

Hector Guerrero

Excel Data Analysis

Modeling and Simulation

Second Edition



Springer

Excel Data Analysis

Hector Guerrero

Excel Data Analysis

Modeling and Simulation

Second Edition



Springer

Hector Guerrero
College of William & Mary
Mason School of Business
Williamsburg, VA, USA

ISBN 978-3-030-01278-6 ISBN 978-3-030-01279-3 (eBook)
<https://doi.org/10.1007/978-3-030-01279-3>

Library of Congress Control Number: 2018958317

© Springer Nature Switzerland AG 2019

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To my parents . . .Paco and Irene

Preface

Why Does the World Need—*Excel Data Analysis, Modeling, and Simulation*?

When spreadsheets first became widely available in the early 1980s, it spawned a revolution in teaching. What previously could only be done with arcane software and large-scale computing was now available to the *common man*, on a desktop. Also, before spreadsheets, most substantial analytical work was done outside the classroom where the tools were; spreadsheets and personal computers moved the work into the classroom. Not only did it change how the data analysis curriculum was taught, but it also empowered students to venture out on their own to explore new ways to use the tools. I can't tell you how many phone calls, office visits, and/or emails I have received in my teaching career from ecstatic students crowing about what they have just done with a spreadsheet model.

I have been teaching courses related to business and data analytics and modeling for over 40 years, and I have watched and participated in the spreadsheet revolution. During that time, I have been a witness to the following important observations:

- Each successive year has led to more and more demand for Excel-based analysis and modeling skills, both from students, practitioners, and recruiters.
- Excel has evolved as an ever more powerful suite of tools, functions, and capabilities, including the recent iteration and basis for this book—Excel 2013.
- The ingenuity of Excel users to create applications and tools to deal with complex problems continues to amaze me.
- Those students who preceded the spreadsheet revolution often find themselves at a loss as to where to go for an introduction to what is commonly taught to most undergraduates in business and sciences.

Each one of these observations has motivated me to write this book. The first suggests that there is no foreseeable end to the demand for the skills that Excel enables; in fact, the need for continuing productivity in all economies guarantees that an individual with proficiency in spreadsheet analysis will be highly prized by an

organization. At a minimum, these skills permit you freedom from *specialists* that can delay or hold you captive while waiting for a solution. This was common in the early days of information technology (IT); you requested that the IT group provide you with a solution or tool and you waited, and waited, and waited. Today if you need a solution you can do it yourself.

The combination of the second and third observations suggests that when you couple bright and energetic people with powerful tools and a good learning environment, wonderful things can happen. I have seen this throughout my teaching career, as well as in my consulting practice. The trick is to provide a teaching vehicle that makes the analysis accessible. My hope is that this book is such a teaching vehicle. I believe that there are three simple factors that facilitate learning—select examples that contain interesting questions, methodically lead students through the rationale of the analysis, and thoroughly explain the Excel tools to achieve the analysis.

The last observation has fueled my desire to lend a hand to the many students who passed through the educational system *before* the spreadsheet analysis revolution: to provide them with a book that points them in the right direction. Several years ago, I encountered a former MBA student in a Cincinnati Airport bookstore. He explained to me that he was looking for a good Excel-based book on data analysis and modeling—“You know it’s been more than 20 years since I was in a Tuck School classroom, and I desperately need to understand what my interns seem to be able to do so easily.” By providing a broad variety of exemplary problems, from *graphical/statistical analysis* to *modeling/simulation* to *optimization*, and the Excel tools to accomplish these analyses, most readers should be able to achieve success in their self-study attempts to master spreadsheet analysis. Besides a good compass, students also need to be made aware of *the possible*. It is not usual to hear from students “Can you use Excel to do *this*?” or “I didn’t know you could do *that* with Excel!”

Who Benefits from This Book?

