

Guided Tour

Business Statistics in Practice, Fourth Edition, has been written with students' needs in mind. Its clear and understandable explanations and use of real world case studies and examples present content that business students can relate to. Because today's students learn in a visual and interactive way, the text is supplemented by a free student CD-ROM, containing a host of updated resources and helpful study aids. In addition, both students and instructors are provided with additional resources on the text website. Thus, students are given a number of statistical tools in a variety of ways and shown how these tools can be used to positively impact business and other organizations.

Chapter Introductions

Each chapter opens with a preview showing how the statistical topics to be discussed apply to real business problems. The continuing case examples that run throughout the book are briefly introduced along with the techniques that will be used to analyze them.

Visual Statistics 2.0

Visual Statistics, described later in the tour, helps students learn statistics through interactive experimentation and visualization. Concepts in the text that are treated in the Visual Statistics software program are identified by icon, with chapter reference, in the margin of the text next to the concept.

In Chapter 1 we saw that although we can sometimes take a census of an entire population, we often must randomly select a sample from a population. When we have taken a census or a sample, we typically wish to describe the observed data set. In particular, we describe a sample in order to make inferences about the sampled population.

In this chapter we learn about **descriptive statistics**, which is the science of describing the important characteristics of a population or sample. Generally, we look at several important aspects of a set of measurements. One such aspect is the **central tendency**, or middle, of the data set. For instance, we might estimate a typical bottle design rating in the marketing research case or a typical city driving mileage in the car mileage case. Another important aspect of a data set is the **variability**, or spread, of the data. For example, we might estimate the largest and smallest gas mileage that would likely be obtained when a new midsize car is purchased. Or, in the marketing research case we

might measure the spread of the bottle design ratings. If the ratings are clustered closely together, consumers' ratings are much the same (or are consistent). If the ratings are spread far apart, then consumers have widely varying opinions of the new bottle design. A third important aspect of a data set is the **shape** of the population or sample. Looking at a data set's shape tells us how the population or sample is **distributed** over various values (more about this later). Still another important aspect is whether **outliers** exist. For instance, if there are outlying bottle design ratings, then several consumers have opinions about the design that are very different from the opinions of most of the sampled consumers. Descriptive statistics also involves using **graphical methods** to depict data sets and to study relationships between different variables.

In this chapter we use a variety of methods to describe the cell phone usages, bottle design ratings, coffee temperatures, and car mileages introduced in the cases of Chapter 1. In addition, we introduce three new cases:

The Payment Time Case: A management consulting firm assesses how effectively a new electronic billing system reduces bill payment times.

The Electronic Article Surveillance Case: A survey is used to study the unintended effects on consumer

attitudes of false electronic article surveillance alarms.

The Marketing Ethics Case: A survey is conducted to study marketing researchers' attitudes toward violating confidentiality in marketing research studies.



CHAPTER 2 Descriptive Statistics

CHAPTER 2

Chapter Outline

2.1 Describing the Shape of a Distribution	2.6 Using Scatter Plots to Study Relationships between Variables (Optional)
2.2 Describing Central Tendency	2.7 Misleading Graphs and Charts (Optional)
2.3 Measures of Variation	2.8 Weighted Means and Grouped Data (Optional)
2.4 Percentiles, Quartiles, and Box-and-Whiskers Displays	2.9 The Geometric Mean (Optional)
2.5 Describing Qualitative Data	

2.1 Describing the Shape of a Distribution

We begin looking at the characteristics of a population by describing the population's overall pattern of variation. That is, we describe the shape of the distribution of population measurements. We often employ a sample of measurements taken from a population in order to infer what the population looks like.

Several graphical methods—the **stem-and-leaf display**, the **histogram**, and the **dot plot**—are often used to portray shapes of distributions.

Stem-and-leaf displays We illustrate how to construct stem-and-leaf displays in the following examples.

