

Business Data Analysis
SCH-MGMT 650

STATISTICS FOR MANAGERS USING Microsoft® Excel

David M. Levine • David F. Stephan
Timothy C. Krehbiel • Mark L. Berenson

Custom Edition for
UMASS-Amherst
Professor Robert Nakosteen

Taken from:

Statistics for Managers: Using Microsoft® Excel, Fifth Edition
by David M. Levine, David F. Stephan, Timothy C. Krehbiel, and Mark L. Berenson



Cover photo taken by Lauren Labrecque.

Taken from:

Statistics for Managers: Using Microsoft® Excel, Fifth Edition
by David M. Levine, David F. Stephan, Timothy C. Krehbiel, and Mark L. Berenson
Copyright © 2008, 2005, 2002, 1999, 1997 by Pearson Education, Inc.
Published by Prentice Hall
Upper Saddle River, New Jersey 07458

All rights reserved. No part of this book may be reproduced, in any form or by any means, without permission in writing from the publisher.

This special edition published in cooperation with Pearson Custom Publishing.

The information, illustrations, and/or software contained in this book, and regarding the above-mentioned programs, are provided "As Is," without warranty of any kind, express or implied, including without limitation any warranty concerning the accuracy, adequacy, or completeness of such information. Neither the publisher, the authors, nor the copyright holders shall be responsible for any claims attributable to errors, omissions, or other inaccuracies contained in this book. Nor shall they be liable for direct, indirect, special, incidental, or consequential damages arising out of the use of such information or material.

All trademarks, service marks, registered trademarks, and registered service marks are the property of their respective owners and are used herein for identification purposes only.

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN 0-536-04080-X

2008600006

KA

Please visit our web site at www.pearsoncustom.com



PEARSON CUSTOM PUBLISHING
501 Boylston Street, Suite 900, Boston, MA 02116
A Pearson Education Company

*To our wives,
Marilyn L., Mary N., Patti K., and Rhoda B.,
and to our children
Sharyn, Mark, Ed, Rudy, Rhonda, Kathy, and Lori*

ABOUT THE AUTHORS

The textbook authors meet to discuss statistics at Shea Stadium for a Mets v. Phillies game. Shown left to right, Mark Berenson, David Stephan, David Levine, Tim Krehbiel.



David M. Levine is Professor Emeritus of Statistics and Computer Information Systems at Bernard M. Baruch College (City University of New York). He received B.B.A. and M.B.A. degrees in Statistics from City College of New York and a Ph.D. degree from New York University in Industrial Engineering and Operations Research. He is nationally recognized as a leading innovator in statistics education and is the co-author of 14 books including such best selling statistics textbooks as *Statistics for Managers using Microsoft Excel*, *Basic Business Statistics: Concepts and Applications*, *Business Statistics: A First Course*, and *Applied Statistics for Engineers and Scientists using Microsoft Excel and Minitab*.

He also recently wrote *Even You Can Learn Statistics* and *Statistics for Six Sigma Green Belts* published by Financial Times-Prentice-Hall. He is coauthor of *Six Sigma for Green Belts and Champions* and *Design for Six Sigma for Green Belts and Champions*, also published by Financial Times-Prentice-Hall, and *Quality Management Third Ed.*, McGraw-Hill-Irwin (2005). He is also the author of *Video Review of Statistics* and *Video Review of Probability*, both published by Video Aided Instruction. He has published articles in various journals including *Psychometrika*, *The American Statistician*, *Communications in Statistics*, *Multivariate Behavioral Research*, *Journal of Systems Management*, *Quality Progress*, and *The American Anthropologist* and given numerous talks at Decision Sciences, American Statistical Association, and Making Statistics More Effective in Schools of Business conferences. While at Baruch College, Dr. Levine received several awards for outstanding teaching and curriculum development.

