R

Notes for Professionals

400+ pages

of professional hints and tricks

GoalKicker.com
Free Programming Books

Disclaimer
This is an unofficial free book created for educational purposes and is not affiliated with official R group(s) or company(s).
All trademarks and registered trademarks are the property of their respective owners

Contents

About ... 1

Chapter 1: Getting started with R Language ... 2

- Section 1.1: Installing R ... 2
- Section 1.2: Hello World! ... 3
- Section 1.3: Getting Help ... 3
- Section 1.4: Interactive mode and R scripts ... 3

Chapter 2: Variables ... 7

- Section 2.1: Variables, data structures and basic Operations ... 7

Chapter 3: Arithmetic Operators ... 10

- Section 3.1: Range and addition ... 10
- Section 3.2: Addition and subtraction ... 10

Chapter 4: Matrices ... 13

- Section 4.1: Creating matrices ... 13

Chapter 5: Formula ... 15

- Section 5.1: The basics of formula ... 15

Chapter 6: Reading and writing strings ... 17

- Section 6.1: Printing and displaying strings ... 17
- Section 6.2: Capture output of operating system command ... 18
- Section 6.3: Reading from or writing to a file connection ... 19

Chapter 7: String manipulation with stringi package ... 21

- Section 7.1: Count pattern inside string ... 21
- Section 7.2: Duplicating strings ... 21
- Section 7.3: Paste vectors ... 22
- Section 7.4: Splitting text by some fixed pattern ... 22

Chapter 8: Classes ... 23

- Section 8.1: Inspect classes ... 23
- Section 8.2: Vectors and lists ... 23
- Section 8.3: Vectors ... 24

Chapter 9: Lists ... 25

- Section 9.1: Introduction to lists ... 25
- Section 9.2: Quick Introduction to Lists ... 25
- Section 9.3: Serialization: using lists to pass information ... 27

Chapter 10: Hashmaps ... 29

- Section 10.1: Environments as hash maps ... 29
- Section 10.2: package:hash ... 32
- Section 10.3: package:listenv ... 33

Chapter 11: Creating vectors ... 35

- Section 11.1: Vectors from build in constants: Sequences of letters & month names ... 35
- Section 11.2: Creating named vectors ... 35
- Section 11.3: Sequence of numbers ... 37
- Section 11.4: seq() ... 37
- Section 11.5: Vectors ... 38
- Section 11.6: Expanding a vector with the rep() function ... 39

Chapter 12: Date and Time ... 41

- Section 12.1: Current Date and Time ... 41

Section 12.2: Go to the End of the Month 41
Section 12.3: Go to First Day of the Month 42
Section 12.4: Move a date a number of months consistently by months 42
Chapter 13: The Date class 44
Section 13.1: Formatting Dates 44
Section 13.2: Parsing Strings into Date Objects 44
Section 13.3: Dates 45
Chapter 14: Date-time classes (POSIXct and POSIXlt) 47
Section 14.1: Formatting and printing date-time objects 47
Section 14.2: Date-time arithmetic 47
Section 14.3: Parsing strings into date-time objects 48
Chapter 15: The character class 50
Section 15.1: Coercion 50
Chapter 16: Numeric classes and storage modes 51
Section 16.1: Numeric 51
Chapter 17: The logical class 53
Section 17.1: Logical operators 53
Section 17.2: Coercion 53
Section 17.3: Interpretation of NAs 53
Chapter 18: Data frames 55
Section 18.1: Create an empty data.frame 55
Section 18.2: Subsetting rows and columns from a data frame 56
Section 18.3: Convenience functions to manipulate data.frames 59
Section 18.4: Introduction 60
Section 18.5: Convert all columns of a data.frame to character class 61
Chapter 19: Split function 63
Section 19.1: Using split in the split-apply-combine paradigm 63
Section 19.2: Basic usage of split 64
Chapter 20: Reading and writing tabular data in plain-text files (CSV, TSV, etc.) 67
Section 20.1: Importing .csv files 67
Section 20.2: Importing with data.table 68
Section 20.3: Exporting .csv files 69
Section 20.4: Import multiple csv files 69
Section 20.5: Importing fixed-width files 69
Chapter 21: Pipe operators (%>% and others) 71
Section 21.1: Basic use and chaining 71
Section 21.2: Functional sequences 72
Section 21.3: Assignment with %<>% 73
Section 21.4: Exposing contents with %$% 73
Section 21.5: Creating side effects with %T>% 74
Section 21.6: Using the pipe with dplyr and ggplot2 75
Chapter 22: Linear Models (Regression) 76
Section 22.1: Linear regression on the mtcars dataset 76
Section 22.2: Using the 'predict' function 78
Section 22.3: Weighting 79
Section 22.4: Checking for nonlinearity with polynomial regression 81
Section 22.5: Plotting The Regression (base) 83
Section 22.6: Quality assessment 85
Chapter 23: data.table 87

