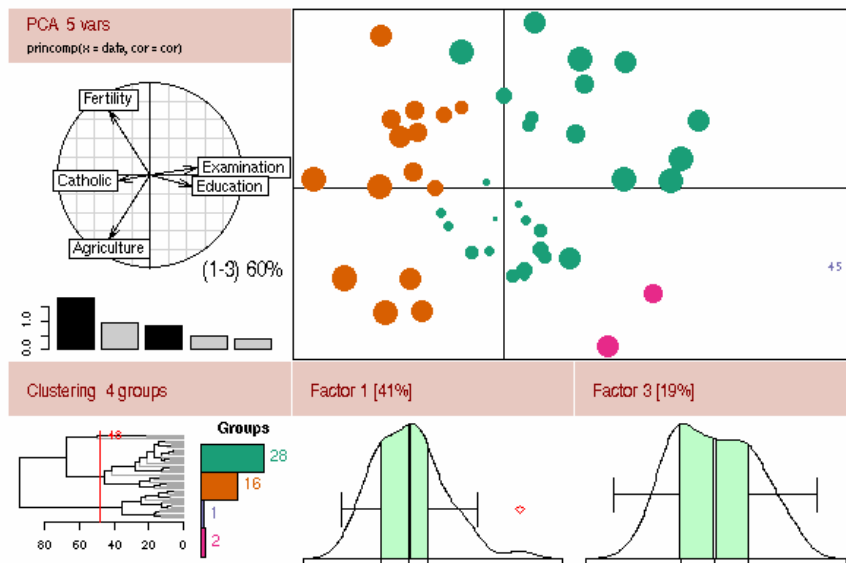


you experienced?

Einführung in Statistik und Grafik mit dem freien Statistikprogramm „R“
und der grafischen Benutzeroberfläche „R-Commander“ (Rcmdr)



Dipl.-Psych. Wilmar Igl

Institut für Psychotherapie u.
Medizinische Psychologie
AB Rehabilitationswissenschaften
Universität Würzburg

I. Einleitung

S, S Plus und R

- **S** wurde als spezielle „Sprache“ für statistische Berechnungen von John Chambers et al. in den AT&T Bell Labs entwickelt (seit 1976).
- **S Plus** (Insightful Inc.) ist ein kommerzielles Programm für Statistik und Visualisierung von Daten basierend auf S.
- **R** (auch „GNU S“) ist ein freies Programm für Statistik und Visualisierung von Daten basierend auf S.
- **Ross Ihaka** and **Robert Gentleman** (University of Auckland, New Zealand) haben R entwickelt.
- R unterliegt der GNU Lizenz und ist daher **Open Source**.
- **Programme in S, S Plus und R sind meist kompatibel.**

Vorteile von R

Anwender

- keine Kosten
- Qualität
- umfangreiche Dokumentation, Support

Statistiker

- Aktualität
- Funktionsvielfalt
- Erweiterbarkeit (u. a. mit grafischen Benutzeroberflächen)

Entwickler

- Objektorientiertheit
- Open Source
- Unabhängigkeit von Betriebssystem

Nachteile von R

- R basiert auf einer **komplexen Sprache**. Die Mächtigkeit von R kann nur mit Hilfe der Sprache vollständig genutzt werden.
- R ist **weniger benutzerfreundlich** als kommerzielle Software.

II. Steuerung von R mit textbasierten Befehlen

Konsole („Taschenrechner“)

- Eingabe von Befehlen auf der Kommando-Zeile (>....)

- Arithmetische Berechnungen

Beispiel:

```
> (3+4)*2  
[1] 14
```

- Statistische Berechnungen

Beispiel:

```
> x<-c(4,6,7,9)  
> mean(x)  
[1] 6.5
```

- Numerische Funktionen

Beispiel:

```
> sqrt(2)  
[1] 1.414214
```

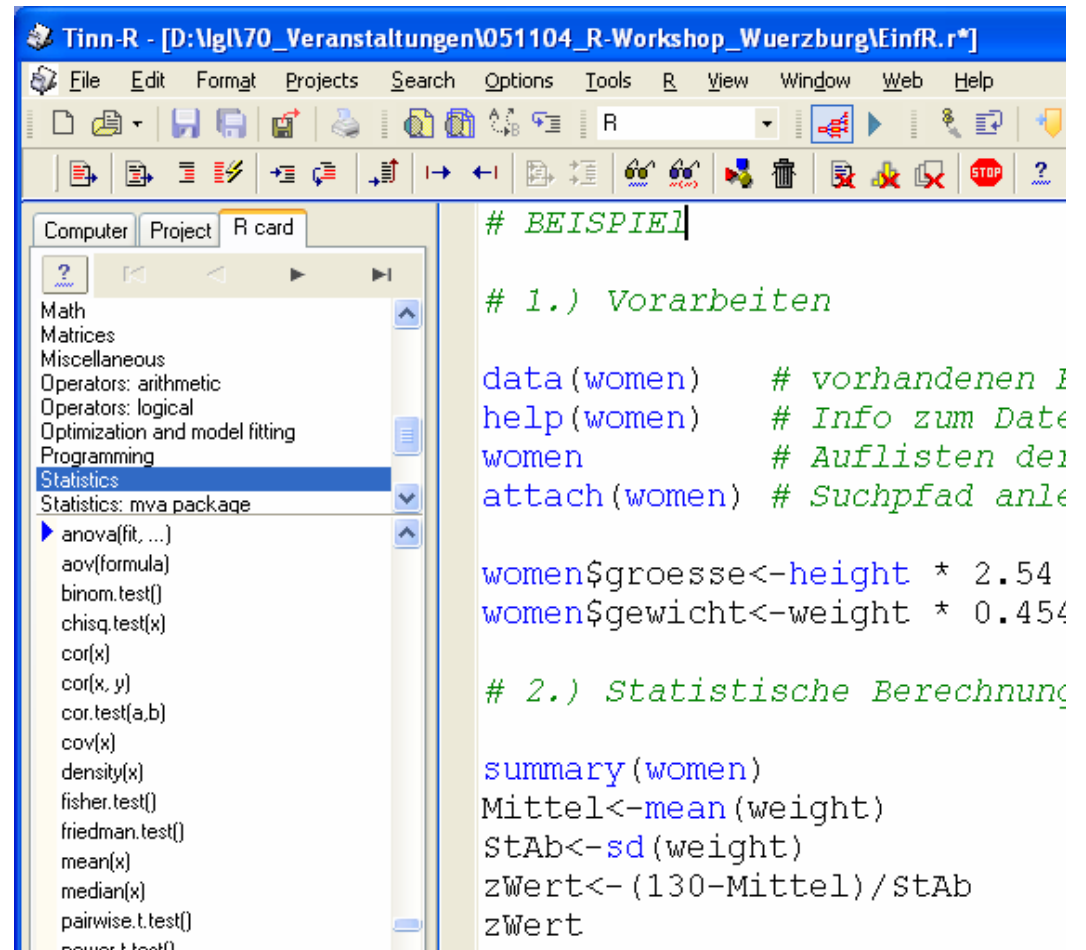
```
R R Console  
Datei Bearbeiten Verschiedenes Pakete Hilfe  
  
R : Copyright 2005, The R Found  
Version 2.2.0 (2005-10-06 r357  
ISBN 3-900051-07-0  
  
R ist freie Software und kommt  
Sie sind eingeladen, es unter b  
Tippen Sie 'license()' or 'lice  
  
R ist ein Gemeinschaftsprojekt  
Tippen Sie 'contributors()' für  
um zu erfahren, wie R oder R pa  
  
Tippen Sie 'demo()' für einige  
'help.start()' für eine HTML Br  
Tippen Sie 'q()', um R zu verla  
  
> (3+4)*2  
[1] 14  
> x<-c(4,6,7,9)  
> mean(x)  
[1] 6.5  
> sqrt(2)  
[1] 1.414214  
> █
```

Editor (hier: Tinn-R)

- **Editor** = einfaches Textprogramm zur Erstellung eines Auswertungsprogramms
- **Auswertungsprogramm** = Reihenfolge von statistischen Befehlen und Kommentaren (#)
- Auswertungsprogramm kann gespeichert, wieder aufgerufen und ausgeführt oder verändert werden

Weitere Eigenschaften:

- Syntaxhighlighting
- Katalog von Befehlen



```
Tinn-R - [D:\Mgl\70_Veranaltungen\051104_R-Workshop_Wuerzburg\EinfR.r*]
File Edit Format Projects Search Options Tools R View Window Web Help
# BEISPIEL
# 1.) Vorarbeiten
data(women) # vorhandenen 1
help(women) # Info zum Date
women # Auflisten der
attach(women) # Suchpfad anle
women$groesse<-height * 2.54
women$gewicht<-weight * 0.454
# 2.) Statistische Berechnung
summary(women)
Mittel<-mean(weight)
StAb<-sd(weight)
zWert<-(130-Mittel)/StAb
zWert
```


III. Steuerung von R mit der grafischen Benutzeroberfläche R-Commander

Rcmdr: Benutzeroberfläche

Menü-Leiste



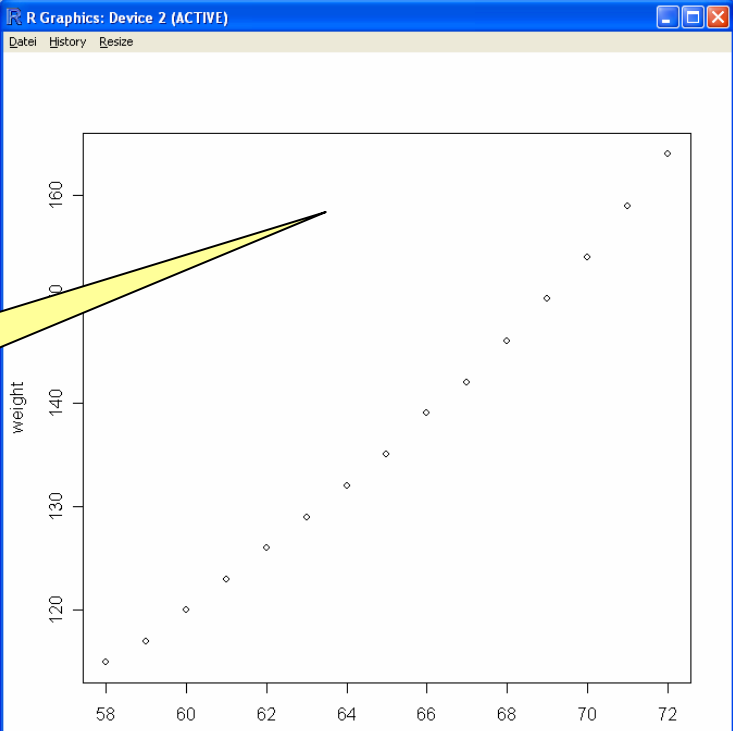
Eingabefenster für Befehle

```
Script Window  
Mittel<-mean(weight)  
StAb<-sd(weight)  
zWert<- (130-mittel)/stab  
zWert  
  
# 3.) Graphische Darstellung  
  
plot(height,weight)
```

Ausgabefenster für stat. Ergebnisse

```
Output Window  
>  
> summary(women)  
      height      weight  
Min.   :58.0   Min.   :115.0  
1st Qu.:61.5   1st Qu.:124.5  
Median :65.0   Median :135.0  
Mean   :65.0   Mean   :136.7  
3rd Qu.:68.5   3rd Qu.:148.0  
Max.   :72.0   Max.   :164.0  
> Mittel<-mean(weight)  
> StAb<-sd(weight)  
> zWert<- (130-mittel)/stab  
> zWert  
[1] -0.4344452  
  
> # 3.) Graphi  
>  
> plot(height,weight)
```

Ausgabefenster für Diagramme



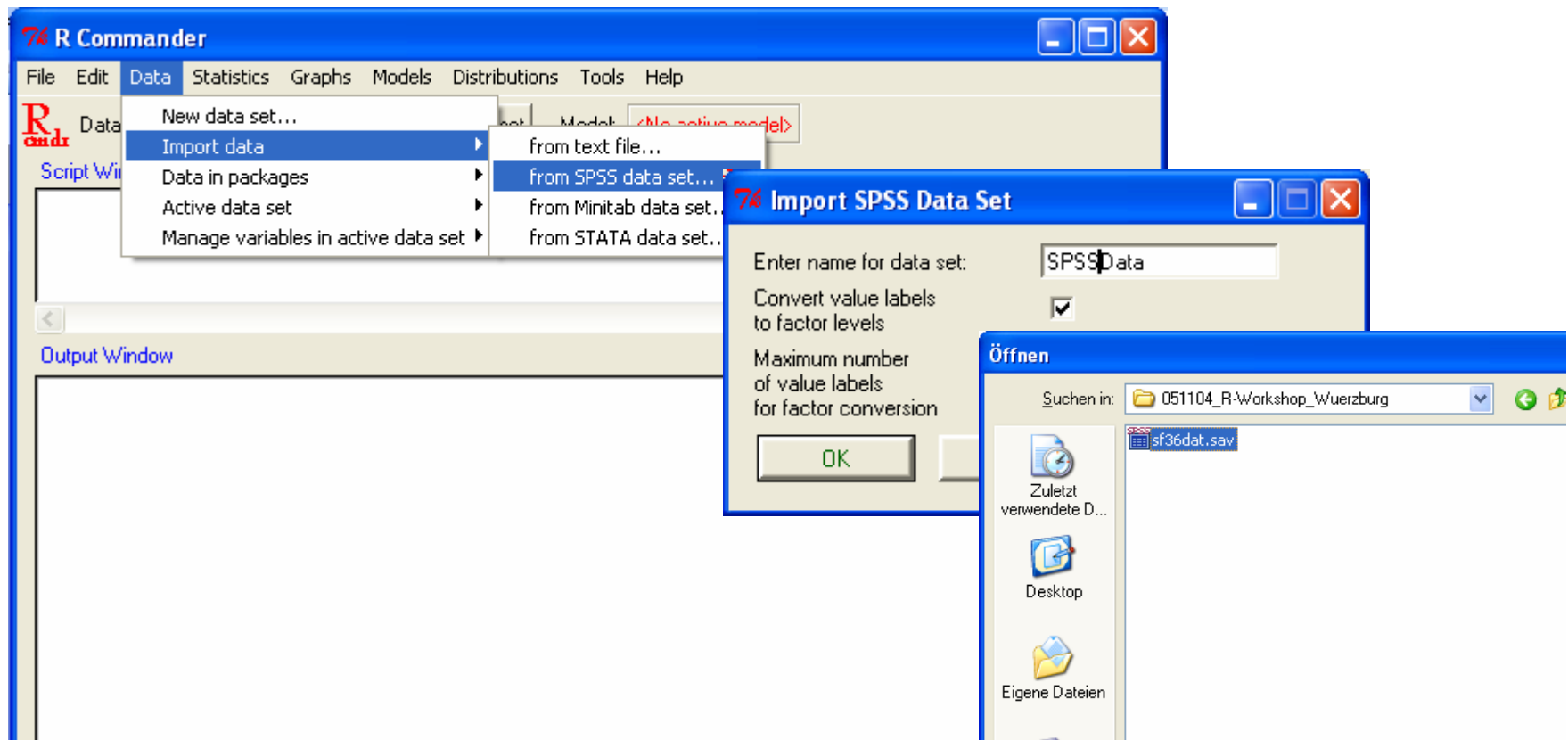
Beispieldaten

- Datensatz „**Women**“
 - Beispieldatensatz in R
 - N=15 Frauen
 - Variablen: height=Körpergröße [inch]; weight=Körpergewicht [pounds];
groesse=Körpergröße [cm]; gewicht=Körpergewicht [kg],
gruppe=Gruppeneinteilung
- Datensatz „**Subjektive Lebensqualität**“
 - Lebensqualitätsdaten von Reha-Patienten 3-4 Wochen vor Reha-Beginn
 - Fragebogen SF-36
 - N=1145 Patienten (Orthopädie, Rheumatologie, Kardiologie)
 - Variablen: Code, Alter, Geschlecht, Indikation, 36 SF-36-Items

III. A) Rcmdr: Statistische Auswertungen

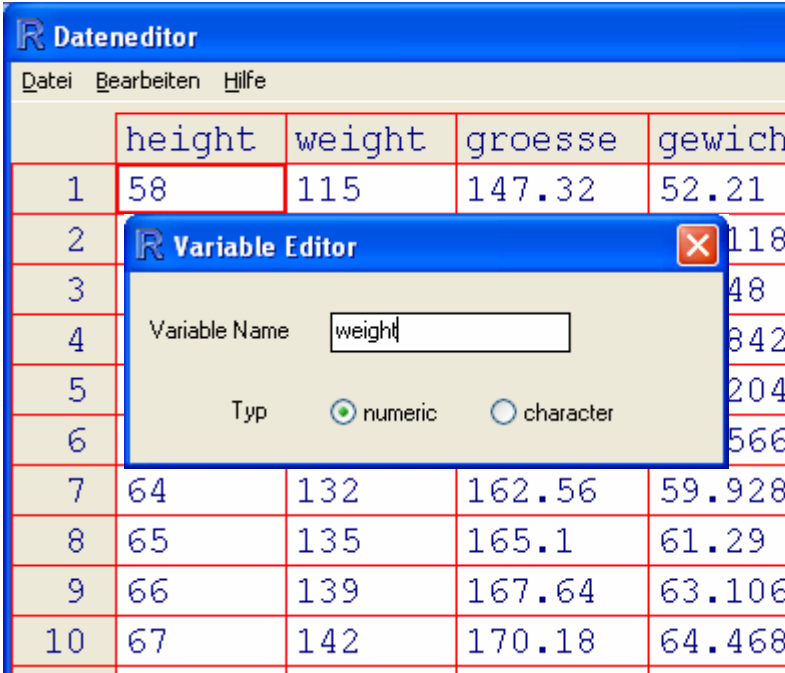
Einlesen von Dateien

- Rcmdr-Menü => Data => Import Data
=> from SPSS data set
=> from text file
=> ...

