Notes

Chapter 1

1. One the best available online resources for brushing up on basic math skills is Khan Academy. The following courses are recommended:

   - Linear Algebra = https://www.khanacademy.org/math/linear-algebra.
   - Differential Calculus = https://www.khanacademy.org/math/differential-calculus.
   - Integral Calculus = https://www.khanacademy.org/math/integral-calculus.
   - Probability and Statistics = https://www.khanacademy.org/math/probability.

2. The Python language is a popular scripting language with a wide variety of data-processing and analytical capabilities, as well as, huge support from academia and industry. The following links provide excellent introductory resources:

   - https://wiki.python.org/moin/BeginnersGuide/NonProgrammers.
   - http://www.youtube.com/show/python101.

   The user WiseOwlTutorials on YouTube has a nice series of introductory videos on Excel/VBA and SQL. Both technologies are ubiquitous in the financial world, and it is worthwhile spending the time to learn how to use them. The link to the YouTube playlists is

   - http://www.youtube.com/user/WiseOwlTutorials/playlists.

3. In ancient Greece, the agora was a central place of assembly, typically at the center of a city, where citizens would gather to listen to speeches by politicians, cast votes, and trade goods. The modern Greek words for "I purchase", *agorazo* and "I publicly speak," *agorevo* are derivates of agora [68, 113].
4. The financial regulation landscape, at least in the United States, is a fairly complicated one, to say the least. For the stock and options exchanges, US Securities and Exchange Commission (SEC) reigns supreme. The SEC was created in 1934 by the Securities Exchange Act. Its introduction followed one of the most tumultuous periods in US Financial history. According to www.sec.gov, the SEC's mission is to protect investors, maintain fair, orderly, and efficient markets and facilitate capital formation [139]. The futures and some options markets are regulated by the US Commodity Futures Trading Commission (CFTC.) The CFTC

was created in 1974 by the Commodity Futures Trading Commission Act [118]. Its main purpose was to replace the US Department of Agriculture's Commodity Exchange Authority. Even though the SEC and the CFTC are different entities, they do collaborate with each other. More information on the CFTC can be found at www.cftc.gov.

5. Broadly speaking, there are three types of actions a trader can take when interacting with an electronic exchange. The first one is to submit market orders to the exchange. A market order, in effect, tells the exchange that immediacy of execution is the most important thing to the trader. Price does not matter as much. The second action is to submit limit orders to the exchange. Here, price matters. Price is more important than the immediacy of execution. The third action is that of modifying or canceling orders on the exchange. Limit orders effectively populate the orderbook with orders and provide a steady supply of liquidity to those who wish to execute market orders.
6. Trading Technologies, Inc. actually owns patents on the concept of a vertically displayed orderbook, specifically, U.S. Patent Nos. 6,766,304 and 6,772,132, which were issued in 2004.

Chapter 2

1. A programming paradigm is a mathematically derived set of concepts that lends itself to the efficient and elegant solution of a specific problem. For example, problems that can be decomposed into an abstraction of hierarchical interacting objects can best be expressed via the object-oriented programming style [85, 137]. Functional programming is well suited for problems that involve the manipulation, transformation, and processing of data. Functional programming is based on the λ-calculus first formulated by Alonzo Church in 1936. Functions are considered as the basic building blocks in this programming paradigm. The main idea is to have functions that produce results based only on the inputs to the function. These functions are not allowed to mutate data external to the function.
2. A programming language is considered to be Turing Complete if it can be used to simulate any single-taped Turing machine. A single-taped Turing machine is a mechanical construct that can be used to emulate any computing machine utilizing limited memory. A great explanation of such a Turing machine is given here:
   http://en.wikipedia.org/wiki/Turing_machine.
   From a practical aspect, Turing completeness is another way of stating that a programming language is capable of computing almost anything that is computable given enough resources.
3. Compilation of a program is the transformation of human readable code into machine-level code. Compiled code is fast. The drawback with code compilation

is the amount of time it takes to compile the code. An alternative mode of operation is for the computer to interpret the human readable code into something more manageable. This interpreted method allows for faster interaction between the developer and the code. The drawback is that it is slower.

