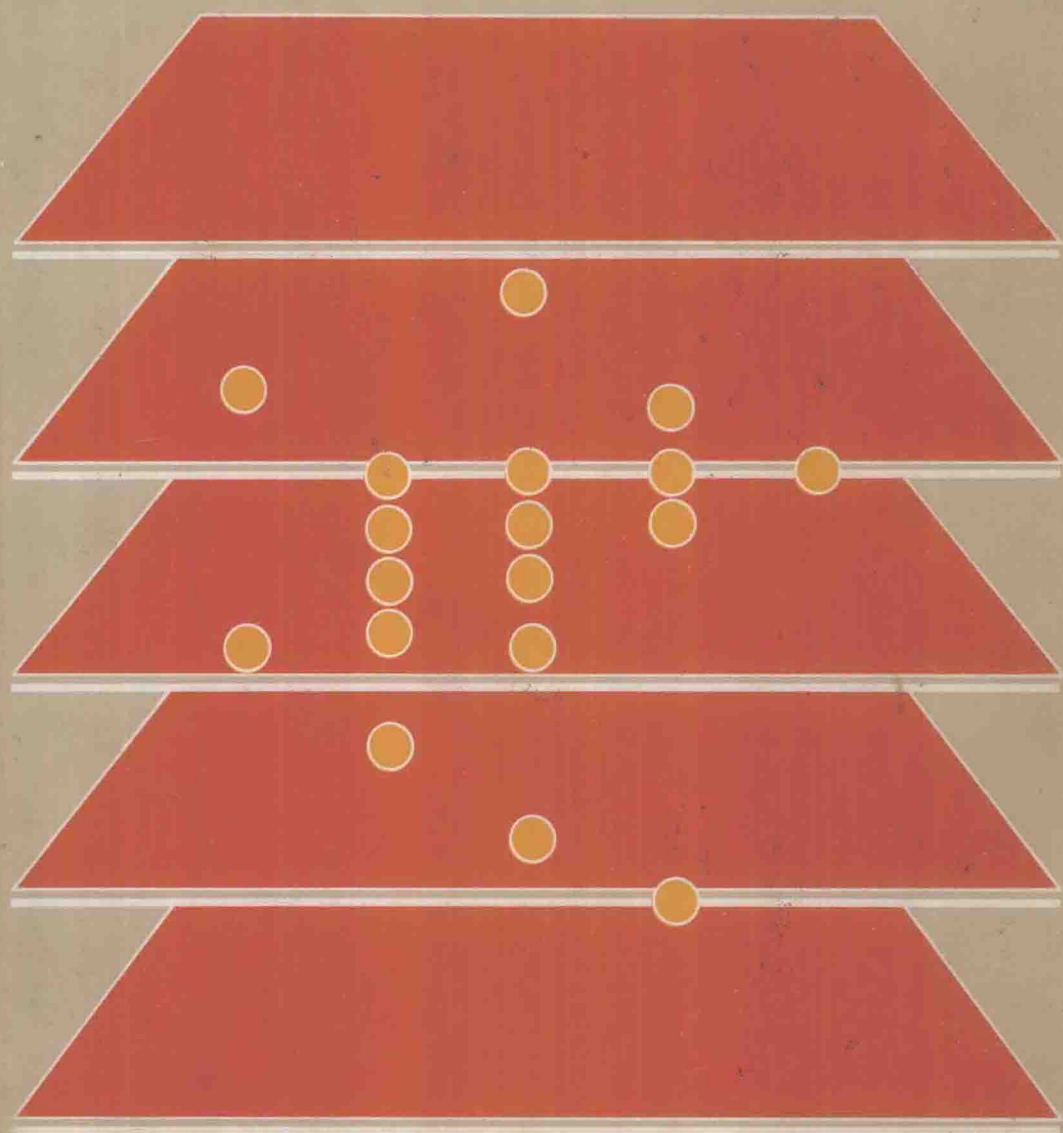


# Statistics for Business and Economics



Van Matre

Gilbreath

# Statistics for Business and Economics

REVISED EDITION



1983

**BUSINESS PUBLICATIONS, INC.**  
Plano, Texas 75075

F

To Mindy, Kelly,  
and My Parents  
J.G.V.

© BUSINESS PUBLICATIONS, INC., 1980 and 1983

*All rights reserved.* No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.

