Contents

List of Figures  xiv
List of Tables  xvi
Acknowledgements  xx

Introduction: About the Book  1

1 Social Research and Data Analysis: Demystifying Basic Concepts  10
2 Data Analysis in Context: Working with Two Data Sets  37
3 Descriptive Analysis – Univariate: Looking for Characteristics  47
4 Descriptive Analysis – Bivariate: Looking for Patterns  89
5 Explanatory Analysis: Looking for Influences  116
6 Inferential Analysis: From Sample to Population  159
7 Data Reduction: Preparing to Answer Research Questions  214
8 Real Data Analysis: Answering Research Questions  249

Glossary  306

Appendix A: Symbols  324
Appendix B: Equations  326
Appendix C: SPSS Procedures  333
Appendix D: Statistical Tables  339

References  344

Index  347

Summary Chart of Methods  353
Detailed Chapter Contents

List of Figures xiv
List of Tables xvi
Acknowledgements xx

Introduction: About the Book 1

Why was it written? 1
Who is it for? 3
What makes it different? 4
What are the controversial issues? 6
What is the best way to read this book? 7
What is needed to cope with it? 8
Notes 9

1 Social Research and Data Analysis: Demystifying Basic Concepts 10

Introduction 10
What is the purpose of social research? 10
The research problem 11
Research objectives 11
Research questions 13
The role of hypotheses 13
What are data? 15
Data and social reality 16
Types of data 17
Forms of data 20
Concepts and variables 22
Levels of measurement 22
Categorical measurement 23
Nominal-level measurement 23
Ordinal-level measurement 23
Metric measurement 24
Interval-level measurement 25
Ratio-level measurement 25
Discrete and continuous measurement 26
Review 26
Transformations between levels of measurement 27
What is data analysis? 28
Types of analysis 29
Univariate descriptive analysis 29
2 Data Analysis in Context: Working with Two Data Sets

Introduction 37
Two samples 37
Descriptions of the samples 39
   Student sample 39
   Resident sample 39
Concepts and variables 40
   Formal definitions 40
   Operational definitions 40
Levels of measurement 43
Data reduction 44
Notes 45

3 Descriptive Analysis – Univariate: Looking for Characteristics 47

Introduction 47
Basic mathematical language 48
Univariate descriptive analysis 51
   Describing distributions 52
      Frequency counts and distributions 53
         Nominal categories 53
         Ordinal categories 54
         Discrete and grouped data 55
   Proportions and percentages, ratios and rates 59
      Proportions 59
      Percentages 59
      Ratios 61
      Rates 62
   Pictorial representations 62
      Categorical variables 63
      Metric variables 64
   Shapes of frequency distributions: symmetrical, skewed and normal 66
   Measures of central tendency 68
The three Ms 68
   Mode 68
   Median 69
   Mean 71
4 Descriptive Analysis – Bivariate: Looking for Patterns

Introduction
Association with nominal-level and ordinal-level variables
   Contingency tables
Forms of association
   Positive and negative
   Linear and curvilinear
   Symmetrical and asymmetrical
Measures of association for categorical variables
   Nominal-level variables
      Contingency coefficient
      Standardized contingency coefficient
      Phi
      Cramér's $V$
   Ordinal-level variables
      Gamma
      Kendall's tau-b
Other methods for ranked data
Combinations of categorical and metric variables
Association with interval-level and ratio-level variables
   Scatter diagrams
   Covariance
   Pearson's $r$
Comparing the measures
Association between categorical and metric variables
   Code metric variable to ordinal categories
   Dichotomize the categorical variable
Summary
Notes
5 Explanatory Analysis: Looking for Influences 116

