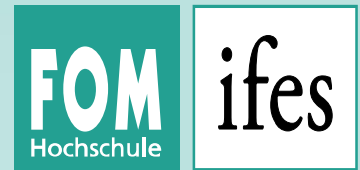




```
> sessionInfo()  
[1] "June 30 - July 3, 2015"  
[2] "Aalborg, Denmark"
```



Teaching R in heterogeneous settings: Lessons learned

useR! 2015
Aalborg, Denmark

Matthias Gehrke, Oliver Gansser, Bianca Krol, Karsten Lübke, Joachim Schwarz

1. **Who are we, whom do we teach: FOM**
2. **Obvious, but: Why we use R**
3. **Technical: What R tools do we use**
4. **Results: Survey**
5. **Lessons learned so far**

1. About FOM

- founded in 1993 by Employers' Associations in Essen
- state-recognised / system-accredited
- non-profit oriented
- more than 32,000 students / 320 professors
- unique network of study centres in Germany (30 locations)
- FOM is "going international" and has developed a global network with universities and educational institutions from e.g. China, USA, GB, Spain, and Hungary.
- FOM cooperates with many companies for dual studies (e.g. Bertelsmann, BP, Siemens, Thyssen Krupp, IBM).

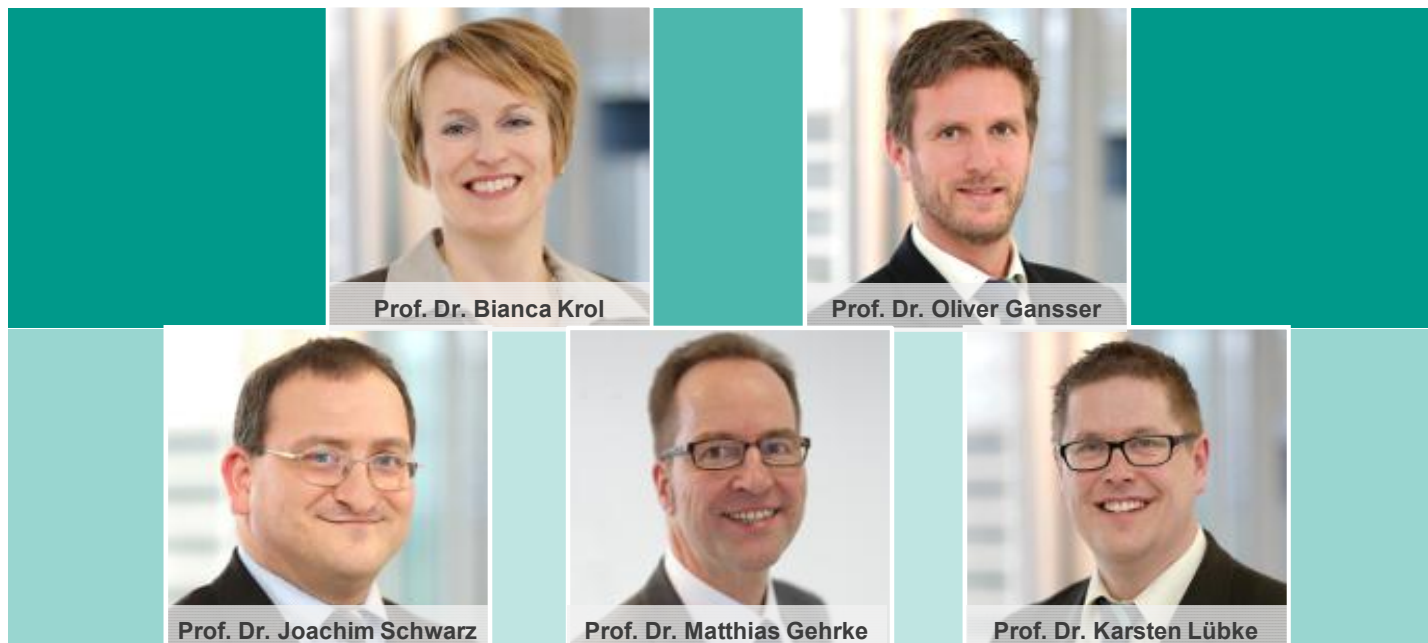


Studying at FOM University means...