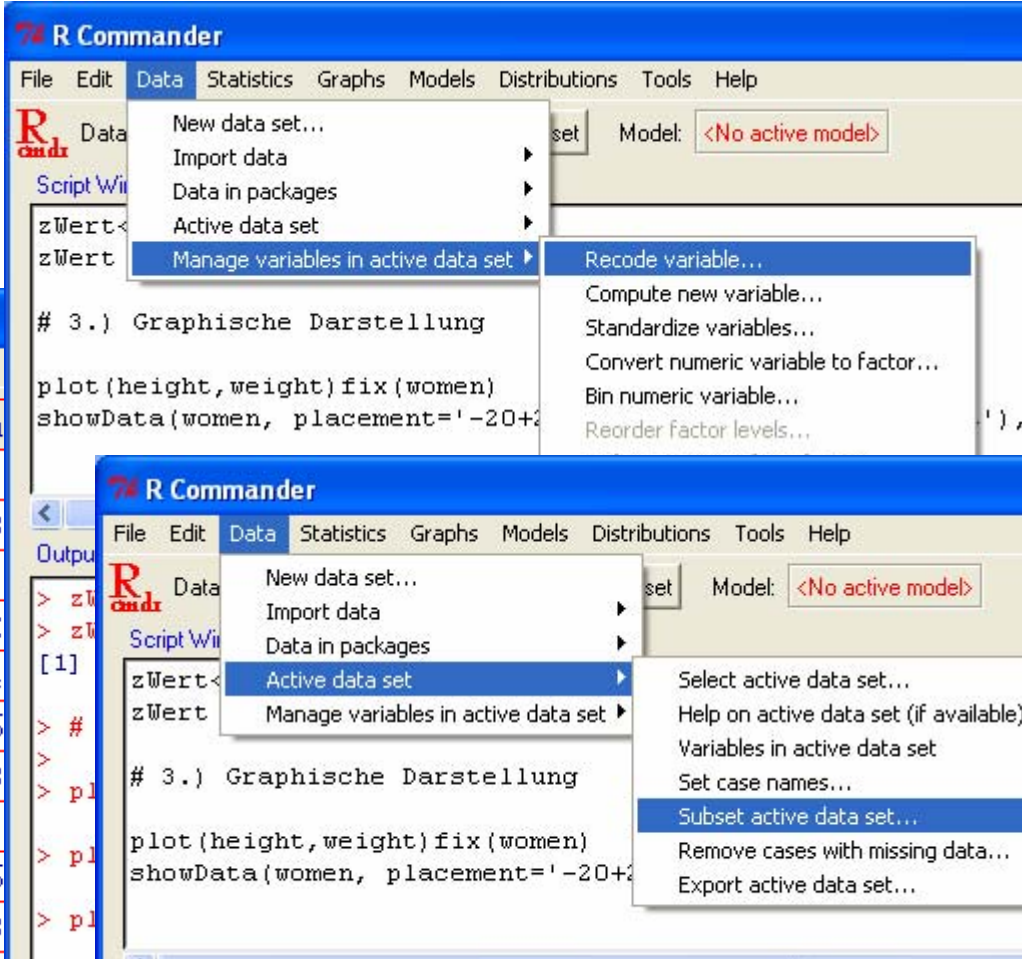


Bearbeiten von Daten

- **Dateneditor**: Rcmdr-Menü => Edit data set
- Unterscheidung zwischen **numerischen** und **String-Variablen**
- **Berechnen neuer Variablen**
- **Umkodieren** von Variablen
- Auswahl von **Substichproben**



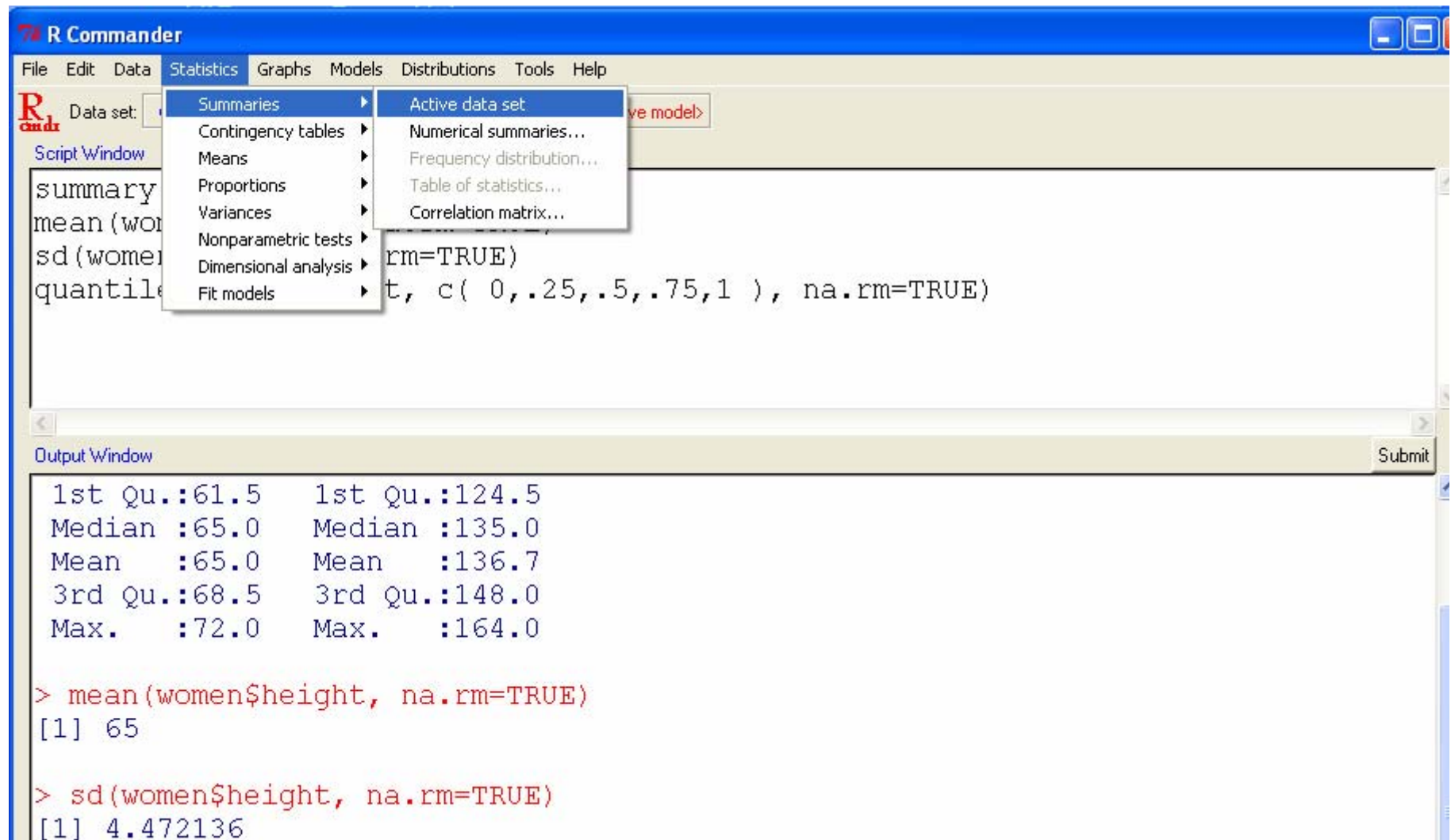
	height	weight	groesse	gewicht
1	58	115	147.32	52.21
2				118
3				48
4				842
5				204
6				566
7	64	132	162.56	59.928
8	65	135	165.1	61.29
9	66	139	167.64	63.106
10	67	142	170.18	64.468



```
# 3.) Graphische Darstellung
plot(height, weight) fix(women)
showData(women, placement='-20+2')
```