4. Julia is a relatively new programming language that is showing a lot of promise in the scientific programming space. Its powerful LLVM-based just-in-time compiler makes it extremely fast (comparable to C) for most numerical applications. More details can be found here: http://julialang.org/.
5. Reference semantics imply that variables do not store objects by value. Rather, they store the address of an object's location in memory. This allows for more efficient code since larger objects do not have to be copied every time they are passed around
6. The Big-O notation in computer science describes the limiting behavior of functions when their arguments approach infinity. The following entry lists the mathematical properties of O(): http://en.wikipedia.org/wiki/Big_O_notation. From a practical standpoint, O(1) means that the execution time will be the same no matter the size of the input. O(n) implies that the time increases linearly with size n.
7. CRAN task views provide a comprehensive list of packages that are related by topic. The following link mentions the more important graphing related ones: http://cran.r-project.org/web/views/Graphics.html.
8. In R, the order of precedence for argument matching is first by name, then by prefix matching, and finally by position: http://adv-r.had.co.nz/Functions.html.
9. Abstraction is one of those ideas in computer science that is not mathematically well defined. It can refer to the decrease in repetitious code by abstracting away all the details of an implementation and by coding up the general idea as a module that can be reused by other functions.
10. Good R resources can be found here:

    - List of contributed documents on CRAN: http://cran.r-project.org/other-docs.html
    - An Introduction to R: http://cran.r-project.org/doc/manuals/R-intro.pdf
    - The R Inferno: http://www.burns-stat.com/pages/Tutor/R_inferno.pdf
    - Advanced R: http://adv-r.had.co.nz/

11. A good explanation on regular expressions is given here: http://en.wikipedia.org/wiki/Regular_expression. In a nutshell, these constructs are sequences of characters that create a search pattern for textual information. Regular expression programs then translate these patterns into nondeterministic finite automata.
12. The three most popular correlation estimators are Pearson's product-moment coefficient, Spearman's rank correlation coefficient, and the Kendall tau rank correlation coefficient [138].

Chapter 3

1. Yahoo Finance: http://finance.yahoo.com/ provides free daily data on multiple securities and indexes. The data can be downloaded into R via the **quantmod** package [99].
2. The Hadley Wickham book on Advanced R Programming has a nice section on package creation here: http://adv-r.had.co.nz/Philosophy.html.
3. Jeffrey Ryan is a Chicago quant/developer. He has created or co-created some of the most popular finance related R packages. These include: **quantmod**, **xts**, **mmap** and **greeks** [49, 99].
4. The `library()` and `require()` both load packages into memory. The first package will throw an error if the package has not been installed, whereas the second method will only produce a warning. The `require()` call also generates a boolean FALSE if the package does not exist
5. The CRAN task-view on web technologies contains some of the most popular packages that facilitate the retrieval and parsing of data from the Web: http://cran.r-project.org/web/views/WebTechnologies.html.
6. NoSQL is an acronym that stands for Not Only SQL. These databases are not relational in type and typically employ a key value, and a graphical or a document storage paradigm. Some popular NoSQL databases include:

    - MongoDB
    - Redis
    - Neo4j
    - Cassandra

7. The Wikipedia entry on SQL lists it as a "special-purpose programming language designed for managing data held in relational database management systems (RDBMS)." It has been around since the early seventies and is based on Edgar Codd's relational model, which he described in a paper titled "A Relational Model of Data for Large Shared Data Banks." The language itself was developed at IBM by Donald D. Chamberlin and Raymond F. Boyce. They originally called it SEQUEL (Structured English Query Language).
8. The following link shows how to install MariaDB, create a data directory, and then instantiate a database: https://github.com/hadley/dplyr/blob/master/vignettes/notes/mysql-setup.Rmd. According to the mariadb.org website, "MariaDB is an enhanced, drop-in replacement for MySQL."
9. This is another wonderful tabular structure that provides extremely efficient lookups and data manipulation operations. This recent presentation by Matt Dowle explains the workings of this package in greater detail: http://www.rinfinance.com/agenda/2014/workshop/MatthewDowle.pdf.

10. POSIX stands for "Portable Operating System Interface." It is a family of standards specified by IEEE for designing operating systems. POSIXct is a representation that represents the signed number of seconds since the beginning of 1970.
11. Polymorphism is an object-oriented programming concept and refers to the ability of different objects to invoke the same function with different behavior. A cat object, when invoking the speak() function will get a different result than the dog object invoking the same function.
12. Some of the popular date-time packages are:

    - **lubridate**
    - **chron**
    - **timeData**

13. A description of these concepts can be found here: http://vita.had.co.nz/papers/layered-grammar.pdf.

Chapter 4

1. Statistics computed from biased samples can have the effect of consistently over or under estimating the true population parameter [30]. For some examples of this effect, visit: http://en.wikipedia.org/wiki/Sampling_bias.
2. In statistics, the degree of freedom is the number of available values that can change in the computation of a statistic. The classic example is that of the sample variance estimator in which only N-1 values can vary since 1 value is taken up in estimating the mean.
3. Jensen's inequality is a mathematical expression stating that a secant line of a convex function lies above the graph of the convex function [125]. It is often used to prove that the expectation of a convex function is greater than the function of the expectation:

$$E(f(x)) \geq f(E(x)) \tag{4.25}$$

4. The dW_t term is referred to as a Wiener process or a Brownian motion process. This is the main building block of many stochastic time series representations: http://en.wikipedia.org/wiki/Wiener_process.
5. A few popular variance estimators include:

    - Andersen et al. RV estimator,
    - Barndorff-Nielsen and Shephard Cov estimator,
    - Zhang's nonsynchronous volatility measure.