- ... studying while working
 - With its concept of enabling students to study while working or completing an apprenticeship, FOM regards itself as a useful addition to the German university system.
 - By creating study conditions that are flexible and geared specifically towards the target group, the university provides employees with numerous opportunities for further development while allowing companies to adjust to the requirements that result from demographic developments and increased demand for qualified employees.

ifes – Institute for Empirical Research & Statistics

- 7 professors, 3 assistants and 4 research fellows
 - bundle empirical skills to support FOM's applied research
 - develop and support statistics and method training in FOM's bachelor and master degree programs, as well as in the PhD program of FOM



2. Reasons for and against R

Pro:

- Functionality: Methods for Fitting Garch Models (Finance) to Structural Equation Models (e.g. HR). From Text Mining (Computer Science) to solving Travelling Salesman Problems (Logistics)
- Documentation & Literature
- Available for Win, Mac, Linux
- Free & Open Source

Contra:

- No intuitive start
- Lecturers more familiar with SPSS

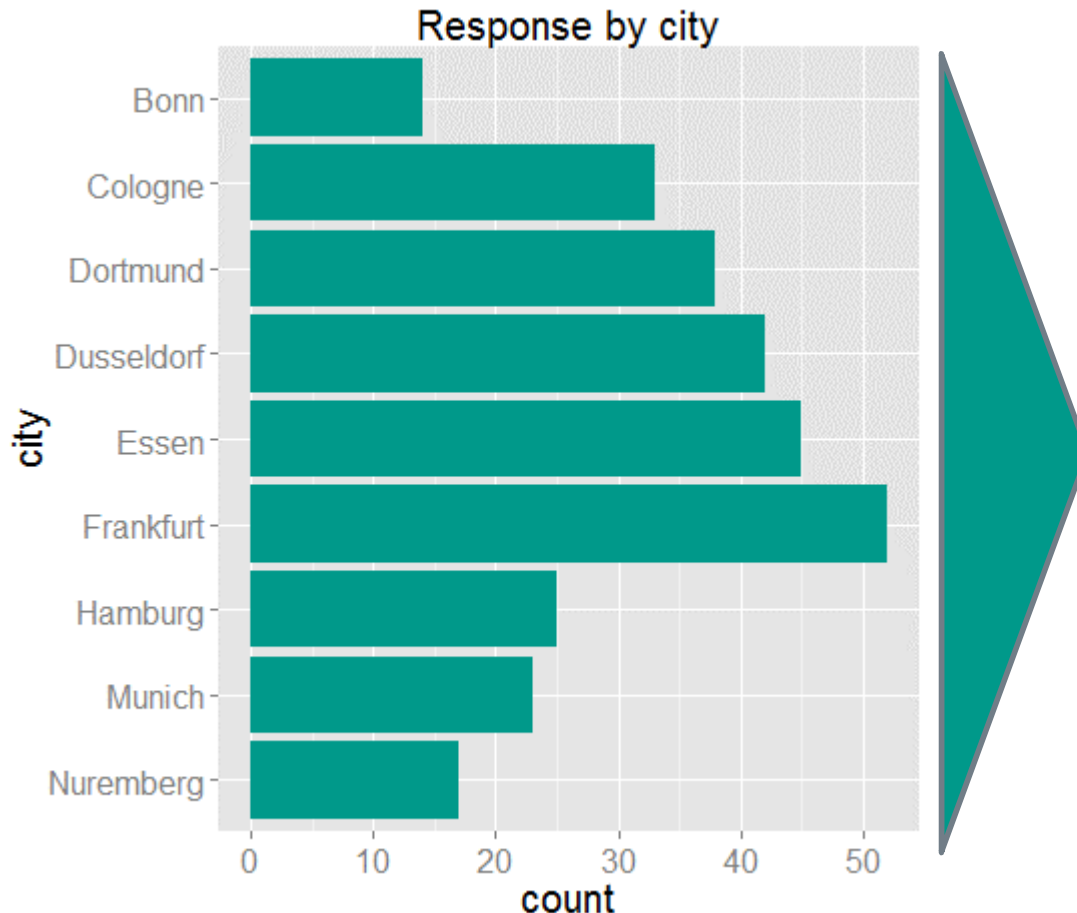
3. The R framework at FOM

- Central links to R and R Documentation
- For Windows: Portable R Version with preinstalled packages and automatic start of the R Commander (Rcmdr).
Customized `Rprofile` and `Renviron` in `etc/`. Batch file for starting R
- For other, e.g. Mac: Detailed installation guide especially for R Commander (requirements etc.)
- Support email address
- Training courses for lecturers: beginners' and advanced level