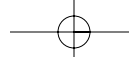
Example 2.1 The Car Mileage Case

Table 2.1 presents the sample of 49 gas mileages that have been obtained by the new midsize model in Example 1.4 (page 15). To graphically portray the pattern of variation in these mileages, we can construct a stem-and-leaf display. In order to do this, we first notice that the sample mileages range from 29.8 to 33.3. For this data we will (somewhat arbitrarily) construct a display having the first two digits of the mileages—the whole numbers 29, 30, 31, 32, and 33—as the **stems**. These are placed in a column on the left side of the display as follows:

- 29
- 30
- 31
- 32
- 33

Case Studies

The text provides a unique use of case studies that span individual chapters and groups of chapters. Cases are used to introduce the concepts, to demonstrate the methods, and to provide students with motivating exercises. These case studies help students see how statistics is used in business and can be used to improve processes.



Student Friendly Presentation

The authors have made improvements throughout the text to make learning easier for students. The following examples highlight some of these improvements.

Step-by-Step Hypothesis Testing Approach

This approach consists of a seven-step procedure designed to break hypothesis testing down into small, easy to understand steps. This procedure is used in almost all the examples in Chapters 8 and 9 and can be applied by students throughout the remainder of the text where hypothesis testing is done.

Greater Accessibility of Continuing Cases

Each time a continuing case is revisited, any needed computer output and, whenever possible, relevant background information is included with the current case discussion. Consequently, students seldom if ever need to refer back to previously covered material in order to grasp the content included in a given case segment.

Business Improvement

Business improvement applications are identified by icons and are highlighted—to indicate when important business improvement conclusions have been reached using statistical analysis.

A 99 percent confidence interval for p is

$$\begin{aligned} \left[\hat{p} \pm z_{.005} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right] &= \left[.063 \pm 2.575 \sqrt{\frac{(.063)(.937)}{1000}} \right] \\ &= [.063 \pm .0198] \\ &= [.0432, .0828] \end{aligned}$$

The upper limits of both the 95 percent and 99 percent intervals are less than .10. Therefore, we have very strong evidence that the true proportion p of all current purchasers who would stop buying the cheese spread is less than .10. Based on this result, it seems reasonable to use the new spout. **BI**

In the cheese spread example, a sample of 1,000 purchasers gives us a 95 percent confidence interval for $p = [.063 \pm .0151]$ —with a reasonably small margin of error of .0151. Generally speaking, quite a large sample is needed in order to make the margin of error in a confidence interval for p reasonably small. The next two examples demonstrate that a sample size of 200, which most people would consider quite large, does not necessarily give a 95 percent confidence interval for p with a small margin of error.

Testing a “greater than” alternative hypothesis by using a rejection point rule In Section 8.1 we explained how to set up appropriate null and alternative hypotheses. We also discussed how to specify a value for α , the probability of a Type I error (also called the **level of significance**) of the hypothesis test, and we introduced the idea of a test statistic. We can use these concepts to begin developing a seven step hypothesis testing procedure. We will introduce these steps in the context of the trash bag case and testing a “greater than” alternative hypothesis.

Step 1: State the null hypothesis H_0 and the alternative hypothesis H_a . In the trash bag case, we will test $H_0: \mu \leq 50$ versus $H_a: \mu > 50$. Here, μ is the mean breaking strength of the new trash bag.

Step 2: Specify the level of significance α . The television network will run the commercial stating that the new trash bag is stronger than the former bag if we can reject $H_0: \mu \leq 50$ in favor of $H_a: \mu > 50$ by setting α equal to .05.

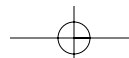
Step 3: Select the test statistic. In order to test $H_0: \mu \leq 50$ versus $H_a: \mu > 50$, we will test the modified null hypothesis $H_0: \mu = 50$ versus $H_a: \mu > 50$. The idea here is that if there is sufficient evidence to reject the hypothesis that μ equals 50 in favor of $\mu > 50$, then there is certainly also sufficient evidence to reject the hypothesis that μ is less than or equal to 50. In order to test $H_0: \mu = 50$ versus $H_a: \mu > 50$, we will randomly select a sample of $n = 40$ new trash bags and calculate the mean \bar{x} of the breaking strengths of these bags. We will then utilize the **test statistic**

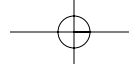
$$z = \frac{\bar{x} - 50}{\sigma_{\bar{x}}} = \frac{\bar{x} - 50}{\sigma / \sqrt{n}}$$

A positive value of this test statistic results from an \bar{x} that is greater than 50 and thus provides evidence against $H_0: \mu = 50$ and in favor of $H_a: \mu > 50$.

