Dynamic Reports with **R** and LaTeX

Uwe Ziegenhagen

June 30, 2010

What is **R**?

- §language, developed by R. Becker & J. Chambers (Bell Labs)
- commercial version S-Plus
- R: GNU implementation of §in 1992 by R. Ihaka und R. Gentleman
- more than 1 000 packages on CRAN
- >500 project members, "'R Core Team"'
- covers all areas of statistics & data analysis
- platforms: Win32, Linux/Unix, MacOS
- `http://www.r-project.org`

Interfaces for **R**

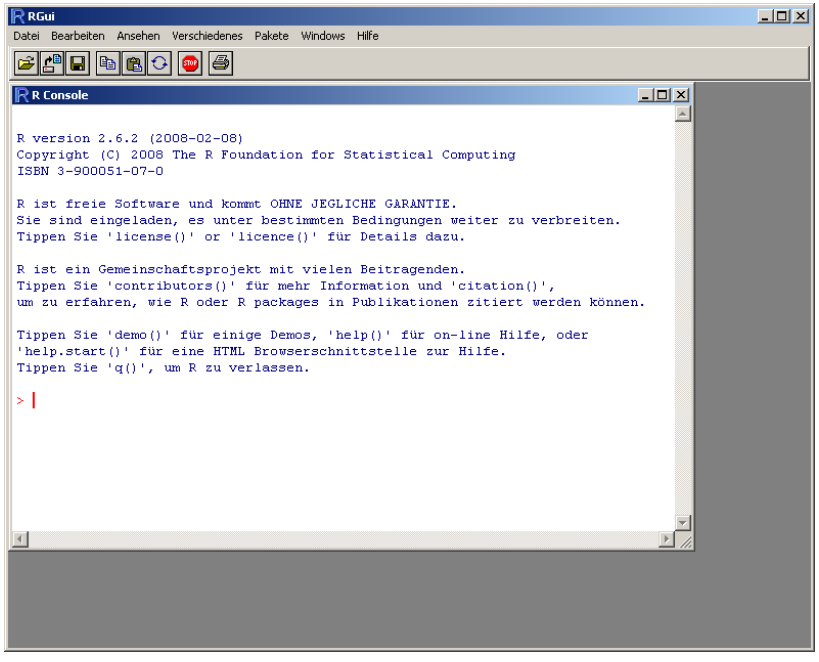


Figure: R on Windows

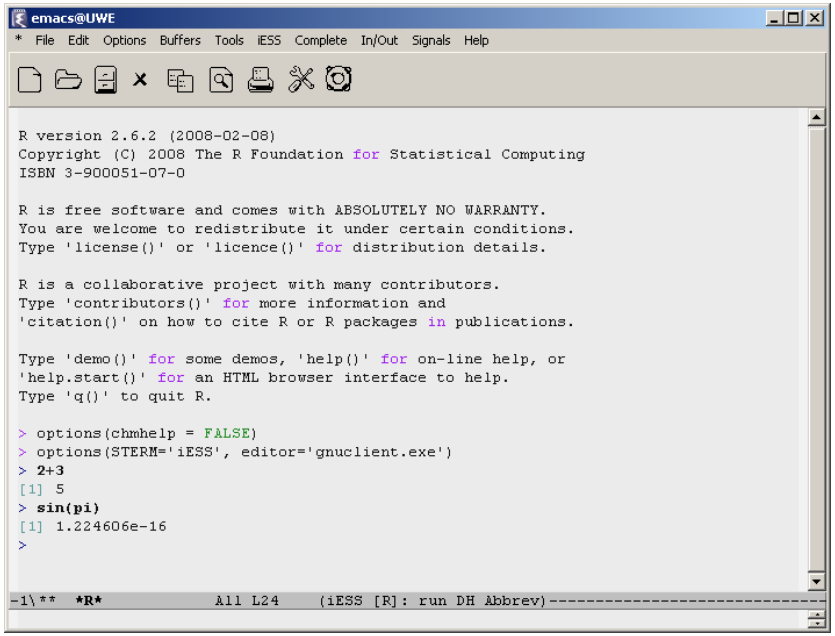


Figure: Emacs with ESS

Various interfaces, among them

- JGR/JRI for Java
- RCOM for COM
- RPY for Python

R as a Calculator

⇒ http://cran.r-project.org/doc/manuals/R-intro.pdf

```
1+2
1*2
1/2
1-2
2^2
sqrt(2)
sin(pi) # cos, tan
trunc(-pi) # -3
round(pi)  # 3
```