- Installed packages (including dependencies):
`Rcmdr`, `klaR`, `conjoint`, `CTT`, `linprog`, `rela`, `Matching`,
`rpart`, `nortest`, `plm`, `psych`, `pwr`, `randomForest`,
`sampleSelection`, `lavaan`, `tseries`, `arules`, `arulesViz`, `tm`,
`sem`, `paran`, `ROCR`, `forecast`, `XLConnect`,
`fGarch`, `strucchange`, `irr`, `coin`

4. A Survey of the acceptance of R

- Two years after the start of the FOM Master of Science programs (with teaching of R), we conducted a survey to get information about the acceptance of R.
- The survey was conducted between May 22th and June 15th.
- We distributed the survey directly to all target students via email with link to a web page, containing a standardized online questionnaire. Furthermore, we sent an accompanying letter to all lecturers requesting assistance.
- Target were all students, whose curriculum required at least one lecture in R, i.e. all Master of Science students (1721 students / summer 2015) plus Bachelor of Business & Psychology students (896 students / summer 2015).
- Methods of evaluation were descriptive statistics, principal component analysis and linear mixed model.
- Total response was 390 (14.9% response rate), 325 were used for the descriptive statistics and the PCA, 209 for the linear mixed model.

Response by city



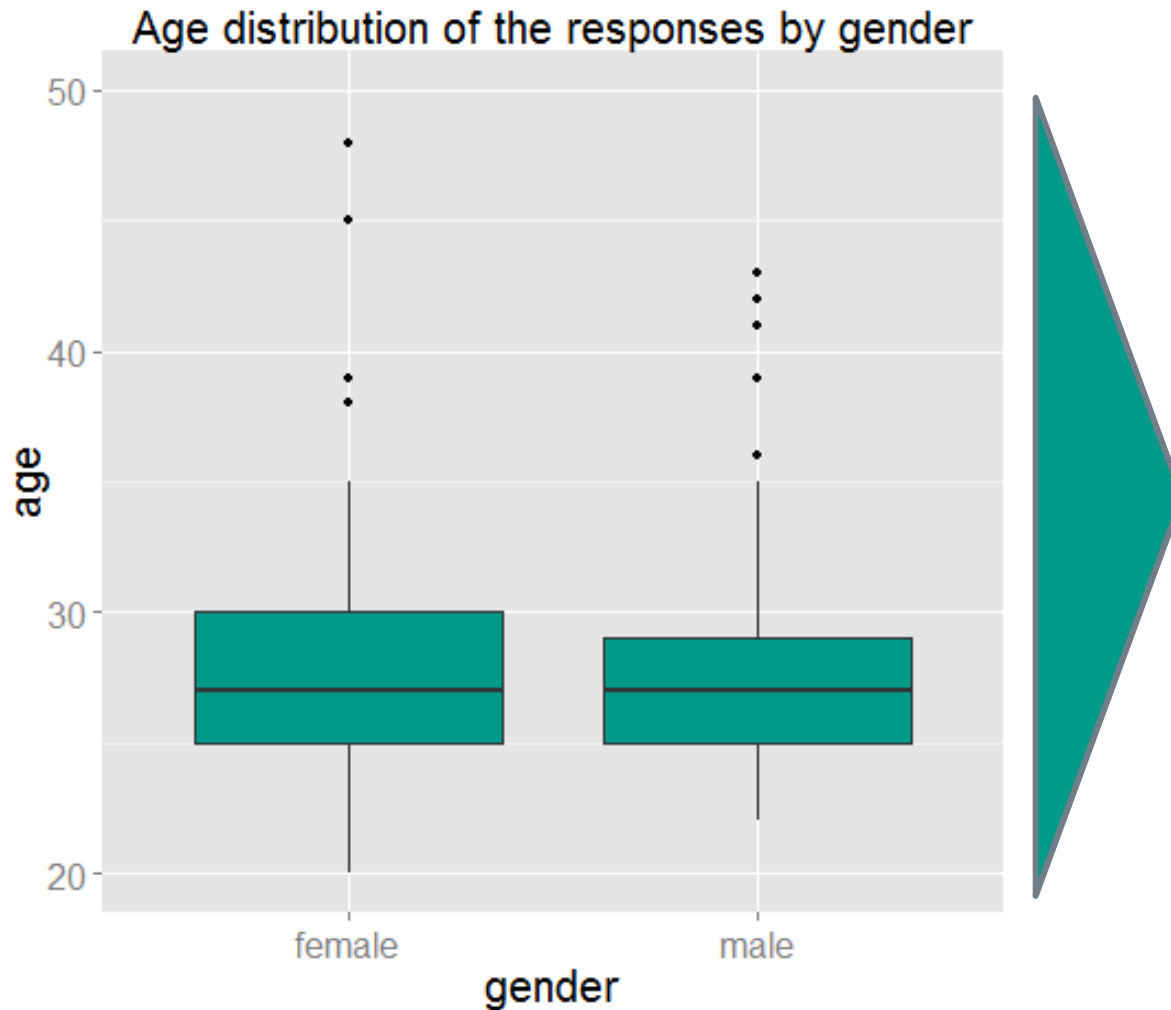
- We only got from 9 of the 31 locations in Germany 10 or more responses.
- Although Essen is the largest location with most of the students, the highest response came from Frankfurt.
- Although Hamburg and Munich are large locations similar to Frankfurt, we got much less response from there.

