The Reproducible Computing package

GUEST EDITORS' INTRODUCTION

Computing in Science & Engineering, January/February 2009, pp. 5-7.

Reproducible Research

by Sergey Fomel and Jon F. Claerbout

Reproducibility is a core principle of science. For computational experiments to become reproducible, one needs to develop a system for linking scientific publications with computational recipes. Articles in this special issue argue in favor of computational reproducibility and describe several practical approaches to reproducible research.

Our generation of computational scientists is living in an exciting time: not only do we get to pioneer important algorithms and computations,we also get to set standards on how computational research should be conducted and published. From Euclid's reasoning and Galileo's experiments, it took hundreds of years for the theoretical and experimental branches of science to develop standards for publication and peer review. Computational science, rightly regarded as the third branch, can walk the same road much faster.

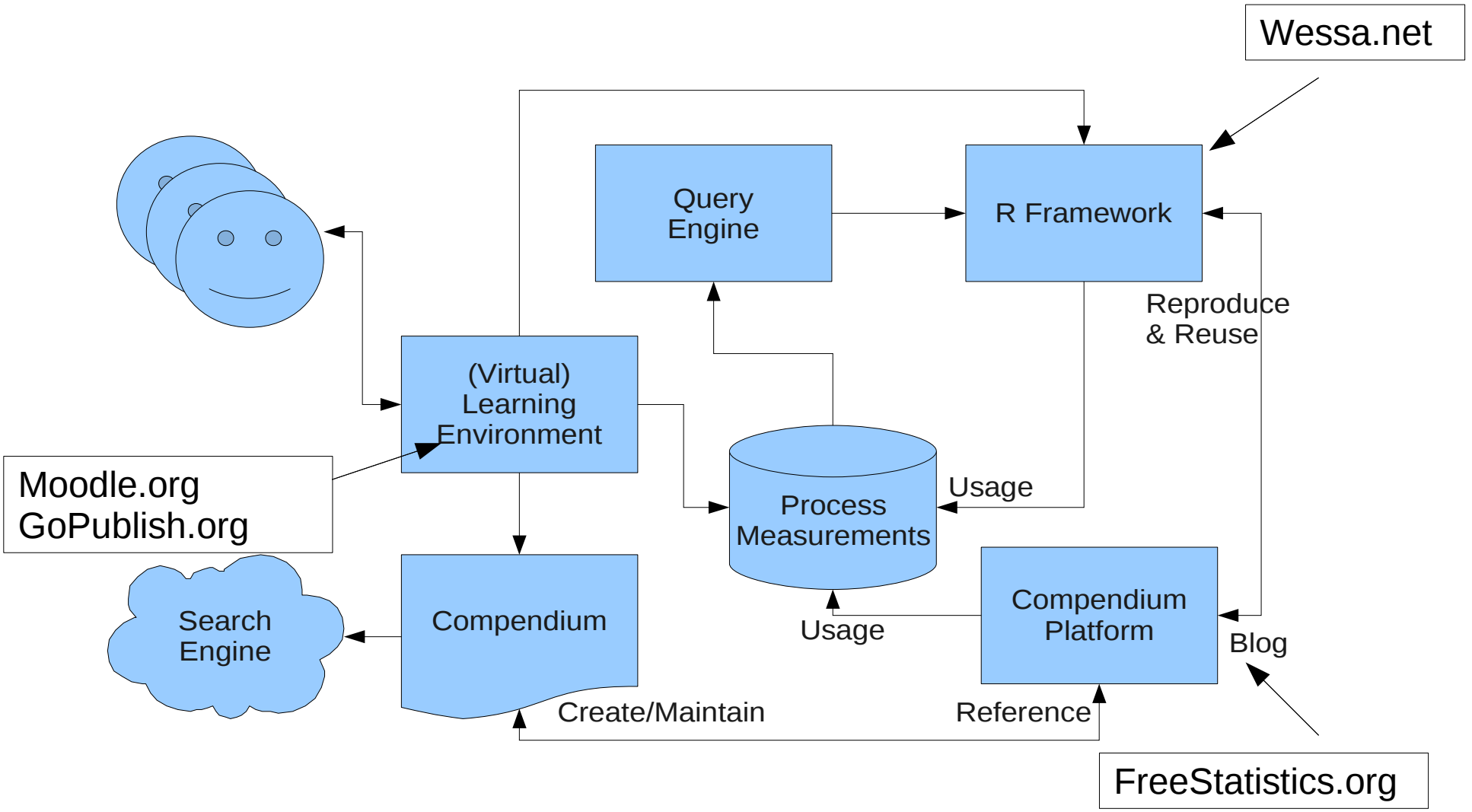
Some References

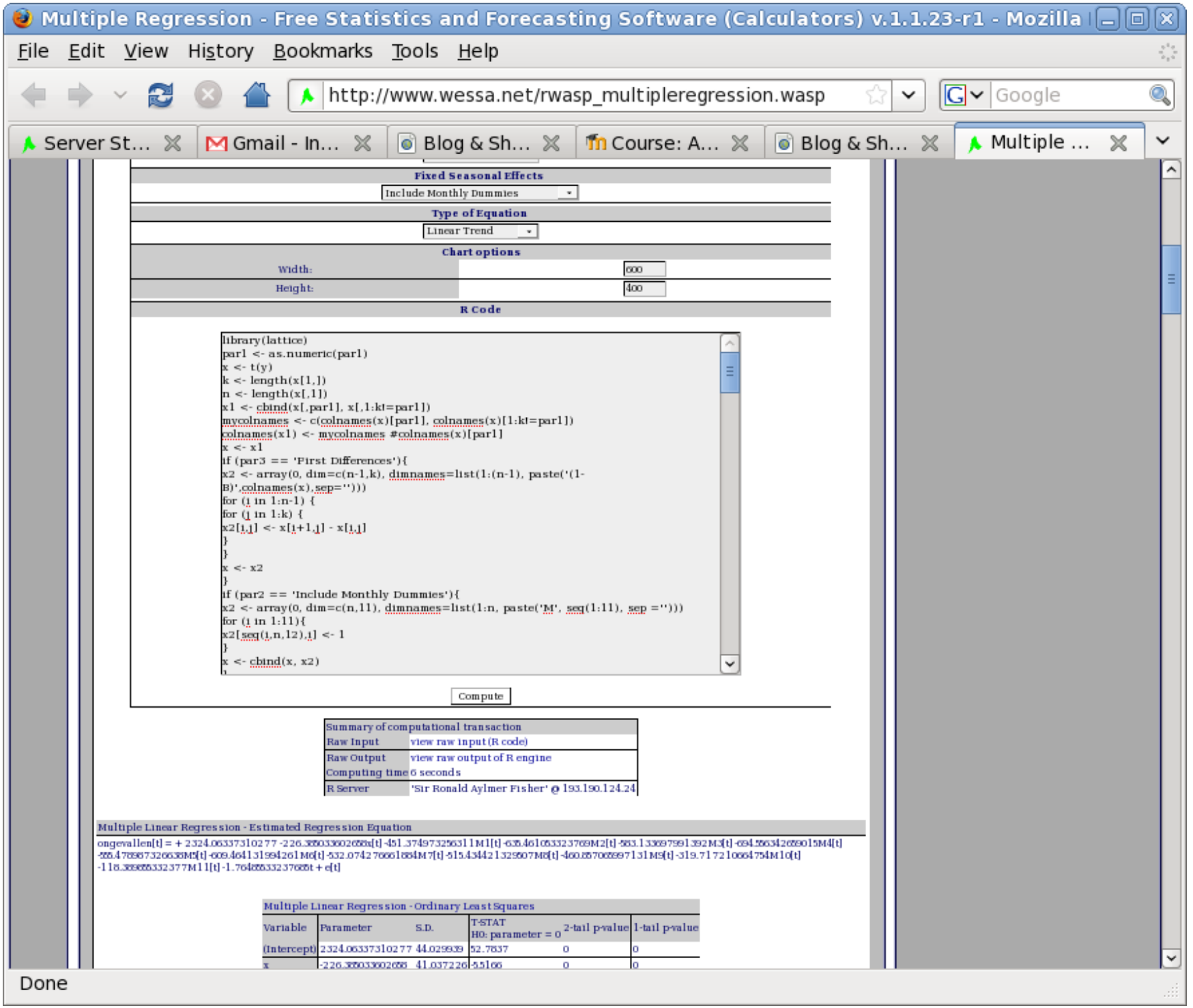
- **J. Buckheit and D. L. Donoho**. Wavelab and reproducible research. In A. Antoniadis, editor, Wavelets and Statistics, 1995.
- **Peter J. Green**. Diversities of gifts, but the same spirit. The Statistician, 2003.
- **T. R. Golub, et al**. Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. Science, 286:531–537, 1999.
