

JMP[®] Start Statistics

A Guide to Statistics and
Data Analysis Using JMP[®]

Sixth Edition



John Sall, Ann Lehman, Mia Stephens, Sheila Loring

The correct bibliographic citation for this manual is as follows: John Sall, Ann Lehman, Mia Stephens, and Sheila Loring. *JMP® Start Statistics: A Guide to Statistics and Data Analysis Using JMP®, Sixth Edition*. Cary, NC: SAS Institute Inc.

JMP® Start Statistics: A Guide to Statistics and Data Analysis Using JMP®, Sixth Edition

Copyright © 2017, SAS Institute Inc., Cary, NC, USA

ISBN 978-1-62960-875-4 (Hardcopy)

ISBN 978-1-62960-876-1 (EPUB)

ISBN 978-1-62960-877-8 (MOBI)

ISBN 978-1-62960-878-5 (Web PDF)

All rights reserved. Produced in the United States of America.

For a hard-copy book: 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, or otherwise, without the prior written permission of the publisher, SAS Institute Inc.

For a Web download or e-book: Your use of this publication shall be governed by the terms established by the vendor at the time you acquire this publication.

The scanning, uploading, and distribution of this book via the Internet or any other means without the permission of the publisher is illegal and punishable by law. Please purchase only authorized electronic editions and do not participate in or encourage electronic piracy of copyrighted materials. Your support of others' rights is appreciated.

U.S. Government Restricted Rights Notice: The Software and its documentation is commercial computer software developed at private expense and is provided with RESTRICTED RIGHTS to the United States Government. Use, duplication or disclosure of the Software by the United States Government is subject to the license terms of this Agreement pursuant to, as applicable, FAR 12.212, DFAR 227.7202-1(a), DFAR 227.7202-3(a) and DFAR 227.7202-4 and, to the extent required under U.S. federal law, the minimum restricted rights as set out in FAR 52.227-19 (DEC 2007). If FAR 52.227-19 is applicable, this provision serves as notice under clause (c) thereof and no other notice is required to be affixed to the Software or documentation. The Government's rights in Software and documentation shall be only those set forth in this Agreement.

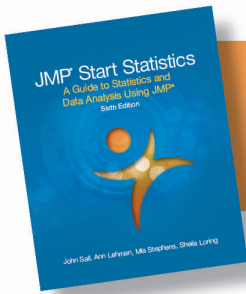
SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513-2414

February 2017

SAS® and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.

SAS software may be provided with certain third-party software, including but not limited to open-source software, which is licensed under its applicable third-party software license agreement. For license information about third-party software distributed with SAS software, refer to <http://support.sas.com/thirdpartylicenses>.



JMP® Start Statistics: A Guide to Statistics and Data Analysis Using JMP®, Sixth Edition. Full book available for purchase [here](#).

Table of Contents

Preface xv

The Software xv

How to Get JMP xvi

JMP Start Statistics, Sixth Edition xvii

SAS xvii

JMP versus JMP Pro xviii

This Book xviii

1 Preliminaries 1

What You Need to Know 1

...about statistics 1

Learning about JMP 1

...on your own with JMP Help 1

...hands-on examples 2

...using Tutorials 2

...reading about JMP 2

Chapter Organization 3

Typographical Conventions 5

2 Getting Started with JMP 7

Hello! 7

First Session 9

Tip of the Day 9

The JMP Starter (Macintosh) 9

The JMP Home Window (Windows) 10

Open a JMP Data Table 12

Launch an Analysis Platform 14

Interact with the Report Surface 15

Special Tools 18

Customize JMP 19

Modeling Type 21

Analyze and Graph 22

Navigating Platforms and Building Context 22

Contexts for a Histogram 23

- Contexts for the *t*-Test 23
- Contexts for a Scatterplot 24
- Contexts for Nonparametric Statistics 24
- The Personality of JMP 25

3 Data Tables, Reports, and Scripts 27

- Overview 27
- The Ins and Outs of a JMP Data Table 29
 - Selecting and Deselecting Rows and Columns 30
 - Mousing around a Data Table: Cursor Forms 30
- Creating a New JMP Table 32
 - Define Rows and Columns 33
 - Enter Data 35
 - The New Column Command 36
 - Plot the Data 37
- Importing Data 39
 - Importing Text Files 41
 - Importing Other File Types 44
 - Copy, Paste, and Drag Data 46
- Moving Data Out of JMP 47
- Saving Graphs and Reports 48
 - Copy and Paste 48
 - Drag Report Elements 49
 - Save JMP Reports and Graphs 49
 - Create Interactive Web Reports 49
 - Pop-up Menu Commands 50
- Juggling Data Tables 51
 - Data Management 51
 - Give New Shape to a Table: Stack Columns 52
- Creating Summary Statistics 55
 - Create Summary Statistics with the Summary Command 55
 - Create Summary Statistics with Tabulate 58
- Working with Scripts 60
 - Creating Scripts 60
 - Running Data Table Scripts 60
 - Opening and Running Stand-alone Scripts 61

