

Calling R Functions from SAS...

....with PROC IML

PhUSE 2011 – 12.10.2011

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 Partner for 15 years
 - Doing data oriented software projects for more than 20 years
 - Focus on life science industry
- 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, R and Microsoft)

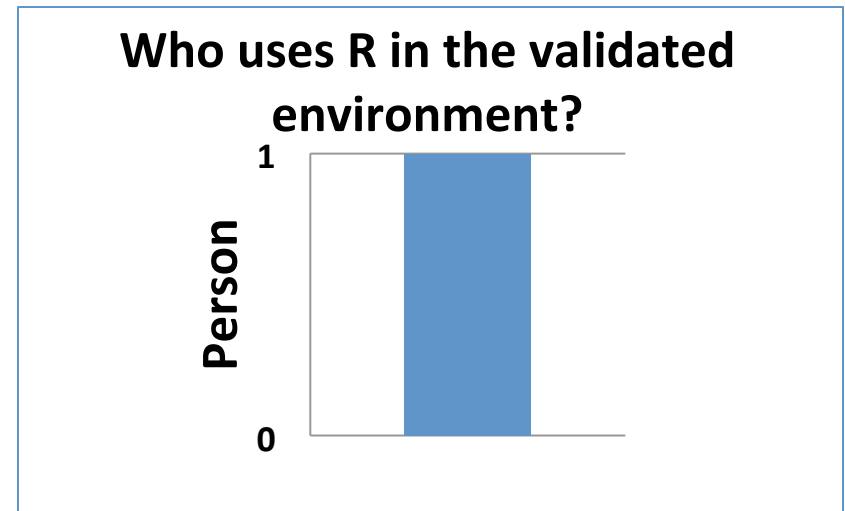
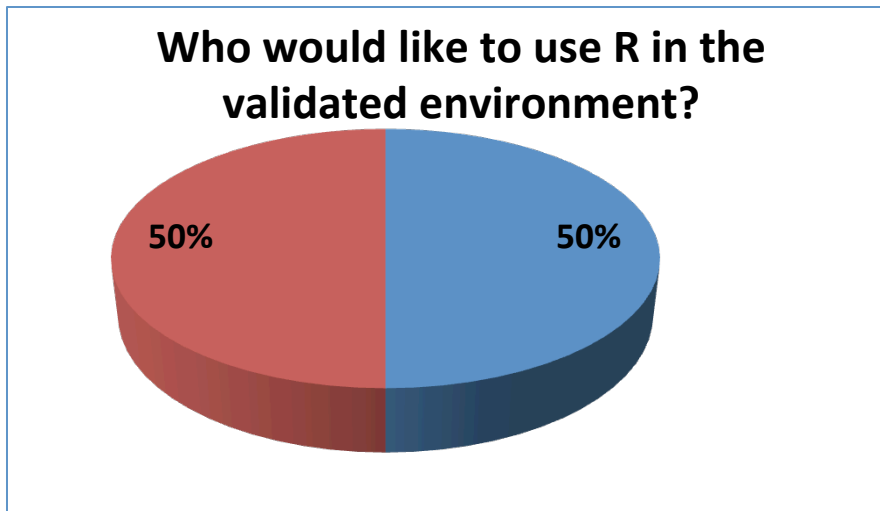
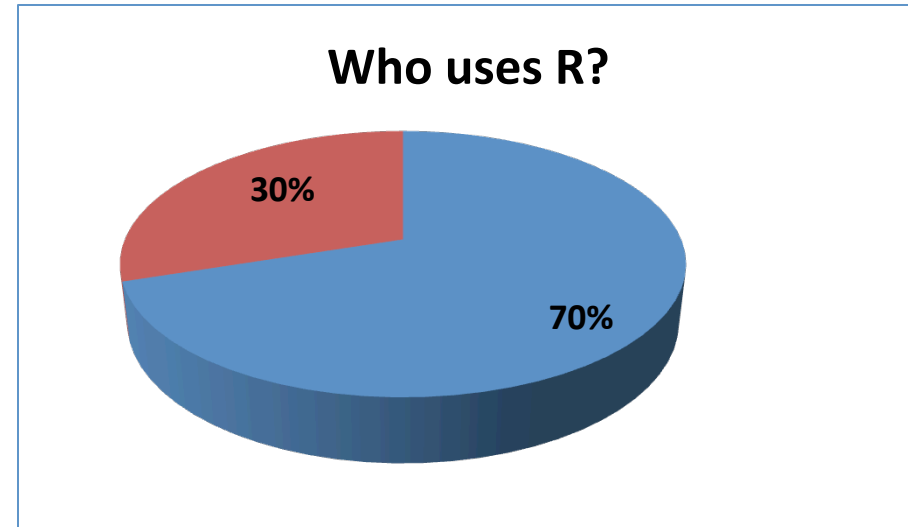
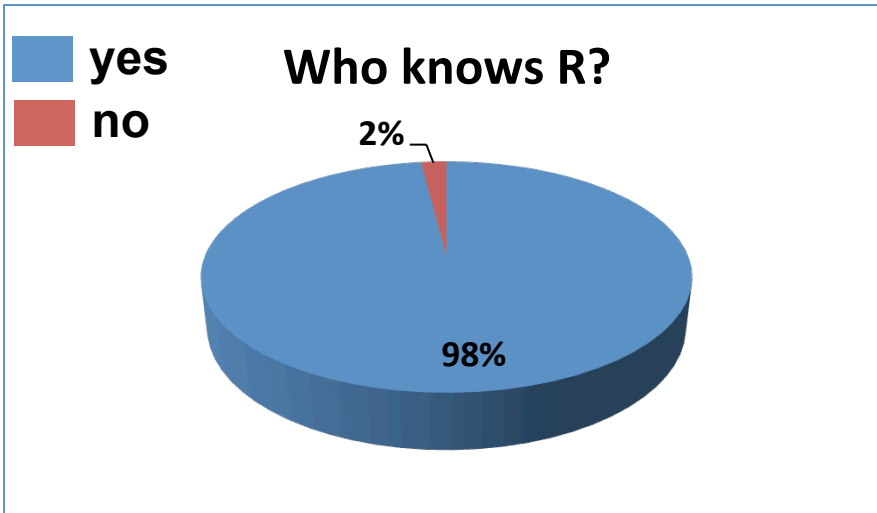
- Independent Consulting
- Programming
- Data Management
- Data Mining / Analysis
- Training and Individual Coaching
- Application Development and Integration
- Software Validation