- **David L. Donoho, Xiaoming Huo**, BeamLab and Reproducible Research, International Journal of Wavelets, Multiresolution and Information Processing, 2004
- **Roger D. Peng, Francesca Dominici, and Scott L. Zeger**, Reproducible Epidemiologic Research, American Journal of Epidemiology, 2006
- **R. Gentleman**, Reproducible Research: A Bioinformatics Case Study, Bioconductor
- **R. Gentleman**, Applying Reproducible Research in Scientific Discovery, BioSilico, 2005
- **Jan de Leeuw**, Reproducible Research: the Bottom Line, 2001, online
- **Roger Koenker, Achim Zeileis**, Reproducible Econometric Research (A Critical Review of the State of the Art), Department of Statistics and Mathematics Wirtschaftsuniversität Wien, Research Report Series, Report 60, November 2007
- **Robert Gentleman, Duncan Temple Lang**, Statistical Analyses and Reproducible Research, http://www.bepress.com/bioconductor/paper2
- **Schwab, M., Karrenbach, N. and Claerbout, J.** Making scientific computations reproducible, Computing in Science & Engineering, 2 (6), pp. 61-67, 2000.
- **Robert Gentleman**, Some Perspectives on Statistical Computing, online
- **Leisch, F.**, “Sweave and beyond: Computations on text documents”, Proceedings of the 3rd International Workshop on Distributed Statistical Computing, 2003, Vienna, Austria, ISSN 1609-395

- mefa package, Solymos P. (2008) (data prcessing/sharing in biogeography)
- http://thedata.org

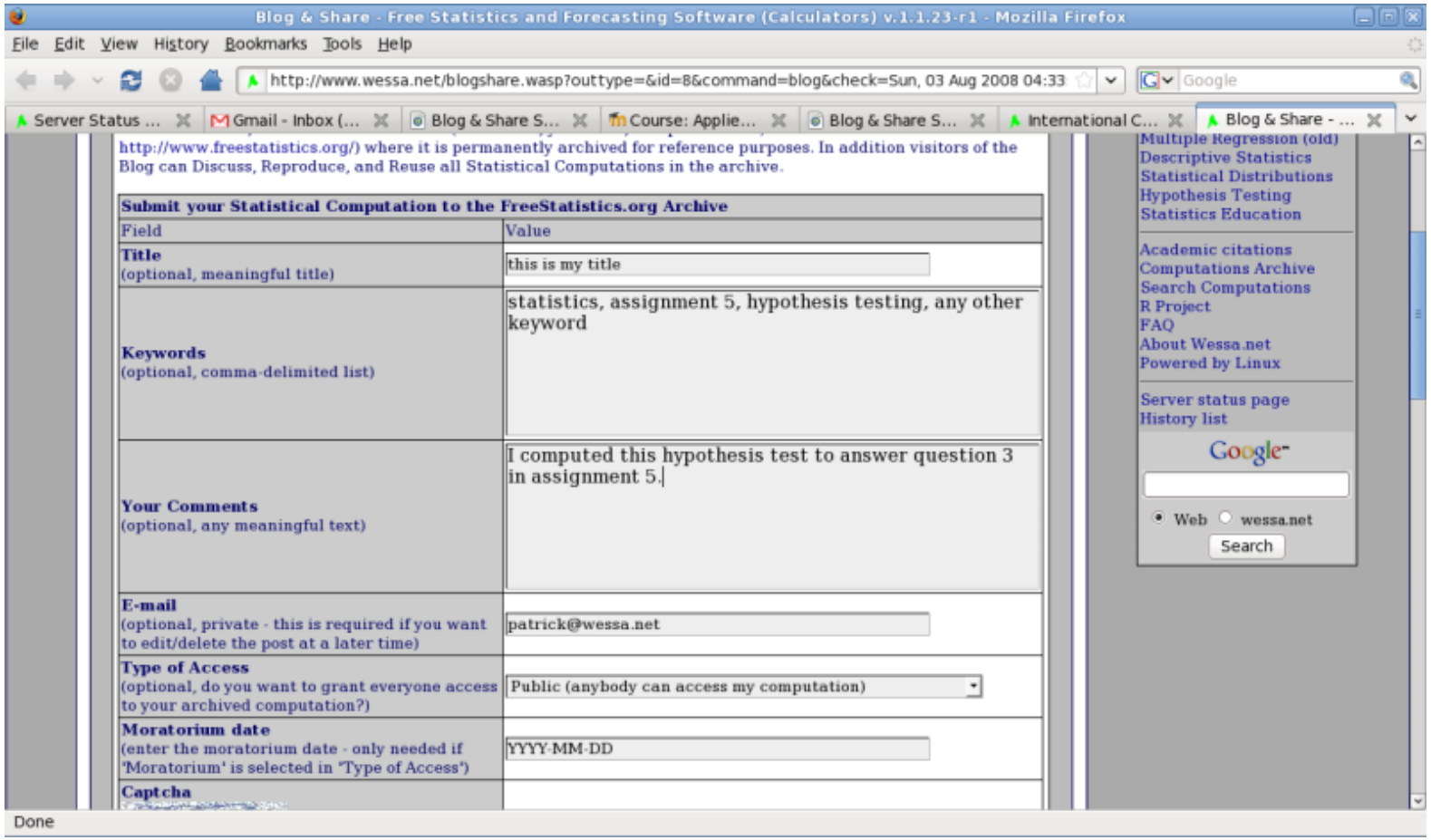
- http://www.FreeStatistics.org/
  -> Publications
  -> Repository
  -> RC package home