This book is targeted at the student or practitioner who is looking for a *single* introductory Excel-based resource that covers three essential business skills—data analysis, business modeling, and simulation. I have successfully used this material with undergraduates, MBAs, and executive MBAs and in executive education programs. For my students, the book has been the main teaching resource for both semester and half-semester long courses. The examples used in the books are sufficiently flexible to guide teaching goals in many directions. For executives, the book has served as a compliment to classroom lectures, as well as an excellent post-program, self-study resource. Finally, I believe that it will serve practitioners, like that former student I met in Cincinnati, who have the desire and motivation to refurbish their understanding of data analysis, modeling, and simulation concepts through self-study.

Key Features of This Book

I have used a number of examples in this book that I have developed over many years of teaching and consulting. Some are brief and to the point; others are more complex and require considerable effort to digest. I urge you to not become frustrated with the more complex examples. There is much to be learned from these examples, not only the analytical techniques, but also *approaches* to solving complex problems. These examples, as is always the case in real world, messy problems, require making reasonable assumptions and some concession to simplification if a solution is to be obtained. My hope is that the approach will be as valuable to the reader as the analytical techniques. I have also taken great pains to provide an abundance of Excel screen shots that should give the reader a solid understanding of the chapter examples.

But, let me vigorously warn you of one thing—this is not an Excel *how-to* book. Excel *how-to* books concentrate on the Excel tools and not on analysis—it is assumed that you will fill in the analysis blanks. There are many excellent Excel *how-to* books on the market and a number of excellent websites (e.g., MrExcel.com) where you can find help with the details of specific Excel issues. I have attempted to write a book that is about analysis, analysis that can be easily and thoroughly handled with Excel. Keep this in mind as you proceed. So in summary, remember that the analysis is the primary focus and that Excel simply serves as an excellent vehicle by which to achieve the analysis.

Second Edition

The second edition of this book has updated to the current version of Excel, 2013. The additions and changes to Excel, since the first publication of the book, have been significant; thus, a revision was requested by many users. Additionally, topics have been extended for a more complete coverage. For example, in Chaps. 2–6 a more in-depth discussion of statistical techniques (sampling, confidence interval analysis, regression, and graphical analysis) is provided. Also, in numerous passages, changes have been made to provide greater ease of understanding.

Williamsburg, VA, USA

Hector Guerrero

Acknowledgements

I would like to thank the editorial staff of Springer for their invaluable support—Christian Rauscher and Barbara Bethke. Thanks to Ms. Elizabeth Bowman and Traci Walker for their invaluable editing effort over many years. Special thanks to the countless students I have taught over the years, particularly Bill Jelen, the World Wide Web’s Mr. Excel who made a believer out of me. Finally, thanks to my family and friends who provided support over the years.

Contents

- 1 Introduction to Spreadsheet Modeling 1**
 - 1.1 Introduction 1
 - 1.2 What’s an MBA to do? 2
 - 1.3 Why Model Problems? 3
 - 1.4 Why Model Decision Problems with Excel? 3
 - 1.5 The Feng Shui of Spreadsheets 5
 - 1.6 A Spreadsheet Makeover 8
 - 1.6.1 Julia’s Business Problem—A Very Uncertain Outcome 8
 - 1.6.2 Ram’s Critique 11
 - 1.6.3 Julia’s New and Improved Workbook 11
 - 1.7 Summary 17
 - Key Terms 18
 - Problems and Exercises 18
- 2 Presentation of Quantitative Data: Data Visualization 21**
 - 2.1 Introduction 21
 - 2.2 Data Classification 22
 - 2.3 Data Context and Data Orientation 23
 - 2.3.1 Data Preparation Advice 26
 - 2.4 Types of Charts and Graphs 29
 - 2.4.1 Ribbons and the Excel Menu System 29
 - 2.4.2 Some Frequently Used Charts 31
 - 2.4.3 Specific Steps for Creating a Chart 35
 - 2.5 An Example of Graphical Data Analysis and Presentation 38
 - 2.5.1 Example—Tere’s Budget for the 2nd Semester of College 39
 - 2.5.2 Collecting Data 40
 - 2.5.3 Summarizing Data 40
 - 2.5.4 Analyzing Data 43
 - 2.5.5 Presenting Data 48
 - 2.6 Some Final Practical Graphical Presentation Advice 51