Deskriptive Statistiken

- Rcmdr-Menü => Statistics => Summaries => Active data set



The screenshot shows the R Commander interface. The 'Statistics' menu is open, and the path 'Summaries' -> 'Active data set' is highlighted. The script window contains the following R code:

```
summary  
mean(women$height, na.rm=TRUE)  
sd(women$height, na.rm=TRUE)  
quantile(women$height, c(0,.25,.5,.75,1), na.rm=TRUE)
```

The output window displays the results of the first three commands:

```
1st Qu.:61.5    1st Qu.:124.5  
Median :65.0    Median :135.0  
Mean   :65.0    Mean   :136.7  
3rd Qu.:68.5    3rd Qu.:148.0  
Max.   :72.0    Max.   :164.0  
  
> mean(women$height, na.rm=TRUE)  
[1] 65  
  
> sd(women$height, na.rm=TRUE)  
[1] 4.472136
```

Korrelationen und Faktorenanalyse

- Rcmdr-Menü => Statistics => Summaries => Correlation matrix
- Rcmdr-Menü => Statistics => Principal-components analysis

The screenshot displays the R Commander interface with several windows and a menu path. The main window shows a correlation matrix for variables SFIO3A through SFIO13A. A secondary window shows the 'Statistics' menu with 'Dimensional analysis' selected, and a sub-menu with 'Principal-components analysis...' highlighted. A third window shows a bar chart of variances for components 1 through 9.

Correlation Matrix (Output Window):

	SFIO9A	SFIO10A
SFIO3A	0.3361204	0.3137215
SFIO4A	0.4215202	0.4092837
SFIO5A	0.4089524	0.4177102
SFIO6A	0.5440894	0.5674543
SFIO7A	0.5108206	0.5269779
SFIO8A	0.3841911	0.4006264
SFIO9A	1.0000000	0.8089935
SFIO10A	0.8089935	1.0000000
SFIO11A	0.6560478	0.7456000
SFIO12A	0.3981472	0.4271169
SFIO13A	0.2378256	0.2632097

Principal Components Analysis (Script Window):

```
cor (sf36dat[,c ("SFIO3A", "SFIO4A", "SFIO5A", "SFIO6A", "SFIO7A", "SFIO8A", "SFIO9A", "SFIO10A", "SFIO11A", "SFIO12A", "SFIO13A")])  
  
> .PC <- factanal(x = ~SFIO1A + SFIO2A + SFIO3A + SFIO4A + SFIO5A + SFIO6A + SFIO7A + SFIO8A + SFIO9A + SFIO10A + SFIO11A + SFIO12A + SFIO13A, nfactors = 9, rotate = "none", weights = "none", variances = c(10, 4, 2, 2, 1, 1, 1, 1, 1))  
  
> .FA  
  
Call:  
factanal(x = ~SFIO1A + SFIO2A + SFIO3A + SFIO4A + SFIO5A + SFIO6A + SFIO7A + SFIO8A + SFIO9A + SFIO10A + SFIO11A + SFIO12A + SFIO13A, nfactors = 9, rotate = "none", weights = "none", variances = c(10, 4, 2, 2, 1, 1, 1, 1, 1))
```

Variance Plot (.PC):

Component	Variance
Comp.1	10
Comp.2	4
Comp.3	2
Comp.4	2
Comp.5	1
Comp.6	1
Comp.7	1
Comp.8	1
Comp.9	1

Einfache Inferenzstatistik (t-Test und Co.)

- Rcmdr-Menü => Statistics => Means => Independent samples t-test
- Rcmdr-Menü => Statistics => Nonparam. tests => Two sample Wilcoxon -Test

The screenshot displays the R Commander interface. The main window shows the output of a t-test:

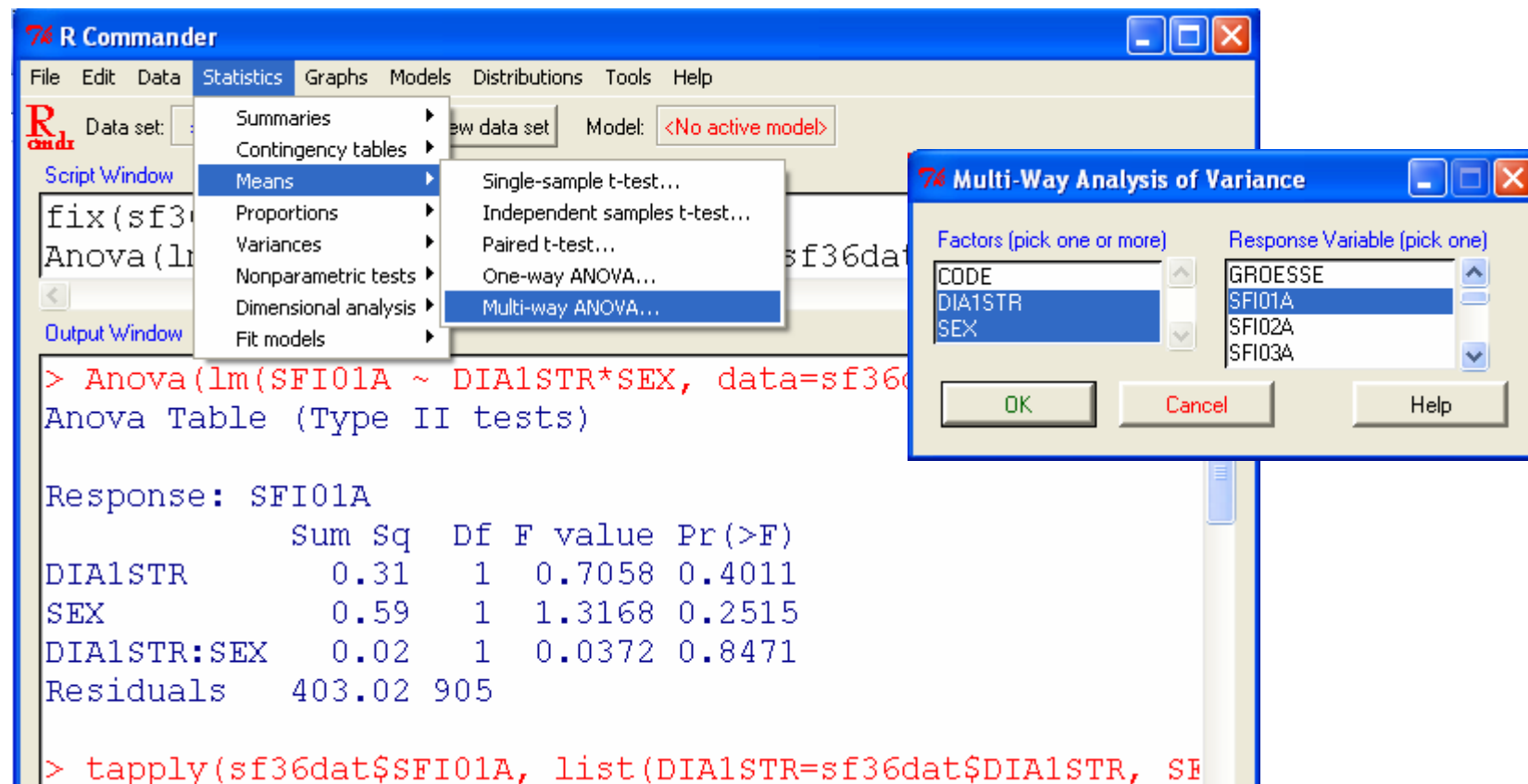
```
data: gewicht by gruppe  
t = -1.1554, df = 12.825, p-value = 0.2  
alternative hypothesis: true difference  
95 percent confidence interval:  
-11.922760 3.621046  
sample estimates:  
mean in group 1 mean in group 2  
59.86314 64.01400
```

Two dialog boxes are overlaid on the main window:

- Independent Samples t-Test:** Shows 'gruppe' as the factor and 'gewicht', 'groesse', 'height', and 'weight' as response variables. The confidence level is set to 0.95, and 'Assume equal variances?' is set to 'Yes'.
- Two-Sample Wilcoxon Test:** Shows 'gruppe' as the factor and the same response variables. The difference is set to '1 - 2', and the 'Type of Test' is set to 'Default'.