David F. Stephan is an instructional designer and lecturer who pioneered the teaching of spreadsheet applications to business school students in the 1980's. He has over 20 years experience teaching at Baruch College, where he developed the first personal computing lab to support statistics and information systems studies and was twice nominated for his excellence in teaching. He is also proud to have been the lead designer and assistant project director of a U.S. Department of Education FIPSE project that brought interactive, multimedia learning to Baruch College.

Today, David focuses on developing materials that help users make better use of the information analysis tools on their computer desktops and is a co-author, with David M. Levine, of *Even You Can Learn Statistics*.

Timothy C. Krehbiel is Professor of Decision Sciences and Management Information Systems at the Richard T. Farmer School of Business at Miami University in Oxford, Ohio. He teaches undergraduate and graduate courses in business statistics. In 1996 he received the prestigious Instructional Innovation Award from the Decision Sciences Institute. In 2000 he received the Richard T. Farmer School of Business Administration Effective Educator Award. He also received a Teaching Excellence Award from the MBA class of 2000.

Krehbiel's research interests span many areas of business and applied statistics. His work appears in numerous journals including *Quality Management Journal*, *Ecological Economics*, *International Journal of Production Research*, *Journal of Marketing Management*, *Communications in Statistics*, *Decision Sciences Journal of Innovative Education*, *Journal of Education for Business*, *Marketing Education Review*, and *Teaching Statistics*. He is a co-author of three statistics textbooks published by Prentice Hall: *Business Statistics: A First Course*, *Basic Business Statistics*, and *Statistics for Managers Using Microsoft Excel*. Krehbiel is also a co-author of the book *Sustainability Perspectives in Business and Resources*.

Krehbiel graduated *summa cum laude* with a B.A. in history from McPherson College in 1983, and earned an M.S. (1987) and Ph.D. (1990) in statistics from the University of Wyoming.

Mark L. Berenson is Professor of Management and Information Systems at Montclair State University (Montclair, New Jersey) and also Professor Emeritus of Statistics and Computer Information Systems at Bernard M. Baruch College (City University of New York). He currently teaches graduate and undergraduate courses in statistics and in operations management in the School of Business and an undergraduate course in international justice and human rights that he co-developed in the College of Humanities and Social Sciences.

Berenson received a B.A. in economic statistics and an M.B.A. in business statistics from City College of New York and a Ph.D. in business from the City University of New York.

Berenson's research has been published in *Decision Sciences Journal of Innovative Education*, *Review of Business Research*, *The American Statistician*, *Communications in Statistics*, *Psychometrika*, *Educational and Psychological Measurement*, *Journal of Management Sciences and Applied Cybernetics*, *Research Quarterly*, *Stats Magazine*, *The New York Statistician*, *Journal of Health Administration Education*, *Journal of Behavioral Medicine*, and *Journal of Surgical Oncology*. His invited articles have appeared in *The Encyclopedia of Measurement & Statistics* and in *Encyclopedia of Statistical Sciences*. He is co-author of 11 statistics texts published by Prentice Hall, including *Statistics for Managers using Microsoft Excel*, *Basic Business Statistics: Concepts and Applications*, and *Business Statistics: A First Course*.

Over the years, Berenson has received several awards for teaching and for innovative contributions to statistics education. In 2005 he was the first recipient of The Catherine A. Becker Service for Educational Excellence Award at Montclair State University.

BRIEF CONTENTS

Preface xix

- 1 INTRODUCTION AND DATA COLLECTION 1**
- 2 PRESENTING DATA IN TABLES AND CHARTS 31**
- 3 NUMERICAL DESCRIPTIVE MEASURES 95**
- 4 BASIC PROBABILITY 147**
- 5 SOME IMPORTANT DISCRETE PROBABILITY DISTRIBUTIONS 179**
- 6 THE NORMAL DISTRIBUTION AND OTHER CONTINUOUS DISTRIBUTIONS 217**
- 7 SAMPLING AND SAMPLING DISTRIBUTIONS 251**
- 8 CONFIDENCE INTERVAL ESTIMATION 283**
- 9 FUNDAMENTALS OF HYPOTHESIS TESTING: ONE-SAMPLE TESTS 327**
- 10 SIMPLE LINEAR REGRESSION 369**
- 11 INTRODUCTION TO MULTIPLE REGRESSION 429**

