

LEARNING MADE EASY



5th Edition

# Statistical Analysis with Excel<sup>®</sup>

for  
**dummies**<sup>®</sup>  
A Wiley Brand



Perform statistical analyses  
on Windows<sup>®</sup>, Mac<sup>®</sup>, and iPad<sup>®</sup>

Download the book's spreadsheets  
to accelerate your learning

Understand common (and  
uncommon) statistical terms

**Joseph Schmuller, PhD**

Teaching statistics through the use  
of Excel since 2005



6167 23971

012585087

122706122

# Statistical Analysis with Excel<sup>®</sup>



by Joseph Schmuller, PhD

for  
**dummies**<sup>®</sup>  
A Wiley Brand

# Table of Contents

<b>INTRODUCTION</b> .....	1
About This Book .....	2
What's New in This Edition .....	2
What's New in Excel (Microsoft 365) .....	3
Foolish Assumptions .....	3
Icons Used in This Book .....	4
Where to Go from Here .....	5
Beyond This Book .....	5
<b>PART 1: GETTING STARTED WITH STATISTICAL ANALYSIS WITH EXCEL: A MARRIAGE MADE IN HEAVEN</b> .....	7
<b>CHAPTER 1: Evaluating Data in the Real World</b> .....	9
The Statistical (and Related) Notions You Just Have to Know .....	9
Samples and populations .....	10
Variables: Dependent and independent .....	11
Types of data .....	12
A little probability .....	13
Inferential Statistics: Testing Hypotheses .....	14
Null and alternative hypotheses .....	15
Two types of error .....	16
Some Excel Fundamentals .....	18
Autofilling cells .....	22
Referencing cells .....	25
<b>CHAPTER 2: Understanding Excel's Statistical Capabilities</b> .....	29
Getting Started .....	30
Setting Up for Statistics .....	32
Worksheet functions .....	32
Quickly accessing statistical functions .....	36
Array functions .....	38
What's in a name? An array of possibilities .....	41
Creating Your Own Array Formulas .....	50
Using data analysis tools .....	51
Additional data analysis tool packages .....	56
Accessing Commonly Used Functions .....	58
The New Analyze Data Tool .....	59
Data from Pictures! .....	60

<b>PART 2: DESCRIBING DATA</b> .....	63
<b>CHAPTER 3: Show-and-Tell: Graphing Data</b> .....	65
Why Use Graphs? .....	65
Examining Some Fundamentals .....	67
Gauging Excel's Graphics (Chartics?) Capabilities .....	68
Becoming a Columnist .....	69
Stacking the Columns .....	73
Slicing the Pie .....	74
A word from the wise .....	76
Drawing the Line .....	77
Adding a Spark .....	80
Passing the Bar .....	82
The Plot Thickens .....	84
Finding Another Use for the Scatter Chart .....	88
<b>CHAPTER 4: Finding Your Center</b> .....	91
Means: The Lore of Averages .....	91
Calculating the mean .....	92
AVERAGE and AVERAGEA .....	93
AVERAGEIF and AVERAGEIFS .....	95
TRIMMEAN .....	99
Other means to an end .....	100
Medians: Caught in the Middle .....	102
Finding the median .....	102
MEDIAN .....	103
Statistics à la Mode .....	104
Finding the mode .....	104
MODE.SNGL and MODE.MULT .....	104
<b>CHAPTER 5: Deviating from the Average</b> .....	107
Measuring Variation .....	108
Averaging squared deviations: Variance and how to calculate it .....	108
VAR.P and VARPA .....	111
Sample variance .....	113
VAR.S and VARA .....	114
Back to the Roots: Standard Deviation .....	114
Population standard deviation .....	115
STDEV.P and STDEVP .....	115
Sample standard deviation .....	116
STDEV.S and STDEVA .....	116
The missing functions: STDEVIF and STDEVIFS .....	117