Response by program



- Although about one third of the population aims at a Bachelor degree in Business&Psychology, their share at the total response is very small.

Age distribution of the responses by gender



- Most of the students are between 25 and 30 years old, with some older than 35 years. This is typical for FOM Master students.
- There is no significant difference between gender.

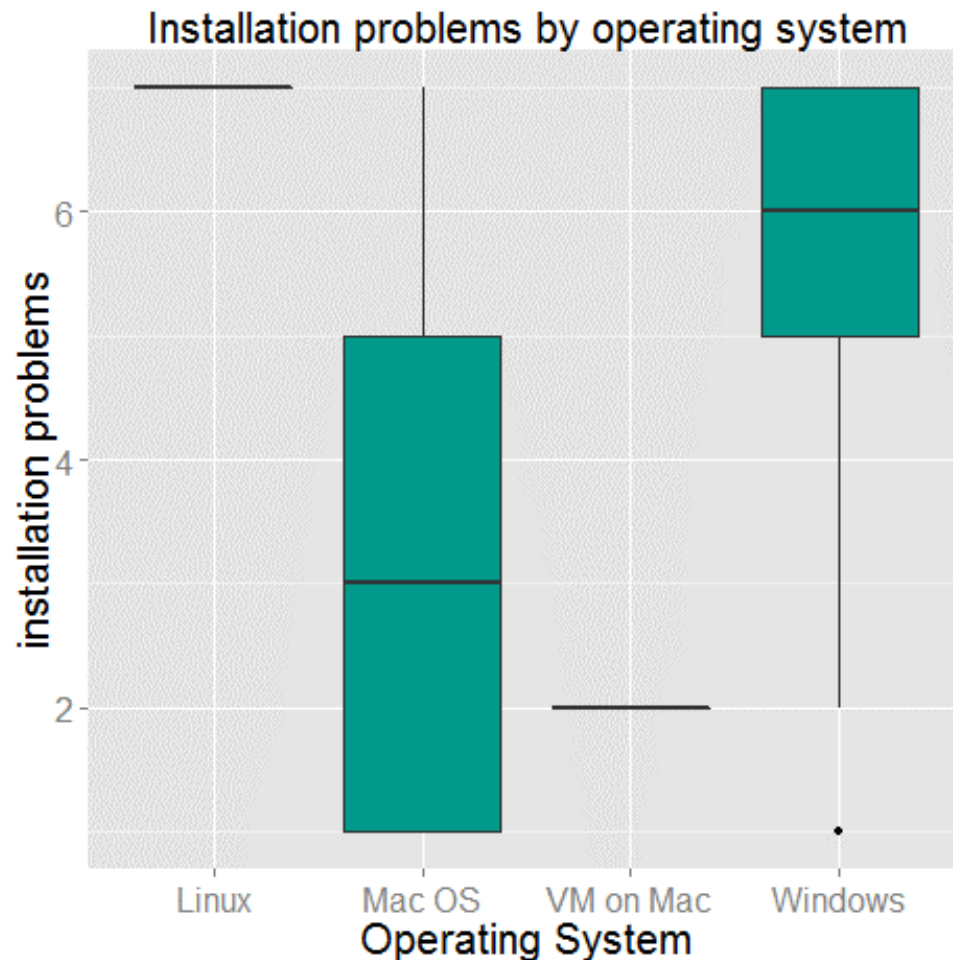
4. A Survey of the acceptance of R

- To catch the acceptance of R, we used altogether 15 items and 7 point Likert scales. The Likert scale ranged from 1 (absolute agreement) to 7 (absolute disagreement).
- We condensed the items using a PCA (package psych) and extracted two principal components (cumulative variance = 0.58). The two principal components could be described as follows:
 - Usability: how easy is it to use R? It contains if students...
 - are often confused using R
 - make often mistakes
 - find the use of R frustrating
 - find R easy to understand
 - a.s.o.
 - Usefulness: how useful is R? It contains if students...
 - consider R useful for their studies
 - consider R useful for their jobs
 - find that R gives useful orientation for statistical analyses
 - find statistical analyses very difficult without R

