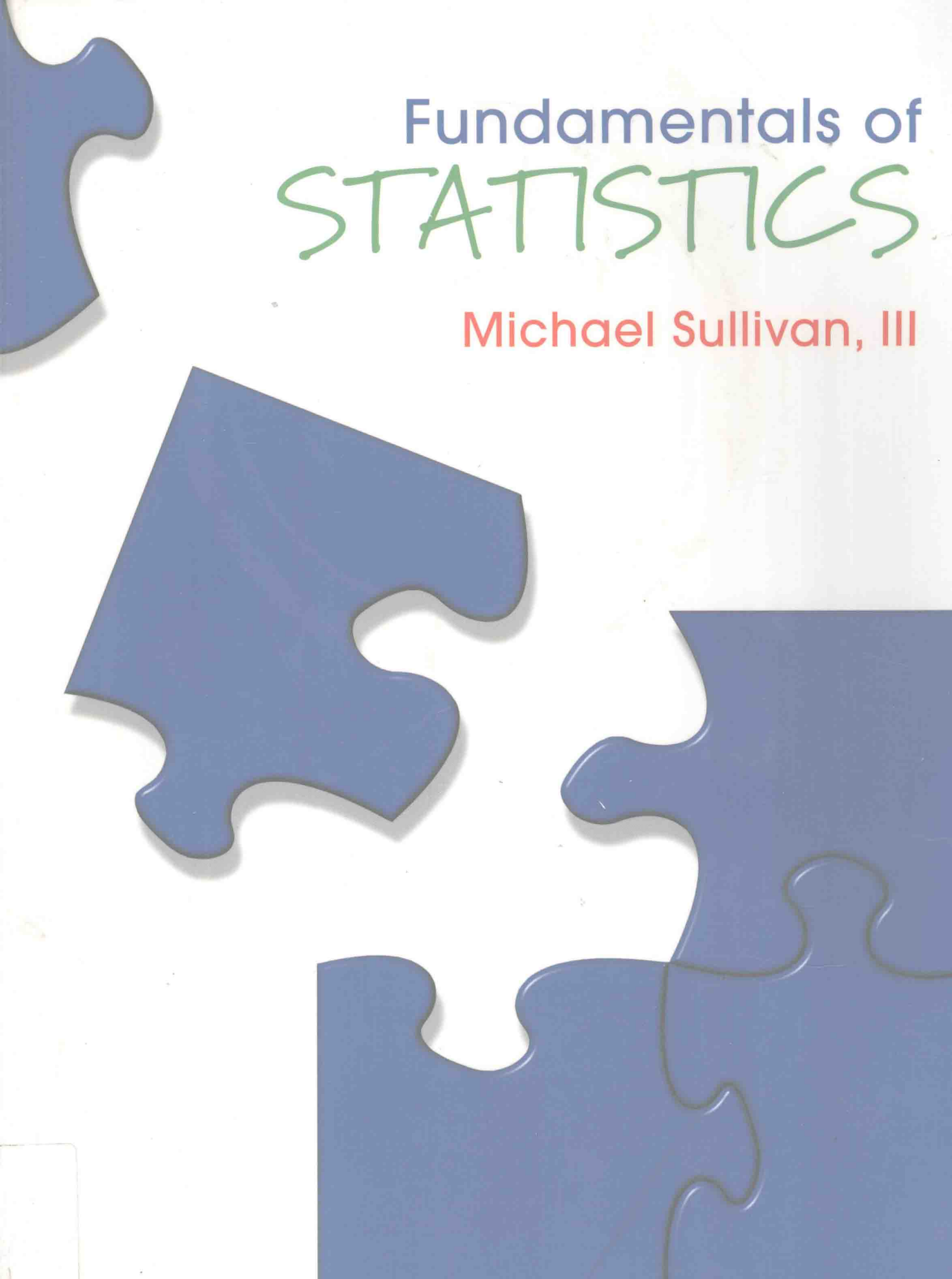
Fundamentals of STATISTICS

Michael Sullivan, III



FUNDAMENTALS OF

Statistics

Michael Sullivan, III
Joliet Junior College

PEARSON Prentice Hall
Upper Saddle River, New Jersey 07458

Library of Congress Cataloging-in-Publication Data

Sullivan, Michael (date)
Fundamentals of statistics / Michael Sullivan, III.
p. cm.
Includes index.
ISBN 0-13-146449-3 (alk. paper)
1. Statistics. I. Title.
QA276.12.S846 2005
519.5--dc22 2003068942

