

## SECTION 1.1

# Exercises

9. **Cool car colors** The most popular colors for cars and light trucks change over time. Silver passed green in 2000 to become the most popular color worldwide, then gave way to shades of white in 2007. Here is the distribution of colors for vehicles sold in North America in 2008.<sup>8</sup>

Color	Percent of vehicles
White	20
Black	17
Silver	17
Blue	13
Gray	12
Red	11
Beige/brown	5
Green	3
Yellow/gold	2

- (a) What percent of vehicles had colors other than those listed?
- (b) Display these data in a bar graph. Be sure to label your axes and title your graph.
- (c) Would it be appropriate to make a pie chart of these data? Explain.

10. **Spam** Email spam is the curse of the Internet. Here is a compilation of the most common types of spam:<sup>9</sup>

Type of spam	Percent
Adult	19
Financial	20
Health	7
Internet	7
Leisure	6
Products	25
Scams	9
Other	??

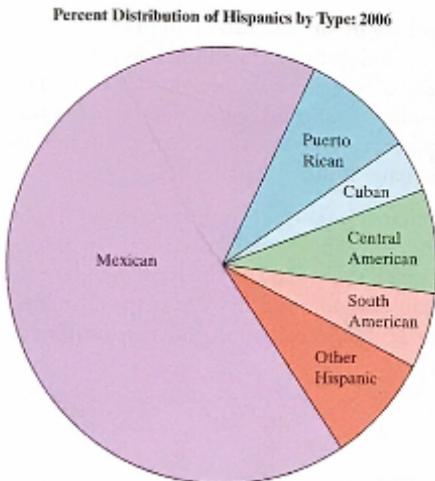
- (a) What percent of spam would fall in the “Other” category?
- (b) Display these data in a bar graph. Be sure to label your axes and title your graph.
- (c) Would it be appropriate to make a pie chart of these data? Explain.

11. **Birth days** Births are not evenly distributed across the days of the week. Here are the average numbers of babies born on each day of the week in the United States in a recent year:<sup>10</sup>

Day	Births
Sunday	7,374
Monday	11,704
Tuesday	13,169
Wednesday	13,038
Thursday	13,013
Friday	12,664
Saturday	8,459

- (a) Present these data in a well-labeled bar graph. Would it also be correct to make a pie chart?
- (b) Suggest some possible reasons why there are fewer births on weekends.
12. Deaths among young people Among persons aged 15 to 24 years in the United States, the leading causes of death and number of deaths in a recent year were as follows: accidents, 15,567; homicide, 5359; suicide, 4139; cancer, 1717; heart disease, 1067; congenital defects, 483.<sup>11</sup>
- (a) Make a bar graph to display these data.
- (b) To make a pie chart, you need one additional piece of information. What is it?

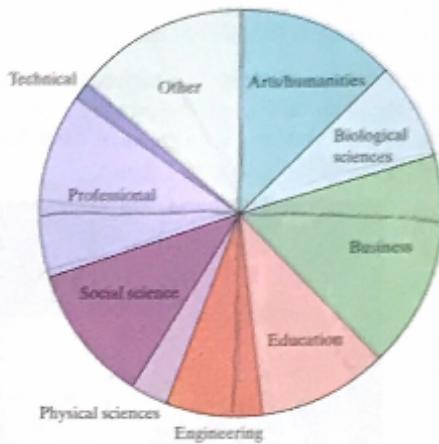
13. Hispanic origins Below is a pie chart prepared by the Census Bureau to show the origin of the more than 43 million Hispanics in the United States in 2006.<sup>12</sup> About what percent of Hispanics are Mexican? Puerto Rican?



*Comment:* You see that it is hard to determine numbers from a pie chart. Bar graphs are much easier to use. (The Census Bureau did include the percents in its pie chart.)

14. Which major? About 1.6 million first-year students enroll in colleges and universities each year. What do they plan to study? The pie chart displays data on the percents of first-year students who plan to major

in several discipline areas.<sup>13</sup> About what percent of first-year students plan to major in business? In social science?



pg 10

15. Buying music online Young people are more likely than older folk to buy music online. Here are the percents of people in several age groups who bought music online in 2006:<sup>14</sup>

Age group	Bought music online
12 to 17 years	24%
18 to 24 years	21%
25 to 34 years	20%
35 to 44 years	16%
45 to 54 years	10%
55 to 64 years	3%
65 years and over	1%

- (a) Explain why it is *not* correct to use a pie chart to display these data.
- (b) Make a bar graph of the data. Be sure to label your axes and title your graph.

16. The audience for movies Here are data on the percent of people in several age groups who attended a movie in the past 12 months:<sup>15</sup>

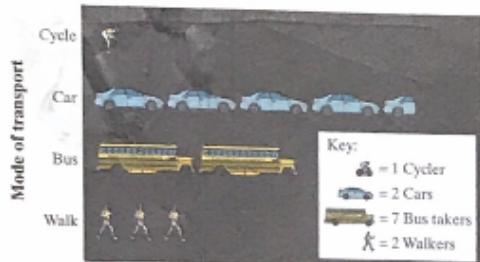
Age group	Movie attendance
18 to 24 years	83%
25 to 34 years	73%
35 to 44 years	68%
45 to 54 years	60%
55 to 64 years	47%
65 to 74 years	32%
75 years and over	20%

- (a) Display these data in a bar graph. Describe what you see.

