Chapter 1
Introduction

Section 1.1

1.1 1) **Statistics** refers to numerical facts such as the age of a student or the income of a family.

2) Statistics refers to the field or discipline of study. Statistics is a group of methods used to collect, analyze, present, and interpret data and to make decisions.

1.2 **Descriptive statistics** consists of methods that help us organize, display, and describe data using tables, graphs, and summary measures. **Inferential statistics** consists of methods that use sample results to help make decisions or predictions about a population.

Section 1.2

1.3 A **population** is the collection of all elements whose characteristics are being studied. A s**ample** is a portion of the population selected for study. A r**epresentative sample** is a sample that represents the characteristics of the population as closely as possible, and a **random sample** is a sample drawn in such a way that each element of the population has a chance of being included in the sample. **Sampling with replacement** refers to a sampling procedure in which the item selected at each selection is put back in the population before the next item is drawn; s**ampling without replacement** is a sampling procedure in which the item selected at each selection is not replaced in the population.

1.4 Consider a standard deck of 52 cards. Suppose we randomly select one card from the deck and record the value and suit. If we put this card back in the deck before we randomly select a second card, this is an example of **sampling with replacement**. If we lay the first card aside and randomly select the second card from the 51 cards remaining in the deck, this is an example of **sampling without replacement**.

1.5 A **census** is a survey that includes every member of the population. A survey based on a portion of the population is called a **sample survey**. A sample survey is preferred over a census for the following reasons:

1) Conducting a census is very expensive because the size of the population is often very large.
2) Conducting a census is very time consuming.
3) In many cases it is impossible to identify each element of the target population.

1.6 **a.** Population **b.** Sample
c. Sample **d.** Population
e. Sample

1.7 **a.** Population **b.** Sample
c. Population **d.** Population
e. Sample

Section 1.3

1.8 An **element** is a specific subject or object about which the information is collected. A **variable** is a characteristic under study that assumes different values for different elements. An **observation** is the value of a variable for a single element. A **data set** is a collection of observations on one or more variables.

1.9 With reference to this table, we have the following definitions:

- Member: Each city included in the table
- Variable: Number of dog bites reported last year
- Measurement: Number of dog bites in a specific city
- Data set: Collection of dog bite numbers for the six cities listed in the table.

1.10 With reference to this table, we have the following definitions:

- Member: Each country included in the table
- Variable: The number of billionaires
- Measurement: The number of billionaires in a specific country
- Data set: Collection of the number of billionaires in each country listed in the table

1.11 **a.** Number of dog bites reported last year

b. Six

c. Six (cities)

1.12 **a.** Number of billionaires

b. Eight

c. Eight (countries)

Section 1.4

1.13 **a.** A **quantitative variable** is a variable that can be measured numerically.

b. A variable that cannot assume a numeric value but can be classified into two or more nonnumeric categories is called a **qualitative variable**.

c. A **discrete variable** is a variable whose values are countable.

d. A variable that can assume any numerical value over a certain interval or intervals is called a **continuous variable**.

e. Data collected on a quantitative variable are called **quantitative data**.

f. **Qualitative data** are data collected on a qualitative variable.

1.14 **a.** Quantitative **b.** Qualitative
c. Qualitative **d.** Quantitative
e. Quantitative

1.15 **a.** Quantitative **b.** Quantitative
c. Qualitative **d.** Qualitative
e. Quantitative

1.16 **a.** Discrete **b.** not applicable
c. not applicable **d.** Continuous
e. Discrete

1.17 **a.** Continuous **b.** Continuous
c. not applicable **d.** not applicable
e. Continuous

Section 1.6

1.18 Data collected on different elements at the same point in time or for the same period of time are called **cross-section data**. Total sales for the 2011 Christmas season at 10 stores in a particular mall is an example of cross-section data. Data collected on the same element for the same variable at different points in time or for different periods of time are called **time-series data**. Total sales for one particular store for the Christmas season for the years 2005 to 2011 is an example of time-series data.

