



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
<i>JMP Start Statistics, Sixth Edition</i>	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