

# Business Statistics for Competitive Advantage with Excel 2016



Cynthia Fraser

# Business Statistics for Competitive Advantage with Excel 2016

Basics, Model Building, Simulation and Cases

Cynthia Fraser  
McIntire School of Commerce  
University of Virginia  
Charlottesville, VA, USA

ISBN 978-3-319-32184-4      ISBN 978-3-319-32185-1 (eBook)  
DOI 10.1007/978-3-319-32185-1

Library of Congress Control Number: 2016939396

© Springer International Publishing Switzerland 2016

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.

Printed on acid-free paper

This Springer imprint is published by Springer Nature  
The registered company is Springer International Publishing AG Switzerland

# Table of Contents

<b>Preface .....</b>	<b>xiii</b>
<b>Acknowledgements .....</b>	<b>xvi</b>
<b>Chapter 1    Statistics for Decision Making and Competitive Advantage .....</b>	<b>1</b>
1.1        Statistical Competences Translate into Competitive Advantages .....	1
1.2        The Path Toward Statistical Competence and Competitive Advantage .....	2
1.3        Use Excel for Competitive Advantage .....	2
1.4        Statistical Competence Is Powerful and Yours .....	3
<b>Chapter 2    Describing Your Data.....</b>	<b>5</b>
2.1        Describe Data with Summary Statistics and Histograms .....	5
2.2        Round Descriptive Statistics.....	9
2.3        Share the Story That Your Graphics Illustrate .....	9
2.4        Data Is Measured with Quantitative or Categorical Scales .....	10
2.5        Continuous Data Are Sometimes Normal.....	11
2.6        The Empirical Rule Simplifies Description.....	12
2.7        Outliers Can Distort the Picture.....	13
2.8        Central Tendency, Dispersion and Skewness Describe Data .....	14
2.9        Describe Categorical Variables Graphically .....	14
2.10       Descriptive Statistics Depend On the Data and Rely on Your Packaging.....	15
Excel 2.1   Produce Descriptive Statistics .....	17
Excel 2.2   Sort to Produce Descriptives Without Outliers .....	25
Excel 2.3   Plot a Cumulative Distribution .....	26
Excel 2.4   Use a PivotTable to Sort by Industry.....	29
Excel 2.5   Produce a Column Chart of a Nominal Variable .....	31
Excel Shortcuts Used in Chapter 2 .....	34
Significant Digits Guidelines.....	37
Lab 2 Description .....	39
Assignment 2.1 Procter & Gamble’s Global Advertising .....	41
Assignment 2.2 Best Practices Survey .....	42
Assignment 2.3 Shortcut Challenge .....	43
Case 2.1 VW Backgrounds.....	43
Case 2.2 Global Smelter Costs at Alcoa.....	43

<b>Chapter 3</b>	<b>Hypothesis Tests, Confidence Intervals to Infer Population Characteristics and Differences .....</b>	<b>47</b>
3.1	Sample Means Are Random Variables.....	47
3.2	Infer Whether a Population Mean Exceeds a Target .....	51
3.3	Critical t Provides a Benchmark .....	53
3.4	Confidence Intervals Estimate the Population Mean.....	54
3.5	Calculate Approximate Confidence Intervals with Mental Math.....	56
3.6	Margin of Error Is Inversely Proportional to Sample Size.....	57
3.7	Determine Whether Two Segments Differ with Student t.....	58
3.8	Estimate the Extent of Difference Between Two Segments.....	62
3.9	Estimate a Population Proportion from a Sample Proportion .....	63
3.10	Conditions for Assuming Approximate Normality .....	65
3.11	Conservative Confidence Intervals for a proportion .....	65
3.12	Assess the Difference Between Alternate Scenarios or Pairs .....	67
3.13	Inference from Sample to Population.....	71
Excel 3.1	Test the Level of a Population Mean with a One Sample t test. ....	73
Excel 3.2	Make a Confidence Interval for a Population Mean.....	74
Excel 3.3	Illustrate Confidence Intervals with Column Charts .....	75
Excel 3.4	Test the Difference Between Two Segment Means with a Two Sample t test.....	80
Excel 3.5	Construct a Confidence Interval for the Difference Between Two Segments .....	81
Excel 3.6	Illustrate the Difference Between Two Segment Means with a Column Chart.....	84
Excel 3.7	Construct a Pie Chart of Shares.....	85
Excel 3.8	Test the Difference in Between Alternate Scenarios or Pairs with a Paired t test.....	87
Excel 3.9	Construct a Confidence Interval for the Difference Between Alternate Scenarios or Pairs .....	88
Lab 3.1	Inference .....	89
	Cingular’s Position in the Cell Phone Service Market .....	89
	Value of a Nationals Uniform .....	89
	Confidence in Chinese Imports .....	90
Lab 3.2	Inference: Dell Smartphone Plans .....	91
Assignment 3.1	The Marriott Difference.....	93
Assignment 3.2	Immigration in the U.S. ....	93
Assignment 3.3	McLattes .....	94
Assignment 3.4	A Barbie Duff in Stuff .....	94
Assignment 3.5	Alcoa Smelters.....	94
Case 3.1	Yankees v Marlins: The Value of a Yankee Uniform.....	97
Case 3.2	Gender Pay.....	97
Case 3.3	Polaski Vodka: Can a Polish Vodka Stand Up to the Russians? .....	98

