Financial Analytics with R Errata Sheet

Mark J. Bennett and Dirk L. Hugen

July 11, 2019

Welcome to Financial Analytics with R: Building a Laptop Laboratory for Data Science
by Mark J. Bennett and Dirk L. Hugen from Cambridge University Press!

It is best to send inquiries directly to us at mjbenntt@uiowa.edu or dirklhugen@gmail.com
rather than Cambridge University Press.

Organization:

The code for each Chapter is in a file called code.R.
Associated datasets are also contained in each directory.

Ticker symbols contained in literal constants in the book do become obsolete occasionally.
The most common event that causes this is when these publicly traded companies are acquired.
The tseries package routine get.hist.quote() will no longer quote prices
after this happens even though the prices existed up until the acquisition.

In order to approximate the book results, it is best to replace the ticker that
was absorbed (and no longer quoted) with the surviving ticker. During 2016-2018 these
ticker changes occurred and the code.R files have been updated accordingly.

http://money.cnn.com/news/newsfeeds/articles/globenewswire/6172457.htm is about PCP -> BRK-B.
Note: for Yahoo's get.hist.quote() we used BRK-B but for Quandl market prices we now use BRK-B.

http://www.kindermorgan.com/content/docs/KMI_Summary_Transaction_Information.pdf
is about KMP -> KMI.

http://investors.danaher.com/2015-08-31-Danaher-Completes-Acquisition-Of-Pall-Corporation
is about PLL -> DHR.

https://www.cnbc.com/2017/06/13/verizon-completes-yahoo-acquisition-marissa-mayer-resigns.html
is about YHOO -> VZ

https://seekingalpha.com/news/3332795-priceline-renames-booking-holdings-new-ticker-coming-soon
is about PCLN -> BKNG from Feb 27 2018 forward

Errata:

```
errata:
p. 20  'factor' should be 'fac' three places in the second boxes block of code.
p. 24  A given B should be B given A
p. 32  middle page Cor(X,Y) definition: denominator should be sqrt(Var(X))*sqrt(Var(Y))
p. 37  Formula 3.35 remove the capital Sigma left of =. Formula 3.36 should be
       Cor(R_j,R_k) = Cov(R_j,R_k)/(sigma_j sigma_k)
p. 62  Top line of code should be wInc <- w/sum(w)*amtToRealloc.
p. 64  Former errata sheet(s) had get.hist.quote() disabled, however, it is again operational.
p. 64  Upper code block replace 'YHOO" with 'AAPL' since YHOO no longer independent.
p. 70  'len <- 1006' becomes 'len <- length(get.hist.quote("NVDA",start="2010-02-17",
```

