

Detailed Contents

Preface: In praise of data literacy xi

Acknowledgments xiii

Section I. Welcome to the world of data 1

Chapter 1. DATA DEFINED 3

Climbing the pyramid 6

A brief history of the data world 8

Data file formats 9

Section II. Identifying and obtaining data 13

Chapter 2. CLUES FOR UNCOVERING DATA 15

How and why agencies collect, analyze and publish data 15

Clues from data entry 19

Clues from reports 23

Tricks to uncover forms and reports 25

On your own 26

Chapter 3. ONLINE DATABASES 29

Destination: Data portals 29

Statistical stockpiles 35

Agency sites 38

Nongovernmental resources 40

Data search tricks 41

Don't forget the road map 42

Downloading, unzipping and inspecting data files 43

On your own 49

Chapter 4. IDENTIFYING AND REQUESTING OFFLINE DATA 51

- Other clues for offline data 51
- Find the data nerd 57
- Requesting the data 58
- Writing the data request 60
- FOIA in action 60
- Negotiating through obstacles 61
- Getting help 62
- On your own 64

Section III. Evaluating and cleaning data 65

Chapter 5. DATA DIRT IS EVERYWHERE 67

- All data are dirty 69
- Detecting dirt in agricultural data 71
- Changed rules = changed data 72
- On your own 73

Chapter 6. DATA INTEGRITY CHECKS 75

- Big-picture checks 76
- Detailed checks 78
- On your own 94

Chapter 7. GETTING YOUR DATA IN SHAPE 95

- Column carving 96
- Concatenate to paste 103
- Date tricks 106
- Power scrubbing with OpenRefine 107
- Extracting data from PDFs 111
- On your own 112

Section IV. Analyzing data 113

Chapter 8. NUMBER SUMMARIES AND COMPARISONS 115

- Simple summary statistics 116
- Compared to what? 117
- Benchmarking 118
- On your own 119

Chapter 9. CALCULATING SUMMARY STATISTICS AND NUMBER COMPARISONS 121

Sum crimes by year 122
Minimum and maximum numbers 122
Amount change 123
Stepping up to percent change 123
Running rates 126
Running ratios 127
Percent of total 128
More summarizing 129
On your own 129

Chapter 10. SPREADSHEETS AS DATABASE MANAGERS 131

Sorting 131
Filtering 134
Grouping and summarizing 138
On your own 141

Section V. Visualizing data 143

Chapter 11. VISUALIZING YOUR DATA 145

Data visualization defined 145
Some best practices 146

Chapter 12. CHARTING CHOICES 149

Visualizing data with charts 149
On your own 156

Chapter 13. CHARTING IN EXCEL 157

Pie chart 157
Horizontal bar charts 158
Column and line charts 160
Scatterplot 164
Stock chart 166
Sparklines 167
On your own 167

Chapter 14. CHARTING WITH WEB TOOLS 169

Online visualization options 171
Evaluating Web visualization platforms 174
Creating Fusion Table charts 175
On your own 180

Chapter 15. TAKING ANALYSIS TO THE NEXT LEVEL 181

Database managers 181
Statistical programs 184

Appendix: DATA TOOLKIT 187

Spreadsheets 187
Text editors 188
Data cleaners 188

Glossary 189
References 195
Index 199
About the Author 205