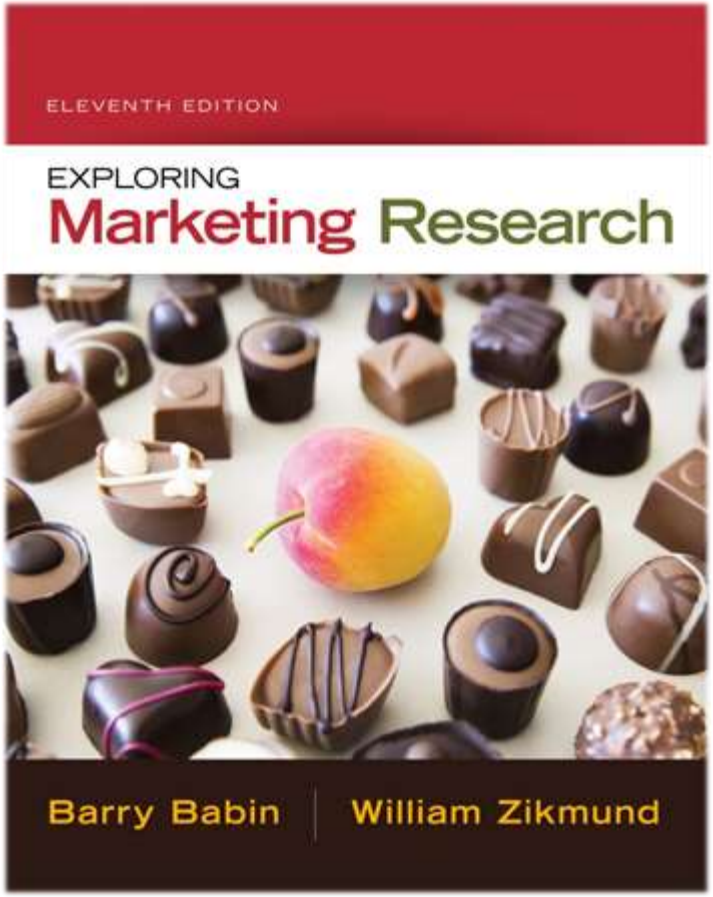


Chapter 14 Basic Data Analysis

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

LEARNING OUTCOMES

After studying this chapter, you should

1. Prepare qualitative data for interpretation or data analysis
2. Know what descriptive statistics are and why they are used
3. Create and interpret tabulation and cross-tabulation tables
4. Perform basic data transformations
5. Understand the basics of testing hypotheses using inferential statistics
6. Be able to use a p-value to make statistical inferences
7. Conduct a univariate *t*-test

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Introduction

- As researchers, we infer whether or not some condition exists in a population based on what we observe in a sample
- Alternatively, the research can be more exploratory and the researcher can use statistics simply to search for some pattern within the data

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Coding Qualitative Responses

- Coding represents the way a specific meaning is assigned to a response within previously edited data
  - Codes represent the meaning in data by assigning some measurement symbol to different categories of responses
  - A code may be a number, letter, or word
  - The proper form of coding relates back to the level of scale measurement
  - Any mistakes in coding can dramatically change the conclusions

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Structured Qualitative Responses and Dummy Variables

- For statistical purposes the research may consider adopting dummy coding for dichotomous responses like yes or no
  - Dummy coding assigns a 0 to one category and a 1 to the other
  - Multiple dummy variables are needed to represent a single qualitative response that can take on more than two categories: if k is the number of categories for a qualitative variable, $k - 1$ dummy variables are needed to represent the variable

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Other Types of Coding

- An alternative to dummy coding is effects coding
  - Effects coding is performed by assigning a + 1 to one value of a dichotomous variable and a –1 to the other
- Class coding is another approach that can be used if the data are not going to be directly used to perform computations
  - Class coding assigns numbers to categories in an arbitrary way merely as a means of identifying some characteristic

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

The Nature of Descriptive Analysis

- Descriptive analysis is the elementary transformation of data to describe basic characteristics such as central tendency, distribution, and variability
  - Averages, medians, modes, variance, range and standard deviation typify descriptive statistics
  - Can summarize responses from large numbers of respondents in a few simple statistics
  - Sample descriptive statistics are used to make inferences about characteristics of the entire population of interest