Agenda

- Servey
- Who needs an interface to R
- Requirements
- How to submit R Code?
- What about data sets and SAS formats?
- Do I get my R Graphics into SAS-Report
- How to handle errors coming from R
- Questions that came up during PhUSE

Degree of Popularity of R compared to PhUSE 2010



The Best of two Worlds

- [-]  SAS Products
 -  SAS Procedures
 - [+]  Base SAS
 - [+]  SAS/ACCESS
 - [+]  SAS/AF
 - [+]  SAS/ASSIST
 - [+]  SAS/CONNECT
 - [+]  SAS Data Quality Server 9.2: Reference
 - [+]  SAS/EIS
 - [+]  SAS/ETS 9.22 User's Guide
 - [+]  SAS/FSP
 - [+]  SAS/GIS
 - [+]  SAS/GRAPH
 - [+]  SAS/IML Studio
 - [+]  SAS/IML 9.2 User's Guide
 - [+]  SAS/INSIGHT
 - [+]  SAS Integration Technologies
 - [+]  SAS/IntrNet
 - [+]  SAS/LAB Software
 - [+]  SAS 9.2 National Language Support Reference Guide
 - [+]  SAS OLAP Server
 - [+]  SAS/OR
 - [+]  SAS/QC
 - [+]  SAS/SECURE
 - [+]  SAS/SHARE
 - [+]  SAS/STAT
 - [+]  SAS Stat Studio
 - [+]  SAS/Warehouse Administrator
 - [+]  SAS/TOOLKIT Help for Windows



