

Table of Contents

Preface	i
Why maps?	i
Why Julia?	i
Why not R?	ii
Why not Python?	ii
Why not QGIS/ArcGIS?	ii
Who this book is for	iii
Table of Contents	iv
1 Getting started with Julia	1
1.1 Overview	1
1.2 Using the terminal	2
1.3 Download and install juliaup	2
1.3.1 Configure	3
1.3.2 Try it out	4
1.3.3 Special characters for comments	4
1.4 Using an editor	5
1.5 Creating a project environment	6
2 Getting geographic boundaries	10
2.1 Two types of data required	10
2.2 Encoding geometry	10
2.3 Shapefiles	11
2.4 GeoJSON files	13
2.5 GeoPackage files	13
3 Bringing in data	14
3.1 Downloaded files and spreadsheets	14
3.2 How functions work	15
3.3 Overdecorated spreadsheets	15
3.4 Fetching from URL	20
3.5 Wikitables and other HTML tables	21
3.6 PDF files	21
4 Geometry plus data makes spatial data	22
4.1 Joins	22
4.2 With default color options	23
5 Working with DataFrames	25
5.1 Reading from a CSV file	25
5.2 Inspecting a DataFrame	26
5.3 Renaming a DataFrame	27
5.4 Copying a DataFrame	27
5.5 Selecting columns	28
5.6 Reordering columns	28

5.7	Renaming columns	28
5.8	Copying columns from one DataFrame to another	28
5.9	Deleting rows	28
5.10	Adding rows	28
5.11	Sorting rows	29
5.12	Dealing with missing data	29
5.13	Addressing specific locations	30
5.14	Editing data	30
5.15	Using the underscore to make numbers easier to enter	30
5.16	Subset(filter) rows	31
5.17	Grouping rows	31
5.18	Summarizing rows	31
5.19	Adding row or column totals	33
5.20	Joining DataFrames	35
	5.20.1 Inner joins	35
	5.20.2 Left joins	36
	5.20.3 Outer joins	36
6	Crosswalks	37
7	Basic Plotting	39
7.1	The basemap	39
7.2	Focus on contiguous United States	40
7.3	Special note for Linux running as a “headless server”	43
8	Typography	44
8.1	Types of fonts	44
8.2	Fonts are functional	44
8.3	Selecting fonts	45
8.4	Getting and installing fonts	45
9	Layout	47
9.1	Data	47
9.2	Setup	48
9.3	Figure	51
9.4	Geometry	51
9.5	Title	52
9.6	Title, Legend, Table and Caption	52
9.7	Completed example	53
10	Colors	54
10.1	Color theory	54
10.2	Colorspaces	55
10.3	Named colors	56
10.4	How many colors to use?	56
10.5	Choosing colors to go together: the three types of palettes	57

10.5.1	Sequential	57
10.5.2	Diverging	57
10.5.3	Qualitative	58
10.6	Pre-selected palettes	58
10.6.1	ColorSchemes	58
10.6.2	ColorBrewer	59
10.6.3	ColorBrewer finding aids	60
10.7	Custom palettes	63
11	Counts data	64
11.1	Dividing a continuous variable into categories	68
11.1.1	Equal interval breaks	70
11.1.2	Quantile breaks	71
11.1.3	Fisher–Jenks natural breaks	71
11.1.4	K–means breaks	71
11.2	Choice of binning strategy	71
11.2.1	Uniformly distributed data	72
11.2.2	Skewed data	73
11.2.3	Clustered data	76
12	When color is not an option	78
12.1	Shades of gray	78
12.2	Patterns	79
13	Insets and outsets	81
13.1	Insetting Alaska and Hawaii	81
13.2	Showing smaller states	82
13.3	Handwork	83
14	Putting it all together	85
14.1	Basic script	85
14.1.1	Data	85
14.1.2	Plotting the basemap	85
14.2	Creating the legend	87
14.3	Adding the caption	88
14.4	Adding the summary table	88
14.5	Adding the explanatory text	88
14.6	Adding a north arrow	88
14.7	Adding a scale bar	89
14.8	Adjusting the relative size of the columns	89
15	Distance	91
15.1	Setup	91
15.2	Workflow	92
15.3	load_and_prepare_data	93
15.4	Distance calculations	93