4 Formula Editor 63

- Overview 63
- The Formula Editor Window 65
 - The Formula Editor and the JMP Scripting Language 66
 - A Quick Example: Standardizing Data 67

	Making a New Formula Column	69
	Using Popular Formula Functions	71
	Writing Conditional Expressions	72
	Summarizing Data with the Formula Editor	77
	Generating Random Data	82
	Local Variables and Table Variables	87
	Working with Dates	89
	Tips on Building Formulas	90
	Examining Expression Values	90
	Cutting, Dragging, and Pasting Formulas	90
	Selecting Expressions	91
	Exercises	91
5	What Are Statistics?	95
	Overview	95
	Ponderings	97
	The Business of Statistics	97
	The Yin and Yang of Statistics	97
	The Faces of Statistics	98
	Don't Panic	99
	Preparations	101
	Three Levels of Uncertainty	101
	Probability and Randomness	102
	Assumptions	102
	Data Mining?	103
	Statistical Terms	104
6	Simulations	109
	Overview	109
	Rolling Dice	111
	Rolling Several Dice	114
	Flipping Coins, Sampling Candy, or Drawing Marbles	114
	Probability of Making a Triangle	115
	Confidence Intervals	120
	Data Table-Based Simulations	121
	Other JMP Simulators	122
	Exercises	123
7	Univariate Distributions: One Variable, One Sample	125
	Overview	125
	Looking at Distributions	128

- Probability Distributions 130
 - True Distribution Function or Real-World Sample Distribution 131
 - The Normal Distribution 133
- Describing Distributions of Values 134
 - Generating Random Data 134
 - Histograms 135
 - Stem-and-Leaf Plots 137
 - Dot Plots 138
 - Outlier and Quantile Box Plots 139
 - Mean and Standard Deviation 141
 - Median and Other Quantiles 142
 - Mean versus Median 142
 - Other Summary Statistics: Skewness and Kurtosis 143
 - Extremes, Tail Detail 143
- Statistical Inference on the Mean 144
 - Standard Error of the Mean 144
 - Confidence Intervals for the Mean 144
 - Testing Hypotheses: Terminology 147
 - The Normal z -Test for the Mean 149
 - Case Study: The Earth's Ecliptic 150
 - Student's t -Test 152
 - Comparing the Normal and Student's t Distributions 153
 - Testing the Mean 154
 - The p -Value Animation 155
 - Power of the t -Test 157
- Practical Significance versus Statistical Significance 159
- Examining for Normality 161
 - Normal Quantile Plots 162
 - Statistical Tests for Normality 165
- Special Topic: Practical Difference 167
- Special Topic: Simulating the Central Limit Theorem 170
- Seeing Kernel Density Estimates 172
- Exercises 173

8 The Difference Between Two Means 177

- Overview 177
- Two Independent Groups 179
 - When the Difference Isn't Significant 179
 - Check the Data 180
 - Launch the Fit Y by X Platform 181
 - Examine the Plot 182
 - Display and Compare the Means 183
 - Inside the Student's t -Test 184

	Equal or Unequal Variances?	185
	One-Sided Version of the Test	187
	Analysis of Variance and the All-Purpose <i>F</i> -Test	188
	How Sensitive Is the Test?	
	How Many More Observations Are Needed?	190
	When the Difference Is Significant	192
	Normality and Normal Quantile Plots	194
	Testing Means for Matched Pairs	196
	Thermometer Tests	197
	Look at the Data	198
	Look at the Distribution of the Difference	199
	Student's <i>t</i> -Test	199
	The Matched Pairs Platform for a Paired <i>t</i> -Test	200
	Optional Topic:	
	An Equivalent Test for Stacked Data	203
	Two Extremes of Neglecting the Pairing Situation: A Dramatization	205
	A Nonparametric Approach	211
	Introduction to Nonparametric Methods	211
	Paired Means: The Wilcoxon Signed-Rank Test	211
	Independent Means: The Wilcoxon Rank Sum Test	213
	Exercises	214
9	Comparing Many Means: One-Way Analysis of Variance	217
	Overview	217
	What Is a One-Way Layout?	219
	Comparing and Testing Means	221
	Means Diamonds: A Graphical Description of Group Means	222
	Statistical Tests to Compare Means	223
	Means Comparisons for Balanced Data	226
	Means Comparisons for Unbalanced Data	227
	Adjusting for Multiple Comparisons	232
	Are the Variances Equal across the Groups?	235
	Testing Means with Unequal Variances	238
	Nonparametric Methods	239
	Review of Rank-Based Nonparametric Methods	239
	The Three Rank Tests in JMP	240
	Exercises	242
10	Fitting Curves through Points: Regression	245
	Overview	245
	Regression	247