Step 4: Determine the rejection point rule for deciding whether to reject H_0 . To decide how large the test statistic must be to reject H_0 in favor of H_a by setting the probability of a Type I error equal to α , we do the following:

- Place the probability of a Type I error, α , in the right-hand tail of the standard normal curve and use the normal table (see Table A.3, page 000) to find the normal point z_α . Here z_α , which we call a **rejection point** (or **critical point**), is the point on the horizontal axis under the standard normal curve that gives a right-hand tail area equal to α .
- **Reject $H_0: \mu = 50$ in favor of $H_a: \mu > 50$ if and only if the test statistic z is greater than the rejection point z_α .** (This is the **rejection point rule**.)





Appendix 11.3 ■ Simple Linear Regression Analysis Using MegaStat

The instructions in this section begin by describing the entry of data into an Excel worksheet. Alternatively, the data may be loaded directly from the data disk included with the text. The appropriate data file name is given at the top of each instruction block. Please refer to Appendix 1.2 for further information about entering data and saving and printing results in Excel. Please refer to Appendix 1.3 for more information about using MegaStat.

Simple linear regression for the service time data in Figure 11.14 on page 476 (data file: SrvTime.xls):

- Enter the service time data (page 000) with the numbers of copiers serviced in column A (with label Copiers) and the service times in column B (with label Minutes).
- Select **MegaStat** : **Correlation/Regression** : **Regression Analysis**.
- In the Regression Analysis dialog box, click in the "Independent variables" box and use the AutoExpand feature to enter the range A1:A12.
- Click in the "Dependent variable" box and use the AutoExpand feature to enter the range B1:B12.
- Check the appropriate Options and Residuals check boxes as follows:
 - 1 Check "Test Intercept" to include a y-intercept and to test its significance.
 - 2 Check "Output Residuals" to obtain a list of the model residuals.
 - 3 Check "Plot Residuals by Observation" and "Plot Residuals by Predicted Y and X" to obtain residual plots versus time, versus the predicted values of y, and versus the values of the independent variable.
 - 4 Check "Normal Probability Plot of Residuals" to obtain a normal plot.
 - 5 Check "Durbin-Watson" to obtain the Durbin-Watson statistic.

To obtain a **point prediction** of y when four computers will be serviced (as well as a confidence interval and prediction interval):

- Click on the drop-down menu above the Predictor Values box and select "Type in predictor values."
- Type the value of the independent variable for which a prediction is desired (here equal to 4) into the "predictor values" box.
- Select a desired level of confidence (here 95%) from the Confidence Level drop-down menu or type in a value.
- Click OK in the Regression Analysis dialog box.

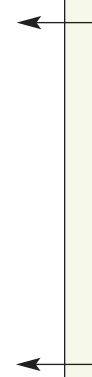


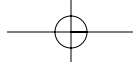
Excel/MINITAB/MegaStat Tutorials

The end of chapter appendices contain helpful tutorials that teach students about the use of Excel, MINITAB, and MegaStat. These tutorials include step-by-step instructions for entering and saving data, retrieving information, and presenting spreadsheets. For additional help, video tutorials for Excel, MINITAB, and MegaStat are provided on the Student CD-ROM.

Excel, MINITAB, and MegaStat Output

Throughout the text, Excel, MINITAB, and MegaStat outputs illustrate how statistical analysis is done electronically.





Exercises...

There are over 1000 exercises in the text. Many use real data from the current business literature. Data sets on the Student CD-ROM are identified by icon in the text. Within each chapter, exercises are broken into two parts—"Concepts" and "Methods and Applications." The methods and applications exercises vary in rigor from routine calculations to fairly sophisticated case study analysis. In addition, there are Internet exercises to help students make use of the Internet for gathering and using real data and supplementary exercises at the ends of chapters.

