**Chapter 1**

**Introduction to Statistics**

**LEARNING OBJECTIVES**

The primary objective of Chapter 1 is to introduce you to the world of

statistics, thereby enabling you to:

1.List quantitative and graphical examples of statistics within a business

context

2.Define important statistical terms, including population, sample, and

parameter, as they relate to descriptive and inferential statistics

3.Explain the difference between variables, measurement, and data.

4. Compare the four different levels of data: nominal, ordinal, interval, and ratio

5. Define important business analytics terms including big data, business analytics,

data mining, and data visualization.

 6. List the four dimensions of big data and explain the differences between them.

 7. Compare and contrast the three categories of business analytics.

# CHAPTER OUTLINE

1.1 Basic Statistical Concepts

1.2 Data Measurement

 Nominal Level

 Ordinal Level

 Interval Level

 Ratio Level

 Comparison of the Four Levels of Data

 1.3 Introduction to Business Analytics

 Big Data

 Business Analytics

 Categories of Business Analytics

 Descriptive Analytics

 Predictive Analytics

 Data Mining

 Data Visualization

 Statistical Analysis Using the Computer: Excel, Minitab, and Tableau

**KEY TERMS**

 Big Data Ordinal Level Data

 Business Analytics Parameter

Census Parametric Statistics

 Data Population

Data Mining Predictive Analytics

Data Visualization Prescriptive Analytics

Descriptive Analysis Ratio Level Data

Descriptive Statistics Sample

 Inferential Statistics Statistic

 Interval Level Data Statistics

 Measurement Variable

Metric Data Variety

 Nominal Level Data Velocity

 Nonmetric Data Veracity

 Nonparametric Statistics Volume

**STUDY QUESTIONS**

 1. A science dealing with the collection, analysis, interpretation, and presentation of numerical data is called \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

 2. One way to subdivide the field of statistics is into the two branches of \_\_\_\_\_\_\_\_\_\_\_\_\_\_ statistics and \_\_\_\_\_\_\_\_\_\_\_\_\_ statistics.

 3. A collection of persons, objects or items of interest is a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

 4. Data gathered from a whole population is called a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

 5. If a population consists of all the radios produced today in the Akron facility and if a

 quality control inspector randomly selects forty of the ratios, the group of forty is referred

 to as \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

 6. If data are used to reach conclusions only about the group from which the data are

 gathered, then the statistics are referred to as \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ statistics.

 7. If data are gathered from a subgroup of a larger group and the data are used to reach

 conclusions about the larger group, then the statistics are said to be \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

 statistics.

 8. Another name for inferential statistics is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ statistics.

 9. Descriptive measures which are usually denoted by Greek letters are called

 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

 10. A characteristic of any entity being studied that is capable of taking on different values is

 a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

 11. When a standard process is used to assign numbers to particular attributes or characteristics

 of a variable, it is called a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

 12. Recorded measurements are \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

 13. The highest level of data measurement is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

 14. The level of data measurement used when ranking items is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

 15. If a number represents the geographic location of a business, then the level of data

 represented by the number is probably \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

 16. If the data being gathered are only ordinal level data, then the researcher should only use

 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ statistics to analyze the data.

 For each of the following (17-28), the data gathered are most likely to be which level of data?

 Nominal, Ordinal, Interval, or Ratio?

 17. The ages of managers of fast-food restaurants.

 18. An employee's identification number.

 19. The number of freight cars per train for five hundred trains.

 20. The elevation of a town.

 21. The number of feet it takes a car to stop going fifty miles per hour.

 22. The number of ounces of orange juice consumed by each Floridian in the morning.

 23. The volume of wheat in each silo in Nebraska in August.

 24. A rating scale of the productivity of each worker which has as its adjectives: very poor,

 poor, average, good, outstanding.

 25. A person's religious preference.

 26. Weights of statistics' textbooks.

 27. Years of experience on the job.

 28. Number representing a worker's assignment to the red team, blue team, or green team at

 work where the red team is considered the top workers, the green team is considered the

 least productive workers, and the blue team is for workers in the middle.

29. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ is the application of processes and techniques that

 transforms raw data into meaningful information to improve decision-making.