Chapter 5

1. Statistical inference is the theory, methods, and practice of determining population parameter properties based off of sample statistic behavior.
2. When talking about a probability distribution, the zeroth moment is the total probability (i.e., one), the first moment is the mean, the second moment is the variance, and the third moment is the skewness: http://en.wikipedia.org/wiki/Moment_(mathematics).
3. Cointegration is a statistical property of time series. When two or more nonstationary time series can be combined in a linear way such that the resulting time series is stationary, then we say that the those time series are cointegrated. The term was coined in a seminal 1987 paper by Robert Engle. The tests used for detecting cointegration are: the Engle-Granger two-step method, the Johansen test, and the Phillips-Ouliaris cointegration test. http://en.wikipedia.org/wiki/Cointegration.
4. A p-value is the probability of observing an event given that the null hypothesis is true. If our goal is to reject the null-hypothesis, then a small p-value gives us confidence in doing so.
5. A helpful tutorial on robust statistics can be found here: http://www.rci.rutgers.edu/~dtyler/ShortCourse.pdf.
6. The rank correlation is a measure of the linear dependence between the rankings of random variables. Ranking requires some preprocessing of the data in order to properly assign the labels *first*, *second*, *third*, etc. to each observation of the data set. The ranking correlation is also fairly robust to outliers [138].
7. References to the Fisher approximate unbiased estimator and the Olkin-Pratt estimator can be found here: http://www.uv.es/revispsi/articulos1.03/9.ZUMBO.pdf.

Chapter 6

1. The pair trading idea was supposedly pioneered by Gerry Bamberger, and later led by Nunzio Tartaglia's quantitative group at Morgan Stanley in the 1980s.
2. The Fama-French three-factor model is an example of a factor model that attempts to decompose returns into a small caps component, a price-to-book ratio component, and a value vs. growth stock component.
3. A good introduction on cointegration can be found here: http://faculty.washington.edu/ezivot/econ584/notes/cointegration.pdf.
4. Total least squares, or orthogonal regression, is a type of errors-in-variables regression in which errors on both dependent and independent variables are taken into account. See http://en.wikipedia.org/wiki/Total_least_squares for more details.

5. According to Wikipedia, "Principal component analysis is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components [133]." In finance, this technique is used often to model the behavior of the yield curve or the dynamics of the implied volatility surface.
6. There have been a total of three Basel Accords thus far. Basel II regulations were first published in June 2004 [114]. The recent Basel III accords have superseded the previous proceedings. Effectively, these accords establish a voluntary regulatory structure on bank capital adequacy, stress testing, and liquidity risk.

Chapter 7

1. Interactive Brokers has some of the lowest fees in the industry for retail investors. Their homepage can be found here: https://www.interactivebrokers.com.
2. A very good introduction to the workings of **quantstrat** is the presentation by Jan Humme and Brian Peterson: http://www.rinfinance.com/agenda/2013/workshop/Humme+Peterson.pdf.
3. Andreas Clenow is a successful entrepreneur, hedge fund trader, quant analyst, and author of *Following the Trend: Diversified Managed Futures Trading*. A link to the trend following approach referenced in this book is: http://www.followingthetrend.com/2014/03/improving-the-free-trend-following-trading-rules/.
4. Laurence Connors has authored and co-authored: *How Markets Really Work*, *Short Term Trading Strategies That Work*, and the *High Probability ETF Trading*". His opinions have been featured in the Wall Street Journal, Bloomberg, and on Dow Jones.
5. The **PerformanceAnalytics** package is a workhorse of quantitative finance. It is highly recommended that you include the functionality offered by this package into your workflow. A great resource is the following: http://braverock.com/brian/R/PerformanceAnalytics/html/PerformanceAnalytics-package.html.

Chapter 8

1. Microstructure refers to the granular market price dynamics in stocks, bonds, options, futures, and other financial instruments. Various correlation and volatility measures that outperform their classical counterparts can be derived by studying the underlying market microstructure. Such analysis is typically conducted by analyzing tick-data on the entire orderbook.

2. This figure is based on the sample files of SPY intra-day data provided by Tick Data, Inc.
3. One of the main criticisms against high-frequency trading is that the algorithms retract their limit orders at the worst possible time, that is, when most institutional investors and retail clients need the liquidity the most. High-frequency algorithms are able to detect market changes faster than other participants and are therefore faster to move out of the way. The argument is that the liquidity that such firms provide is simply an illusion.
4. Some third-party tick-databases are:

   - **OneTick** from OneMarketData: http://www.onetick.com,
   - **kdb** from Kx Systems: http://kx.com,
   - **SciDB** from: http://www.scidb.org/.