...And More Exercises

are found on the student text's website.

406 Chapter 10 Experimental Design and Analysis of Variance

TABLE 10.3 Display Panel Study Data

Display Panel	A	B	C
21	24	40	
27	21	36	
24	18	35	
26	19	32	

TABLE 10.4 Bottle Design Study Data

Bottle Design	A	B	C
16	33	23	
18	31	27	
19	37	21	
17	29	28	
13	34	25	

METHODS AND APPLICATIONS

10.3 A study compared three different display panels for use by air traffic controllers. Each display panel was tested in a simulated emergency condition; 12 highly trained air traffic controllers took part in the study. Four controllers were randomly assigned to each display panel. The time (in seconds) needed to stabilize the emergency condition was recorded. The results of the study are given in Table 10.3. For this situation, identify the response variable, factor of interest, treatments, and experimental units.

10.4 A consumer preference study compares the effects of three different bottle designs (A, B, and C) on sales of a popular fabric softener. A completely randomized design is employed. Specifically, 15 supermarkets of equal sales potential are selected, and 5 of these supermarkets are randomly assigned to each bottle design. The number of bottles sold in 24 hours at each supermarket is recorded. The data obtained are displayed in Table 10.4. For this situation, identify the response variable, factor of interest, treatments, and experimental units.

Boxed Equations, Formulas, and Definitions

Each chapter contains easy-to-find boxes that will help students identify and understand the key ideas in the chapter.

Estimation in One-Way ANOVA

- 1 Consider the pairwise difference $\mu_i - \mu_h$ which can be interpreted to be the change in the mean value of the response variable associated with changing from using treatment h to using treatment i . Then, a point estimate of the difference $\mu_i - \mu_h$ is $\bar{x}_i - \bar{x}_h$, where \bar{x}_i and \bar{x}_h are the sample treatment means associated with treatments i and h .
- 2 An individual $100(1 - \alpha)$ percent confidence interval for $\mu_i - \mu_h$ is

$$(\bar{x}_i - \bar{x}_h) \pm t_{\alpha/2} \sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_h} \right)}$$
 Here the $t_{\alpha/2}$ point is based on $n - p$ degrees of freedom, and MSE is the previously defined error mean square found in the ANOVA table.
- 3 A Tukey simultaneous $100(1 - \alpha)$ percent confidence interval for $\mu_i - \mu_h$ is

$$(\bar{x}_i - \bar{x}_h) \pm q_{\alpha} \sqrt{\frac{MSE}{m}}$$
 Here the value q_{α} is obtained from Table A.9 (page 000), which is a table of percentage points of the studentized range. In this table q_{α} is listed corresponding to values of p and $n - p$. Furthermore, we assume that the sample sizes n_i and n_h are equal to the same value, which we denote as m . If n_i and n_h are not equal, we replace $q_{\alpha} \sqrt{MSE/m}$ by $(q_{\alpha}/\sqrt{2}) \sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_h} \right)}$.
- 4 A point estimate of the treatment mean μ_i is \bar{x}_i and an individual $100(1 - \alpha)$ percent confidence interval for μ_i is

$$\bar{x}_i \pm t_{\alpha/2} \sqrt{\frac{MSE}{n_i}}$$
 Here the $t_{\alpha/2}$ point is based on $n - p$ degrees of freedom.

Chapter Summary

In this chapter we studied **probability**. We began by defining an **event** to be an experimental outcome that may or may not occur and by defining the **probability of an event** to be a number that measures the likelihood that the event will occur. We learned that a probability is often interpreted as a **long-run relative frequency**, and we saw that probabilities can be found by examining

sample spaces and by using **probability rules**. We learned several important probability rules—**addition rules**, **multiplication rules**, and **the rule of complements**. We also studied a special kind of probability called a **conditional probability**, which is the probability that one event will occur given that another event occurs, and we used probabilities to define **independent events**.

