## 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**  

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why was it written?</td>
<td>1</td>
</tr>
<tr>
<td>Who is it for?</td>
<td>3</td>
</tr>
<tr>
<td>What makes it different?</td>
<td>4</td>
</tr>
<tr>
<td>What are the controversial issues?</td>
<td>6</td>
</tr>
<tr>
<td>What is the best way to read this book?</td>
<td>7</td>
</tr>
<tr>
<td>What is needed to cope with it?</td>
<td>8</td>
</tr>
<tr>
<td>Notes</td>
<td>9</td>
</tr>
</tbody>
</table>

**1 Social Research and Data Analysis: Demystifying Basic Concepts** 10  

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

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>37</td>
</tr>
<tr>
<td>Two samples</td>
<td>37</td>
</tr>
<tr>
<td>Descriptions of the samples</td>
<td>39</td>
</tr>
<tr>
<td>Student sample</td>
<td>39</td>
</tr>
<tr>
<td>Resident sample</td>
<td>39</td>
</tr>
<tr>
<td>Concepts and variables</td>
<td>40</td>
</tr>
<tr>
<td>Formal definitions</td>
<td>40</td>
</tr>
<tr>
<td>Operational definitions</td>
<td>40</td>
</tr>
<tr>
<td>Levels of measurement</td>
<td>43</td>
</tr>
<tr>
<td>Data reduction</td>
<td>44</td>
</tr>
<tr>
<td>Notes</td>
<td>45</td>
</tr>
</tbody>
</table>

3 Descriptive Analysis – Univariate: Looking for Characteristics

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>47</td>
</tr>
<tr>
<td>Basic mathematical language</td>
<td>48</td>
</tr>
<tr>
<td>Univariate descriptive analysis</td>
<td>51</td>
</tr>
<tr>
<td>Describing distributions</td>
<td>52</td>
</tr>
<tr>
<td>Frequency counts and distributions</td>
<td>53</td>
</tr>
<tr>
<td>Nominal categories</td>
<td>53</td>
</tr>
<tr>
<td>Ordinal categories</td>
<td>54</td>
</tr>
<tr>
<td>Discrete and grouped data</td>
<td>55</td>
</tr>
<tr>
<td>Proportions and percentages, ratios and rates</td>
<td>59</td>
</tr>
<tr>
<td>Proportions</td>
<td>59</td>
</tr>
<tr>
<td>Percentages</td>
<td>59</td>
</tr>
<tr>
<td>Ratios</td>
<td>61</td>
</tr>
<tr>
<td>Rates</td>
<td>62</td>
</tr>
<tr>
<td>Pictorial representations</td>
<td>62</td>
</tr>
<tr>
<td>Categorical variables</td>
<td>63</td>
</tr>
<tr>
<td>Metric variables</td>
<td>64</td>
</tr>
<tr>
<td>Shapes of frequency distributions: symmetrical, skewed and normal</td>
<td>66</td>
</tr>
<tr>
<td>Measures of central tendency</td>
<td>68</td>
</tr>
<tr>
<td>The three Ms</td>
<td>68</td>
</tr>
<tr>
<td>Mode</td>
<td>68</td>
</tr>
<tr>
<td>Median</td>
<td>69</td>
</tr>
<tr>
<td>Mean</td>
<td>71</td>
</tr>
</tbody>
</table>
Mean of means
Comparing the mode, median and mean
Comparative analysis using percentages and means
Measures of dispersion
Categorical data
  Interquartile range
  Percentiles
Metric data
  Range
  Mean absolute deviation
  Standard deviation
  Variance
Characteristics of the normal curve
Summary
Notes

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
8 Real Data Analysis: Answering Research Questions