Related Functions	121
DEVSQ	121
Average deviation	122
AVEDEV	123
<b>CHAPTER 6: Meeting Standards and Standings</b>	<b>125</b>
Catching Some Z's	126
Characteristics of z-scores	126
Bonds versus the Bambino	127
Exam scores	128
STANDARDIZE	128
Where Do You Stand?	131
RANK.EQ and RANK.AVG	131
LARGE and SMALL	133
PERCENTILE.INC and PERCENTILE.EXC	134
PERCENTRANK.INC and PERCENTRANK.EXC	137
Data analysis tool: Rank and Percentile	138
<b>CHAPTER 7: Summarizing It All</b>	<b>141</b>
Counting Out	141
COUNT, COUNTA, COUNTBLANK, COUNTIF, COUNTIFS	141
The Long and Short of It	144
MAX, MAXA, MIN, and MINA	144
Getting Esoteric	145
SKEW and SKEW.P	146
KURT	148
Tuning In the Frequency	150
FREQUENCY	150
Data analysis tool: Histogram	152
Can You Give Me a Description?	154
Data analysis tool: Descriptive Statistics	154
Be Quick About It!	156
Instant Statistics	159
<b>CHAPTER 8: What's Normal?</b>	<b>161</b>
Hitting the Curve	161
Digging deeper	162
Parameters of a normal distribution	163
NORM.DIST	165
NORM.INV	167
A Distinguished Member of the Family	168
NORM.S.DIST	169
NORM.S.INV	170
PHI and GAUSS	170
Graphing a Standard Normal Distribution	171



<b>PART 3: DRAWING CONCLUSIONS FROM DATA</b> .....	173
<b>CHAPTER 9: The Confidence Game: Estimation</b> .....	175
Understanding Sampling Distributions .....	176
An EXTREMELY Important Idea: The Central Limit Theorem .....	177
(Approximately) simulating the Central Limit Theorem.....	178
The Limits of Confidence .....	183
Finding confidence limits for a mean.....	183
CONFIDENCE.NORM .....	186
Fit to a t.....	187
CONFIDENCE.T.....	188
<b>CHAPTER 10: One-Sample Hypothesis Testing</b> .....	189
Hypotheses, Tests, and Errors.....	190
Hypothesis Tests and Sampling Distributions.....	191
Catching Some Z's Again.....	193
Z.TEST .....	196
t for One .....	197
T.DIST, T.DIST.RT, and T.DIST.2T .....	198
T.INV and T.INV.2T .....	200
Visualizing a t-Distribution .....	201
Testing a Variance .....	203
CHISQ.DIST and CHISQ.DIST.RT.....	205
CHISQ.INV and CHISQ.INV.RT .....	206
Visualizing a Chi-Square Distribution .....	208
<b>CHAPTER 11: Two-Sample Hypothesis Testing</b> .....	211
Hypotheses Built for Two .....	211
Sampling Distributions Revisited .....	212
Applying the Central Limit Theorem.....	213
Z's once more .....	215
Data analysis tool: z-Test: Two Sample for Means .....	216
t for Two .....	219
Like peas in a pod: Equal variances .....	220
Like p's and q's: Unequal variances.....	221
T.TEST .....	222
Data analysis tool: t-Test: Two Sample.....	223
A Matched Set: Hypothesis Testing for Paired Samples .....	227
T.TEST for matched samples .....	228
Data analysis tool: t-Test: Paired Two Sample for Means .....	230
t-tests on the iPad with StatPlus .....	232
Testing Two Variances .....	235
Using F in conjunction with t .....	237
F.TEST .....	238
F.DIST and F.DIST.RT .....	240

	F.INV and F.INV.RT . . . . .	241
	Data analysis tool: F-test: Two Sample for Variances. . . . .	242
	Visualizing the F-Distribution . . . . .	244
<b>CHAPTER 12:</b>	<b>Testing More Than Two Samples . . . . .</b>	<b>247</b>
	Testing More than Two. . . . .	247
	A thorny problem . . . . .	248
	A solution . . . . .	249
	Meaningful relationships . . . . .	253
	After the F-test. . . . .	254
	Data analysis tool: Anova: Single Factor. . . . .	258
	Comparing the means . . . . .	260
	Another Kind of Hypothesis, Another Kind of Test. . . . .	262
	Working with repeated measures ANOVA. . . . .	262
	Getting trendy . . . . .	264
	Data analysis tool: Anova: Two-Factor Without Replication . . . . .	268
	Analyzing trend . . . . .	271
	ANOVA on the iPad . . . . .	272
	ANOVA on the iPad: Another Way. . . . .	274
	Repeated Measures ANOVA on the iPad . . . . .	277
<b>CHAPTER 13:</b>	<b>Slightly More Complicated Testing . . . . .</b>	<b>281</b>
	Cracking the Combinations . . . . .	281
	Breaking down the variances. . . . .	282
	Data analysis tool: Anova: Two-Factor Without Replication . . . . .	284
	Cracking the Combinations Again. . . . .	286
	Rows and columns . . . . .	286
	Interactions. . . . .	287
	The analysis . . . . .	288
	Data analysis tool: Anova: Two-Factor With Replication . . . . .	289
	Two Kinds of Variables — at Once . . . . .	292
	Using Excel with a Mixed Design. . . . .	293
	Graphing the Results. . . . .	298
	After the ANOVA . . . . .	300
	Two-Factor ANOVA on the iPad. . . . .	300
<b>CHAPTER 14:</b>	<b>Regression: Linear and Multiple . . . . .</b>	<b>303</b>
	The Plot of Scatter. . . . .	303
	Graphing a line. . . . .	305
	Regression: What a Line! . . . . .	307
	Using regression for forecasting . . . . .	309
	Variation around the regression line . . . . .	309
	Testing hypotheses about regression . . . . .	311

