Chapter 3: Describing Categorical Data

Mix and Match

In each case, unless noted, bar charts are better to emphasize counts whereas pie charts are better to communicate the relative share of the total amount.

1. Proportion of autos: pie chart is the most common; a bar chart or Pareto chart can also be used.

2. Types of defects: Pareto chart (a bar chart with the categories sorted in order of the most common defect)

3. Coupons: bar chart or Pareto chart (these are counts) or perhaps a table (only three values)

4. Type of automobile: bar chart or Pareto chart (counts) or pie chart (shares)

5. Destination: bar chart or Pareto chart (counts) or pie chart (shares)

6. Hanging up: Pareto chart (counts)

7. Excuses: Pareto chart (counts)

8. Brand of computer: bar chart (counts) or pie chart (shares)

9. Software: pie chart (shares) or perhaps a table (only three values)

10. Camera: bar chart (counts), pie chart (shares), or a table (only three values)

11. Ratings: Bar chart or table (only four values). Because the values are ordinal, avoid a pie chart.

12. Loans: Bar chart or table (only three values). Because the data is ordinal, it should not be put into a pie chart – even though the plot shows shares.
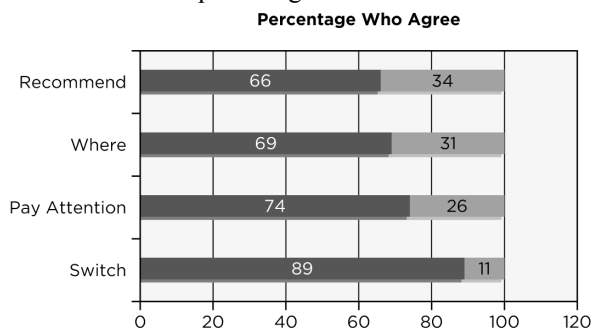
True/False

13. True, but only in general. For variables with few categories, a frequency table is often better, particularly when the analysis requires knowing the detailed frequencies.

14. False. The measure of variation has to be within one category.

15. False. The frequency is the count of the items.

16. False. A relative frequency is a proportion.

Copyright © 2014 Pearson Education, Inc.

17. True. It would be false if the variable were ordinal; you should not put the shares of an ordinal variable into a pie chart.

18. False. The proportion must match the relative frequency.

19. True.

20. False. It has fewer bars.

21. True.

22. False. The median only applies to ordinal variables and identifies the category of the middle value.

Think About It

23. The message is that customers tend to stick with manufacturers from the same region. Someone trading in a domestic car tends to get another domestic car whereas someone who trades in an Asian car tends to buy another Asian car. There's not a lot of switching of loyalties. The more subtle message, one that is disturbing to domestic car makers, is that those who own Asian cars are more loyal (78% buy another Asian car compared to 69% who stick with a domestic car). That makes it hard for domestic manufacturers to win back customers, even if they improve the quality of their cars.

24. The answer is yes. Since lighting makes up 37% of the use of electricity, reducing the demand for electricity by using more efficient bulbs can have a substantial impact. Compact fluorescent blubs produce the same amount of light with much less, say one-quarter, of the electricity used by an incandescent bulbs. Less energy also implies less heat and lower cooling costs. That said, the benefit of these savings for utilities is less pronounced because these savings happen mostly at night, not during the times of peak load that occur during the daytime.

25. This is a bar chart if you think about the underlying data as labeling the dollars held in these countries. The intent of the plot is to show the relative sizes of these counts, comparing the shares of U.S. debt held in these countries.

26. No, this is not a bar chart in the sense of this chapter. The chart uses bars to show a very short time series with two data points, the annual revenue in 2004 and 2005. The three other bars are projections of future revenue (the footnote indicates that the article was published in December of 2006). It's really a time plot.

27. (a) No, these categories are not mutually exclusive. These percentages summarize four dichotomous variables, not one variable.

(b) Divided bars such as these might work well. This style is commonly used in reporting opinion poll results in the news. Sorting the values so that the percentages are in order also makes for a cleaner presentation.

Percentage Who Agree

Category	Agree	Disagree
Recommend	66	34
Where	69	31
Pay Attention	74	26
Switch	89	11

(horizontal axis: 0, 20, 40, 60, 80, 100, 120)

28. (a) No. Each customer could report several of these items, so the categories are not mutually exclusive.
  (b) A figure such as the divided bars used in Exercise 27 would be useful to illustrate the varying shares.

29. No. These percentages only list the percent of executives that report each problem. The categories are not mutually exclusive; some of the executives listed several issues.

30. The percentages do not add to 100; we need an Other category (which has a 9% share of the market).

31. A bar chart would not be appropriate for this situation since the variable, the amount spent by the last 200 customers, is a continuous variable.