Learning System or Educational Laboratory?




Multiple Regression - Free Statistics and Forecasting Software (Calculators) v.1.1.23-r1 - Mozilla
File Edit View History Bookmarks Tools Help
http://www.wessa.net/rwasp_multipleregression.wasp
Google
Server St...
Gmail - In...
Blog & Sh...
Course: A...
Blog & Sh...
Multiple ...
Fixed Seasonal Effects
Include Monthly Dummies
Type of Equation
Linear Trend
Chart options
Width: 600
Height: 400
R Code
library(lattice)
par1 <- as.numeric(par1)
x <- t(y)
k <- length(x[1,])
n <- length(x[,1])
x1 <- cbind(x[,par1], x[,1:k!=par1])
mycolnames <- c(colnames(x)[par1], colnames(x)[1:k!=par1])
colnames(x1) <- mycolnames #colnames(x)[par1]
x <- x1
if (par3 == 'First Differences'){
x2 <- array(0, dim=c(n-1,k), dimnames=list(1:(n-1), paste('(1-
B)',colnames(x),sep='')))
for (i in 1:n-1) {
for (j in 1:k) {
x2[i,j] <- x[i+1,j] - x[i,j]
}
}
x <- x2
}
if (par2 == 'Include Monthly Dummies'){
x2 <- array(0, dim=c(n,11), dimnames=list(1:n, paste('M', seq(1:11), sep ='')))
for (i in 1:11){
x2[seq(i,n,12),i] <- 1
}
x <- cbind(x, x2)
Compute
Summary of computational transaction
Raw Input view raw input (R code)
Raw Output view raw output of R engine
Computing time 6 seconds
R Server 'Sir Ronald Aylmer Fisher' @ 193.190.124.24
Multiple Linear Regression - Estimated Regression Equation
ongevallen[t] = + 2324.06337310277 -226.385033602688x[t] -451.374973256311M1[t] -635.461053323769M2[t] -583.133697591392M3[t] -694.566342688015M4[t]
-555.478887326638M5[t] -609.464131594261M6[t] -532.074276661884M7[t] -515.434421329507M8[t] -460.857068997131M9[t] -319.717210664754M10[t]
-118.388668332377M11[t] -1.76488533237688t + e[t]
Multiple Linear Regression - Ordinary Least Squares
Variable Parameter S.D. T-STAT H0: parameter = 0 2-tail p-value 1-tail p-value
(Intercept) 2324.06337310277 44.029939 52.7837 0 0
x -226.385033602688 41.037226 -5.5106 0 0
Done

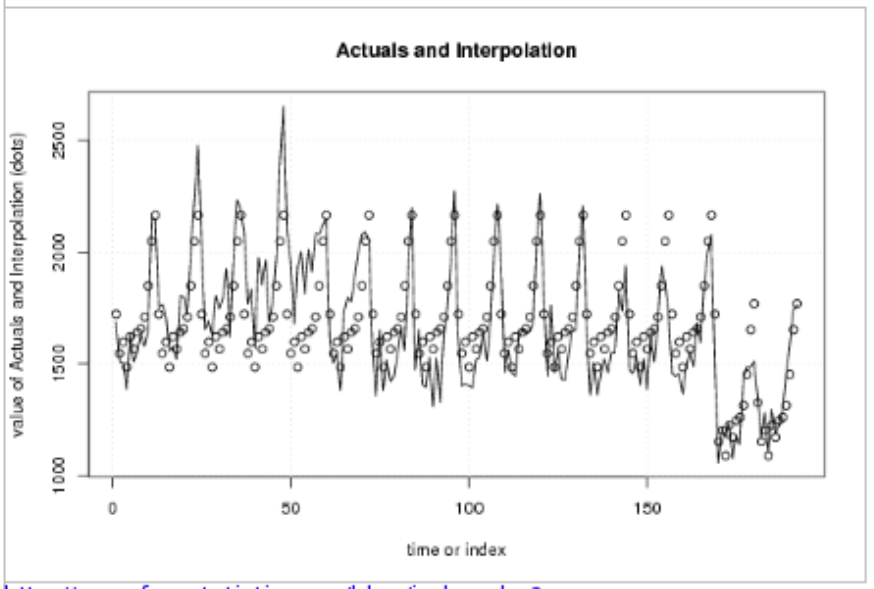
Computations are “blogged” (not archived)



Question 2: Investigate the prediction errors of the model that you used in question 1. Are the underlying regression assumptions satisfied?

De Adjusted R-squared is gelijk aan 0,6412. Dit wil zeggen dat we 64% van de wijzigingen van het aantal verkeersslachtoffers kunnen verklaren. Het resterende gedeelte (36%) kunnen we niet verklaren aan de hand van ons model maar zijn bijvoorbeeld te wijten aan uitzonderlijke weersomstandigheden.

Het volgende waar we heen zullen kijken is de Interpolation Plot. De stippellijn op deze grafiek geeft het werkelijke aantal slachtoffers weer. De volle lijn geeft het aantal verkeersslachtoffers weer die voorspeld zijn door het model.



http://www.freestatistics.org/blog/index.php?v=date/2007/Nov/14/t1195074007ni07puuvacjlu0w.htm

We zien ook hier een dalende trend op lange termijn. Op het einde zien we een structurele breuk, dit is het effect van de seatbelt law (-395). Er treedt ook een

Novelty about RC package?

