

---

# Teaching Basic Statistics with R

**Bernhard Spangl,**  
*Universität für Bodenkultur, Wien*

joint work with  
**Lydia Matiasch and Robert Wiedermann,**  
*Universität für Bodenkultur, Wien,*

Vienna, Jan. 16th 2009

- \* many students (~700 in winter term, ~250 in summer)
- \* early in the curricula (1st, 2nd or 3rd semester)  
depending on the studies
- \* mostly Windows user
- \* not familiar with 'open source' or 'mirrors'  
(even not familiar with as graduate students)
- \* no further resources (time, staff, lecture rooms, etc.)  
to give them an additional introduction to R  
↳ only provide them with a modified version of R  
+ lecture notes

- \* download & install R-2.8.1-win32.exe
- \* Custom installation, only select
  - \* Main Files
  - \* Compiled HTML Help Files
  - \* Support Files for Package tcltk
  - \* Message Translation
- \* install additional packages, e.g. foreign, xlsReadWrite
- \* create directory 'eigene R-Dateien'
- \* create proper batch program to start R
- \* create `.First` & save default workspace '.RData'
- \* compress to archive 'R-Ordner.zip'

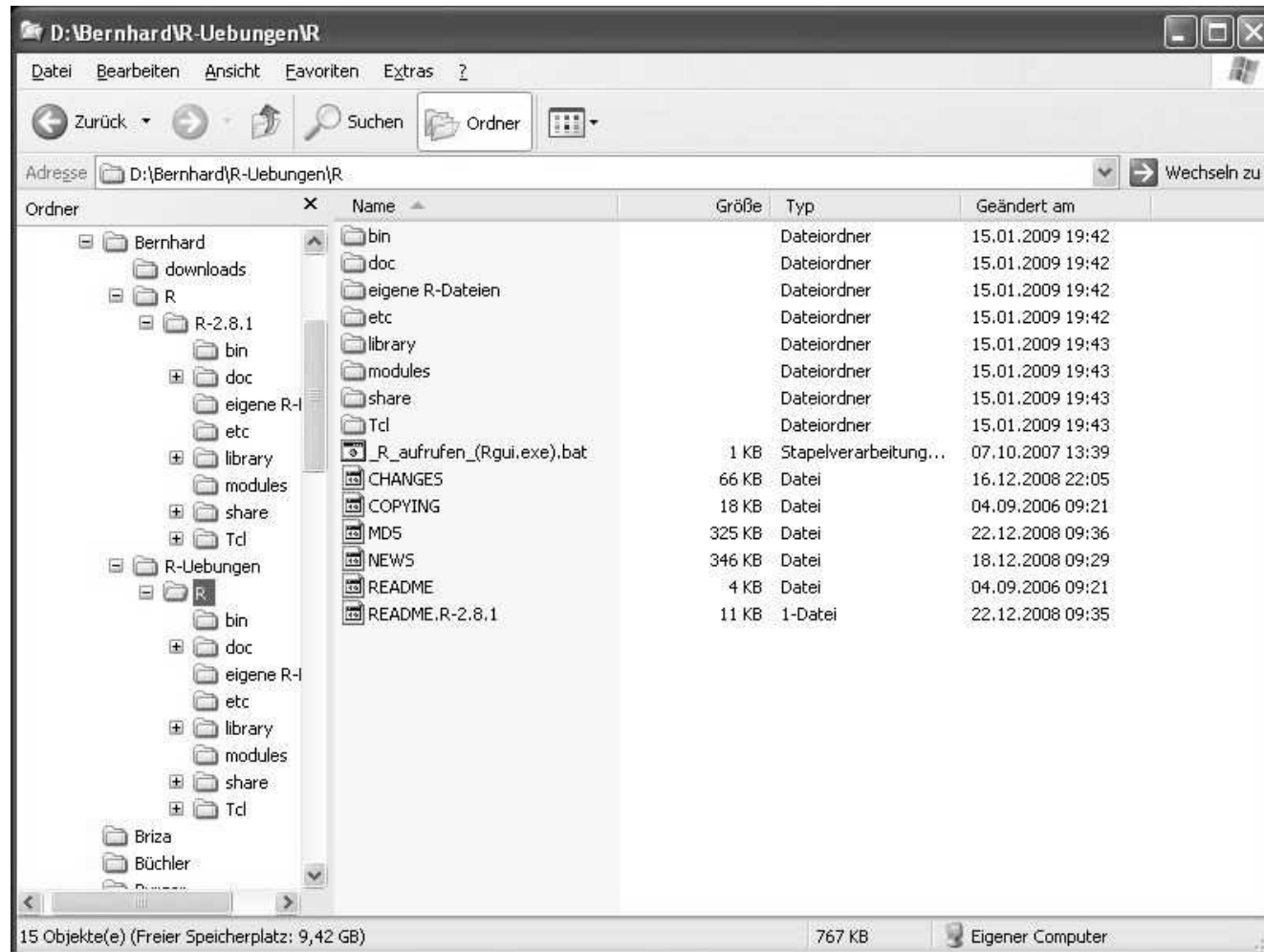
## \* batch program:

```
@echo off
echo Rgui wird gestartet ...
cd "eigene R-Dateien"
start ..\bin\Rgui.exe --sdi
```