ISBN 0-256-02885-0

Library of Congress Catalog Card No. 82-74066

*Printed in the United States of America*

1 2 3 4 5 6 7 8 9 0 V 0 9 8 7 6 5 4 3

Statistics for Business  
and Economics

**JOSEPH G. VAN MATRE**

University of Alabama  
in Birmingham

**GLENN H. GILBREATH**

Virginia Commonwealth University

# Preface

*Statistics for Business and Economics*, revised edition, is designed primarily for undergraduate students in business, economics, and management. The introductory statistics course provides the sole exposure to statistics for many students, while other students may continue to study in this and related areas such as operations research or econometrics. To meet the needs of both groups, this book emphasizes a conceptual approach with both mathematical and verbal explanations. Selected derivations are presented, often in footnotes. Illustrations and exercises are plentiful; the amount of “busy” work is kept to a minimum.

We assume the reader’s mathematical maturity is college algebra. Students with a knowledge of elementary calculus will benefit from the supplementary explanations given for certain concepts. However, calculus-based material is relegated to footnotes or optional sections so that no loss of continuity occurs if this material is omitted. Optional sections which contain calculus, selected derivations, or advanced concepts are clearly marked with asterisks.

The book may be used for courses of varying lengths. For two quarters or two semesters, we suggest that the book be covered much in its entirety. Several chapters stand alone, however, and may be considered optional. These include Chapter 3, Index Numbers; Chapter 7, Statistical Decision Theory; Chapter 14, Time Series Analysis; and Chapter 15, Nonparametric Methods. A one-quarter or one-semester course could be built around the chapters remaining after excluding these optional chapters. Chapters 9, 10, 11, and 13 may also be considered optional for these shorter courses. Note that Chapter 12, Simple Linear Regression, does *not* require any material from Chapters 10 and 11. Optional sections of chapters are marked with asterisks. (More detailed suggestions may be found in the Instructor’s Manual.)

The use of computers and statistical software packages is integrated throughout the text as a natural component of statistical analysis rather than as something unique. Illustrations are presented throughout the chapters. Some exercises are designated as computer based while others are left to the discretion of the instructor or student. This approach allows the instructor complete freedom in structuring students’ computer involvement. While widely available computer routines provide efficient ways to perform the more laborious calculations, we caution the reader that the use of a computer routine is not a substitute for sound statistical reasoning; it is merely a computational aid.

## FEATURES OF EACH CHAPTER

Each chapter is introduced with a quotation that relates to the chapter content. Biographical sketches of major and minor figures in statistics are included to provide a historical frame of reference and to give human interest to the subject. Cartoons are used to interject humor; reprints from various sources are used to relate text discussion to current practice. Each chapter concludes with a summary of major topics. Following the chapter summary is a section that provides references to the applications appropriate to the chapter. (The Instructor's Manual also provides detailed classifications of this material, allowing instructors to use the material in a variety of ways.) Also included at the end of every chapter are convenient lists of important terms, symbols, and formulas, followed by true-false and fill-in-the blank questions, as well as a generous number of carefully designed exercises. Answers to the odd-numbered exercises are given at the end of the book.

## OVERVIEW OF THE BOOK

All topics under the general heading of "descriptive statistics" are presented in a single chapter, Chapter 2. This includes frequency distributions, averages, and measures of variation. The result is a more unified and complete treatment of these closely related topics. Index numbers are presented as descriptive measures in Chapter 3. However, this topic stands alone and can be placed at any point in the sequence of topics for a given course. This material is organized so that the topic may be covered only briefly or in some depth.

Probabilistic concepts are essential for an understanding of statistical inference. A careful presentation of the basic concepts is given in Chapter 4, while probability mass functions and probability density functions are given in Chapter 5. Sampling and sampling distributions, covered in Chapter 6, complete the preparation for estimation and testing. For hypothesis testing, the modern approach to reporting the *p-value* is provided in addition to discussions of the traditional decision rules. The use of various decision rules is illustrated throughout the text.

Separate chapters are devoted to nonparametrics, multiple regression, chi-square, and analysis of variance. Chapter 7 covers decision theory through normal priors and the two-action problem without extended digressions into utility theory and the assessment of prior distributions. An expanded coverage of time series analysis, Chapter 14, includes exponential smoothing and a discussion of quadratic and exponential trend equations.

