High-Frequency Data Modelling with R software

Piotr Wójcik

University of Warsaw, Faculty of Economic Sciences

QFRG-Eurex Conference
2012-12-05

What is R?

- freely available language and environment for statistical computing and graphics (very similar to S in S-Plus),
- a well developed intuitive, flexible and convenient programming language,
- initially written by Ross Ihaka and Robert Gentleman at the Department of Statistics of the University of Auckland in New Zealand,
- current R is the result of a collaborative effort with contributions from all over the world,

R is hot!

A good analyst/trader has to be a good programmer !!!

The choice includes a number of alternatives (used in financial modelling):

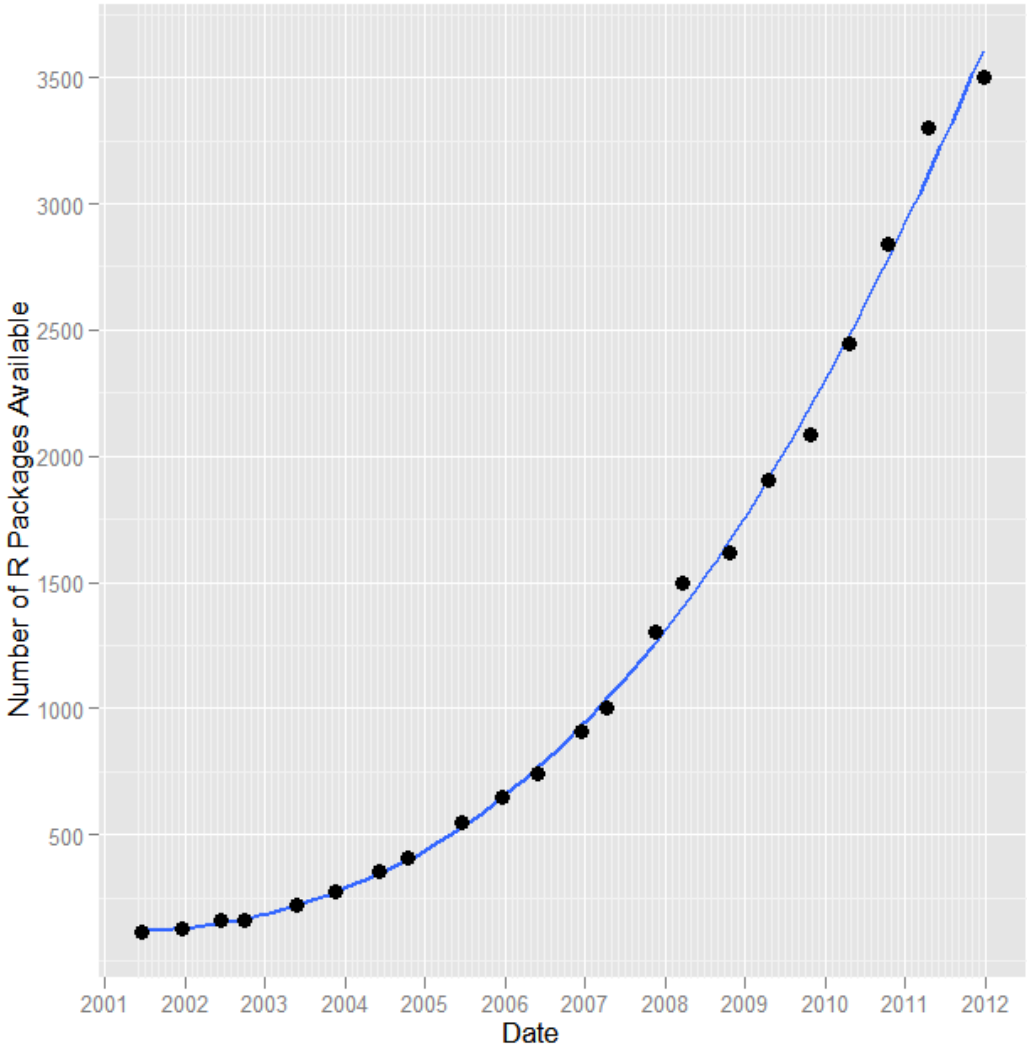
- R,
- SAS,
- Matlab,
- Mathematica,
- S-Plus,
- C++,
- Python,
- others.

Why use R in high-frequency data modelling? (1)

- open-source, free,

Why use R in high-frequency data modelling? (2)

- good to know it, universal, including many additional packages not only for statistics or quantitative finance,



Why use R in high-frequency data modelling? (3)

- at least 10 years ahead of commercial packages,

„A key benefit of R is that it provides nearinstant availability of new and experimental methods created by its user base – without waiting for the development/release cycle of commercial software. SAS recognizes the value of R to our customer base. . . ”

– Michael Gilliland, Product Marketing Manager SAS Institute, Inc.

„There are very few things that SAS or SPSS will do that R cannot, while R can do a wide range of things that the others cannot.”

– Robert A. Muenchen, author, R for SAS and SPSS Users

Why use R in high-frequency data modelling? (4)

- (world-) wide R community – easy to learn get help online.