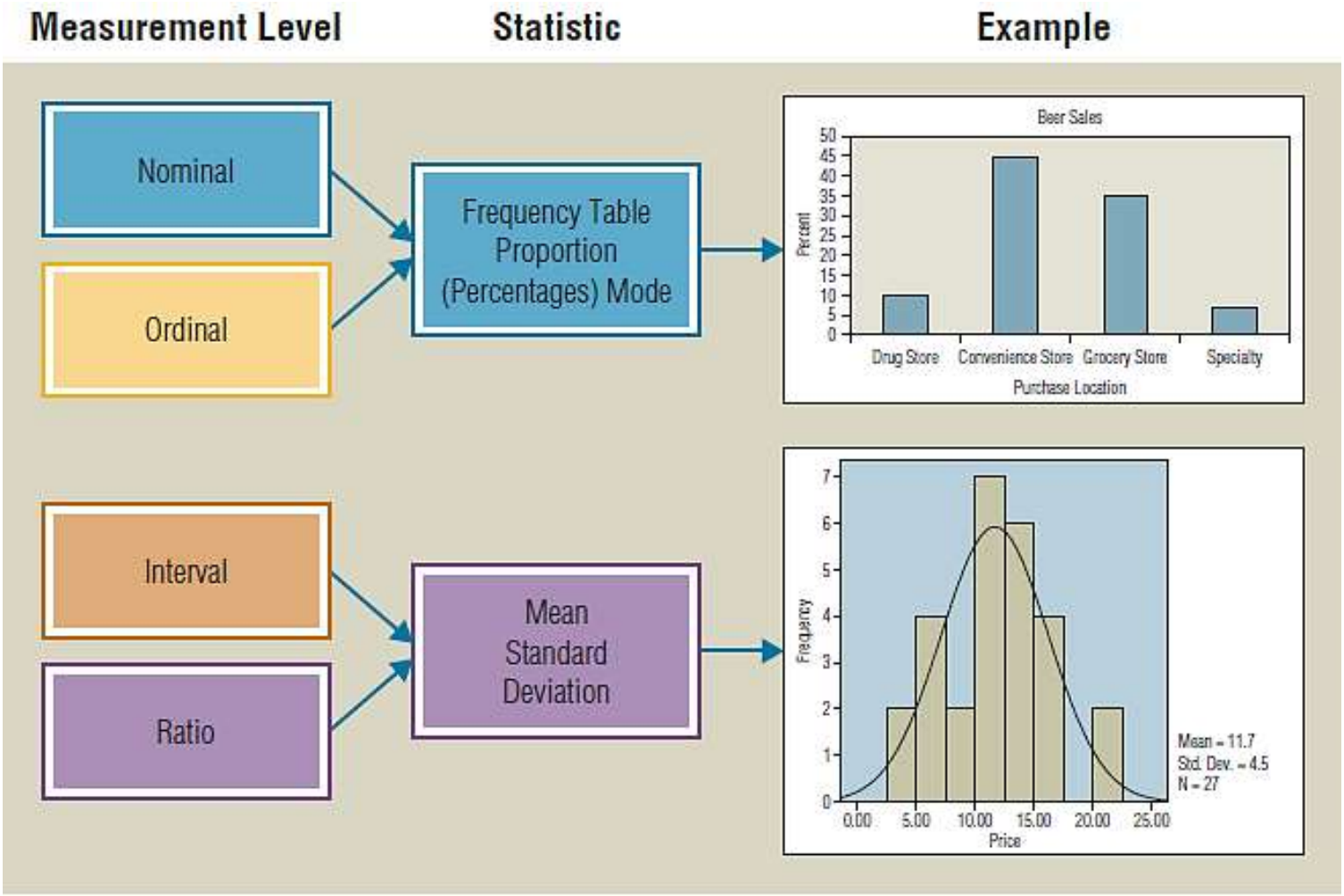
© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

The Nature of Descriptive Analysis (cont'd.)

- Uses univariate statistics
- Simple but powerful and are used very widely
- The level of scale measurement influences the choice of descriptive statistics
- A histogram is a graphical way of showing a frequency distribution in which the height of a bar corresponds to the frequency of a category

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

EXHIBIT 14.1 Levels of Scale Measurement and Suggested Descriptive Statistics



© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Creating and Interpreting Tabulation

- Tabulation refers to the orderly arrangement of data in a table or other summary format
  - Tallying refers to tabulation done by hand
- Counting the different ways respondents answered a question and arranging them in a simple tabular form yields a frequency table
  - The actual number of responses to each category is a variable's frequency distribution
  - A simple tabulation of this type is sometimes called a marginal tabulation

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Creating And Interpreting Tabulation (cont'd.)