Introduction 116
The use of controlled experiments 117
Explanation in cross-sectional research 118
Bivariate analysis 120
  Influence between categorical variables 120
    Nominal-level variables: lambda 120
    Ordinal-level variables: Somer’s d 124
Influence between metric variables: bivariate regression 125
  Two methods of regression analysis 128
    Coefficients 130
    An example 132
    Points to watch for 133
Influence between categorical and metric variables 134
  Coding to a lower level 134
  Means analysis 134
  Dummy variables 135
Multivariate analysis 136
  Trivariate analysis 136
    Forms of relationships 136
    Interacting variables 137
    The logic of trivariate analysis 138
  Influence between categorical variables 141
    Three-way contingency tables 141
      An example 141
    Other methods 145
  Influence between metric variables 146
    Partial correlation 146
    Multiple regression 146
      An example 148
    Collinearity 150
      Multiple-category dummy variables 150
  Other methods 153
    Dependence techniques 153
      Analysis of variance 154
    Multiple analysis of variance 154
    Logistic regression 154
    Logit logistic regression 154
    Multiple discriminant analysis 154
    Structural equation modelling 154
  Interdependence techniques 155
    Factor analysis 155
    Cluster analysis 155
    Multidimensional scaling 155
Summary 156
Notes 158
6 Inferential Analysis: From Sample to Population

Introduction 159
Sampling 160
- Populations and samples 160
- Probability samples 161
- Probability theory 163
- Sample size 166
- Response rate 167
- Sampling methods 168
Parametric and non-parametric tests 171
Inference in univariate descriptive analysis 172
- Categorical variables 173
- Metric variables 175
Inference in bivariate descriptive analysis 177
- Testing statistical hypotheses 178
  - Null and alternative hypotheses 179
  - Type I and type II errors 180
  - One-tailed and two-tailed tests 181
  - The process of testing statistical hypotheses 182
  - Testing hypotheses under different conditions 183
  - Some critical issues 185
- Categorical variables 189
  - Nominal-level data 189
  - Ordinal-level data 191
- Metric variables 192
  - Comparing means 192
    - Group t test 193
    - Mann–Whitney U test 197
    - Analysis of variance 201
    - Test of significance for Pearson’s r 204
Inference in explanatory analysis 205
- Nominal-level data 205
- Ordinal-level data 206
- Metric variables 208
  - Bivariate regression 208
  - Multiple regression 209
Summary 210
Notes 212

7 Data Reduction: Preparing to Answer Research Questions 214

Introduction 214
Scales and indexes 214
Creating scales 215
  - Environmental Worldview scales and subscales 215
    - Pre-testing the items 216
    - Item-to-item correlations 217
Item-to-total correlations 217
Cronbach’s alpha 219
Factor analysis 220
Willingness to Act scale 238
Indexes 239
Avoidance of environmentally damaging products 240
Support for environmental groups 240
Recycling behaviour 240
Recoding to different levels of measurement 241
Environmental Worldview scales and subscales 242
Recycling index 243
Age 243
Characteristics of the samples 244
Summary 246
Notes 248

8 Real Data Analysis: Answering Research Questions 249

Introduction 249
Univariate descriptive analysis 249
Environmental Worldview 250
Environmentally Responsible Behaviour 252
Bivariate descriptive analysis 257
Environmental Worldview and Environmentally Responsible Behaviour 258
Metric variables 258
Categorical variables 260
Comparing metric and categorical variables 262
Conclusion 263
Age, Environmental Worldview and Environmentally Responsible Behaviour 264
Metric variables 264
Categorical variables 266
Gender, Environmental Worldview and Environmentally Responsible Behaviour 268
Explanatory analysis 270
Bivariate analysis 273
Categorical variables 274
Categorical and metric variables: means analysis 276
Metric variables 277
Multivariate analysis 278
Categorical variables 278
EWVGSC and WILLACT with ERB 279
WILLACT, Age and Gender with ERB 282
Categorical and metric variables: means analysis 285
EWVGSC and WILLACT with ERB 286
WILLACT and Gender with ERB 287
<table>
<thead>
<tr>
<th>Metric variables</th>
<th>292</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partial correlation</td>
<td>292</td>
</tr>
<tr>
<td>Multiple regression</td>
<td>293</td>
</tr>
<tr>
<td>Conclusion</td>
<td>303</td>
</tr>
<tr>
<td>Notes</td>
<td>304</td>
</tr>
<tr>
<td><strong>Glossary</strong></td>
<td>306</td>
</tr>
<tr>
<td><strong>Appendix A: Symbols</strong></td>
<td>324</td>
</tr>
<tr>
<td><strong>Appendix B: Equations</strong></td>
<td>326</td>
</tr>
<tr>
<td><strong>Appendix C: SPSS Procedures</strong></td>
<td>333</td>
</tr>
<tr>
<td><strong>Appendix D: Statistical Tables</strong></td>
<td>339</td>
</tr>
<tr>
<td><strong>References</strong></td>
<td>344</td>
</tr>
<tr>
<td><strong>Index</strong></td>
<td>347</td>
</tr>
<tr>
<td><strong>Summary Chart of Methods</strong></td>
<td>353</td>
</tr>
</tbody>
</table>
List of Figures

