# Chapter 1
# Data and Statistics

**Learning Objectives**

1. Obtain an appreciation for the breadth of statistical applications in business and economics.
2. Understand the meaning of the terms elements, variables, and observations as they are used in statistics.
3. Obtain an understanding of the difference between qualitative, quantitative, crossectional and time series data.
4. Learn about the sources of data for statistical analysis both internal and external to the firm.
5. Be aware of how errors can arise in data.
6. Know the meaning of descriptive statistics and statistical inference.
7. Be able to distinguish between a population and a sample.
8. Understand the role a sample plays in making statistical inferences about the population.

Chapter 1
**Solutions:**

1. Statistics can be referred to as numerical facts. In a broader sense, statistics is the field of study dealing with the collection, analysis, presentation and interpretation of data.

2. a. Average number of rooms = 808/9 = 89.78 or approximately 90 rooms

   b. Average score = 732.1/9 = 81.3

   c. 2 of 9 are located in England; approximately 22%

   d. 4 of 9 have a room rate of $$; approximately 44%

3. a. 9

   b. 4

   c. Country and room rate are qualitative variables; number of rooms and the overall score are quantitative variables.

   d. Country is nominal; room rate is ordinal; number of rooms is ratio and overall score is interval.

4. a. 5

   b. Price, CD capacity, and the number of tape decks are quantitative. Sound quality and FM tuning sensitivity and selectivity are qualitative.

   c. Average CD capacity = 30/10 = 3.

   d. $\frac{7}{10}(100) = 70\%$

   e. $\frac{4}{10}(100) = 40\%$

5. a. 10

   b. All brands of minisystems manufactured.

   c. Average price = 3140/10 = $314

   d. $314

6. a. 1005

   b. Qualitative
   c. Percentages

   d. .29(1005) = 291.45 or approximately 291.

7. a. The population is all visitors coming to the state of Hawaii.

   b. Since airline flights carry the vast majority of visitors to the state, the use of questionnaires for passengers during incoming flights is a good way to reach this population. The questionnaire actually appears on the back of a mandatory plants and animals declaration form that passengers

must complete during the incoming flight. A large percentage of passengers complete the visitor information questionnaire.

c. Questions 1 and 4 provide quantitative data indicating the number of visits and the number of days in Hawaii. Questions 2 and 3 provide qualitative data indicating the categories of reason for the trip and where the visitor plans to stay.

8. a. Qualitative

b. 30 of 71; 42.3%

9. a. Quantitative; ratio

b. Qualitative; ordinal

c. Qualitative; ordinal (assuming employees can be ranked by classification)

d. Quantitative; ratio

e. Qualitative; nominal

10. a. Although the data are recorded as numbers, the numbers are codes for the ratings of Fair (1), Average (2), Good (3) and Excellent (4). Thus the variables are qualitative with each data value corresponding to a rating category for the variable.

b. The data may also be ranked in order of the quality. A higher number indicates a higher rating on a scale from Fair (1) to Excellent (4). Since the data can be ranked or ordered, the scale of measurement is ordinal.

11. a. Quantitative; ratio scale of measurement

b. Qualitative; nominal scale of measurement

c. Qualitative; ordinal scale of measurement since the responses can be ordered from earliest (high school) to latest (retirement)

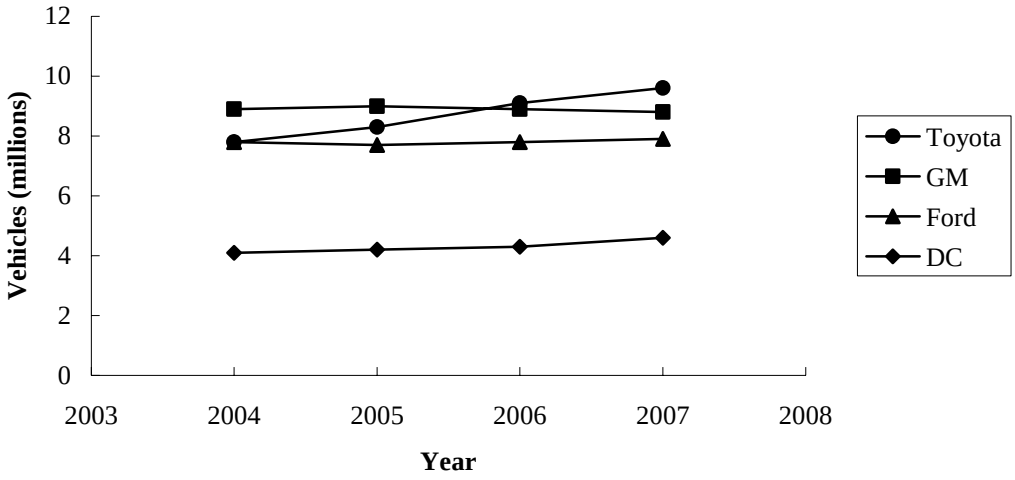
d. Quantitative; ratio scale of measurement