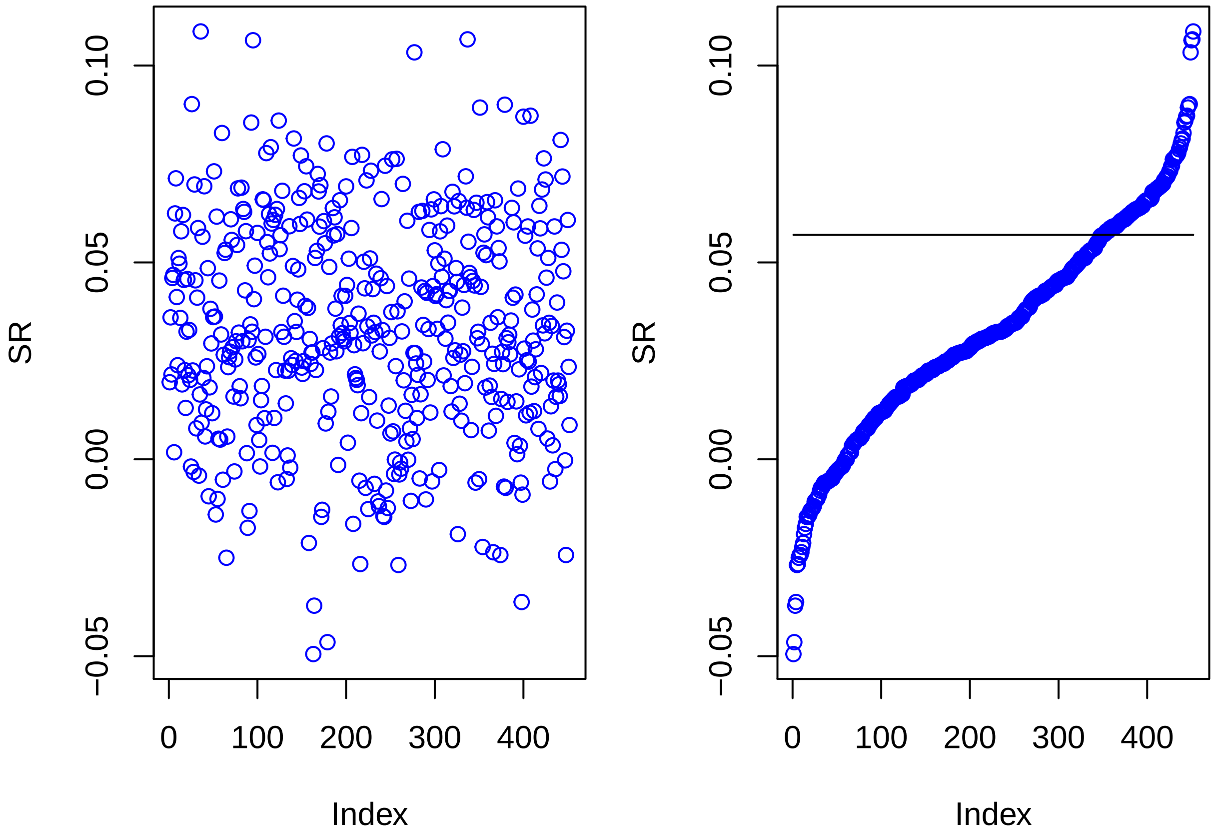
```
                                                end="2014-02-14",quote="AdjClose"))'
p. 70  Middle of page: start='2010-02-18' becomes start='2010-02-17' two places.
p. 75  Repace code block final 2 lines with:
       #unit test:
       library(tseries)
       get.hist.quote("BRK-B",start="2016-05-01",end="2017-05-01",quote="Adj")
       pdf <- getHistPrices(c('BKNG','MCD'),c(.5,.5),251,
              start="2013-02-15",end="2014-02-14")
       pdf
p. 90  Vector c(-0.0339, -0.0351, 0.0000, 0.0690, 0.0645 -0.0317) with missing comma should be
       c(-0.0339, -0.0351, 0.0000, 0.0690, 0.0645, -0.0317)
p. 90  Exercise 5.2 sigmaZ should be .0007
p. 92 to 97: * these code blocks no longer run because getSymbol() from library(quantmod)
       because the facility, ichart.finance.yahoo.com, no longer operates. Meanwhile
       the output of these code blocks are still seen in the book.
       No other code in Chapter 6 is affected from the remainder of Section 6.1 onward.
p. 98  Middle page equation Y_t-1 + Y_t-1 should be Y_t-1 + Y_t-2 and Y_t on left side
       should have a hat on it.
p. 101 Line 3 the number 15 should be 25.
p. 110 Middle page 'then we say that Y_t is an' should be
       'then we say that W_t is an'.
p. 110 Middle page equation Y_t-1 + Y_t-1 should be Y_t-1 + Y_t-2 and W_t on left
       side should have a hat on it.
p. 110 Middle page equations Y_t should have a hat on it two places.
p. 111 Line 13 up from bottom of page replace p = 1 with q = 1 for MA component
p. 111 End of 1st paragraph: remove "at the alpha = 1 percent level but not at
       the alpha = 5 percent level."
p. 127 Formula 6.14 should be 0.00005246/(1-.1397-.7698) = 0.0005796685
p. 128 Exercise 6.1 (e) should be: pacf() instead of acf()
p. 128 Exercise 6.1 (g) should be: Examine the model residuals with a histogram
       hist(residuals(m.boardings),breaks=20) and test
       for normalityshapiro.test(residuals(m.boardings))
p. 129 Exercise 6.3 (g) should be Plot the fitted model conditional
       variance: plot((fitted(m2)[,1])^2,type='l'
p. 130 line 8 'pass' should be 'passer'
p. 133 after l <- stockdata$info[1:D,1] add:
       for(i in 1:D)  #Need to splitAdjust():
         p[,i] <- splitAdjust(p[,i],l[i])
       This change is necessary because findR() only side affects a matrix named
       prices via the '<<-' global assignment operator and so does not side affect p.
p. 133 replace .035 with .057 in code block 6th line from the bottom
p. 134 replace Figure 7.2 with the following:
```