2.7	Summary	55
	Key Terms	56
	Problems and Exercises	56
3	Analysis of Quantitative Data	59
3.1	Introduction	59
3.2	What Is Data Analysis?	60
3.3	Data Analysis Tools	61
3.4	Data Analysis for Two Data Sets	64
3.4.1	Time Series Data: Visual Analysis	66
3.4.2	Cross-Sectional Data: Visual Analysis	68
3.4.3	Analysis of Time Series Data: Descriptive Statistics	71
3.4.4	Analysis of Cross-Sectional Data: Descriptive Statistics	72
3.5	Analysis of Time Series Data: Forecasting/Data Relationship Tools	75
3.5.1	Graphical Analysis	76
3.5.2	Linear Regression	80
3.5.3	Covariance and Correlation	86
3.5.4	Other Forecasting Models	87
3.5.5	Findings	88
3.6	Analysis of Cross-Sectional Data: Forecasting/Data Relationship Tools	89
3.6.1	Findings	96
3.7	Summary	97
	Key Terms	98
	Problems and Exercises	99
4	Presentation of Qualitative Data—Data Visualization	103
4.1	Introduction—What Is Qualitative Data?	103
4.2	Essentials of Effective Qualitative Data Presentation	104
4.2.1	Planning for Data Presentation and Preparation	104
4.3	Data Entry and Manipulation	107
4.3.1	Tools for Data Entry and Accuracy	108
4.3.2	Data Transposition to Fit Excel	112
4.3.3	Data Conversion with the Logical IF	115
4.3.4	Data Conversion of Text from Non–Excel Sources	118
4.4	Data Queries with Sort, Filter, and Advanced Filter	121
4.4.1	Sorting Data	122
4.4.2	Filtering Data	124
4.4.3	Filter	125
4.4.4	Advanced Filter	129
4.5	An Example	133
4.6	Summary	139
	Key Terms	140
	Problems and Exercises	141

5	Analysis of Qualitative Data	145
5.1	Introduction	145
5.2	Essentials of Qualitative Data Analysis	147
5.2.1	Dealing with Data Errors	147
5.3	PivotChart or PivotTable Reports	151
5.3.1	An Example	151
5.3.2	PivotTables	153
5.3.3	PivotCharts	164
5.4	TiendaMía.com Example: Question 1	168
5.5	TiendaMía.com Example: Question 2	169
5.6	Summary	175
	Key Terms	175
	Problems and Exercises	176
6	Inferential Statistical Analysis of Data	179
6.1	Introduction	180
6.2	Let the Statistical Technique Fit the Data	181
6.3	χ^2 —Chi-Square Test of Independence for Categorical Data	181
6.3.1	Tests of Hypothesis—Null and Alternative	182
6.4	z-Test and t-Test of Categorical and Interval Data	186
6.5	An Example	186
6.5.1	z-Test: 2 Sample Means	189
6.5.2	Is There a Difference in Scores for SC Non-prisoners and EB Trained SC Prisoners?	190
6.5.3	t-Test: Two Samples Unequal Variances	193
6.5.4	Do Texas Prisoners Score Higher than Texas Non-prisoners?	193
6.5.5	Do Prisoners Score Higher Than Non-prisoners Regardless of the State?	195
6.5.6	How Do Scores Differ Among Prisoners of SC and Texas Before Special Training?	196
6.5.7	Does the EB Training Program Improve Prisoner Scores?	198
6.5.8	What If the Observations Means Are the Same, But We Do Not See Consistent Movement of Scores?	199
6.5.9	Summary Comments	199
6.6	Confidence Intervals for Sample Statistics	201
6.6.1	What Are the Ingredients of a Confidence Interval?	202
6.6.2	A Confidence Interval Example	203
6.6.3	Single Sample Hypothesis Tests Are Similar to Confidence Intervals	204
6.7	ANOVA	207
6.7.1	ANOVA: Single Factor Example	207
6.7.2	Do the Mean Monthly Losses of Reefers Suggest That the Means Are Different for the Three Ports?	209