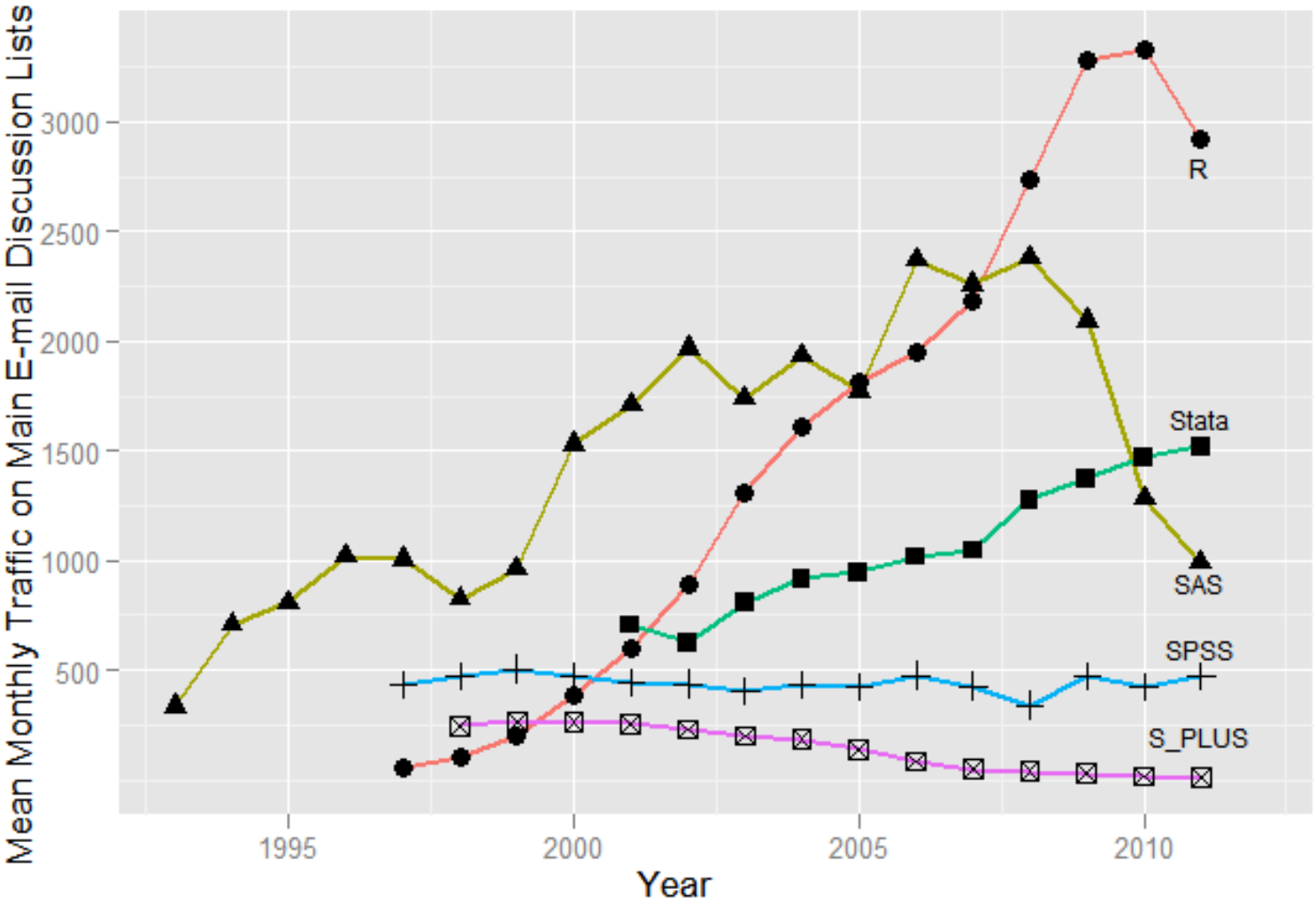
software	# of blogs*
R	365
SAS	40
Stata	8
Others	0-3

* in March 2012

Source: `http://r4stats.com/articles/popularity/`

R is becoming the *lingua franca* of analytic statistics.

Similar relations hold for discussion lists subscribers, stackoverflow discussions. . .



R strengths

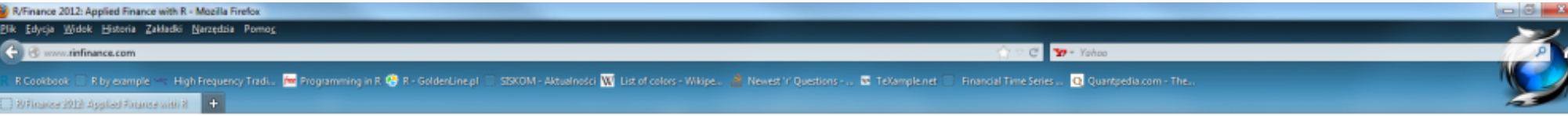
- data manipulation and management,
- statistics,
- graphics,
- programming language,
- user friendly interface – RStudio,
- thousands of available packages with specialized functions,
- active user community (forums, discussion groups),

R weaknesses

- standard interface not very user friendly,
- lack of commercial support,
- usually several possible ways to do the same,
- slower than programming languages (like C++, Perl, Java).

Where to find R?

- `http://cran.r-project.org/`
- `http://www.r-project.org/`
- `http://www.rstudio.com/`



R/Finance 2012: Applied Finance with R

May 11 & 12, Chicago, IL, USA

home agenda register travel committee 2009 Conference 2010 Conference 2011 Conference

UIC ICFD

REVOLUTION ANALYTICS

MS-Computational Finance
UNIVERSITY of WASHINGTON

Google lemnica OpenGamma ONE TICK RStudio TICK DATA

> home(2012)

The fourth annual R/Finance conference for applied finance using **R**, the premier free software system for statistical computation and graphics, will be held this spring in Chicago, IL, USA on **Friday May 11 and Saturday May 12, 2012**. The two-day conference will cover portfolio management, time series analysis, advanced risk tools, high-performance computing, econometrics and more. All will be discussed within the context of using **R** as a primary tool for financial risk management, analysis and trading.

We invite you to submit complete papers or one-page abstracts (in txt or pdf format) for consideration. Academic and practitioner proposals related to R are encouraged. We welcome submissions for full talks, abbreviated "lightning talks", and for a limited number of (longer) pre-conference seminar sessions.

Presenters are strongly encouraged to provide working R code to accompany the presentation/paper. Data sets should also be made public for the purposes of reproducibility (though we realize this may be limited due to contracts with data vendors). Preference may be given to presenters who have released R packages.

Travel and accommodation grants may be available for selected presenters at the discretion of the committee. In addition, the conference will award prizes for best papers. To be eligible for a best paper award, a submission must be a full paper. Extended abstracts, even if a full paper by conference time, are not eligible for a best paper award.

The submission deadline was January 31, 2012.

The **draft agenda** is now available, as is **information about registration**.

The 2012 conference will build upon the success of the three previous events. We expect invited keynote lectures by:

- Blair Hull,
- Paul Gilbert,
- Rob McCulloch and
- Simon Urbanek.

plus presentations of contributed papers, short "lightning-style" presentations, and optional pre-conference tutorials. It also presents a unique opportunity to meet fellow R users and developers, and a thus provides a chance to discuss the future of R in Finance.

The inaugural 2009 conference featured keynotes by Patrick Burns, Robert Grossman, David Kane, Roger Koenker, David Ruppert, Diethelm Wuertz, and Eric Zivot, as well as a number of excellent presentations. The 2010 conference followed up with keynotes by Bernhard Pfaff, Ralph Vince, Marc Wildi, and Achim Zeileis. Last year's conference featured keynotes by Meb Faber, Stefano Iacus, John Bollinger and Louis Kates.

Complete programs of the previous conferences, along with downloadable presentation slides, are available via the links above and below.

The R/Finance 2012 conference is again organized by a local group of R package authors and community contributors, and hosted by the International Center for Futures and Derivatives [ICFD] at the University of Illinois at Chicago. Limited sponsorship opportunities are available.