- Least Squares 247
- Seeing Least Squares 248
- Fitting a Line and Testing the Slope 250
- Testing the Slope By Comparing Models 252
- The Distribution of the Parameter Estimates 255
- Confidence Intervals on the Estimates 256
- Examine Residuals 258
- Exclusion of Rows 258
- Time to Clean Up 260
- Polynomial Models 260
 - Look at the Residuals 261
 - Higher-Order Polynomials 261
 - Distribution of Residuals 262
- Transformed Fits 263
 - Spline Fit 265
- Are Graphics Important? 266
- Why It's Called Regression 269
- What Happens When X and Y Are Switched? 271
- Curiosities 274
 - Sometimes It's the Picture That Fools You 274
 - High-Order Polynomial Pitfall 275
 - The Pappus Mystery on the Obliquity of the Ecliptic 276
- Exercises 277

11 Categorical Distributions 281

- Overview 281
- Categorical Situations 283
- Categorical Responses and Count Data: Two Outlooks 283
- A Simulated Categorical Response 286
 - Simulating Some Categorical Response Data 287
 - Variability in the Estimates 288
 - Larger Sample Sizes 290
 - Monte Carlo Simulations for the Estimators 291
 - Distribution of the Estimates 292
- The X^2 Pearson Chi-Square Test Statistic 293
- The G^2 Likelihood-Ratio Chi-Square Test Statistic 294
 - Likelihood Ratio Tests 295
 - The G^2 Likelihood Ratio Chi-Square Test 296
- Univariate Categorical Chi-Square Tests 296
 - Comparing Univariate Distributions 297
 - Charting to Compare Results 299

Exercises 301

12 Categorical Models 303

Overview 303

Fitting Categorical Responses to Categorical Factors: Contingency Tables 305

Testing with G^2 and X^2 Statistic 305

Looking at Survey Data 306

Car Brand by Marital Status 310

Car Brand by Size of Vehicle 311

Two-Way Tables: Entering Count Data 312

Expected Values under Independence 313

Entering Two-Way Data into JMP 314

Testing for Independence 314

If You Have a Perfect Fit 316

Special Topic: Correspondence Analysis— Looking at Data with Many Levels 318

Continuous Factors with Categorical Responses: Logistic Regression 321

Fitting a Logistic Model 321

Degrees of Fit 325

A Discriminant Alternative 326

Inverse Prediction 327

Polytomous (Multinomial) Responses: More Than Two Levels 330

Ordinal Responses: Cumulative Ordinal Logistic Regression 331

Surprise: Simpson's Paradox: Aggregate Data versus Grouped Data 334

Generalized Linear Models 337

Exercises 342

13 Multiple Regression 345

Overview 345

Parts of a Regression Model 347

Regression Definitions 347

A Multiple Regression Example 348

Residuals and Predicted Values 351

The Analysis of Variance Table 354

The Whole Model F -Test 354

Whole-Model Leverage Plot 355

Details on Effect Tests 356

Effect Leverage Plots 356

Collinearity 358

Exact Collinearity, Singularity, and Linear Dependency 362

The Longley Data: An Example of Collinearity 364

The Case of the Hidden Leverage Point 366
Mining Data with Stepwise Regression 369
Exercises 373