e. Qualitative; nominal scale of measurement

12. Questions a, c, and d provide quantitative data.

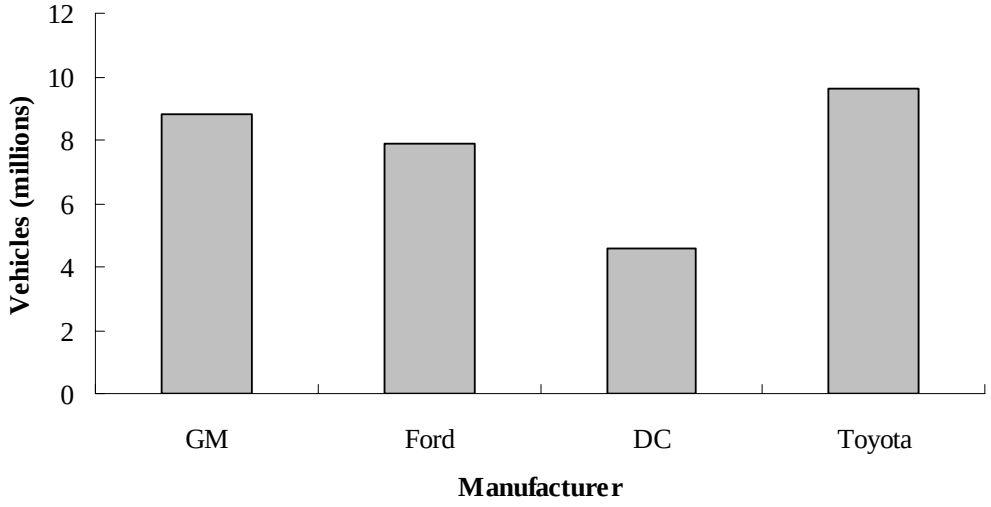
Questions b and e provide qualitative data.

13. a.




b. According to the CSM data, Toyota surpasses General Motors as the biggest auto manufacturer in the world in 2006. In 2006, Toyota manufactured approximately (9.1 – 8.9) = .2 million or 200,000 more vehicles than General Motors. The gap is expected to widen to 800,000 vehicles in 2007. General Motors is the only manufacturer showing a decline in vehicle production over the four year period.

c. The following is a bar graph of cross-sectional data as it shows the number of vehicles manufactured in 2007.




14. a. Earning in $ billions are quantitative data

b. Time series for the years 1997 to 2005

c. Earnings for Volkswagen

d. Earnings were relatively low from 1997 to 1999, but showed excellent growth in 2000 and 2001. 2002 was a relatively strong year, but declines in 2003 to 2005 brought earnings back in line with the 1997 to 1999 levels. The decline in earnings trend from 2002 to 2005 suggests the $600 million projected earnings for 2006 is reasonable.

e. In July 2001, the earnings tread was positive. Assuming the trend in earnings would continue, Volkswagen would have been a promising investment in 2001.

f. Be careful when projecting time series data into the future. Trends in past data may or may not continue into the future. Assuming that past trends will continue can end up fooling an investor.

15. a. We would like to see data from product taste tests and test marketing the product.

b. Such data would be obtained from specially designed statistical studies.

16. a. Quantitative – number of new drugs approved

b. Time series from 1996 to 2003

c. 18

d. 2002; 16 new drugs

e. Over the eight-year period, the number of new drugs approved by the FDA declined. From approximately 50 new drugs approved in 1996, the most recent years are showing only 16 to 18 new drugs approved.

17. a. This is a statistically correct descriptive statistic for the sample.

b. An incorrect generalization since the data was not collected for the entire population.

c. An acceptable statistical inference based on the use of the word “estimate.”

d. While this statement is true for the sample, it is not a justifiable conclusion for the entire population.

e. This statement is not statistically supportable. While it is true for the particular sample observed, it is entirely possible and even very likely that at least some students will be outside the 65 to 90 range of grades.

18. a. 43% of managers were bullish or very bullish.

21% of managers expected health care to be the leading industry over the next 12 months.

b. We estimate the average 12-month return estimate for the population of investment managers to be 11.2%.

c. We estimate the average over the population of investment managers to be 2.5 years.

19. a. Nielsen is attempting to measure the popularity of each television program by showing the percentage of households that are watching the program.

b. All households with televisions in the United States.

c. A census of the population is impossible. A sample provides timely information in that the ratings and share data can be obtained weekly. In addition, the sample saves data collection costs.

d. The cancellation or renewal of television programs, advertising cost rates for the television programs and the scheduling of television programs are often based on the Nielsen information.

20. a. All subscribers of Business Week in North America at the time the survey was conducted.

b. Quantitative

c. Qualitative (yes or no)

d. Crossectional - all the data relate to the same time.

e. Using the sample results, we could infer or estimate 59% of the population of subscribers have an annual income of $75,000 or more and 50% of the population of subscribers have an American Express credit card.

21. a. All registered voters in the state of California

b. Registered voters contacted during the Policy Institute of California survey.

c. A sample was used both to save time and money. The Policy Institute wanted to publish a current estimate of voter support. Contacting all registered voters, even if possible, would have taken so long it is doubtful that the Institute could have obtained the results prior to the election. Also the sample saved money compared to the cost of contacting the entire population of voters.