6.8	Experimental Design	210
6.8.1	Randomized Complete Block Design Example	213
6.8.2	Factorial Experimental Design Example	216
6.9	Summary	219
	Key Terms	221
	Problems and Exercises	221
7	Modeling and Simulation: Part 1	225
7.1	Introduction	225
7.1.1	What Is a Model?	227
7.2	How Do We Classify Models?	229
7.3	An Example of Deterministic Modeling	231
7.3.1	A Preliminary Analysis of the Event	232
7.4	Understanding the Important Elements of a Model	235
7.4.1	Pre-modeling or Design Phase	236
7.4.2	Modeling Phase	236
7.4.3	Resolution of Weather and Related Attendance	240
7.4.4	Attendees Play Games of Chance	241
7.4.5	Fr. Efiia's What-if Questions	243
7.4.6	Summary of OLPS Modeling Effort	244
7.5	Model Building with Excel	245
7.5.1	Basic Model	246
7.5.2	Sensitivity Analysis	248
7.5.3	Controls from the Forms Control Tools	255
7.5.4	Option Buttons	256
7.5.5	Scroll Bars	259
7.6	Summary	261
	Key Terms	261
	Problems and Exercises	262
8	Modeling and Simulation: Part 2	265
8.1	Introduction	265
8.2	Types of Simulation and Uncertainty	267
8.2.1	Incorporating Uncertain Processes in Models	267
8.3	The Monte Carlo Sampling Methodology	268
8.3.1	Implementing Monte Carlo Simulation Methods	269
8.3.2	A Word About Probability Distributions	274
8.3.3	Modeling Arrivals with the Poisson Distribution	278
8.3.4	VLOOKUP and HLOOKUP Functions	280
8.4	A Financial Example—Income Statement	282
8.5	An Operations Example—Autohaus	286
8.5.1	Status of Autohaus Model	291
8.5.2	Building the Brain Worksheet	292
8.5.3	Building the Calculation Worksheet	294

8.5.4	Variation in Approaches to Poisson Arrivals: Consideration of Modeling Accuracy	296
8.5.5	Sufficient Sample Size	297
8.5.6	Building the Data Collection Worksheet	298
8.5.7	Results	303
8.6	Summary	307
	Key Terms	307
	Problems and Exercises	308
9	Solver, Scenarios, and Goal Seek Tools	311
9.1	Introduction	311
9.2	Solver–Constrained Optimization	313
9.3	Example–York River Archaeology Budgeting	314
9.3.1	Formulation	316
9.3.2	Formulation of YRA Problem	318
9.3.3	Preparing a Solver Worksheet	318
9.3.4	Using Solver	322
9.3.5	Solver Reports	323
9.3.6	Some Questions for YRA	328
9.4	Scenarios	334
9.4.1	Example 1—Mortgage Interest Calculations	334
9.4.2	Example 2—An Income Statement Analysis	337
9.5	Goal Seek	338
9.5.1	Example 1—Goal Seek Applied to the PMT Cell	339
9.5.2	Example 2—Goal Seek Applied to the CUMIPMT Cell	341
9.6	Summary	342
	Key Terms	343
	Problems and Exercises	344

About the Author



Hector Guerrero is a Professor Emeritus at Mason School of Business at the College of William and Mary, in Williamsburg, Virginia. He teaches in the areas of business analytics, decision making, statistics, operations, and business quantitative methods. He has previously taught at the Amos Tuck School of Business at Dartmouth College and the College of Business of the University of Notre Dame. He is well known among his students for his quest to bring clarity to complex decision problems.

He earned a PhD in Operations and Systems Analysis at the University of Washington and a BS in Electrical Engineering and an MBA at the University of Texas. He has published scholarly work in the areas of operations management, product design, and catastrophic planning.

Prior to entering academe, he worked as an engineer for Dow Chemical Company and Lockheed Missiles and Space Co. He is also very active in consulting and executive education with a wide variety of clients—U.S. Government, international firms, as well as many small and large U.S. manufacturing and service firms.

It is not unusual to find him relaxing on a quiet beach with a challenging Excel workbook and an excellent cabernet.