4. A Survey of the acceptance of R

- We investigated if installation problems were dependent on operating system used.
- For further analyses we used the results of the PCA to calculate factor values for the two principal components. The lower the values of the principal components, the higher the usability / usefulness.
- First, we were interested if there were differences in the evaluation of the two principal components concerning gender or study course.
- Then, we investigated if there were any developments over time. For this, we evaluated the term in which R was taught to the students.
- Finally we investigated if there were differences in usefulness in regard to studies or work usage.
- We addressed these questions with descriptive analyses and a linear mixed model with two fixed effects (gender and study course) and two random effects (location and term) using the lme4-package.

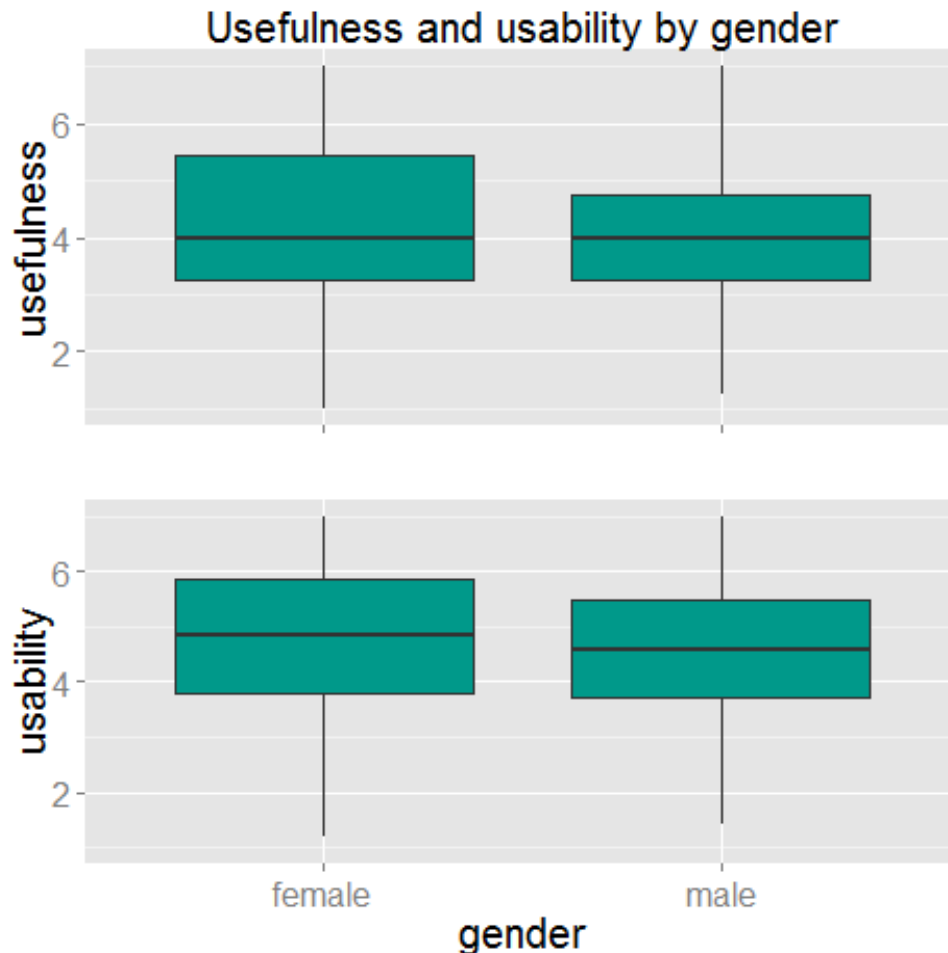
Installation problems



- There are significant differences between operating systems.
- Linux users did not face any installation problems.
 - However this result might not be representative (2 answers only).
- Pre-configured windows package (FOM portable version) helped a lot.