Executive Acquisitions Editor: Petra Recter
Editor in Chief: Sally Yagan
Project Manager: Jacquelyn Riotto
Vice President/Director of Production and Manufacturing: David W. Riccardi
Executive Managing Editor: Kathleen Schiaparelli
Senior Managing Editor: Linda Mihatov Behrens
Production Editor: Barbara Mack
Assistant Managing Editor, Math Media Production: John Matthews
Media Production Editor: Donna Crilly
Assistant Manufacturing Manager/Buyer: Michael Bell
Manufacturing Manager: Trudy Pisciotti
Marketing Manager: Krista M. Bettino
Marketing Assistant: Annett Uebel
Development Editors: Anne Scanlan-Rohrer and Frank Purcell
Editor in Chief, Development: Carol Trueheart
Editorial Assistant/Print Supplements Editor: Joanne Wendelken
Art Director: Jonathan Boylan
Interior Designers: Joseph Sengotta, Judy Matz-Coniglio
Assistants to Art Director: Geoffrey Cassar, Kristine Carney
Cover Designer: Jonathan Boylan
Art Editor: Thomas Benfatti
Creative Director: Carole Anson
Director of Creative Services: Paul Belfanti
Director, Image Resource Center: Melinda Reo
Manager, Rights and Permissions: Zina Arabia
Interior Image Specialist: Beth Boyd-Brenzel
Image Permission Coordinator: Debbie Hewitson
Photo Researcher: Melinda Alexander
Art Studio: Artworks
Managing Editor, AV Management and Production: Patricia Burns
Production Manager: Ronda Whitson
Manager, Production Technologies: Matt Haas
Project Coordinator: Jessica Einsig
Illustrators: Jackie Ambrosius, Kathryn Anderson, Christopher Kayes, Mark Landis, Bill Neff, Audrey Simonetti, Stacy Smith
Quality Assurance: Pamela Taylor, Kenneth Mooney, Timothy Nguyen
Composition: Lithokraft

Pearson Prentice Hall
Pearson Education, Inc.
Upper Saddle River, New Jersey 07458

Printed in the United States of America
10 9 8 7 6 5 4

ISBN 0-13-146449-3

Pearson Education LTD., London
Pearson Education Australia PTY, Limited, Sydney
Pearson Education Singapore, Pte. Ltd
Pearson Education North Asia Ltd, Hong Kong
Pearson Education Canada, Ltd., Toronto
Pearson Educación de Mexico, S.A. de C.V.
Pearson Education—Japan, Tokyo
Pearson Education Malaysia, Pte. Ltd

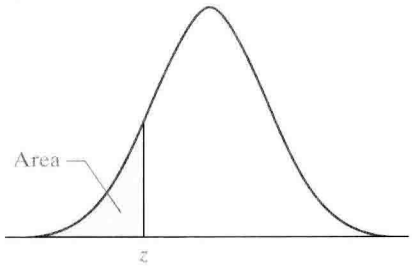


TABLE II

Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
−3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
−3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
−3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

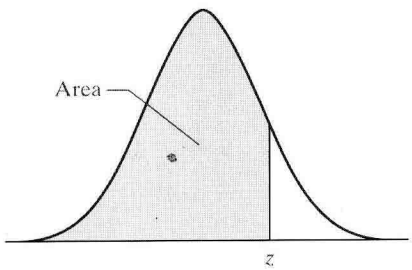


TABLE II *(continued)*

Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

To My Wife
Yolanda

and

My Children
Michael, Kevin, and Marissa

PREFACE TO THE INSTRUCTOR

Fundamentals of Statistics is a brief version of *Statistics: Informed Decisions Using Data.* This shorter text is intended to meet the needs of those schools that offer a one-semester introductory statistics course that do not wish or do not have the time to cover some of the more advanced topics found in *Statistics: Informed Decisions Using Data.* Some of these more advanced topics that appear in *Statistics: Informed Decisions Using Data,* such as residual analysis, nonlinear regression, the Poisson distribution, and nonparametric statistics, have been removed from this briefer text. The coverage of topics has been adapted to make it more suitable for usage in a one-semester course, but this does not compromise the depth of coverage.

Fundamentals of Statistics retains the strong pedagogy, features, and writing style that have made *Statistics: Informed Decisions Using Data* so successful. In addition, it retains the popular problems and discussion that encourage problem solving and statistical thinking.

Having taught introductory statistics for the past 10 years at both 2-year and 4-year institutions has brought to my attention some basic needs and issues of the Introductory Statistics course. The diversity of both the students taking the course and the instructors teaching the course drives many of the challenges in presenting statistics at this level. Statistics is a powerful subject and one that I hold great passion for. It is the coupling of my passion for the subject with the desire for a text that would work for me, my students, and my school that led me to pursue writing this textbook.