Glossary of Terms

complement (of an event): If A is an event, the complement of A is the event that A will not occur. (page 136)
conditional probability: The probability that one event will occur given that we know that another event occurs. (page 143)
dependent events: When the probability of one event is influenced by whether another event occurs, the events are said to be dependent. (page 146)
event: A set of sample space outcomes. (page 130)
experiment: A process of observation that has an uncertain outcome. (page 127)
independent events: When the probability of one event is not influenced by whether another event occurs, the events are said to be independent. (page 146)

mutually exclusive events: Events that have no sample space outcomes in common, and, therefore, cannot occur simultaneously. (page 139)
probability (of an event): A number that measures the chance, or likelihood, that an event will occur when an experiment is carried out. (page 131)
sample space: The set of all possible experimental outcomes (sample space outcomes). (page 129)
sample space outcome: A distinct outcome of an experiment (that is, an element in the sample space). (page 129)
subjective probability: A probability assessment that is based on experience, intuitive judgment, or expertise. (page 128)

Important Formulas

Probabilities when all sample space outcomes are equally likely: page 133
 The rule of complements: page 136
 The addition rule for two events: page 139
 The addition rule for two mutually exclusive events: page 140
 The addition rule for N mutually exclusive events: page 141

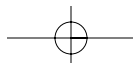
Conditional probability: page 144
 The general multiplication rule: page 145
 Independence: page 146
 The multiplication rule for two independent events: page 147
 The multiplication rule for N independent events: page 147

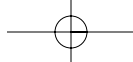
Supplementary Exercises

Exercises 3.34 through 3.37 are based on the following situation: An investor holds two stocks, each of which can rise (R), remain unchanged (U), or decline (D) on any particular day.

Chapter Ending Material


The end of each chapter includes a chapter summary, a comprehensive glossary of terms, and important formula references. The examples shown here are from Chapter 3, Probability.





Exercises for Section 1.4

CONCEPTS

 **1.20** Discuss the difference between a ratio variable and an interval variable.
1.21 Discuss the difference between an ordinal variable and a nominative variable.

METHODS AND APPLICATIONS

1.22 Classify each of the following qualitative variables as ordinal or nominative. Explain your answers.

Qualitative Variable	Categories
Statistics course letter grade	A B C D F
Door choice on Let's Make A Deal	Door #1 Door #2
Television show classifications	TV-G TV-PG TV-14 TV-MA
Personal computer ownership	Yes No
Restaurant rating	***** ** * ** *
Income tax filing status	Married filing jointly Married filing separately Single Head of household Qualifying widow(er)

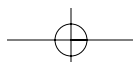
1.23 Classify each of the following qualitative variables as ordinal or nominative. Explain your answers.

Qualitative Variable	Categories
Personal computer operating system	DOS Windows 98 Windows 2000 Windows NT Other
Motion picture classifications	G PG PG-13 R NC-17 X
Level of education	Elementary Middle school High school College Graduate school
Rankings of top 10 college football teams	1 2 3 4 5 6 7 8 9 10
Exchange on which a stock is traded	AMEX NYSE NASDAQ Other
Zip code	45056 90015 etc.

McGraw-Hill's Homework Manager™ and Homework Manager Plus™

McGraw-Hill's Homework Manager is an online electronic homework system and tutor customized to the text and available as an option for students. The system uses the exercises from the text both statically, in a "one problem at a time" fashion as well as algorithmically, where problems can generate multiple data possibilities and answers. You choose the problems. Assignments are graded automatically and the results are stored in your private gradebook. Detailed results show you at a glance how each student does on an assignment or even on an individual problem. You can also monitor progress to see which students need extra help. A Homework Manager icon appears in the text next to those exercise sections available in Homework Manager.

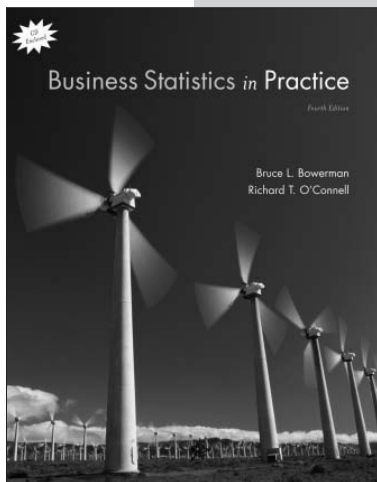
Homework Manager Plus™ is available with the interactive online version of the textbook. It is integrated with the text so that students can review the relevant parts of the text while doing homework. The online text is identical to the printed version, including all figures and exhibits, and provides an interactive learning experience.