The lower the values, the more installation problems were observed.

Usefulness and usability by gender



- There are no significant differences between gender (result of linear mixed model).
- Usefulness centers around 4, i.e. the middle category, and it is slightly left skewed for females.
- Usability is for both males and females above the middle category.

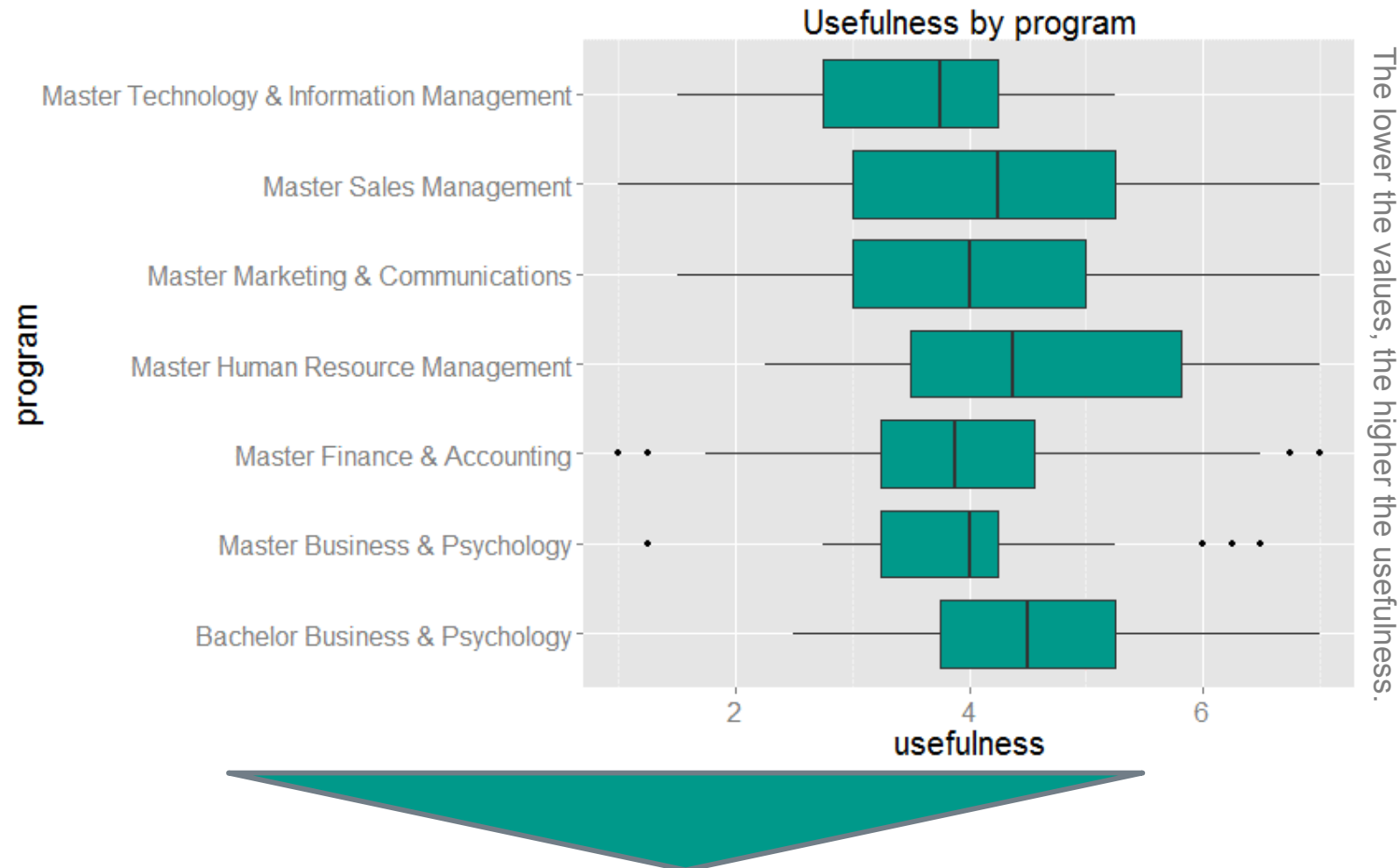
The lower the values, the higher the usefulness resp. usability.