Before getting started on this text, I spent many hours reflecting on the goals of the project. Overall, I wanted a textbook that instilled in students my enthusiasm and appreciation for the subject and supported the variety of approaches to teaching the subject. In particular, the following were my goals for the text:

- Students see and appreciate the usefulness of the subject
- Students can be more successful in the course
- Students are provided ample practice
- Instructors can present the material using multiple philosophies

Keeping the Students' Interest

An immediate challenge for any instructor teaching this course is the motivation of the students. I always ask my students to pass in their class notes at the end of the semester. By studying these notes, I have been able to determine areas where my students lost interest and where they became excited about the subject. In addition, I was able to determine topics where the concept was lost and topics where the concept was completely understood. I also noticed that my students appreciated when I took a formal definition or theorem and expressed it in everyday language that they could understand. This view into how students read texts, attend lectures, and take notes was invaluable to me in writing this text.

Today's students demand that the material they study have relevance in their everyday lives. One of the joys of teaching a statistics course is that the question, "When will I ever use this stuff?" rarely comes up. However, students still say, "Okay, I can see the usefulness of this if I was a medical researcher, but that is not my goal in life!" Introductory statistics texts need to have interesting, real data sets that students can relate to. They need to see themselves analyzing the data and using the conclusions to make informed decisions that impact their lives. For example, what type of car should I buy? Where should I invest my retirement money? Where should I live? What major should I choose?

Student Success

It is absolutely true that if students experience success early in a course, they will be inclined to work harder as the course progresses. However, if students struggle early, it will lead to disenchanted students who are more willing to "give up" when a difficult concept is introduced.

How can success be experienced while still maintaining the integrity of the course? After all, statistics is a discipline whose theory is deep and mathematically intensive. In addition, understanding the concepts of statistics can be elusive. This does not mean, however, that an introductory textbook should "spare students details" of statistical thought. It is possible, and indeed, desirable, to expose students to the intricacies of statistics. This can be done without delving deep into mathematical thought and losing most of your students.

The text has been written using a conversational tone that "speaks" to the student. This is accomplished through explanations that don't overly complicate the material. Because the text was written using the class notes obtained from my students, the presentation flows like a lecture. The pedagogy that is found in the text mirrors the presentation that I use in my lectures. When my lecture begins, I tie concepts learned earlier in the course to concepts that I am about to present. This is the "Putting It All Together" feature found at the beginning of each chapter. Before tackling a new section, I review topics that are going to be needed in order to succeed in the section. This is the "Preparing for This Section" feature. I then list the objectives that we are going to cover within the section. As I present the material, I provide students with statistically accurate definitions/theorems. I then restate the definitions/theorems using everyday language that doesn't compromise the definition or theorem. This is the "In Your Own Words" feature. I also mention some of the pitfalls in statistical analysis. These pitfalls are displayed in the "Caution" feature. After going over an example in class, I like to have my students attempt to solve a similar problem in class so that I am confident that they understand the concept before continuing. This is the "Now Work Problem xx" feature.

In addition, the text has been written to appeal to a variety of learning styles. There are many, many graphs and figures that allow students to visualize results. For students who learn best through discovery, there are Explorations that guide students through a series of questions that develop statistical concepts. For auditory learners there is a lecture series in which I present all the examples in the text on compact disk or video.

Practice, Practice, Practice

The only way that students will learn statistics is by doing statistics. The exercise sets of a strong statistics text will help develop a student's confidence in, and understanding of, the material. This is accomplished in a twofold manner: (1) through graduated exercise sets, and (2) through problems that ask the student to think about concepts and encourage statistical thinking.

All the exercise sets begin with Concepts and Vocabulary, which are open-ended questions that ask students to define words and explain concepts using their own words. The next portion of the exercise sets is Skill Building. These are drill and practice type problems that develop computational skills that increase student understanding of formulas and concepts. These problems also serve as confidence building problems so that students experience early success. Finally, the Applying the Concepts problems are real data-based problems that ask a variety of questions that help develop solid statistical analysis. Not only do these problems ask the standard questions such as find the mean or compute a 95% confidence interval, they also ask for students to explain the results. In addition, there are questions that ask students to consider some additional questions. For example, what is the impact of an outlier on the arithmetic mean? Is the linear correlation coefficient resistant? What happens if we have outliers when constructing confidence intervals from small data sets? These "higher level of thinking" problems truly develop a student's understanding of statistical thinking.

