

# Table of Contents

<b>Preface</b>	1
<b>Chapter 1: Pandas Foundations</b>	15
<b>Introduction</b>	15
<b>Dissecting the anatomy of a DataFrame</b>	16
Getting ready	16
How to do it...	16
How it works...	17
There's more...	18
See also	18
<b>Accessing the main DataFrame components</b>	18
Getting ready	18
How to do it...	19
How it works...	20
There's more...	21
See also	21
<b>Understanding data types</b>	22
Getting ready	23
How to do it...	23
How it works...	23
There's more...	24
See also	24
<b>Selecting a single column of data as a Series</b>	24
Getting ready	24
How to do it...	25
How it works...	25
There's more...	26
See also	27
<b>Calling Series methods</b>	27
Getting ready	28
How to do it...	28
How it works...	32
There's more...	33
See also	34
<b>Working with operators on a Series</b>	34

Getting ready	35
How to do it...	35
How it works...	37
There's more...	37
See also	38
<b>Chaining Series methods together</b>	38
Getting ready	38
How to do it...	39
How it works...	40
There's more...	41
<b>Making the index meaningful</b>	42
Getting ready	42
How to do it...	42
How it works...	43
There's more...	43
See also	43
<b>Renaming row and column names</b>	43
Getting ready	43
How to do it...	44
How it works...	44
There's more...	45
<b>Creating and deleting columns</b>	45
Getting ready	46
How to do it...	46
How it works...	48
There's more...	49
See also	49
<b>Chapter 2: Essential DataFrame Operations</b>	51
<b>Introduction</b>	51
<b>Selecting multiple DataFrame columns</b>	52
Getting ready	52
How to do it...	52
How it works...	53
There's more...	53
<b>Selecting columns with methods</b>	54
Getting ready	54
How it works...	54
How it works...	57
There's more...	57

See also	58
<b>Ordering column names sensibly</b>	59
Getting ready	59
How to do it...	59
How it works...	62
There's more...	62
See also	63
<b>Operating on the entire DataFrame</b>	63
Getting ready	63
How to do it...	63
How it works...	65
There's more...	66
<b>Chaining DataFrame methods together</b>	66
Getting ready	66
How to do it...	66
How it works...	68
There's more...	68
See also	69
<b>Working with operators on a DataFrame</b>	69
Getting ready	70
How to do it...	71
How it works...	72
There's more...	73
See also	73
<b>Comparing missing values</b>	74
Getting ready	74
How to do it...	75
How it works...	77
There's more...	77
<b>Transposing the direction of a DataFrame operation</b>	78
Getting ready	78
How to do it...	78
How it works...	80
There's more...	81
See also	81
<b>Determining college campus diversity</b>	81
Getting ready	82
How to do it...	82
How it works...	86

There's more...	86
See also	87
<b>Chapter 3: Beginning Data Analysis</b>	<b>89</b>
<b>Introduction</b>	<b>89</b>
<b>Developing a data analysis routine</b>	<b>90</b>
Getting ready	90
How to do it...	90
How it works...	93
There's more...	94
Data dictionaries	95
See also	96
<b>Reducing memory by changing data types</b>	<b>96</b>
Getting ready	96
How to do it...	96
How it works...	99
There's more...	101
See also	102
<b>Selecting the smallest of the largest</b>	<b>103</b>
Getting ready	103
How to do it...	103
How it works...	104
There's more...	105
<b>Selecting the largest of each group by sorting</b>	<b>105</b>
Getting ready	105
How to do it...	105
How it works...	107
There's more...	107
<b>Replicating nlargest with sort_values</b>	<b>108</b>
Getting ready	108
How to do it...	109
How it works...	110
There's more...	111
<b>Calculating a trailing stop order price</b>	<b>111</b>
Getting ready	112
How to do it...	112
How it works...	113
There's more...	114
See also	114
<b>Chapter 4: Selecting Subsets of Data</b>	<b>115</b>

<b>Introduction</b>	115
<b>Selecting Series data</b>	116
Getting ready	116
How to do it...	117
How it works...	118
There's more...	119
See also	119
<b>Selecting DataFrame rows</b>	119
Getting ready	120
How to do it...	120
How it works...	122
There's more...	122
See also	122
<b>Selecting DataFrame rows and columns simultaneously</b>	123
Getting ready	123
How to do it...	124
How it works...	125
There's more...	126
<b>Selecting data with both integers and labels</b>	126
Getting ready	126
How to do it...	126
How it works...	127
There's more...	127
See also	128
<b>Speeding up scalar selection</b>	128
Getting ready	128
How to do it...	129
How it works...	130
There's more...	130
<b>Slicing rows lazily</b>	130
Getting ready	130
How to do it...	131
How it works...	132
There's more...	133
<b>Slicing lexicographically</b>	133
Getting ready	134
How to do it...	134
How it works...	135
There's more...	135