Appendices A-F 471

Self-Test Solutions and Answers to Selected Even-Numbered Problems 513

Index 535

CD-ROM TOPICS

- 4.5 COUNTING RULES CD4-1**
- 5.6 USING THE POISSON DISTRIBUTION TO APPROXIMATE THE BINOMIAL DISTRIBUTION CD5-1**
- 6.6 THE NORMAL APPROXIMATION TO THE BINOMIAL DISTRIBUTION CD6-1**
- 7.6 SAMPLING FROM FINITE POPULATIONS CD7-1**
- 8.7 ESTIMATION AND SAMPLE SIZE DETERMINATION FOR FINITE POPULATIONS CD8-1**
- 9.7 THE POWER OF A TEST CD9-1**

CONTENTS

Preface **xix**

1 INTRODUCTION AND DATA COLLECTION 1

- 1.1 Why Learn Statistics 2**
- 1.2 Statistics for Managers 2**
 - How This Text is Organized 3
- Using Statistics @ Good Tunes 4**
- 1.3 Basic Vocabulary of Statistics 4**
- 1.4 Data Collection 6**
- 1.5 Types of Variables 8**
 - Levels of Measurement and Measurement Scales 9
- 1.6 Microsoft Excel Worksheets 11**
 - Worksheet Cells 11
 - Designing Effective Worksheets 12
- Summary 13**
- Key Terms 13**
- Chapter Review Problems 14**
- End-of-Chapter Cases 15**
- Learning with the Web Cases 16**
- References 17**
- Excel Companion to Chapter 1 18**
- Key Terms 30**

2 PRESENTING DATA IN TABLES AND CHARTS 31

- Using Statistics @ Choice Is Yours, Part I 32**
- 2.1 Tables and Charts for Categorical Data 32**
 - The Summary Table 33
 - The Bar Chart 33
 - The Pie Chart 34
 - The Pareto Diagram 35
- 2.2 Organizing Numerical Data 40**
 - The Ordered Array 41
 - The Stem-and-Leaf Display 41
- 2.3 Tables and Charts for Numerical Data 44**
 - The Frequency Distribution 44
 - The Relative Frequency Distribution and the Percentage Distribution 46
 - The Cumulative Distribution 47
 - The Histogram 48
 - The Polygon 50
 - The Cumulative Percentage Polygon (Ogive) 51
- 2.4 Cross Tabulations 54**
 - The Contingency Table 55
 - The Side-by-Side Bar Chart 56