- “RC.blog” R code from your console
- “RC.reproduce” computations in your console
- “RC.ls” computations (by keyword)
- reuse “RC.meta.data” of computations
- build a “RC.tree” of computations based on parent-child relationships (and “RC.print.tree” it)
- ... and much more in the near future...

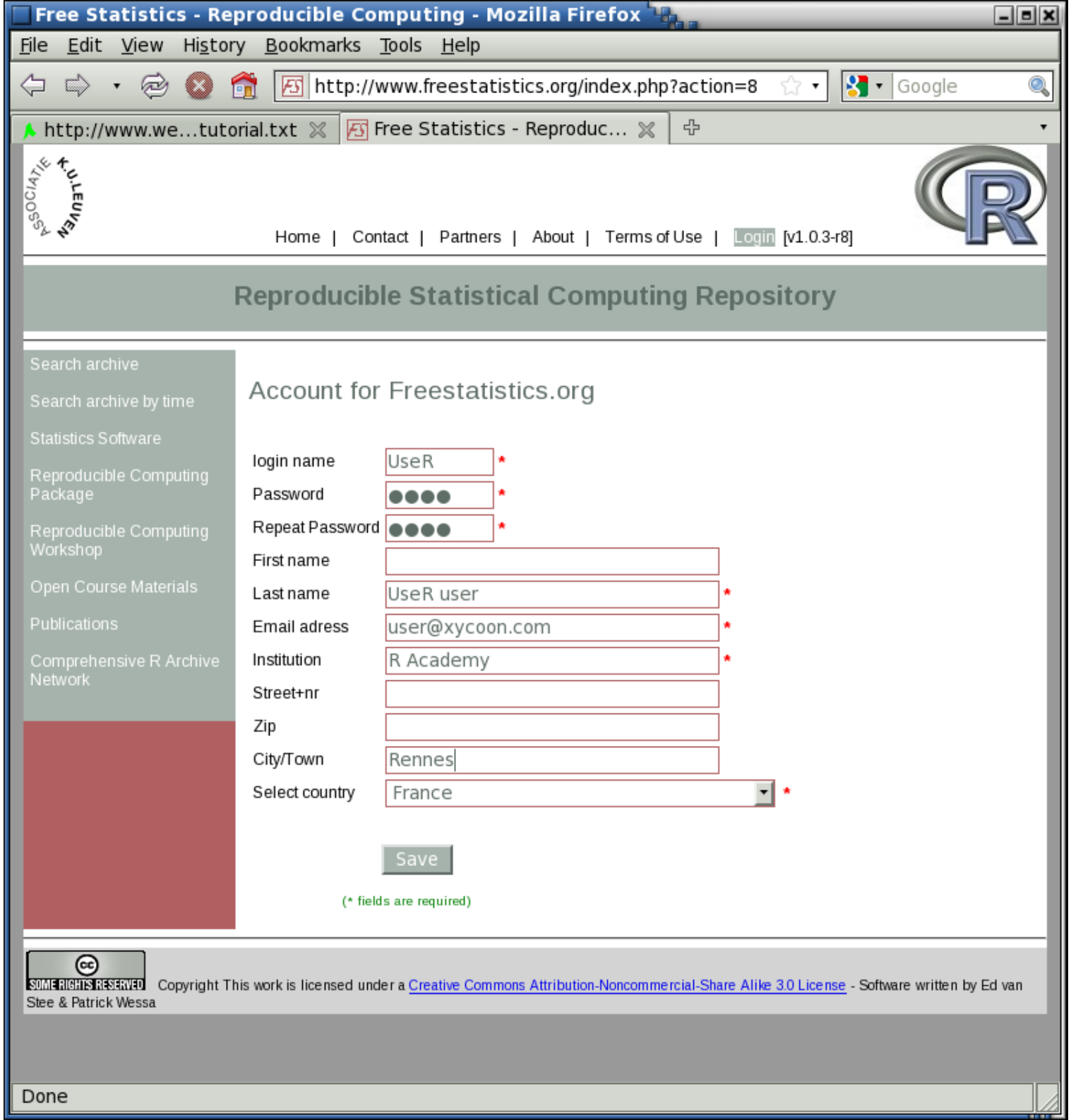
saving/loading image files

```
#extremely slow
> RC.save.image(keywords="testuser2009")
HTTP/1.1 200 OK
Date: Mon, 06 Jul 2009 14:57:56 GMT
Server: Apache/2.2.8 (Fedora)
X-Powered-By: PHP/5.2.6
Content-Length: 376
Connection: close
Content-Type: text/html

Submission to R Framework completed. Waiting for reply from FreeStatistics.org...

Your submission to FreeStatistics.org is complete. Thank you for sharing your computations & comments!
You can view your submission at
http://www.freestatistics.org/blog/date/2009/Jul/06/t1246892281gxgeiltqrwcs57j.htm.
Warning message:
In RC.save.image(keywords = "testuser2009") : No title was specified.

#very fast
> RC.load("http://www.freestatistics.org/blog/date/2009/Jul/06/t1246892281gxgeiltqrwcs57j/Rimage.RData")
```


Free Statistics - Reproducible Computing - Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://www.freestatistics.org/index.php?action=8
Google
http://www.we...tutorial.txt
Free Statistics - Reproduc...
Home | Contact | Partners | About | Terms of Use | Login [v1.0.3-r8]
Reproducible Statistical Computing Repository
Search archive
Search archive by time
Statistics Software
Reproducible Computing Package
Reproducible Computing Workshop
Open Course Materials
Publications
Comprehensive R Archive Network
Account for Freestatistics.org
login name UseR *
Password *
Repeat Password *
First name
Last name UseR user *
Email adress user@xycoon.com *
Institution R Academy *
Street+nr
Zip
City/Town Rennes
Select country France *
Save
(* fields are required)
SOME RIGHTS RESERVED Copyright This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 License - Software written by Ed van Stee & Patrick Wessa
Done

Say hello to RC network

```
#library(RC) fetches fresh code from internet
#use at own risk:
> source("http://Send me an e-mail if you want to know the URL")
> RC.hello()

[1] "Calling R Framework server network. This may take a while..."
HTTP/1.1 200 OK
Date: Sun, 05 Jul 2009 18:54:04 GMT
Server: Apache/2.2.8 (Fedora)
X-Powered-By: PHP/5.2.6
Content-Length: 576
Connection: close
Content-Type: text/html

R Framework is online.
 Main webserver             system capacity   : EXCELLENT
'Herman Ole Andreas Wold'  system capacity   : EXCELLENT     response time :  0.42455697059631 seconds
'Gwilym Jenkins'           system capacity   : EXCELLENT     response time :  0.22293996810913 seconds
'George Udny Yule'         system capacity   : EXCELLENT     response time :  0.32254195213318 seconds
'Sir Ronald Aylmer Fisher' system capacity   : EXCELLENT     response time :  0.42430806159973 seconds

Note: response times are measured between the main webserver and each R server.

   user  system elapsed
  0.003   0.000   1.996

>
```