R Data Structures

Vectors	vectors of length m, one type
Matrices	$m \times n$ array, one type
Dataframes	List of objects of various types

```
1 a <- 2 # assign scalar
2 d <- c(1,2,3,4) # assign vector
3 a
4 d
```

Variables, vectors and matrices

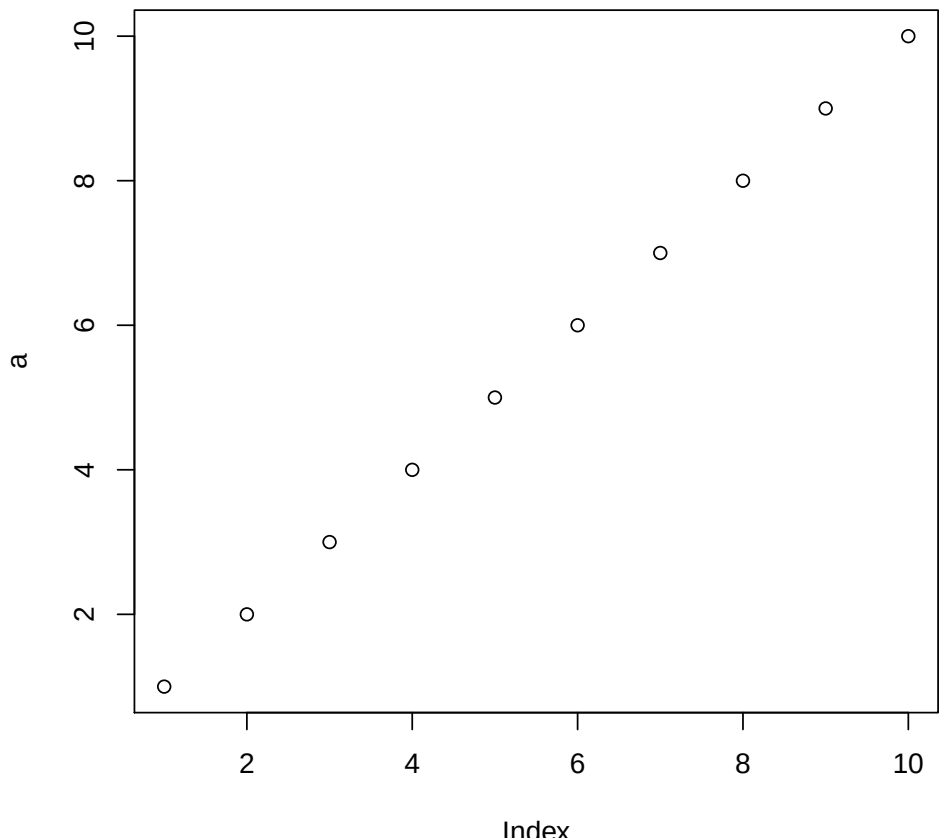
```
1 a = 1:3
2 b = 2:4
3 c(a,b) # [1] 1 2 3 2 3 4
4 seq(1,2,by=0.1) [1] 1.1 1.2 1.3 1.4 1.5 ...
5 rep(1:4,2) # [1] 1 2 3 4 1 2 3 4
```

Simple Linear Models

```
> x<-1:10
> y=rnorm(10)*x
> lm(y~x)

Call:
lm(formula = y ~ x)

Coefficients:
(Intercept)            x
     0.1079       1.0697
```

Simple Graphics

```
1 a<- c(1:10)
2 plot(a)
```



R Graphics Drivers

- x11()/X11(), windows(), quartz() for screen
- postscript(), pdf(), pictex(), xfig(), win.metafile()
- devGTK(), devJava(), devSVG()
- tikzdevice()