Finance | R-statistics blog - Mozilla Firefox
Plik Edycja Widok Historia Zakładki Narzędzia Pomoc
www.r-statistics.com/tag/finance/
r graph
Revolutions: Milktrader: Quantitative ...
FOSS Trading
Finance | R-statistics blog
www.r-statistics.com
R-statistics blog
Writing about statistics with R, and open source stuff (software, data, community)
Search
About
All articles
Contact me
Posts Tagged 'Finance'
Tips for the R beginner (a 5 page overview)
Posted in R, statistics on August 23rd, 2010 by Tal Galili – 4 Comments
In this post I publish a PDF document titled "A collection of tips for R in Finance". It is a basic 5 page introduction to R in finances by Arnaud Amsellem (linked in profile).
The article offers tips related to the following points:
Code Editor
Organizing R code
Update packages
Getting external data into R
Communicating with external applications
Optimizing R code
This article is well articulated, and offers a perspective of someone who is experienced in the field and touches points that I can imagine beginners might otherwise overlook. I hope publishing it here will be of use to some readers out there.
Update: as some readers have noted to me (by e-mail, and by commenting), this document touches very lightly on the topic of "finances" in R. I therefore decided to update the title from "R in finance – some tips for beginners", to it's current form.
Lastly: if you (a reader of this blog) feel you have an article ("post") to contribute, but don't
Subscribe: Entries | Comments
e-mail subscription
Enter your email address:
Subscribe
173 readers
BY FEEDBURNER
Categories
R
R and the web
R community
R links
statistics
visualization
wordpress
WordPress and Statistics
Recent Posts
Comparison of ave, ddply and data.table
Calling R lovers and bloggers – to work together on "The R Programming wikibook"
Engineering Data Analysis (with R and ggplot2) – a Google Tech Talk given by Hadley Wickham

Tutorial: How to make NYT-style bar charts with R | (R news & tutorials) - Mozilla Firefox

Plik Edycja Widok Historia Zakładki Narzędzia Pomoc

Tutorial: How to make NYT-style bar ch...

www.r-bloggers.com/tutorial-how-to-make-nyt-style-bar-charts-with-r/

forumSoft4U.org • Do... Często odwiedzane LaTeX/Print version - ... Getting to Grips with L... http://www-stary.gust... R Links Kalendarz rozwoju dzie... SISKOM - Aktualności Nieruchomości mieszk... LaTeX Search - Mathe...

http://www.r-bloggers.com

R-bloggers

R news and tutorials contributed by (425) R bloggers

Home | About | add your blog! | Contact us | RSS

WELCOME!

Here you will find daily **news and tutorials about R**, contributed by over 400 bloggers. You can subscribe for e-mail updates:

Your e-mail here

Subscribe

10818 readers

BY FEEDBURNER

And get updates to your Facebook: **If you are an R blogger yourself** you are invited to add your own R content feed to this site (**Non-English** R bloggers should add themselves-here)

TOPICS OF INTEREST

ggplot2 rstats graphics r-project programming R Language visualization Books Data R-english events tutorial Big Data LaTex finance packages ggplot applications Software

Tutorial: How to make NYT-style bar charts with R

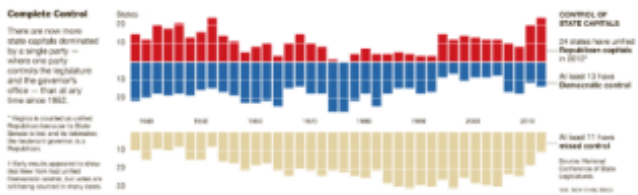
December 3, 2012

By David Smith

Lubię to! Sign Up, aby zobaczyć co lubią Twoi znajomi.

(This article was first published on **Revolutions**, and kindly contributed to R-bloggers)

New York Times columnist Charles Blow needed a chart to accompany his op-ed piece Lincoln, Liberty and Two Americas (about one-party control in state legislatures). So he turned to resident graphic editor Kevin Quealy, who found the source data and used R to create the chart below:



If you'd like to create similar charts yourself, Kevin provides a useful tutorial on his behind-the-scenes blog, *chartsnthings*. It's just a few lines of base graphics in R, beginning with line plots and then moving on to bar plots. The abline function is used to create the gridlines. You can also use the same technique to create "waffle charts" (the display technique often used by the NYT, where small blocks represent data elements). You can find the full tutorial at the link below.

chartsnthings: R tutorial: Simple charts

To **leave a comment** for the author, please follow the link and comment on his blog: **Revolutions**.

Search & Hit Enter

TOP 7 ARTICLES OF THE WEEK

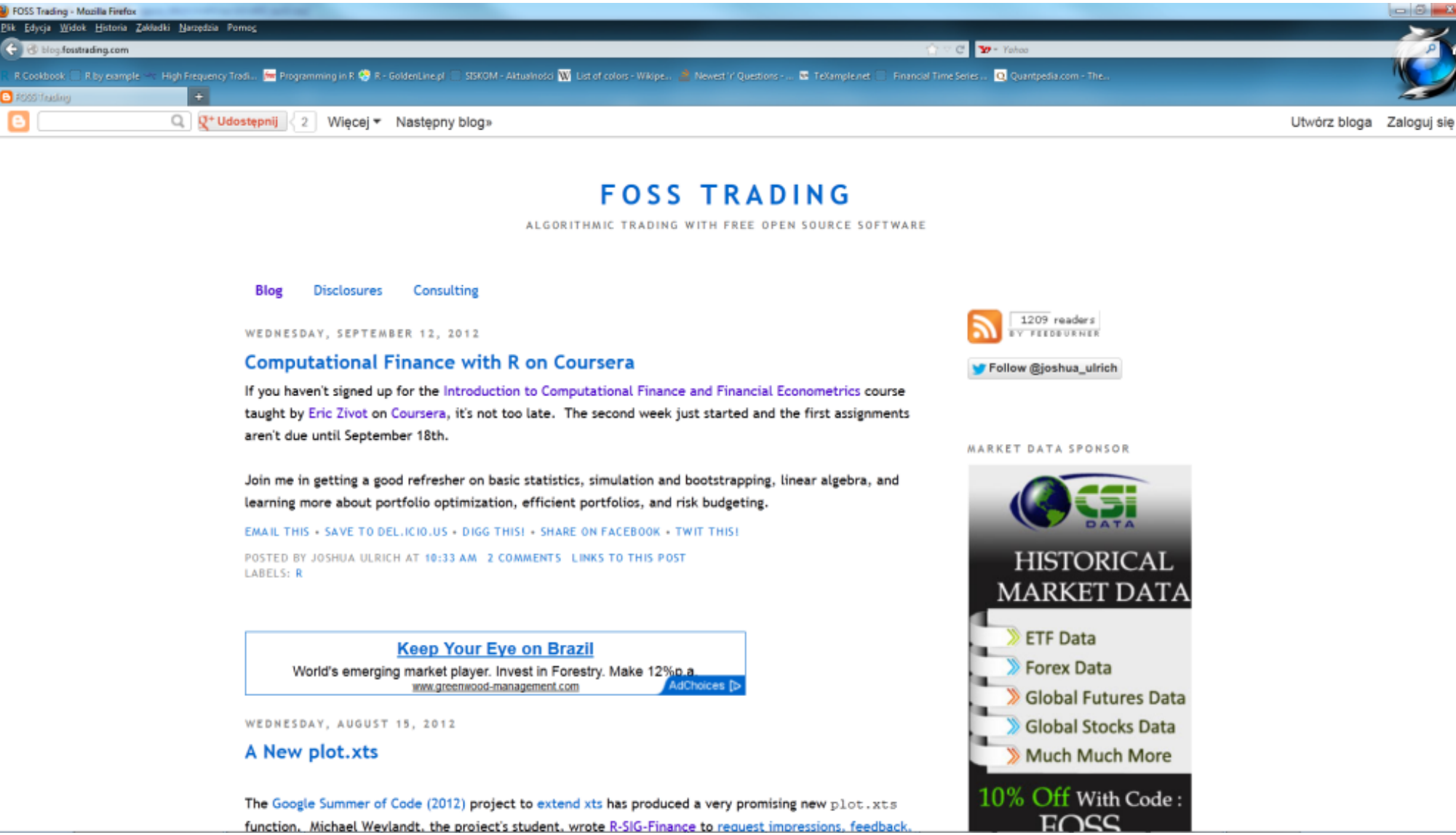
1. Data.table rocks! Data manipulation the fast way in R
2. How to: network animation with R and the iGraph package & Meaning in data viz
3. Hadley's guide to high-performance R with Rcpp
4. Images as Voronoi tesselations
5. Select operations on R data frames
6. RStudio and Rcpp
7. Spatial Data visualization with R