1.19 **Internal sources** of data are sources inside the organization conducting the study which make needed data available**. External sources** of data are the sources outside the organization from which data is available**.**

1.20 **a.** Time-series data

b. Time-series data

c. Cross-section data

d. Cross-section data

1.21 **a.** Cross-section data

b. Cross-section data

c. Time-series data

d. Time-series data

Section 1.7

1.22

m	f	f^2	mf	m^2f
5	12	144	60	300
10	8	64	80	800
17	6	36	102	1734
20	16	256	320	6400
25	4	16	100	2500
$\Sigma m = 77$	$\Sigma f = 46$	$\Sigma f^2 = 516$	$\Sigma mf = 662$	$\Sigma m^2 f = 11{,}734$

a. $\Sigma m = 77$ **b.** $\Sigma f^2 = 516$ **c.** $\Sigma mf = 662$ **d.** $\Sigma m^2 f = 11{,}734$

1.23

m	f	m^2	mf	m^2f
3	16	9	48	144
6	11	36	66	396
25	16	625	400	10,000
12	8	144	96	1152
15	4	225	60	900
18	14	324	252	4536
$\Sigma m = 79$	$\Sigma f = 69$	$\Sigma m^2 = 1363$	$\Sigma mf = 922$	$\Sigma m^2 f = 17{,}128$

a. $\Sigma f = 69$ **b.** $\Sigma m^2 = 1363$ **c.** $\Sigma mf = 922$ **d.** $\Sigma m^2 f = 17{,}128$

1.24

x	y	xy	y^2
54	1	54	1
28	2	56	4
16	3	48	9
11	4	44	16
7	5	35	25
6	6	36	36
1	7	7	49
5	8	40	64
1	9	9	81
3	11	33	121
$\Sigma x = 132$	$\Sigma y = 56$	$\Sigma xy = 362$	$\Sigma y^2 = 406$

a. $\Sigma x = 132$ **b.** $\Sigma y = 56$

c. $\Sigma xy = 362$ **d.** $\Sigma y^2 = 406$

e. $(\Sigma y)^2 = 56^2 = 3136$

1.25

x	y	xy	y^2
63	1	63	1
29	2	58	4
13	3	39	9
9	4	36	16
1	5	5	25
2	6	12	36
1	7	7	49
1	8	8	64
1	9	9	81
$\Sigma x = 120$	$\Sigma y = 45$	$\Sigma xy = 237$	$\Sigma y^2 = 285$

a. $\Sigma x = 120$ **b.** $\Sigma y = 45$

c. $\Sigma xy = 237$ **d.** $\Sigma y^2 = 285$

e. $(\Sigma y)^2 = 45^2 = 2025$

1.26 **a.** $\Sigma y = 216 + 184 + 35 + 92 + 144 + 175 + 11 + 57 = \$914$

b. $(\Sigma y)^2 = (914)^2 = 835{,}396$

c. $\Sigma y^2 = (216)^2 + (184)^2 + (35)^2 + (92)^2 + (144)^2 + (175)^2 + (11)^2 + (57)^2 = 144{,}932$

1.27 **a.** $\Sigma x = 48 + 103 + 95 + 188 + 286 + 136 = 856$ pizzas

b. $(\Sigma x)^2 = (856)^2 = 732{,}736$

c. $\Sigma x^2 = (48)^2 + (103)^2 + (95)^2 + (188)^2 + (286)^2 + (136)^2 = 157{,}574$

1.28 **a.** $\Sigma y = 975 + 520 + 1560 + 872 + 1105 + 437 + 910 + 785 + 1335 = 8499$ calories

b. $(\Sigma y)^2 = (8499)^2 = 72{,}233{,}001$

c. $\Sigma y^2 = (975)^2 + (520)^2 + (1560)^2 + (872)^2 + (1105)^2 + (437)^2 + (910)^2 + (785)^2 + (1335)^2 = 9{,}053{,}553$

1.29 **a.** $\Sigma x = 387 + 414 + 404 + 396 + 410 + 422 + 414 = 2847$ miles

b. $(\Sigma x)^2 = (2847)^2 = 8{,}105{,}409$

c. $\Sigma x^2 = (387)^2 + (414)^2 + (404)^2 + (396)^2 + (410)^2 + (422)^2 + (414)^2 = 1{,}158{,}777$

Supplementary Exercises

1.30 With reference to this table, we have the following definitions:

- Member: Each airline included in the table
- Variable: Mishandled baggage reports
- Measurement: The number of mishandled baggage reports for an individual airline
- Data Set: Collection of the number of mishandled baggage reports for the airlines listed in the table

1.31 With reference to this table, we have the following definitions

- Member: Each CEO included in the table
- Variable: The total compensation
- Measurement: The total compensation for a specific CEO
- Data Set: Collection of the amounts of total compensation for the CEOs listed in the table

1.32 Since the data set in Exercise 1.30 contains measurements on different elements for the same period of time (the first nine months of 2010 for six airlines), it is cross-section data. The data set in Exercise 1.31 contains measurements on different elements for the same period of time (year 2010 for nine CEOs), it is cross-section data.