3.1 Religion (Students): bar chart 63
3.2 Religion (both samples): bar chart 64
3.3 Religiosity (both samples): bar chart 64
3.4 Religion (Students): pie chart 65
3.5 Religiosity (Students): pie chart 65
3.6 Age (both samples): line graphs 66
3.7 Examples of symmetrical distributions 67
3.8 Median to one decimal place 71
3.9 Environmental Worldview (both samples): line graphs 77
3.10 Environmental Worldview (combined samples): line graph 77
3.11 Area covered under the normal curve by one to three standard deviations 86

4.1 Parts of a table 92
4.2 Scatter diagram: Environmental Worldview by Age (Residents) 107
4.3 Scatter diagram: Environmental Worldview by Age (subsample of Residents) 109

5.1 Scatter plot of weekly hours worked by weekly wages 127
5.2 Residuals from a regression line (hypothetical data) 131
5.3 Possible forms of relationships between three variables 137

6.1 Distributions of mean ages of 20 samples 164
6.2 Types and methods of sampling 170
6.3 Confidence intervals for mean Age by sample size (Resident sample) 177

7.1 Scree plot of eigenvalues for 24 items (combined samples) 223
7.2 Scree plot of eigenvalues for 14 items (combined samples) 227
7.3 Scree plot of eigenvalues for nine items (combined samples) 229
7.4 EWVGSC mean scores (combined samples) 233
7.5 HUSENV mean scores (combined samples) 233
7.6 GOVCONT mean scores (combined samples) 233
7.7 ECGROW mean scores (combined samples) 234
7.8 SCITEK mean scores (combined samples) 234
7.9 IMPACT mean scores (combined samples) 234
7.10 ALTENGY mean scores (combined samples) 235
7.11 WILLACT mean scores (combined samples) 239
8.1 EWVGSC categories (both samples) 253
8.2 WILLACT categories (both samples) 255
8.3 Support Groups (both samples) 256
8.4 Avoid Products (both samples) 256
8.5 Recycling index (both samples) 257
8.6 Support Groups by WILLACT controlled for Gender (Students) 288
8.7 Avoid Products by WILLACT controlled for Gender (Students) 289
8.8 Support Groups by WILLACT controlled for Gender (Residents) 290
8.9 Avoid Products by WILLACT controlled for Gender (Residents) 290
## List of Tables

1.1 Research questions and objectives 14
1.2 Levels of measurement 27

3.1 Raw data on Religion (Students) 53
3.2 Distribution by Religion (both samples) 53
3.3 Distribution by Religiosity (both samples) 54
3.4 Age distribution in years (Students) 55
3.5 Age distribution in five categories (Students) 56
3.6 Age distribution in six categories (Residents) 56
3.7 Number of children (Residents) 57
3.8 Number of children (subsample of Residents) 57
3.9 Comparison of Student and Resident samples by Age 58
3.10 Comparison of Gender proportions (both samples) 60
3.11 Age in years (Residents) 70
3.12 Calculation of mean Age in years (Residents) 73
3.13 Mean of Age distributed in ten categories (Residents) 73
3.14 Mean of two means (both samples) 74
3.15 Mean of two Age category percentages (both samples) 75
3.16 Deviations from the mean of Age in years (Residents) 81–82