Varianzanalyse

- Rcmdr-Menü => Statistics => Means => Multi-way ANOVA
- **Beispiel:** „Gesundheitsbezogene Lebensqualität“ (SF-36, 1. Item)
UV: „Unterschiede abhängig von Geschlecht und/oder Indikation?“
AV: „Wie würden Sie Ihren Gesundheitszustand im Allgemeinen beschreiben?“



The screenshot shows the R Commander interface. The 'Statistics' menu is open, and 'Multi-way ANOVA...' is selected. The 'Multi-Way Analysis of Variance' dialog box is open, showing 'CODE', 'DIA1STR', and 'SEX' as factors, and 'GROESSE', 'SFI01A', 'SFI02A', and 'SFI03A' as response variables. The 'OK' button is highlighted.

```
> Anova(lm(SFI01A ~ DIA1STR*SEX, data=sf36dat))
Anova Table (Type II tests)

Response: SFI01A
          Sum Sq Df F value Pr(>F)
DIA1STR    0.31  1  0.7058 0.4011
SEX        0.59  1  1.3168 0.2515
DIA1STR:SEX 0.02  1  0.0372 0.8471
Residuals 403.02 905

> tapply(sf36dat$SFI01A, list(DIA1STR=sf36dat$DIA1STR, SE
```

Diagnostik

- Rcmdr-Menü => Statistics => Dimensional Analysis => Scale Reliability
- **Beispiel:** Interne Konsistenz der SF-36, Skala „Körperl. Funktionsfähigkeit“

The screenshot shows the R Commander interface. The 'Statistics' menu is open, and 'Scale reliability...' is selected. The 'Scale Reliability' dialog box is open, showing a list of variables: SF10A, SF11A, SF12A, and SF13A. The 'Output Window' displays the results of the reliability analysis.

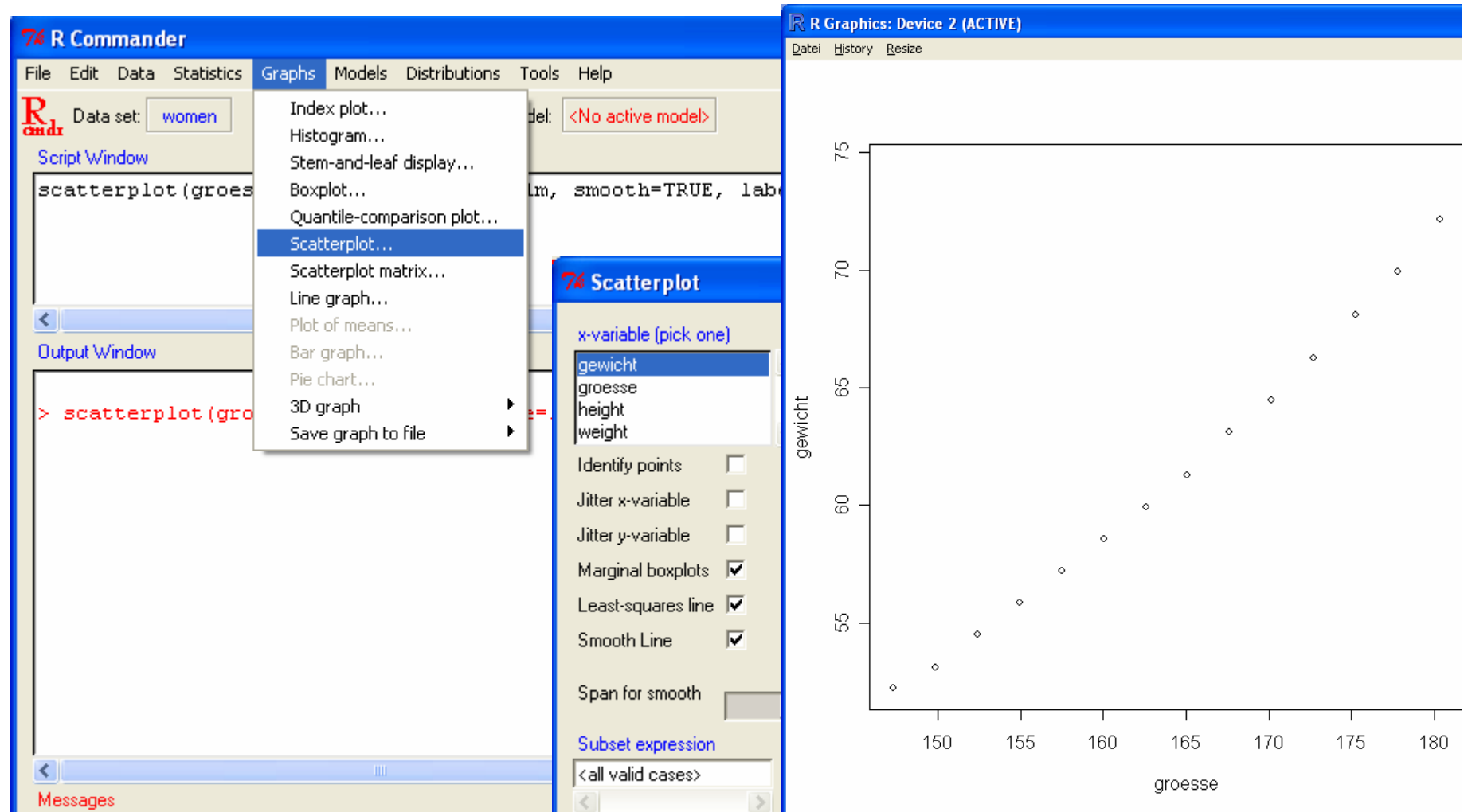
```
> reliability(cov(sf36dat[,c("SF10A", "SF11A", "SF12A", "SF13A", "SF14A", "SF15A", "SF16A", "SF17A", "SF18A", "SF19A", "SF20A", "SF21A", "SF22A", "SF23A", "SF24A", "SF25A", "SF26A", "SF27A", "SF28A", "SF29A", "SF30A", "SF31A", "SF32A", "SF33A", "SF34A", "SF35A", "SF36A")], method="alpha")
Alpha reliability = 0.8851
Standardized alpha = 0.8834

Reliability deleting each item in turn:
      Alpha Std.Alpha r(item, total)
SF10A 0.8814    0.8800    0.4909
SF11A 0.8731    0.8706    0.6308
SF12A 0.8753    0.8729    0.5977
SF13A 0.8713    0.8693    0.6581
SF14A 0.8721    0.8695    0.6570
SF15A 0.8788    0.8764    0.5431
SF16A 0.8677    0.8666    0.7069
```