Introduction
Univariate descriptive analysis
  Environmental Worldview
  Environmentally Responsible Behaviour
Bivariate descriptive analysis
  Environmental Worldview and Environmentally Responsible Behaviour
  Metric variables
  Categorical variables
  Comparing metric and categorical variables
  Conclusion
Age, Environmental Worldview and Environmentally Responsible Behaviour
  Metric variables
  Categorical variables
Gender, Environmental Worldview and Environmentally Responsible Behaviour
Explanatory analysis
Bivariate analysis
  Categorical variables
  Categorical and metric variables: means analysis
  Metric variables
Multivariate analysis
  Categorical variables
    EWVGSC and WILLACT with ERB
    WILLACT, Age and Gender with ERB
  Categorical and metric variables: means analysis
    EWVGSC and WILLACT with ERB
    WILLACT and Gender with ERB
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Religion (Students): bar chart</td>
<td>63</td>
</tr>
<tr>
<td>3.2</td>
<td>Religion (both samples): bar chart</td>
<td>64</td>
</tr>
<tr>
<td>3.3</td>
<td>Religiosity (both samples): bar chart</td>
<td>64</td>
</tr>
<tr>
<td>3.4</td>
<td>Religion (Students): pie chart</td>
<td>65</td>
</tr>
<tr>
<td>3.5</td>
<td>Religiosity (Students): pie chart</td>
<td>65</td>
</tr>
<tr>
<td>3.6</td>
<td>Age (both samples): line graphs</td>
<td>66</td>
</tr>
<tr>
<td>3.7</td>
<td>Examples of symmetrical distributions</td>
<td>67</td>
</tr>
<tr>
<td>3.8</td>
<td>Median to one decimal place</td>
<td>71</td>
</tr>
<tr>
<td>3.9</td>
<td>Environmental Worldview (both samples): line graphs</td>
<td>77</td>
</tr>
<tr>
<td>3.10</td>
<td>Environmental Worldview (combined samples): line graph</td>
<td>77</td>
</tr>
<tr>
<td>3.11</td>
<td>Area covered under the normal curve by one to three standard deviations</td>
<td>86</td>
</tr>
<tr>
<td>4.1</td>
<td>Parts of a table</td>
<td>92</td>
</tr>
<tr>
<td>4.2</td>
<td>Scatter diagram: Environmental Worldview by Age (Residents)</td>
<td>107</td>
</tr>
<tr>
<td>4.3</td>
<td>Scatter diagram: Environmental Worldview by Age (subsample of Residents)</td>
<td>109</td>
</tr>
<tr>
<td>5.1</td>
<td>Scatter plot of weekly hours worked by weekly wages</td>
<td>127</td>
</tr>
<tr>
<td>5.2</td>
<td>Residuals from a regression line (hypothetical data)</td>
<td>131</td>
</tr>
<tr>
<td>5.3</td>
<td>Possible forms of relationships between three variables</td>
<td>137</td>
</tr>
<tr>
<td>6.1</td>
<td>Distributions of mean ages of 20 samples</td>
<td>164</td>
</tr>
<tr>
<td>6.2</td>
<td>Types and methods of sampling</td>
<td>170</td>
</tr>
<tr>
<td>6.3</td>
<td>Confidence intervals for mean Age by sample size (Resident sample)</td>
<td>177</td>
</tr>
<tr>
<td>7.1</td>
<td>Scree plot of eigenvalues for 24 items (combined samples)</td>
<td>223</td>
</tr>
<tr>
<td>7.2</td>
<td>Scree plot of eigenvalues for 14 items (combined samples)</td>
<td>227</td>
</tr>
<tr>
<td>7.3</td>
<td>Scree plot of eigenvalues for nine items (combined samples)</td>
<td>229</td>
</tr>
<tr>
<td>7.4</td>
<td>EWVGSC mean scores (combined samples)</td>
<td>233</td>
</tr>
<tr>
<td>7.5</td>
<td>HUSENV mean scores (combined samples)</td>
<td>233</td>
</tr>
<tr>
<td>7.6</td>
<td>GOVCONT mean scores (combined samples)</td>
<td>233</td>
</tr>
<tr>
<td>7.7</td>
<td>ECGROW mean scores (combined samples)</td>
<td>234</td>
</tr>
<tr>
<td>7.8</td>
<td>SCITEK mean scores (combined samples)</td>
<td>234</td>
</tr>
<tr>
<td>7.9</td>
<td>IMPACT mean scores (combined samples)</td>
<td>234</td>
</tr>
<tr>
<td>7.10</td>
<td>ALTENGY mean scores (combined samples)</td>
<td>235</td>
</tr>
<tr>
<td>7.11</td>
<td>WILLACT mean scores (combined samples)</td>
<td>239</td>
</tr>
</tbody>
</table>
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
5.6 A means analysis of Education and Environmental Worldview (Residents) 135
5.7 Forms of relationships between three variables 139
5.8 Environmental Worldview and Age (Residents) 142
5.9 Environmental Worldview and Age controlled for Education (Residents) 143
5.10 Environmental Worldview and Age controlled for Gender (Residents) 144
5.11 Regression of Environmental Worldview on Age, Gender and Education (Residents) 148
5.12 Regression of Environmental Worldview on Age, Gender and Education in five categories (Residents) 151
5.13 Correlation matrix for Age, Gender and six Education dummy variables (Residents) 152
5.14 Regression of Environmental Worldview on Age, Gender and Education, Marital Status, Religion and Political Party Preference (Residents) 152

6.1 Hypothetical sampling 163
6.2 Variations in confidence intervals of mean Age by confidence level and sample size (Residents) 176
6.3 Type I and type II errors 181
6.4 Ranked Environmental Worldview scores by Gender (subsample of Students) 200
6.5 Cells and their 'diagonals' in Table 4.2 208

7.1 Correlation matrix of 24 items (both samples) 218
7.2 Unidimensionality, reliability and commonalities of 24 items (combined samples) 219
7.3 Commonalities and unrotated factors with 24 items (combined samples) 222
7.4 Rotated solution for five factors with 24 items (combined samples) 225
7.5 Rotated solution for six factors with 24 items (combined samples) 226
7.6 Unrotated and rotated solutions with 14 retained items (combined samples) 228
7.7 Unidimensionality and reliability of 10 rejected items (combined samples) 228
7.8 Unrotated and rotated solutions with nine rejected items (combined samples) 230
7.9 Distributions on the 24 items (combined samples) 231
7.10 Distributions on scales and subscales (combined samples) 232
7.11 Reliability of scales and subscales (combined samples) 236
7.12 Correlation matrix of EWV scales and subscales (combined samples) 237
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)