32. This grouping data into categories is one step in constructing a frequency table. A histogram should be used since it reports a continuous variable, not a bar chart (a bar chart should be used when you deal with category variable).

33. The bar chart would have one very long bar (height 900) and five shorter bars of height 20 each. The plot would not be very useful, other than to show the predominance of one category.

34. A pie chart would devote 90% of its area to the main category and divide the remaining area into five small slices, each with equal area of 2%.

35. The bar chart would have five bars, each of the same height.

36. The bar chart. It would be hard to tell in the pie chart that the slices were of the same size (however, if the slices were labeled with the percentages it would be the same).

37. A bar chart is preferred because the categorical nature of the variable. A pie chart also can be used to present this data. A frequency table is not appropriate in this situation.

38. With so many categories (the 51 states, including Washington D.C.), some aggregation by region might be useful. Alternatively, it might be good to highlight the most common states, and combine the rest together into a separate, Other category. A bar chart or pie chart could be used, and a frequency table would be fine if there were only a few states represented.

39. The mode is Public. There's no median for this chart since this is nominal data.

40. The East is the modal location. To find the median size, notice there are 50 sizes given, so the median is the size in position 25 or 26. Both lie in the category 10,000 to 19,999. The percentages of enrollment categories, not counts, should be shown in a pie chart.

41. The manufacturers want to know the modal preference because it identifies the most common color preference. Color preferences cannot be ordered, the median color preference can't be defined.

42. A median rating of Excellent implies that at least one-half rated the service as Excellent. A modal rating of Excellent implies that this is the most common rating, but far fewer than half might have picked this rating.

43. The radius of the circle for consumer electronics, for example, would have to be sqrt (10.5/9.1) times the radius of the circle for cell phones, sqrt(10.5/6.3) times the radius for chips, and sqrt(10.5/5.7) times the radius for LCD panels.

44. For instance, render as dollar bills with area determined by the amounts.

You Do It

45. (a) The underlying data table probably accumulates case sales by brand to some degree, such as the number sold by different retail chains and perhaps in different months. All we have here are the aggregated totals.

    (b) To show the shares, a pie chart is most natural because it implies that we are dividing up the total amount. With all the brands shown, however, the small categories make it hard to show the labels. This version accumulates the 3 smallest brands into an "Other" category.

Share of Soft Drink Market
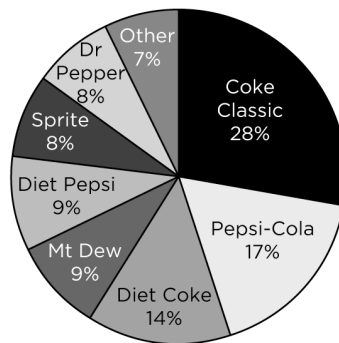
    (c) Here is the pie chart, organized by company rather than brand. If we were artists, then a nicer way to prepare this chart would be to show at the same time the division of Pepsi into brands and so forth. The 38% share of Pepsi comes from Pepsi-Cola, Mountain Dew, Diet Pepsi, and Sierra Mist. These divisions could then be used to slice up Pepsi's share of the chart.
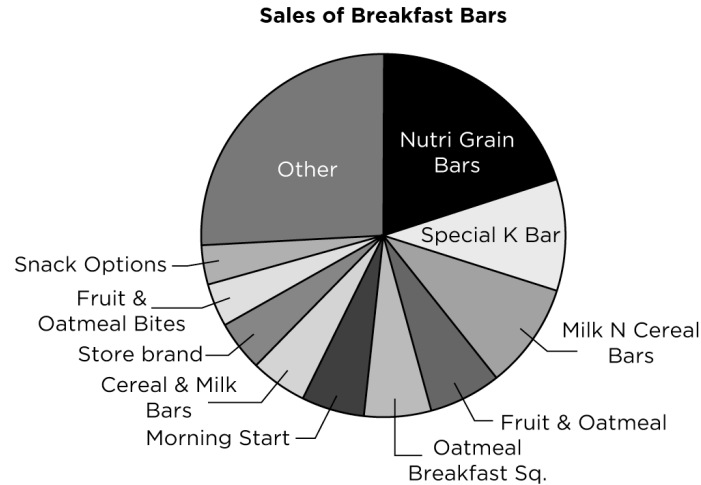
Market Share
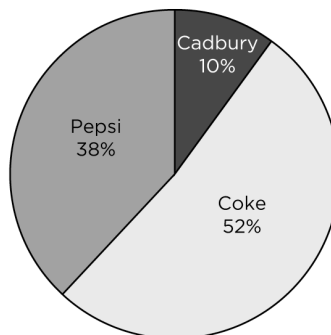
(d) To show the amounts sold, here's a bar chart with the two accumulated types.