- Simple tabulation tells how frequently each response occurs, and this starting point for analysis requires the counting of responses or observations for each category or code
  - Frequency column: shows the tally result or the number of respondents for each category
  - Percent column: shows the total percentage in each category
  - Cumulative percentage: shows the percentage indicating either a particular category or any preceding category

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Cross Tabulation

- Tabulation of data may answer many research questions, and as long as a question deals with only one categorical variable, it is probably the best approach
  - Cross-tabulation is the appropriate technique for addressing research questions involving relationships among multiple less than interval variables
  - One key to interpreting a cross-tabulation table is comparing the observed table values with hypothetical values that would result from pure chance

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

EXHIBIT 14.2 Cross-Tabulation from Consumer Ethics Survey

Generation	Purchase	Download	Total
Echo Boomer	18	35	53
Gen X	41	12	53
Boomer	54	1	55
Mature	53	0	53
	166	48	214

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Contingency Tables

- Data matrices that display the frequency of some combination of possible responses to multiple variables
  - Two-way contingency tables (i.e., involve two less than interval variables) are used most often
  - Beyond three variables, contingency tables become difficult to analyze and explain
  - The row and column totals are often called marginals, because they appear in the table's margin

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Contingency Tables (cont'd.)

- Researchers usually are more interested in the inner cells of a contingency table, which display conditional frequencies (combinations)
- Any cross-tabulation table may be classified according to the number of rows by the number of columns (*R* by *C)*
  - *2* x *2 table*: two variables with two levels each
  - *3* x *4 table*: two variables, one with three levels and the other with four

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

EXHIBIT 14.3 Different Ways of Depicting the Cross-Tabulation of Biological Sex and Target Patronage

(A) Cross-Tabulation of Question "Do you shop at Target?" by Sex of Respondent

	Yes	No	Total
Men	150	75	225
Women	180	45	225
Total	330	120	450

(B) Percentage Cross-Tabulation of Question "Do you shop at Target?" by Sex of Respondent, Row Percentage

	Yes	No	Total (Base)
Men	66.7%	33.3%	100% (225)
Women	80.0%	20.0%	100% (225)

(C) Percentage Cross-Tabulation of Question "Do you shop at Target?" by Sex of Respondent, Column Percentage

	Yes	No
Men	45.5%	62.5%
Women	54.5%	37.5%
Total	100%	100%
(Base)	(330)	(120)

Source: © Cengage Learning 2013

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Percentage Cross-Tabulations

- When data from a survey are cross-tabulated, percentages help the researcher understand the nature of the relationship by making relative comparisons simpler
- The total number of respondents or observations may be used as a statistical base for computing the percentage in each cell

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Percentage Cross-Tabulations (cont'd.)

- One of the questions is commonly chosen as a base for computing percentages
  - Selecting either the row percentages or the column percentages will emphasize a particular comparison or distribution
  - The nature of the problem the researcher wishes to answer will determine which marginal total will serve as a base for computing percentages
  - The margin total of the independent variable should be used as the base for computing the percentages

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Elaboration and Refinement

- Once the basic relationship between two variables has been examined, the researcher may wish to investigate this relationship under a variety of different conditions
  - Typically, a third variable is introduced into the analysis to elaborate and refine the researcher's understanding
    - Specifies the conditions under which the relationship between the first two variables is strongest and weakest
  - Elaboration analysis involves the basic cross-tabulation within various subgroups of the sample

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Elaboration and Refinement (cont'd.)

- The researcher breaks down the analysis for each level of another variable
- Interactions between variables examine moderating variables
  - Moderator variable: a third variable that changes the nature of a relationship between the original independent and dependent variables
- In some cases, adding a third variable to the analysis may lead us to reject the original conclusion about the relationship

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

EXHIBIT 14.4 Cross-Tabulation of Marital Status, Sex, and Responses to the Question "Do You Shop at Target?"

	Single		Married	
	Men	Women	Men	Women
"Do you shop at Target?"				
Yes	55%	80%	86%	80%
No	45%	20%	14%	20%

Source: © Cengage Learning 2013

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

How Many Cross-Tabulations?

- Surveys may ask dozens of questions and hundreds of categorical variables can be stored in a data warehouse
  - Computer-assisted marketing researchers can “fish” for relationships by cross-tabulating every categorical variable with every other categorical variable
  - A researcher addressing an exploratory research question may find some benefit in such a fishing expedition

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

How Many Cross-Tabulations? (cont'd.)