- (b) Would it be correct to make a pie chart of these data? Why or why not?  
 (c) A movie studio wants to know what percent of the total audience for movies is 18 to 24 years old. Explain why these data do not answer this question.

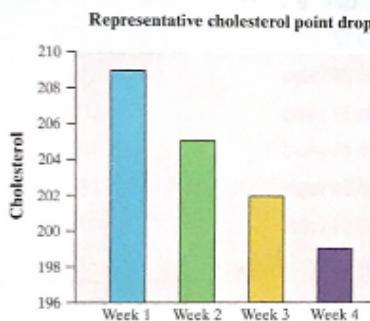
- 17. Going to school** Students in a high school statistics class were given data about the primary method of transportation to school for a group of 30 students. They produced the pictograph shown.

pg 11



- (a) How is this graph misleading?  
 (b) Make a new graph that isn't misleading.  
**18. Oatmeal and cholesterol** Does eating oatmeal reduce cholesterol? An advertisement included the following graph as evidence that the answer is "Yes."

pg 15



- (a) How is this graph misleading?  
 (b) Make a new graph that isn't misleading. What do you conclude about the effect of eating oatmeal on cholesterol reduction?

pg 13

- 19. Attitudes toward recycled products** Recycling is supposed to save resources. Some people think recycled products are lower in quality than other products, a fact that makes recycling less practical. People who actually use a recycled product may have different opinions from those who don't use it. Here are data on attitudes toward coffee filters made of recycled paper among people who do and don't buy these filters:<sup>16</sup>

Think the quality of the recycled product is:

	Higher	The same	Lower
Buyers	20	7	9
Nonbuyers	29	25	43

- (a) How many people does this table describe? How many of these were buyers of coffee filters made of recycled paper?

- (b) Give the marginal distribution of opinion about the quality of recycled filters. What percent think the quality of the recycled product is the same or higher than the quality of other filters?

- 20. Smoking by students and parents** Here are data from a survey conducted at eight high schools on smoking among students and their parents:<sup>17</sup>

	Neither parent smokes	One parent smokes	Both parents smoke
Student does not smoke	1168	1823	1380
Student smokes	188	416	400

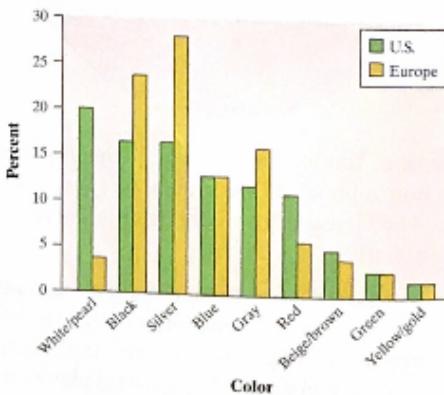
- (a) How many students are described in the two-way table? What percent of these students smoke?

- (b) Give the marginal distribution of parents' smoking behavior, both in counts and in percents.

- 21. Attitudes toward recycled products** Exercise 19 gives data on the opinions of people who have and have not bought coffee filters made from recycled paper. To see the relationship between opinion and experience with the product, find the conditional distributions of opinion (the response variable) for buyers and nonbuyers. What do you conclude?

- 22. Smoking by students and parents** Refer to Exercise 20. Calculate three conditional distributions of students' smoking behavior: one for each of the three parental smoking categories. Describe the relationship between the smoking behaviors of students and their parents in a few sentences.

- 23. Popular colors—here and there** Favorite vehicle colors may differ among countries. The side-by-side bar graph shows data on the most popular colors of cars in 2008 for the United States and Europe. Write a few sentences comparing the two distributions.



24. Comparing car colors Favorite vehicle colors may differ among types of vehicle. Here are data on the most popular colors in 2008 for luxury cars and for SUVs, trucks, and vans.

Color	Luxury cars (%)	SUVs, trucks, vans (%)
Black	22	13
Silver	16	16
White pearl	14	1
Gray	12	13
White	11	25
Blue	7	10
Red	7	11
Yellow/gold	6	1
Green	3	4
Beige/brown	2	6

- (a) Make a graph to compare colors by vehicle type.  
 (b) Write a few sentences describing what you see.

- pg 18 STEP 4 25. Snowmobiles in the park Yellowstone National Park surveyed a random sample of 1526 winter visitors to the park. They asked each person whether they owned, rented, or had never used a snowmobile. Respondents were also asked whether they belonged to an environmental organization (like the Sierra Club). The two-way table summarizes the survey responses.

Environmental Clubs			
	No	Yes	Total
Never used	445	212	657
Snowmobile renter	497	77	574
Snowmobile owner	279	16	295
Total	1221	305	1526

Do these data provide convincing evidence of an association between environmental club membership and snowmobile use for the population of visitors to Yellowstone National Park? Follow the four-step process.