14 Fitting Linear Models 377

Overview 377
The General Linear Model 379
Types of Effects in Linear Models 380
Coding Scheme to Fit a One-Way ANOVA as a Linear Model 381
Regressor Construction 384
Interpretation of Parameters 385
Predictions Are the Means 385
Parameters and Means 385
Analysis of Covariance: Continuous and Categorical Terms in the Same Model 386
The Prediction Equation 389
The Whole-Model Test and Leverage Plot 390
Effect Tests and Leverage Plots 391
Least Squares Means 393
Lack of Fit 394
Separate Slopes: When the Covariate Interacts with a Categorical Effect 396
Two-Way Analysis of Variance and Interactions 400
Optional Topic: Random Effects and Nested Effects 406
Nesting 407
Repeated Measures 409
Method 1: Random Effects-Mixed Model 409
Method 2: Reduction to the Experimental Unit 414
Method 3: Correlated Measurements-Multivariate Model 416
Varieties of Analysis 418
Closing Thoughts 418
Exercises 419

15 Design of Experiments 421

Overview 421
Introduction 424
Key Concepts 424
JMP DOE 425
A Simple Design 426
The Experiment 426
Enter the Response and Factors 427
Define the Model 429
Is the Design Balanced? 432
Perform Experiment and Enter Data 432

Analyze the Model	433
Flour Paste Conclusions	439
Details of the Design: Confounding Structure	439
Using the Custom Designer	440
How the Custom Designer Works	440
Choices in the Custom Designer	441
An Interaction Model: The Reactor Data	442
Analyzing the Reactor Data	444
Where Do We Go from Here?	450
Some Routine Screening Examples	452
Main Effects Only (a Review)	452
All Two-Factor Interactions Involving a Single Factor	453
Alias Optimal Designs	455
Response Surface Designs	456
The Odor Experiment	456
Response Surface Designs in JMP	456
Analyzing the Odor Response Surface Design	458
Plotting Surface Effects	461
Specifying Response Surface Effects Manually	462
The Custom Designer versus the Response Surface Design Platform	463
Split-Plot Designs	464
The Box Corrosion Split-Plot Experiment	465
Designing the Experiment	465
Analysis of Split-Plot Designs	467
Design Strategies	471
DOE Glossary of Key Terms	472
Exercises	476

16 Bivariate and Multivariate Relationships 479

Overview	479
Bivariate Distributions	481
Density Estimation	481
Bivariate Density Estimation	482
Mixtures, Modes, and Clusters	484
The Elliptical Contours of the Normal Distribution	485
Correlations and the Bivariate Normal	487
Simulating Bivariate Correlations	487
Correlations across Many Variables	490
Bivariate Outliers	491
Outliers in Three and More Dimensions	494
Identify Variation with Principal Components Analysis	496
Principal Components for Six Variables	499

- How Many Principal Components? 501
- Discriminant Analysis 502
 - Canonical Plot 503
 - Discriminant Scores 504
 - Stepwise Discriminant Variable Selection 507
- Cluster Analysis 508
 - Hierarchical Clustering: How Does It Work? 508
 - A Real-World Example 511
- Some Final Thoughts 514
- Exercises 515

17 Exploratory Modeling 517

- Overview 517
- Recursive Partitioning (Decision Trees) 519
 - Growing Trees 521
 - Exploratory Modeling with Partition 528
 - Saving Columns and Formulas 530
- Neural Nets 531
 - A Simple Example 532
 - Modeling with Neural Networks 535
 - Saving Columns 535
 - Profiles in Neural 537
- Exercises 541

18 Control Charts and Capability 545

- Overview 545
- What Does a Control Chart Look Like 548
- Types of Control Charts 549
 - Variables Charts 550
 - Attributes Charts 551
 - Specialty Charts 551
- Control Chart Basics 551
- Control Charts for Variables Data 552
- Variables Charts Using Control Chart Builder 553
 - The Control Chart Builder Work Space 553
 - Control Chart Builder Examples 554
- Control Charts for Attributes Data 557
- Specialty Charts 560
 - Presummarize Charts 560
 - Levey-Jennings Charts 561
 - Uniformly Weighted Moving Average (UWMA) Charts 561

Exponentially Weighted Moving Average (EWMA) Chart	563
Capability Analysis	564
What Is Process Capability?	564
Capability for One Process Measurement	567
Capability for Many Process Measurements	569
Capability for Time-Ordered Data	572
A Few Words about Measurement Systems	574
Exercises	574

19 Mechanics of Statistics 577

Overview	577
Springs for Continuous Responses	579
Fitting a Mean	579
Testing a Hypothesis	580
One-Way Layout	581
Effect of Sample Size Significance	581
Effect of Error Variance on Significance	582
Experimental Design's Effect on Significance	583
Simple Regression	584
Leverage	585
Multiple Regression	586
Summary: Significance and Power	586
Mechanics of Fit for Categorical Responses	586
How Do Pressure Cylinders Behave?	587
Estimating Probabilities	588
One-Way Layout for Categorical Data	589
Logistic Regression	591