22. a. $\frac{155}{430} = .36$ or 36%

b. 44% of 430 = .44(430) = 189 business travelers

c. Qualitative data with categories online travel site, travel agent, direct with airline/hotel, other.

23. a. The two populations are the population of women whose mothers took the drug DES during pregnancy and the population of women whose mothers did not take the drug DES during pregnancy.

b. It was a survey.

c. 63 / 3.980 = 15.8 women out of each 1000 developed tissue abnormalities.

d. The article reported “twice” as many abnormalities in the women whose mothers had taken DES during pregnancy. Thus, a rough estimate would be 15.8/2 = 7.9 abnormalities per 1000 women whose mothers had *not* taken DES during pregnancy.

e. In many situations, disease occurrences are rare and affect only a small portion of the population. Large samples are needed to collect data on a reasonable number of cases where the disease exists.

24. Internal data on salaries of other employees can be obtained from the personnel department. External data might be obtained from the Department of Labor or industry associations.

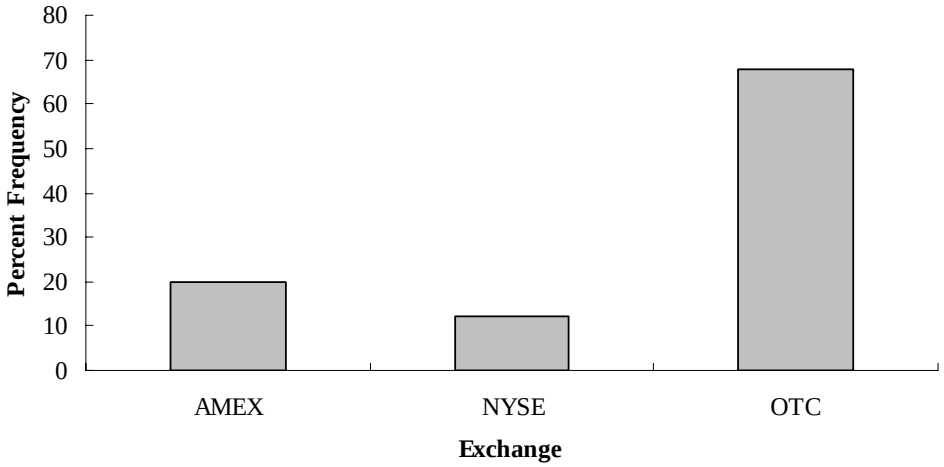
25. a. There are five variables: Exchange, Ticker Symbol, Market Cap, Price/Earnings Ratio and Gross Profit Margin.

b. Qualitative variables: Exchange and Ticker Symbol

Quantitative variables: Market Cap, Price/Earnings Ratio, Gross Profit Margin

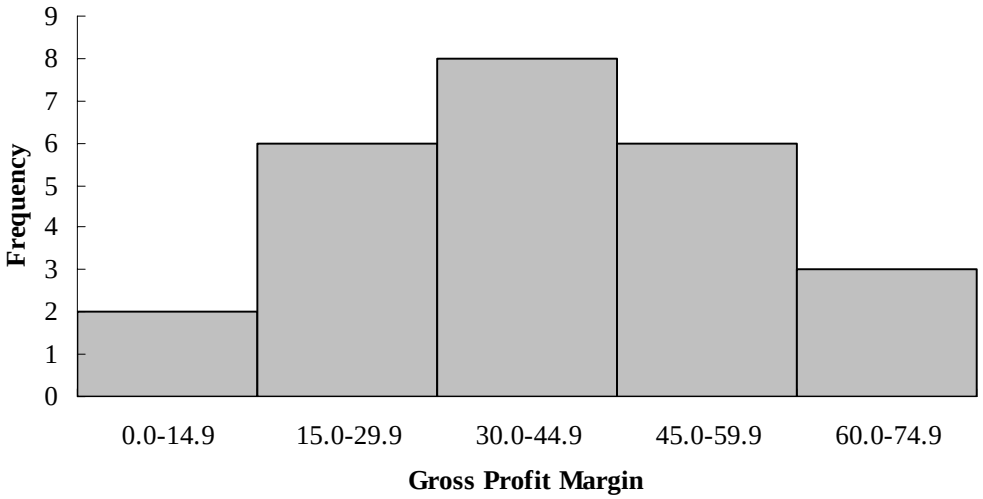
c. Exchange variable:

| Exchange | Frequency | Percent Frequency |
|---|---|---|
| AMEX | 5 | (5/25) 20% |
| NYSE | 3 | (3/25) 12% |
| OTC | 17 | (17/25) 68% |
| | 25 | |




d. Gross Profit Margin variable:

| Gross Profit Margin | Frequency |
|---|---|
| 0.0 – 14.9 | 2 |
| 15.0 – 29.9 | 6 |
| 30.0 – 44.9 | 8 |
| 45.0 – 59.9 | 6 |
| 60.0 – 74.9 | 3 |




e. Sum the Price/Earnings Ratio data for all 25 companies.

Sum = 505.4
Average Price/Earnings Ratio = Sum/25 = 505.4/25 = 20.2