Search & Hit Enter

SPONSORS

MANGOSOLUTIONS
data analysis that delivers

blog.fosstrading.com



Milk Trader: Chop, Slice and Dice Your Returns in R - Mozilla Firefox

Plik Edycja Widok Historia Zakładki Narzędzia Pomoc

www.milktrader.net/2011/04/chop-slice-and-dice-your-returns-in-r.html

Revolutions: Milktrader: Quantitative ... | Milk Trader: Chop, Slice and Dice You... | Quick-R: Home Page

Udostępnij Zgłoś nadużycie Następny blog» Utwórz bloga Zaloguj się

$$y = \varphi\left(\sum_{i=1}^{n} w_i x_i + b\right)$$

$$P(t) = \frac{1}{1+e^{-t}}.$$

www.milktrader.net

$$\text{Geo. Mean} = \left(\prod_{i=1}^{N}(1 + f \cdot (-\text{trade}_i/\text{biggest loss}))\right)\text{^}(1/N)$$

Milktrader

Iterating Until Convergence

SUNDAY, APRIL 24, 2011

Chop, Slice and Dice Your Returns in R

I have a knife rack on my kitchen wall with all my kitchen knives easily identifiable and accessible. I also have small scars on my hand where each knife can claim to have left a mark. It's not the knife's fault, of course. They hardly like being suddenly dropped and cursed at. They have no control over who gets picked on a given day. The choice is really mine.

What's good in the kitchen is good at the trade desk. We like choice as traders. We choose markets, trading styles and excuses for our sub-optimal performances. On those occasions when we need to crunch some numbers, we also like some choice. More than any other curve-fitting software, R is best suited for trading precisely because of its "quote" - diversity.

In fact, I'm sure this is on purpose. The Unix geniuses did a lot of thinking about