Worksheet Functions for Regression . . . . .	317
SLOPE, INTERCEPT, STEYX . . . . .	318
FORECAST.LINEAR . . . . .	319
Array function: TREND . . . . .	319
Array function: LINEST . . . . .	323
Data Analysis Tool: Regression . . . . .	325
Working with tabled output . . . . .	327
Opting for graphical output . . . . .	329
Juggling Many Relationships at Once: Multiple Regression . . . . .	330
Excel Tools for Multiple Regression . . . . .	331
TREND revisited . . . . .	331
LINEST revisited . . . . .	333
Regression data analysis tool revisited . . . . .	336
Regression Analysis on the iPad . . . . .	338
<b>CHAPTER 15: Correlation: The Rise and Fall of Relationships . . . . .</b>	<b>341</b>
Scatterplots Again . . . . .	341
Understanding Correlation . . . . .	342
Correlation and Regression . . . . .	345
Testing Hypotheses about Correlation . . . . .	347
Is a correlation coefficient greater than zero? . . . . .	348
Do two correlation coefficients differ? . . . . .	349
Worksheet Functions for Correlation . . . . .	350
CORREL and PEARSON . . . . .	350
RSQ . . . . .	351
COVARIANCE.P and COVARIANCE.S . . . . .	352
Data Analysis Tool: Correlation . . . . .	353
Tabled output . . . . .	354
Multiple correlation . . . . .	355
Partial correlation . . . . .	356
Semipartial correlation . . . . .	357
Data Analysis Tool: Covariance . . . . .	358
Using Excel to Test Hypotheses about Correlation . . . . .	358
Worksheet functions: FISHER, FISHERINV . . . . .	359
Correlation Analysis on the iPad . . . . .	360
<b>CHAPTER 16: It's About Time . . . . .</b>	<b>363</b>
A Series and Its Components . . . . .	363
A Moving Experience . . . . .	364
Lining up the trend . . . . .	365
Data analysis tool: Moving Average . . . . .	365
How to Be a Smoothie, Exponentially . . . . .	368
One-Click Forecasting . . . . .	369
Working with Time Series on the iPad . . . . .	374



<b>CHAPTER 17: Nonparametric Statistics</b> .....	379
Independent Samples .....	380
Two samples: Mann-Whitney U test .....	380
More than two samples: Kruskal-Wallis one-way ANOVA .....	382
Matched Samples .....	383
Two samples: Wilcoxon matched-pairs signed ranks .....	384
More than two samples: Friedman two-way ANOVA .....	386
More than two samples: Cochran's Q .....	387
Correlation: Spearman's $r_s$ .....	389
A Heads-Up .....	391
<b>PART 4: PROBABILITY</b> .....	393
<b>CHAPTER 18: Introducing Probability</b> .....	395
What Is Probability? .....	395
Experiments, trials, events, and sample spaces .....	396
Sample spaces and probability .....	396
Compound Events .....	397
Union and intersection .....	397
Intersection, again .....	398
Conditional Probability .....	399
Working with the probabilities .....	400
The foundation of hypothesis testing .....	400
Large Sample Spaces .....	400
Permutations .....	401
Combinations .....	402
Worksheet Functions .....	403
FACT .....	403
PERMUT and PERMUTONA .....	403
COMBIN and COMBINA .....	404
Random Variables: Discrete and Continuous .....	405
Probability Distributions and Density Functions .....	405
The Binomial Distribution .....	407
Worksheet Functions .....	409
BINOM.DIST and BINOM.DIST.RANGE .....	409
NEGBINOM.DIST .....	411
Hypothesis Testing with the Binomial Distribution .....	412
BINOM.INV .....	413
More on hypothesis testing .....	414
The Hypergeometric Distribution .....	415
HYPGEOM.DIST .....	416