PDF Output Example

```
pdf(file = "c:/points.pdf",width = 6, height = 6,
onefile = FALSE, family = "Helvetica",
title = "R Graphics Output", fonts = NULL,
version = "1.4",paper = "special")

a<- c(1:10)
plot(a)

dev.off()
```

tikzdevice

- `http://cran.r-project.org/web/packages/tikzDevice/index.html`
- by Charlie Sharpsteen and Cameron Bracken
- R graphics code is converted to TikZ primitives
- creates embedded or oder self-contained documents
- uses the fonts from the document
- allows math code in captions

TikZ-Device Example

```
1 library(tikzDevice)
2 tikz(file = "c:/test2.tex",standAlone=F)
3 # StandAlone=T
4 plot(1:10)
5
6 dev.off()
```

Generated Code (Excempt)

```
% Created by tikzDevice
\begin{tikzpicture}[x=1pt,y=1pt]
\draw[color=white,opacity=0] (0,0) rectangle (505.89,505.89);
\begin{scope}
\path[clip] ( 49.20, 61.20) rectangle (480.69,456.69);
\definecolor[named]{drawColor}{rgb}{0.56,0.96,0.51}
\definecolor[named]{fillColor}{rgb}{0.13,0.09,0.52}
\definecolor[named]{drawColor}{rgb}{0.00,0.00,0.00}
\draw[color=drawColor,line cap=round,line join=round,fill opacity=0.00,] (
  65.18, 75.85) circle (  2.25);
\draw[color=drawColor,line cap=round,line join=round,fill opacity=0.00,]
  (109.57,116.54) circle (  2.25);
\draw[color=drawColor,line cap=round,line join=round,fill opacity=0.00,]
  (153.97,157.22) circle (  2.25);
\draw[color=drawColor,line cap=round,line join=round,fill opacity=0.00,]
  (198.36,197.91) circle (  2.25);
\draw[color=drawColor,line cap=round,line join=round,fill opacity=0.00,]
  (242.75,238.60) circle (  2.25);
\draw[color=drawColor,line cap=round,line join=round,fill opacity=0.00,]
  (287.14,279.29) circle (  2.25);
\draw[color=drawColor,line cap=round,line join=round,fill opacity=0.00,]
  (331.53,319.98) circle (  2.25);
\draw[color=drawColor,line cap=round,line join=round,fill opacity=0.00,]
  (375.92,360.67) circle (  2.25);
\draw[color=drawColor,line cap=round,line join=round,fill opacity=0.00,]
  (420.32,401.35) circle (  2.25);
\draw[color=drawColor,line cap=round,line join=round,fill opacity=0.00,]
  (464.71,442.04) circle (  2.25);
\end{scope}
\begin{scope}
```

Integrating **R** and LaTeX with Sweave

- developed by Friedrich Leisch, now at LMU Munich
- Part of `utils` package (standard)
- LaTeX-Dokument contains TeX and **R**
- **R** Code is embedded in noweb-Syntax
- noweb: literate programming tool by Norman Ramsey
- TeX-file stored with `.nw`
- in **R**: call Sweave("<filename.nw>") [1]
- use latex/pdflatex on the generated LaTeX-file

[1] `Stangle` extracts the **R** code

Sweave: example1.nw

```
\documentclass{scrartcl}
\title{Sweave}
\author{Uwe Ziegenhagen}
\begin{document}

\maketitle

<<>>=
1+1
@
\end{document}
```

Generated LaTeX-Code

```
\documentclass{scrartcl}

\title{Sweave}
\author{Uwe Ziegenhagen}
\usepackage{Sweave}
\begin{document}

\maketitle

\begin{Schunk}
\begin{Sinput}
> 1 + 1
\end{Sinput}
\begin{Soutput}
[1] 2
\end{Soutput}
\end{Schunk}
\end{document}
```

Result of the Example

Sweave

Uwe Ziegenhagen

February 28, 2010

```
> 1 + 1
```

```
[1] 2
```

Sweave: Options for the «»= Part

echo=false	suppresses **R**-source
results=hide	suppresses results
results=tex	suppresses `verbatim`-output
fig=true	output is **one** graphics ⇒ PDF/EPS files
pdf=true	Create PDF-version
eps=true	Create EPS-version
width= <x>	width in inches
height= <y>	height in inches