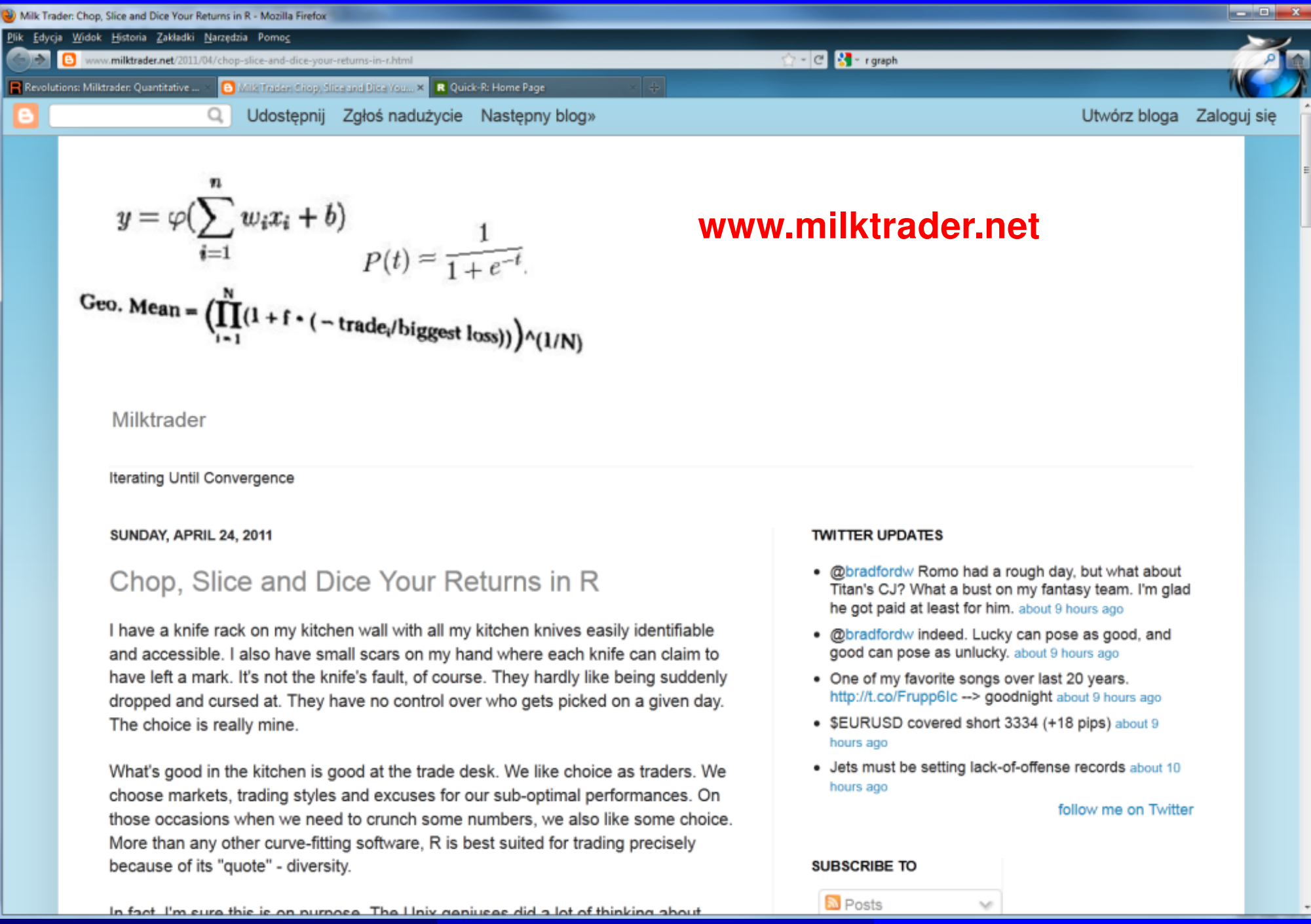
TWITTER UPDATES

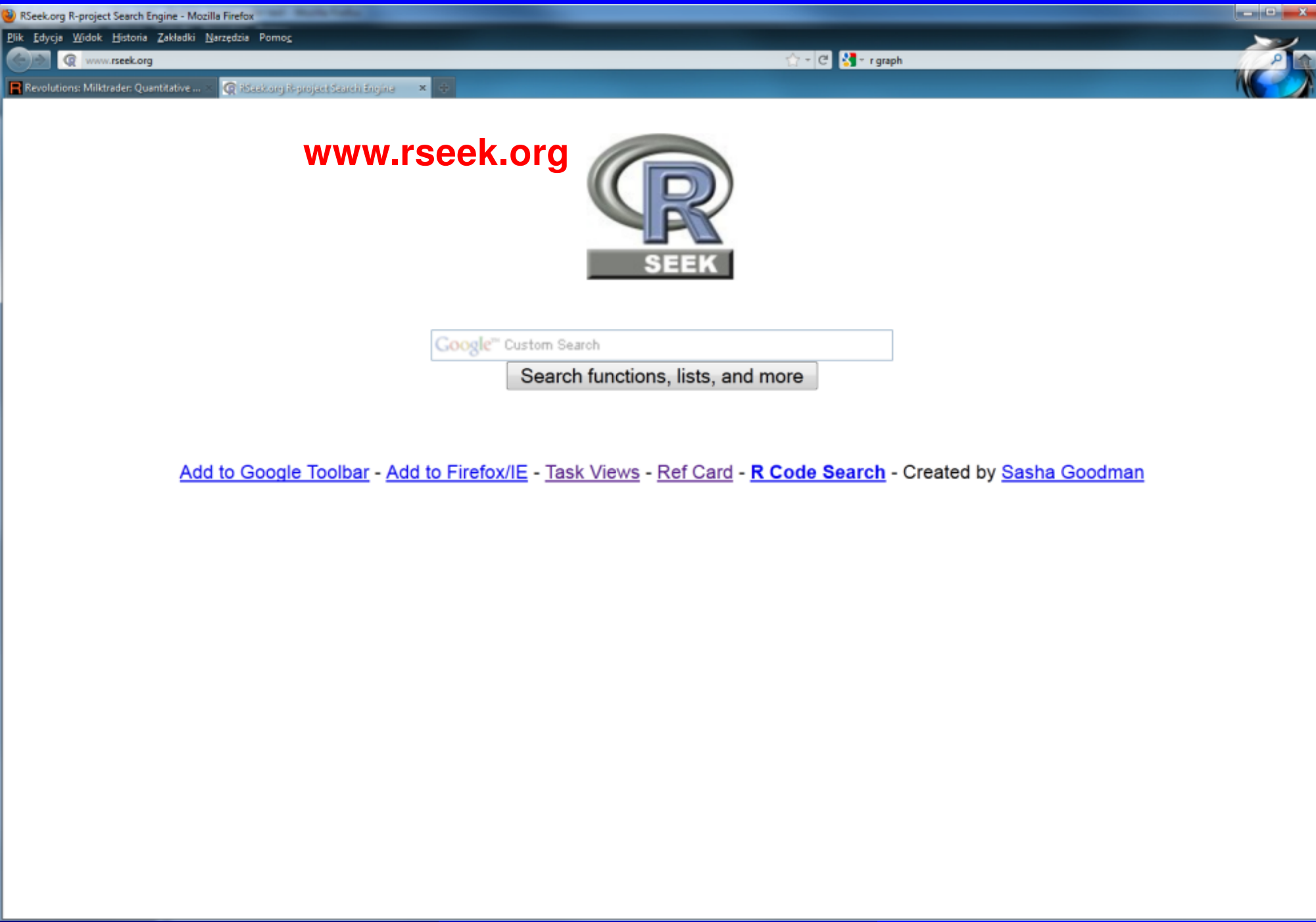
- @bradfordw Romo had a rough day, but what about Titan's CJ? What a bust on my fantasy team. I'm glad he got paid at least for him. about 9 hours ago
- @bradfordw indeed. Lucky can pose as good, and good can pose as unlucky. about 9 hours ago
- One of my favorite songs over last 20 years. http://t.co/Frupp6Ic --> goodnight about 9 hours ago
- $EURUSD covered short 3334 (+18 pips) about 9 hours ago
- Jets must be setting lack-of-offense records about 10 hours ago

follow me on Twitter

SUBSCRIBE TO

Posts




RSeek.org R-project Search Engine - Mozilla Firefox
Plik Edycja Widok Historia Zakładki Narzędzia Pomoc
www.rseek.org
r graph
Revolutions: Milktrader: Quantitative ...
RSeek.org R-project Search Engine
www.rseek.org
SEEK
Google Custom Search
Search functions, lists, and more
Add to Google Toolbar - Add to Firefox/IE - Task Views - Ref Card - R Code Search - Created by Sasha Goodman

data type	package	description
ts	base	regularly spaced time series
mts	base	regularly spaced multipletime series
zoo	zoo	regular/irregular and arbitrary time stamp classes
xts	xts	an extension of the zoo class

High-frequency begins when indexing only by date alone becomes insufficient (intra-day, tick data).

A time series object in R includes a data matrix and a vector of associated date/time stamps.