Code snippet 1

```
x <- rnorm(150)
y <- rnorm(150)
cor.test(x,y)
plot(x,y)
```

the above code snippet is wrapped into a function, and the graphics device is opened/closed

```
my.fun <- function() {
x <- rnorm(150)
y <- rnorm(150)
print(cor.test(x,y))
RC.start.plot
plot(x,y)
RC.end.plot
}
```

now we “blog” the function:

```
> RC.blog(title='my first computation', keywords='tutorial test', comments='This is the first time that
UseR is blogging a computation.', uid='UseR', pwd='UseR', typeofaccess='public', rcode=my.fun)
HTTP/1.1 200 OK
Date: Mon, 06 Jul 2009 06:49:57 GMT
Server: Apache/2.2.8 (Fedora)
X-Powered-By: PHP/5.2.6
Content-Length: 376
Connection: close
Content-Type: text/html

Submission to R Framework completed. Waiting for reply from FreeStatistics.org...

Your submission to FreeStatistics.org is complete. Thank you for sharing your computations & comments!
You can view your submission at
http://www.freestatistics.org/blog/date/2009/Jul/06/t1246862999odwh34bz66dnt0p.htm.
[1] "http://www.freestatistics.org/blog/date/2009/Jul/06/t1246862999odwh34bz66dnt0p.htm"
```

```
RC.browse("http://www.freestatistics.org/blog/date/2009/Jul/06/t1246862999odwh34bz66dnt0p.htm")
```


```
> source("http://www.freestatistics.org/blog/index.php?v=date/2009/Jul/06/t1246862999odwh34bz66dnt0p.htm&rcode=T")

Pearson's product-moment correlation

data:  x and y
t = 0.3299, df = 148, p-value = 0.742
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.1337382  0.1865555
sample estimates:
       cor
0.02710428

> r <- RC.ls(keyword='tutorial*')
[1] "Fetching list from FreeStatistics.org archive..."
[1] "Number of valid cases found: 26."
> r$user
 [1] Truyts Kevin         Engels Kevin         Machiels Romina
 [4] Machiels Romina      Van Riet Jan         Van Riet Jan
 [7] Van Riet Jan         De Wilde Natalie     Van Ham Ellen
[10] Van den Heuvel Koen Van den Heuvel Koen Geudens Gert-Jan
[13] Sergoynne Sofie      Van Ham Ellen        Claes Stéphanie
[16] Claassens Jens       Moons Bert           Machiels Romina
[19] Machiels Romina      Moons Bert           Moons Bert
[22] Moons Bert           Van Dooren Leen      Moons Bert
[25] Michel Jeroen        UseR user
15 Levels: Claassens Jens Claes Stéphanie De Wilde Natalie ... Van Riet Jan
```

```
> r[26,]
                                                                      url
26 http://www.freestatistics.org/blog/date/2009/Jul/06/t1246862999odwh34bz66dnt0p.htm
                          key                   folder                date
26 t1246862999odwh34bz66dnt0p /blog/date/2009/Jul/06/ 2009-07-06 06:49:57
      module                 title      keywords    course       user parent
26 R console my first computation tutorial test R console UseR user
   message
26       0
> (md <- RC.meta.data(r$url[26]))
$type
[1] "Rscript"
$date
[1] "Mon, 06 Jul 2009 00:49:57 -0600"
$rmodulecode
[1] "\n{\n    x <- rnorm(150)\n    y <- rnorm(150)\n    print(cor.test(x, y))\n    \n    plot(x, y)\n    \n}"
$rawinput
[1] "\n{\n    x <- rnorm(150)\n    y <- rnorm(150)\n    print(cor.test(x, y))\n    \n    plot(x, y)\n    \n}"
$rawoutput
[1] "\n> {\n+     x <- rnorm(150)\n+     y <- rnorm(150)\n+     print(cor.test(x, y))\n+     plot(x,
y)\n+ }\n\n\tPearson's product-moment correlation\n\ndata:  x and y \nt = -1.5048, df = 148, p-value =
0.1345\nalternative hypothesis: true correlation is not equal to 0 \n95 percent confidence interval:\n
-0.27755888  0.03825629 \nsample estimates:\n       cor \n-0.1227579 \n\n\n"

> labels(RC.meta.data(RC.ls(keyword="growth")$url[3]))
[1] "Fetching list from FreeStatistics.org archive..."
[1] "Number of valid cases found: 10."
 [1] "type"        "date"        "uid"         "title"       "target"
 [6] "rawinput"    "rawoutput"   "output"      "ylimmax"     "ylimmin"
[11] "chartxlab"   "chartylab"   "chartheight" "chartwidth"  "par1"
[16] "par2"        "par3"        "par4"        "par5"        "par6"
[21] "par7"        "par8"        "par9"        "par10"       "par11"
[26] "par12"       "par13"       "par14"       "par15"       "par16"
[31] "par17"       "par18"       "par19"       "par20"       "parent"
[36] "data"        "newformula"


TODO: return pictures in postscript (already available on the website)
```

Code snippet 2

```
RCx <- data.frame(array(rnorm(100),dim=c(50,2)))
RCxnames <- c("X1","X2")
RC.sample.1 <- function(first_number=5,second_number=7,strpar="main title") {
  myfun <- function(x,y) {x+y}
  RC.start.plot
  plot(RCx,main=strpar,xlab='my xlab',ylab='my ylab')
  RC.end.plot
  RC.start.plot
  hist(RCx[,1],main="my histogram")
  RC.end.plot
  RC.start.plot
  pairs(RCx,main="pairs plot")
  RC.end.plot
  print(myfun(first_number,second_number))
}

> RC.blog(title='fixed data', keywords='UseR1', comments='', uid='UseR', pwd='UseR', typeofaccess='public',
rcode=RC.sample.1)
HTTP/1.1 200 OK
Date: Mon, 06 Jul 2009 08:50:03 GMT
Server: Apache/2.2.8 (Fedora)
X-Powered-By: PHP/5.2.6
Content-Length: 376
Connection: close
Content-Type: text/html

Submission to R Framework completed. Waiting for reply from FreeStatistics.org...

Your submission to FreeStatistics.org is complete. Thank you for sharing your computations & comments!
You can view your submission at
http://www.freestatistics.org/blog/date/2009/Jul/06/t1246870205dca8pzlyzslfrvk.htm.
[1] "http://www.freestatistics.org/blog/date/2009/Jul/06/t1246870205dca8pzlyzslfrvk.htm"
```

