Contents

About New Edition **11**
Revision 2021 12

Preface **13**
Supplement Material 15
Content for Instructors 16

1 Introduction **17**
1.1 What is R 18
1.2 Why Choose R 18
1.3 What Can You Do With R and RStudio? 19
1.4 Installing R 21
1.5 Installing RStudio 24
1.6 Resources in the Web 25
1.7 Structure and Organization 25
1.8 Exercises 27

2 Basic Operations in R **31**
2.1 Working With R 31
2.2 Objects in R 33
2.3 International and Local Formats 34
2.4 Types of Files in R 35
2.5 Explaining the RStudio Screen 36
2.6 R Packages 39
2.6.1 Installing Packages from CRAN 42
2.6.2 Installing Packages from Github 42
2.6.3 Loading Packages 43
2.6.4 Upgrading Packages 45
2.7 Running Scripts from RStudio 46

2.7.1 RStudio shortcuts 48
2.8 Testing and Debugging Code 49
2.9 Creating Simple Objects 51
2.10 Creating Vectors 52
2.11 Knowing Your Environment and Objects 53
2.12 Displaying and Formatting Output 55
2.12.1 Customizing the Output 57
2.13 Finding the Size of Objects 59
2.14 Selecting Elements from an Atomic Vector 62
2.15 Removing Objects from the Memory 65
2.16 Displaying and Setting the Working Directory 66
2.17 Canceling Code Execution 68
2.18 Code Comments 69
2.19 Looking for Help 71
2.20 Using Code Completion with *tab* 72
2.21 Interacting with Files and the Operating System 76
2.21.1 Listing Files and Folders 76
2.21.2 Deleting Files and Directories 79
2.21.3 Downloading Files from the Internet 80
2.21.4 Using Temporary Files and Directories 81
2.22 Exercises 82

3 Writing Research Scripts 85
3.1 Stages of Research 85
3.2 Folder Structure 87
3.3 Important Aspects of a Research Script 90
3.4 Exercises 92

4 Importing Data from Local Files 93
4.1 *csv* files 94
4.1.1 Importing Data 97
4.1.2 Exporting Data 101
4.2 *Excel* Files (*xls* and *xlsx*) 102
4.2.1 Importing Data 103
4.2.2 Exporting Data 104
4.3 *RData* and *rds* Files 107
4.3.1 Importing Data 107
4.3.2 Exporting Data 108
4.4 *fst* files 109
4.4.1 Importing Data 110
4.4.2 Exporting Data 110
4.4.3 Timing the *fst* format 110
4.5 SQLite Files 112

4.5.1 Importing Data . . . 113
4.5.2 Exporting Data . . . 114
4.6 Unstructured Data and Other Formats . . . 115
4.6.1 Importing Data . . . 116
4.6.2 Exporting Data . . . 117
4.7 How to Select a Format . . . 118
4.8 Exercises . . . 119

5 Importing Data from the Internet 121
5.1 Package `GetQuandlData` . . . 121
5.2 Package `BatchGetSymbols` . . . 127
5.3 Package `simfinR` . . . 133
5.3.1 Example 01 - Apple Inc Annual Profit . . . 134
5.3.2 Example 02 - Quarterly Net Profit of Many Companies 137
5.3.3 Example 03: Fetching price data . . . 139
5.4 Package `tidyquant` . . . 140
5.5 Package `Rbitcoin` . . . 142
5.6 Other Packages . . . 144
5.7 Accessing Data from Web Pages (*webscraping*) . . . 144
5.7.1 Scraping the Components of the SP500 Index from Wikipedia . . . 145
5.7.2 Scraping the Website of the Reserve Bank of Australia 147
5.8 Exercises . . . 149

6 Dataframes and other objects 153
6.1 `Dataframes` . . . 154
6.1.1 Creating `dataframes` . . . 154
6.1.2 Inspecting a Dataframe . . . 156
6.1.3 The *pipeline* Operator (`%>%`) . . . 158
6.1.4 Accessing Columns . . . 159
6.1.5 Modifying a `dataframe` . . . 162
6.1.6 Filtering rows of a `dataframe` . . . 164
6.1.7 Sorting a `dataframe` . . . 165
6.1.8 Combining and Aggregating `dataframes` . . . 168
6.1.9 Extensions of the `dataframe` Class . . . 172
6.1.10 Other Useful Functions for Handling `dataframes` . . . 175
6.2 `Lists` . . . 177
6.2.1 Creating `lists` . . . 177
6.2.2 Accessing the Elements of a `list` . . . 179
6.2.3 Adding and Removing Elements from a `list` . . . 182
6.2.4 Processing the Elements of a `list` . . . 184
6.2.5 Other Useful Functions . . . 185
6.3 `Matrices` . . . 186