Classes for high-frequency time series:

	low frequency	high frequency
time series	zoo	xts
time index	Date	POSIXlt/POSIXct

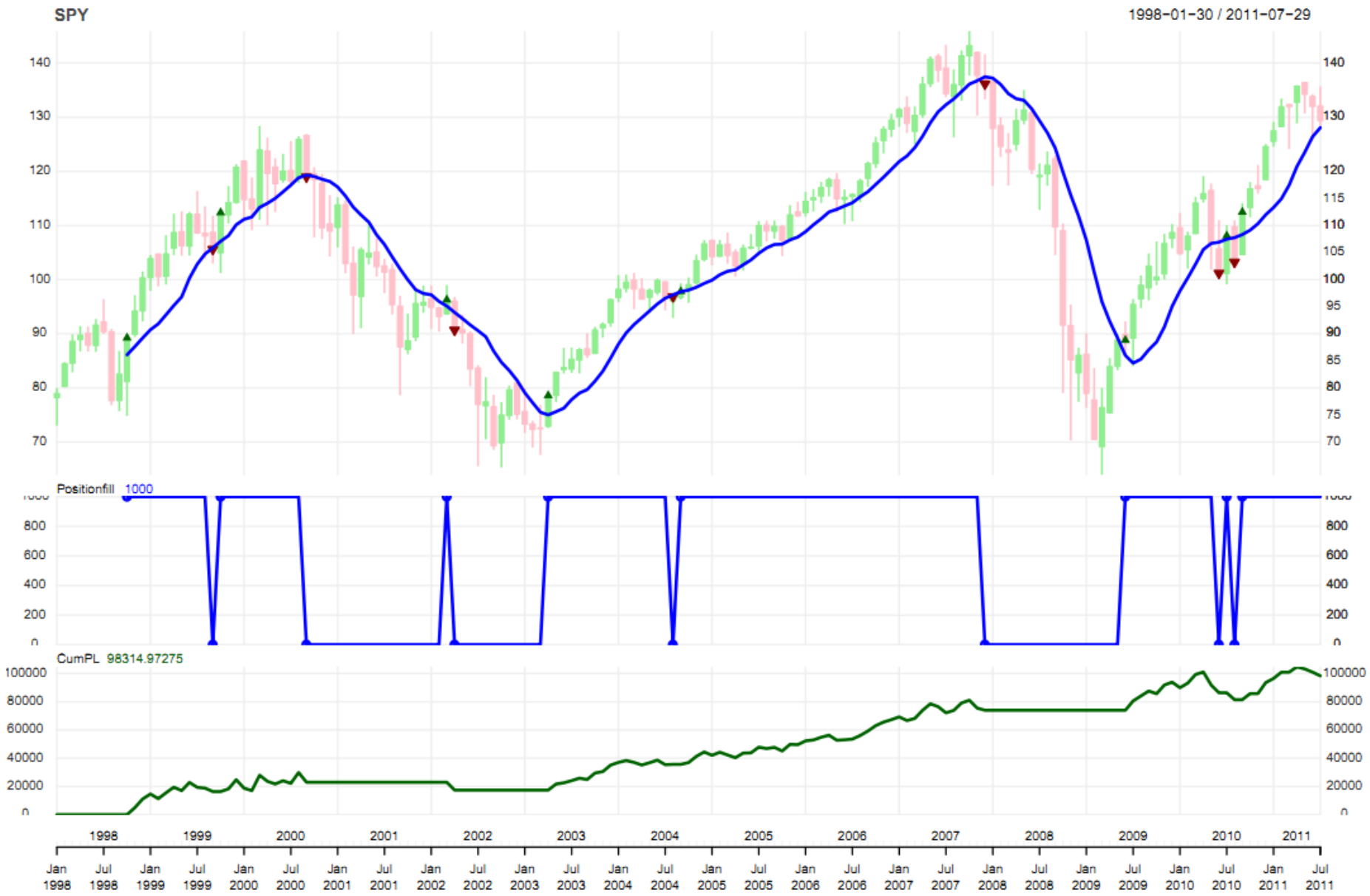
Useful packages

- **quantmod** – getting the historical and (delayed) real-time data automatically, many TA tools, easy analyses of OHLC type data,
- **TTR, caTools** – efficient functions for rolling analyses,
- **performanceAnalytics** – measures of strategy performance (eg. Sharpe Ratio, MDD, Calmar Ratio, Sortino Ratio, Upside Potential Ratio, Treynor Ratio),
- **realized** – package for realized variance analyses,
- **IBrokers, RTAQ** – direct connection to commercial HFD providers,
- and many many others. . .