30. A collection of large and complex datasets from different sources that are difficult to

 process using traditional data management and processing applications is referred to

 as \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

31. The collecting, exploring, and analyzing of large volumes of data in an effort to

uncover hidden patterns and/or relationships that can be used to enhance business decision-making is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

 32. The study of the visual representation of data intended to convey data or information

 by imparting it as visual objects displayed in graphics is called

 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

**ANSWERS TO STUDY QUESTIONS**

 1. Statistics 17. Ratio

 2. Descriptive, Inferential 18. Nominal

 3. Population 19. Ratio

 4. Census 20. Interval

 5. Sample 21. Ratio

 6. Descriptive 22. Ratio

 7. Inferential 23. Ratio

 8. Inductive 24. Ordinal

 9. Parameters 25. Nominal

 10. Variable 26. Ratio

 11. Measurement 27. Ratio

 12. Data 28. Ordinal

 13. Ratio 29. Business Analytics

 14. Ordinal 30. Big Data

 15. Nominal 31. Data Mining

 16. Nonparametric 32. Data Visualization

**SOLUTIONS TO THE ODD-NUMBERED PROBLEMS IN CHAPTER 1**

1.1 Examples of data in functional areas:

accounting - cost of goods, salary expense, depreciation, utility costs, taxes, equipment inventory, etc.

finance - World bank bond rates, number of failed savings and loans, measured risk of common stocks, stock dividends, foreign exchange rate, liquidity rates for a single-family, etc.

human resources - salaries, size of engineering staff, years experience, age of employees, years of education, etc.

 marketing - number of units sold, dollar sales volume, forecast sales, size of sales force, market share, measurement of consumer motivation, measurement of consumer frustration, measurement of brand preference, attitude measurement, measurement of consumer risk, etc.

information systems - CPU time, size of memory, number of work stations, storage capacity, percent of professionals who are connected to a computer network, dollar assets of company computing, number of “hits” on the Internet, time spent on the Internet per day, percentage of people who use the Internet, retail dollars spent in e-commerce, etc.

production - number of production runs per day, weight of a product; assembly time, number of defects per run, temperature in the plant, amount of inventory, turnaround time, etc.

management - measurement of union participation, measurement of employer support, measurement of tendency to control, number of subordinates reporting to a manager, measurement of leadership style, etc.

1.3 Descriptive statistics in recorded music industry -

 1) RCA total sales of compact discs this week, number of artists under contract to a company at a given time.

 2) total dollars spent on advertising last month to promote an album.

 3) number of units produced in a day.

 4) number of retail outlets selling the company's products.

Inferential statistics in recorded music industry -

 1) measure the amount spent per month on recorded music for a few consumers then use that figure to infer the amount for the population.

 2) determination of market share for rap music by randomly selecting a sample of 500 purchasers of recorded music.

 3) Determination of top ten single records by sampling the number of requests at a few radio stations.

 4) Estimation of the average length of a single recording by taking a sample of records and measuring them.

The difference between descriptive and inferential statistics lies mainly in the usage of the data. These descriptive examples all gather data from every item in the population about which the description is being made. For example, RCA measures the sales on all its compact discs for a week and reports the total.

In each of the inferential statistics examples, a sample of the population is taken and the population value is estimated or inferred from the sample. For example, it may be practically impossible to determine the proportion of buyers who prefer rap music. However, a random sample of buyers can be contacted and interviewed for music preference. The results can be inferred to population market share.

 1.5 1) Size of sale ($) per customer in men’s formal wear. Either by taking a sample or using a census, management could compute the average sale in men’s formal wear of a weekly period and compare the number to the same average taken a year ago or a month ago to determine if more is being sold per customer. Other variables might include number of sales per hour, number of people entering the department per day, number of dress shirts sold per day, etc.

 2) Number of employees working per day. This variable could indicate the day of the week (certain days have more or less sales), sales activity (how sales are doing overall), or even health of associates. Other variables might include percent of employees absent due to illness, average number of hours worked per week per employee, number of open positions, etc.

 3) Inventory turnover rate. How fast are items in the store selling? Other variables migh include reorder rate, percent of storage space utilized, number of stockouts per week, etc.

 4) Number of customers that enter the store per hour. This figure will vary by day, time of day, and season. Compare figures on this variable from period to period can give some indication of sales trends which can help drive human resource planning, etc. Other variables might include amount of time spent per customer in the store per visit, distance that customers travel to shop in the store, number of referrals that customers make to other people annually, etc.

 5) Percentage of people paying with cash. Percentage of people using credit cards. These can be used to expedite pay systems, investigate employee theft, calculate surcharges associated with credit cards, etc. Other variables might include average time per checkout, average wait time in pay line, etc.

1.7 a) ratio

 b) ratio

 c) ordinal

 d) nominal

 e) ratio

 f) ratio

 g) nominal

 h) ratio

 1.9 a) The population for this study is the 900 electric contractors who purchased Rathburn wire.

 b) The sample is the randomly chosen group of thirty-five contractors.

 c) The statistic is the average satisfaction score for the sample of thirty-five contractors.

 d) The parameter is the average satisfaction score for all 900 electric contractors in the population.