III. B) Rcmdr: Visualisierung von Daten

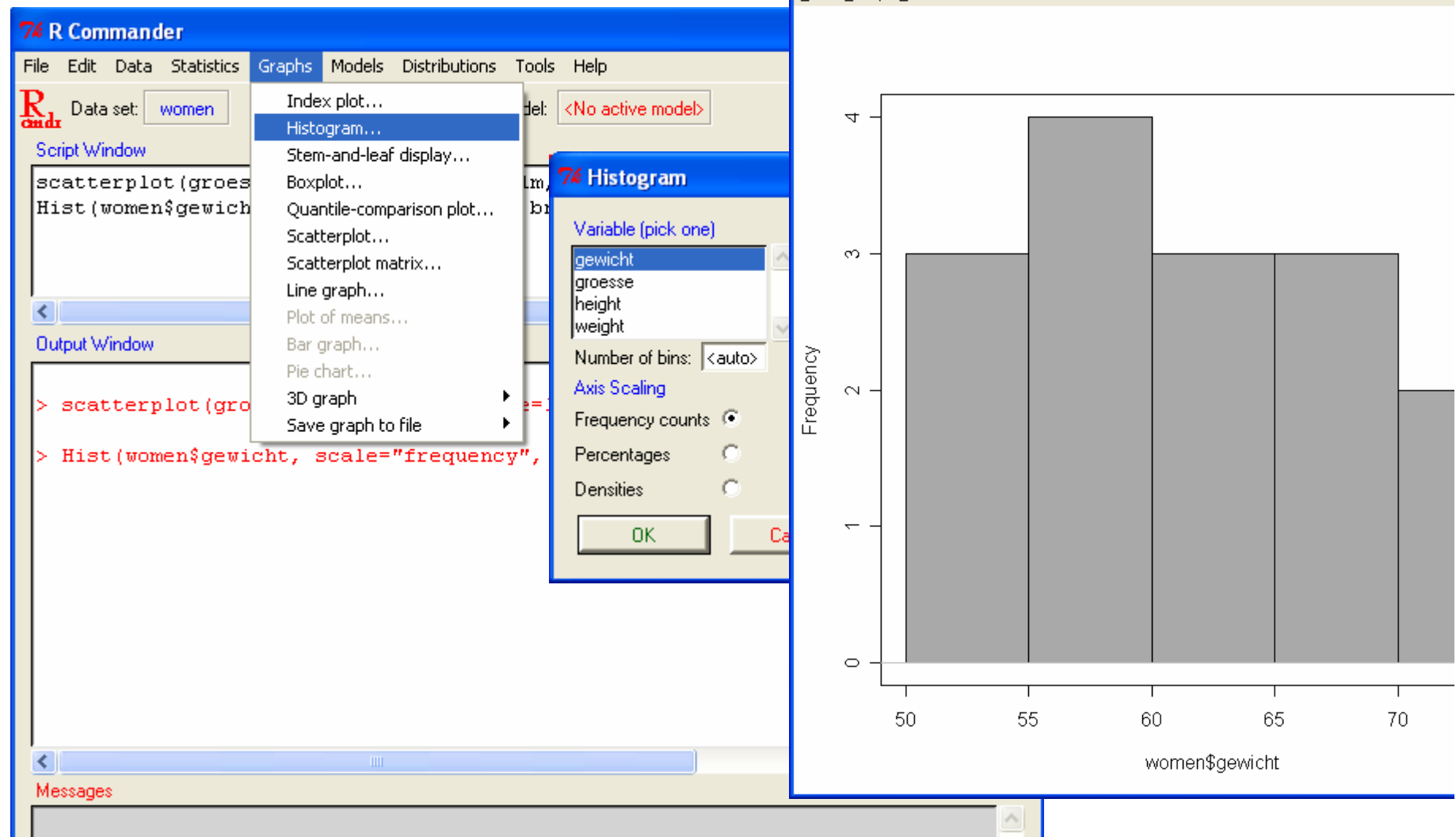
Streudiagramme

- Rcmdr-Menü => Graphs => Scatterplot



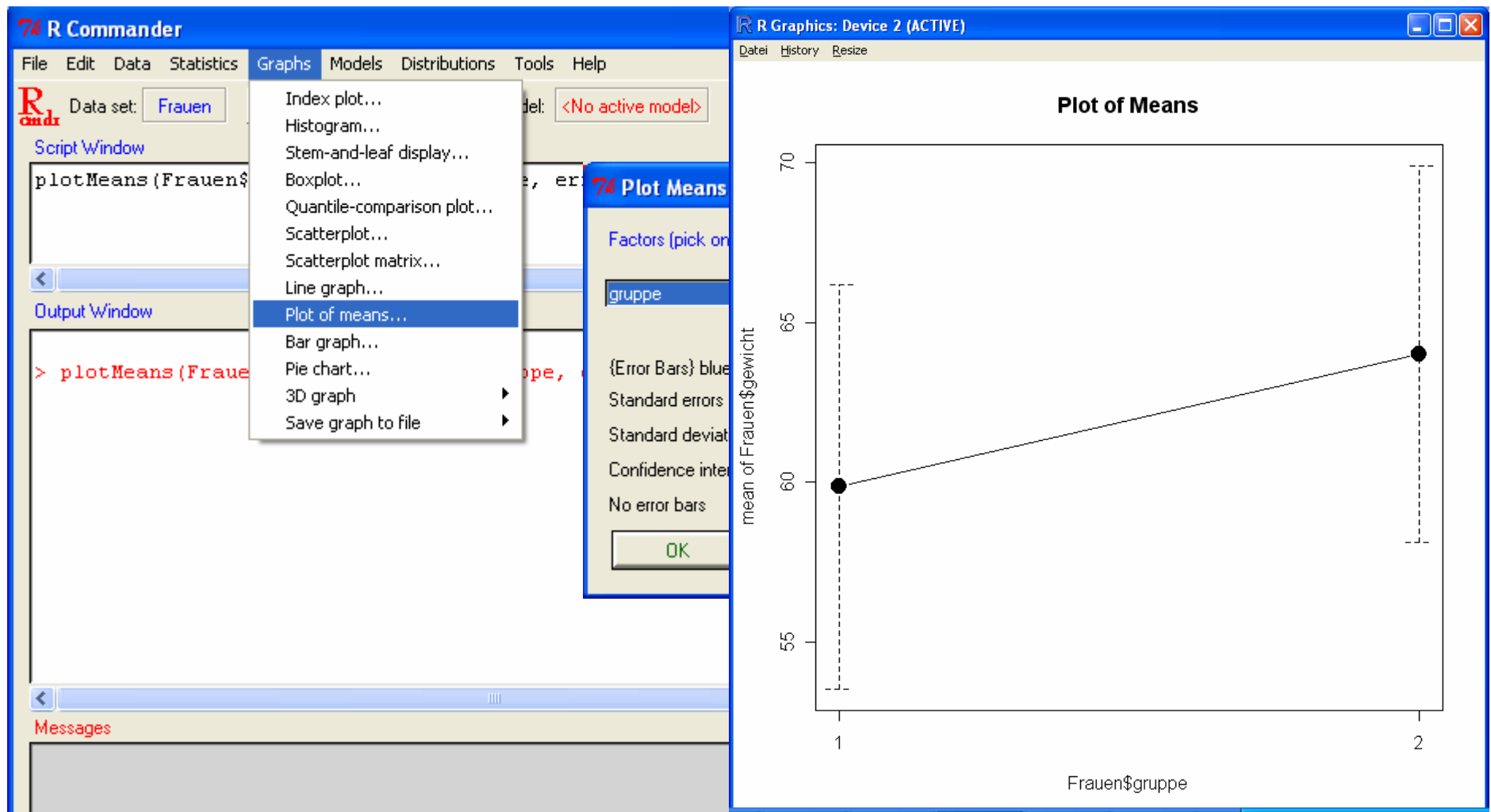
Histogramme

- Rcmdr => Graphs => Histogramm



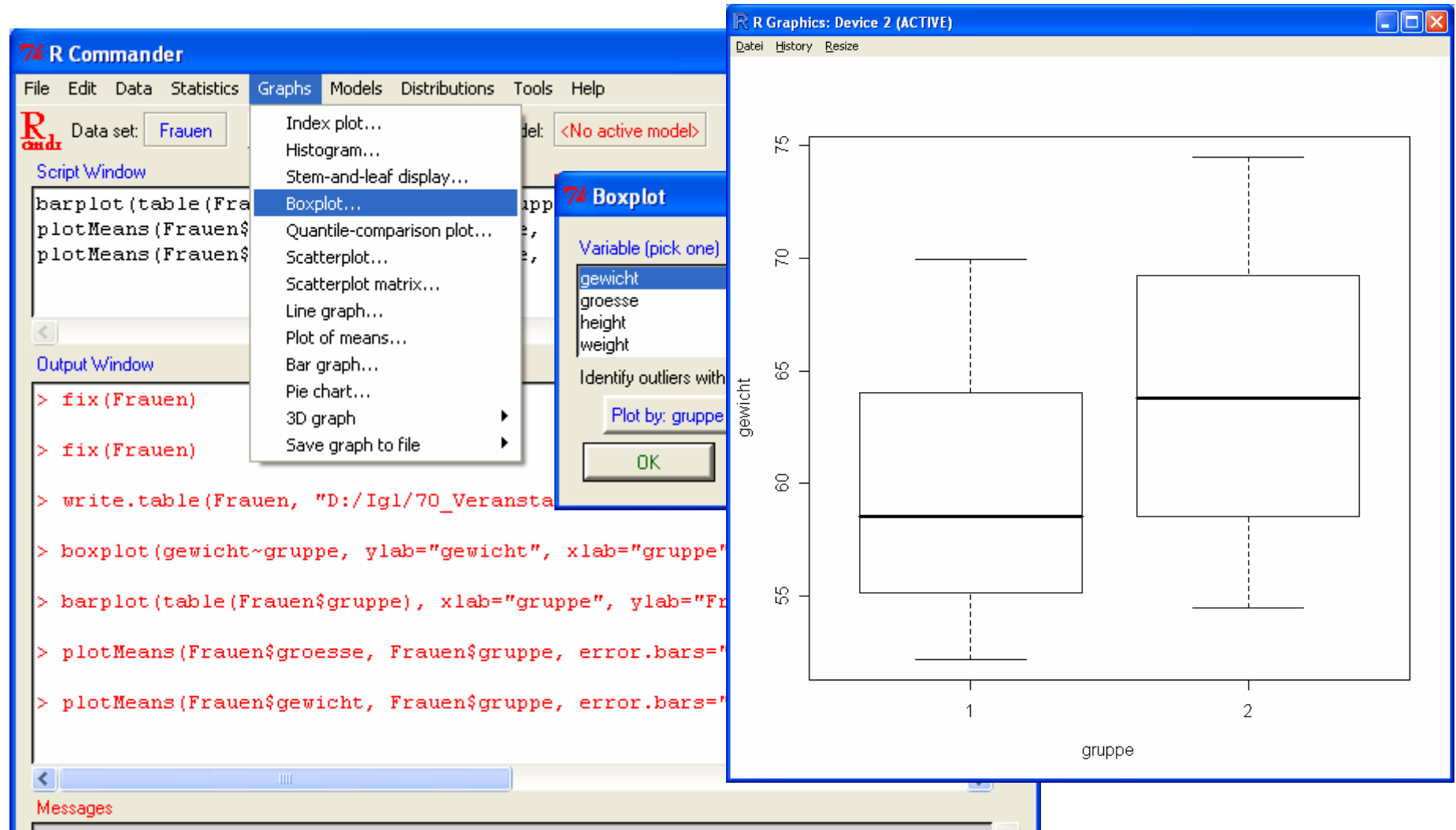
Fehlerbalkendiagramm

- Rcmdr-Menü => Graphs => Plot of Means



Box-and-Whisker-Diagramme

- Rcmdr-Menü => Graphs => Boxplot



IV. Installation

R-Project Homepage

- <http://www.r-project.org> bzw. <http://cran.r-mirror.de/>

The screenshot shows the R Project for Statistical Computing homepage in Mozilla Firefox. The browser window title is "The R Project for Statistical Computing - Mozilla Firefox". The address bar shows the URL <http://www.r-project.org/>. The page content includes the R logo, a navigation menu on the left, and several data visualization plots.