Sales of Diet and Regular Soft Drinks
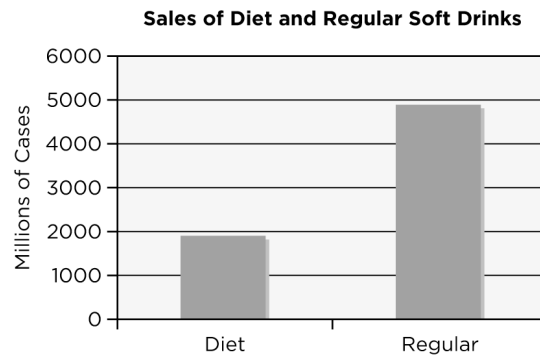
46. (a) The underlying data would be organized by the brand of breakfast bars sold, not by the company.

Sales of Breakfast Bars

(b) The pie chart, including a category labeled Other for the rest of the market, is a bit cluttered by so many smaller brands. The chart makes the shares of the leading brands rather clear. A bar chart might be more useful with so many categories.

Sales of Breakfast Bars
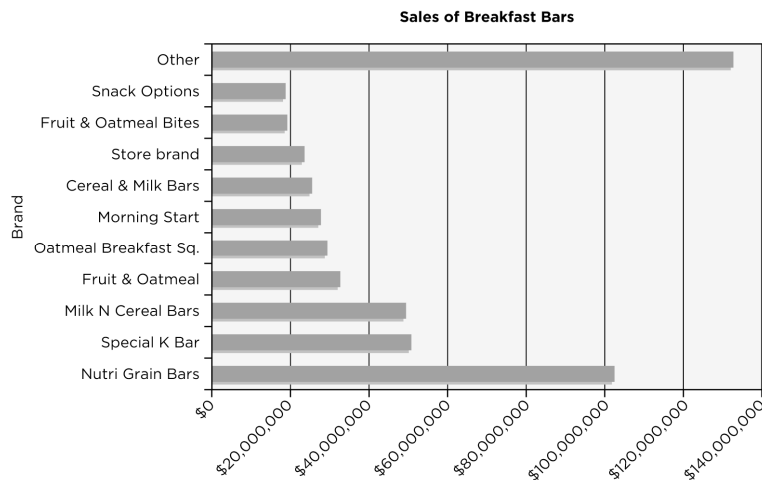
(c) This plot shows the sales of products grouped by the six named manufacturers rather than by brand.

Sales of Top 10 Breakfast Bars
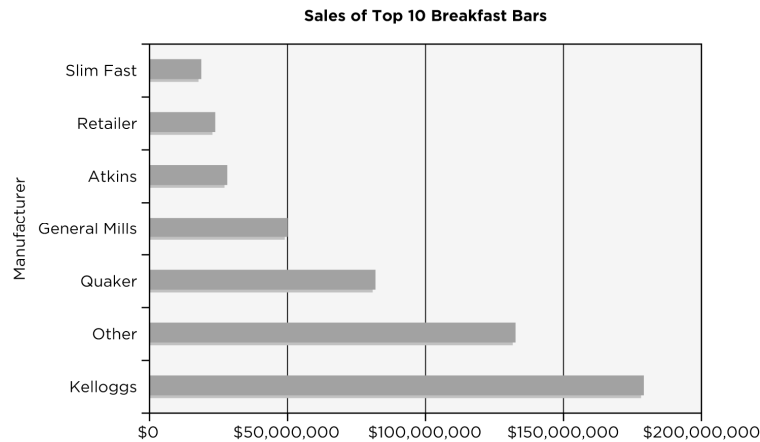
47. (a) The Other category forms an additional row in the tables so that each column adds up to 100%. The addition of this extra row makes up a big part of both pie charts.
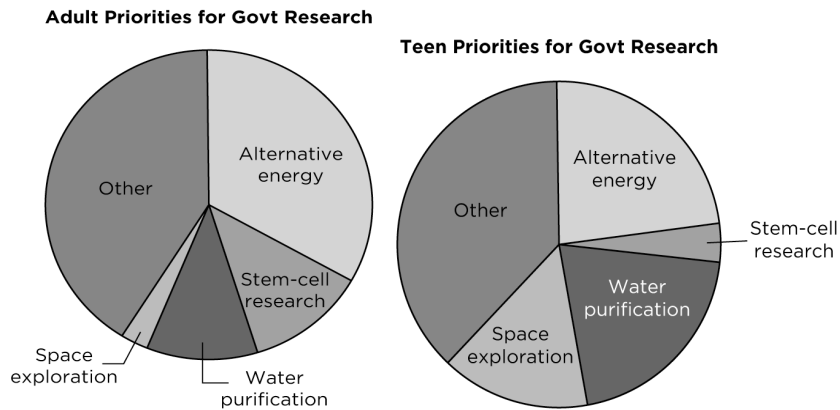

(b) The side-by-side bar chart works well for this. Notice that we no longer need the Other category that dominates the pie charts.