Flexibility

Just as there are many different learning styles, there are many different teaching styles. One of the challenges in writing a text is to create a product that "appeals to the masses." As I survey the halls of the Mathematics Department at Joliet Junior College, I see many high-quality instructors who present the same material in many different ways. Some like to incorporate as much technology as possible into their classroom, some prefer to minimize the technology. Some instructors prefer to use collaborative learning in order to present material, while others utilize lecture. For the Introductory Statistics course, some of the instructors are trained in the discipline of statistics, while others are trained in mathematics, but teach statistics. With these varied backgrounds, it is clear that a text needs to meet the needs of all these backgrounds and teaching philosophies.

Let's consider how technology is presented in the course. Every example in the text is presented using the "by hand" approach. The reason for this is twofold. First, and probably most importantly, by presenting a solution by hand, the student's ability to understand the concept is enhanced. How else can a student understand the concept of linear correlation, except by seeing the product of z-scores? Second, it allows for flexibility in philosophies. If your particular philosophy is to present statistics by utilizing "by hand" solutions, it is there for you. If you are more apt to use technology, then you can utilize the "Using Technology" feature. Following virtually every example is a gray "Using Technology" screen, which provides the output from a TI83+, Minitab, or Excel. In addition, problems that have 15 or more observations in the data set have a CD icon, which indicates to the instructor that the data set is available on the data CD.

Using the Text Efficiently with Your Syllabus

To meet the varied needs of diverse syllabi, this book has been organized with flexibility of use in mind. When structuring your syllabus, notice the topics listed in the "Preparing for This Section" material at the beginning of the section, which will tip you off to dependencies within the course.

The two most common variations within an Introductory Statistics course are the treatment of Regression Analysis and the treatment of Probability.

- **Coverage of Correlation and Regression** The text was written with the descriptive portion of bivariate data (Chapter 4) presented after the descriptive portion of univariate data (Chapter 3). For instructors who prefer to postpone the discussion of bivariate data until later in the course, simply skip Chapter 4 and return to it prior to covering Sections 11.3 and 11.4. Within Chapter 4, an instructor may skip Section 4.3 without loss of continuity.
- **Coverage of Probability** The text allows for a course to present an extensive introduction to probability or light coverage of probability. For instructors wishing to present light coverage of probability, they may cover Section 5.1 and skip the remaining sections. A word of caution is in order, however. Instructors who will be covering the Chi-Square Test for Independence will want to cover Sections 5.1 through 5.3. In addition, any instructor who will be covering Binomial Probabilities will want to cover Independence in Section 5.3 and Combinations in Section 5.5.

Chapter by Chapter Content

Chapter 1 Data Collection

This chapter deals with the methods of obtaining data. There is a detailed presentation of the various sampling techniques along with circumstances under which each is used. In addition, there is an entire section dedicated to nonsampling errors and how to control them. The chapter ends with a detailed discussion of experimental design.

Chapter 2 Organizing and Summarizing Data

This chapter addresses methods for summarizing qualitative data (Section 2.1) and quantitative data (Section 2.2). The chapter ends with a discussion of graphical misrepresentations of data. This section can be covered as a reading assignment.

Chapter 3 Numerically Summarizing Data

Sections 3.1 and 3.2 present numerical measures of central tendency and dispersion. Section 3.3 is optional and can be skipped without loss of continuity. However, if it is skipped and Section 6.1 is covered, proceed slowly through the mean and standard deviation of a discrete random variable. Section 3.4 discusses measures of position including the z-score, percentiles, and outliers. Section 3.5 presents exploratory data analysis.

Chapter 4 Describing the Relation between Two Variables

Section 4.1 introduces scatter diagrams and correlation. Section 4.2 presents the least-squares regression line. Section 4.3 presents the coefficient of determination. In addition, the material in this chapter can be postponed until before Section 11.3.

Chapter 5 Probability

Section 5.1 introduces the basic concepts of probability and unusual events. Section 5.2 presents the Addition Rule. Section 5.3 presents the Multiplication Rule. Section 5.4 presents Conditional Probability and can be skipped without loss of continuity. Section 5.5 presents Counting Techniques and can be skipped without loss of continuity with the exception of Combinations.

Chapter 6 The Binomial Probability Distribution

Section 6.1 introduces the concept of a random variable and discrete probability distributions along with expected value. Section 6.2 presents the binomial probability formula and an introduction to inference using binomial probabilities. If you intend to cover Section 6.2, then it is a good idea to also cover Section 5.3.