Available CRAN Packages By Name

[A](#)[B](#)[C](#)[D](#)[E](#)[F](#)[G](#)[H](#)[I](#)[J](#)[K](#)[L](#)[M](#)[N](#)[O](#)[P](#)[Q](#)[R](#)[S](#)[T](#)[U](#)[V](#)[W](#)[X](#)[Y](#)[Z](#)

abc	Tools for Approximate Bayesian Computation (ABC)
abd	The Analysis of Biological Data
abind	Combine multi-dimensional arrays
abn	Multidimensional Bayesian Regression
AcceptanceSampling	Creation and evaluation of Acceptance Sampling Plans
ACCLMA	ACC & LMA Graph Plotting
accuracy	Tools for testing and improving accuracy of statistical results
acepack	ace() and avas() for selecting regression transformations
aCGH.Spline	Robust spline interpolation for dual color array comparative genomic hybridisation data
ACNE	Affymetrix SNP probe-summarization using non-negative matrix factorization
actuar	Actuarial functions
ada	ada: an R package for stochastic boosting
adabag	Applies Adaboost.M1 and Bagging
ADaCGH	Analysis of data from aCGH experiments
AdaptFit	Adaptive Semiparametric Regression
adaptivetau	Tau-leaping stochastic simulation
adaptTest	Adaptive two-stage tests
ade4	Analysis of Ecological Data : Exploratory and

The Best of two Worlds



Client Tier

Management Console
Data Integration Studio
Information Map
Studio
Enterprise Guide
....

Web Tier

Web Report Studio
Information Delivery
Portal
STP Web Application

Metadata Server

Server Tier

Connect Server
SAS Foundation
OLAP Server
STP Server
....

Data Tier

Data Sets
OLAP Cubes
ERP data structure
RDBMS Tables
....



BioConductor

Bioinformatic community

Rgeo

Spatial Statistics

gR

Graphical Models

Robus

Robus statistics

Rmetrics

Financial Market Analysis

Omega

Distributed Statistical Computing

Who needs an Interface to R?

- People who don't want a "religious" decision between SAS or R. They want both!
- People who want to use one of the huge amount of add-on-packages of R
- Lazy people, who don't want to convert their existing R code to SAS
- People who switch from R to SAS
- People who want to validate their SAS programs (Nikhil Abhyankar, yesterday talk)

SAS/IML as an Interface to R

- SAS Interactive **M**atrix **L**anguage (IML) is a programming language for explorative data analysis.
- wide range of most standard matrix operations
- direct access to SAS data sets
- create dynamic graphics



it has the an interface to R



PROC IML use this interface

Requirements to Use PROC IML as an Interface to R

	SAS 9.22	SAS 9.3
32-bit and 64-bit Windows	X	X
32-bit and 64-bit Linux	-	X
32-bit version of R	X	X
64-bit version of R	-	X
Supported R Versions	R-2.11.1 or before	??

- R must be installed on the same computer / server
- access to the SAS workspace server via Enterprise Guide®

How to submit R Code

1. Start SAS with the RLANG-option

(C:\...\sas.exe -RLANG).

