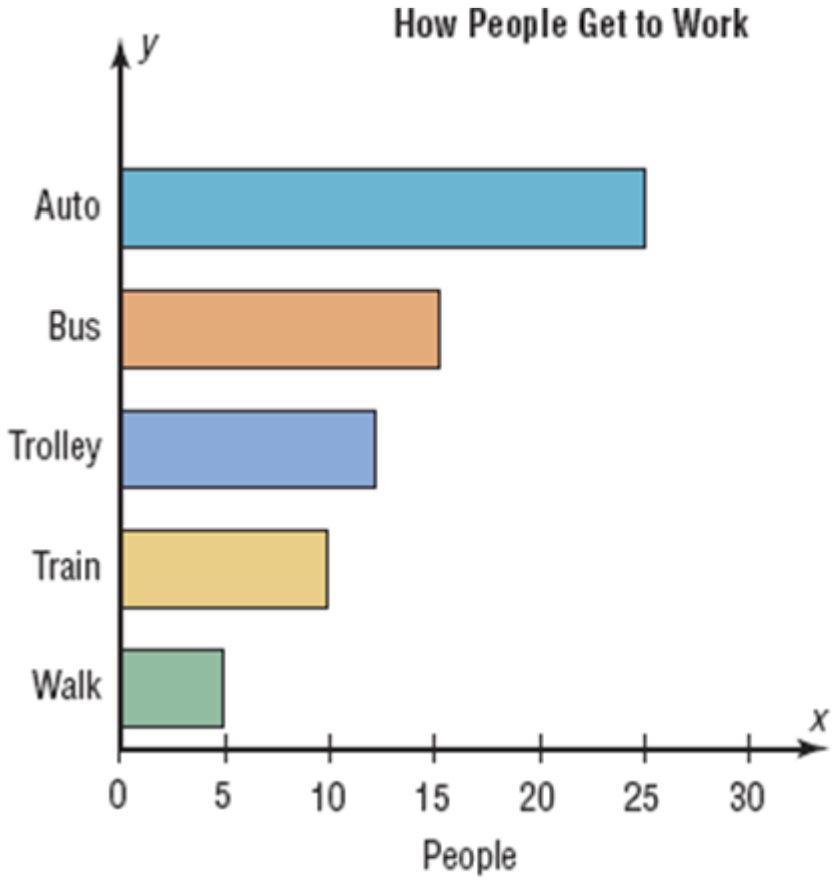




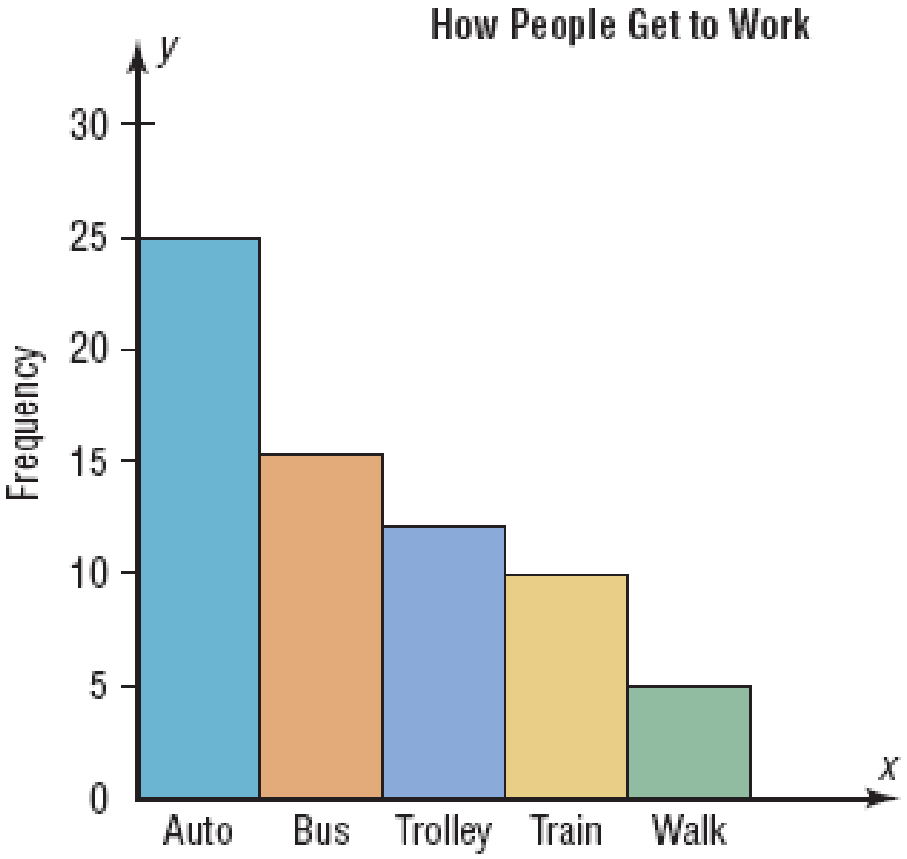




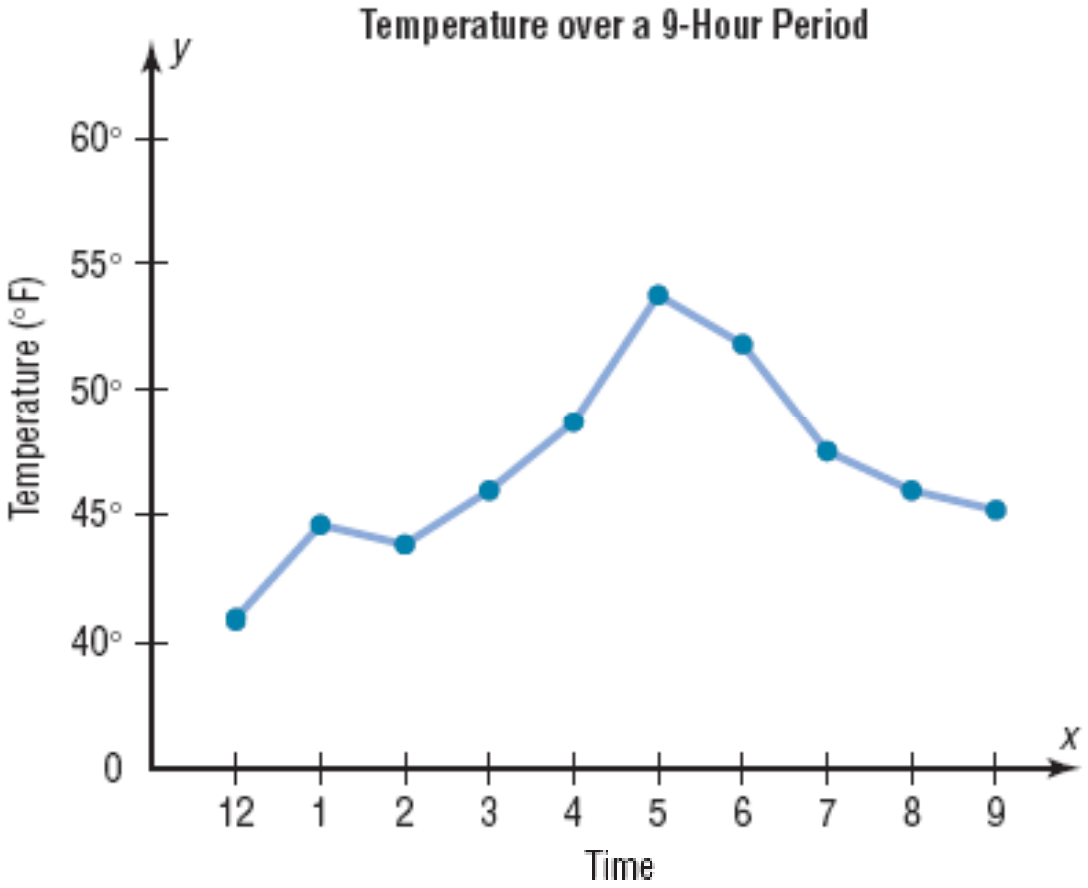
2.3 Other Types of Graphs
Bar Graphs



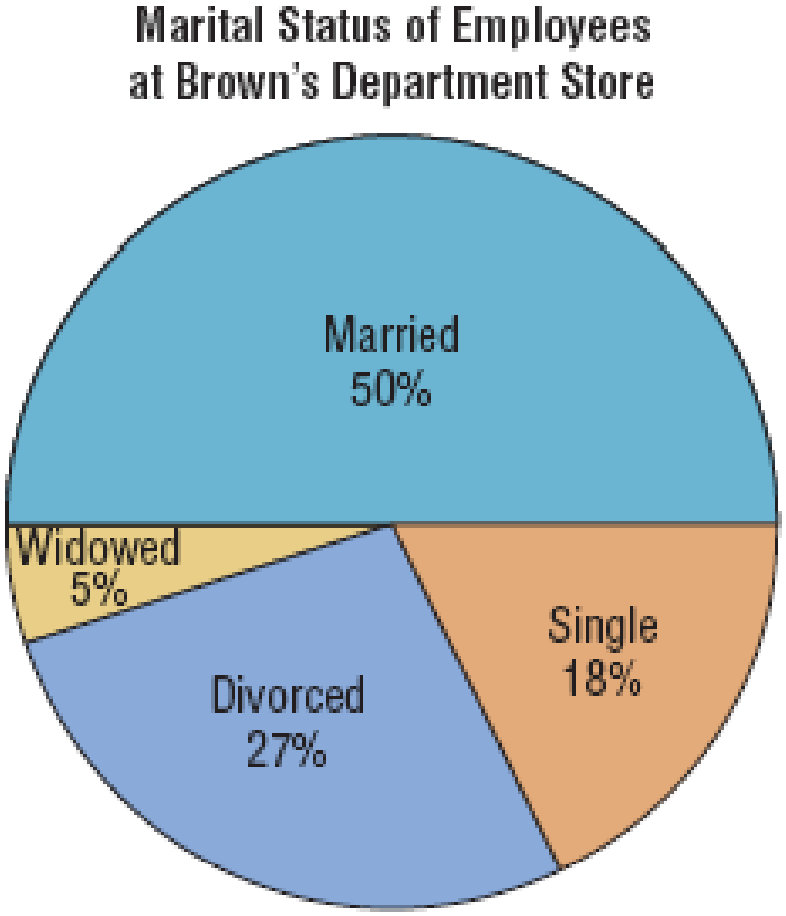
2.3 Other Types of Graphs
Pareto Charts



2.3 Other Types of Graphs
Time Series Graphs



2.3 Other Types of Graphs
Pie Graphs



2.3 Other Types of Graphs
Stem and Leaf Plots

A ***stem and leaf plots*** is a data plot that uses part of a data value as the stem and part of the data value as the leaf to form groups or classes.

It has the advantage over grouped frequency distribution of retaining the actual data while showing them in graphic form.

Chapter 2
Frequency Distributions and Graphs

Section 2-3

Example 2-13

Page #80

At an outpatient testing center, the number of cardiograms performed each day for 20 days is shown. Construct a stem and leaf plot for the data.

25	31	20	32	13
14	43	2	57	23
36	32	33	32	44
32	52	44	51	45

25	31	20	32	13
14	43	2	57	23
36	32	33	32	44
32	52	44	51	45

Unordered Stem Plot

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

Ordered Stem Plot

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