5. The bid-ask bounce is the description given to trades that occur either on the bid or the ask. For some liquid products, multiple trades can occur on the inside market without the price moving at all. This effect gives the impression that there is movement, when in effect, there is none. Calculations for intra-day volatility that rely on such prices will exaggerate the estimate.
6. An autoregressive model attempts to relate the output variable of a time-series as a function of its previous input variables. The general form looks something like this:

$$X_t = \alpha + \sum_{i=1}^{p} \beta_i X_{t-i} + \epsilon_t \tag{8.2}$$

7. A tutorial on using the **highfrequency** package can be found here: http://cran.r-project.org/web/packages/highfrequency/vignettes/highfrequency.pdf.

Chapter 9

1. This is also known as the Black-Scholes-Merton model. Merton made extensive contributions to the mathematical theory of options pricing, and he coined the term: "Black-Scholes options pricing model". http://en.wikipedia.org/wiki/Black-Scholes_model.
2. Dirk Eddelbuettel has created an R wrapper around QuantLib [2]. Examples, installation instructions, and more information can be found here: http://dirk.eddelbuettel.com/code/rquantlib.html.

3. The formulas for the sensitivities of the option price to: underlying, volatility, dividends, interest rates, and strike can be found here: http://en.wikipedia.org/wiki/Black-Scholes_model.

Chapter 10

1. R is jam-packed with optimization routines that implement some of the latest research in the field. The following CRAN task-view lists some of the available packages: http://cran.r-project.org/web/views/Optimization.html.
2. Gradient descent methods utilize the gradient (derivative) of a function and take steps proportional to the negative of this derivative at the points of interest. These techniques are fast but require the computation of a gradient or partial gradient. http://en.wikipedia.org/wiki/Gradient_descent.
3. Arbitrage-free pricing implies that the relationship between instruments of various maturities (for fixed-income instruments) and strikes (for options) needs to be defined in such as way as to not allow the creation of a riskless profit opportunity. In other words, it should be unlikely to create a portfolio with a zero or negative cash outlay that has the potential of yielding a positive return in the future.
4. Overfitting is a nasty condition in which a fitted model produces output that can be completely erroneous or have no predictive power whatsoever. It is very hard to detect overfitting, especially in very high dimensional models. Great care has to be exercised when fitting models to data. Techniques such as regularization and shrinkage are just some of the things that can be applied to limiting this problem.

Chapter 11

1. Luke Tierney gives a nice presentation on the R-engine performance: http://www.rinfinance.com/agenda/2014/talk/LukeTierney.pdf.
2. The **Rtools** package was originally created by Brian Ripley. It is currently being maintained by Duncan Murdoch. The relevant installation files and release notes can be referenced from: http://cran.r-project.org/bin/windows/Rtools/.
3. An explanatory overview of TDD can be found here: http://www.agiledata.org/essays/tdd.html.

References

[1] Clenow A.F. *Following the trend: diversified managed futures trading*. Wiley Trading. Wiley, New York, 2012.

[2] Ferdinando Ametrano and Luigi Ballabio. Quantlib - a free/open-source library for quantitative finance, 2003.

[3] Clenow Andreas. Twelve months momentum trading rules - part 2. http://www.followingthetrend.com/2014/03/improving-the-free-trend-following-trading-rules/, March 2014.

[4] Pfaff B. *Analysis of Integrated and Cointegrated Time Series with R*. Springer, New York, second edition, 2008. ISBN 0-387-27960-1.

[5] Johnson Barry. *Algorithmic Trading and DMA: An introduction to direct access trading strategies*. 4Myeloma Press, February 2010.

[6] Ripley Brian D. *Spatial statistics*. Wiley-Interscience, Hoboken, NJ, 2004.

[7] Bacon Carl R. *Practical portfolio performance: measurement and attribution*. Wiley finance. Wiley, Chichester, England, 2nd edition, 2008.

[8] Bartholomew Daniel. *Getting started with MariaDB*. Packt Publishing, Birmingham, UK, 2013.

[9] Ardia David, Ospina Arango Juan, and Gomez Norman Giraldo. Jump-diffusion calibration using Differential Evolution. *Wilmott Magazine*, 55:76–79, 2011.

[10] Ardia David, Mullen Katharine M., Peterson Brian G., and Joshua Ulrich. *DEoptim: Differential Evolution in R*, 2013. version 2.2-2.

[11] Ardia David, Boudt Kris, Carl Peter, Mullen Katharine M., and Peterson Brian G. Differential Evolution with DEoptim: An application to non-convex portfolio optimization. *The R Journal*, 3(1):27–34, 2011.