- \SweaveOpts<Option> sets options globally.
- with «*name*, opt=. . . »= names code-parts
- accessing these parts with «*name*»

Sweave: \Sexpr<command>

For scalar results: `\Sexpr`

- \Sexpr<command>
- **R**-return value must be string or be convertible to string
- useful for the output of results in the normal text

Sweave Example 2

```
\documentclass[a4paper]{scrartcl}

\begin{document}

<<echo=false,results=hide>>=
data(iris) # load iris data
@

The data has \Sexpr{ncol(iris)} columns and \Sexpr{nrow(iris)} rows.

<<echo=false>>=
summary(iris$Petal.Length)
@

<<echo=false,results=tex>>=
xtable(lm(iris$Sepal.Width~iris$Petal.Length),
caption="Linear model of Sepal.Width and Petal.Length")
@

\begin{center}
\begin{figure}[h]
<<fig=true,echo=false>>=
pch.vec <- c(16,2,3)[iris$Species]
col.vec <- c(16,2,3)[iris$Species]
plot(iris$Sepal.Width,iris$Petal.Length,
col = col.vec,pch=pch.vec)
@
\caption{Plot of iris\$Petal.Length vs. iris\$Sepal.Width}
\end{figure}
\end{center}
\end{document}
```

Sweave Example 2 - Part A

```
1  \documentclass[a4paper]{scrartcl}
2
3  \begin{document}
4
5  <<echo=false,results=hide>>=
6  data(iris) # load iris data
7  @
8
9  The data has \Sexpr{ncol(iris)} columns
10 and \Sexpr{nrow(iris)} rows.
```

Sweave Example 2 - Part B

```
1
2 <<echo=false>>=
3 summary(iris$Petal.Length)
4 @
5
6 <<echo=false,results=tex>>=
7 xtable(lm(iris$Sepal.Width~iris$Petal.Length),
8 caption="Lin. model Sepal.Width and Petal.Length")
9 @
```

Sweave Example 2 - Part C

```
\begin{center}
\begin{figure}[h]
<<fig=true,echo=false>>=
pch.vec <- c(16,2,3)[iris$Species]
col.vec <- c(16,2,3)[iris$Species]
plot(iris$Sepal.Width,iris$Petal.Length,
col = col.vec,pch=pch.vec)
@
\caption{iris\$Petal.Length vs. iris\$Sepal.Width}
\end{figure}
\end{center}
\end{document}
```