For Students

Study Guide

for use with



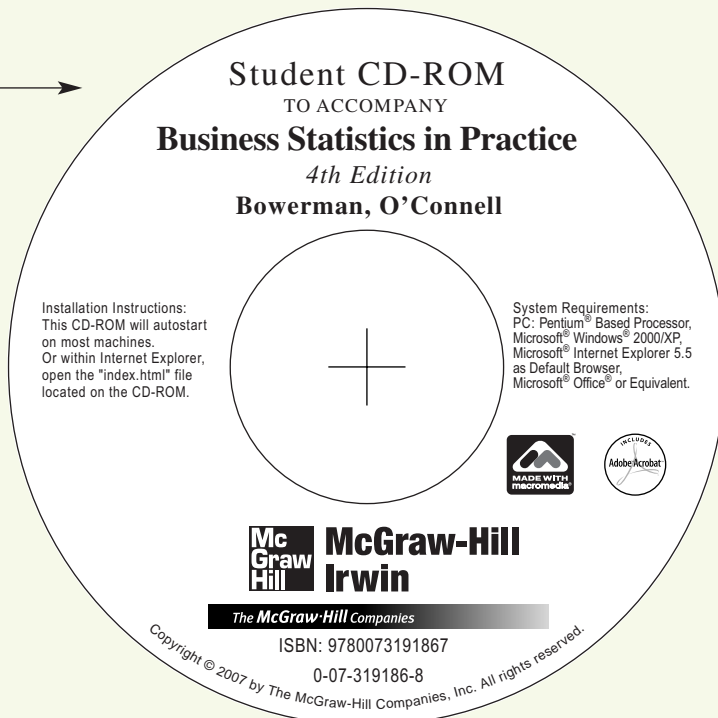
Prepared by
Sandra Strasser

Student Study Guide— 0-07-319188-4

This supplement for the student has been completely updated by Sandra Strassar, Valparaiso University, with the goal of helping students master the course content. It highlights the important ideas in the text and contains detailed tutorial solutions to some of the more challenging exercises. Each chapter includes key ideas, true/false questions, multiple choice questions, and problems, as well as solutions to these. A key feature is solutions using the new version of the Excel add-in MegaStat, which is included on the Student CD-ROM.

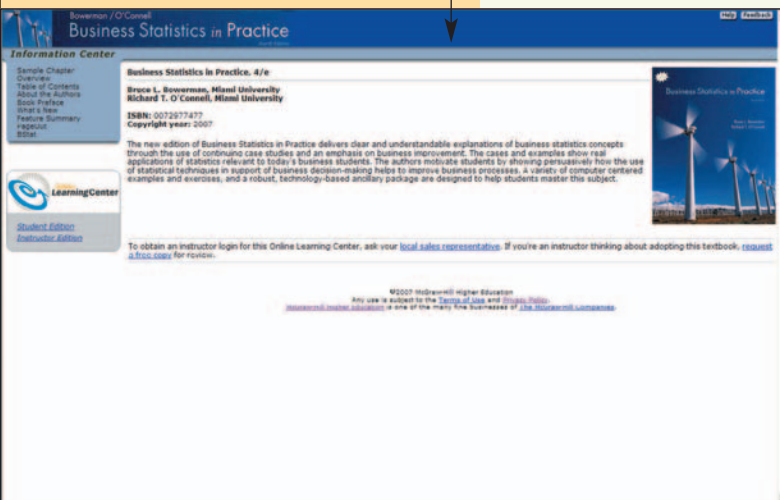
Student CD-ROM

Every new copy of the Fourth Edition is packaged with a free Student CD-ROM featuring self-quizzes, data sets identified by icons in the text, video tutorials in Excel, MINITAB, and MegaStat, PowerPoint presentation slides for easy note taking, MegaStat version 10 and the “Getting Started” Manual developed by J. B. Orris of Butler University, Visual Statistics 2.0, which is described later, Advanced Topics Appendices, Excel Spreadsheets and Workbook, and web links to the text website and the McGraw-Hill Business Statistics website (BSC).