One of the unique features of this book is the devotion of a full chapter to the actual use of statistics in today's business world. Chapter 16, Statistics in Business, provides a look at how some of the basic concepts are used in accounting, banking, finance, marketing, personnel management, and other areas of application. The intent of the chapter is to help show the reader why there is a need for statistical concepts and to provide a brief look at how these concepts are used in decision making. This material can be used in at least three ways. The entire chapter, or selected applications, could be read as a motivational tool in an introduction to the course, or as each topic is introduced. Second, study of one or more methods in a given chapter could be followed by a look at a specific application. Finally, the entire chapter could serve as a capstone to the course.

**SUPPLEMENTS**

A comprehensive Instructor's Manual and an optional Study Guide/Workbook are available. The guide may be used independently by students as a supplement or assigned by the instructor as an integral component of the course. The Instructor's Manual provides detailed solutions to all exercises in the text as well as the solutions to the unanswered exercises in the Study Guide. The Instructor's Manual also contains suggestions on using the text in a variety of course situations. Transparency masters for selected material are included in the manual.

**CHANGES IN  
THIS EDITION**

The objectives, basic structure, and level of presentation of the revised edition remain the same as in the first edition. However, users of the first edition will find a number of minor, yet significant, changes. These include:

1. All chapters contain additional exercises.
2. The definition and use of the sample variance have been revised throughout the text to reflect the popular usage of the unbiased estimator
 
$$s^2 = \frac{\sum (x - \bar{x})^2}{n - 1}.$$
3. New sections have been added to incorporate (a) hypothesis testing for differences in means with paired data and (b) the runs test in nonparametrics.
4. Expanded binomial tables now include values for  $\pi > 0.5$ .
5. Some notation is revised to promote easier reader comprehension.
6. Transparency masters for selected material have been added to the Instructor's Manual.
7. The Study Guide/Workbook has been expanded and updated to reflect changes in the text.

**ACKNOWLEDG-  
MENTS**

To all those individuals who took the time to send comments and suggestions about the first edition, we sincerely appreciate your thoughtfulness and efforts. Be assured that all comments, regardless of their nature, were welcomed and carefully studied. Without doubt, this helped to improve the content. We also benefitted from several thoughtful reviews of the entire manuscript. While it is impossible to cite all, we must publicly thank the following individuals: Robert L. Andrews, Virginia Commonwealth University; Aaron H. Brown, Corpus Christi State University; Manfred W. Hopfe, California State University, Sacramento; Ronald S. Koot, Pennsylvania State University; David W. Pentico, Virginia Commonwealth University; Ernest M. Scheuer, California State University, Northridge; Lynette K. Solomon, Stephen F. Austin State University; Betty J. Whitten, University of Georgia; and T. Hillman Willis, Louisiana State University, Shreveport. Thanks also to M. Gene Newport, University of Alabama in Birmingham; to U.A.B. colleagues Seung-Dong Lee, Bob Stanford, and Herb Tsang; and to U.A.B. graduate assistants Ed Gillenwater and Ellen Hiller.

For typing and clerical support, we are very grateful to K. W. Hall and Jill Kramer, both of Virginia Commonwealth University, and to Kaki Culpepper, University of Alabama in Birmingham.



We are grateful to the Literary Executor of the late Sir Ronald A. Fisher, F.R.S., to Dr. Frank Yates, F.R.S., and to Longman Group Ltd., London, for permission to reprint Table III from their book *Statistical Tables for Biological, Agricultural and Medical Research* (6th ed., 1974). To the individuals associated with the University of New Hampshire Statistics Teaching Project, we appreciate permission to reproduce and adapt material from the Statistical Test Item Collection System (STICS). We also thank the other authors and publishers cited throughout the text who allowed us to reprint various figures, tables, and excerpts from articles.

We trust our colleagues will find this book helpful in meeting their classroom needs; we welcome any comments and suggestions.

