

Table of Contents

Preface.....	xiii
--------------	------

Part I. The R Language

1. Introduction.....	3
Chapter Goals	3
What Is R?	3
Installing R	4
Choosing an IDE	5
Emacs + ESS	5
Eclipse/Architect	6
RStudio	6
Revolution-R	7
Live-R	7
Other IDEs and Editors	7
Your First Program	8
How to Get Help in R	8
Installing Extra Related Software	11
Summary	11
Test Your Knowledge: Quiz	12
Test Your Knowledge: Exercises	12
2. A Scientific Calculator.....	13
Chapter Goals	13
Mathematical Operations and Vectors	13
Assigning Variables	17
Special Numbers	19
Logical Vectors	20
Summary	22

Test Your Knowledge: Quiz	22
Test Your Knowledge: Exercises	23
3. Inspecting Variables and Your Workspace.....	25
Chapter Goals	25
Classes	25
Different Types of Numbers	26
Other Common Classes	27
Checking and Changing Classes	30
Examining Variables	33
The Workspace	36
Summary	37
Test Your Knowledge: Quiz	37
Test Your Knowledge: Exercises	37
4. Vectors, Matrices, and Arrays.....	39
Chapter Goals	39
Vectors	39
Sequences	41
Lengths	42
Names	42
Indexing Vectors	43
Vector Recycling and Repetition	45
Matrices and Arrays	46
Creating Arrays and Matrices	46
Rows, Columns, and Dimensions	48
Row, Column, and Dimension Names	50
Indexing Arrays	51
Combining Matrices	51
Array Arithmetic	52
Summary	54
Test Your Knowledge: Quiz	55
Test Your Knowledge: Exercises	55
5. Lists and Data Frames.....	57
Chapter Goals	57
Lists	57
Creating Lists	57
Atomic and Recursive Variables	60
List Dimensions and Arithmetic	60
Indexing Lists	61
Converting Between Vectors and Lists	64

Combining Lists	65
NULL	66
Pairlists	70
Data Frames	70
Creating Data Frames	71
Indexing Data Frames	74
Basic Data Frame Manipulation	75
Summary	77
Test Your Knowledge: Quiz	77
Test Your Knowledge: Exercises	78
6. Environments and Functions.	79
Chapter Goals	79
Environments	79
Functions	82
Creating and Calling Functions	82
Passing Functions to and from Other Functions	86
Variable Scope	89
Summary	91
Test Your Knowledge: Quiz	91
Test Your Knowledge: Exercises	91
7. Strings and Factors.	93
Chapter Goals	93
Strings	93
Constructing and Printing Strings	94
Formatting Numbers	95
Special Characters	97
Changing Case	98
Extracting Substrings	98
Splitting Strings	99
File Paths	100
Factors	101
Creating Factors	101
Changing Factor Levels	103
Dropping Factor Levels	103
Ordered Factors	104
Converting Continuous Variables to Categorical	105
Converting Categorical Variables to Continuous	106
Generating Factor Levels	107
Combining Factors	107
Summary	108

Test Your Knowledge: Quiz	108
Test Your Knowledge: Exercises	108
8. Flow Control and Loops.	111
Chapter Goals	111
Flow Control	111
if and else	112
Vectorized if	114
Multiple Selection	115
Loops	116
repeat Loops	116
while Loops	118
for Loops	120
Summary	122
Test Your Knowledge: Quiz	122
Test Your Knowledge: Exercises	122
9. Advanced Looping.	125
Chapter Goals	125
Replication	125
Looping Over Lists	127
Looping Over Arrays	132
Multiple-Input Apply	135
Instant Vectorization	136
Split-Apply-Combine	136
The plyr Package	138
Summary	141
Test Your Knowledge: Quiz	141
Test Your Knowledge: Exercises	141
10. Packages.	143
Chapter Goals	143
Loading Packages	144
The Search Path	146
Libraries and Installed Packages	146
Installing Packages	148
Maintaining Packages	150
Summary	150
Test Your Knowledge: Quiz	151
Test Your Knowledge: Exercises	151
11. Dates and Times.	153

Chapter Goals	153
Date and Time Classes	154
POSIX Dates and Times	154
The Date Class	155
Other Date Classes	156
Conversion to and from Strings	156
Parsing Dates	156
Formatting Dates	157
Time Zones	158
Arithmetic with Dates and Times	160
Lubridate	161
Summary	165
Test Your Knowledge: Quiz	165
Test Your Knowledge: Exercises	166
Part II. The Data Analysis Workflow	
12. Getting Data.....	169
Chapter Goals	169
Built-in Datasets	169
Reading Text Files	170
CSV and Tab-Delimited Files	170
Unstructured Text Files	175
XML and HTML Files	175
JSON and YAML Files	176
Reading Binary Files	179
Reading Excel Files	179
Reading SAS, Stata, SPSS, and MATLAB Files	181
Reading Other File Types	181
Web Data	182
Sites with an API	182
Scraping Web Pages	184
Accessing Databases	185
Summary	188
Test Your Knowledge: Quiz	189
Test Your Knowledge: Exercises	189
13. Cleaning and Transforming.....	191
Chapter Goals	191
Cleaning Strings	191
Manipulating Data Frames	196

Adding and Replacing Columns	196
Dealing with Missing Values	197
Converting Between Wide and Long Form	198
Using SQL	200
Sorting	201
Functional Programming	202
Summary	204
Test Your Knowledge: Quiz	205
Test Your Knowledge: Exercises	205
14. Exploring and Visualizing.....	207
Chapter Goals	207
Summary Statistics	207
The Three Plotting Systems	211
Scatterplots	212
Take 1: base Graphics	213
Take 2: lattice Graphics	218
Take 3: ggplot2 Graphics	224
Line Plots	230
Histograms	238
Box Plots	249
Bar Charts	253
Other Plotting Packages and Systems	260
Summary	261
Test Your Knowledge: Quiz	261
Test Your Knowledge: Exercises	262
15. Distributions and Modeling.....	263
Chapter Goals	263
Random Numbers	264
The sample Function	264
Sampling from Distributions	265
Distributions	266
Formulae	267
A First Model: Linear Regressions	268
Comparing and Updating Models	271
Plotting and Inspecting Models	276
Other Model Types	280
Summary	282
Test Your Knowledge: Quiz	282

Test Your Knowledge: Exercises	282
16. Programming.....	285
Chapter Goals	285
Messages, Warnings, and Errors	286
Error Handling	289
Debugging	292
Testing	294
RUnit	295
testthat	298
Magic	299
Turning Strings into Code	299
Turning Code into Strings	301
Object-Oriented Programming	302
S3 Classes	303
Reference Classes	305
Summary	310
Test Your Knowledge: Quiz	310
Test Your Knowledge: Exercises	311
17. Making Packages.....	313
Chapter Goals	313
Why Create Packages?	313
Prerequisites	313
The Package Directory Structure	314
Your First Package	315
Documenting Packages	317
Checking and Building Packages	320
Maintaining Packages	321
Summary	323
Test Your Knowledge: Quiz	323
Test Your Knowledge: Exercises	324

Part III. Appendixes

A. Properties of Variables.....	327
B. Other Things to Do in R.....	331
C. Answers to Quizzes.....	333