The version of pruneBySharpe() below is invoked with a p (prices) matrix that has been split-adjusted so the threshold changes from .035 to .057 to maintain our top 102 stocks in terms of Sharpe Ratio. The revised pruneBySharpe() also returns all three major result objects in a list().

```r
pruneBySharpe <- function(prices,lab,meanv,sdevv,threshSR,mufree=0) {
  par(mar=c(4,4,1,1))
  par(mfrow=c(1,2))
  indepSharpes <- (meanv-mufree)/sdevv
  len = length(indepSharpes)
  plot(indepSharpes,ylab="SR",col=4)
  plot(sort(indepSharpes),ylab="SR",col=4)
  lines(1:len,rep(threshSR,len))
```

```r
  indHighSharpes <- (indepSharpes > threshSR)
  #clean up NAs
  for(d in 1:length(indHighSharpes)) #clean up NAs
    if(is.na(indHighSharpes[d]))
      indHighSharpes[d] <- FALSE
  len = dim(prices)[1]
  wid = dim(prices)[2]
  smallerSz = sum(indHighSharpes)
  newPrices <- matrix(rep(0,len*smallerSz),
                      nrow=len,ncol=smallerSz)
  newLab    <- vector(length=smallerSz)
  e <- 1
  for(d in 1:wid) {
    if(indHighSharpes[d]) {
      if(F) print(paste("e",e))
      newPrices[,e] <- prices[,d]
      newLab[e] <- lab[d]
      e <- e + 1
    }
  }
  print("completed Sharpe pruning")
  list(newPrices,newLab,indepSharpes)
}
#unit test:
suppressPackageStartupMessages(library(huge))
data(stockdata)
D <- length(stockdata$data[1,])
p <- stockdata$data[,1:D]
l <- stockdata$info[1:D,1]
###added 2017-09-16:
for(i in 1:D)  #Need to splitAdjust():
  p[,i] <- invisible(splitAdjust(p[,i],l[i]))
###
r <- findR(p)
res <- findCovMat(r)
meanv    <- res[[1]]
cov_mat  <- res[[2]]
diag_cov_mat <- res[[3]]
sdevv <- res[[4]]

res <- pruneBySharpe(p,l,meanv,sdevv,.057)
```