Now we have fixed data

```
> r <- RC.ls(keyword="UseR1")
[1] "Fetching list from FreeStatistics.org archive..."
[1] "Number of valid cases found: 1."
> cat(RC.meta.data(r$url[1])$rawinput)
x <- array(c(0.0327570747625087, -1.01260220468867, 0.987781241007297, -0.04686368515551,
-0.474607692103688, -0.0372435023825232, [...truncated...] -0.708516781545271,
0.899414776157957),dim=c(2,50),dimnames=list(c("X1", "X2"), 1:50))
y <- array(NA,dim=c(2,50),dimnames=list(c("X1", "X2"), 1:50))
for (i in 1:dim(x)[1]) {
for (j in 1:dim(x)[2]) {
y[i,j] <- as.numeric(x[i,j])
}
}
x <- t(y)
{
    myfun <- function(x, y) {
        x + y
    }
    plot(x, main = par3, xlab = "my xlab", ylab = "my ylab")
    hist(x[, 1], main = "my histogram")
    pairs(x, main = "pairs plot"
    print(myfun(par1, par2))

}

> RC.browse(r$url[1])
```

-1.05108305897369	-0.415106662193027
0.0692836959108357	1.17212983625136
-0.711709414856647	0.594084948288943
-0.399536337143537	0.902343374942135
-1.45135708457178	0.472871502086415
-0.417792663937469	-0.572838513095986
-1.07273200120520	-0.519766499597772
2.36115292254887	0.825067154141609
0.155621956518056	-0.716691498447023
0.124334984296978	0.50346277269306

Names of X columns:

X1 X2

Sample Range:
(leave blank to include all observations)

From:	
To:	

Chart options

Width:	600
Height:	400
Label y-axis:	Y
Label x-axis:	X

Compute

Summary of computational transaction	
Raw Input	view raw input (R code)
Raw Output	view raw output of R engine
Computing time	0 seconds
R Server	'Gwilym Jenkins' @ 72.249.127.135

Two easy steps to reproduce

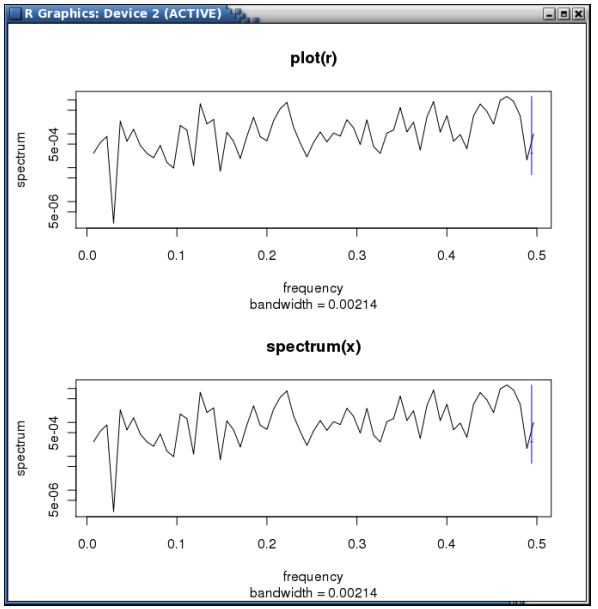
```
#First you obtain the URL of the computation
> r <- RC.ls(keyword='tutorial*test')
[1] "Fetching list from FreeStatistics.org archive..."
[1] "Number of valid cases found: 1."

#or you simply copy&paste it from a publication
```


```
#reproduce the computation in your console [output not shown]
> RC.reproduce(r$url[1])
> RC.reproduce('http://www.freestatistics.org/blog/date/2009/Jul/06/t1246862999odwh34bz66dnt0p.htm/')
> source("http://www.freestatistics.org/blog/index.php?v=date/2009/Jul/06/t1246862999odwh34bz66dnt0p.htm&rcode=T")
#note: picture is also generated on the default graphics device on your local machine
```

Warning

```
> r <- RC.ls(keyword="AS2009")
[1] "Fetching list from FreeStatistics.org archive..."
[1] "Number of valid cases found: 2."
> md <- RC.meta.data(r$url[2])
> cat(RC.prepare.input(md$rawinput))
[...truncated...]
r <- spectrum(x,main="Raw Periodogram")
[...truncated...]
> RC.reproduce(r$url[2])
[1] "> x <- c(112, 118,  .... [TRUNCATED] \n> a <- table.end(a)\n"
> #r does not contain the search results anymore because the reproduced R script uses the variable
r to hold the results of the spectral analysis about x
> op <- par(mfrow=c(2,1))
> plot(r,main="plot(r)")
> spectrum(x,main="spectrum(x)")
> par(op)
```