## \* R function .First:

```
.First <- function () {
library(xlsReadWrite)
[...]
cat(getwd(), "\n")
[...]
}
```

# Screenshot



## ANOVA

Die entsprechenden Nullhypothesen für die Varianzanalyse lauten:

$H_{\text{Sorte}}$ : Die Sorte hat keinen Einfluss auf den theoretischen mittleren Ertrag.

$H_{\text{Steinmehl}}$ : Die Zugabe von Steinmehl zur Gülle hat keinen Einfluss auf den theoretischen mittleren Ertrag.

$H_{\text{Steinmehl:Sorte}}$ : Es gibt keine Wechselwirkungen zwischen den beiden Faktoren Sorte und Steinmehl.

Die Alternativhypothesen lauten: "Sorte und Zugabe von Steinmehl zur Gülle haben Einfluss auf den theoretischen mittleren Ertrag und es gibt Wechselwirkungen zwischen den beiden Faktoren." Das Risiko 1. Art wird mit  $\alpha = 0.05$  festgelegt.

```
> anova(lm(roggen$ertrag ~ rogggen$steinmehl * rogggen$sorte))
```

```
'+' ..... Modell ohne Wechselwirkungen
```

```
'*' ..... Modell mit Wechselwirkungen
```

```
'*' steht kurz für das Modell 'sorte + steinmehl + sorte: steinmehl'
```

Analysis of Variance Table

Response: rogggen\$ertrag

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
rogggen\$steinmehl	1	11.900	11.900	7.3731	0.014181	*
rogggen\$sorte	2	73.091	36.545	22.6424	1.218e-05	***
rogggen\$steinmehl:rogggen\$sorte	2	25.556	12.778	7.9168	0.003414	**
Residuals	18	29.052	1.614			

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Die p-Werte für alle drei Hypothesen liegen unter dem gewählten Risiko 1. Art  $\alpha = 0.05$ , daher müssen alle Hypothesen verworfen werden. Es haben daher sowohl die Zugabe von Steinmehl zur Gülle als auch die Sorte Einfluss auf den Ertrag und es gibt Wechselwirkungen zwischen diesen beiden Faktoren.

Die Mittlere Quadratsumme des Fehlers (1.614) entspricht der Fehlervarianz  $s_e^2$ .

## \* Benefit:

- \* combination of modified R + lecture notes works remarkable well
- \* portable, simply copy it to an USB stick
- \* form an educational point of view:  
students question their results  
(They had never done that when using SPSS!)

## \* Drawback:

- \* only apply recipes, but gain no real insight into R