[12] Smith David. Fast and easy data munging, with dplyr. http://blog.revolutionanalytics.com/2014/01/fast-and-easy-data-munging-with-dplyr.html.

[13] Belsley David A. and Kontoghiorghes Erricos John. *Handbook of computational econometrics*. Wiley, 2009.

[14] James David A. and DebRoy Saikat. Package rmysql. `http://cran.r-project.org/web/packages/RMySQL/RMySQL.pdf`, July 2014.

[15] Tyler David E. A short course on robust statistics. `http://www.rci.rutgers.edu/~dtyler/ShortCourse.pdf`, July 2014.

[16] Kwiatkowski Denis and Philips Peter C. B. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54:159–178, 1991.

[17] Derryberry DeWayne R. *Basic data analysis for time series with R*. Wiley, 2014.

[18] Eddelbuettel Dirk. *Seamless R and C++ integration with RCPP*. Springer, New York, 2013.

[19] Eddelbuettel Dirk. Rcpp overview. `http://dirk.eddelbuettel.com/code/rcpp.html`, July 2014.

[20] Eddelbuettel Dirk and Francois Romain. Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, 40(8):1–18, 2011.

[21] Eddelbuettel Dirk and Francois Romain. *RInside: C++ classes to embed R in C++ applications*, 2014. R package version 0.2.11.

[22] Chance Don. A brief history of derivatives. http://husky1.stmarys.ca/~gye/derivativeshistory.pdf, July 2014.
[23] Knuth Donald E. Structured programming with go to statements. *ACM Comput. Surv.*, 6(4):261–301, December 1974.
[24] Zimmerman Donald W., Zumbo Bruno D., and Williams Richard H. Bias in estimation and hypothesis testing of correlation. *Psicologica*, 24:133–158, 2003.
[25] Zivot Eric and Wang Jiahui. *Modeling financial time series with S-plus*. Springer, New York, NY, 2nd edition, 2006.
[26] Chan Ernie. How useful is order flow and vpin? http://epchan.blogspot.com/2013/10/how-useful-is-order-flow-and-vpin.html, October 2013.
[27] Black Fischer and Scholes Myron S. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy, University of Chicago Press*, 81(3):637–54, May-June 1973.
[28] James Gareth, Witten Daniela, Hastie Trevor, and Tibshirani Robert. *An Introduction to Statistical Learning: with Applications in R*. Springer Texts in Statistics. Springer New York, 2014.
[29] A.K.I.I. Gary, J.U. Schluetter, and H. Brumfield. Click based trading with intuitive grid display of market depth, August 3, 2004. US Patent 6,772,132.
[30] Henry Gary T. *Practical Sampling*. SAGE Publications, Inc., 1990.
[31] Box George E. P. and Draper Norman Richard. *Empirical model-building and response surfaces*. Wiley, New York, 1987.
[32] Mirai Solutions GmbH. Xlconnect 0.2-0. http://www.r-bloggers.com/xlconnect-0-2-0/, July 2012.
[33] Mirai Solutions GmbH. Package xlconnect. http://cran.r-project.org/web/packages/XLConnect/XLConnect.pdf, July 2014.
[34] CME Group. Leading products q1 2014. http://www.cmegroup.com/education/files/cme-group-leading-products-2014-q1.pdf.
[35] CME Group. Growth of cme globex platform: A retrospective. *CME Group*, page 2, 2012.
[36] CME Group. Daily exchange volume and open interest. *CME Group*, 2014.
[37] Zaner Group. A study in platform volume. *CME Group*, page 2, 2010.
[38] Lebanon Guy. Bias, variance, and mse of estimators. http://www.cc.gatech.edu/~lebanon/notes/estimators1.pdf.
[39] Yollin Guy. R tools for portfolio optimization. http://www.rinfinance.com/RinFinance2009/presentations/yollin_slides.pdf, April 2009.
[40] Wickham Hadley. *ggplot2: elegant graphics for data analysis*. Springer New York, 2009.
[41] Wickham Hadley. testhat: Get started with testing. *The R Journal*, 3/1, June 2011.
[42] Wickham Hadley. Advanced r. http://adv-r.had.co.nz/, July 2014.
[43] Wickham Hadley. Tidy data. *The Journal of Statistical Software*, Submitted.
[44] Wickham Hadley and Francois Romain. Introduction to dplyr. http://cran.rstudio.com/web/packages/dplyr/vignettes/introduction.html.
[45] Wickham Hadley and Francois Romain. *dplyr: dplyr: a grammar of data manipulation*, 2014. R package version 0.2.
[46] Allaire J. Rstudio and inc. (2014), rmarkdown: R markdown document conversion, r package. github.com/rstudio/rmarkdown, 2014.
[47] Maindonald J. H. and Braun John. *Data analysis and graphics using R: an example-based approach*, volume 10 of *Cambridge series in statistical and probabilistic mathematics*. Cambridge University Press, 3rd edition, 2010.