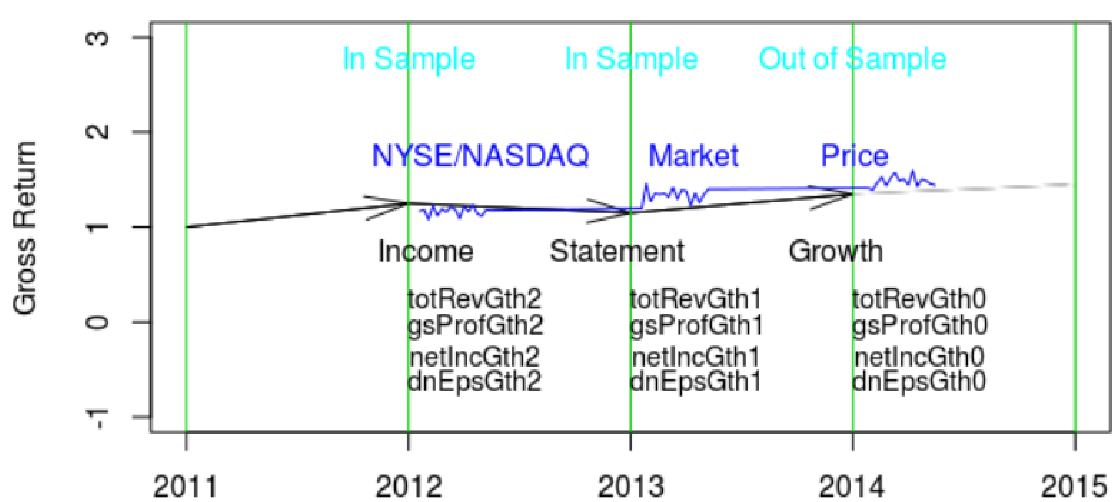

```
## [1] "completed Sharpe pruning"
```

```
p <- res[[1]]
l <- res[[2]]
D <- length(l)
indepSharpes <- res[[3]]
print(paste('D =',D))
```

```
## [1] "D = 102"
```

```
p. 134 isnaCheckCovMat: cor replaced by cov 6 places.
p. 137 5 lines at top of page obsolete due to package not being supported.
p. 150 'WIT' for Wipro Limited cannot be quoted so Figure 7.7 will vary for reader.
p. 151 Top line of code 'PLL' should be 'DHR' due to merger.
p. 153 Middle page, middle code block, due to mergers, replace topSRlab with
       topSRlab[9]<-"KMI"; topSRlab[15]<-"BRK_B"; topSRlab
p. 158 Paragraph 3 5th formula demonimator sigma_e^2 should be (sigma_e)^2.
p. 158 Paragraph 3 Cov(X,Y) sigma_XY = sigma_X sigmaY rho should be
       Cov(X,Y) = sigma_XY = sigma_X sigmaY rho .
p. 159 Top of page, 'vectors to record 100' shoud be 'vectors to record 101'
p. 161 'b, a vector of solution variables,' should be b0
p. 162 First full paragraph 'Formula 6.1' should be 'Formula 8.1'
p. 171 last code block insert after daysPerYr=252; mufree = 0
       for(i in 1:D)  #Need to splitAdjust():
         prices[,i] <- splitAdjust(prices[,i],l[i])
       change last line to R <- findR(prices)
p. 172 remove line: isSplitAdjusted=isSplitAdjusted) #side affects prices
p. 180 to 181: large code block is largely repeat from p. 179. Remove.
p. 182 weightportOOS() has been updated in new FinAnalytics.zip
```

```
      file on book web site. '{' needed on code block line 17 and '} else {'
      replaces 'else' on line 20 and need to insert '}' after line 22.
p. 182 Bottom of code block replace 252 with 251 three places.
p. 182 Bottom of code block replace 'PCLN' with 'BKNG' two places.
p. 185 middle page .72 should be .0456
p. 187 len <- 754 should be automatically set with:
      len <- length(get.hist.quote("QQQ",quote="Adj",start=start,end=end))
p. 189 pruneBySharpe() last param should be .0827
p. 191 Top code block, top lines of code, add 'prices=prices,' at end `after 'w,'
      in order to reuse cacheXXX.csv files for ETF tickers XXX.
p. 191 Top code block, last line writeWeights() should be writeWeights("ETF",lab,w)
p. 194 Bottom of bottom shaded block: .1893 replaces .2014
p. 197 2nd paragraph from bottom m observations should be n observations
p. 200 total return due to 'Formula 7.3' should be 'Formula 9.3'
p. 212 Figure 9.5 caption: 'based upon prices' insert 'from last 252 days (one year)'
p. 219 Figure 9.7 caption: add to end 'Prices are from 1512 days (six years)
      from 2008 to 2014'
p. 225 Formula 9.16 argmin should be argmax two places.
p. 238 first line of code: mode=2 should be mode=3.
p. 255 third line of code should diff(log(ec)) > 0.
p. 271 Figure 11.6 has duplicated arcs from state -1 to states -.75 and -.50
      that should be combined into single arc named 'exit position expected gain'
      and 'exit position expected loss'.
p. 271 Figure 11.6 has duplicated arcs from state +1 to states +.75 and +.50
      that should be combined into single arc named 'exit position expected gain'
      and 'exit position expected loss'.
p. 285 In the 5th line of the first boxed codes on p.285, ???diskdb???
      should be ???con??? (it is ???con??? in the code.R file). 'command' should be 'query'.
n the 3rd line of the last boxed codes on p.288, ???ylab??? should be ???main???.
p. 288 In the 4th line of the first boxed codes, delete '&'.
p. 291 In the second boxed codes, the whole set of code should be preceded by 'query <-'
p. 293 to 294: * Last line of code in bottom code block to 'plot(UFS)' is
      same situation as p. 92 to 95 see above.
p. 294 In the last boxed codes, 'diskdb' should be 'con'.
p. 299 Figure 13.2 caption: replace (2012,2013,2014,2015) with (2011,2012,2013,2014) and
      add to the end 'over an 82 day period (January to May) following December 31st'.
p. 299 Replace Figure 13.2 with:
```