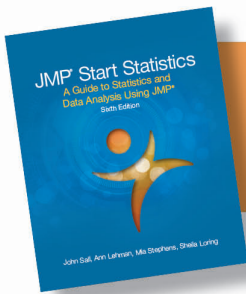
A Answers to Selected Exercises 593

Chapter 4, "Formula Editor"	593
Chapter 7, "Univariate Distributions: One Variable, One Sample"	596
Chapter 8, "The Difference Between Two Means"	600
Chapter 9, "Comparing Many Means: One-Way Analysis of Variance"	602
Chapter 10, "Fitting Curves through Points: Regression"	606
Chapter 11, "Categorical Distributions"	609
Chapter 12, "Categorical Models"	610
Chapter 13, "Multiple Regression"	612
Chapter 14, "Fitting Linear Models"	613
Chapter 15, "Design of Experiments"	614
Chapter 16, "Bivariate and Multivariate Relationships"	615
Chapter 17, "Exploratory Modeling"	617
Chapter 18, "Control Charts and Capability"	617

B References and Data Sources 619

Technology License Notices 625

Index 627



JMP® Start Statistics: A Guide to Statistics and Data Analysis Using JMP®, Sixth Edition. Full book available for purchase [here](#).

Preliminaries

What You Need to Know

...about statistics

This book is designed to help you learn about statistics. Even though JMP has many advanced features, you do not need a background of formal statistical training to use it. All analysis platforms include graphical displays with options that help you review and interpret the results. Each platform also includes access to Help that offers general guidance and appropriate statistical details.

Learning about JMP

...on your own with JMP Help

If you are familiar with Macintosh or Microsoft Windows software, you might want to proceed on your own. After you install JMP, you can open any of the JMP sample data files and experiment with analysis tools. Help is available for most menus, options, and reports.

There are several ways to access JMP Help:

- Select **JMP Help** from the **Help** menu.
- You can click the **Help** button in launch windows whenever you launch an analysis or graph platform.
- After you generate a report, click the Help tool (?) on the **Tools** menu or toolbar and click the report surface. Context-sensitive help tells about the items that you click.

...hands-on examples

This book describes JMP features and is reinforced with hands-on examples. By following these step-by-step examples, you can quickly become familiar with JMP menus, options, and report windows.

☞ Steps for example analyses begin with the mouse symbol in the margin, like this paragraph.

...using Tutorials

Tutorials interactively guide you through some common tasks in JMP and are accessible from the **Help > Tutorials** menu. We recommend that you complete the Beginners Tutorial as a quick introduction to the report features found in JMP.

...reading about JMP

JMP is accompanied by a series of built-in reference manuals, a menu reference card and a quick reference card. The newest in the series of guides, *Discovering JMP*, provides a general introduction to JMP. It contains basic examples and descriptions that give you a feel for JMP and can get you started.

Discovering JMP is followed by *Using JMP*, which helps new users understand JMP data tables and how to perform basic operations. *Using JMP* is followed by several books that document all of the JMP analysis and graph platforms. In addition, there are specialty books for design of experiments and the JMP scripting language. These references cover all the commands and options in JMP and have extensive examples of the **Analyze**, **Graph**, and **DOE** platforms.

The documentation is available in the following formats:

- In-product help (Select the **Help > JMP Help** menu.)
- PDF files (Select the **Help > Books** menu.)
- e-books (Amazon.com, Google Books, the Apple App Store, or the SAS Bookstore)
- Help at <http://jmp.com/support/help>
- Print books (Amazon.com or the SAS Bookstore)

Chapter Organization

The chapters of this book are supported by guided actions that you can take to become familiar with JMP.

The first five chapters get you quickly started with information about JMP tables, how to use the JMP formula editor, and give an overview of how to obtain results from the **Analyze** and **Graph** menus.

- Chapter 1, “Preliminaries,” is this introductory material.
- Chapter 2, “Getting Started with JMP,” tells you how to start and stop JMP, how to open data tables, and takes you on a short guided tour. You are introduced to the general personality of JMP. You see how data is handled by JMP. There is an overview of all analysis and graph commands; information about how to navigate a platform of results; and a description of the tools and options available for all analyses.
- Chapter 3, “Data Tables, Reports, and Scripts,” focuses on using the JMP data table. It shows how to create tables, subset, sort, and manipulate them with built-in menu commands, and how to get data and results out of JMP and into a report.
- Chapter 4, “Formula Editor,” covers the formula editor and quick ways to create formulas and derived variables. There is a description of the formula editor components and an overview of the extensive functions available for calculating column values.
- Chapter 5, “What Are Statistics?,” gives you some things to ponder about the nature and use of statistics. It also attempts to dispel statistical fears and phobias that are prevalent among students and professionals alike.