- CHAID (chi-square automatic interaction detection) software makes searches through large numbers of variables possible
  - Data-mining can be conducted in a similar fashion and may suggest relationships that are worth considering further
- Outside of exploratory research, researchers should conduct cross-tabs that address specific research questions or hypotheses
  - When categorical variables are involved, cross-tabs are the right tool

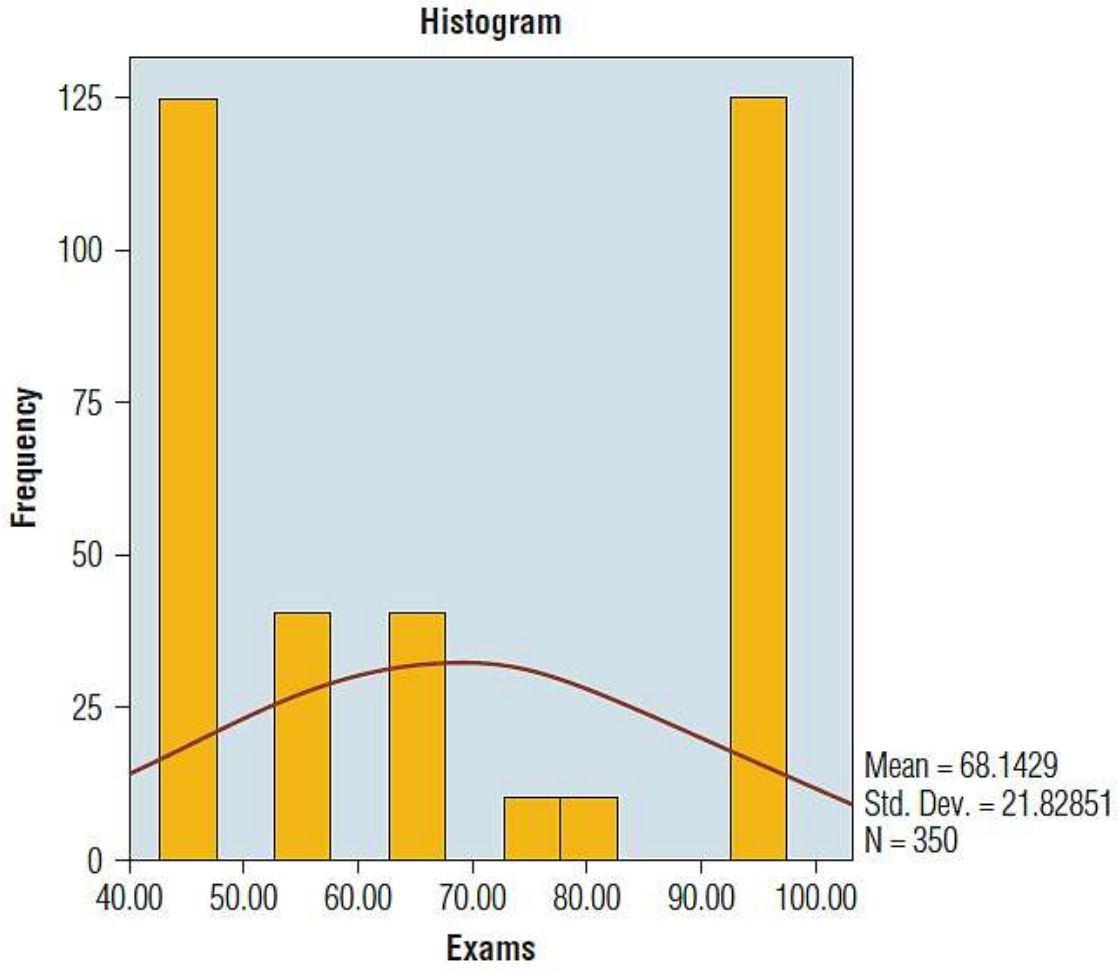
© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Data Transformation

- Simple transformations
  - Data transformation (i.e., data conversion) is the process of changing data from their original form to a format suitable for performing a data analysis that will achieve the research objectives
  - Researchers often recode the raw responses into modified or new variables
  - Collapsing or combining adjacent categories of a variable is a common form of data transformation used to reduce the number of categories

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

EXHIBIT 14.5 Bimodal Distributions Are Consistent with Transformations into Categorical Values



Source: Adapted from 1987 Nielsen Television Report.