<b>CHAPTER 19: More on Probability</b> .....	419
Discovering Beta .....	419
BETA.DIST .....	421
BETA.INV .....	423
Poisson .....	424
POISSON.DIST .....	425
Working with Gamma .....	427
The gamma function and GAMMA .....	427
The gamma distribution and GAMMA.DIST .....	428
GAMMA.INV .....	430
Exponential. ....	431
EXPON.DIST .....	431
<b>CHAPTER 20: Using Probability: Modeling and Simulation</b> .....	433
Modeling a Distribution .....	434
Plunging into the Poisson distribution .....	434
Visualizing the Poisson distribution .....	435
Working with the Poisson distribution .....	436
Using POISSON.DIST again .....	437
Testing the model's fit .....	437
A word about CHISQ.TEST .....	440
Playing ball with a model .....	441
A Simulating Discussion .....	444
Taking a chance: The Monte Carlo method .....	444
Loading the dice .....	444
Data analysis tool: Random Number Generation .....	445
Simulating the Central limit Theorem .....	448
Simulating a business .....	452
<b>CHAPTER 21: Estimating Probability: Logistic Regression</b> .....	457
Working Your Way Through Logistic Regression .....	458
Mining with XLMiner .....	460
<b>PART 5: THE PART OF TENS</b> .....	465
<b>CHAPTER 22: Ten (12, Actually) Statistical and Graphical Tips and Traps</b> .....	467
Significant Doesn't Always Mean Important .....	467
Trying to Not Reject a Null Hypothesis Has a Number of Implications .....	468
Regression Isn't Always Linear .....	468
Extrapolating Beyond a Sample Scatterplot Is a Bad Idea .....	469
Examine the Variability Around a Regression Line .....	469
A Sample Can Be Too Large .....	470

Consumers: Know Your Axes. . . . .	470
Graphing a Categorical Variable as a Quantitative Variable Is Just Plain Wrong. . . . .	471
Whenever Appropriate, Include Variability in Your Graph . . . . .	472
Be Careful When Relating Statistics Textbook Concepts to Excel . . . . .	472
It's Always a Good Idea to Use Named Ranges in Excel. . . . .	472
Statistical Analysis with Excel on the iPad Is Pretty Good!. . . . .	473
<b>CHAPTER 23: Ten Topics (Thirteen, Actually) That Just Don't Fit Elsewhere . . . . .</b>	<b>475</b>
Graphing the Standard Error of the Mean. . . . .	475
Probabilities and Distributions . . . . .	479
PROB . . . . .	479
WEIBULL.DIST. . . . .	479
Drawing Samples. . . . .	480
Testing Independence: The True Use of CHISQ.TEST. . . . .	481
Logarithmica Esoterica. . . . .	484
What is a logarithm? . . . . .	484
What is e? . . . . .	486
LOGNORM.DIST. . . . .	489
LOGNORM.INV. . . . .	490
Array Function: LOGEST . . . . .	491
Array Function: <i>GROWTH</i> . . . . .	494
The logs of Gamma . . . . .	497
Sorting Data . . . . .	498
<b>PART 6: APPENDICES. . . . .</b>	<b>501</b>
<b>APPENDIX A: When Your Data Live Elsewhere . . . . .</b>	<b>503</b>
<b>APPENDIX B: Tips for Teachers (and Learners). . . . .</b>	<b>507</b>
Augmenting Analyses Is a Good Thing. . . . .	507
Understanding ANOVA. . . . .	508
Revisiting regression. . . . .	510
Simulating Data Is Also a Good Thing. . . . .	512
When All You Have Is a Graph. . . . .	514
<b>APPENDIX C: More on Excel Graphics. . . . .</b>	<b>515</b>
Tasting the Bubbly. . . . .	515
Taking Stock . . . . .	516
Scratching the Surface . . . . .	518
On the Radar . . . . .	519
Growing a Treemap and Bursting Some Sun . . . . .	520
Building a Histogram . . . . .	521

Ordering Columns: Pareto .....	522
Of Boxes and Whiskers.....	523
3D Maps .....	524
Filled Maps .....	527
<b>APPENDIX D: The Analysis of Covariance .....</b>	<b>529</b>
Covariance: A Closer Look .....	529
Why You Analyze Covariance.....	530
How You Analyze Covariance .....	531
ANCOVA in Excel .....	532
Method 1: ANOVA .....	533
Method 2: Regression.....	537
After the ANCOVA .....	540
And One More Thing.....	542
<b>INDEX.....</b>	<b>545</b>