**Joseph G. Van Matre**

**Glenn H. Gilbreath**

# Contents

## **1. Statistics: An Overview, 1**

- 1.1 Introduction, 2
- 1.2 Statistics Defined, 2
- 1.3 Statistics in Business, 5
- 1.4 Historical Review, 7
- 1.5 Statistics Courses, 8
- 1.6 Terms, 9

## **2. Descriptive Statistics, 11**

- 2.1 Introduction, 12
- 2.2 Data Sources, 13
- 2.3 Frequency Distributions, 13
- 2.4 Preparing a Frequency Distribution, 15
- 2.5 Types of Data and Forms of Frequency Distributions, 18
- 2.6 Cumulative and Relative Frequencies, 20
- 2.7 Charts of Frequency Distributions, 22
- 2.8 Notation, Symbols, and Summation Operations, 25
- 2.9 Measures of Central Tendency, 28
- 2.10 Arithmetic Mean, 28
- 2.11 Median, 34
- 2.12 Mode, 37
- 2.13 Percentiles, 38
- 2.14 Measures of Variation, 40
- 2.15 Range, 40
- 2.16 Variance and Standard Deviation, 40
- 2.17 Coefficient of Variation, 49
- 2.18 Skewness, 50
- 2.19 Comparison of Measures, 51
- 2.20 Use of Computers, 53
- 2.21 Summary, 54
- 2.22 Application, 54
- 2.23 Terms and Formulas, 55

## **3. Index Numbers, 69**

- 3.1 Introduction, 70
- 3.2 Simple Index, 72

- 3.3 Composite Index, 73
- 3.4 Relative of Aggregates Index, 73
- 3.5 Average of Relatives Index, 74
- 3.6 Relative of Weighted Aggregates Index, 76
- 3.7 Average of Weighted Relatives Index, 77
- 3.8 Revision of Indexes, 79
- 3.9 Change of Base Period, 80
- 3.10 Use of Consumer Price Index, 81
- 3.11 Summary, 84
- 3.12 Terms and Formulas, 84

#### **4. Introduction to Probability, 93**

- 4.1 Measuring Uncertainty, 94
- 4.2 Experiments and Events, 94
- 4.3 Counting, 98
- 4.4 Probability Fundamentals, 104
- 4.5 Theory of Probability, 108
- 4.6 Application of Basic Concepts, 110
- 4.7 Classification of Probabilities, 115
- 4.8 Probability Trees, 120
- 4.9 The Birthday Problem, 122
- 4.10 Bayes' Theorem, 123
- 4.11 Summary, 126
- 4.12 Terms and Formulas, 127

#### **5. Probability Distributions, 139**

- 5.1 Introduction, 140
- 5.2 Random Variables, 140
- 5.3 Probability Mass Functions, 144
- 5.4 Uniform PMF, 146
- 5.5 Binomial PMF, 148
- 5.6 Hypergeometric PMF, 153
- 5.7 Binomial Approximation of Hypergeometric PMF, 155
- 5.8 Poisson PMF, 156
- 5.9 Poisson Approximation of Binomial PMF, 159
- 5.10 Probability Density Functions, 160
- \*\*5.11 Use of Integral Calculus, 163
- 5.12 Uniform PDF, 166
- 5.13 Normal PDF, 168
- 5.14 Normal Approximation of Binomial PMF, 176
- 5.15 Linear Combinations of Random Variables, 178
- 5.16 Summary, 181

---

\*\*Optional material.

5.17	Applications, 183
5.18	Terms and Formulas, 183
<b>6.</b>	<b>Sampling and Sampling Distributions, 195</b>
6.1	Introduction, 196
6.2	Simple Random Samples, 196
6.3	Obtaining Simple Random Samples, 197
6.4	Other Probability Samples, 199
6.5	Sampling Distribution of the Mean (without Replacement), 201
**6.6	Sampling Distribution of the Mean (with Replacement), 204
6.7	Central Limit Theorem, 206
6.8	Sampling Distribution of the Proportion, 209
6.9	Sampling Distribution of $(\bar{X}_1 - \bar{X}_2)$ , 213
6.10	Summary, 217
6.11	Applications, 217
6.12	Terms and Formulas, 217
<b>7.</b>	<b>Statistical Decision Theory, 225</b>
7.1	Introduction, 226
7.2	Elements of the Decision Problem, 227
7.3	The McDougal's Example, 228
7.4	Decision Criteria, 229
7.5	Expected Value of Perfect Information, 231
7.6	Opportunity Loss, 232
7.7	Revision of the Discrete Prior, 234
7.8	The Normal Prior, 236
7.9	The Two-Action Problem with Linear Payoff Functions, 238
7.10	Summary, 243
7.11	Application, 244
7.12	Terms and Formulas, 244
<b>8.</b>	<b>Introduction to Statistical Inference, 253</b>
8.1	Introduction, 254
8.2	Estimation, 255
8.3	Criteria for Point Estimators, 256
8.4	Point Estimator for $\mu$ , 257
8.5	Point Estimator for $\sigma^2$ and $\sigma$ , 259
8.6	Point Estimator for $\pi$ , 261
8.7	Interval Estimation, 262
8.8	Interval Estimator for $\mu$ , 262
8.9	Interval Estimator for $\pi$ , 266
8.10	Determination of Sample Size, 268
8.11	Testing, 272

---

\*\*Optional material.