Priorities for Govt Research
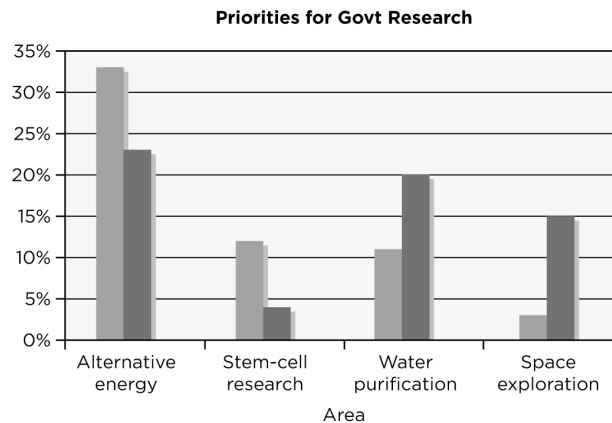
(c) No, because the categories would no longer partition the cases into distinct, non-overlapping subsets. A pie chart should only be used to summarize mutually exclusive groups.

48. (a) The totals within a row do not sum to 100%. The columns give the proportion with different types of media, and these can sum to more than 100% as well. The represented categories do not divide the homes into different groups; a bedroom could have all of these media.

(b) The side-by-side bar chart shows more of every type of media in the rooms of older children.

Media in Bedrooms
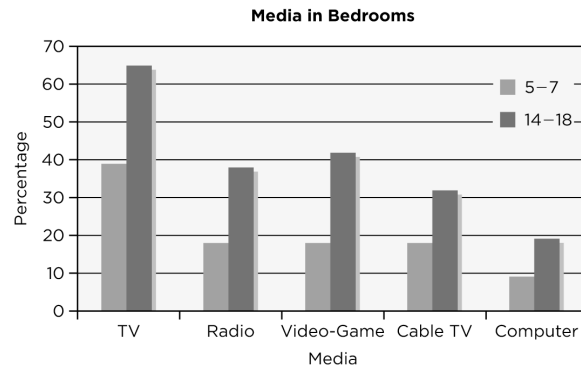
(c) The big adoption of games appears to happen in the 8-13 age range.

49. It's hard to beat this combined bar chart, though some would rather see two pie charts placed side by side. From the bar chart, you can quickly tell the shares each year by the colors, and the adjacent bars show that Splenda gained shares while the others fell.

Share of Sugar Market
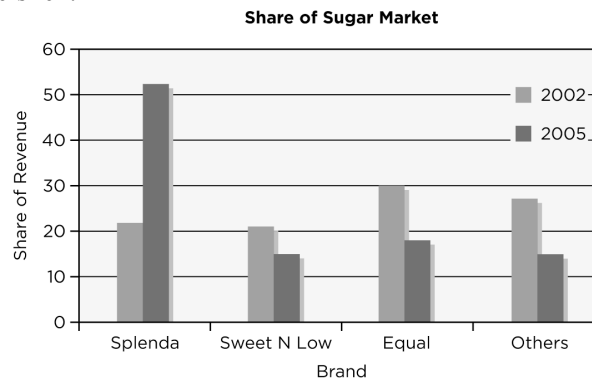
50. (a) Side by side bars don't seem as useful in this example because of the very small shares given to satellite and fixed wireless, categories that were not present in 2003. The bars do make it easy to read off the shares, however. It might be better to omit the smaller categories or lump them as Other in order to focus the attention on the larger, more active categories.
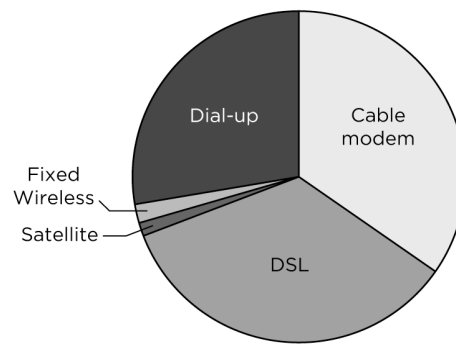
Home Internet Access
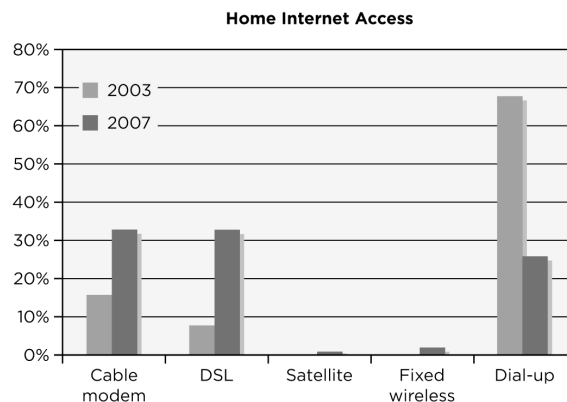
Pie charts put that emphasis more clearly on the larger categories, without leaving so much vacant space in the chart.