Chapters 6–19 cover the array of analysis techniques offered by JMP. Chapters begin with simple-to-use techniques and gradually work toward more complex methods. Emphasis is on learning to think about these techniques and on how to visualize data analysis at work. JMP offers a graph for almost every statistic and supporting tables for every graph. Using highly interactive methods, you can learn more quickly and discover what your data has to say.

- Chapter 6, “Simulations,” introduces you to some probability topics by using the JMP scripting language. You learn how to open and execute these scripts and to see other ways of simulating data in JMP.

- Chapter 7, “Univariate Distributions: One Variable, One Sample,” covers distributions of continuous and categorical variables and statistics to test univariate distributions.
- Chapter 8, “The Difference Between Two Means,” covers t tests of independent groups and tells how to handle paired data. The nonparametric approach to testing related pairs is also shown.
- Chapter 9, “Comparing Many Means: One-Way Analysis of Variance,” covers one-way analysis of variance, with standard statistics and a variety of graphical techniques.
- Chapter 10, “Fitting Curves through Points: Regression,” shows how to fit a regression model for a single factor.
- Chapter 11, “Categorical Distributions,” discusses how to think about the variability in single batches of categorical data. It covers estimating and testing probabilities in categorical distributions, shows Monte Carlo methods, and introduces the Pearson and Likelihood ratio chi-square statistics.
- Chapter 12, “Categorical Models,” covers fitting categorical responses to a model, starting with the usual tests of independence in a two-way table, and continuing with graphical techniques and logistic regression.
- Chapter 13, “Multiple Regression,” describes the parts of a linear model with continuous factors, talks about fitting models with multiple numeric effects, and shows a variety of examples, including the use of stepwise regression to find active effects.
- Chapter 14, “Fitting Linear Models,” is an advanced chapter that continues the discussion of Chapter 12. The chapter moves on to categorical effects and complex effects, such as interactions and nesting.
- Chapter 15, “Design of Experiments,” looks at the built-in commands in JMP used to generate specified experimental designs. It also looks at examples of how to analyze common screening and response-level designs are covered.
- Chapter 16, “Bivariate and Multivariate Relationships,” looks at ways to examine two or more response variables using correlations, scatterplot matrices, three-dimensional plots, principal components, and other techniques. Discriminant and Cluster Analysis discuss methods that group data into clumps. Outliers are discussed.
- Chapter 17, “Exploratory Modeling,” illustrates common data mining techniques—Neural Nets and Recursive Partitioning.

- Chapter 18, “Control Charts and Capability,” discusses common types of control charts for both continuous and attribute data, and introduces process capability studies.
- Chapter 19, “Mechanics of Statistics,” is an essay about statistical fitting that might prove enlightening to those who enjoy mechanics.

Typographical Conventions

The following conventions help you relate written material in this book to information that you see on your screen.

- Reference to menu names (**File** menu) or menu items (**Save** command), and buttons on windows (**OK**), appear in the **Helvetica bold** font.
- When you are asked to select a command from a submenu, such as **File > Save As**, go to the **File** menu and select the **Save As** command.
- Likewise, items on menus in reports are shown in **Helvetica bold**, but you are given a more detailed instruction about where to find the command or option. For example, you might be asked to select the **Show Points** option from the red triangle menu on the analysis title bar. You might select the **Save Predicted** command from the **Fitting** menu on the scatterplot title bar. Each menu is always visible as a small red triangle on the platform or on its outline title bars, as circled below.



- References to variable names, data table names, and some items in reports appear in Helvetica but can appear in illustrations in either a plain or boldface font. These items show on your screen as you have specified in your JMP Preferences.
- Words or phrases that are important, new, or have definitions specific to JMP are in *italics* the first time you see them.

- When there is an action statement, you can do the example yourself by following the instructions. These statements are preceded by a mouse symbol (⌘) in the margin. An example of an action statement is:
 - ⌘ Highlight the Month column by clicking the area above the column name, and then select **Cols > Column Info**.
- Occasionally, special information is in a boxed side bar in Helvetica to help distinguish them from the text flow.