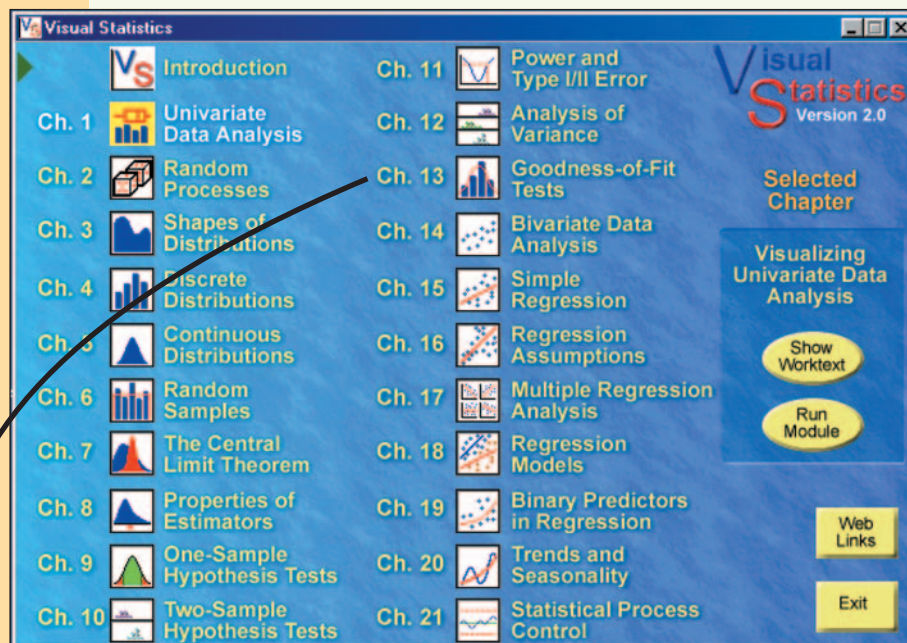
Business Statistics in Practice 4/e website www.mhhe.com/bowerman4e

This text website contains a convenient collection of resources for students to help master course content. Included are additional exercises, self-quizzes, video tutorials, narrated PowerPoint slides, Excel templates and workbook, data sets, link to ALEKS, a self-paced tutorial, Homework Management, a coursework management tool, and more. A password protected section is also available to instructors with downloadable supplements and links to resources, such as PageOut, for creating a professional, interactive course website for students and instructors to access.





Visual Statistics, 2.0, by Doane, Mathieson, and Tracy This software program for Windows is designed for teaching and learning statistics concepts. It is included on the CD-ROM in both the student and instructor editions. The program is unique in that it is intended to help the instructor teach and students learn the concepts through interactive experimentation and visualization. Active learning is promoted through competency building exercises and individual and team projects. Visual Statistics contains 21 software modules and coordinating student worktext all on CD. A printed version of the worktext is also available. An icon appears in the text next to concepts that are further explained and illustrated through Visual Statistics. The samples below show a display of the content and a portion of the tutorial help included in the worktext.



Orientation to Basic Features

This module does goodness-of-fit tests. You can analyze a variety of different data sets by selecting them from the Notebook or create your own using the data editor.

1. Opening screen Start the module by pressing its view button in the *Visual Statistics* menu. When the module is loaded, you will be in the introduction page of the Notebook. Read the questions this module enables you to answer and then click the **Concepts** tab to see the concepts that are covered. Click on the **Examples** tab, select **Class Projects**, select **Weight of D'Anjou Pears**, and press **OK**. Read the Hint that appears in the middle of the display and press **OK**. The upper left shows a frequency histogram with a normal distribution (the default). The Control Panel appears on the right. On the bottom left is a statistical