Household Internet Access, 2003
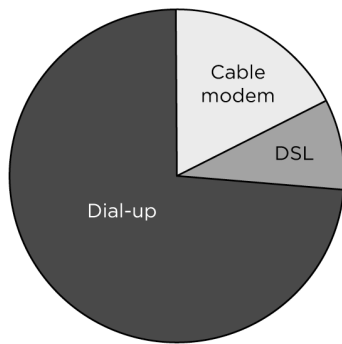
Household Internet Access, 2007

(b) No. These are percentages, not counts. If the number of households with Internet access grew enough during these years, 68% of a small number would be smaller than 26% of a much larger number. That's not the case, but remember that percentages conceal the total counts.

51. (a) Visually, the distributions in the two bar charts seem similar. You should not use one bar chart because the scale required to show all businesses would hide the counts for the businesses owned by women.


Women-Owned Businesses in 1992
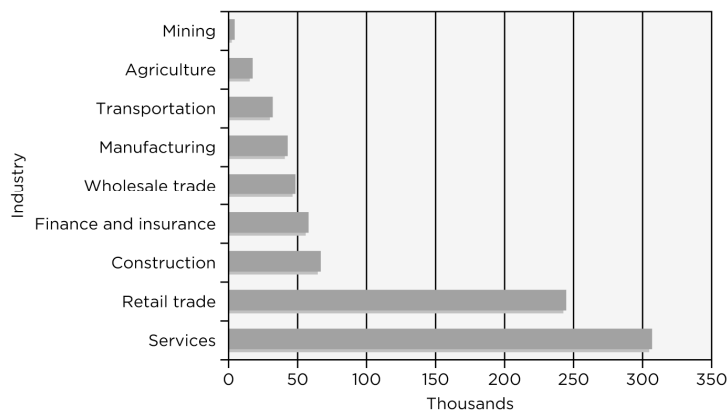
All Businesses in 1995
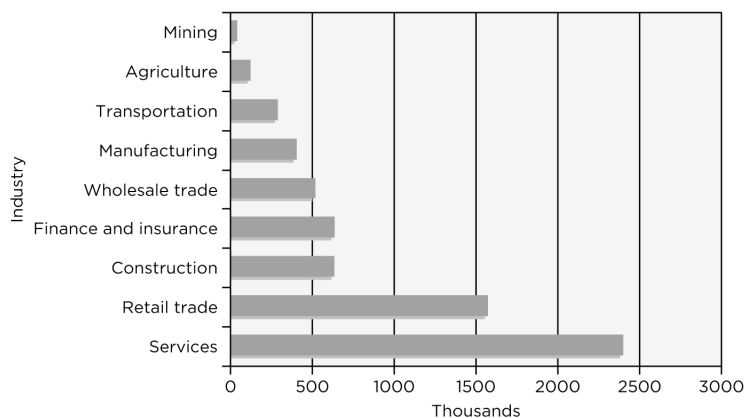
(b) A pie chart shows which industries have the most women-owned businesses, but it does not show the percentage of women-owned businesses *within* each industry. It is necessary to form a column that shows

the proportion of women-owned businesses within each industry. This plot shows the percentages within each industry. While some have more and some less, women own about 10-15% across the board, with none particularly standing out.

Percentage of Women-Owned Businesses
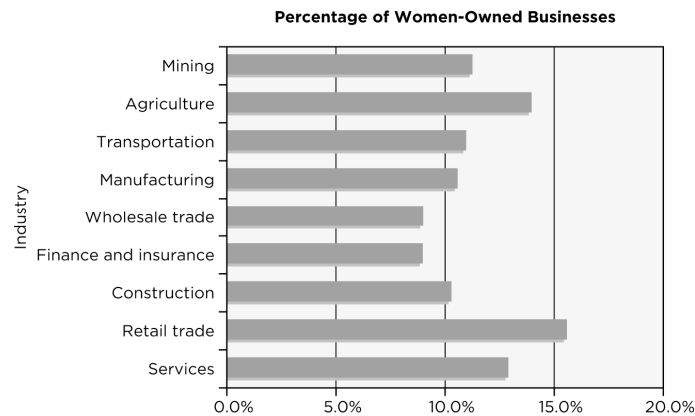
(c) Yes, but slight. The problem is that some industries might have grown or shrunk in number during the three intervening years. The broad nature of these categories suggests that any change was relatively small. If the data tracked more specialized industries, there might have been a larger proportional change over the three years.