[48] Humme Jan and Peterson Brian G. Using quantstrat to evaluate intraday trading strategies. http://www.rinfinance.com/agenda/2013/workshop/Humme+Peterson.pdf, 2013.
[49] Ryan Jeffrey A. and Ulrich Joshua M. *xts: eXtensible Time Series*, 2013. R package version 0.9-7.
[50] Knight J.L. and Satchell S. *Forecasting Volatility in the Financial Markets*. Butterworth-Heinemann Finance. Butterworth-Heinemann, 2002.
[51] Weisenthal Joe. The story of the first-ever options trade in recorded history. http://www.businessinsider.com/the-story-of-the-first-ever-options-trade-in-recorded-history-2012-3, March 2012.
[52] Fox John and Weisberg Sanford. *An R companion to applied regression*. SAGE Publications, Thousand Oaks, CA, 2nd edition, 2011.
[53] Miyamoto John. Demo 02-1: Using r to think about bayesian inference. https://faculty.washington.edu/jmiyamot/p548/demo02-1.p548.w14.pdf, July 2014.
[54] Mount John. Frequentist inference only seems easy. http://www.win-vector.com/blog/2014/07/frequenstist-inference-only-seems-easy/, July 2014.
[55] O'Connor John J. and Robertson Edmund F. The mactutor history of mathematics archive. http://www-history.mcs.st-and.ac.uk/.
[56] Kruschke John K. *Doing bayesian data analysis: a tutorial with R and BUGS*. Academic Press, Burlington, MA, 2011.
[57] Cornelissen Jonathan, Boudt Kris, and Payseur Scott. *highfrequency: highfrequency*, 2013. R package version 0.2.
[58] Ulrich Joshua. How can i view the source code for a function? http://stackoverflow.com/questions/19226816/how-can-i-view-the-source-code-for-a-function, October 2013.
[59] Ulrich Joshua. *TTR: Technical Trading Rules*, 2013. R package version 0.22-0.
[60] Mullen Katharine, Ardia David, Gil David, Windover Donald, and Cline James. DEoptim: An R package for global optimization by differential evolution. *Journal of Statistical Software*, 40(6):1–26, 2011.
[61] Price Kenneth and Storn Rainer. Differential evolution (de) for continuous function optimization. http://www1.icsi.berkeley.edu/~storn/code.html, July 2014.
[62] Price Kenneth V., Storn Rainer M., and Lampinen Jouni A. *Differential Evolution - A Practical Approach to Global Optimization*. Natural Computing. Springer-Verlag, January 2006. ISBN 540209506.
[63] Andrey N. Kolmogorov. *Foundations of the Theory of Probability*. Chelsea Pub Co., 2 edition, June 1960.
[64] Boudt Kris, Cornelissen Jonathan, and Payseur Scott. Highfrequency: Toolkit for the analysis of highfrequency financial data in r. http://cran.r-project.org/web/packages/highfrequency/vignettes/highfrequency.pdf, July 2014.
[65] Connors L.A. and Alvarez C. *Short Term Trading Strategies that Work: A Quantified Guide to Trading Stocks and ETFs*. Tradingmarkets Publishing Group, 2009.
[66] Torgo Luis. *Data mining with R: learning with case studies*. Chapman and Hall/CRC data mining and knowledge discovery series. Chapman and Hall/CRC, Boca Raton, FL, 2011.
[67] Tierney Luke. A byte code compiler for r. http://homepage.stat.uiowa.edu/~luke/R/compiler/compiler.pdf, March 2012.
[68] Lang Mabel L. *Life, death, and litigation in the Athenian Agora*, volume no. 23. American School of Classical Studies at Athens, Princeton, NJ, 1994.
[69] Blais Marcel and Protter Philip. Signing trades and an evaluation of the leeready algorithm. *Annals of Finance*, 8(1):1–13, 2012.

