Contents

Part I Getting Started

1 Introduction 3
1.1 Welcome to ggplot2 3
1.2 What Is the Grammar of Graphics? 4
1.3 How Does ggplot2 Fit in with Other R Graphics? 5
1.4 About This Book 7
1.5 Installation 8
1.6 Other Resources 8
1.7 Colophon 8
References 10

2 Getting Started with ggplot2 11
2.1 Introduction 11
2.2 Fuel Economy Data 12
2.2.1 Exercises 12
2.3 Key Components 13
2.3.1 Exercises 14
2.4 Colour, Size, Shape and Other Aesthetic Attributes 14
2.4.1 Exercises 16
2.5 Facetting 16
2.5.1 Exercises 17
2.6 Plot Geoms 18
2.6.1 Adding a Smoother to a Plot 18
2.6.2 Boxplots and Jittered Points 20
2.6.3 Histograms and Frequency Polygons 22
2.6.4 Bar Charts 24
2.6.5 Time Series with Line and Path Plots 25
2.6.6 Exercises 26
2.7 Modifying the Axes 27

2.8 Output 29
2.9 Quick Plots 30

3 Toolbox 33
3.1 Introduction 33
3.2 Basic Plot Types 34
3.2.1 Exercises 36
3.3 Labels 37
3.4 Annotations 42
3.5 Collective Geoms 46
3.5.1 Multiple Groups, One Aesthetic 46
3.5.2 Different Groups on Different Layers 47
3.5.3 Overriding the Default Grouping 49
3.5.4 Matching Aesthetics to Graphic Objects 50
3.5.5 Exercises 53
3.6 Surface Plots 54
3.7 Drawing Maps 55
3.7.1 Vector Boundaries 55
3.7.2 Point Metadata 57
3.7.3 Raster Images 59
3.7.4 Area Metadata 59
3.8 Revealing Uncertainty 61
3.9 Weighted Data 62
3.10 Diamonds Data 64
3.11 Displaying Distributions 65
3.11.1 Exercises 69
3.12 Dealing with Overplotting 69
3.13 Statistical Summaries 72
3.14 Add-on Packages 73
References 74

Part II The Grammar

4 Mastering the Grammar 77
4.1 Introduction 77
4.2 Building a Scatterplot 78
4.2.1 Mapping Aesthetics to Data 78
4.2.2 Scaling 80
4.3 Adding Complexity 82
4.4 Components of the Layered Grammar 83
4.4.1 Layers 85
4.4.2 Scales 85
4.4.3 Coordinate System 86
4.4.4 Facetting 87
4.5 Exercises 87
References 88

5 Build a Plot Layer by Layer 89
5.1 Introduction 89
5.2 Building a Plot 89
5.3 Data 91
5.3.1 Exercises 94
5.4 Aesthetic Mappings 94
5.4.1 Specifying the Aesthetics in the Plot vs. in the Layers. 95
5.4.2 Setting vs. Mapping 96
5.4.3 Exercises 98
5.5 Geoms 99
5.5.1 Exercises 101
5.6 Stats 102
5.6.1 Generated Variables 103
5.6.2 Exercises 105
5.7 Position Adjustments 105
5.7.1 Exercises 107

6 Scales, Axes and Legends 109
6.1 Introduction 109
6.2 Modifying Scales 109
6.2.1 Exercises 111
6.3 Guides: Legends and Axes 111
6.3.1 Scale Title 112
6.3.2 Breaks and Labels 113
6.3.3 Exercises 117
6.4 Legends 118
6.4.1 Layers and Legends 118
6.4.2 Legend Layout 120
6.4.3 Guide Functions 122
6.4.4 Exercises 125
6.5 Limits 126
6.5.1 Exercises 129
6.6 Scales Toolbox 129
6.6.1 Continuous Position Scales 130
6.6.2 Colour 133
6.6.3 The Manual Discrete Scale 141
6.6.4 The Identity Scale 143
6.6.5 Exercises 144
References 145

7 Positioning 147
7.1 Introduction 147
7.2 Facetting 147
7.2.1 Facet Wrap 148
7.2.2 Facet Grid 149

7.2.3 Controlling Scales 151
7.2.4 Missing Facetting Variables 154
7.2.5 Grouping vs. Facetting 155
7.2.6 Continuous Variables 157
7.2.7 Exercises 159
7.3 Coordinate Systems 159
7.4 Linear Coordinate Systems 160
7.4.1 Zooming into a Plot with `coord_cartesian()` 160
7.4.2 Flipping the Axes with `coord_flip()` 161
7.4.3 Equal Scales with `coord_fixed()` 162
7.5 Non-linear Coordinate Systems 162
7.5.1 Transformations with `coord_trans()` 165
7.5.2 Polar Coordinates with `coord_polar()` 166
7.5.3 Map Projections with `coord_map()` 167

8 Themes 169
8.1 Introduction 169
8.2 Complete Themes 172
8.2.1 Exercises 174
8.3 Modifying Theme Components 175
8.4 Theme Elements 178
8.4.1 Plot Elements 178
8.4.2 Axis Elements 179
8.4.3 Legend Elements 180
8.4.4 Panel Elements 182
8.4.5 Facetting Elements 183
8.4.6 Exercises 184
8.5 Saving Your Output 184
References 186

Part III Data Analysis

9 Data Analysis 189
9.1 Introduction 189
9.2 Tidy Data 190
9.3 Spread and Gather 191
9.3.1 Gather 192
9.3.2 Spread 194
9.3.3 Exercises 195
9.4 Separate and Unite 195
9.4.1 Exercises 196
9.5 Case Studies 196
9.5.1 Blood Pressure 197
9.5.2 Test Scores 198
9.6 Learning More 200
References 201

10 Data Transformation 203
10.1 Introduction 203
10.2 Filter Observations 204
10.2.1 Useful Tools 206
10.2.2 Missing Values 207
10.2.3 Exercises 208
10.3 Create New Variables 208
10.3.1 Useful Tools 210
10.3.2 Exercises 211
10.4 Group-wise Summaries 211
10.4.1 Useful Tools 214
10.4.2 Statistical Considerations 214
10.4.3 Exercises 217
10.5 Transformation Pipelines 218
10.5.1 Exercises 219
10.6 Learning More 220
Reference 220

11 Modelling for Visualisation 221
11.1 Introduction 221
11.2 Removing Trend 222
11.2.1 Exercises 226
11.3 Texas Housing Data 226
11.3.1 Exercises 230
11.4 Visualising Models 230
11.5 Model-Level Summaries 232
11.5.1 Exercises 234
11.6 Coefficient-Level Summaries 235
11.6.1 Exercises 237
11.7 Observation Data 238
11.7.1 Exercises 239
Reference 240

12 Programming with ggplot2 241
12.1 Introduction 241
12.2 Single Components 242
12.2.1 Exercises 243
12.3 Multiple Components 244
12.3.1 Plot Components 245
12.3.2 Annotation 245
12.3.3 Additional Arguments 246
12.3.4 Exercises 247
12.4 Plot Functions 247
12.4.1 Indirectly Referring to Variables 249

12.4.2 The Plot Environment 251
12.4.3 Exercises 251
12.5 Functional Programming 252
12.5.1 Exercises 253

Index 255

R Code index 259