52. (a) No, the percentages do not add; some respondents gave more than one reason.
(b) No, unless you think there is a natural way to prioritize the reasons.
(c) A bar chart, with the length given by the percentage.

**53.** (a) Use a table with the two rows and the percentages (or proportions)

Unexpected illness	4,463	15.8%
Planned leave	23,735	84.2%

(b) A Pareto chart shows the categories in order of size.
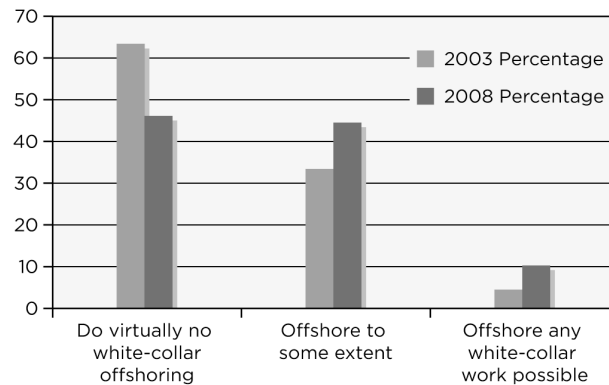
Days Missed from lines
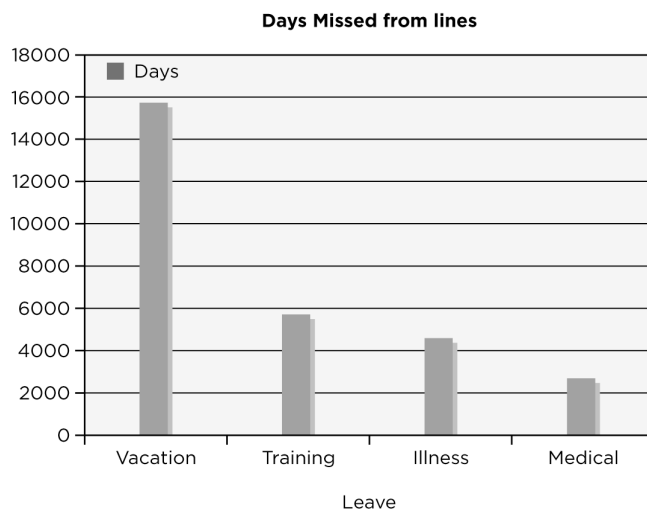
54. (a) Kraft plus Cadbury (15.2%) becomes the mode.
(b) Yes, if the Hershey items are much less expensive than those of other brands.

**55.** (a) Yes, pie charts are fine because the responses are mutually exclusive and sum to 100%.
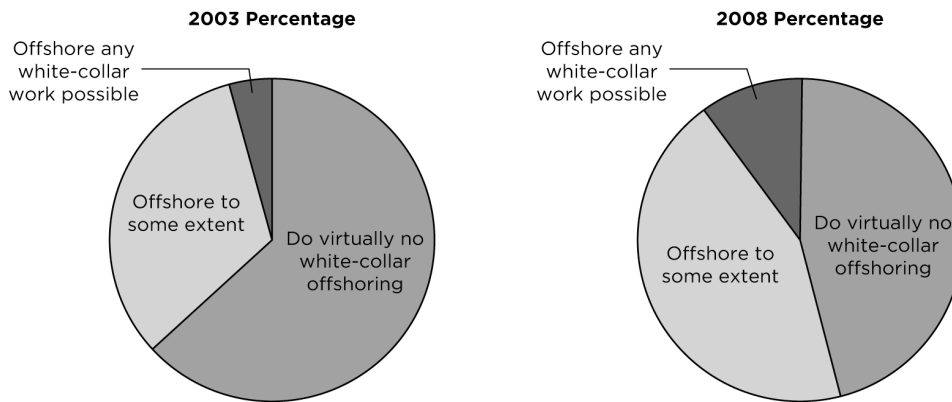
(b) Various answers are possible. The following layout is reasonable.

(c) The bar chart facilitates comparison. The pie chart makes the relative shares more apparent. For example, the 2003 pie shows a predominant share for taking no action, the only choice anticipated to fall in 2008.

(d) The mode and median agree (virtually none) in 2003, but differ in 2008 as responses shift from the more consistent response to a tendency to do more off shoring.