2.5 Scatter Plots and Time-Series Plots 58

The Scatter Plot 58

The Time-Series Plot 59

2.6 Misusing Graphs and Ethical Issues 62

Microsoft Excel Graphs 64

Summary 66

Key Terms 66

Chapter Review Problems 67

Managing the *Springville Herald* 73

Web Case 74

References 74

Excel Companion to Chapter 2 75

3 NUMERICAL DESCRIPTIVE MEASURES 95

Using Statistics @ Choice Is Yours, Part II 96

3.1 Measures of Central Tendency 96

The Mean 97

The Median 99

The Mode 100

Quartiles 101

The Geometric Mean 103

3.2 Variation and Shape 105

The Range 105

The Interquartile Range 106

The Variance and the Standard Deviation 106

The Coefficient of Variation 110

Z Scores 111

Shape 112

Visual Explorations: Exploring Descriptive Statistics 113

Microsoft Excel Descriptive Statistics Results 114

3.3 Numerical Descriptive Measures for a Population 118

The Population Mean 118

The Population Variance and Standard Deviation 119

The Empirical Rule 120

The Chebyshev Rule 120

3.4 Exploratory Data Analysis 122

The Five-Number Summary 123

The Box-and-Whisker Plot 124

3.5 The Covariance and the Coefficient of Correlation 127

The Covariance 127

The Coefficient of Correlation 128

3.6 Pitfalls in Numerical Descriptive Measures and Ethical Issues 133

Ethical Issues 133

Summary 134

Key Equations 134

Key Terms 135

Chapter Review Problems 135

Managing the *Springville Herald* 142

Web Case 142

References 142

Excel Companion to Chapter 3 143

4 BASIC PROBABILITY 147

Using Statistics @ The Consumer Electronics Company 148

4.1 Basic Probability Concepts 149

Events and Sample Spaces 150

Contingency Tables 151

Simple (Marginal) Probability 151

Joint Probability 152

General Addition Rule 154

4.2 Conditional Probability 157

Computing Conditional Probabilities 157

Decision Trees 159

Statistical Independence 161

Multiplication Rules 162

Marginal Probability Using the General Multiplication Rule 163

4.3 Bayes' Theorem 166

4.4 Ethical Issues and Probability 171

4.5 (CD-ROM Topic) Counting Rules 172

Summary 172

Key Equations 172

Key Terms 172

Chapter Review Problems 173

Web Case 176

References 176

Excel Companion to Chapter 4 177

5 SOME IMPORTANT DISCRETE PROBABILITY DISTRIBUTIONS 179

Using Statistics @ Saxon Home Improvement 180

5.1 The Probability Distribution for a Discrete Random Variable 180

Expected Value of a Discrete Random Variable 181

Variance and Standard Deviation of a Discrete Random Variable 182

5.2 Covariance and Its Application in Finance 184

Covariance 184

Expected Value, Variance, and Standard Deviation of the Sum of Two Random Variables 186

Portfolio Expected Return and Portfolio Risk 186

5.3 Binomial Distribution 189

5.4 Poisson Distribution 197

5.5 Hypergeometric Distribution 201


5.6 (CD-ROM Topic) Using the Poisson Distribution to Approximate the Binomial Distribution 204

Summary 204

Key Equations 204


Key Terms 205
Chapter Review Problems 206
Managing the *Springville Herald* 209
Web Case 209
References 210
Excel Companion to Chapter 5 211

6 THE NORMAL DISTRIBUTION AND OTHER CONTINUOUS DISTRIBUTIONS 217


Using Statistics @ OurCampus! 218
6.1 Continuous Probability Distributions 218
6.2 The Normal Distribution 219
 Visual Explorations: Exploring the Normal Distribution 229
6.3 Evaluating Normality 234
 Comparing Data Characteristics to Theoretical Properties 234
 Constructing the Normal Probability Plot 236
6.4 The Uniform Distribution 238
6.5 The Exponential Distribution 241
6.6  (CD-ROM Topic) The Normal Approximation to the Binomial Distribution 243
Summary 243
Key Equations 243
Key Terms 243
Chapter Review Problems 244
Managing the *Springville Herald* 246
Web Case 246
References 246
Excel Companion to Chapter 6 247

7 SAMPLING AND SAMPLING DISTRIBUTIONS 251

Using Statistics @ Oxford Cereals 252
7.1 Types of Sampling Methods 252
 Simple Random Samples 253
 Systematic Samples 256
 Stratified Samples 256
 Cluster Samples 257
7.2 Evaluating Survey Worthiness 258
 Survey Error 259
 Ethical Issues 260
7.3 Sampling Distributions 261
7.4 Sampling Distribution of the Mean 262
 The Unbiased Property of the Sample Mean 262
 Standard Error of the Mean 264
 Sampling from Normally Distributed Populations 265
 Sampling from Non-Normally Distributed Populations—
 The Central Limit Theorem 268
 Visual Explorations: Exploring Sampling Distributions 270
7.5 Sampling Distribution of the Proportion 272

7.6	 (CD-ROM Topic) Sampling From Finite Populations	275
	Summary	275
	Key Equations	276
	Key Terms	276
	Chapter Review Problems	276
	Managing the <i>Springville Herald</i>	279
	Web Case	279
	References	280
	Excel Companion to Chapter 7	281