<b>Chapter 4</b>	<b>Simulation to Infer Future Performance Levels Given Assumptions .....</b>	<b>101</b>
4.1	Specify Assumptions Concerning Future Performance Drivers.....	101
4.2	Compare Best and Worst Case Performance Outcomes.....	105
4.3	Spread and Shape Assumptions Influence Possible Outcomes .....	106
4.4	Monte Carlo Simulation of the Distribution of Performance Outcomes .....	107
4.5	Monte Carlo Simulation Reveals Possible Outcomes Given Assumptions .....	112
Excel 4.1	Set Up a Spreadsheet to Link Simulated Performance Components .....	113
Excel 4.2	View a Simulated Sample with a Histogram.....	115
	Lab 4 Inference: Dell Android Smartphone Plans.....	131
	Case 4.1 American Girl in Starbucks.....	133
	Case 4.2 Can Whole Foods Hold On? .....	133
	Case 4.3 Chipotle's Ambitions to Triple Share of Top 100 Chain Sales in the Recession Rebound.....	135
<b>Chapter 5</b>	<b>Simple Regression for Long Range Forecasts.....</b>	<b>137</b>
5.1	The Simple Linear Regression Equation Describes the Line Relating an Independent Variable to Performance .....	138
5.2	Hide the Two Most Recent Datapoints to Validate a Time Series Model.....	138
5.3	Test and Infer the Slope .....	141
5.4	The Regression Standard Error Reflects Model Precision .....	144
5.5	Prediction Intervals Estimate Average Population Response.....	145
5.6	<i>Rsquare</i> Summarizes Strength of the Hypothesized Linear Relationship and <i>F</i> Tests Its Significance.....	146
5.7	Assess Residuals to Learn Whether Assumptions Are Met .....	149
5.8	Recalibrate to Update a Valid Model .....	151
5.9	Present Regression Results in Concise Format .....	153
5.10	Assumptions We Make When We Use Linear Regression .....	154
5.11	Correlation Reflects Linear Association.....	154
5.12	Correlation Coefficients Are Key Components of Regression Slopes .....	157
5.13	Correlation Complements Regression .....	158
5.14	Linear Regression Is Doubly Useful .....	158
Excel 5.1	Build a Simple Linear Regression Model .....	159
Excel 5.2	Assess Residuals .....	160
Excel 5.3	Construct Prediction Intervals to Validate .....	162
Excel 5.4	Recalibrate and Present Fit and Forecast in a Scatterplot.....	165
Excel 5.5	Find Correlations Between Variable Pairs .....	170
	Lab 5 Forecast Concha y Toro Exports to Latin America .....	171
	Assignment 5.1 Forecast Concha y Toro Exports to Europe and Asia.....	173
<b>Chapter 6</b>	<b>Consolidating Multiple Naïve Forecasts with Monte Carlo .....</b>	<b>175</b>
6.1	Use Monte Carlo to Integrate Multiple Uncertain Naïve Forecasts .....	176
6.2	Monte Carlo Offers Likely Possibilities from Consolidated Multiple Naïve Forecasts .....	177