- 8.12 Formal Steps of Hypothesis Testing for  $\mu$ , 275
- 8.13 Decision Rules, 279
- 8.14 Type I and Type II Errors, 281
- 8.15 One-Tailed versus Two-Tailed Tests, 286
- 8.16 Hypothesis Test for  $\pi$ , 290
- 8.17 Summary, 292
- 8.18 Applications, 293
- 8.19 Terms and Formulas, 293

## **9. Additional Topics in Statistical Inference, 303**

- 9.1 Introduction, 304
- 9.2 Student's  $t$  Distribution, 304
- 9.3 Interval Estimator for  $\mu$ : Small Samples, 308
- 9.4 Hypothesis Test for  $\mu$ : Small Samples, 309
- 9.5 Hypothesis Test for  $\mu_1 - \mu_2$ : Large Samples, 312
- 9.6 Hypothesis Test for  $\mu_1 - \mu_2$ : Small Samples, 314
- 9.7 Hypothesis Test for  $\mu_1 - \mu_2$ : Paired Samples, 318
- 9.8 Hypothesis Test for  $\pi_1 - \pi_2$ : Large Samples, 321
- 9.9 Summary, 324
- 9.10 Applications, 324
- 9.11 Terms and Formulas, 324

## **10. The Chi-Square Distribution, 333**

- 10.1 Introduction, 334
- 10.2 The Chi-Square Distribution, 335
- \*\***10.3 Tests on a Variance, 338
- 10.4 Goodness-of-Fit Tests, 342
- 10.5 Tests for Independence, 345
- 10.6 Summary, 348
- 10.7 Terms and Formulas, 349

## **11. The $F$ Distribution and Analysis of Variance, 357**

- 11.1 Introduction, 358
- 11.2 The  $F$  Distribution, 359
- \*\***11.3 Testing for the Equality of Variances, 361
- 11.4 The Completely Randomized Design (One-Way ANOVA), 363
- 11.5 Testing for the Equality of Means, 366
- \*\***11.6 Confidence Intervals, 370
- 11.7 The Randomized Block Design (Two-Way ANOVA), 372
- 11.8 Summary, 377
- 11.9 Applications, 377
- 11.10 Terms and Formulas, 377

---

**\*\***Optional material.

**12. Simple Linear Regression, 385**

- 12.1 Introduction, 386
- 12.2 Review of Linear (Straight-Line) Relationships, 387
- 12.3 The Scatter Diagram, 389
- 12.4 The Simple Linear Regression Model, 390
- 12.5 Fitting the Model, 392
- 12.6 The Sampling Distributions of  $b_0$  and  $b_1$ , 394
- 12.7 Testing and Estimation, 397
- 12.8 Other Topics, 402
- 12.9 Comprehensive Example, 405
- 12.10 Precautionary Notes, 408
- 12.11 Applications, 411
- 12.12 Terms and Formulas, 411

**13. Extensions of Regression Analysis, 421**

- 13.1 Introduction, 422
- 13.2 The  $F$  Test in Regression, 423
- 13.3 Multiple Regression, 427
- 13.4 Inference in Multiple Regression, 432
- \*\*13.5 Qualitative Variables, 439
- 13.6 The Coefficient of Multiple Determination, 443
- 13.7 Other Topics, 444
- 13.8 Applications, 446
- 13.9 Terms and Formulas, 446

**14. Time Series Analysis, 455**

- 14.1 Introduction, 456
- 14.2 Components of Time Series, 457
- 14.3 Smoothing Concepts, 460
- 14.4 Moving Averages, 461
- 14.5 Exponential Smoothing, 464
- 14.6 Trend Analysis, 467
- 14.7 Least Squares Trend Equations, 468
- 14.8 Linear Trend Equation, 469
- 14.9 Quadratic Trend Equation, 472
- 14.10 Exponential Trend Equation, 475
- 14.11 Selecting the Best Model, 480
- 14.12 Use of the Computer, 480
- 14.13 Forecasting, 482
- 14.14 Cyclical Variation, 485
- 14.15 Seasonal Variation, 487
- 14.16 Summary, 494

---

\*\*Optional material.