Chapter 7 The Normal Probability Distribution

Sections 7.1 through 7.3 introduce the normal probability distribution. Section 7.4 presents normal probability plots as a means for assessing normality and is required in order to cover topics presented in Chapters 8–11. Section 7.4 does not require the use of technology because the output generated by technology is presented. This section is necessary in order to help students see that verifying normality is necessary before proceeding with inference for small samples. Section 7.5 introduces sampling distributions. Section 7.6 discusses the normal approximation to the binomial and is optional.

Chapter 8 Confidence Intervals about a Single Parameter

Section 8.1 introduces the construction of a confidence interval when the population standard deviation is known, while Section 8.2 constructs confidence intervals when the population standard deviation is unknown. This approach is different from that in some other texts, but it is logical. It's as simple as "σ known, use z; σ unknown, use t." In both cases, small samples require that the population from which the sample was drawn must be normal–we check this requirement with a normal probability plot. Section 8.3 covers confidence intervals about a population proportion.

Chapter 9 Hypothesis Testing

Section 9.1 provides an introduction to the language of hypothesis testing. Sections 9.2 and 9.3 test hypotheses regarding a population mean, again segmented by "σ known, use z; σ unknown, use t." Section 9.4 presents hypothesis testing about a population proportion. An interesting feature in this section is that it includes how to use the binomial probability distribution to compute exact P-values. This is especially important if the requirement for using the normal approximation to the binomial is not satisfied.

Chapter 10 Inferences on Two Samples

Section 10.1 presents the analysis required for matched-pairs design. Section 10.2 presents the analysis for comparing two means from independent samples. Notice that the discussion regarding pooled estimates of σ is absent. This is because the pooled estimate approach requires that the two populations have a common variance and this is an extremely difficult requirement to test. Because "pooling" versus "not pooling" often provide the same results, it is not necessary at this level to cloud the students' thought process any further. Section 10.3 discusses comparing two population proportions.

Chapter 11 Additional Inferential Procedures

Section 11.1 presents the chi-square goodness of fit. Section 11.2 discusses chi-square tests for independence and homogeneity. Again, if Sections 5.3 and/or 5.4 were skipped, proceed slowly. Sections 11.3 and 11.4 are independent of Sections 11.1 and 11.2. Sections 11.3 and 11.4 require that Sections 4.1 and 4.2 were covered.

Acknowledgments

Textbooks evolve into their final form through the efforts and contributions of many people. First and foremost, I would like to thank my family, whose dedication to this project was just as much as mine: my wife, Yolanda, whose words of encouragement and support were unabashed, and my children, Michael, Kevin, and Marissa, who would come and greet me every morning with smiles that only children can deliver. I owe each of them my sincerest gratitude.

I would also like to thank the entire Mathematics Department at Joliet Junior College, who provided support, ideas, and encouragement to help me complete this project. Countless ideas were bounced off of each of them and their responses are present throughout the text. Special thanks also go to the Joliet Junior College community, who also supported this project.

I sincerely appreciate the comments and suggestions of Erica Egizio, who had the courage to use the text in its nascent stages. Her comments and suggestions have proven to be invaluable. Erica is also to be commended for her "Note to the Instructor" annotations. Special thanks to Michael McCraith, who helped locate interesting data and also made the PowerPoint presentations friendly for both PCs and Macs. Thank you to Faye Dang and Priscilla Gathoni for checking all the answers. Thank you to Elena Catoiu, Kathleen Miranda, Gigi Williams, and Kevin Bodden for proofreading the text to help ensure its accuracy. Thanks to Lindsay Packer, who took on the daunting task of writing the solutions manuals; and to Beverly J. Dretzke, Kate McLaughlin, and Dorothy Wakefield, who wrote the technology guides; and David Lane, who wrote the applets workbook. I would also like to thank Max Linn, who wrote the Case Studies at the end of each chapter. A big thank you to the development editors: Frank Purcell of Twin Prime Editorial and Anne Scanlan-Rohrer. I would also like to thank my father, who has taught me the ropes. It is always nice to learn from the best.

There are many colleagues I would like to thank for their input. I owe each of them my sincerest thanks. I apologize for any omissions.