Meta Data – Data Mining

```
> r <- RC.ls(keyword = "retail sales")
[1] "Fetching list from FreeStatistics.org archive..."
[1] "Number of valid cases found: 12."
> mytree <- RC.tree(r$url[5])
> RC.print.tree(mytree)
Univariate Data Series[HPC Retail Sales][2008-03-02 15:42:48][]
|    Structural Time Series Models[ HPC Retail Sales][2008-03-06 16:52:55][]***
|    |    Structural Time Series Models[ HPC Retail Sales][2008-03-08 11:12:03][]
|    |    Structural Time Series Models[ HPC Retail Sales][2008-03-08 11:33:35][]
> mytree <- RC.tree(mytree$url[1])
> RC.print.tree(mytree)
Univariate Data Series[ HPC Retail Sales][2008-03-02 15:42:48][]***
|    Structural Time Series Models[ HPC Retail Sales][2008-03-06 16:52:55][]
|    |    Structural Time Series Models[ HPC Retail Sales][2008-03-08 11:12:03][]
|    |    Structural Time Series Models[ HPC Retail Sales][2008-03-08 11:33:35][]
|    Classical Decomposition[ Multiplicative mo...][2008-04-03 10:35:14][]
|    Classical Decomposition[ decomp verkoop][2008-04-28 12:19:26][]
...[truncated]
```

```
> r <- RC.ls(keyword = "Exercise")
[1] "Fetching list from FreeStatistics.org archive..."
[1] "Number of valid cases found: 724."
> mytree <- RC.tree(r$url[6])
> (mytab <- table(mytree$level))

  1  10  11  12  13   2   3   4   5   6   7   8   9
  1   2   1   1   3 532 122  47  12   8   4   8   2
> RC.print.tree(mytree)
...[truncated]
|     |     |     |     |     |    Exercise 1.13[ test user][2008-10-16 10:11:31][edje]
|     |     |     |     |     Exercise 1.13[ vraag 1 poging 1][2008-10-16 10:38:19][Van den Eynde Evelin]
|     |     |     |     |     Exercise 1.13[ vraag 2 pog 1][2008-10-16 10:44:32][Van den Eynde Evelin]
|     |     |     |     |     Univariate Data Series[ oiokok][2008-10-16 10:54:35][Van den Eynde Evelin]
|     |     |     |     |     Exercise 1.13[ Aantal geboortes ...][2008-10-17 15:48:05][Blondeau Matthieu]
|     |     |     |     |     Univariate Data Series[ Tijdreeks 1: Huur...][2008-10-20 15:52:32][Jackers Veerle]
|     |     |     |     |     Univariate Data Series[ Tijdreeks 2: Gaso...][2008-10-20 15:56:05][Jackers Veerle]
|     |     |     |     |     |    Variance Reduction Matrix[ Identification/es...][2008-12-03 21:31:10][Jackers
|     |     |     |     |     |    (Partial) Autocorrelation Function[ Identification/es...][2008-12-03 21:36..
|     |     |     |     |     |    |    Spectral Analysis[ Identification/es...][2008-12-03 21:43:51][Jackers Ve
|     |     |     |     |     |    |    |    Spectral Analysis[ Identification/es...][2008-12-03 21:47:18][Jacker
|     |     |     |     |     |    |    |    |    Standard Deviation-Mean Plot[ Identification/es...][2008-12-05
|     |     |     |     |     |    |    |    |    |    (Partial) Autocorrelation Function[ Identification/es...]
|     |     |     |     |     |    |    |    |    |    |    ARIMA Backward Selection[ Identification/es...][200
|     |     |     |     |     |    |    |    |    |    |    |    ARIMA Backward Selection[ Identification/es...]
|     |     |     |     |     |    |    (Partial) Autocorrelation Function[ Identification/es...][2008-12-03 21:39:45
|     |     |     |     |     Univariate Data Series[ Tijdreeks 1: Huur...][2008-10-20 15:59:35][s0800838]
|     |     |     |     |     Univariate Data Series[ Tijdreeks 3: Euro...][2008-10-20 16:02:06][Jackers Veerle]
|     |     |     |     |     Univariate Data Series[ Tijdreeks 4: Prij...][2008-10-20 16:04:30][Jackers Veerle]
|     |     |     |     |     |    Univariate Data Series[ Extra tijdreeks v...][2008-10-27 17:24:30][Jackers Veer
|     |     |     |     |     Exercise 1.13[ ex 1,13 vraag 1][2008-10-20 18:52:50][ ]
|     |     |     |     |     Exercise 1.13[ Q1 reproductie 1][2008-12-04 18:31:14][Melgers Peter]
|     |     |     |     |     |    Exercise 1.13[ Q1 reproductie 2][2008-12-04 18:34:30][Melgers Peter]
|     |     |     |     |     |    |    Exercise 1.13[ Q1 reproductie 3][2008-12-04 18:36:35][Melgers Peter]
|     |     |     |     |     |    |    Exercise 1.13[ Q1 aantal dagen 365][2008-12-04 18:42:39][Melgers Peter]
|     |     |     |     |     |    |    |    Exercise 1.13[ Q1 aantal dagen 730][2008-12-04 18:44:45][Melgers Pet
|     |     |     |     |     |    |    |    |    Exercise 1.13[ Q1 aantal dagen 1095][2008-12-04 18:47:01][Melge
|     |     |     |     |     |    Exercise 1.13[ Q2 reproductie 1][2008-12-04 18:54:00][Melgers Peter]
```