## Chapter 5: Boolean Indexing

<b>Introduction</b>	137
<b>Calculating boolean statistics</b>	138
Getting ready	138
How to do it...	138
How it works...	140
There's more...	141
See also	141
<b>Constructing multiple boolean conditions</b>	141
Getting ready	141
How to do it...	141
How it works...	142
There's more...	143
See also	144
<b>Filtering with boolean indexing</b>	144
Getting ready	144
How to do it...	144
How it works...	146
There's more...	147
See also	147
<b>Replicating boolean indexing with index selection</b>	147
Getting ready	147
How to do it...	147
How it works...	149
There's more...	149
<b>Selecting with unique and sorted indexes</b>	150
Getting ready	150
How to do it...	150
How it works...	152
There's more...	152
See also	153
<b>Gaining perspective on stock prices</b>	153
Getting ready	154
How to do it...	154
How it works...	156
There's more...	156
See also	157
<b>Translating SQL WHERE clauses</b>	157
Getting ready	158

How to do it...	159
How it works...	160
There's more...	161
See also	162
<b>Determining the normality of stock market returns</b>	162
Getting ready	162
How to do it...	162
How it works...	165
There's more...	165
See also	166
<b>Improving readability of boolean indexing with the query method</b>	167
Getting ready	167
How to do it...	167
How it works...	168
There's more...	168
See also	169
<b>Preserving Series with the where method</b>	169
Getting ready	169
How to do it...	169
How it works...	172
There's more...	173
See also	173
<b>Masking DataFrame rows</b>	173
Getting ready	173
How to do it...	174
How it works...	176
There's more...	177
See also	177
<b>Selecting with booleans, integer location, and labels</b>	177
Getting ready	177
How to do it...	178
How it works...	181
There's more...	181
See also	182
<b>Chapter 6: Index Alignment</b>	183
<b>Introduction</b>	183
<b>Examining the Index object</b>	184
Getting ready	184
How to do it...	184

How it works...	185
There's more...	186
See also	186
<b>Producing Cartesian products</b>	187
Getting ready	187
How to do it...	187
How it works...	188
There's more...	189
See also	190
<b>Exploding indexes</b>	190
Getting ready	190
How to do it...	190
How it works...	192
There's more...	193
<b>Filling values with unequal indexes</b>	193
Getting ready	193
How to do it...	194
How it works...	196
There's more...	197
<b>Appending columns from different DataFrames</b>	199
Getting ready	199
How to do it...	200
How it works...	201
There's more...	202
See also	203
<b>Highlighting the maximum value from each column</b>	203
Getting ready	203
How to do it...	203
How it works...	207
There's more...	209
See also	209
<b>Replicating idxmax with method chaining</b>	209
Getting ready	210
How to do it...	210
How it works...	214
There's more...	215
<b>Finding the most common maximum</b>	216
Getting ready	216
How to do it...	216

How it works...	217
There's more...	218
<b>Chapter 7: Grouping for Aggregation, Filtration, and Transformation</b>	<b>219</b>
<b>Introduction</b>	219
<b>Defining an aggregation</b>	220
Getting ready	221
How to do it...	221
How it works...	222
There's more...	223
See also	223
<b>Grouping and aggregating with multiple columns and functions</b>	224
Getting ready	224
How to do it...	224
How it works...	226
There's more...	226
<b>Removing the MultiIndex after grouping</b>	227
Getting ready	227
How to do it...	227
How it works...	230
There's more...	230
<b>Customizing an aggregation function</b>	231
Getting ready	232
How to do it...	232
How it works...	233
There's more...	234
<b>Customizing aggregating functions with *args and **kwargs</b>	235
Getting ready	236
How to do it...	236
How it works...	237
There's more...	238
See also	239
<b>Examining the groupby object</b>	240
Getting ready	240
How to do it...	240
How it works...	243
There's more...	244
See also	244
<b>Filtering for states with a minority majority</b>	245
Getting ready	246

How to do it...	246
How it works...	247
There's more...	248
See also	248
<b>Transforming through a weight loss bet</b>	249
Getting ready	249
How to do it...	249
How it works...	254
There's more...	255
See also	256
<b>Calculating weighted mean SAT scores per state with apply</b>	256
Getting ready	256
How to do it...	257
How it works...	259
There's more...	260
See also	261
<b>Grouping by continuous variables</b>	262
Getting ready	262
How to do it...	262
How it works...	264
There's more...	264
See also	266
<b>Counting the total number of flights between cities</b>	266
Getting ready	266
How to do it...	266
How it works...	268
There's more...	269
See also	270
<b>Finding the longest streak of on-time flights</b>	270
Getting ready	271
How to do it...	271
How it works...	274
There's more...	276
See also	278
<b>Chapter 8: Restructuring Data into a Tidy Form</b>	279
<b>Introduction</b>	279
<b>Tidying variable values as column names with stack</b>	281
Getting ready	282
How to do it...	282