Preface

JMP is statistical discovery software. JMP helps you explore data, fit models, discover patterns, and discover points that don't fit patterns. This book is a guide to statistics using JMP.

The Software

As statistical discovery software, JMP emphasizes working interactively with data and graphics in a progressive structure to make discoveries.

- With graphics, you are more likely to make discoveries. You are also more likely to understand the results.
- With interactivity, you are encouraged to dig deeper and try out more things that might improve your chances of discovering something important. With interactivity, one analysis leads to a refinement, and one discovery leads to another discovery.
- With a progressive structure, you build a context that maintains a live analysis. You don't have to redo analyses and plots to make changes in them, so details come to attention at the right time.

The purpose of JMP software is to create a virtual workplace. The software has facilities and platforms where the tools are located and the work is performed. JMP provides the workplace that we think is best for the job of analyzing data. With the right software workplace, researchers embrace computers and statistics, rather than avoid them.

JMP aims to present a graph with every statistic. You should always see the analysis in both ways, with statistical text and graphics, without having to ask for it. The text and graphs stay together.

JMP is controlled largely through point-and-click mouse manipulation. If you place the pointer over a point, JMP identifies it. If you click on a point in a plot, JMP highlights the point in the plot and highlights the point in the data table. In fact, JMP highlights the point everywhere it is represented.

JMP has a progressive organization. You begin with a simple report at the top, and as you analyze, more and more depth is revealed. The analysis is alive, and as you dig deeper into the data, more and more options are offered according to the context of the analysis.

In JMP, completeness is not measured by the “feature count,” but by the range of possible applications, and the orthogonality of the tools. In JMP, you get a feeling of being in more control despite your having less awareness of the control surface. You also get a feeling that statistics is an orderly discipline that makes sense, rather than an unorganized collection of methods.

A statistical software application is often the point of entry into the practice of statistics. JMP strives to offer fulfillment rather than frustration, empowerment rather than intimidation.

If you give someone a large truck, they will find someone to drive it for them. But if you give them a sports car, they will learn to drive it themselves. We believe that statistics can be interesting and reachable so that people will want to drive that vehicle.

How to Get JMP

There are several ways to get JMP:

- JMP is available through department or campus licenses at most colleges and universities and through site licenses in many organizations. See your software IT administrator for availability and download information.
- Individual copies of JMP for academic use are also available from <http://onthehub.com/jmp>. If you would like more information about academic licensing or would like to request an evaluation copy of JMP for classroom use, email academic@jmp.com.

- If you do not qualify for an academic license, a trial version of JMP is available at <http://jmp.com/trial>. Read license information at <http://jmp.com/buy>.

JMP Start Statistics, Sixth Edition

JMP Start Statistics has been updated and revised to feature JMP 13. Major enhancements have been made to JMP since the fifth edition, which was based on JMP 10. The new enhancements include DOE (Design Evaluation, new Custom Design options, and Definitive Screening designs), analysis and modeling (Generalized Regression, Partition enhancements, Model Comparison, and Formula Depot), data preparation (handling missing values and outliers, and model validation), and graphics (continued development of the interactive Graph Builder), most of which are covered in this book. In addition, the menus have been restructured, and we've added functionality for getting data into JMP (Query Builder) and sharing results (saving as Microsoft PowerPoint, saving as HTML, and creating interactive web reports).

JMP 13 also continues our focus on enhancing the user experience, with new daily Tips of the Day and expanded documentation.

We include discussion of many of these new features throughout this text.

SAS

SAS, or the SAS System, is an integrated statistical software system used by universities, research institutions, and industries across the globe. JMP Statistical Discovery Software is desktop software from SAS that runs natively on Mac and Windows. JMP was originally designed as a personal analysis tool for engineers and scientists, but is now used in a variety of applications and industries worldwide.

JMP versus JMP Pro

JMP was first released by SAS in 1989 to run on a Macintosh operating system, and became available on Windows in the early 1990s. Since then, JMP has grown into a family of products, each designed to meet particular needs.

In this book we use JMP Pro, which includes advanced tools for analytics and predictive modeling. However, JMP Pro is not required to take full advantage of the methods covered. Unless otherwise specified, the features that we discuss are available in both JMP and JMP Pro.

This Book

Software Manual and Statistics Text

This book is a mix of software manual and statistics text. It is designed to be a complete and orderly introduction to analyzing data. It is a teaching text, but is especially useful when used in conjunction with a standard statistical textbook.