Usability by master program



- Usability got worst ratings from Master HRM and Bachelor BP students. However, differences are not significant.

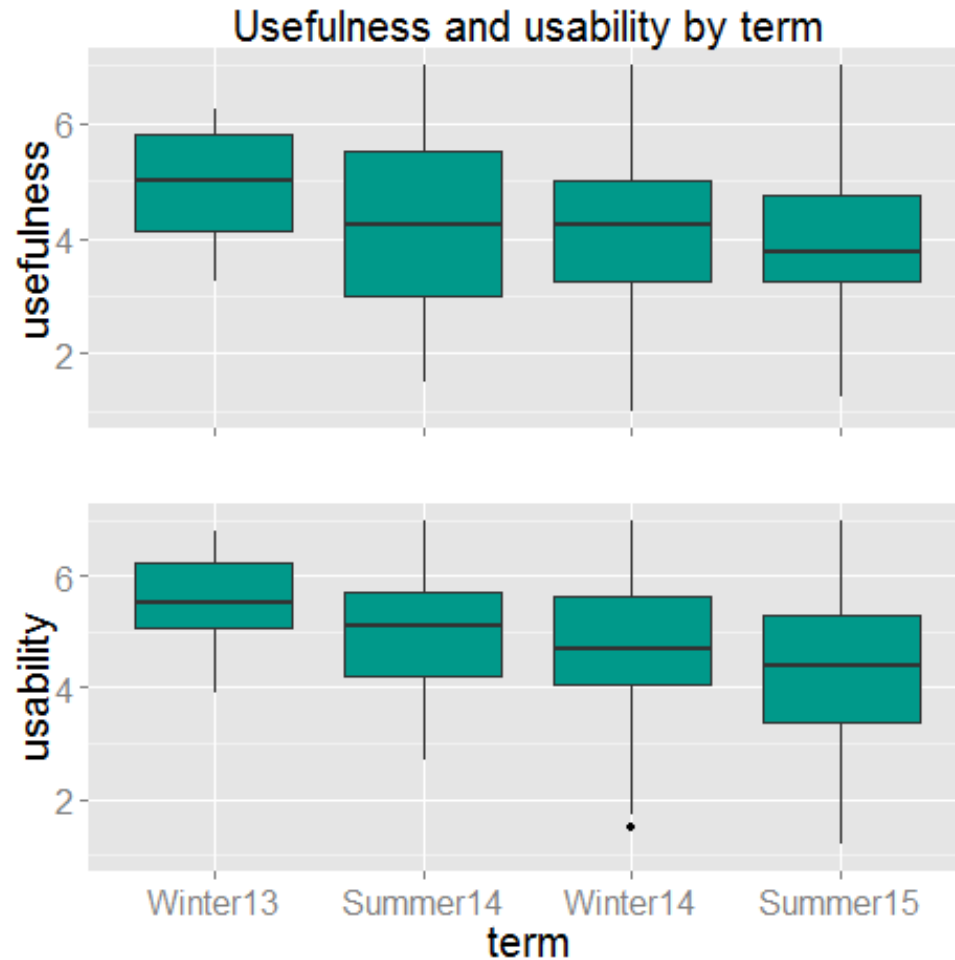
Usefulness by master program



- Usefulness got best ratings from Master TIM students. The differences are significant on 5% level.