8 CONFIDENCE INTERVAL ESTIMATION 283

	Using Statistics @ Saxon Home Improvement	284
8.1	Confidence Interval Estimation for the Mean (σ Known)	285
8.2	Confidence Interval Estimation for the Mean (σ Unknown)	290
	Student's t Distribution	290
	Properties of the t Distribution	290
	The Concept of Degrees of Freedom	291
	The Confidence Interval Statement	292
8.3	Confidence Interval Estimation for the Proportion	296
8.4	Determining Sample Size	299
	Sample Size Determination for the Mean	300
	Sample Size Determination for the Proportion	302
8.5	Applications of Confidence Interval Estimation in Auditing	306
	Estimating the Population Total Amount	307
	Difference Estimation	308
	One-Sided Confidence Interval Estimation of the Rate of Noncompliance with Internal Controls	311
8.6	Confidence Interval Estimation and Ethical Issues	313
8.7	 (CD-ROM Topic) Estimation and Sample Size Determination for Finite Populations	314
	Summary	314
	Key Equations	314
	Key Terms	315
	Chapter Review Problems	315
	Managing the <i>Springville Herald</i>	320
	Web Case	321
	References	321
	Excel Companion to Chapter 8	322

9 FUNDAMENTALS OF HYPOTHESIS TESTING: ONE-SAMPLE TESTS 327

	Using Statistics @ Oxford Cereals, Part II	328
9.1	Hypothesis-Testing Methodology	328
	The Null and Alternative Hypotheses	328
	The Critical Value of the Test Statistic	330
	Regions of Rejection and Nonrejection	330
	Risks in Decision Making Using Hypothesis-Testing Methodology	331


9.2 ZTest of Hypothesis for the Mean (σ Known) 334
 The Critical Value Approach to Hypothesis Testing 334
 The p -Value Approach to Hypothesis Testing 337
 A Connection Between Confidence Interval Estimation
 and Hypothesis Testing 340

9.3 One-Tail Tests 342
 The Critical Value Approach 342
 The p -Value Approach 343

9.4 t Test of Hypothesis for the Mean (σ Unknown) 346
 The Critical Value Approach 347
 The p -Value Approach 349
 Checking Assumptions 349

9.5 ZTest of Hypothesis for the Proportion 353
 The Critical Value Approach 354
 The p -Value Approach 355

9.6 Potential Hypothesis-Testing Pitfalls and Ethical Issues 357

9.7  (CD-ROM Topic) The Power of a Test 359

Summary 359

Key Equations 360

Key Terms 360

Chapter Review Problems 360

Managing the *Springville Herald* 363

Web Case 363

References 363

Excel Companion to Chapter 9 364

10 SIMPLE LINEAR REGRESSION 369

Using Statistics @ Sunflowers Apparel 370

10.1 Types of Regression Models 370

10.2 Determining the Simple Linear Regression Equation 372
 The Least-Squares Method 373
 Visual Explorations: Exploring Simple Linear Regression Coefficients 376
 Predictions in Regression Analysis: Interpolation Versus Extrapolation 377
 Computing the Y Intercept, b_0 , and the Slope, b_1 377

10.3 Measures of Variation 382
 Computing the Sum of Squares 382
 The Coefficient of Determination 384
 Standard Error of the Estimate 386

10.4 Assumptions 387

10.5 Residual Analysis 388
 Evaluating the Assumptions 388

10.6 Measuring Autocorrelation: The Durbin-Watson Statistic 392
 Residual Plots to Detect Autocorrelation 392
 The Durbin-Watson Statistic 394