Not Just the Basics

A few of the techniques in this book are not found in most introductory statistics courses, but are accessible in basic form using JMP. These techniques include logistic regression, correspondence analysis, principal components with biplots, leverage plots, and density estimation. All these techniques are used in the service of understanding other, more basic methods. Where appropriate, supplemental material is labeled as “Special Topics” so that it is recognized as optional material.

JMP also includes several advanced methods not covered in this book, such as nonlinear regression, multivariate analysis of variance, tools for predictive modeling and data mining, consumer research methods, text mining, and some advanced design of experiments capabilities. If you are planning to use these features extensively, it is recommended that you refer to the Help system or the JMP documentation for the professional version of JMP.

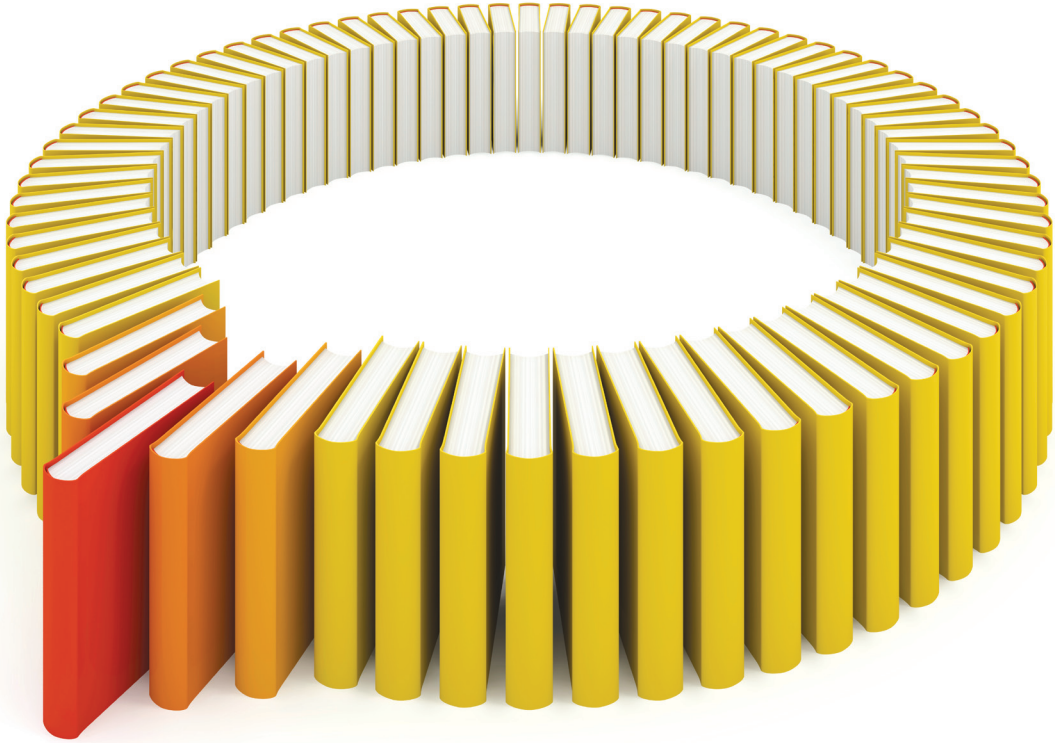
Examples Both Real and Simulated

Most examples are real-world applications. A few simulations are included too, so that the difference between a true value and its estimate can be discussed, along with the variability in the estimates. Some examples are unusual and are calculated to emphasize an important concept. The data for the examples are installed with JMP, with step-by-step instructions in the text. The same data are

also available on the Internet at <http://support.sas.com/stephens>. JMP can also import data from files that are distributed with other textbooks. See Chapter 3, “Data Tables, Reports, and Scripts,” for details about importing various types of data.

Acknowledgments

Thank you to the JMP testers as well as the contributors and reviewers of earlier versions of *JMP Start Statistics*: Bradley Jones, Chris Gotwalt, Lou Valente, Tom Donnelly, Michael Benson, Avignor Cahaner, Howard Yetter, David Ikle, Robert Stine, Andy Mauromoustkos, Al Best, Jacques Goupy, and Chris Olsen for contributions to earlier versions of the book. Special thanks to Curt Hinrichs for invaluable support to the JMP Start Statistics project.



Gain Greater Insight into Your SAS[®] Software with SAS Books.

Discover all that you need on your journey to knowledge and empowerment.

 support.sas.com/bookstore
for additional books and resources.


THE POWER TO KNOW.®