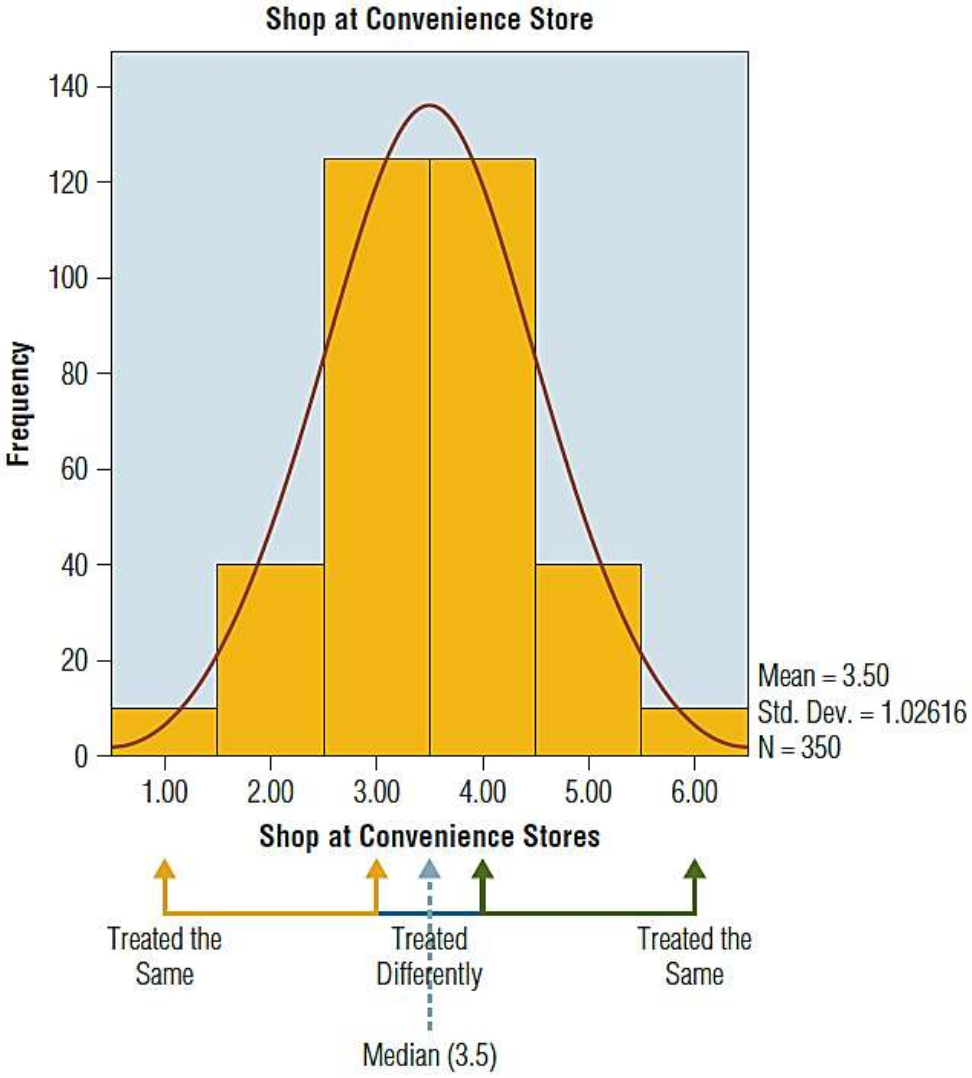
© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Problems with Data Transformations

- Researchers often perform a median split to collapse a scale with multiple response points into two categories
  - The median split means respondents below the observed median go into one category and respondents above the median go into another
  - This approach is best applied only when data exhibit bimodal characteristics
  - When a sufficient number of responses exist and a variable is ratio, the researcher may choose to delete one-fourth to one-third of the responses around the median to effectively ensure a bimodal distribution

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

EXHIBIT 14.6 The Problem with Median Splits with Unimodal Data



Frequency Distribution: X1 = I Do Most of My Shopping at Convenience stores.

Response Category (Code)	Counts	Cumulative Percentage
Strongly Disagree (1)	10	2.86%
Disagree (2)	40	14.29%
Slightly Disagree (3)	125	50.00%
Slightly Agree (4)	125	85.71%
Agree (5)	40	97.14%
Strongly Agree (6)	10	100.00%

Median = 3.5

Recode to Complete Data Transformation:

Old Values	1	2	3	4	5	6
New Values	1	1	1	2	2	2

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Index Numbers

- Represent simple data transformations that allow researchers to track a variable's value over time and compare a variable(s) with other variables
  - Recalibration allows scores or observations to be related to a certain base period or base number
  - If the data are time-related, a base year is chosen
  - Index numbers are computed by dividing each year's activity by the base-year activity and multiplying by 100
  - Require ratio measurement scales

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Tabular and Graphic Methods of Displaying Data

- Tables, graphs and charts can simplify and clarify data
  - Graphical representations of data may take a number of forms, ranging from a computer printout to an elaborate pictograph
  - Today's researcher has many convenient tools to quickly produce charts, graphs, or tables
  - All forms facilitate summarization and communication
  - Bar charts (histograms), pie charts, curve/line diagrams and scatter plots are among the most widely used tools

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Hypothesis Testing Using Basic Statistics

- Empirical testing involves inferential statistics
  - An inference can be made about some population based on observations of a sample representing that population

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Hypothesis Testing Using Basic Statistics (cont'd.)