Excel 6.1	Use Monte Carlo to Produce a 95% Prediction Interval of Consolidated Possibilities from Multiple Naïve Forecasts .....	178
	Lab 6 Forecast Concha y Toro Consolidated Exports to the New World.....	181
	Assignment 6 Forecast Concha y Toro Consolidated Exports Worldwide .....	183
	Case 6 Can Arcos Dorados Hold On? .....	185
<b>Chapter 7</b>	<b>Presenting Statistical Analysis Results to Management .....</b>	<b>187</b>
7.1	Use PowerPoints to Present Statistical Results for Competitive Advantage.....	187
7.2	Write Memos that Encourage Your Audience to Read and Use Results .....	194
	MEMO Re: Worldwide exports forecast to grow modestly through 2016.....	196
	Case 7 Segmentation of the Market for Preemie Diapers.....	199
	The Market for Preemie Diapers .....	200
	Preemie Parent Segments.....	200
	The Concept Test.....	201
	Data Recoding .....	202
<b>Chapter 8</b>	<b>Finance Application: Portfolio Analysis with a Market Index as a Leading Indicator in Simple Linear Regression .....</b>	<b>207</b>
8.1	Rates of Return Reflect Expected Growth of Stock Prices .....	207
8.2	Investors Trade Off Risk and Return.....	209
8.3	Beta Measures Risk .....	209
8.4	A Portfolio Expected Return, Risk and Beta Are Weighted Averages of Individual Stocks.....	213
8.5	Better Portfolios Define the Efficient Frontier.....	214
	MEMO Re: Recommended Portfolio is Diversified.....	216
8.6	Portfolio Risk Depends on Correlations with the Market and Stock Variability .....	217
Excel 8.1	Estimate Portfolio Expected Rate of Return and Risk.....	218
Excel 8.2	Plot Return by Risk to Identify Dominant Portfolios and the Efficient Frontier.....	220
	Lab 8 Portfolio Risk and Return .....	225
	Assignment 8 Portfolio Risk and Return .....	227
<b>Chapter 9</b>	<b>Association Between Two Categorical Variables: Contingency Analysis with Chi Square.....</b>	<b>229</b>
9.1	When Conditional Probabilities Differ from Joint Probabilities, There Is Evidence of Association .....	229
9.2	Chi Square Tests Association Between Two Categorical Variables .....	231
9.3	Chi Square Is Unreliable If Cell Counts Are Sparse .....	233
9.4	Simpson’s Paradox Can Mislead.....	235
	MEMO Re.: Country of Assembly Does Not Affect Older Buyers’ Choices .....	240
9.5	Contingency Analysis Is Demanding .....	241
9.6	Contingency Analysis Is Quick, Easy, and Readily Understood.....	241

Excel 9.1	Construct Crosstabulations and Assess Association Between Categorical Variables with PivotTables and PivotCharts .....	242
Excel 9.2	Use Chi Square to Test Association .....	244
Excel 9.3	Conduct Contingency Analysis with Summary Data .....	246
	Lab 9 Skype Appeal.....	251
	Assignment 9.1 Wine Preferences by Global Region.....	253
	Assignment 9.2 Fit Matters.....	253
	Assignment 9.3 Netbooks in Color.....	253
	Case 9.1 Hybrids for American Car.....	255
	Case 9.2 Tony's GREAT Advertising .....	255
	Case 9.3 Hybrid Motivations .....	256
<b>Chapter 10</b>	<b>Building Multiple Regression Models.....</b>	<b>259</b>
10.1	Explanatory Multiple Regression Models Identify Drivers and Forecast .....	259
10.2	Use Your Logic to Choose Model Components.....	260
10.3	Multicollinear Variables Are Likely When Few Variable Combinations Are Popular in a Sample .....	263
10.4	<i>F</i> Tests the Joint Significance of the Set of Independent Variables .....	263
10.5	Insignificant Parameter Estimates Signal Multicollinearity .....	265
10.6	Combine or Eliminate Collinear Predictors.....	267
10.7	Decide Whether Insignificant Drivers Matter .....	272
10.8	Sensitivity Analysis Quantifies the Marginal Impact of Drivers.....	274
	MEMO Re: Light, responsive, fuel efficient cars with smaller engines are cleanest.....	277
10.9	Model Building Begins With Logic and Considers Multicollinearity.....	278
Excel 10.1	Build and Fit a Multiple Linear Regression Model .....	279
Excel 10.2	Use Sensitivity Analysis to Compare the Marginal Impacts of Drivers .....	284
	Lab 10 Model Building with Multiple Regression: Pricing Dell's Navigreat.....	293
	Assignment 10.1 Sakura Motor's Quest for Fuel Efficiency.....	297
	Case 10.1 Fast Food Nations .....	299
	Case 10.2 Chasing Chipotle's Success .....	299
	Case 10.3 Costco's Warehouse Location Scheme.....	301
<b>Chapter 11</b>	<b>Indicator Variables.....</b>	<b>303</b>
11.1	Indicators Modify the Intercept to Account for Segment Differences .....	303
11.2	Indicators Estimate the Value of Product Attributes .....	306
11.3	Indicators Estimate Segment Mean Differences.....	310
11.4	Analysis of Variance Offers an Alternative to Regression with Indicators.....	314
11.5	ANOVA and Regression with Indicators Are Complementary Substitutes .....	318
11.6	ANOVA and Regression in Excel .....	319
Excel 11.1	Use Indicators to Find Part Worths and Attribute Importances.....	320
Excel 11.2	Use ANOVA to Test Equivalence of Mean Interest Ratings .....	325
	Lab 11.1 Revere Bank Profits.....	329
	Lab 11.2 Power PowerPoints.....	331