4.1 Religion by Gender (Residents; observed and expected frequencies, and percentages) 92
4.2 Environmental Worldview by Age (Residents; observed frequencies and percentages) 94
4.3 Environmental Worldview by Age (percentages) 95
4.4 Religion by Gender (Residents; observed frequencies) 99
4.5 Calculation of gamma (from Table 4.2) 103
4.6 Mean deviation method for computing \( r \) (subsample of Residents) 110
4.7 Raw score method for computing \( r \) (subsample of Residents) 110
4.8 Education by Age (percentages; Residents) 113

5.1 Occupation by Religion (Residents; observed frequencies and percentages) 122
5.2 Occupation by Religion (subsample of Residents) 123
5.3 Occupation by Religion (subsample of Residents; 2 by 2 table) 124
5.4 Working hours per week and weekly wage 126
5.5 Unexplained variation and standard error of the estimate (subsample of Residents) 132
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.6</td>
<td>A means analysis of Education and Environmental Worldview (Residents)</td>
<td>135</td>
</tr>
<tr>
<td>5.7</td>
<td>Forms of relationships between three variables</td>
<td>139</td>
</tr>
<tr>
<td>5.8</td>
<td>Environmental Worldview and Age (Residents)</td>
<td>142</td>
</tr>
<tr>
<td>5.9</td>
<td>Environmental Worldview and Age controlled for Education (Residents)</td>
<td>143</td>
</tr>
<tr>
<td>5.10</td>
<td>Environmental Worldview and Age controlled for Gender (Residents)</td>
<td>144</td>
</tr>
<tr>
<td>5.11</td>
<td>Regression of Environmental Worldview on Age, Gender and Education (Residents)</td>
<td>148</td>
</tr>
<tr>
<td>5.12</td>
<td>Regression of Environmental Worldview on Age, Gender and Education in five categories (Residents)</td>
<td>151</td>
</tr>
<tr>
<td>5.13</td>
<td>Correlation matrix for Age, Gender and six Education dummy variables (Residents)</td>
<td>152</td>
</tr>
<tr>
<td>5.14</td>
<td>Regression of Environmental Worldview on Age, Gender and Education, Marital Status, Religion and Political Party Preference (Residents)</td>
<td>152</td>
</tr>
<tr>
<td>6.1</td>
<td>Hypothetical sampling</td>
<td>163</td>
</tr>
<tr>
<td>6.2</td>
<td>Variations in confidence intervals of mean Age by confidence level and sample size (Residents)</td>
<td>176</td>
</tr>
<tr>
<td>6.3</td>
<td>Type I and type II errors</td>
<td>181</td>
</tr>
<tr>
<td>6.4</td>
<td>Ranked Environmental Worldview scores by Gender (subsample of Students)</td>
<td>200</td>
</tr>
<tr>
<td>6.5</td>
<td>Cells and their 'diagonals' in Table 4.2</td>
<td>208</td>
</tr>
<tr>
<td>7.1</td>
<td>Correlation matrix of 24 items (both samples)</td>
<td>218</td>
</tr>
<tr>
<td>7.2</td>
<td>Unidimensionality, reliability and commonalities of 24 items (combined samples)</td>
<td>219</td>
</tr>
<tr>
<td>7.3</td>
<td>Commonalities and unrotated factors with 24 items (combined samples)</td>
<td>222</td>
</tr>
<tr>
<td>7.4</td>
<td>Rotated solution for five factors with 24 items (combined samples)</td>
<td>225</td>
</tr>
<tr>
<td>7.5</td>
<td>Rotated solution for six factors with 24 items (combined samples)</td>
<td>226</td>
</tr>
<tr>
<td>7.6</td>
<td>Unrotated and rotated solutions with 14 retained items (combined samples)</td>
<td>228</td>
</tr>
<tr>
<td>7.7</td>
<td>Unidimensionality and reliability of 10 rejected items (combined samples)</td>
<td>228</td>
</tr>
<tr>
<td>7.8</td>
<td>Unrotated and rotated solutions with nine rejected items (combined samples)</td>
<td>230</td>
</tr>
<tr>
<td>7.9</td>
<td>Distributions on the 24 items (combined samples)</td>
<td>231</td>
</tr>
<tr>
<td>7.10</td>
<td>Distributions on scales and subscales (combined samples)</td>
<td>232</td>
</tr>
<tr>
<td>7.11</td>
<td>Reliability of scales and subscales (combined samples)</td>
<td>236</td>
</tr>
<tr>
<td>7.12</td>
<td>Correlation matrix of EWV scales and subscales (combined samples)</td>
<td>237</td>
</tr>
</tbody>
</table>
7.13 Unrotated and rotated solutions with Willingness to Act items (combined samples) 238
7.14 Reliability of behavioural scales (combined samples) 239
7.15 Characteristics of both samples 245