- Statistical analysis can be divided into several groups based on how many variables are involved
  - Univariate statistical analysis tests hypotheses involving only one variable.
  - Bivariate statistical analysis tests hypotheses involving two variables
  - Multivariate statistical analysis tests hypotheses and models involving multiple (three or more) variables or sets of variables and potentially involving multiple equations

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Hypothesis Testing Procedure

1. The hypothesis is derived from the research objectives and should be stated as specifically as possible be theoretically sound
2. A sample is obtained and the relevant variables are measured
3. The measured value obtained in the sample is compared to the value either stated explicitly or implied in the hypothesis
   - The hypothesis is *supported* or *not supported* by the measured value

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Hypothesis Testing Procedure (cont'd.)

- The exact point where the hypothesis changes from not being supported to being supported depends on how much risk the researcher is willing to accept
  - Univariate hypotheses are typified by tests comparing some observed sample mean against a benchmark value
    - *Is the sample mean truly different from the benchmark?*
  - A statistical test's significance level or p-value becomes a key indicator of whether or not a hypothesis can be supported

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Significance Levels and *p*-Values

- A significance level is a critical probability associated with a statistical hypothesis test
  - Indicates how likely it is that an inference supporting a difference between an observed value and some statistical expectation is true
- *p*-value: probability value, or the observed or computed significance level
  - Low *p*-values mean there is little likelihood that the statistical expectation is true

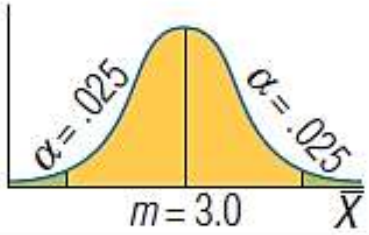
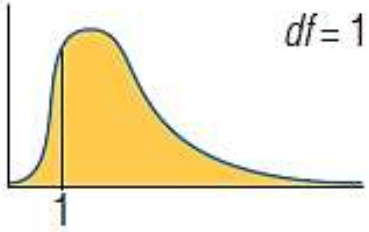
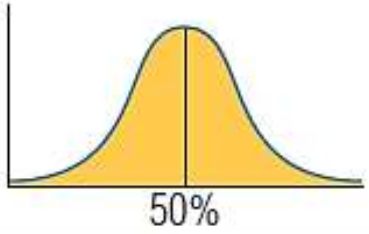
© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Significance Levels and p-Values (cont'd.)

- Traditionally, researchers have specified an acceptable significance level for a test prior to the analysis
  - Most typically, researchers set the acceptable amount of error, and therefore the acceptable significance level, at 0.1, 0.05, or 0.01
  - If the *p*-value resulting from a statistical test is less than the pre-specified significance level, the results support a hypothesis implying differences

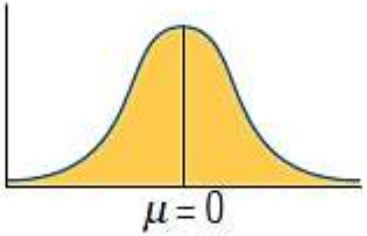
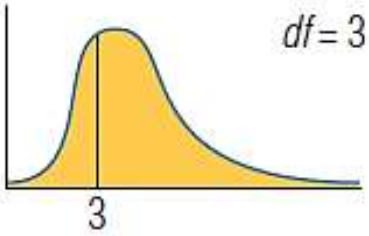
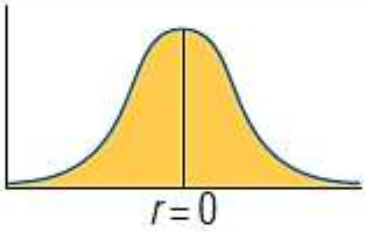
© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

EXHIBIT 14.7 p-Values and Statistical Tests

Test Description	Test Statistic	
Compare an Observed Mean with Some Predetermined Value	*Z* or *t*-test—Low p-Values Indicate the Observed Mean Is Different Than Some Predetermined Value (Often 0)	
Compare an Observed Frequency with a Predetermined Value	X^2—Low p-Values Indicate That Observed Frequency Is Different Than Predetermined Value	
Compare an Observed Proportion with Some Predetermined Value	*Z* or *t*-Test for Proportions—Low p-Values Indicate That the Observed Proportion Is Different Than the Predetermined Value	

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

EXHIBIT 14.7 p-Values and Statistical Tests (cont'd.)

