STATISTICS BY EXAMPLE

FIFTH EDITION

TERRY SINCICH

STATISTICS BY EXAMPLE

FIFTH EDITION

TERRY SINCICH
University of South Florida

PRENTICE HALL
Upper Saddle River, NJ 07458

Library of Congress Cataloging-in-Publication Data

Sincich, Terry.
Statistics by example / Terry Sincich.—5th ed.
p. cm.
Includes index.
ISBN 0-02-410981-9
1. Statistics. I. Title.
QA276.12S56 1993
519.5—dc20 92—42330
CIP

On the cover: "Aftershock" is a painting measuring 56" x 105" by San Francisco artist Alex Baker. Its colored patterns, which incorporate fragments of organic derivation, are composed with oils, acrylics, and copper leaf on wood. Baker studied at New York's Art Students' League, the School for Visual Arts, and Parsons School of Design. Galleries in New York, New Orleans, Aspen, San Francisco, and Mexico City have shown his paintings. The artist also has done murals for clubs, restaurants, and public spaces in New York, Miami, and Aspen.

Cartoons by Tom Barnett. Technical illustrations by Kevin Tucker and Caroline Jumper.

A Simon & Schuster Company
Upper Saddle River, New Jersey 07458

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

ISBN 0-02-410981-9

Prentice-Hall International (UK) Limited, *London*
Prentice-Hall of Australia Pty. Limited, *Sydney*
Prentice-Hall Canada Inc., *Toronto*
Prentice-Hall Hispanoamericana, S.A., *Mexico*
Prentice-Hall of India Private Limited, *New Delhi*
Prentice-Hall of Japan, Inc., *Tokyo*
Simon & Schuster Asia Pte. Ltd., *Singapore*
Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*

PREFACE

This introductory college statistics text is designed for students who have only a high school background in mathematics as a prerequisite. It differs from most other texts in two ways:

1. **Real data sets.** Explanations of basic statistical concepts and methodology are based on and motivated by the use of real data sets.
2. **Teaching by example.** Concepts and statistical methods are explained in examples. Many examples arise as questions posed about the data sets.

This practical orientation helps the student to relate statistics to real-life problems, and develops a pattern of thought that will persist after the student enters the working world. The American Statistical Association (ASA) sponsors an annual conference on "Making Statistics More Effective in Schools of Business," where there is a substantial consensus among the participants on the following point: Students are most effectively motivated by seeing statistics at work in real applications, problems, cases, and projects. In much of the current teaching of statistics, there is a limited opportunity for students to work with real data or to make serious use of statistical computing.

The text contains six real data sets to help motivate students. These data sets are also available on diskette (at no charge to adopters).

APPENDIX A The first 100 observations from the set of actual starting salaries, majors, and colleges of 1,795 University of Florida graduates during the period from Fall 1989 to Spring 1991. (The complete data set is available on diskette.)

APPENDIX B The starting salaries (extracted from Appendix A) of bachelor's degree graduates of the Colleges of Business Administration, Engineering, Liberal Arts and Sciences, Journalism, and Nursing.

APPENDIX C The per-member per-month costs accrued by physicians in a managed-care health maintenance organization (HMO).

APPENDIX D Length, weight, and DDT measurements for various species of fish collected from the Tennessee River, Alabama, and its creek tributaries.

APPENDIX E Checkout times for 500 customers at a Publix supermarket.

APPENDIX F The tar, nicotine, and carbon monoxide rankings of 372 domestic cigarette brands, as determined by the Federal Trade Commission.

Through examples and exercises, these data sets are used to develop the notion of a population and a sample, to demonstrate the need for data description, to develop the notion of a sampling distribution, and to motivate the inferential methods commonly studied in an introductory statistics course.

In addition to teaching using data sets and examples, the fifth edition retains the following features of the earlier editions:

1. **Case studies** Four case studies that detail specific and interesting current events are used in each chapter to pose questions for the student. These case studies are exctracted from journals, newspapers, and magazines. The first case study in each chapter, presented as a section in the text, demonstrates how to apply the statistical techniques learned in the chapter to solve a practical problem. The student is left to answer the questions posed by the three case studies given at the end of the chapter. (Refer to the Table of Contents for a list of the case studies.)
2. **Examples** The text, as its name implies, employs the "teaching by example" method. Each section contains several worked examples to demonstrate how to solve various types of statistical problems encountered in the real world.
3. **Key concepts highlighted** Throughout the text, key concepts are highlighted.
   a. **Definitions** are boxed.
   b. **Steps** for constructing bar graphs, performing statistical calculations, and conducting statistical tests are listed and boxed for each procedure.
   c. **Key words**, which must be added to a student's vocabulary, are listed at the end of each chapter.
   d. **Key formulas** are listed at the end of each chapter.
   e. **Warnings**, indicating situations where a student might misuse a statistical technique, are presented in boxed form. The student is directed to specific alternative methods.
4. **Real data (referenced) exercises** Since most students learn best by doing, the text contains a large number (over 1,000) of exercises. The answers for odd-numbered exercises are provided at the end of the text. Each chapter contains exercises at the end of each section and a set of Supplementary Exercises at the chapter's end. The exercises are of two types:
   a. **Learning the Mechanics** These exercises are intended to be straightforward applications of the new concepts presented in the section. They are introduced in a few words and are unhampered by a barrage of background information designed to make them "practical," but which often detracts from instructional objectives. Thus, with a minimum of labor, the student can recheck his or her ability to comprehend a concept or definition.
   b. **Applying the Concepts** The mechanical exercises are followed by realistic exercises that allow the student to see applications of statistics to the solutions of a variety of real-world problems. Almost all of these exercises contain data or information extracted from newspaper articles, magazines, and journals. Once the mechanics are mastered, these exercises develop students' skills at comprehending realistic problems that describe situations to which the techniques may be applied.

5. **Computer printouts** To allow the instructor to emphasize interpretation of the statistical results, printouts generated from statistical computer software packages are presented throughout the text. The computer printouts for three commercial software packages—Minitab, SAS, and SPSS—are displayed.
6. **Reduced emphasis on probability formulas** Based on the recommendations of the ASA, the emphasis on probability formulas is reduced in this introductory text. The probability chapter (Chapter 4) presents only the essential probability concepts (e.g., mutually exclusive events, conditional probability, and independent events) needed to apply the statistical inferential techniques in later chapters. Problem solving for the sake of problem solving is avoided.
7. **Reduced emphasis on analysis of variance formulas** The text adopts a computerized rather than a "cookbook" approach to ANOVA in Chapter 13. Tedious calculation formulas are relegated to an optional section; the emphasis is on understanding designed experiments and interpretation of computer printouts.

The fifth edition contains several substantial modifications, additions, and enhancements.

1. **Computer labs** Easy-to-follow instructions on how to use the statistical analysis commands of SAS, SPSS, and Minitab are provided at the end of most chapters. Except where noted, the commands are appropriate for both the mainframe and PC versions of the packages.
2. **Computer activities** The text is now accompanied by ASP, a user-friendly, totally menu-driven statistical software package designed to run on IBM-compatible PCs. To encourage the use of computers in the statistical analysis of data, the student is asked to analyze a particular data set with ASP (or any other package) at the end of the relevant chapters. An ASP tutorial is provided in Appendix H.
3. **Collecting data (Chapter 1)** The importance of collecting data with random (and representative) samples is emphasized by moving this material from its original location in Chapter 7 to Chapter 1 (Section 1.5).
4. **Descriptive methods for assessing normality (Chapter 6)** A new section (Section 6.3) on determining whether a data set is approximately normal has been added to Chapter 6, A Continuous Probability Distribution: The Normal Distribution. In addition to the traditional graphical methods (histogram, stem-and-leaf display), we present the ratio of the interquartile range to the standard deviation as a check on normality. The emphasis on these techniques early in the text makes the student aware of the importance of checking assumptions in later chapters.
5. **The z statistic versus the t statistic in statistical inference (Chapter 8)** We have expanded the discussion of the proper use of z and t in Chapter 8, Estimation of Population Parameters: Confidence Intervals. The importance of knowing the value of the population variance, σ^2, as well as the size of the sample (large or small), is emphasized.

6. **The relationship between confidence intervals and tests (Chapter 9)** A new section (Section 9.1) emphasizing the similarities and differences between a confidence interval and a two-tailed hypothesis test has been added to Chapter 9, General Concepts of Hypothesis Testing.
7. **Earlier presentation of *p*-values (Chapter 9)** The section on Reporting Test Results: *p*-Values (Section 9.6) is now presented much earlier in the pair of chapters on hypothesis testing. This enables the instructor to emphasize, if desired, the interpretation of results from computer printouts in Chapter 10, Hypothesis Testing: Applications.
8. **Power curves (Chapter 9)** Power curves have been added to the optional section on calculating β and the power of a test (Section 9.7).
9. ***t*-Test for the population coefficient of correlation (Chapter 11)** Earlier editions presented a test for zero correlation in the population based on a table of sample correlation coefficients. In this new edition, we replace the *r* test with the more familiar and equivalent *t*-test (Section 11.7) in Chapter 11, Simple Linear Regression and Correlation.
10. **Multicollinearity in multiple regression (Chapter 12)** A new section on multicollinearity (Section 12.15) has been added to Chapter 12, Multiple Regression and Model Building. The emphasis is on how to detect multicollinearity and its associated problems, and remedial measures.
11. **Analysis of variance chapter moved and expanded (Chapter 13)** The ANOVA chapter now follows the two regression chapters. This move enables us to add two new sections to Chapter 13. In new optional Section 13.9, we demonstrate the regression modeling approach to ANOVA; in new Section 13.10, we use regression residuals to check the ANOVA assumptions.
12. **Logistic regression (Chapter 14)** As an optional section (Section 14.5), we have added material on using logistic regression to model categorical probabilities in Chapter 14, Categorical Data Analysis.
13. **Nonparametric regression (Chapter 15)** In addition to Spearman's test for rank correlation, we also cover a new nonparametric test in Chapter 15: Theil's test for zero slope in regression (Section 15.8).
14. **Case studies updated** The case studies are updated, where necessary, and several new case studies are included.
15. **More exercises with "real" data** Many new "real-life" exercises have been added to each chapter. These exercises, like the case studies, are extracted from news articles, magazines, and professional journals to give students the opportunity to apply their knowledge of statistics to current practical problems.

