
Contents

1	Statistics and Empirical Research	1
1.1	Do Statistics Lie?	1
1.2	Two Types of Statistics	3
1.3	The Generation of Knowledge Through Statistics	4
1.4	The Phases of Empirical Research	6
1.4.1	From Exploration to Theory	6
1.4.2	From Theories to Models	7
1.4.3	From Models to Business Intelligence	11
2	Disarray to Dataset	13
2.1	Data Collection	13
2.2	Level of Measurement	15
2.3	Scaling and Coding	18
2.4	Missing Values	19
2.5	Outliers and Obviously Incorrect Values	21
2.6	Chapter Exercises	22
3	Univariate Data Analysis	23
3.1	First Steps in Data Analysis	23
3.2	Measures of Central Tendency	29
3.2.1	Mode or Modal Value	30
3.2.2	Mean	30
3.2.3	Geometric Mean	34
3.2.4	Harmonic Mean	36
3.2.5	The Median	38
3.2.6	Quartile and Percentile	41
3.3	The Boxplot: A First Look at Distributions	42
3.4	Dispersion Parameters	45
3.4.1	Standard Deviation and Variance	46
3.4.2	The Coefficient of Variation	48
3.5	Skewness and Kurtosis	49
3.6	Robustness of Parameters	52
3.7	Measures of Concentration	52

3.8	Using the Computer to Calculate Univariate Parameters	55
3.8.1	Calculating Univariate Parameters with SPSS	55
3.8.2	Calculating Univariate Parameters with Stata	56
3.8.3	Calculating Univariate Parameters with Excel 2010	57
3.9	Chapter Exercises	58
4	Bivariate Association	61
4.1	Bivariate Scale Combinations	61
4.2	Association Between Two Nominal Variables	61
4.2.1	Contingency Tables	61
4.2.2	Chi-Square Calculations	63
4.2.3	The Phi Coefficient	67
4.2.4	The Contingency Coefficient	70
4.2.5	Cramer's V	70
4.2.6	Nominal Associations with SPSS	72
4.2.7	Nominal Associations with Stata	75
4.2.8	Nominal Associations with Excel	76
4.2.9	Chapter Exercises	77
4.3	Association Between Two Metric Variables	80
4.3.1	The Scatterplot	80
4.3.2	The Bravais-Pearson Correlation Coefficient	83
4.4	Relationships Between Ordinal Variables	86
4.4.1	Spearman's Rank Correlation Coefficient (Spearman's rho)	88
4.4.2	Kendall's Tau (τ)	92
4.5	Measuring the Association Between Two Variables with Different Scales	97
4.5.1	Measuring the Association Between Nominal and Metric Variables	98
4.5.2	Measuring the Association Between Nominal and Ordinal Variables	100
4.5.3	Association Between Ordinal and Metric Variables	100
4.6	Calculating Correlation with a Computer	101
4.6.1	Calculating Correlation with SPSS	101
4.6.2	Calculating Correlation with Stata	102
4.6.3	Calculating Correlation with Excel	104
4.7	Spurious Correlations	105
4.7.1	Partial Correlation	107
4.7.2	Partial Correlations with SPSS	109
4.7.3	Partial Correlations with Stata	109
4.7.4	Partial Correlation with Excel	110
4.8	Chapter Exercises	110
5	Regression Analysis	115
5.1	First Steps in Regression Analysis	115
5.2	Coefficients of Bivariate Regression	116

5.3	Multivariate Regression Coefficients	122
5.4	The Goodness of Fit of Regression Lines	123
5.5	Regression Calculations with the Computer	125
5.5.1	Regression Calculations with Excel	125
5.5.2	Regression Calculations with SPSS and Stata	126
5.6	Goodness of Fit of Multivariate Regressions	128
5.7	Regression with an Independent Dummy Variable	129
5.8	Leverage Effects of Data Points	131
5.9	Nonlinear Regressions	132
5.10	Approaches to Regression Diagnostics	135
5.11	Chapter Exercises	140
6	Time Series and Indices	147
6.1	Price Indices	148
6.2	Quantity Indices	155
6.3	Value Indices (Sales Indices)	157
6.4	Deflating Time Series by Price Indices	158
6.5	Shifting Bases and Chaining Indices	159
6.6	Chapter Exercises	161
7	Cluster Analysis	163
7.1	Hierarchical Cluster Analysis	164
7.2	K-Means Cluster Analysis	176
7.3	Cluster Analysis with SPSS and Stata	177
7.4	Chapter Exercises	177
8	Factor Analysis	183
8.1	Factor Analysis: Foundations, Methods, Interpretations	183
8.2	Factor Analysis with SPSS and Stata	191
8.3	Chapter Exercises	193
9	Solutions to Chapter Exercises	197
	References	211
	Index	213

Exploratory Data Analysis in Business and Economics

An Introduction Using SPSS, Stata, and Excel

Cleff, Th.

2014, XXII, 215 p. 130 illus., 11 illus. in color.,

ISBN: 978-3-319-01517-0