Bivariate Tests:		
Compare Whether Two Observed Means Are Different from One Another	*Z* or *t*-Test—Low p-Values Indicate the Means Are Different	
Compare Whether Two Less-Than Interval Variables Are Related Using Cross-Tabs	X^2—Low p-Values Indicate the Variables Are Related to One Another	
Compare Whether Two Interval or Ratio Variables Are Correlated to One Another	*t*-Test for Correlation—Low p-Values Indicate the Variables Are Related to One Another	

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Type I and Type II Errors

- Because we cannot make any statement about a sample with complete certainty, there is always the chance that an error will be made
  - When a researcher makes the observation using a census, meaning that every unit (person or object) in a population is measured, then conclusions are certain
  - Researchers very rarely use a census; they are susceptible to two types of inferential errors

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Type I Error

- A Type I error occurs when a condition that is true in the population is rejected based on statistical observations
  - When a researcher sets an acceptable significance level a priori (α), he or she is determining how much tolerance he or she has for a Type I error
  - A Type I error occurs when the researcher concludes that there is a statistical difference based on a sample result when in reality one does not exist in the population

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Type II Error

- A Type II error is the probability of failing to reject a false hypothesis
  - This incorrect decision is called beta (β)
  - A sample does not show a difference between an observed mean and a benchmark when in fact the difference does exist in the population
- For correlation type relationships, the sample data suggests that a relationship does not exist when in fact a relationship does exist
  - Such an occurrence is related to statistical power

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Type II Error (cont'd.)

- Unfortunately, without increasing sample size, the researcher cannot simultaneously reduce Type I and Type II errors – as they are inversely related
  - Thus, reducing the probability of a Type II error increases the probability of a Type I error
- In marketing problems, Type I errors generally are considered more serious than Type II errors
  - Thus, more emphasis is placed on determining the significance level, α, than in determining β

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Type II Error (cont'd.)

		Conclusion about null hypothesis from statistical test	
		Accept Null	Reject Null
Truth about null hypothesis in population	True	Correct	Type I error Observe difference when none exists
	False	Type II error Fail to observe difference when one exists	Correct

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Type I & II Error (cont'd.)

Type I error, also known as a "false positive": the **error** of rejecting a null. hypothesis when it is actually true. In other words, this is the **error** of accepting an. alternative hypothesis (the real hypothesis of interest) when the results can be. attributed to chance.

Type II error, also known as a "false negative": the error of not rejecting a null hypothesis when the alternative hypothesis is the true state of nature. In other words, this is the error of failing to accept an alternative hypothesis when you don't have adequate power. Plainly speaking, it occurs when we are failing to observe a difference when in truth there is one.

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Univariate Tests of Means

- A univariate *t*-test is appropriate for testing hypotheses involving some observed mean against some specified value such as a sales target
  - The *t*-distribution, like the standardized normal curve, is a symmetrical, bell-shaped distribution with a mean of 0 and a standard deviation of 1.0
  - When sample size (*n*) is larger than 30, the *t*-distribution and *Z*-distribution are almost identical
  - The shape of the *t*-distribution is influenced by its degrees of freedom (*df*)

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

Univariate Tests of Means (cont'd.)

- The degrees of freedom are determined by the number of distinct calculations that are possible given a set of information
  - In the case of a univariate *t*-test, the degrees of freedom are equal to the sample size (*n*) minus one
  - Today, with computerized software packages, the number of degrees of freedom is provided automatically for most tests