26. Angry people and heart disease People who get angry easily tend to have more heart disease. That's the conclusion of a study that followed a random sample of 12,986 people from three locations for about four years. All subjects were free of heart disease at the beginning of the study. The subjects took the Spielberger Trait Anger Scale test, which measures how prone a person is to sudden anger. Here are data for the 8474 people in the sample who had normal blood pressure. CHD stands for "coronary heart disease."

This includes people who had heart attacks and those who needed medical treatment for heart disease.

	Low anger	Moderate anger	High anger	Total
CHD	53	110	27	190
No CHD	3057	4621	606	8284
Total	3110	4731	633	8474

Do these data support the study's conclusion about the relationship between anger and heart disease?  
 Follow the four-step process.

**Multiple choice: Select the best answer.**

Exercises 27 to 32 refer to the following setting. The National Survey of Adolescent Health interviewed several thousand teens (grades 7 to 12). One question asked was "What do you think are the chances you will be married in the next ten years?" Here is a two-way table of the responses by gender:<sup>18</sup>

	Female	Male
Almost no chance	119	103
Some chance, but probably not	150	171
A 50-50 chance	447	512
A good chance	735	710
Almost certain	1174	756

27. The percent of females among the respondents was  
 (a) 2625. (c) about 46%. (e) None of these.  
 (b) 4877. (d) about 54%.
28. Your percent from the previous exercise is part of  
 (a) the marginal distribution of females.  
 (b) the marginal distribution of gender.  
 (c) the marginal distribution of opinion about marriage.  
 (d) the conditional distribution of gender among adolescents with a given opinion.  
 (e) the conditional distribution of opinion among adolescents of a given gender.
29. What percent of females thought that they were almost certain to be married in the next ten years?  
 (a) About 16% (c) About 40% (e) About 61%  
 (b) About 24% (d) About 45%
30. Your percent from the previous exercise is part of  
 (a) the marginal distribution of gender.  
 (b) the marginal distribution of opinion about marriage.  
 (c) the conditional distribution of gender among adolescents with a given opinion.

- (d) the conditional distribution of opinion among adolescents of a given gender.  
 (e) the conditional distribution of "Almost certain" among females.
31. What percent of those who thought they were almost certain to be married were female?  
 (a) About 16% (c) About 40% (e) About 61%  
 (b) About 24% (d) About 45%
32. Your percent from the previous exercise is part of  
 (a) the marginal distribution of gender.  
 (b) the marginal distribution of opinion about marriage.  
 (c) the conditional distribution of gender among adolescents with a given opinion.  
 (d) the conditional distribution of opinion among adolescents of a given gender.  
 (e) the conditional distribution of females among those who said "Almost certain."
33. Marginal distributions aren't the whole story Here are the row and column totals for a two-way table with two rows and two columns:

<i>a</i>	<i>b</i>	50
<i>c</i>	<i>d</i>	50
60	40	100

Find two different sets of counts *a*, *b*, *c*, and *d* for the body of the table that give these same totals. This shows that the relationship between two variables cannot be obtained from the two individual distributions of the variables.

34. **Baseball paradox** Most baseball hitters perform differently against right-handed and left-handed pitching. Consider two players, Joe and Moe, both of whom bat right-handed. The table below records their performance against right-handed and left-handed pitchers:

Player	Pitcher	Hits	At-bats
Joe	Right	40	100
	Left	80	400
Moe	Right	120	400
	Left	10	100

- (a) Use these data to make a two-way table of player (Joe or Moe) versus outcome (hit or no hit).

- (b) Show that Simpson's paradox holds: one player has a higher overall batting average, but the other player hits better against both left-handed and right-handed pitching.

- (c) The manager doesn't believe that one player can hit better against both left-handers and right-handers yet have a lower overall batting average. Explain in simple language why this happens to Joe and Moe.

35. **Race and the death penalty** Whether a convicted murderer gets the death penalty seems to be influenced by the race of the victim. Here are data on 326 cases in which the defendant was convicted of murder:<sup>19</sup>

	White Defendant		Black Defendant	
	White victim	Black victim	White victim	Black victim
Death	19	0	Death	11
Not	132	9	Not	52

- (a) Use these data to make a two-way table of defendant's race (white or black) versus death penalty (yes or no).  
 (b) Show that Simpson's paradox holds: a higher percent of white defendants are sentenced to death overall, but for both black and white victims a higher percent of black defendants are sentenced to death.  
 (c) Use the data to explain why the paradox holds in language that a judge could understand..

36. **Fuel economy** (Introduction) Here is a small part of a data set that describes the fuel economy (in miles per gallon) of model year 2009 motor vehicles:

Make and model	Vehicle type	Transmission type	Number of cylinders	City mpg	Highway mpg
Aston Martin Vantage	Two-seater	Manual	8	12	19
Honda Civic	Subcompact	Automatic	4	25	36
Toyota Prius	Midsize	Automatic	4	48	45
Chevrolet Impala	Large	Automatic	6	18	29

- (a) What are the individuals in this data set?  
 (b) What variables were measured? Identify each as categorical or quantitative.

<sup>19</sup>These exercises relate to the optional content on Simpson's paradox.