- 14.17 Application, 496
- 14.18 Terms and Formulas, 496

**15. Nonparametric Methods, 509**

- 15.1 Introduction, 510
- 15.2 Levels of Measurement, 510
- 15.3 Ranked Data, 512
- 15.4 Wilcoxon Signed Rank Test, 514
- 15.5 Mann-Whitney-Wilcoxon Test, 519
- 15.6 Kruskal-Wallis Test, 522
- 15.7 Spearman's Rank Correlation Coefficient, 524
- 15.8 The Runs Test, 527
- 15.9 Summary, 531
- 15.10 Terms and Formulas, 531

**16. Statistics in Business, 541**

- 16.1 Concepts Revisited, 542
- 16.2 Personnel Management, 543
- 16.3 Marketing, 548
- 16.4 Finance, 550
- 16.5 Real Estate, 554
- 16.6 Accounting, 555
- 16.7 Banking, 558
- 16.8 Surveying Public Opinion, 562
- 16.9 Summary, 566

**Appendixes, 571**

- A. Greek Alphabet, 572
- B. Binomial Probability Mass Function, 573
- C. Combinations and Permutations, 579
- D. Poisson Probability Mass Function, 580
- E. Areas of the Normal Distribution, 585
- F. Values of the Unit Normal Loss Function,  $L(D)$ , 586
- G. Student's  $t$  Distribution, 587
- H. Critical Values for the Chi-Square Distribution, 588
- I. Critical Values of the  $F$  Distribution, 589
- J. Critical Values of  $T$  in the Wilcoxon Signed Rank Test, 595
- K. Critical Values of  $T_L$  and  $T_U$  for the Mann-Whitney-Wilcoxon Test, 596
- L. Four-Place Logarithms, 597
- M. Critical Values of  $R$  for the Runs Test, 599

**Answers to Odd-Numbered Questions and Exercises, 601**

**Index, 629**

# Statistics: An Overview

- 1.1 Introduction
- 1.2 Statistics Defined
  - Descriptive Statistics
  - Inferential Statistics
  - Decision Theory
- 1.3 Statistics in Business
  - Marketing
  - Management
  - Finance
  - Accounting
- 1.4 Historical Review
- 1.5 Statistics Courses
- 1.6 Terms



... 'I wonder what Latitude or Longitude I've got to?' (Alice had not the slightest idea what Latitude was, or Longitude either, but she thought they were nice grand words to say).

Lewis Carroll

*Alice's Adventures in Wonderland*

## 1.1 INTRODUCTION

As the student begins this study of statistics, he must be apprised that many new concepts lie ahead. One can understand these concepts only by "speaking the language" of statistics. This is an accepted fundamental for an introductory course in any subject, because the particular terms involved must be mastered before they can be used. Leafing through this text, the student will observe a variety of new symbols (e.g.,  $\sigma$ ,  $\mu$ ) and words (e.g., median, regression). The Greek alphabet printed in Appendix A indicates how these symbols are pronounced ( $\sigma$ : sigma and  $\mu$ : mu), but the statistical meaning of these symbols must be learned. We trust that when a student completes this book his statistical knowledge cannot be characterized by Alice's use of *longitude* and *latitude*.

## 1.2 STATISTICS DEFINED

The word *statistics* calls forth different things to different people. For example, statistics are Tony Dorsett's yards per carry and season touchdowns, or the price-earnings ratio of a stock. However, as we shall see, the study of statistics possesses great utility in capacities other than providing a means of description, although this aspect should not be ignored.

A somewhat traditional definition of the discipline is:

***statistics: a body of methods dealing with the collection, description, analysis, and interpretation of information that can be given in numerical form.***

Since the definition above is rather broad, we consider the subject in more detail by discussing the major divisions within the field; namely, descriptive statistics, inferential statistics, and statistical decision theory.

## Descriptive Statistics

Descriptive statistics are measures that are used to characterize (describe) a mass of numerical data. For example, suppose one were given the SAT scores of the 300 freshmen of the School of Business of some university and then later were asked, "How do the B-school students' entrance scores look?" Probably, no one would reply by beginning to read the 300 scores. An obvious possible descriptor is what the statistical layman calls the average. This and other descriptive measures will be