15.5	Removing duplicates	93
16	Radius maps	95
17	Network data	97
17.1	Graph theory in brief	97
17.2	Undirected	97
17.3	Unidirectional	99
17.4	Characterizing connections	99
17.4.1	Betweenness (upper map)	101
17.4.2	Degree centrality (lower map)	101
17.4.3	Spatial communities	102
18	Bivariate maps	103
19	Dot density maps	109
19.1	Rationale for dot density maps	109
19.2	Application of the 80/20 rule	110
20	Geo-referenced rasters	111
21	Catchment areas	115
22	Cartograms	117
22.1	Bubble plots	117
22.2	Bin plots	120
22.3	Data only	121
23	Small multiples	122
24	Spatial reasoning	123
24.1	Modern analysis	125
24.1.1	Spatial distribution pattern	125
24.1.2	Temporal dynamics	126
24.1.3	Distance relationships	127
24.1.4	Statistical significance testing	128
24.1.5	Point pattern analysis	129
24.1.6	Density surface analysis	130
24.2	Comparative pump analysis	130
24.2.1	Methodological strengths	131
24.2.2	Historical context and implications	131
24.2.3	Limitations	131
24.2.4	Conclusion	131
24.3	The data	132
24.4	The script	132
25	Geometries	133
25.1	Census	133
25.2	Multiagency	133
25.3	Federal Reserve Bank	133
25.4	Environmental Protection Agency EnviroAtlas	134

25.5	US Geological Survey National Map	134
25.6	Other	134
25.7	Coordinate reference systems	134
26	Official data	135
26.1	European Union	135
26.2	United Kingdom	136
26.3	United States	137
26.3.1	Census Bureau	137
26.3.2	Federal Reserve Bank of St. Louis	138
26.3.3	US government agency data available through data.gov .	138
26.4	Other nations	139
26.5	States	141
26.6	Local	142
27	Research Data	144
28	Common error messages	148
28.1	Top tips	148
28.2	Typos, capitalization, punctuation	149
28.3	“Error on line ...”	150
28.4	Method errors	150
28.5	Type mismatch	150
28.6	Other immutable objects	153
28.7	Missing broadcast operator	153
28.8	Unquoted string interpolation expression	153
28.9	Incorrect array indexing	154
28.10	Assignment vs comparison confusion	154
28.11	Wrong function call syntax	155
28.12	Type annotation syntax error	155
28.13	Missing commas in function arguments	155
28.14	Incorrect comprehension syntax	155
28.15	Missing packages	156
28.16	Updating package failures	157
29	Functions, macros and scripts	159
29.1	Functions	159
29.2	Macros	160
29.3	Scripts	163
30	Regular expressions	166
30.1	Basic	166
30.2	Character classes and quantifiers	167
30.3	Capture groups and extraction	167
30.4	Multiline and complex document parsing	167
30.5	Working with Unicode and international text	168

30.6	Using with replace	169
31	Traditional statistics	170
31.1	Core statistics packages	170
31.2	Regression and modeling	171
31.3	Hypothesis testing and analysis	171
31.4	Bayesian statistics	171
31.5	Time series	172
31.6	Specialized areas	172
31.7	Integration with data ecosystem	172
32	juliaup	173
33	Environmental stewardship	174
33.1	What is an environment?	174
33.2	The two environments	174
33.3	Creating the project	175
33.4	Adding dependencies	175
33.5	Making dependencies active	175
33.6	Updating dependencies	175
33.7	Precompilation errors	176
33.8	Conservative approach to updating	178
34	Documentation	179
34.1	Line comments	179
34.2	Docstrings	179
34.3	Module documentation	180
35	Curating data and programs	181
35.1	Original data	181
35.2	Derived data	183
35.3	Local repository on a removable drive	183
35.4	Restoring files	185
36	Using databases for persistent storage and spatial calculations	188
36.1	PostgreSQL relational database	188
36.2	Setting up the server	189
37	Spatial databases	190
37.1	PostGIS capabilities	190
37.2	How to set up	191
38	Using Artificial intelligence effectively	192
38.1	Documenting functions	192
38.2	Debugging	193
38.3	Planning	193
38.4	Implementation	195
39	Sites	196
40	Books and Journals	197

