

How much is it? Validation of Open-Source-Software Using the example of R

PhUSE 2010 – 20.10.2010

Company

- HMS Analytical Software is a specialist for Information Technology in the field of Data Analysis and Business Intelligence Systems
- Profile
 - 40 employees in Heidelberg, Germany
 - SAS Institute Silver Consulting Partner for 14 years
 - Doing data oriented software projects for more than 20 years
- Technologies
 - Analytics and Data Management:
SAS, JMP, R, Microsoft SQL Server
 - Application Development: Microsoft .NET, Java

Our IT Services for the Life Science Industry (SAS, JMP and R)

- Independent Consulting
- Programming
- Data Management
- Training and Individual Coaching
- Application Development and Integration
- Software Validation



What does Open Source mean?

- no support hotline or customer service for urgent requests
- often do not focus on user interfaces and lacks good documentation (not user friendly)
- open source libraries focus on the developer issues, and not on those of the users
- often no information is available about which module has been tested and how much of the code is covered

General CONS of Open Source Software

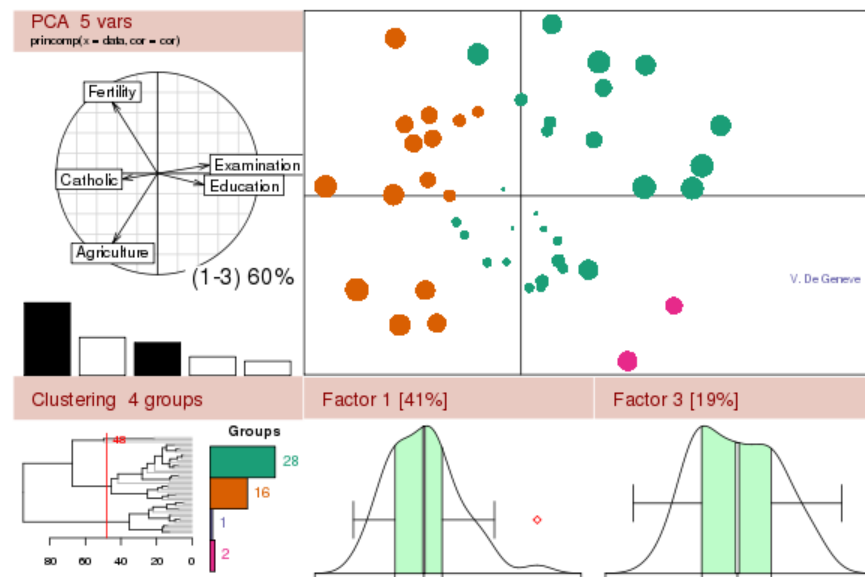
- no support hotline or customer service for urgent requests
- often do not focus on user interfaces and lacks good documentation (not user friendly)
- open source libraries focus on the developer issues, and not on those of the users
- often no information is available about which module has been tested and how much of the code is covered

General PROS of Open Source Software

- the source code is available to all and can be modified
- many open source projects are free of charge
- no need to start from scratch, but rather use existing open source libraries
- the amount of widely spread developers increases the probability to identify bugs
- generally, a community of developers are able to provide support
- makes license management easier

What is R?

- R is an environment for statistical computing and graphics
- R is a scripted based software language
- Open Source Software



Why Use R?

R under the General Public Licence

- R is available free of charge
- everybody is allowed to use the software for any purpose
- copies of the source code can be made free of charge or for a fee
- the source code is open to everybody and may be modified
- it is allowed to distribute the modified code for a fee, but the source code must be open for the customer

Principles Aspects of Using R in a Validated Environment



„We can not use R because it is not validated“

- Would you validate Microsoft Excel?
- What should be validated?
- The FDA does not allow the use of R?

The FutuRe

- ◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡

SLIDE: 21

Dr. Peter Beyer
© 2009 HMS Analytical Software GmbH

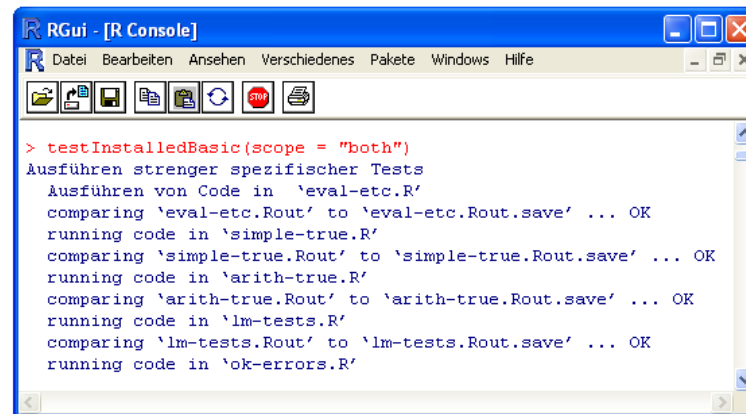
Provided IQ-Tests of R

- **Basic packages:**

includes functions for arithmetic operations, complex numbers, R expressions, linear models, regular expressions etc.

- Call for this test routine:

```
testInstalledBasic(scope = "both")
```



```
RGui - [R Console]
Datei Bearbeiten Ansehen Verschiedenes Pakete Windows Hilfe

> testInstalledBasic(scope = "both")
Ausführen strenger spezifischer Tests
Ausführen von Code in 'eval-etc.R'
  comparing 'eval-etc.Rout' to 'eval-etc.Rout.save' ... OK
  running code in 'simple-true.R'
  comparing 'simple-true.Rout' to 'simple-true.Rout.save' ... OK
  running code in 'arith-true.R'
  comparing 'arith-true.Rout' to 'arith-true.Rout.save' ... OK
  running code in 'lm-tests.R'
  comparing 'lm-tests.Rout' to 'lm-tests.Rout.save' ... OK
  running code in 'ok-errors.R'
```

Provided IQ-Tests of R

- **Base packages:**
includes functions for dates, grid graphics, regression splines, ANOVA, data set examples etc.
- Call for this test routine:

```
testInstalledPackage ("base")
```

Provided IQ-Tests of R

- **Recommended packages:**
includes functions for bootstrapping classification, cluster analysis, kernel smoothing survival analysis etc.
- Call for this test routine:

```
testInstalledPackage ("recommended")
```

IQ of R

What is tested?

OQ of R

OQ of R

Disadvantages of this test approach

OOQ of R

Evaluation of this test approach

OQ of R

What should be done?

Calling R from other Software Systems

Calling R from other Software Systems

R (D) COM Server

Calling R from other Software Systems

RWebServices

Calling R from other Software Systems

Rserve

Calling R from other Software Systems

.NET WCF Services

Conclusion

Thank you for your Attention

Dr. Peter Bowerunge

Software Engineer

HMS Analytical Software GmbH

Rohrbacher Str. 26 • 69115 Heidelberg

Telefon +49 6221 6051-0

Peter.Bowerunge@analytical-software.de

www.analytical-software.de