How it works...	284
There's more...	285
See also	286
<b>Tidying variable values as column names with melt</b>	286
Getting ready	287
How to do it...	287
How it works...	289
There's more...	289
See also	290
<b>Stacking multiple groups of variables simultaneously</b>	291
Getting ready	291
How to do it...	292
How it works...	293
There's more...	294
See also	295
<b>Inverting stacked data</b>	295
Getting ready	295
How to do it...	296
How it works...	299
There's more...	300
See also	301
<b>Unstacking after a groupby aggregation</b>	301
Getting ready	301
How to do it...	301
How it works...	303
There's more...	303
See also	305
<b>Replicating pivot_table with a groupby aggregation</b>	305
Getting ready	305
How to do it...	305
How it works...	306
There's more...	307
<b>Renaming axis levels for easy reshaping</b>	308
Getting ready	308
How to do it...	308
How it works...	312
There's more...	313
<b>Tidying when multiple variables are stored as column names</b>	314
Getting ready...	314

How to do it...	314
How it works...	317
There's more...	318
See also	319
<b>Tidying when multiple variables are stored as column values</b>	319
Getting ready	319
How to do it...	319
How it works...	323
There's more...	324
See also	324
<b>Tidying when two or more values are stored in the same cell</b>	324
Getting ready...	325
How to do it..	325
How it works...	326
There's more...	327
<b>Tidying when variables are stored in column names and values</b>	328
Getting ready	328
How to do it...	328
How it works...	330
There's more...	331
<b>Tidying when multiple observational units are stored in the same table</b>	331
Getting ready	331
How to do it...	332
How it works...	337
There's more...	338
See also	339
<b>Chapter 9: Combining Pandas Objects</b>	341
<b>Introduction</b>	341
<b>Appending new rows to DataFrames</b>	341
Getting ready	342
How to do it...	342
How it works...	348
There's more...	349
<b>Concatenating multiple DataFrames together</b>	350
Getting ready	350
How to do it...	350
How it works...	353
There's more...	353
<b>Comparing President Trump's and Obama's approval ratings</b>	354

Getting ready	354
How to do it...	354
How it works...	367
There's more...	369
See also	370
<b>Understanding the differences between concat, join, and merge</b>	371
Getting ready	372
How to do it...	372
How it works...	379
There's more...	381
See also	382
<b>Connecting to SQL databases</b>	382
Getting ready	382
How to do it...	383
How it works...	385
There's more...	386
See also	388
<b>Chapter 10: Time Series Analysis</b>	389

<b>Introduction</b>	389
<b>Understanding the difference between Python and pandas date tools</b>	390
Getting ready	390
How to do it...	390
How it works...	395
There's more...	396
See also	396
<b>Slicing time series intelligently</b>	396
Getting ready	397
How to do it...	397
How it works...	400
There's more...	401
See also	402
<b>Using methods that only work with a DatetimeIndex</b>	402
Getting ready	402
How to do it...	402
How it works...	406
There's more...	407
See also	407
<b>Counting the number of weekly crimes</b>	408
Getting ready	408

How to do it...	408
How it works...	410
There's more...	411
See also	411
<b>Aggregating weekly crime and traffic accidents separately</b>	412
Getting ready	412
How to do it...	412
How it works...	414
There's more...	415
<b>Measuring crime by weekday and year</b>	416
Getting ready	416
How to do it...	416
How it works...	423
There's more...	425
See also	427
<b>Grouping with anonymous functions with a DatetimeIndex</b>	427
Getting ready	427
How to do it...	427
How it works...	429
There's more...	430
See also	431
<b>Grouping by a Timestamp and another column</b>	431
Getting ready	431
How to do it...	431
How it works...	435
There's more...	436
<b>Finding the last time crime was 20% lower with merge_asof</b>	437
Getting ready	437
How to do it...	438
How it works...	440
There's more...	441
<b>Chapter 11: Visualization with Matplotlib, Pandas, and Seaborn</b>	443
<b>Introduction</b>	443
<b>Getting started with matplotlib</b>	445
Getting ready	445
Object-oriented guide to matplotlib	446
How to do it...	448
How it works...	452
There's more...	455

See also	456
<b>Visualizing data with matplotlib</b>	456
Getting ready	456
How to do it...	456
How it works...	460
There's more...	463
See also	465
<b>Plotting basics with pandas</b>	465
Getting ready	465
How to do it...	466
How it works...	468
There's more...	469
See also	469
<b>Visualizing the flights dataset</b>	470
Getting ready	470
How to do it...	470
How it works...	478
See also	481
<b>Stacking area charts to discover emerging trends</b>	481
Getting ready	481
How to do it...	481
How it works...	484
There's more...	485
<b>Understanding the differences between seaborn and pandas</b>	486
Getting ready	487
How to do it...	487
How it works...	494
See also	495
<b>Doing multivariate analysis with seaborn Grids</b>	495
Getting ready	496
How to do it...	496
How it works...	499
There's more...	499
<b>Uncovering Simpson's paradox in the diamonds dataset with seaborn</b>	502
How to do it...	503
How it works...	507
There's more...	508
<b>Index</b>	509