40.1	Books	197
40.1.1	Textbooks	197
40.1.2	Map design	197
40.1.3	Census	197
40.1.4	Spatial databases	197
40.1.5	Visualization	198
40.2	Journals	198
41	Appendix: Coordinate reference systems	200
	Acknowledgements	202
	Bibliography	204
	Index	208
	Colophon	212
	About the author	213

List of Figures

Figure 1	<i>The Map</i> by Mary Cassatt, collection of the Metropolitan Museum of Art	i
Figure 2	Example Thematic Map	1
Figure 3	Julia REPL	4
Figure 4	Example of an Overdecorated Spreadsheet	16
Figure 5	Map of US Population Taking Default Colors	23
Figure 6	Typical Plot Accepting All the Defaults of a TIGER Shapefile	39
Figure 7	Some of the Far-flung Geographies	40
Figure 8	Contiguous US	41
Figure 9	Contiguous US Without Graticule	42
Figure 10	Contiguous US with Albers Equal Area Projection	43
Figure 11	Layout Elements	47
Figure 12	Data Layer	48
Figure 13	Pink Named Colors	56
Figure 14	Sequential Color Palette Example of Water Area by State	57
Figure 15	Diverging Color Palette Example of Water Area by State	58
Figure 16	Qualitative Color Palette Example of Water Area by State	58
Figure 17	Some ColorSchemes	59
Figure 18	Blues_5, Reds_5 and Greens_5 ColorBrewer Palettes	62
Figure 19	Population and Death Counts, GDP and Deportations	64
Figure 20	Death Counts Compared to Death Rates	65
Figure 21	Raw Population at County Scale	67
Figure 22	Logarithmic Transformation of Population	68
Figure 23	Population Histogram of 2023 Estimated US Population	69
Figure 24	Application of Four Binning Algorithms	70
Figure 25	Sizes of Counties under 10,000 Population	72

Figure 26	Small Counties with Equal Interval Coloring	73
Figure 27	Distributions of Small County Margins	74
Figure 28	Small Counties with Quantile Coloring	76
Figure 29	Small Counties with Fisher-Jenks Coloring	77
Figure 30	Shades of Gray	78
Figure 31	Patterned Map	80
Figure 32	Alaska and Hawaii Insets	81
Figure 33	Use of Regional Map to Show Detail	82
Figure 34	Finished in Vector Editor Application	84
Figure 35	Plot with Optional Elements	90
Figure 36	Example of a Radius Map	96
Figure 37	Undirected Graph Map Example	98
Figure 38	Flow Map Example	99
Figure 39	Betweenness and Centrality	100
Figure 40	Regional Clustering	102
Figure 41	Linear Regression Model of Ohio Election	104
Figure 42	Density Plot of Ohio Election	105
Figure 43	Bivariate Map Example	106
Figure 44	Dot Density Map Example	110
Figure 45	Shaded Relief Map	111
Figure 46	Contiguous US at Night	112
Figure 47	Example of a Geo-referenced Raster Background	113
Figure 48	Example of Voronoi Plot	116
Figure 49	Bubble Chart	118
Figure 50	Example of a Bin Plot Cartogram	120
Figure 51	Symbolic Cartogram Example	121
Figure 52	Small Multiples Map Illustration	122
Figure 53	John Snow's Map of the Golden Square Cholera Outbreak	124
Figure 54	Stylized Map of Cholera Deaths	125
Figure 55	Time Series of the Outbreak	126
Figure 56	Heatmap of Simulated Distribution of First 70 Cases	127
Figure 57	Permutation Test Result	128
Figure 58	Permutation Analysis	129
Figure 59	Density Surface	130
Figure 60	A chart invented by Playfair.	213