Fusan Akman, *Coastal Carolina Community College*
Grant Alexander, *Joliet Junior College*
Randall Allbritton, *Daytona Beach Community College*
Diana Joseph Asmus, *Greenville Technical College*
Margaret M. Balachowski, *Michigan Technological University*
Kari Beaty, *Midlands Technical College*
Charles Biles, *Humboldt College*
Kevin Bodden, *Lewis and Clark College*
Ken Bonee, *University of Tennessee*
Joanne Brunner, *Joliet Junior College*
Jim Butterbach, *Joliet Junior College*
Roxanne Byrne, *University of Colorado–Denver*
Joe Castillo, *Broward Community College*
Elena Catoiu, *Joliet Junior College*
Pinyuen Chen, *Syracuse University*
Faye Dang, *Joliet Junior College*
Julie DePree, *University of New Mexico—Valencia Campus*
Erica Egizio, *Joliet Junior College*
Richard Einsporn, *The University of Akron*
Jill Fanter, *Walters State Community College*
Franco Fedele, *University of West Florida*
William P. Fox, *Francis Marion University*
Joe Gallegos, *Salt Lake City Community College*
Frieda Ganter, *California State University–Fresno*
Cheryl B. Hawkins, *Greenville Technical College*
Marilyn M. Hixson, *Brevard Community College*
Kim Jones, *Virginia Commonwealth University*
Maryann E. Justinger, *Erie Community College–South Campus*
Iraj Kalantari, *Western Illinois University*
Donna Katula, *Joliet Junior College*
Mohammad Kazemi, *University of North Carolina–Charlotte*
Rita R. Kolb, *Community College of Baltimore County*
Melody Lee, *Joliet Junior College*
Mike Mays, *West Virginia University*
Jean McArthur, *Joliet Junior College*
Marilyn McCollum, *North Carolina State University*
Michael McCraith, *Joliet Junior College*
Jill McGowan, *Howard University*
Judy Meckley, *Joliet Junior College*
Jane Blake Millar, *Santa Rosa Junior College*
Glenn Miller, *Borough of Manhattan Community College*
Rachel L. Miller, *Portland State University*
Kathleen Miranda, *SUNY at Old Westbury*
Lindsay Packer, *College of Charleston*
Linda Padilla, *Joliet Junior College*
Ingrid Peterson, *University of Kansas*
Nancy Pevey, *Pellissippi State Technical Community College*
Philip Pina, *Florida Atlantic University*
David Ray, *University of Tennessee–Martin*
Mike Rosenthal, *Florida International University*
Robert Sackett, *Erie Community College–North Campus*
Karen Spike, *University of North Carolina–Wilmington*
John Squires, *Cleveland State Community College*
Timothy D. Stebbins, *Kalamazoo Valley Community College*
Patrick Stevens, *Joliet Junior College*
Sharon Stokero, *Michigan Technological University*
Cathleen M. Zucco Teveloff, *Trinity College*
Kirk Trigsted, *University of Idaho*
Robert Tuskey, *Joliet Junior College*
Farroll Tim Wright, *University of Missouri–Columbia*
James Zimmer, *Chattanooga State Technical Community College*

Recognition and thanks are due particularly to the following individuals for their valuable assistance in the preparation of this text: Petra Recter for her editorial skill and attention to detail; Sally Yagan for her support, interest, and friendship; Patrice Jones for his sage advice and marketing skills; Krista M. Bettino for her innovative marketing efforts; Barbara Mack for her organizational skills as a production supervisor; and the entire Prentice Hall sales staff for their confidence.

Michael Sullivan, III

PREFACE TO THE STUDENT

As you begin your study of statistics, you may feel overwhelmed with the number of theorems, definitions, and formulas that confront you. Many students enter their first statistics course with a sense of anxiety. While these feelings are normal, I want to reassure you that statistics is an exciting discipline that can be learned, appreciated, and most importantly, applied, so that you can make informed decisions throughout your life.

My goal in writing this text was to provide students with a "how to" manual for studying and learning statistics. I have written the textbook using clear, uncomplicated language. In fact, the first draft of the text was written using my students' class notes. There are many learning aids included in the text to help make your study of the material easier and more rewarding. In addition, there are many supplements that accompany the text, including: a student solutions manual that presents completely worked out solutions to all the odd-numbered exercises; a CD lecture series that presents worked out solutions to all the examples in the text; and tutorial software (SullivanTutor) that will generate and grade as many problems as you like along with hints and worked out solutions. The questions are based on the objectives presented in the text. The goal is to minimize the number of frustrating experiences that inevitably come up in the learning process.

To help make your study of statistics more enjoyable and rewarding, I have included a list of time-tested "study tips" that promote successful learning. This list is a sequence of events that should occur prior to, during, and after each class meeting. While this list may seem overwhelming, personal experience indicates that it works.

Before Class Begins

1. Read the section(s) to be covered in the upcoming class meeting.
2. Prepare a list of questions that you have based upon your reading. In many cases, your questions will be answered through the lecture. Those that are not answered completely can then be asked.