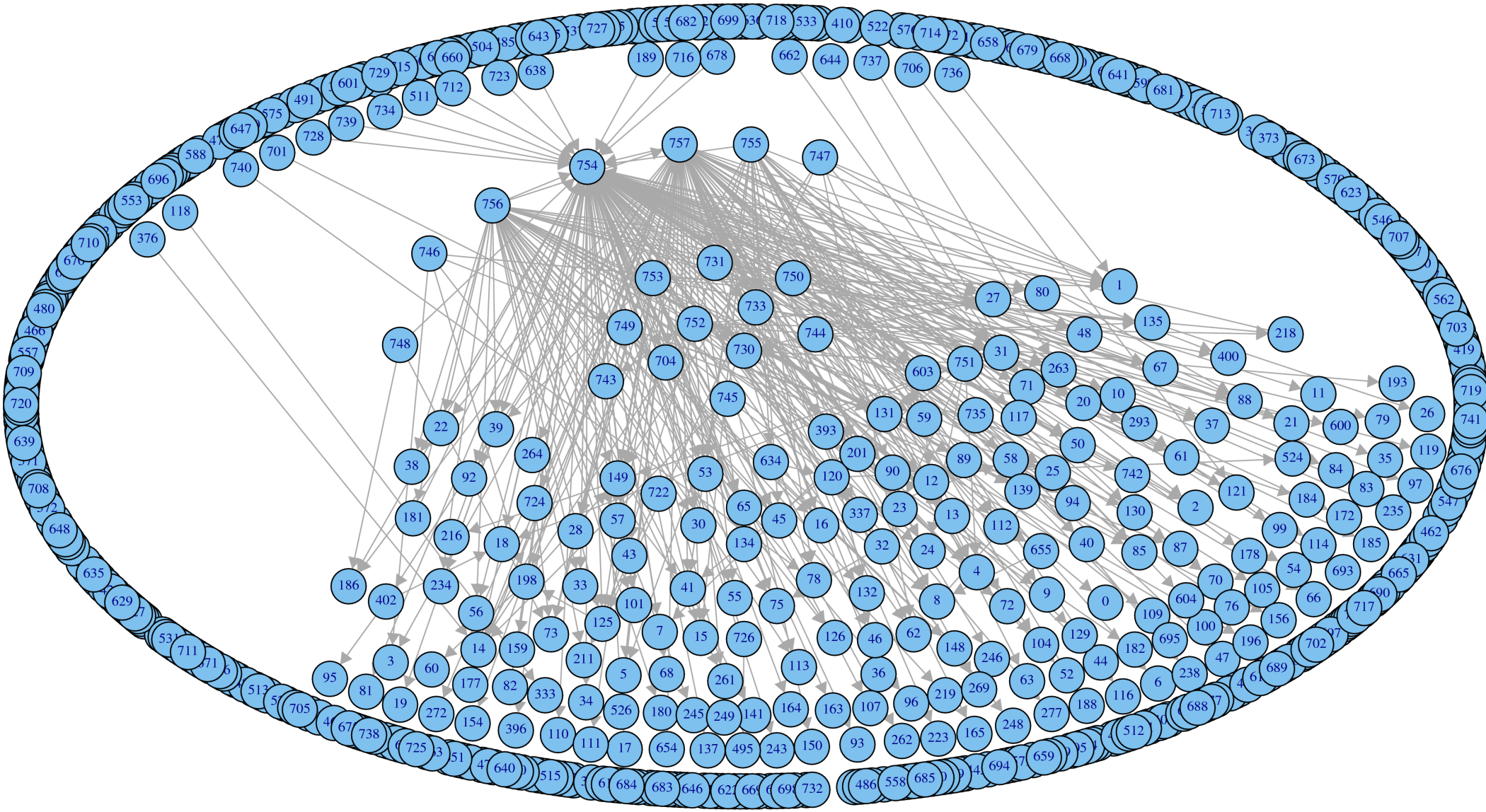
Tracking assignments

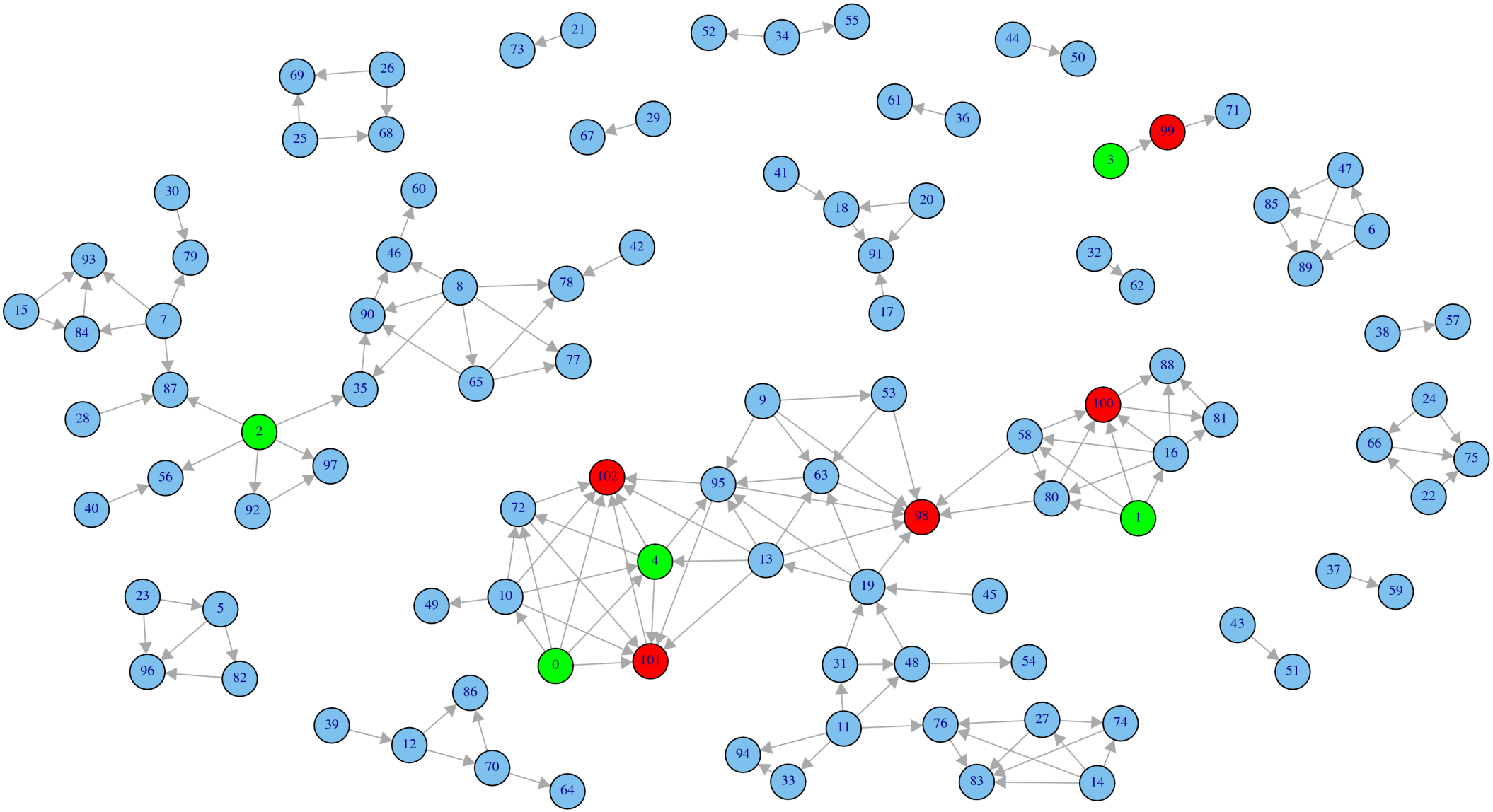
```
> #the assignment deadline was October 14th 2008
> table(substr(mytree$date,1,10))

2008-10-01 2008-10-08 2008-10-09 2008-10-10 2008-10-11 2008-10-12 2008-10-13
         1         15         54         88         50        146        256
2008-10-14 2008-10-15 2008-10-16 2008-10-17 2008-10-18 2008-10-19 2008-10-20
        28          8          5          2         20         11         15
2008-10-27 2008-11-11 2008-11-21 2008-11-30 2008-12-03 2008-12-04 2008-12-05
         1          3         10          5          5         10          3
2008-12-08 2008-12-13
         1          6

> mytab <- table(mytree$user, mytree$forum)
> mytab[78:86, ]

                        -  F
  Tubbax Julie          1  3
  Van den Eynde Evelin  3  0
  Van den Heuvel Ken    9 15
  Van den Heuvel Koen   1  8
  Van Gheluwe Dries     5  3
  Van Ham Ellen         7  5
  Van Isveldt Steffi    1  3
  van Keken Bas         3  0
  Van Opstal Siem      13  1
```




0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102

Summary

- First release of RC (Sep/Oct 2009)
- Workshop @ Applied Statistics (resources available online)
- FreeStatistics.org,
  Wessa.net (computations),
  GoPublish.org (future project on publishing & peer review)

- Questions, Comments & Complaints
  patrick@wessa.net