6.3.1 Selecting Elements from a `matrix` 189
6.3.2 Other Useful Functions 191
6.4 Exercises 193

7 Basic Object Classes **195**
7.1 `Numeric` Objects 195
7.1.1 Creating and Manipulating `numeric` Objects 196
7.1.2 Creating a `numeric` Sequence 198
7.1.3 Creating Vectors with Repeated Elements 199
7.1.4 Creating Vectors with Random Numbers 199
7.1.5 Accessing the Elements of a `numeric` Vector 204
7.1.6 Modifying and Removing Elements of a `numeric` Vector 206
7.1.7 Creating Groups 207
7.1.8 Other Useful Functions 209
7.2 `Character` Objects 211
7.2.1 Creating a Simple `character` Object 211
7.2.2 Creating Structured `character` Objects 212
7.2.3 `character` Constants 213
7.2.4 Selecting Pieces of a Text Object 214
7.2.5 Finding and Replacing Characters of a Text 215
7.2.6 Splitting Text 218
7.2.7 Finding the Number of Characters in a Text 219
7.2.8 Generating Combinations of Text 220
7.2.9 Encoding of `character` Objects 221
7.2.10 Other Useful Functions 222
7.3 `Factor` Objects 223
7.3.1 Creating `factors` 223
7.3.2 Modifying `factors` 224
7.3.3 Converting `factors` to Other Classes 226
7.3.4 Creating Contingency Tables 227
7.3.5 Other Useful Functions 228
7.4 `Logical` Objects 229
7.4.1 Creating `logical` Objects 229
7.5 Date and Time 232
7.5.1 Creating Simple Dates 232
7.5.2 Creating a Sequence of `Dates` 234
7.5.3 Operations with `Dates` 236
7.5.4 Dealing with Time 238
7.5.5 Customizing the Format of Dates and Times 239
7.5.6 Extracting Elements of a Date 241
7.5.7 Find the Current Date and Time 242
7.5.8 Other Useful Functions 243

7.6 Missing Data - NA (*Not available*) . . . 246
7.6.1 Defining NA Values . . . 246
7.6.2 Finding and Replacing NA . . . 247
7.6.3 Other Useful Functions . . . 249
7.7 Exercises . . . 249

8 Programming and Data Analysis 253
8.1 R Functions . . . 253
8.2 Using for Loops . . . 263
8.3 Conditional Statements (if, else, switch) . . . 270
8.4 Using apply Functions . . . 274
8.4.1 Using lapply . . . 274
8.4.2 Using sapply . . . 277
8.4.3 Using tapply . . . 280
8.4.4 Using mapply . . . 283
8.4.5 Using apply . . . 284
8.4.6 Using by . . . 285
8.5 Using package purrr . . . 287
8.5.1 Function map_* . . . 287
8.5.2 Function safely . . . 288
8.5.3 Function pmap . . . 291
8.6 Data Manipulation with Package dplyr . . . 294
8.6.1 Group Operations with dplyr . . . 294
8.6.2 Complex Group Operations with dplyr . . . 297
8.7 Exercises . . . 299

9 Cleaning and Structuring Data 301
9.1 The Format of a dataframe . . . 301
9.1.1 Converting a dataframe Structure (long and wide) . . . 303
9.2 Converting lists into dataframes . . . 305
9.3 Removing Outliers . . . 310
9.3.1 Treating Outliers in dataframes . . . 313
9.4 Inflation and Price Data . . . 317
9.5 Modifying Time Frequency and Aggregating Data . . . 319
9.6 Exercises . . . 320

10 Creating and Saving Figures with ggplot2 323
10.1 The ggplot2 Package . . . 324
10.2 Using Graphics Windows . . . 325
10.3 Creating Figures with Function qplot . . . 326
10.4 Creating Figures with Function ggplot . . . 328
10.4.1 The US Yield Curve . . . 334
10.5 Using Themes . . . 339

10.6 Creating Panels with `facet_wrap` . . . 344
10.7 Using the Pipeline . . . 348
10.8 Creating Statistical Graphics . . . 351
10.8.1 Creating Histograms . . . 351
10.8.2 Creating *boxplot* Figures . . . 355
10.8.3 Creating *QQ* Plots . . . 358
10.9 Saving Graphics to a File . . . 361
10.10Exercises . . . 362

11 Financial Econometrics with R **365**
11.1 Linear Models (OLS) . . . 366
11.1.1 Simulating a Linear Model . . . 366
11.1.2 Estimating a Linear Model . . . 368
11.1.3 Statistical Inference in Linear Models . . . 376
11.2 Generalized Linear Models (GLM) . . . 382
11.2.1 Simulating a GLM Model . . . 383
11.2.2 Estimating a GLM Model . . . 384
11.3 Panel Data Models . . . 389
11.3.1 Simulating Panel Data Models . . . 389
11.3.2 Estimating Panel Data Models . . . 392
11.4 Arima Models . . . 398
11.4.1 Simulating Arima Models . . . 399
11.4.2 Estimating Arima Models . . . 400
11.4.3 Forecasting Arima Models . . . 403
11.5 GARCH Models . . . 404
11.5.1 Simulating Garch Models . . . 404
11.5.2 Estimating Garch Models . . . 407
11.5.3 Forecasting Garch Models . . . 410
11.6 Regime Switching Models . . . 411
11.6.1 Simulating Regime Switching Models . . . 413
11.6.2 Estimating Regime Switching Models . . . 417
11.6.3 Forecasting Regime Switching Models . . . 420
11.7 Dealing with Several Models . . . 421
11.7.1 Using `tapply` and `sapply` . . . 421
11.7.2 Using `by` . . . 422
11.7.3 Using `dplyr::group_by` . . . 425
11.8 Exercises . . . 426

12 Reporting Results **429**
12.1 Reporting Tables . . . 430
12.2 Reporting Models . . . 433
12.3 Creating Reports with *RMarkdown* . . . 435
12.4 Exercises . . . 441

13 Optimizing Code **443**
13.1 Optimizing your Programming Time 443
13.2 Optimizing Code Speed 447
13.2.1 Profiling Code 448
13.2.2 Simple Strategies to Improve Code Speed 452
13.2.3 Using C++ code (package `Rcpp`) 454
13.2.4 Using cache (package `memoise`) 456
13.3 Exercises 463