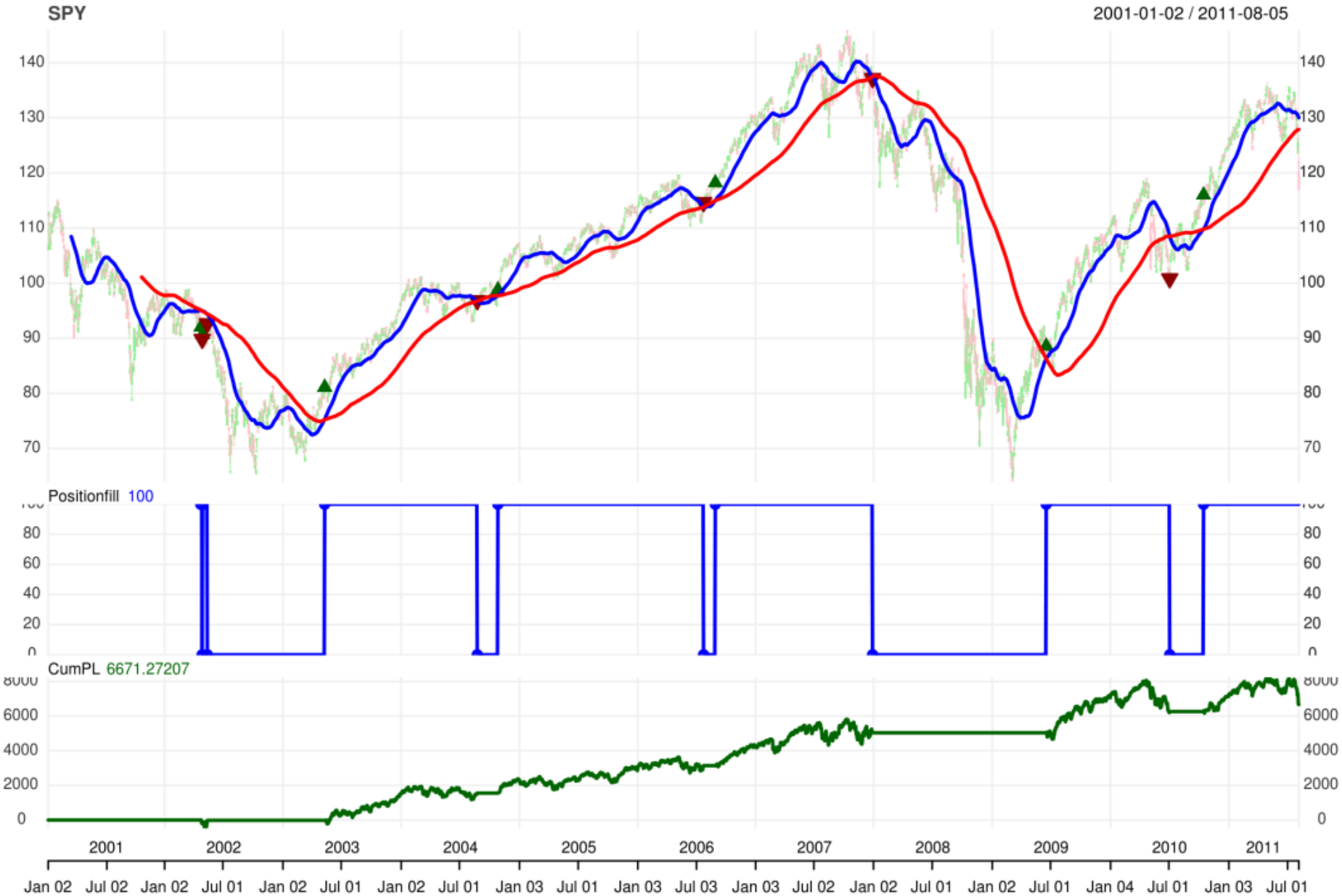
Sample plot

AAPL
[2012-03-01/2012-09-28]
Last 667.1
700
650
600
550
50
40
30
20
10
Volume (millions):
19,111,100
mar 01 2012
kwi 02 2012
maj 01 2012
cze 01 2012
lip 02 2012
sie 01 2012
wrz 04 2012
wrz 28 2012

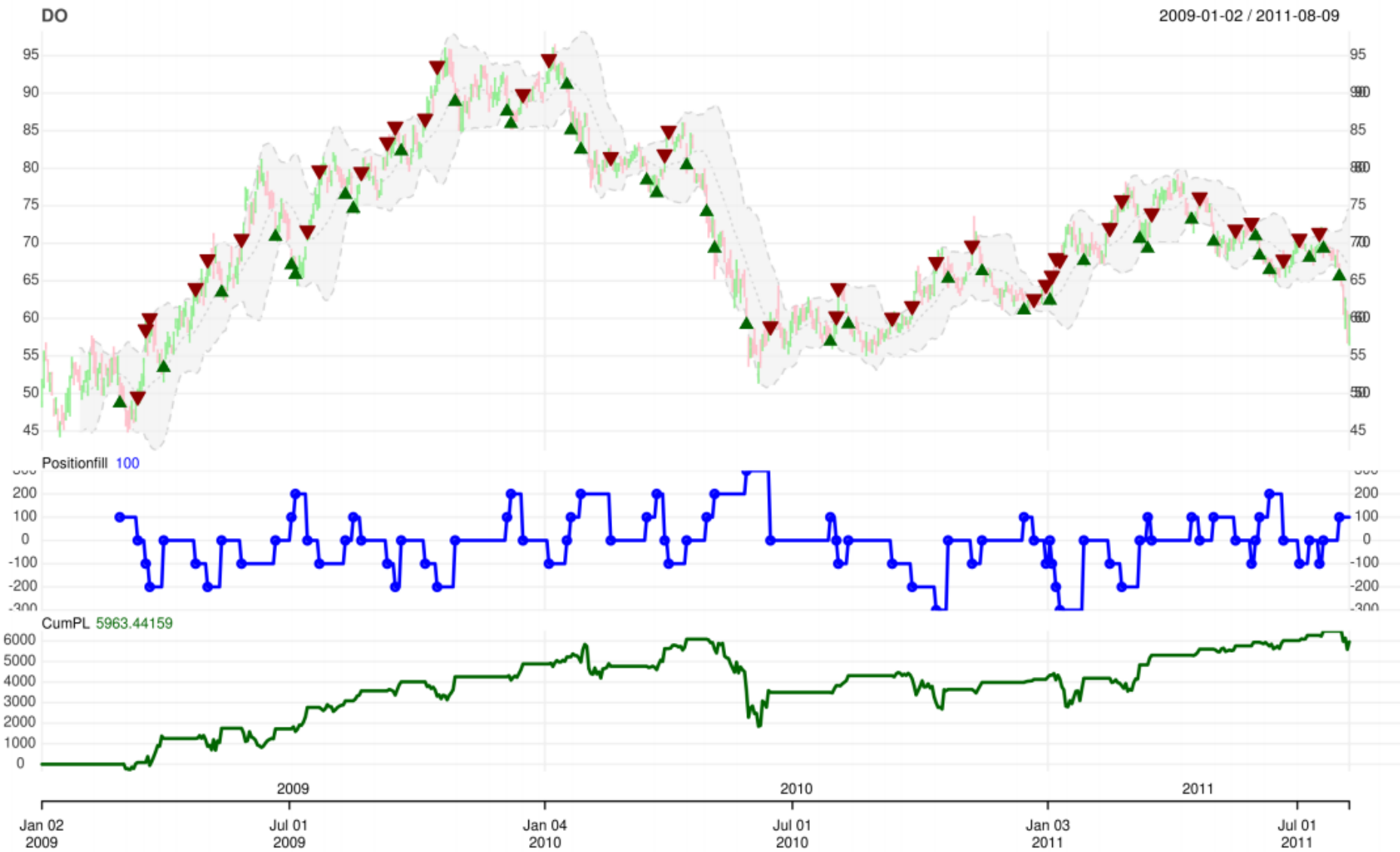
Strategy based on a single moving average