During Class

1. Arrive early enough to prepare your mind and material for the lecture.
2. Stay alert. Do not doze off or daydream during class. It will be very difficult to understand the lecture when you "return to the class."
3. Take thorough notes. It is normal that certain topics will not be completely understood during the lecture. However, this does not mean that you throw your hands up in despair. Rather, continue to write your class notes. You can ask questions when appropriate. In my personal experience, I was amazed how often my confusion during class was alleviated after studying the in-class notes later when I had more time for reflection.

After Class

1. Reread (and possibly rewrite) your class notes.
2. Reread the section.
3. Do your homework. Homework is not optional. There is an old Chinese proverb that says:

   I hear ... and I forget
   I see ... and I remember
   I do ... and I understand

This proverb applies to anything where one wants to succeed. Would a pianist expect to be the best if she didn't practice? The only way you are going to learn statistics is by doing statistics. When you get problems wrong, ask for help.

4. If you have questions, visit your professor during office hours.

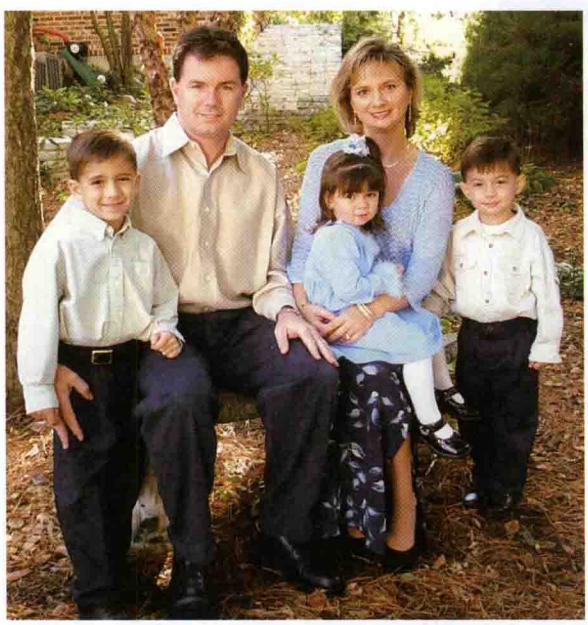
Just like your previous mathematics courses, statistics is a building process. This means the material is used throughout the course. If a topic is not understood, then this lack of understanding could come back to haunt you later in the class. This is the source of a lot of frustration in learning statistics or mathematics. You need to build a strong foundation before you continue to build the house.

How to Use This Book Effectively and Efficiently

First, and foremost, this text was written to be read. You will find that the text has additional explanations and examples that will help you. As mentioned previously, be sure to read the text before attending class.

Many sections begin with "Preparing for This Section," a list of concepts that will be used in the section. Take the short amount of time required to refresh your memory. This will make the section easier to understand and will actually save you time and effort. Each section that has the "Preparing for This Section" feature will have a "Preparing for This Section" quiz on the companion Web site (www.prenhall.com/sullivanstats). The quiz asks questions related to the concepts that are listed. Objectives are provided at the beginning of each section. Read them. They will help you recognize the important ideas and skills developed in the section.

After a concept has been introduced and an example given, you will see NW *Now Work Problem xx.* Go to the exercises at the end of the section, work the problem cited, and check your answer in the back of the book. If you get it right, you can be confident in continuing on in the section. If you don't get it right, go back over the explanations and examples to see what you might have missed. Then rework the problem. Ask for help if you miss it again.



From left to right, Michael IV, Michael III, Marissa, Yolanda, and Kevin

I have included an "In Your Own Words" feature that explains definitions, theorems, and concepts using everyday language. This is meant to help you understand the concepts presented in the text. There are also "Caution" statements. These are meant to make you aware of common errors that occur in statistics, so that you don't make these mistakes.

The chapter review contains a list of formulas and vocabulary introduced in the chapter. Be sure you understand how to use the formulas and that you know the definitions of the vocabulary. There is also a list of objectives along with review problems that correspond to the objective. If you can't do the problems listed for a particular objective, go back to the page indicated and review the material.

Please do not hesitate to contact me, through Prentice Hall, with any suggestions or comments that would improve the text. I look forward to hearing from you.

Michael Sullivan, III

A USER'S GUIDE TO THE TEXT

Utilizing the features found throughout this text will help you learn and study the content for this course. Many of these features were developed by the author's students for students.