Navigation Menu (Left):

- About R
 - [What is R?](#)
 - [Contributors](#)
 - [Screenshots](#)
 - [What's new?](#)
- Download
 - [CRAN](#)
- R Project
 - [Foundation](#)
 - [Members & Donors](#)
 - [Mailing Lists](#)
 - [Bug Tracking](#)
 - [Developer Page](#)
 - [Search](#)
- Documentation
 - [Manuals](#)
 - [FAQs](#)
 - [Newsletter](#)
 - [Books](#)
 - [Other](#)

Main Content:

The R Project for Statistical Computing

PCA 5 vars
`princomp(x = data, cor = cor)`

Clustering 4 groups

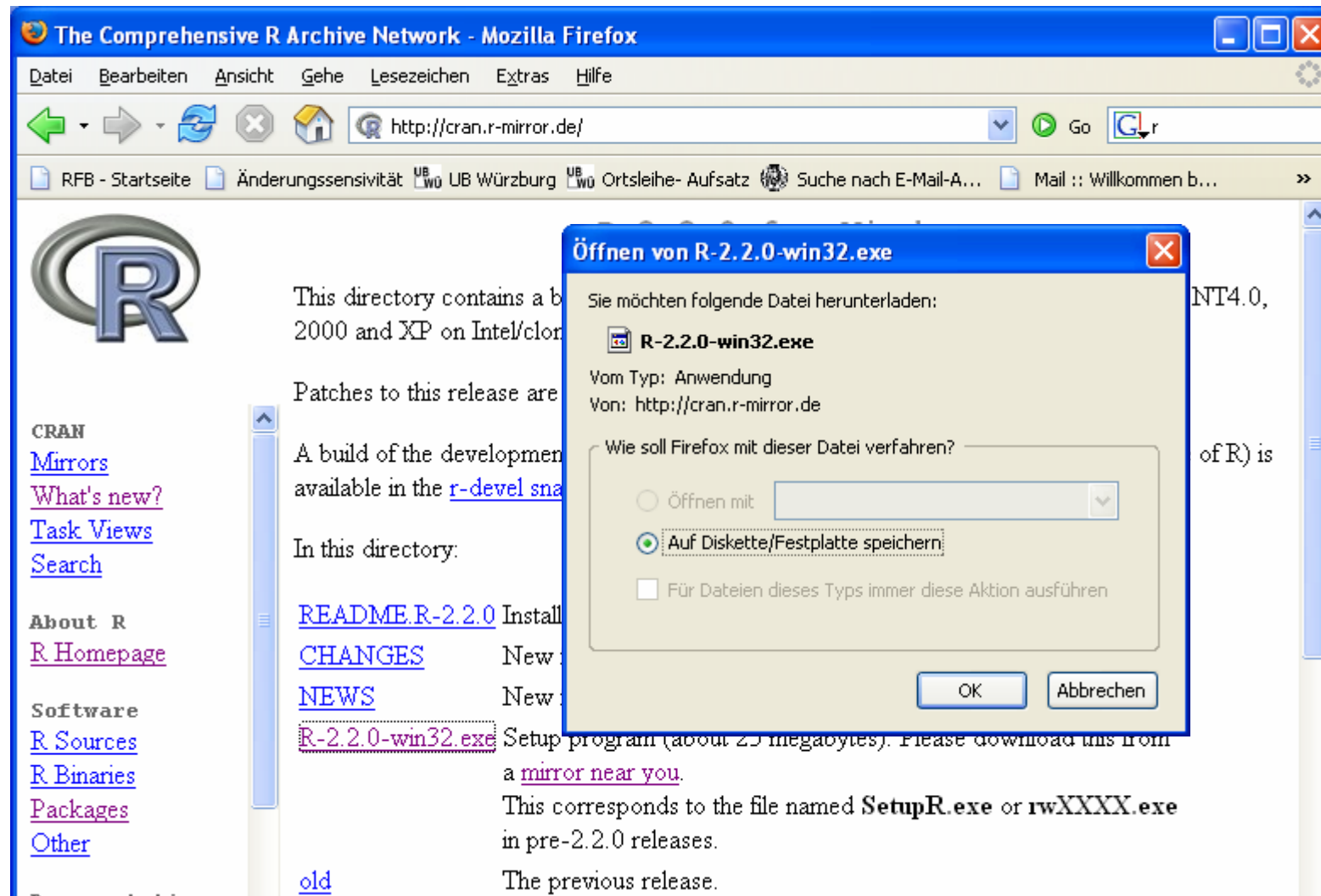
Factor 1 [41%] **Factor 3 [19%]**

Getting started:

- R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To download R, please choose your preferred [CRAN mirror](#).