1.33 **a.** Sample **b.** Population

c. Sample **d.** Population

1.34 **a.** Population **b.** Sample

c. Sample **d.** Population

1.35 **a.** This is an example of sampling without replacement because once a patient is selected, he/she will not be replaced before the next patient is selected.

b. This is an example of sampling with replacement because both times the selection is made from the same group of professors.

1.36 **a.** This is an example of sampling without replacement because once a city is selected it will not be replaced before the next city is selected.

b. This is an example of sampling with replacement because both times the selection is made from the same group of high school teachers.

1.37 **a.** $\Sigma x = 8 + 14 + 3 + 7 + 10 + 5 = 47$ shoe pairs

b. $(\Sigma x)^2 = (47)^2 = 2209$

c. $\Sigma x^2 = (8)^2 + (14)^2 + (3)^2 + (7)^2 + (10)^2 + (5)^2 = 443$

1.38 **a.** $\Sigma y = 4 + 12 + 8 + 10 + 5 = 39$ restaurants

b. $(\Sigma y)^2 = (39)^2 = 1521$

c. $\Sigma y^2 = (4)^2 + (12)^2 + (8)^2 + (10)^2 + (5)^2 = 349$

1.39

m	f	f^2	mf	m^2f	m^2
3	7	49	21	63	9
16	32	1024	512	8192	256
11	17	289	187	2057	121
9	12	144	108	972	81
20	34	1156	680	13,600	400
$\Sigma m = 59$	$\Sigma f = 102$	$\Sigma f^2 = 2662$	$\Sigma mf = 1508$	$\Sigma m^2 f = 24{,}884$	$\Sigma m^2 = 867$

a. $\Sigma m = 59$ **b.** $\Sigma f^2 = 2662$ **c.** $\Sigma mf = 1508$ **d.** $\Sigma m^2 f = 24{,}884$ **e.** $\Sigma m^2 = 867$

1.40

x	y	x^2	xy	x^2y	y^2
7	5	49	35	245	25
11	15	121	165	1815	225
8	7	64	56	448	49
4	10	16	40	160	100
14	9	196	126	1764	81
28	19	784	532	14,896	361
$\Sigma x = 72$	$\Sigma y = 65$	$\Sigma x^2 = 1230$	$\Sigma xy = 954$	$\Sigma x^2y = 19{,}328$	$\Sigma y^2 = 841$

a. $\Sigma y = 65$ **b.** $\Sigma x^2 = 1230$ **c.** $\Sigma xy = 954$ **d.** $\Sigma x^2y = 19{,}328$ **e.** $\Sigma y^2 = 841$

1.41 *Draft round*: quantitative, discrete
40-yard dash speed: quantitative, continuous
position: qualitative
drafting team's current payroll: quantitative, continuous
power ratio: quantitative, continuous
quality starter: qualitative
standing high jump: quantitative, continuous

Self-Review Test

1. b **2.** c

3. **a.** Sample without replacement

b. Sample with replacement

4. **a.** Qualitative

b. Quantitative; continuous

c. Quantitative; discrete

d. Qualitative

5. With reference to this table, we have the following definitions:

- Member: Each category listed in the table
- Variable: The amount spent on different categories of products by all people in Canada during May 2011
- Measurement: The amount spent on a specific category of products by all people in Canada during May 2011
- Data set: Collection of the amounts spent on different categories of products by all people in Canada during May 2011 for the eleven categories listed in the table

6. **a.** $\Sigma x = 6 + 11 + 3 + 5 + 6 + 2 = 33$ types of cereal

b. $(\Sigma x)^2 = (33)^2 = 1089$

c. $\Sigma x^2 = (6)^2 + (11)^2 + (3)^2 + (5)^2 + (6)^2 + (2)^2 = 231$

7.

m	f	m^2	mf	m^2f	f^2
2	82	4	164	328	6724
3	278	9	834	2502	77284
4	43	16	172	688	1849
5	16	25	80	400	256
6	6	36	36	216	36
7	3	49	21	147	9
8	1	64	8	64	1
$\Sigma m = 35$	$\Sigma f = 429$	$\Sigma m^2 = 203$	$\Sigma mf = 1315$	$\Sigma m^2 f = 4345$	$\Sigma(f^2) = 86{,}159$

a. $\Sigma m = 35$ **b.** $\Sigma f = 429$ **c.** $\Sigma m^2 = 203$ **d.** $\Sigma mf = 1315$ **e.** $\Sigma m^2 f = 4345$ **f.** $(\Sigma f)^2 = 429^2 = 184{,}041$