	Lab 11.3 ANOVA and Regression with Indicators: Powerful PowerPoints .....	333
	Assignment 11 Forecasting Chipotle Revenue in the Long Range .....	335
	Case 11 Store24 (A): Managing Employee Retention and Store24 (B): Service Quality and Employee Skills .....	337
<b>Chapter 12</b>	<b>Model Building and Forecasting with Multicollinear Time Series .....</b>	<b>339</b>
12.1	Time Series Models Include Decision Variables, External Forces, and Leading Indicators .....	342
12.2	Indicators of Economic Prosperity Lead Business Performance .....	343
12.3	Hide the Two Most Recent Datapoints to Validate a Time Series Model.....	343
12.4	Compare Scatterplots to Choose Driver Lags: Visual Inspection .....	344
12.5	Assess Residuals to Identify Unaccounted for Trend or Cycles.....	347
12.6	Forecast the Recent, Hidden Points to Assess Predictive Validity.....	352
12.7	Add the Most Recent Datapoints to Recalibrate.....	352
12.8	Compare Part Worths to Assess Driver Importances .....	354
	MEMO Re: Slow, Stable Growth Forecast in Next Four Quarters .....	355
12.9	Leading Indicator Components Are Powerful Drivers and Often Multicollinear .....	356
Excel 12.1	Build and Fit a Multiple Regression Model with Multicollinear Time Series.....	358
Excel 12.2	Create Potential Driver Lags .....	360
Excel 12.3	Select the Most Promising Driver .....	362
Excel 12.4	Plot Residuals to Identify Unaccounted for Trend, Cycles, or Seasonality and Assess Autocorrelation .....	364
Excel 12.5	Test the Model's Forecasting Validity.....	371
Excel 12.6	Recalibrate to Forecast.....	373
Excel 12.7	Illustrate the Fit and Forecast.....	374
Excel 12.8	Assess the Impact of Drivers. ....	375
	Lab 12.1 What Is Driving WFM Revenues... and What Revenues Can WFM Expect Next Year? .....	379
	Lab 12.2 What Is Driving WFM Revenues... and What Revenues Can WFM Expect Next Year? .....	383
	Case 12 McDonalds Revenue Drivers and Future Prospects.....	385
	Case 12.1 Chipotle Quarterly Revenues Model and Forecast .....	390
<b>Chapter 13</b>	<b>Nonlinear Multiple Regression Models .....</b>	<b>395</b>
13.1	Consider a Nonlinear Model When Response Is Not Constant.....	395
13.2	Skewness Signals Nonlinear Response .....	395
13.3	Rescaling $y$ Builds in Interactions .....	399
13.4	The Margin of Error Is Not Constant with a Nonlinear Model .....	404
13.5	Sensitivity Analysis Enables Scenario Comparisons .....	404
13.6	Nonlinear Models Inform Monte Carlo Simulation .....	410
13.7	Gains from Nonlinear Rescaling Are Significant.....	411
13.8	Nonlinear Models Offer the Promise of Better Fit and Better Behavior .....	412
Excel 13.1	Rescale to Build and Fit Nonlinear Regression Models with Linear Regression .....	413
Excel 13.2	Compare Scenarios with Sensitivity Analysis.....	427