56. (a) The radius for each region should be proportional to the total sales in that region, as suggested in this table:

	Japan	Europe	US
Playstation 3	6.3	19.7	12
Wii	11.5	24.9	30
Xbox 360	1.4	13.7	18.6
	19.2	58.3	60.6
radius	1	1.742543639	1.7765838
diamenter	2	3.485087278	3.553167601
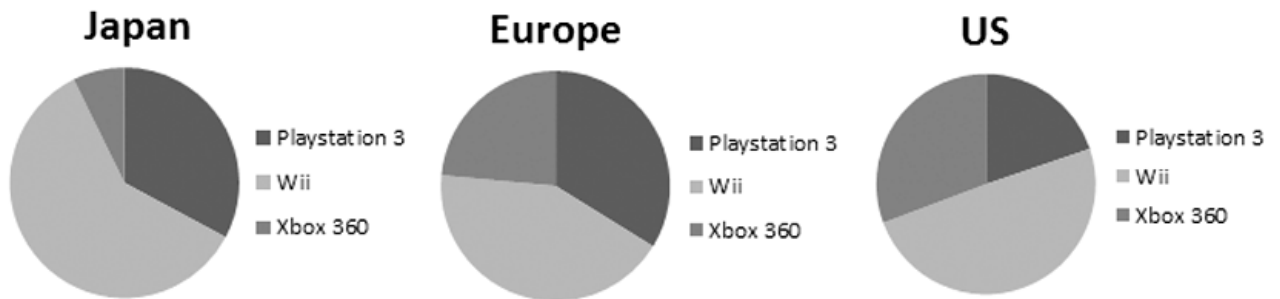
(b)
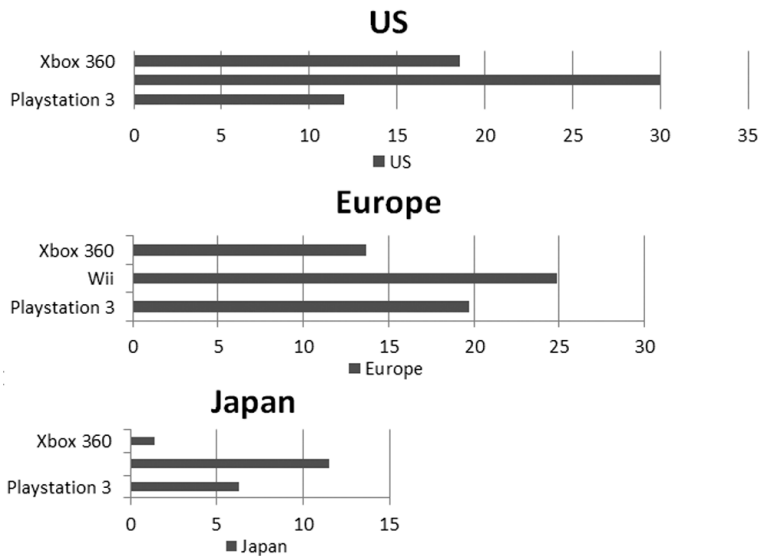
57. (a)

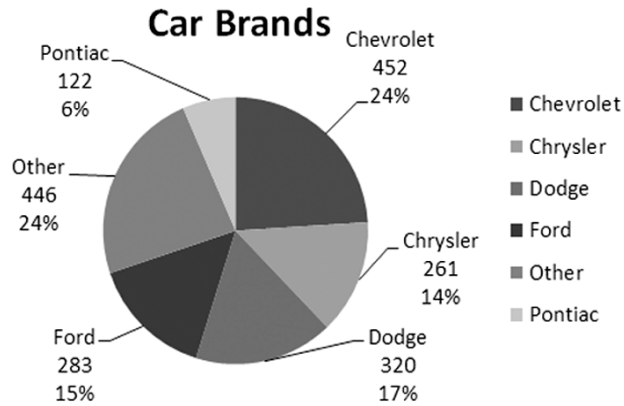|	Frequency	Relative Freq
BUICK	16	0.00849
CHEVROLET	452	0.23992
CHRYSLER	261	0.13854
DODGE	320	0.16985
FORD	283	0.15021

