Contents

About the Authors v

Preface vii

Supplementary Material xi

1. Introduction 1

1.1 Why Should We Learn R? 2
1.2 R: Getting Started 3
1.2.1 Installing R on windows 3
1.2.2 RStudio: A better way to run R 4
1.2.3 Installing RStudio for windows 4
1.2.4 RStudio GUI/IDE 5
1.3 Proxy Setup 6
1.4 R Packages 7
1.4.1 Installing packages 8
1.5 R Core Packages 8
1.6 Task Views in R: Introduction and Installation 9
1.7 Getting Help 10
1.7.1 Getting help from the web 13
1.8 Summary 14

2. Data Objects in R 15

2.1 Introduction 15
2.2 Data Types 15
2.2.1 Double 16
2.2.2 Integer 16
2.2.3 Complex 16
2.2.4 Logical 17
2.2.5 Character 17

2.2.6 Factor . . . 18
2.2.7 Date and time . . . 18
2.2.8 Missing data in R . . . 20
2.3 Data Structures in R . . . 20
2.3.1 Vector . . . 20
2.3.2 Matrices . . . 21
2.3.2.1 Matrix manipulations . . . 22
2.3.3 Arrays . . . 23
2.3.4 Data frames . . . 25
2.3.5 Lists . . . 27
2.4 Summary . . . 28

3. Data Handling in R 29
3.1 Introduction . . . 29
3.2 Importing and Exporting Tabular Data in R . . . 29
3.2.1 Reading data from a text file . . . 29
3.2.2 Reading data from CSV file . . . 32
3.2.3 Reading data from excel files . . . 33
3.2.4 Reading data from databases . . . 33
3.2.5 Reading from data files from other statistical systems . . . 33
3.2.5.1 Reading SPSS data file . . . 34
3.2.5.2 Reading Stata data file . . . 36
3.2.5.3 Reading Matlab data files . . . 37
3.3 Data Preprocessing in R . . . 37
3.3.1 Extracting data . . . 37
3.3.2 Combining data frames . . . 39
3.3.3 Sub-setting and logical data selection . . . 41
3.4 Summary . . . 43

4. R Programming and Control Flow 45
4.1 Introduction . . . 45
4.2 Control Flow . . . 46
4.2.1 If-else conditional statements . . . 46
4.2.2 Using `switch` . . . 47
4.2.3 Loops . . . 48
4.3 Functions in R . . . 50
4.4 Summary . . . 52

5. Data Exploration 53
5.1 Introduction . . . 53
5.2 Summary Statistics . . . 53
5.3 Example: Descriptive Statistics of Stock Returns . . . 59
5.3.1 Introduction . . . 59
5.3.2 Importing the data . . . 61
5.3.2.1 Using the `describe` function . . . 66
5.3.2.2 Using `stat.desc` from package pastecs . . . 67
5.3.3 Some basic plots . . . 71
5.4 Summary . . . 74

6. Graphics in R 75
6.1 Introduction . . . 75
6.2 Basic Plots in R . . . 76
6.3 Exporting Graphics . . . 86
6.4 R Graphical Parameters . . . 88
6.5 Introduction to ggplot2 . . . 91
6.5.1 Getting started with `qplot` . . . 92
6.5.2 Layered graphics using `ggplot` . . . 93
6.6 Transforming Data from Wide to Long Format . . . 100
6.7 Summary . . . 102

7. Regression Analysis-I 103
7.1 Introduction . . . 103
7.2 OLS . . . 103
7.3 QR . . . 107
7.3.1 Estimating QR . . . 110
7.4 Example: Fama French Factor Model and Multiple Regression . . . 114
7.4.1 Introduction . . . 114
7.4.2 Data . . . 116
7.4.2.1 Data preprocessing . . . 117
7.4.3 OLS regression analysis of the Fama French three factor model . . . 121
7.4.4 Quantile analysis of the Fama French three factor model . . . 124
7.5 Summary . . . 129

8. Regression Analysis-II 131
8.1 Introduction . . . 131
8.2 Panel Data Linear Regression . . . 131
8.2.1 Fixed and Random effects using the *plm* package . . . 133
8.2.1.1 Fixed effect estimation . . . 134
8.2.1.2 Random effect estimation . . . 136
8.2.1.3 Panel or OLS . . . 137
8.2.1.4 Fixed effect or random effect . . . 139
8.3 Logistic Regression . . . 139
8.4 Example: Economic Growth and Unemployment — A Panel Analysis . . . 142
8.4.1 Data and methodology . . . 142
8.4.2 Data preprocessing . . . 143
8.4.3 Linear panel regression analysis . . . 145
8.5 Summary . . . 151

9. Time Series Analysis 153
9.1 Introduction . . . 153
9.2 Time Series-Some Properties . . . 153
9.2.1 Stochastic process . . . 153
9.2.2 Stationarity . . . 154
9.2.3 Autocorrelation function (ACF) . . . 157
9.2.4 White noise . . . 159
9.3 Autoregressive Moving Average Model (ARMA) . . . 159
9.3.1 Fitting an ARMA model . . . 161
9.4 Volatility Modelling using Generalised Autoregressive Conditional Heteroskedasticity (GARCH) . . . 166
9.4.1 Fitting a GARCH(1,1) model using the *rugarch* package . . . 167
9.5 Example: Modelling and Forecasting Daily VaR using GARCH . . . 175
9.5.1 Data and methodology . . . 175
9.5.2 VaR forecasts . . . 178
9.6 Summary . . . 182

10. Extreme Value Theory Modelling 183
10.1 Introduction . . . 183

10.2 EVT and Financial Risk Modelling 183
10.2.1 GPD and POT methods 185
10.2.1.1 Sample mean excess plot 186
10.2.1.2 Estimation of GPD 187
10.2.1.3 VaR and expected shortfall 191
10.3 EVT and Tail Dependence 195
10.3.1 Measures of tail dependence 195
10.3.1.1 Asymptotic dependence — the conventional approach 197
10.3.1.2 Asymptotic independence — an alternative measure of dependence 198
10.3.2 Estimating χ and $\bar{\chi}$: nonparametric method 202
10.4 Example: EVT VaR — A Dynamic Approach 204
10.4.1 Data and methodology 206
10.4.2 VaR forecasts using dynamic EVT model 207
10.5 Summary 209

11. Introduction to Multivariate Analysis using Copulas 211

11.1 Introduction 211
11.2 Copula 212
11.2.1 Types of copula 213
11.2.1.1 Elliptical copulas 213
11.2.1.2 Archimedean copulas 217
11.2.2 Copula selection 222
11.3 Multivariate Vine Copulas 222
11.3.1 Estimating R-Vine copulas using R 224
11.4 Example: Portfolio VaR Estimation using Vine Copula 229
11.4.1 Data and methodology 229
11.4.2 VaR forecasts using R 230
11.5 Summary 234

Bibliography 235

Index 243