[70] Bogard Matt. Regression via gradient descent in r. http://econometricsense.blogspot.gr/2011/11/regression-via-gradient-descent-in-r.html, July 2014.
[71] Golder Matt and Golder Sona. Lecture 8: Estimation. https://files.nyu.edu/mrg217/public/lecture8_handouts.pdf, July 2014.
[72] Dowle Matthew. Introduction to the data.table package in r. http://cran.r-project.org/web/packages/data.table/vignettes/datatable-intro.pdf, February 2014.
[73] Kuhn Max and Johnson Kjell. *Applied predictive modeling*. Springer, New York, 2013.
[74] Way Michael J. *Advances in machine learning and data mining for astronomy*. Chapman and Hall/CRC data mining and knowledge discovery series. Chapman and Hall/CRC, 2012.
[75] Naguez Naceur and Prigent Jean-Luc. Kappa performance measures with johnson distributions. *International Journal of Business*, 16(3):210, 2011.
[76] Sommacal Nicola Sturaro. Read excel file from r. http://www.milanor.net/blog/?p=779, July 2013.
[77] Ross Noam. Faster! higher! stronger! - a guide to speeding up r code for busy people. http://www.noamross.net/blog/2013/4/25/faster-talk.html, April 2013.
[78] Matloff Norman S. *The art of R programming: tour of statistical software design*. No Starch Press, San Francisco, 2011.
[79] Valavanes Panos and Delevorrias Angelos. *Great moments in Greek archaeology*. J. Paul Getty Museum, Los Angeles, CA, 2007.
[80] Teetor Paul. *25 recipes for getting started with R*. O'Reilly Media, Beijing, 1st ed edition, 2011.
[81] Teetor Paul. *R cookbook*. O'Reilly, Beijing, 1st edition, 2011.
[82] Cowpertwait Paul S. P. and Metcalfe Andrew V. *Introductory time series with R*. Use R! Springer, Dordrecht, 2009.
[83] Carl Peter and Peterson Brian G. *PerformanceAnalytics: Econometric tools for performance and risk analysis*, 2013.
[84] Diggle Peter and Chetwynd Amanda. *Statistics and scientific method: an introduction for students and researchers*. Oxford University Press, 2011.
[85] Roy Peter V. Programming paradigms for dummies: What every programming paradigms for dummies: What every programmer should know. http://www.info.ucl.ac.be/~pvr/VanRoyChapter.pdf.
[86] Spector Phil. *Data manipulation with R*. Springer, New York, 2008.
[87] Qusma. Equity curve straightness measures. http://qusma.com/2013/09/23/equity-curve-straightness-measures/, September 2013.
[88] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2014.
[89] Subba Rao, T., Subba Rao, S., and Radhakrishna Rao C. *Time series analysis: methods and applications*, volume v. 30 of *Handbook of statistics*. North Holland, Amsterdam, 1st edition, 2012.
[90] Rebonato Riccardo. *Volatility and correlation: the perfect hedger and the fox*. J. Wiley, Chichester, West Sussex, England, 2nd edition, 2004.
[91] Becker Richard A. and Chambers John M. *S: An interactive environment for data analysis and graphics*. Wadsworth Advanced Book Program, Belmont, CA, 1984.
[92] Bookstaber Richard M. *A demon of our own design: markets, hedge funds, and the perils of financial innovation*. J. Wiley, Hoboken, NJ, 2007.
[93] Becker Rick. A brief history of s. *ATT Bell Laboratories*, 11, 1994.

[94] Peng Roger D. Interacting with data using the filehash package. *The Newsletter of the R Project*, 6/4, 2006.
[95] Winston Rory. Newton's method in r. http://www.theresearchkitchen.com/archives/642, July 2014.
[96] Ihaka Ross and Gentleman Robert. R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5(3):299–314, 1996.
[97] Tsay Ruey S. *Analysis of financial time series*. Wiley series in probability and statistics. Wiley, Cambridge, MA, 3rd edition, 2010.
[98] Higginbottom Ryan. Introduction to scientific typesetting. lesson 12: Verbatim text and drawing in latex. http://www2.washjeff.edu/users/rhigginbottom/latex/resources/lecture12.pdf, January 2012.
[99] Jeffrey A. Ryan. *quantmod: Quantitative Financial Modeling Framework*, 2013. R package version 0.4-0.
[100] Meyers Scott. *Effective STL: 50 specific ways to improve your use of the standard template library*. Addison-Wesley, Boston, 2001.
[101] Meyers Scott. *Effective C++: 55 specific ways to improve your programs and designs*. Addison-Wesley, Upper Saddle River, NJ, 3rd edition, 2005.
[102] Ao Sio-Iong, Rieger Burghard B., and Amouzegar Mahyar A. *Advances in machine learning and data analysis*, volume 48 of *Lecture notes in electrical engineering*. Springer Science+Business Media, Dordrecht, 2010.
[103] Prata Stephen. *C++ primer plus*. Addison-Wesley, Upper Saddle River, NJ, 6th ed edition, 2012.
[104] R Core Team. R language definition. http://cran.r-project.org/doc/manuals/r-devel/R-lang.pdf.
[105] Stan Development Team. Rstan: the r interface to stan, version 2.3. http://mc-stan.org/rstan.html, 2014.
[106] Stan Development Team. Stan; a c++ library for probability and sampling, version 2.3. http://mc-stan.org, 2014.
[107] Andersen Torben G. *Handbook of financial time series*. Springer, Berlin, 2009.
[108] Hastie Trevor, Tibshirani Robert, and Friedman Jerome. H. *The elements of statistical learning: data mining, inference, and prediction*. Springer series in statistics. Springer, New York, 2nd edition, 2009.
[109] Durden Tyler. Here is how high frequency trading hurts everyone. http://www.zerohedge.com/news/2014-02-20/here-how-high-frequency-trading- hurts-everyone, February 2014.
[110] Zoonekynd Vincent. Optimization. http://zoonek.free.fr/blosxom/R/2012-06-01_Optimization.html, July 2014.
[111] Venables W. N. and Ripley Brian D. *Modern applied statistics with S*. Springer, New York, 4th edition, 2002.
[112] Brainerd Walter S. and Landweber Lawrence H. *Theory of computation*. Wiley, New York, 1974.
[113] Wikipedia. Agora—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Agora, July 2014.
[114] Wikipedia. Basel committee on banking supervision—Wikipedia, the free encyclopedia. wikipedia.org/wiki/Basel_Committee_on_Banking_Supervision, July 2014.
[115] Wikipedia. Bayes theorem—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Bayes'_theorem, July 2014.
[116] Wikipedia. Black scholes model—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Black-Scholes_model, July 2014.