4. A Survey of the acceptance of R

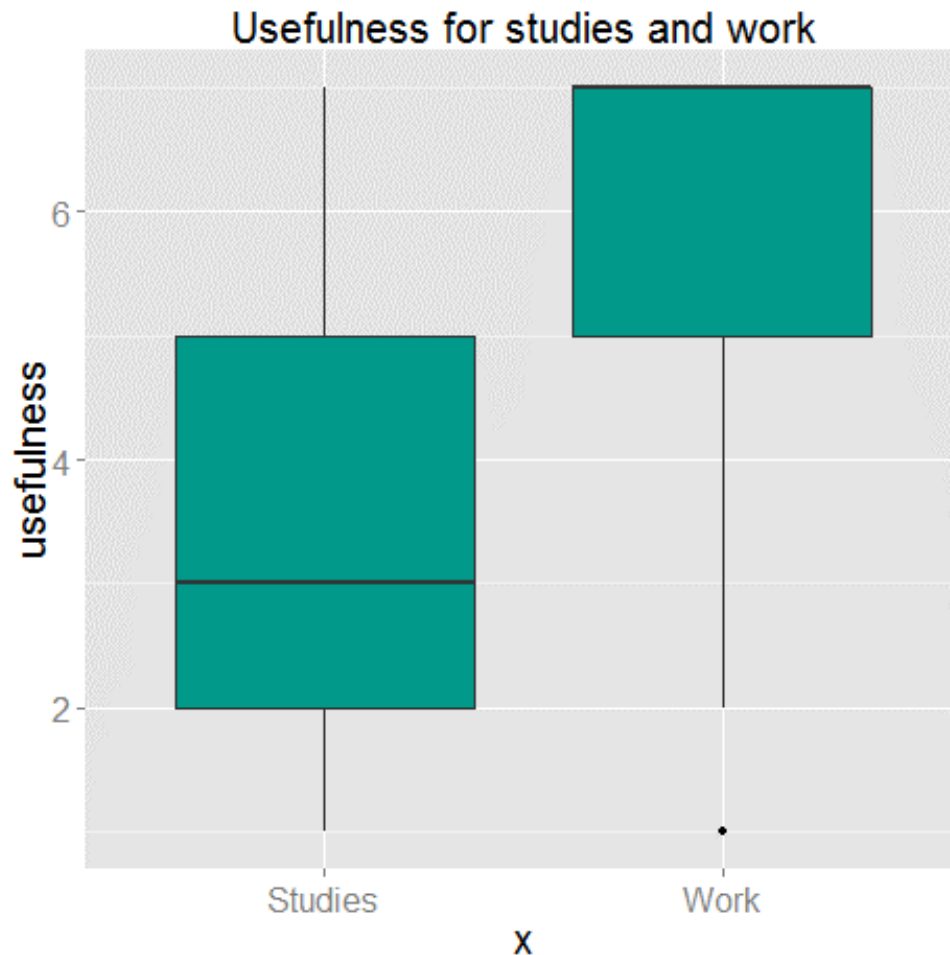
Usefulness and usability by term



- Both usefulness and especially usability show a trend to better ratings.
- Still, even with a positive trend, ratings are not below the middle category mostly.

The lower the values, the higher the usefulness resp. usability.

Usefulness for studies and work



- Differentiated by usefulness for studies and work the acceptance to use R for studies is much higher.
- Median for usefulness for work is 7, meaning students do not see any usefulness for work at all.

The lower the values, the higher the usefulness.

4. A Survey of the acceptance of R

Key findings

- After all, almost 15% of the population responded to the questionnaire. However, descriptive distributions show that the respondents are most likely not representative.
- There are no differences in gender, but usability got worse ratings than usefulness.
- The ratings of usefulness improve over time. This could mean that students more and more accept and understand, that a Master of Science requires statistical analyses and a tool which is able to do so.
- The ratings of usability improve over time. This could mean that teaching R becomes better as lecturers learn to handle R better.

5. Lessons Learned

- Using a GUI like the R Commander helps in getting started.
- Train the lecturers – there is still (even with Rcmdr) a gap for SPSS users.
- Support email-address: more than 500 emails (in & out, including administrative) within 18 month.
- Update portable Win Version every term to synchronize with actual versions on CRAN.
- Write detailed installation guides and try to encourage everybody to read them – still ...
- Quantitative empirical research by students can increase.

Thank you for your attention!

Prof. Dr. Matthias Gehrke
matthias.gehrke@fom-ifes.de

Prof. Dr. Joachim Schwarz
joachim.schwarz@fom-ifes.de

Contact

ifes Institute for Empirical Research & Statistics
FOM University of Applied Science
Leimkugelstraße 6 | 45141 Essen | GERMANY

E-Mail **R@fom.de** | Web **www.fom-ifes.de/R**