Numerous less obvious changes in details have been made throughout the text in response to suggestions by current users of the earlier editions.

The text is also accompanied by the following supplementary material:

1. **Student's solutions manual** (by Nancy S. Boudreau) A student's exercise solutions manual presents the full solutions for half (the odd) exercises contained in the text.

2. **Instructor's solutions manual** (by Mark Dummeldinger) The instructor's exercise solutions manual presents the solutions to the other half (the even) exercises and all case studies contained in the text. For adopters, the manual is complimentary from the publisher.
3. **ASP statistical software diskette** New to this edition, the text includes a $5\frac{1}{4}''$ diskette containing the ASP program, *A Statistical Package for Business, Economics, and the Social Sciences*. ASP, from DMC Software, Inc., is a user-friendly, totally menu-driven program that contains all of the major statistical applications covered in the text, plus many more. ASP runs on any IBM-compatible PC with at least 512K of memory and two disk drives. With ASP, students with no knowledge of computer programming can create and analyze data sets easily and quickly. Appendix H contains start-up procedures and a short tutorial on the use of ASP. Full documentation is provided complimentary to adopters of the text.
4. **Minitab supplement** (by David D. Krueger and Ruth K. Meyer) The Minitab computer supplement was developed to be used with Minitab Release 6.1, a general-purpose statistical computing system. The supplement, written expecially for the student with no previous experience with computers, provides step-by-step descriptions of how to use Minitab effectively as an aid in data analysis. Each chapter begins with a list of new commands introduced in the chapter. Brief examples are then given to explain new commands, followed by examples from the text illustrating the new and previously learned commands. Where appropriate, simulation examples are included. Exercises, many of which are drawn from the text, conclude each chapter. A special feature of the supplement is a chapter describing a survey sampling project.
5. **Appendix data sets on diskette** For adopters, the complete data sets in the appendices are available on either a $3\frac{1}{2}''$ or a $5\frac{1}{4}''$ IBM PC diskette (ASCII format), complimentary from the publisher.
6. **Exercise data sets on diskette** Also for adopters, the data for all exercises containing 20 or more observations are available on either a $3\frac{1}{2}''$ or a $5\frac{1}{4}''$ IBM PC diskette (ASCII format). A list of these exercises follows this preface on pages xvii and xviii.
7. **Test bank** (by Mark Dummeldinger) This manual provides a large number of test items utilizing real data.
8. **DellenTest** This unique computer-generated random test system is complimentary to adopters. Utilizing an IBM (or compatible) PC and printer, the system will generate an almost unlimited number of quizzes, chapter tests, final examinations, and drill exercises. At the same time, the system produces an answer key and student worksheet with an answer column that exactly matches the column on the answer key.
9. **ASP Tutorial and Student Guide** (by George Blackford) Most students have little trouble learning to use ASP without documentation. Some, however, may want to purchase the *ASP Tutorial and Student Guide*. Bookstores can order the tutorial from DMC Software, Inc., 6169 Pebbleshire Drive, Grand Blanc, MI 48439.

I wish to acknowledge the many individuals who provided their invaluable assistance during the preparation of the original text and subsequent revisions. Their efforts are much appreciated. The reviewers are listed here:

Edwin F. Baumgartner (Le Moyne College)
Randal Beck (Millikin University)
Jim Bellis (Rochester Institute of Technology)
Elaine Bohanon (Bemidji State University)
John S. Bowdidge (Southwest Missouri State University)
Barbara J. Bulmahn (Indiana University at Fort Wayne)
John Cameron (Rockhurst College)
P. L. Claypool (Oklahoma State University)
Ronald L. Coccari (Cleveland State University)
Robert Cochran (University of Wyoming)
Murray Cohen (University of South Florida)
Joyce Curry-Daly (California Polytechnical)
Paul Dussere (SUNY–Oswego)
Janice Marie Dykacz (Essex Community College)
Joan Girard (Edison Community College)
Damodar Golhar (Western Michigan University)
Gavin Gregory (University of Texas at El Paso)
Derek Hart (McGill University)
Burt Holland (Temple University)
Geoffrey B. Holmewood (Hudson Valley Community College)
Rod Hurley (Hillsborough County Community College)
Thomas B. Laase (University of Southern Colorado)
Maria Ligas (Nova University)
Carolyn Likins (Millikin University)
James T. McClave (University of Florida)
William Mendenhall (University of Florida)
LaVern J. Meyer (Millikin University)
Daniel Mijalko (Western Michigan University)
Glenn W. Milligan (Ohio State University)
S. Mishra (University of South Alabama)
Amitava Mitra (Auburn University)
Maurice Monahan (South Dakota State University)
Kris K. Moore (Baylor University)
Paul Nelson (Kansas State University)
Bill Redmond (Bismarck State College)
Larry J. Ringer (Texas A&M University)
John B. Rushton (Metropolitan State College)
Dale G. Sauers (York College)
Susan Schott (University of Central Florida)
Fay Sewell (Montgomery Community College)
Brian E. Smith (McGill University)
William L. Toth (Henry Ford Community College)
Manel Wijesinha (Pennsylvania State University–York)
Charles W. Zimmerman (Robert Morris College)
Cathleen Zucco (Le Moyne College)

Caroline Jumper deserves special recognition for managing the production of this edition.

Finally, I owe very special thanks to my wife, Faith Sincich, who not only provided the necessary moral support one needs when undertaking a project like this, but also did an excellent job of typing, cutting, pasting, proofing, editing, and solving exercises. Without her, the fifth edition of this text could not have been completed.

EXERCISE DATA SETS AVAILABLE ON DISK

EXERCISE	PAGE	
11.16	542	AUTOMARK and instructor grades for FORTRAN students ($n = 33$)
11.69	572	Breast-height diameters and heights of white spruce trees ($n = 36$)
11.74	583	Factors and length of stay for coronary care patients ($n = 50$)
11.77	585	P/E ratios and R&D expenditures for companies ($n = 20$)
12.10	620	Man-hours, boiler capacity, design pressure, boiler type, and drum type for manufactured boiler drums ($n = 35$)
12.14	624	Price, microprocessor speed, and chip type for PCs ($n = 60$)
12.25	636	Prices, rates of return, and dividend rates for nuclear and nonnuclear stocks ($n = 28$)
12.62	677	Weight percentile and cigarettes smoked/day for children with cystic fibrosis ($n = 25$)
12.66	680	Leased fee values and sizes of Hawaiian properties ($n = 20$)
12.68	683	1984 and 1985 PCB levels in water samples of bays ($n = 37$)
12.72	696	Tar, nicotine, weight, and carbon monoxide measurements for cigarettes ($n = 25$)
12.83	704	Cash compensation of CEOs in four industries ($n = 32$)
12.85	706	Daily high temperatures and peak power loads ($n = 25$)
13.11	741	Numbers of decoding trials for four groups of abbreviated word decoders ($n = 20$)
13.15	749	Prices of grocery items at three supermarkets ($n = 60$)
13.22	764	Learning the Mechanics ($n = 24$)
13.26	765	Antimony amount, cooling method, and shear strength for solder joints ($n = 48$)
13.64	793	Burn-in time, inspection level, and failures for transformer parts ($n = 81$)
14.19	828	Most important value responses of consumers in nine U.S. "nations" ($n = 72$)
14.24	838	PC ownership status and annual income of heads of households ($n = 20$)
14.25	839	Hiring status, education, experience, and sex of applicants ($n = 28$)
15.3	864	Numbers of fish in left compartment of aquarium ($n = 30$)
15.20	876	Lives of air-conditioning systems of airplanes ($n = 20$)
15.36	891	Performance ratings of four types of modems ($n = 20$)
15.38	892	Financial leverages of industries in four countries ($n = 31$)
15.43	897	Rank orderings of stress categories for four levels of managers ($n = 60$)
15.45	898	Usage of computer software applications for three company sizes ($n = 30$)
15.53	906	*Business Week* rank and average starting pay for graduates for top business schools ($n = 20$)
15.59	911	Bond price changes for four firms ($n = 24$)
15.65	913	Weekly radioactive air particulate at four power plants ($n = 40$)

CASE STUDY	PAGE	
3.1	154	Return-to-pay ratios of CEOs ($n = 36$)
3.3	158	Low-bid-estimate ratios for construction contracts ($n = 279$)
8.2	411	IQ scores of identical twins reared apart ($n = 32$)
12.14	686	Precipitation, altitude, latitude, and distance from coast for meteorological stations ($n = 30$)
13.2	799	Candidate, rater, and performance for in-tray tasks ($n = 21$)
15.3	917	Public perception of risk for several dangerous activities ($n = 90$)

CONTENTS