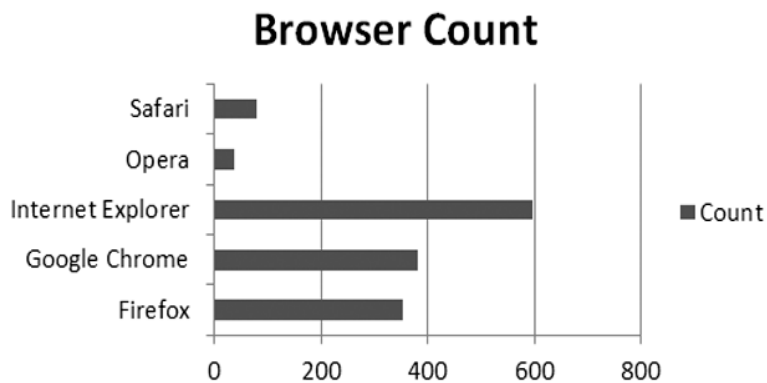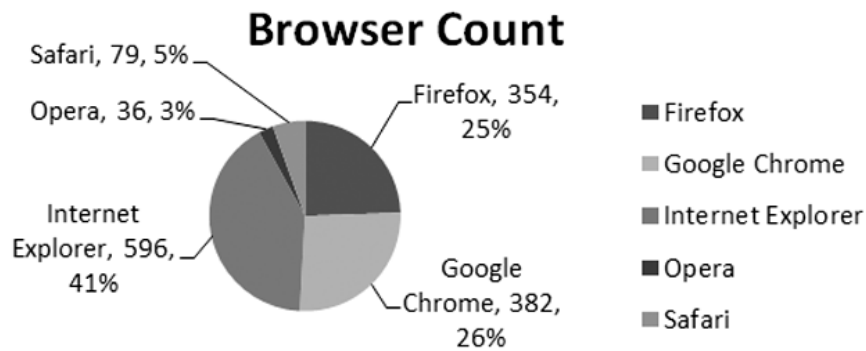(b) Chevrolet.

(c) The pie chart has too many slices.

(d)

Brand	Count
Chevrolet	452
Chrysler	261
Dodge	320
Ford	283
Other	446
Pontiac	122

58. (a) Frequency table

Browser	Count
Firefox	354
Google Chrome	382
Internet Explorer	596
Opera	36
Safari	79

(b) The pie chart is easiest since it shows the percentages.

(c) Internet Explorer is losing share and Google Chrome has grown.

59. 4M Growth Industries

Motivation

(a) It could use trends suggested by the table to indicate how to shift its sales force from declining industries to those that appear stronger and growing.

Method

(b) A bar chart would show the counts and make it simpler to compare the counts within a year. A pie chart would emphasize the shares among industries in the two years.

(c) By looking at the changes from 2000 to 2010, you can see which industries are growing and which are shrinking. You could also use a side-by-side chart, but a chart of the differences does the subtraction for us.

Mechanics

(d) Grouped bar chart, order by size in 2010 to emphasize the change in manufacturing.
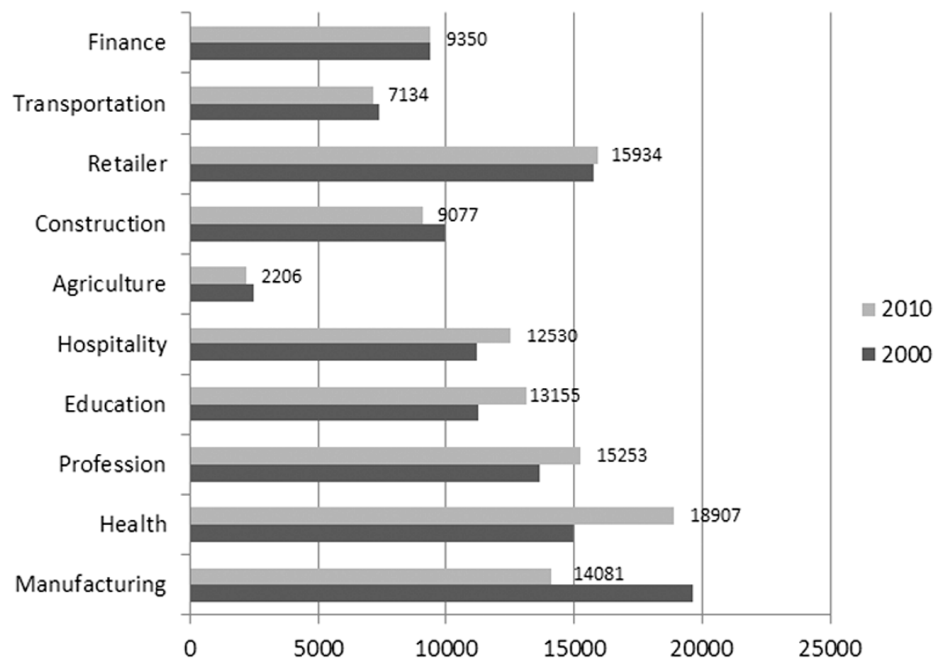
(e) Changes are shown in the differences in adjacent bar heights.

(f) One might choose to show just those that change, but also useful to see those that remained steady. Ten works fine.

Message

(g) One can see the increase in health care and services and the big loss in manufacturing and in construction (mortgage debt crisis).

(h) Percentage shares are not evident, and the plot hides other types of employment.

Statistics For Business Decision Making And Analysis 2nd Edition Stine Solutions Manual

Full Download: https://testbanklive.com/download/statistics-for-business-decision-making-and-analysis-2nd-edition-stine-solution