8.1 Sample comparisons of Environmental Worldview metric variables 250
8.2 Sample comparisons of Environmental Worldview categorical variables (percentages) 252
8.3 Sample comparison of Environmentally Responsible Behaviour metric variables 253
8.4 Sample comparison of Environmentally Responsible Behaviour categorical variables (percentages) 254
8.5 Correlation matrix for EWV and ERB variables (Pearson’s $r$; Students) 258
8.6 Correlation matrix for EWV and ERB variables (Pearson’s $r$; Residents) 259
8.7 Cross-tabulations between EWVGSC and WILLACT, Support Groups, Avoid Products and Recycling (percentages; both samples) 260
8.8 Correlation matrix for EWV and ERB variables (gamma; Students) 261
8.9 Correlation matrix for EWV and ERB variables (gamma; Residents) 262
8.10 Cross-tabulations of Support Groups with WILLACT (percentages; both samples) 263
8.11 EWV and ERB by Age (Pearson’s $r$ and gamma; Residents) 265
8.12 EWV and ERB means and standard deviations by Age (Residents) 265
8.13 Cross-tabulation for Age with EWVGSC, IMPACT, WILLACT, Recycling, Support Groups and Avoid Products (percentages; Residents) 267
8.14 EWV and ERB by Gender (Pearson’s $r$ and $G$; both samples) 268
8.15 EWV and ERB means and standard deviations by Gender (both samples) 269
8.16 Cross-tabulation of Gender with EWVGSC, SCITEK, WILLACT, Recycling, Support Groups and Avoid Products (percentages; both samples) 271
8.17 Influence of EWVGSC and WILLACT on Support Groups and Avoid Products (percentages; both samples) 275
8.18 Means analysis of Gender and Religion (Students), and Age, Gender and Religion (Residents), with Support Groups and Avoid Products 276
8.19 Regression of ERB variables on WILLACT and EWVGSC (both samples) 277
8.20 Influence of EWVGSC on Support Groups and Avoid Products controlled for WILLACT (percentages; Students)  
8.21 Influence of WILLACT on Support Groups and Avoid Products controlled for EWVGSC (percentages; Students)  
8.22 Influence of EWVGSC and WILLACT on Support Groups and Avoid Products with controls for WILLACT and EWVGSC (Residents)  
8.23 Influence of WILLACT on Support Groups and Avoid Products controlled for Gender (percentages; both samples)  
8.24 Influence of WILLACT on Support Groups and Avoid Products controlled for Age (Residents)  
8.25 Means analysis of EWVGSC on Support Groups and Avoid Products controlled for WILLACT (Students)  
8.26 Means analysis of WILLACT on Support Groups and Avoid Products controlled for EWVGSC (Students)  
8.27 Means analysis of WILLACT on Support Groups and Avoid Products controlled for Gender (Students)  
8.28 Means analysis of WILLACT on Support Groups and Avoid Products controlled for Gender (Residents)  
8.29 Means analysis of WILLACT on Support Groups and Avoid Products controlled for Age (Residents)  
8.30 WILLACT by Support Groups and Avoid Products controlled for EWVGSC (Pearson's r; both samples)  
8.31 Regression of ERB variables on EWVGSC, WILLACT and Gender (Students)  
8.32 Regression of ERB variables on EWVGSC, WILLACT, Age and Gender (Residents)  
8.33 Correlation matrix of potential predictor variables (Pearson's r; Residents)  
8.34 Regression of Support Groups on selected predictor variables (Residents)  
8.35 Regression of Avoid Products on selected predictor variables (Residents)