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

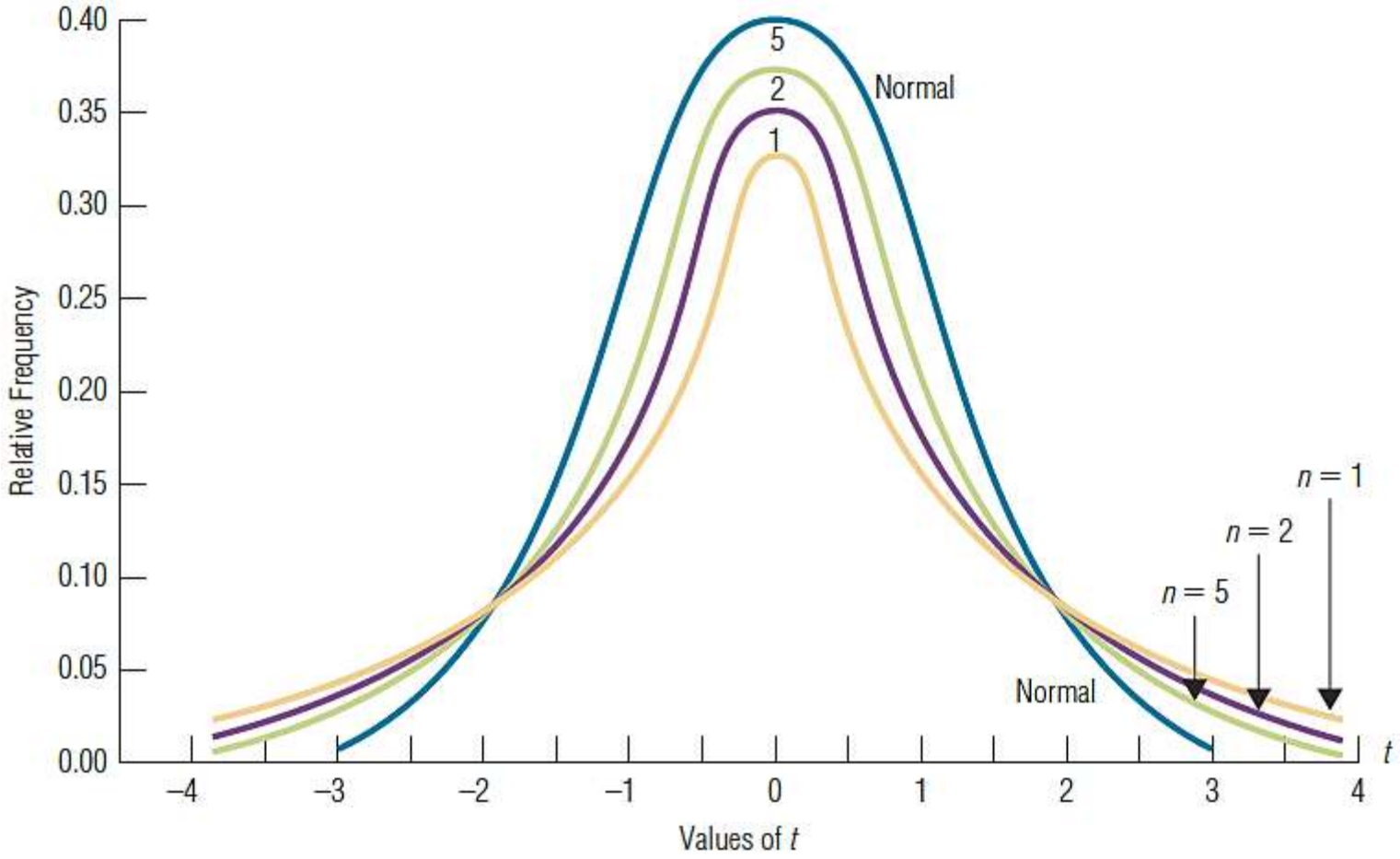
Calculating a *t*-Statistic

$$t = \frac{\overline{X} - \mu}{S_{\overline{X}}}$$

- Calculate sample mean and standard deviation
- The standard error is computed
- Find the *t*-value associated with the desired level of confidence level or statistical significance
  - A 95% confidence level; the significance level is 0.05
- Find the critical values for the *t*-test by locating the upper and lower limits of the confidence interval ► defines the regions of rejection

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

EXHIBIT 14.8



Source: From ZIKMUND/BABIN/CARR/GRIFFIN, *Business Research Methods* (with Qualtrics Card), 8E.
© 2010 Cengage Learning.

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.

The *Z*-Distribution

- The *Z*-distribution and the *t*-distribution are highly similar
  - Provide much the same result in most situations
- Which conditions lend themselves to each?
  - When the population standard deviation (σ) is known, the Z-test is most appropriate
  - When σ is unknown (the situation in most business research studies), and the sample size greater than 30, the *Z*-test also can be used
  - When σ is unknown and the sample size is small, the *t*-test is most appropriate

© 2016 Cengage Learning. All Rights Reserved. May not be scanned, copied or duplicated, in whole or in part, except for use as permitted in a license distributed with a certain product or service or otherwise on a password-protected website for classroom use.