Excel 13.3	Use Nonlinear Regression Estimates with Monte Carlo Simulation.....	431
Lab 13.1	Nonlinear Forecasting LAN Airlines Passenger Revenues: Building the Model.....	437
Lab 13.2	Nonlinear Forecasting LAN Airlines Passenger Revenues: Describe the Model.....	439
Lab 13.3	Forecasting with Uncertain Drivers: LAN Passenger Revenues.....	441
Assignment 13.1	Billionaires in 2020.....	443
Assignment 13.2	Primary Aluminum Production in 2020.....	445
<b>Chapter 14</b>	<b>Nonlinear Explanatory Multiple Regression Models.....</b>	<b>447</b>
14.1	Sensitivity Analysis Reveals the Relative Strength of Drivers.....	451
14.2	Sensitivity Analysis with Nonlinear Models Reveals Interactions.....	453
Excel 14.1	Build a Nonlinear Model with Cross Sectional Data.....	454
Excel 14.2	Sensitivity Analysis of Scenarios and Driver Influence.....	458
Lab 14	Mattel’s Acquisition of Radica.....	463
Assignment 14	Identifying Promising Global Markets.....	465
Case 14.1	Promising Global Markets for EVs.....	467
Case 14.2	Chasing Whole Foods’ Success.....	469
Case 14.3	Promising Global Markets for Water Purification.....	471
<b>Index</b>	.....	<b>473</b>



# Preface

Exceptional managers know that they can create competitive advantages by basing decisions on performance response under alternative scenarios. To create these advantages, managers need to understand how to use statistics to provide information on performance response under alternative scenarios. Statistics are created to make better decisions. Statistics are essential and relevant. Statistics must be easily and quickly produced using widely available software, Excel. Then results must be translated into general business language and illustrated with compelling graphics to make them understandable and usable by decision makers. This book helps students master this process of using statistics to create competitive advantages as decision makers.

Statistics are essential, relevant, easy to produce, easy to understand, valuable, and a powerful source of competitive advantage.

## The examples, assignments, and cases used to illustrate statistics for decision making come from business problems

McIntire Corporate Sponsors and Partners, such as Alcoa, Rolls-Royce, Procter & Gamble, and Dell, and the industries that they do business in, provide many realistic examples. The book also features a number of examples of global business problems, including those from important emerging markets in China, India, and Chile. Students are excited when statistics are used to study real and important business problems. This makes it easy to see how they will use statistics to create competitive advantages in their internships and careers.

## Learning is hands on with Excel and shortcuts

Each type of analysis is introduced with one or more examples. Following is an example of how to create the statistics in Excel, and what the numbers mean in English.

Included in Excel sections are screenshots which allow students to easily master Excel. Featured are a number of popular Excel shortcuts, which are, themselves, a competitive advantage.

Powerful PivotTables and PivotCharts are introduced early and used throughout the book. Results are illustrated with graphics from Excel.

In each chapter, assignments or cases are included to allow students to practice using statistics for decision making and competitive advantage. Beginning in Chapter 11, Harvard Business School cases are suggested which provide additional opportunities to use statistics to advantage.

## Focus is on what statistics mean to decision makers and how to communicate results

From the beginning, results are translated into English. In Chapter 7, results are condensed and summarized in PowerPoints and memos, the standards of communication in businesses. Later

chapters include example memos for students to use as templates, making communication of statistics for decision making an easy skill to master.

Instructors, give your students the powerful skills that they will use to create competitive advantages as decision makers. Students, be prepared to discover that statistics are a powerful competitive advantage. Your mastery of the essential skills of creating and communicating statistics for improved decision making will enhance your career and make numbers fun.

## New in the Fourth Edition

The financial and economic events of 2008–2010 changed business dramatically. Examples have been updated to illustrate how the impacts of recent changes can be acknowledged to build powerful, valid models.

Global examples include analysis of several emerging markets multinationals in Chile. New emerging markets create unique opportunities for global business, and the Fourth Edition moves beyond the BRICs to explore these.

## Acknowledgements

First, Second and Third editions of *Business Statistics for Competitive Advantage* were used in the Integrated Core Curriculum at The McIntire School, University of Virginia, and I thank the many bright, motivated and enthusiastic students who provided comments and suggestions.

*Cynthia Fraser*  
*Charlottesville, VA*