Section 23.1: Creating a data.table 87
Section 23.2: Special symbols in data.table 88
Section 23.3: Adding and modifying columns 89
Section 23.4: Writing code compatible with both data.frame and data.table 91
Section 23.5: Setting keys in data.table 93

Chapter 24: Pivot and unpivot with data.table 95

Section 24.1: Pivot and unpivot tabular data with data.table - I 95
Section 24.2: Pivot and unpivot tabular data with data.table - II 96

Chapter 25: Bar Chart 98

Section 25.1: barplot() function 98

Chapter 26: Base Plotting 104

Section 26.1: Density plot 104
Section 26.2: Combining Plots 105
Section 26.3: Getting Started with R_Plots 107
Section 26.4: Basic Plot 108
Section 26.5: Histograms 111
Section 26.6: Matplot 113
Section 26.7: Empirical Cumulative Distribution Function 119

Chapter 27: boxplot 121

Section 27.1: Create a box-and-whisker plot with boxplot() {graphics} 121
Section 27.2: Additional boxplot style parameters 125

Chapter 28: ggplot2 128

Section 28.1: Displaying multiple plots 128
Section 28.2: Prepare your data for plotting 131
Section 28.3: Add horizontal and vertical lines to plot 133
Section 28.4: Scatter Plots 136
Section 28.5: Produce basic plots with qplot 136
Section 28.6: Vertical and Horizontal Bar Chart 138
Section 28.7: Violin plot 140

Chapter 29: Factors 143

Section 29.1: Consolidating Factor Levels with a List 143
Section 29.2: Basic creation of factors 144
Section 29.3: Changing and reordering factors 145
Section 29.4: Rebuilding factors from zero 150

Chapter 30: Pattern Matching and Replacement 152

Section 30.1: Finding Matches 152
Section 30.2: Single and Global match 153
Section 30.3: Making substitutions 154
Section 30.4: Find matches in big data sets 154

Chapter 31: Run-length encoding 156

Section 31.1: Run-length Encoding with `rle` 156
Section 31.2: Identifying and grouping by runs in base R 156
Section 31.3: Run-length encoding to compress and decompress vectors 157
Section 31.4: Identifying and grouping by runs in data.table 158

Chapter 32: Speeding up tough-to-vectorize code 159

Section 32.1: Speeding tough-to-vectorize for loops with Rcpp 159
Section 32.2: Speeding tough-to-vectorize for loops by byte compiling 159

Chapter 33: Introduction to Geographical Maps 161

Section 33.1: Basic map-making with map() from the package maps 161

Section 33.2: 50 State Maps and Advanced Choropleths with Google Viz 164
Section 33.3: Interactive plotly maps 165
Section 33.4: Making Dynamic HTML Maps with Leaflet 167
Section 33.5: Dynamic Leaflet maps in Shiny applications 168
Chapter 34: Set operations 171
Section 34.1: Set operators for pairs of vectors 171
Section 34.2: Cartesian or "cross" products of vectors 171
Section 34.3: Set membership for vectors 172
Section 34.4: Make unique / drop duplicates / select distinct elements from a vector 172
Section 34.5: Measuring set overlaps / Venn diagrams for vectors 173
Chapter 35: tidyverse 174
Section 35.1: tidyverse: an overview 174
Section 35.2: Creating tbl_df's 175
Chapter 36: Rcpp 176
Section 36.1: Extending Rcpp with Plugins 176
Section 36.2: Inline Code Compile 176
Section 36.3: Rcpp Attributes 177
Section 36.4: Specifying Additional Build Dependencies 178
Chapter 37: Random Numbers Generator 179
Section 37.1: Random permutations 179
Section 37.2: Generating random numbers using various density functions 179
Section 37.3: Random number generator's reproducibility 181
Chapter 38: Parallel processing 182
Section 38.1: Parallel processing with parallel package 182
Section 38.2: Parallel processing with foreach package 183
Section 38.3: Random Number Generation 184
Section 38.4: mcparallelDo 184
Chapter 39: Subsetting 186
Section 39.1: Data frames 186
Section 39.2: Atomic vectors 187
Section 39.3: Matrices 188
Section 39.4: Lists 190
Section 39.5: Vector indexing 191
Section 39.6: Other objects 192
Section 39.7: Elementwise Matrix Operations 192
Chapter 40: Debugging 194
Section 40.1: Using debug 194
Section 40.2: Using browser 194
Chapter 41: Installing packages 196
Section 41.1: Install packages from GitHub 196
Section 41.2: Download and install packages from repositories 197
Section 41.3: Install package from local source 198
Section 41.4: Install local development version of a package 198
Section 41.5: Using a CLI package manager -- basic pacman usage 199
Chapter 42: Inspecting packages 200
Section 42.1: View Package Version 200
Section 42.2: View Loaded packages in Current Session 200
Section 42.3: View package information 200
Section 42.4: View package's built-in data sets 200

Click here to download full PDF material