Strategy based on two intersecting moving averages

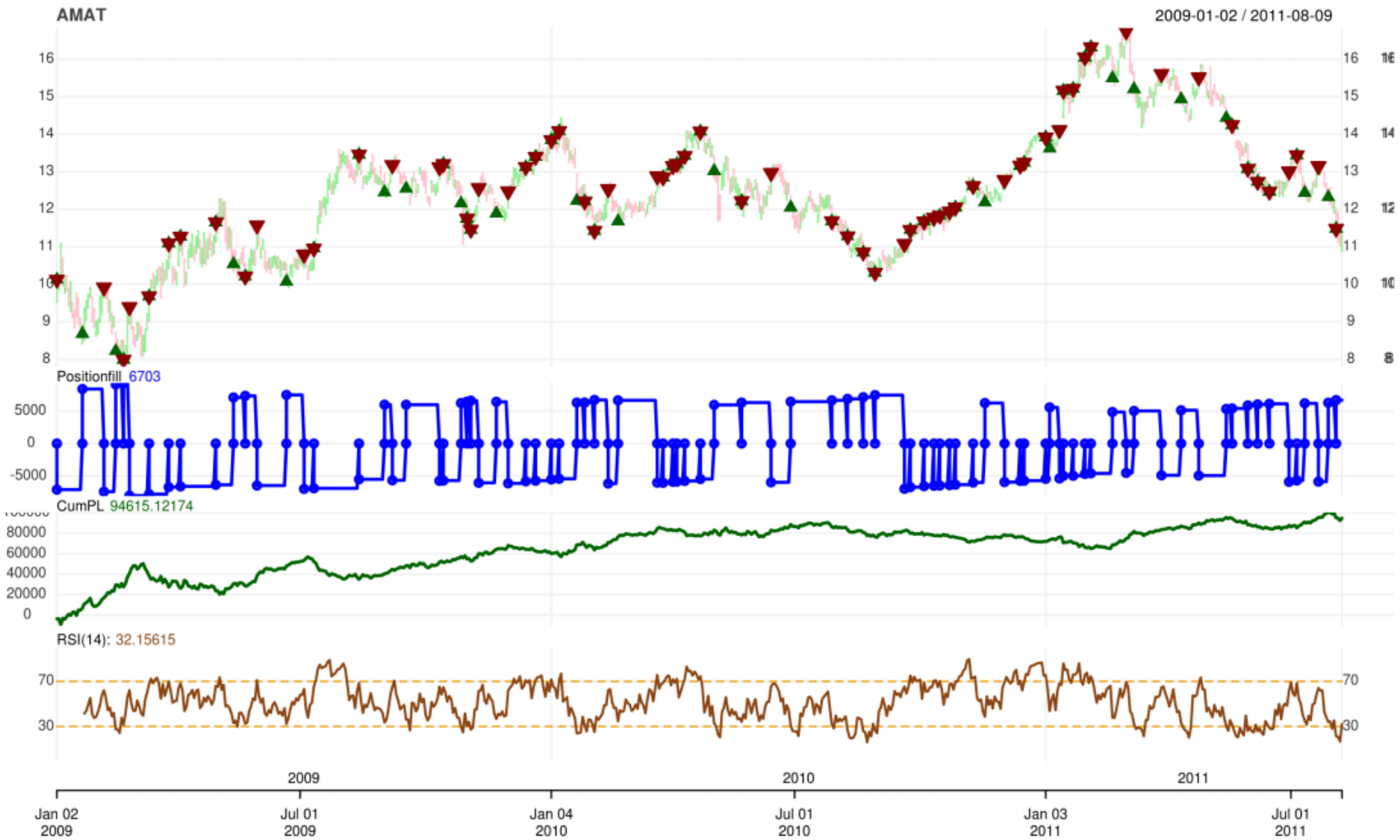


Strategy based on Bollinger Bands

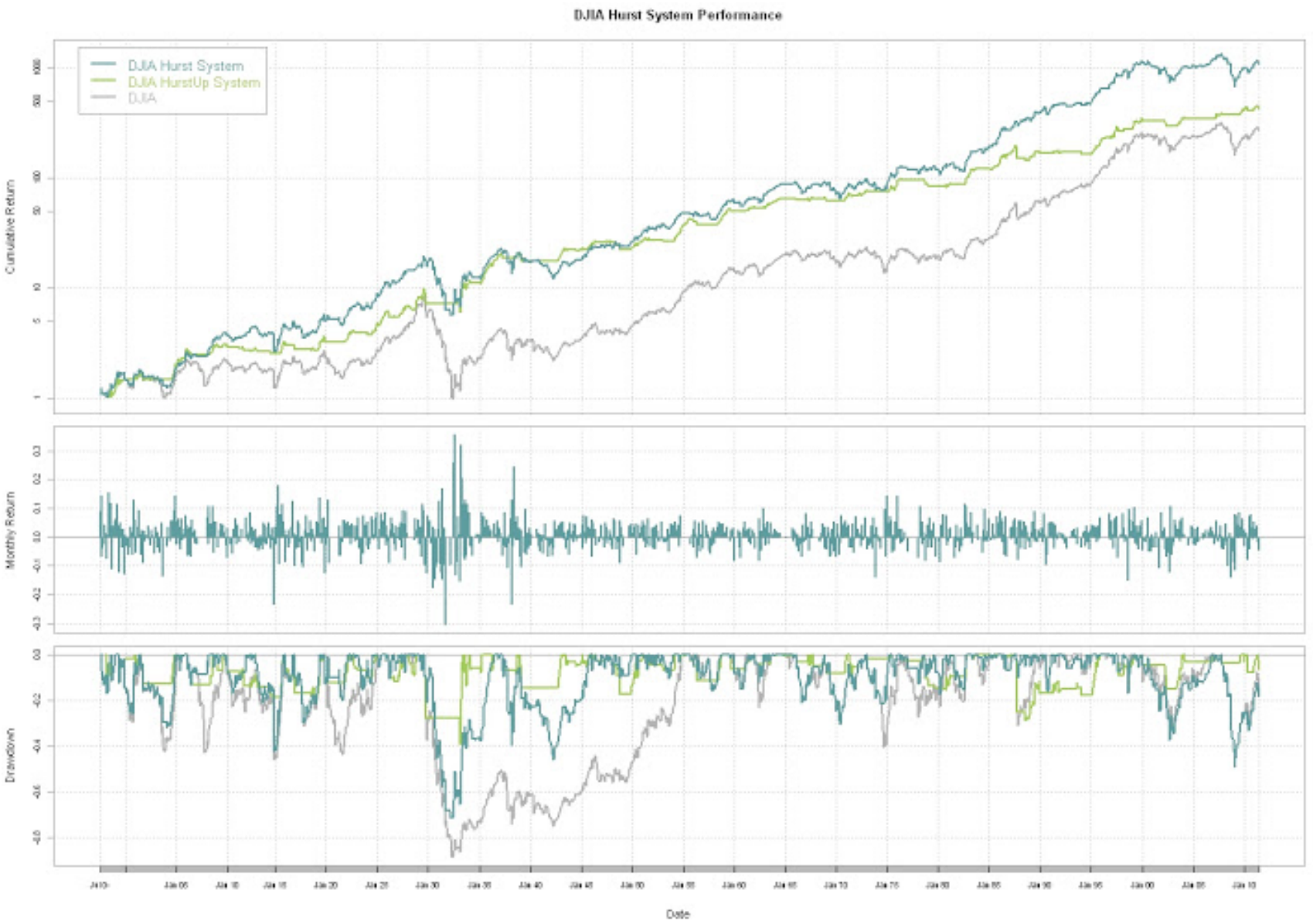


Strategy based on MACD oscillator

Strategy based on Relative Strength Index



Strategy based on a Hurst expnent



Thank you for your attention!