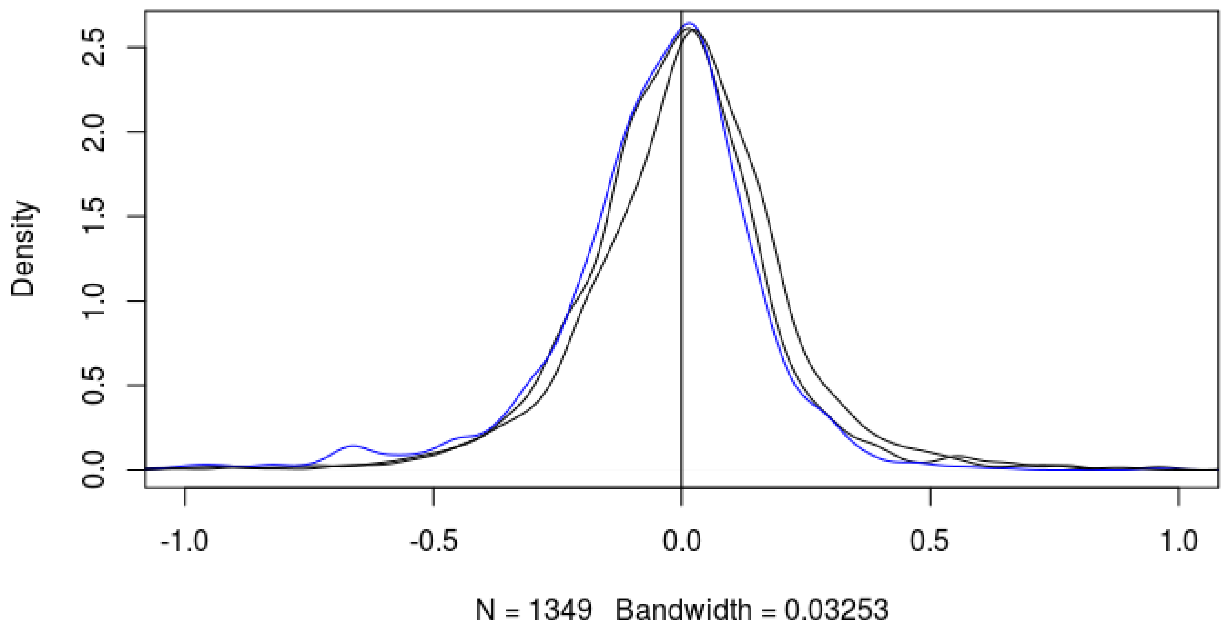
p. 300 middle of page code block replace first two lines with:

```
idxs = which(ISptrnDF$symbol=="UNP")
ISptrnDF[idxs,]
```

p. 302 findCachedPrices() and find3YrsBackCachedPrices() are replaced with acquireCoalescedPrices() and findAllPrices() and findIndexes() per new code.R file.

p. 303 findSPprices() is replaced with calls to getHistPrices() per new code.R file.

p. 306 Figure 13.3 replace with this generated by new code.R file:



p. 306 augPtrn() function is replaced by augPtrnByYrsBack() function as seen in code.R file.

p. 307 bottom of page code block replace first two lines with:

```
idxs = which(ISptrnDF$symbol=="NSC")
ISptrnDFcln[idxs,]
```

p. 312 Figure 13.6 no longer applies.

p. 314 Section 13.6 code block is replaced by pertinent code from code.R file.

p. 316 Table in middle of page: replace Winning Percentages with those from five runs of runClassifier() from latest copy of code.R for Chapter 13, focusing on precision because precision is the winning percentage of long trades.

p. 318 'makes sense' should be 'make sense'.

p. 323 Paragraph 2 replace 'annualized rate of return' with 'annual interest rate'.

p. 324 Remove last two sentences of first paragraph.

p. 324 Add to bottom of page: 'where b = r_A, the annual interest rate'.

p. 329 1+r+d_n just after Formula 14.13 should be 1+r-d_n to match Formula 14.7.

p. 331 middle of page, the reference to (Black and Myron, 1973) should be (Black and Scholes, 1973)

p. 333 middle of page, the last term on line 5 in the expression for dY should be (sigma)^2/2 multiplied by dt instead of t.

p. 334 and 335: replace 'muD,sigmaD,rD' in sim() argument list with 'rA,sigmaA', two occurrences.

p. 335 In the 4th line of code from the bottom of the second boxed block of code remove 'z'.

p. 338 Formula 15.9 two occurrences of 't' should be 'T' like in Formula 15.11 and 15.12.

p. 372 middle of page, 'Myron, S.' should be 'Scholes, M.'.