CHAPTER 2

Organizing and Summarizing Data

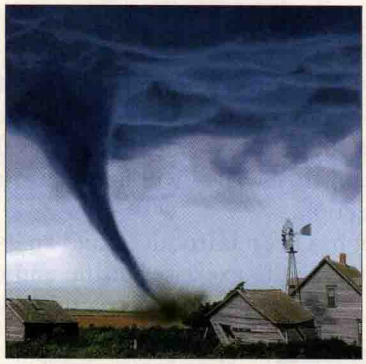
Outline

2.1 Organizing Qualitative Data
2.2 Organizing Quantitative Data I
2.3 Graphical Misrepresentations of Data
- Chapter Review

For additional study help, go to www.prenhall.com/sullivanstats

Materials include

- Projects:
  - Case Study: The Day the Sky Roared
  - Decisions: Tables or Graphs?
  - Consumer Reports Project
- Self-Graded Quizzes
- "Preparing for This Section" Quizzes
- PowerPoint Downloads
- Step-by-Step Technology Guide
- Graphing Calculator Help



Putting It All Together

In Chapter 1, we learned that statistics is a process. The process begins with asking a research question. In order to determine the answer to the question, information (data) must be collected. The information is obtained from a census, existing data sources, surveys or designed experiments. When data are collected from a survey or designed experiment, they must be organized into a manageable form. Data that are not organized are referred to as **raw data**.

Methods for organizing raw data include the creation of tables or graphs, which allow for a quick overview of the information collected. The organization of data is the third step in the statistical process. The procedures used in organizing data into tables and graphs depend upon whether the data are qualitative, discrete, or continuous.

PAGE 39

Putting It All Together

Each chapter opens with a discussion of topics that were covered and how they relate to topics that are about to be discussed. This allows students to see the "big picture" of how the topics relate to each other.

Companion Web Site

Many additional free resources for this text including end of chapter case study, decisions and consumer reports can be found at: *www.prenhall.com/sullivanstats*

Preparing for This Section

Most sections open with a referenced list (by section and page number) of key items to review in preparation for the section ahead. This provides a just-in-time review for the students.

2.1 Organizing Qualitative Data

Preparing for This Section Before getting started, review the following:

✓ Qualitative data (Section 1.1, p. 4)

PAGE 40

Objectives

A numbered list of key objectives appears in the beginning of each section. As the topic corresponding to the objective is addressed, the number appears in the column.

Objectives

1. Construct frequency and relative frequency distributions from qualitative data
2. Construct bar graphs
3. Construct pie charts

In this section we will concentrate on tabular and graphical summaries of qualitative data. Section 2.2 discusses methods for summarizing quantitative data.

PAGE 40

In Your Own Words

When a definition or concept is presented, the "In Your Own Words" feature presents the definition or concept using everyday language while maintaining accuracy.

Definition

In Your Own Words

A frequency distribution shows how often each category occurs. A relative frequency distribution shows the percent of the observations that belong in each category.

The **relative frequency** is the proportion or percent of observations within a category and is found using the formula

$$\text{Relative frequency} = \frac{\text{frequency}}{\text{sum of all frequencies}} \quad (1)$$

A **relative frequency distribution** lists the relative frequency of each category of data.

PAGE 41

▶ EXAMPLE 3 Constructing a Frequency and Relative Frequency Bar Graph

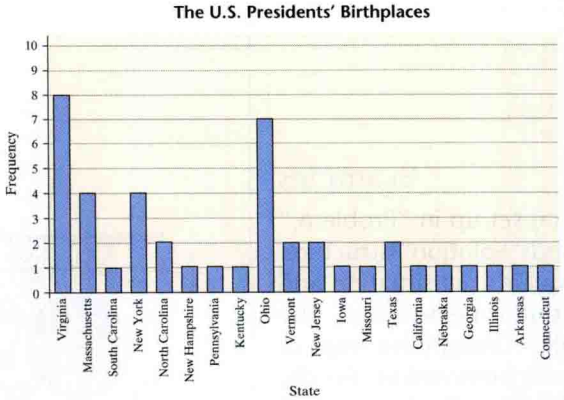
Problem: Use the data summarized in Table 3 to construct (a) a frequency bar graph and (b) a relative frequency bar graph.

Approach: A horizontal axis is used to indicate the category of the data (states, in this case), and a vertical axis is used to represent the frequency or relative frequency (in this case, number of presidential births). Draw bars or rectangles of equal width for each category whose height is the frequency or relative frequency. The bars do not touch each other.

Solution:

(a) Figure 1 shows the frequency bar graph.

Figure 1



Caution

Whenever constructing bar graphs, be sure to include labels for the axes as well as a title for the graph!

PAGE 43

Caution

This alerts students to some of the pitfalls in statistical analysis.