[117] Wikipedia. Central limit theorem—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Central_limit_theorem, July 2014.
[118] Wikipedia. Cftc—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Commodity_Futures_Trading_Commission, August 2014.
[119] Wikipedia. Checking whether a coin is fair—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Checking_whether_a_coin_is_fair, July 2014.
[120] Wikipedia. Chicago board of trade—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Chicago_Board_of_Trade, July 2014.
[121] Wikipedia. Fama-french three-factor model—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Fama%E2%80%93French_three-factor_model, July 2014.
[122] Wikipedia. Frequentist probability—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Frequentist_probability, July 2014.
[123] Wikipedia. Gaussmarkov theorem—wikipedia, the free encyclopedia, 2014. [Online; accessed 13-September-2014].
[124] Wikipedia. Hello world programs—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/List_of_Hello_world_program_examples, July 2014.
[125] Wikipedia. Jensen's inequality—wikipedia, the free encyclopedia, 2014. [Online; accessed 13-September-2014].
[126] Wikipedia. Law of large numbers—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Law_of_large_numbers, July 2014.
[127] Wikipedia. Markdown—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Markdown, July 2014.
[128] Wikipedia. Omega ratio—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Omega_ratio, July 2014.
[129] Wikipedia. Operational risk—Wikipedia, the free encyclopedia. wikipedia.org/wiki/Operational_risk, July 2014.
[130] Wikipedia. Option (finance)—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Option_(finance), July 2014.
[131] Wikipedia. Pairs trade—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Pairs_trade, July 2014.
[132] Wikipedia. Posix—wikipedia, the free encyclopedia, 2014. [Online; accessed 13-September-2014].
[133] Wikipedia. Principal component analysis—wikipedia, the free encyclopedia, 2014. [Online; accessed 14-September-2014].
[134] Wikipedia. Probability axioms—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Probability_axioms, July 2014.
[135] Wikipedia. Probability interpretations—wikipedia, the free encyclopedia, 2014. [Online; accessed 13-September-2014].
[136] Wikipedia. Program / optimization—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Program/optimization, July 2014.
[137] Wikipedia. Programming paradigm—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Programming_paradigm, July 2014.
[138] Wikipedia. Rank correlation—Wikipedia, the free encyclopedia. en.m.wikipedia.org/wiki/Rank_correlation, July 2014.
[139] Wikipedia. Sec—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Securities_and_Exchange_Commission, August 2014.
[140] Wikipedia. Sharpe ratio—Wikipedia, the free encyclopedia. wikipedia.org/wiki/Sharpe_ratio, July 2014.

[141] Wikipedia. Test driven development—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Test-driven_development, July 2014.
[142] Wikipedia. Ulcer index—Wikipedia, the free encyclopedia. wikipedia.org/wiki/Ulcer_index, July 2014.
[143] Wikipedia. Unbiased estimation of standard deviation—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Unbiased_estimation_of_standard_deviation, July 2014.
[144] Wikipedia. Vpin—Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/VPIN, July 2014.
[145] Xie Yihui. *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2013. ISBN 978-1482203530.
[146] Xie Yihui. *knitr: A general-purpose package for dynamic report generation in R*, 2014. R package version 1.6.
[147] Xie Yihui. knitr reference card. http://cran.at.r-project.org/web/packages/knitr/vignettes/knitr-refcard.pdf, May 2014.
[148] Ross Zev. Four reasons why you should check out the r package dplyr. http://zevross.com/blog/2014/03/26/four-reasons-why-you-should-check-out-the-r-package-dplyr-3/.

Index