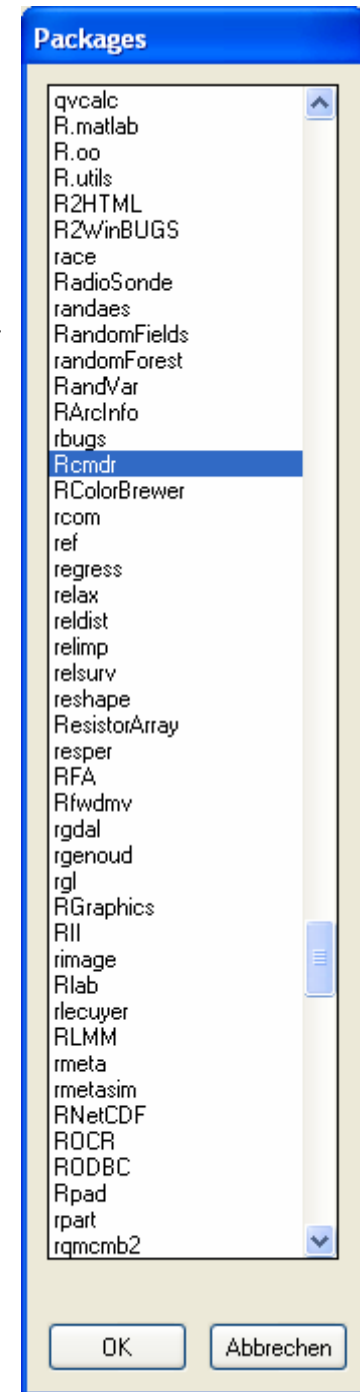
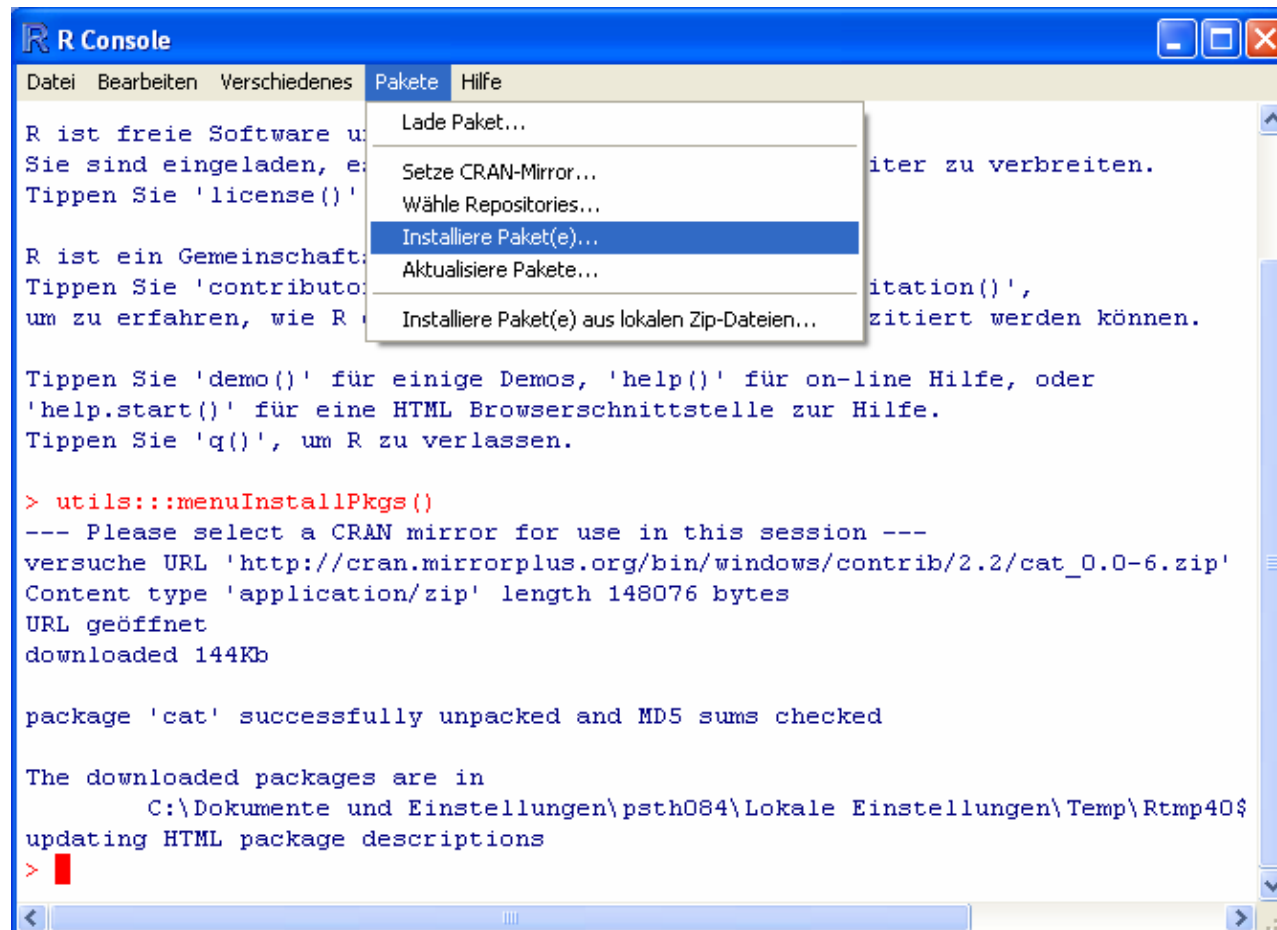
R: Download

- <http://cran.r-mirror.de/bin/windows/base/R-2.2.0-win32.exe>



Installation von R-Commander

- R-Konsole => Pakete => Installiere Paket(e)... => Rcmdr
- R-Konsole => Pakete => Lade Paket(e)... => Rcmdr



IV. Dokumentation, Support, Links

Dokumentation

Freie Handbücher:

- “[R for Beginners](http://cran.r-mirror.de/doc/contrib/rdebutts_en.pdf)” von Emmanuel Paradis
http://cran.r-mirror.de/doc/contrib/rdebutts_en.pdf
- “[Notes on the use of R for psychology experiments and questionnaires](http://cran.r-project.org/doc/contrib/Baron-rpsych.pdf)” by Jonathan Baron and Yuelin Li
<http://cran.r-project.org/doc/contrib/Baron-rpsych.pdf>

Kommerzielle Handbücher:

- Ligges, U. (2005). [Programmieren mit R](#). Heidelberg: Springer Verlag.
- Dolic, D. (2003). [Statistik mit R - Einführung für Wirtschafts- und Sozialwissenschaftler](#). München: Oldenbourg.

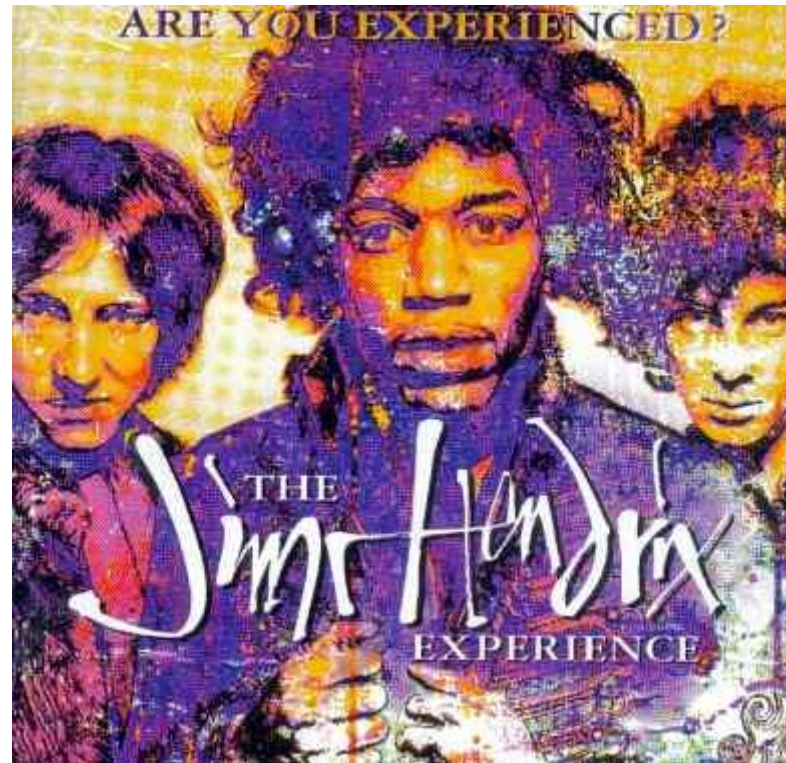
Support

- R-Console => Hilfe => HTML Hilfe
- R-Cmdr => Help (auch im aktiven Fenster einer Funktion verfügbar)
- The R Help Center
<http://www.stat.ucl.ac.be/ISdidactique/Rhelp/>
- R-Help-Mailing-Liste
<http://www.r-project.org/mail.html>
- Internet
<http://www.google.de>

Links

- [The R-Project](http://www.r-project.org)
<http://www.r-project.org>
- [The R Commander: A Basic-Statistics GUI for R](http://socserv.socsci.mcmaster.ca/jfox/Misc/Rcmdr/)
<http://socserv.socsci.mcmaster.ca/jfox/Misc/Rcmdr/>
- [Statistical Analysis on the Web](http://bayes.math.montana.edu/Rweb/Rweb.general.html)
RWeb: <http://bayes.math.montana.edu/Rweb/Rweb.general.html>
R-Online: <http://user.cs.tu-berlin.de/~ulfi/cgi-bin/r-online/r-online.cgi>
- [Statistiklabor](http://www.statistiklabor.de) – Statistik verstehen und anwenden:
<http://www.statistiklabor.de>
- [Editoren](http://www.sciviews.org/Tinn-R/)
TINN-R: <http://www.sciviews.org/Tinn-R/>
- [Weitere Hinweise der Methodenberatung](http://www.rehawissenschaft.uni-wuerzburg.de/methodenberatung/software.html)
<http://www.rehawissenschaft.uni-wuerzburg.de/methodenberatung/software.html>

R you experienced (with R)?



Yes!