10.7 Inferences About the Slope and Correlation Coefficient 397
 t Test for the Slope 397
 F Test for the Slope 398

Confidence Interval Estimate of the Slope (β_1)	400
t Test for the Correlation Coefficient	400
10.8 Estimation of Mean Values and Prediction of Individual Values	404
The Confidence Interval Estimate	404
The Prediction Interval	405
10.9 Pitfalls in Regression and Ethical Issues	408
Summary	412
Key Equations	413
Key Terms	414
Chapter Review Problems	414
Managing the <i>Springville Herald</i>	420
Web Case	421
References	421
Excel Companion to Chapter 10	422

11 INTRODUCTION TO MULTIPLE REGRESSION 429

Using Statistics @ OmniFoods	430
11.1 Developing a Multiple Regression Model	430
Interpreting the Regression Coefficients	431
Predicting the Dependent Variable Y	433
11.2 r^2, Adjusted r^2, and the Overall F Test	435
Coefficient of Multiple Determination	436
Adjusted r^2	436
Test for the Significance of the Overall Multiple Regression Model	437
11.3 Residual Analysis for the Multiple Regression Model	439
11.4 Inferences Concerning the Population Regression Coefficients	441
Tests of Hypothesis	441
Confidence Interval Estimation	443
11.5 Testing Portions of the Multiple Regression Model	445
Coefficients of Partial Determination	448
11.6 Using Dummy Variables and Interaction Terms in Regression Models	450
Interactions	453
Summary	460
Key Equations	462
Key Terms	462
Chapter Review Problems	463
Managing the <i>Springville Herald</i>	466
Web Case	466
References	466
Excel Companion to Chapter 11	467

Appendices 471

A. Review of Arithmetic, Algebra, and Logarithms	472
B. Summation Notation	474
C. Statistical Symbols and Greek Alphabet	477
D. Student CD-ROM Contents	478

E. Tables 485

F. FAQs About Using Microsoft Excel and PHStat2 510

Self-Test Solutions and Answers to Select Even-Numbered Problems 513

Index 535

CD-ROM Topics

4.5 Counting Rules CD4-1

5.6 Using the Poisson Distribution To Approximate the Binomial Distribution CD5-1

6.6 The Normal Approximation to the Binomial Distribution CD6-1

7.6 Sampling From Finite Populations CD7-1

8.7 Estimation and Sample Size Determination for Finite Populations CD8-1

9.7 The Power of a Test CD9-1

He is nationally recognized as a leading innovator in statistics education and is the co-author of 14 books, including such best-selling statistics textbooks as *Statistics for Managers Using Microsoft Excel*, *Basic Business Statistics: Concepts and Applications*, *Business Statistics: A First Course*, and *Applied Statistics for Engineers and Scientists Using Microsoft Excel and Minitab*. He was an Instructor/Lecturer of Computer Information Systems at Baruch College (City University of New York) for over 20 years and also served as an Assistant to the Provost and to the Dean of the School of Business & Public Administration for computing. *Using Microsoft Excel*, 5e 2008 Prentice-Hall, Inc. Linear Regression Example Data

House Price in \$1000s (Y) Square Feet (X) 245 1400 312 1600 279 1700 308 1875 199 1100 219 1550 405 2350 324 2450 319 1425 255

Statistics for Managers Using Microsoft Excel, 5e 2008 Prentice-Hall, Inc. Linear Regression Example Scatterplot

House price model:scatter. Total 9 Coefficients t Stat P-value Lower 95% Upper 95% Intercept Square Feet *Statistics for Managers Using Microsoft Excel*, 5e 2008 Prentice-Hall, Inc. Comparing Standard Errors

SYX is a measure of the variation of observed Y values from the regression line $\hat{Y} = Y - X X$. The magnitude of SYX should always be judged relative to the size of the. This text is the gold standard for learning how to use Microsoft Excel® in business statistics, helping students gain the understanding they need to be successful in their careers. The authors present statistics in the context of specific business fields; full chapters on business analytics further prepare students for success in their professions. Current data throughout the text lets students practice analyzing the types of data they will see in their professions. The friendly writing style include tips throughout to encourage learning. The book also integrates PHStat, an add-in that bolster