Result of example2.nw

Der Datensatz hat 5 Spalten und 150 Zeilen.

```
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
1.000   1.600   4.350   3.758   5.100   6.900
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.4549	0.0761	45.40	0.0000
iris$Petal.Length	-0.1058	0.0183	-5.77	0.0000

Table 1: Lineares Model von Sepal.Width und Petal.Length

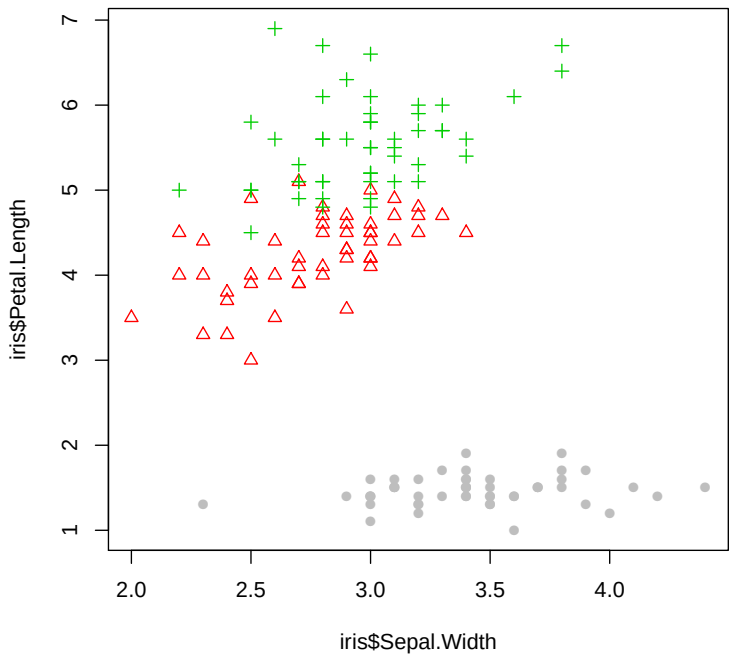


Figure 1: Plot von iris$Petal.Length vs. iris$Sepal.Width

Sweave: Dynamic Example

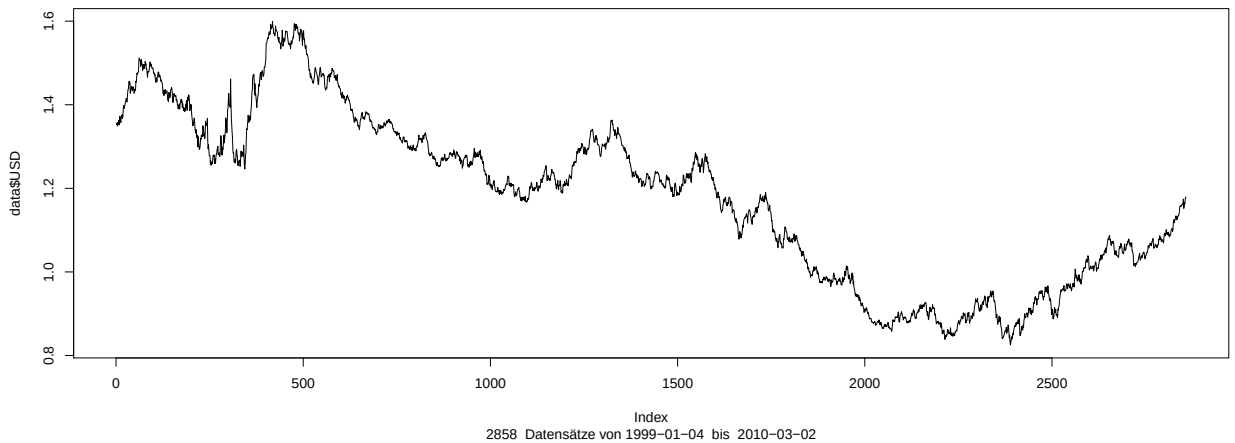
```
\documentclass{scrartcl}
\begin{document}

<<echo=f,results=hide>>=
windows(width = 8, height = 4)
system("wget -O d.zip http://www.ecb.int/stats/eurofxref/eurofxref-hist.zip")
zip.file.extract(file="eurofxref-hist.csv",zip="d.zip",unzip="",dir=getwd())
data= read.csv("eurofxref-hist.csv",sep=",",header=TRUE)
@

The data contains \Sexpr{nrow(data)} prices, the current price (\Sexpr{data$
   Date[1]}) is \Sexpr{data$USD[1]}

\begin{center}
\begin{figure}[h]
<<fig=true,echo=false,width=15,height=6>>=
plot(data$USD,t="l", sub=paste(nrow(data)," rows from ",data$Date[nrow(data)],
   " until ",data$Date[1]),asp=)
@
\end{figure}
\end{center}
\end{document}
```

Sweave: Dynamic Example - Data Retrieval

```
<<echo=f,results=hide>>=
windows(width = 8, height = 4)
system("wget -O d.zip
http://www.ecb.int/stats/eurofxref/eurofxref-hist.
  zip")
zip.file.extract(file="eurofxref-hist.csv",
zip="d.zip",unzip="",dir=getwd())
data= read.csv("eurofxref-hist.csv",sep=",",header=
  TRUE)
@
```

Result

Der Datensatz enthält 2858 Kurse, der aktuelle Kurs (2010-03-02) lautet 1.3548



Literature on **R**

R Core Team
*An Introduction to **R***
http://cran.r-project.org/doc/manuals/R-intro.pdf

Uwe Ligges
*Programmieren mit **R*** (in German)
Springer-Verlag

Michael J. Crawley
*Statistics – An Introduction using **R***
Wiley

John Maindonald und John Brown
*Data Analysis and Graphics Using **R***
Cambridge

Peter Dalgaard
*Introductory Statistics with **R***
Springer-Verlag