Contents

1	Introduction	page 1
1.1	Some Assumptions	2
1.2	R and Python Environments	4
1.3	The Rest of the Book	6
1.4	How to Read This Book	7
1.5	Acknowledgements	8
	PART ONE GETTING STARTED	
2	Data for Plotting	11
2.1	Data-Types	11
2.2	Data Types in R and Python	13
2.3	Dataset Structure	18
3	Visualization Basics	24
3.1	Elements of Information Visualization	24
3.2	Beyond Default	40
3.3	Python's Grammar of Graphics	61
	PART TWO VISUALIZING TABULAR DATA	
4	Insights from ONE Variable	69
4.1	Information from ONE Variable	69
4.2	Categorical I: Visualizing Gaps	70
4.3	Categorical II: Visualizing Representativity	78
4.4	Categorical III: Diversity and Inequality	83
4.5	Categorical IV: Symmetry and Position	90
4.6	Numerical I: Dispersion	96
4.7	Numerical II: Shape	102
4.8	Numerical III: Numerical Inequality	105

vi Contents

5	Insights from TWO Variables	110
5.1	Cat-Cat Relationships	111
5.2	Cat-Num Relationship	131
5.3	Num-Num Relationship	140
6	Insights from THREE or More Variables	167
6.1	Partitioning with Facets	167
6.2	Color Intensity	170
6.3	Simultaneity	177
6.4	Area	182
	PART THREE BEYOND TABULAR DATA	
7	Geospatial Data	193
7.1	Data Suitable for Maps	193
7.2	Map Files	194
7.3	Combining Maps and Data	200
7.4	Focusing on Map Zones	214
7.5	Customizing Color and Size	217
8	Data from the Social Network	236
8.1	Getting Access	237
8.2	Getting Texts	238
8.3	Visuals Based on Text	240
8.4	Visuals Based on Actors	255
	Bibliography	276
	Index	282