2. Check access to R

```
PROC OPTIONS OPTION=RLANG;  
RUN;
```

NORLANG:	Do not support access to R language interfaces
ERROR:	The RLANG system option must be specified in the SAS configuration file or on the SAS invocation command line to enable the submission of R language statements.
WARNING:	SAS option RLANG is not supported on this host.
RLANG:	Support access to R language interfaces

How to submit R Code

3. Use PROC IML as the interface to R

```
PROC IML;
```

```
SUBMIT / R;
```

```
a <- 2  
paste("a:", a)  
b <- 3  
paste("b:", b)  
c <- a * b  
paste("c:", c)
```

```
ENDSUBMIT;
```

```
QUIT;
```

```
[1] "a: 2"  
[1] "b: 3"  
[1] "b: 3"
```

Sending Data from SAS to R and back

- IML built-in subroutines-

From SAS to R:

Subroutine	SAS Source	R Destination
ExportDataSetToR	SAS data set	R data frame
ExportMatrixToR	SAS/IML matrix	R matrix

From R to SAS:

Subroutine	R Source	SAS Destination
ImportDataSetFromR	R expression	SAS data set
ImportMatrixFromR	R expression	SAS/IML matrix

Sending SAS Data Sets to R

```
PROC IML;  
  RUN ExportDataSetToR( "SASHELP.Class"  
                        , "RDataFrame" );  
  
  SUBMIT / R;  
    RDataFrame[1:4, 1:5]  
  ENDSUBMIT;  
QUIT;
```

	Name	Sex	Age	Height	Weight
1	Alfred	M	14	69.0	112.5
2	Alice	F	13	56.5	84.0
3	Barbara	F	13	65.3	98.0
4	Carol	F	14	62.8	102.5

Getting SAS Data Sets back from R

```
PROC IML;
  SUBMIT / R;
    iris.setosa = iris[(iris$Species=="setosa"),]
  ENDSUBMIT;
RUN ImportDataSetFromR("WORK.IrisSetosa"
                        , "iris.setosa");
QUIT;
```

```
PROC MEANS data = WORK.IrisSetosa N SUM MEAN;
RUN;
```

The MEANS Procedure

Variable	N	Sum	Mean
Sepal_Length	50	250.3000000	5.0060000
Sepal_Width	50	171.4000000	3.4280000
Petal_Length	50	73.1000000	1.4620000
Petal_Width	50	12.3000000	0.2460000

What about Date, Times and Datetime Values?

From SAS to R:

SAS	R
Date (e.g. DATEw.d)	Class „Date“
Datetime (e.g. DATETIMEw.d)	Classes that inherit from Class „POSIXT“
Time (e.g. TIMEw.d)	Classes inherit from Class „POSIXT“
All other formats	Class „numeric“

What about Date, Times and Datetime Values?

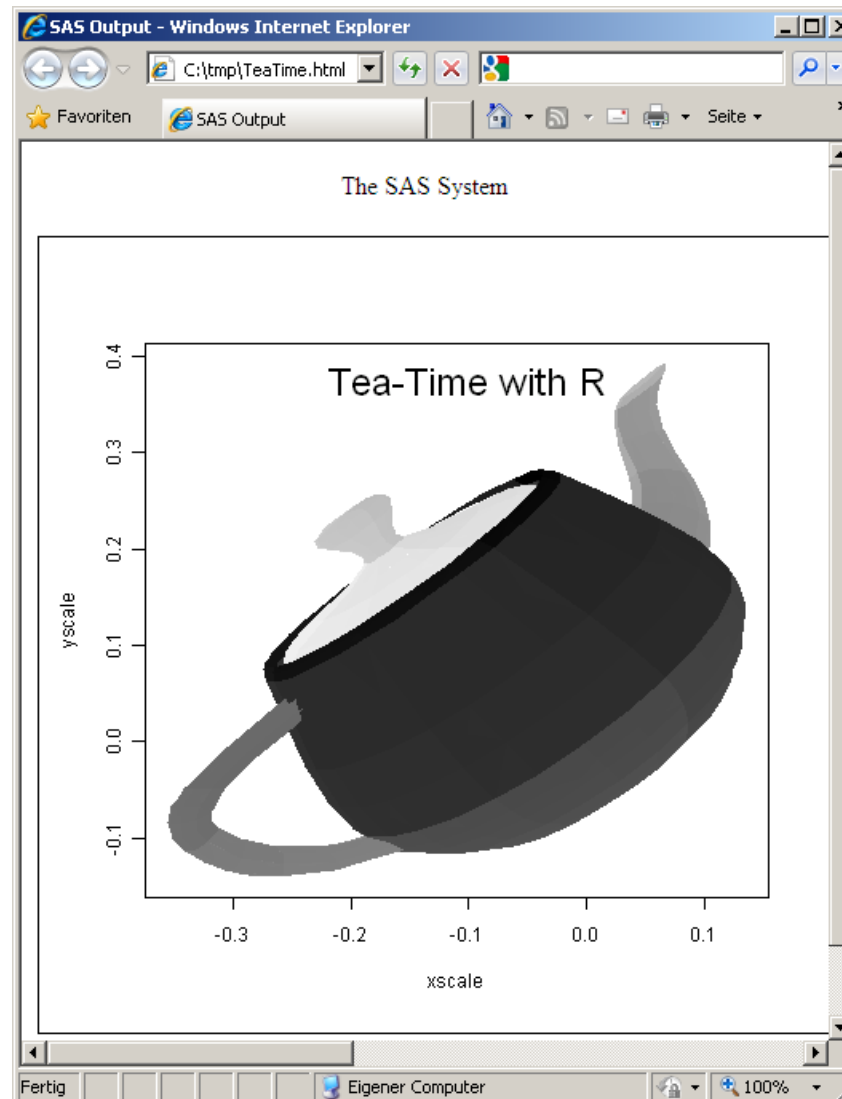
From R to SAS:

R	SAS
Class „Date“	DATE9.
Class „POSIXT“	TIME19.
In all other cases	no SAS format is assigned

Get Graphics from R to SAS

```
PROC IML;  
    SUBMIT / R;  
        png(file ="c:/tmp/TeaTime.png"  
            ,bg ="transparent")  
            <R Code for a graphic>  
        dev.off()  
    ENDSUBMIT;  
QUIT;  
  
ODS ESCAPECHAR='^';  
ODS HTML FILE='c:\tmp\TeaTime.html' STYLE=minimal  
GPATH='c:\temp\' GTITLE GFOOTNOTE;  
  
DATA _NULL_;  
    FILE PRINT;  
    PUT "<IMG SRC='c:\tmp\TeaTime.png' BORDER='0'>";  
RUN;
```

Get Graphics from R to SAS



Handling Errors from R

Error: Division by a alphanumeric variable:

```
PROC IML;  
  SUBMIT / R;  
    1/"5"  
    print("Here I am in R.")  
  ENDSUBMIT;  
QUIT;
```

1. have a look to the SAS LOG

Handling Errors from R

2. read out the global SAS macro variable
SYSERRORTXT (holds the last error message)

```
DATA _NULL_;  
    IF ("&SYSERRORTXT." NE "") THEN DO;  
        PUT "Error in R: &SYSERRORTXT.";  
    END;  
RUN;
```

```
Error in R: R: non-numeric argument to binary operator
```

Handling Errors from R

3. Use tryCatch-block in R and write the error message to a SAS data set

```
PROC IML;
  SUBMIT / R;
    errMess = ""
    tryCatch({
      x = 1/"5";
    },
    error = function(ex) {
      errMess = geterrmessage()
      errMess <- as.data.frame(errMess)
    },
    finally = {
      cat("The error was handled \n")
    }
  )
ENDSUBMIT;
RUN ImportDataSetFromR("WORK.ErrMessFromR", "errMess");
QUIT;
```

Handling Errors from R

3. Use tryCatch-block in R and write the error message to a SAS data set

```
DATA _NULL_ ;  
    SET WORK.ErrMessFromR END=LAST;  
    IF LAST THEN DO;  
        CALL SYMPUT('errMessInSAS', errMess);  
    END;  
RUN;  
%PUT Error-Message from R: &errMessInSAS.;
```

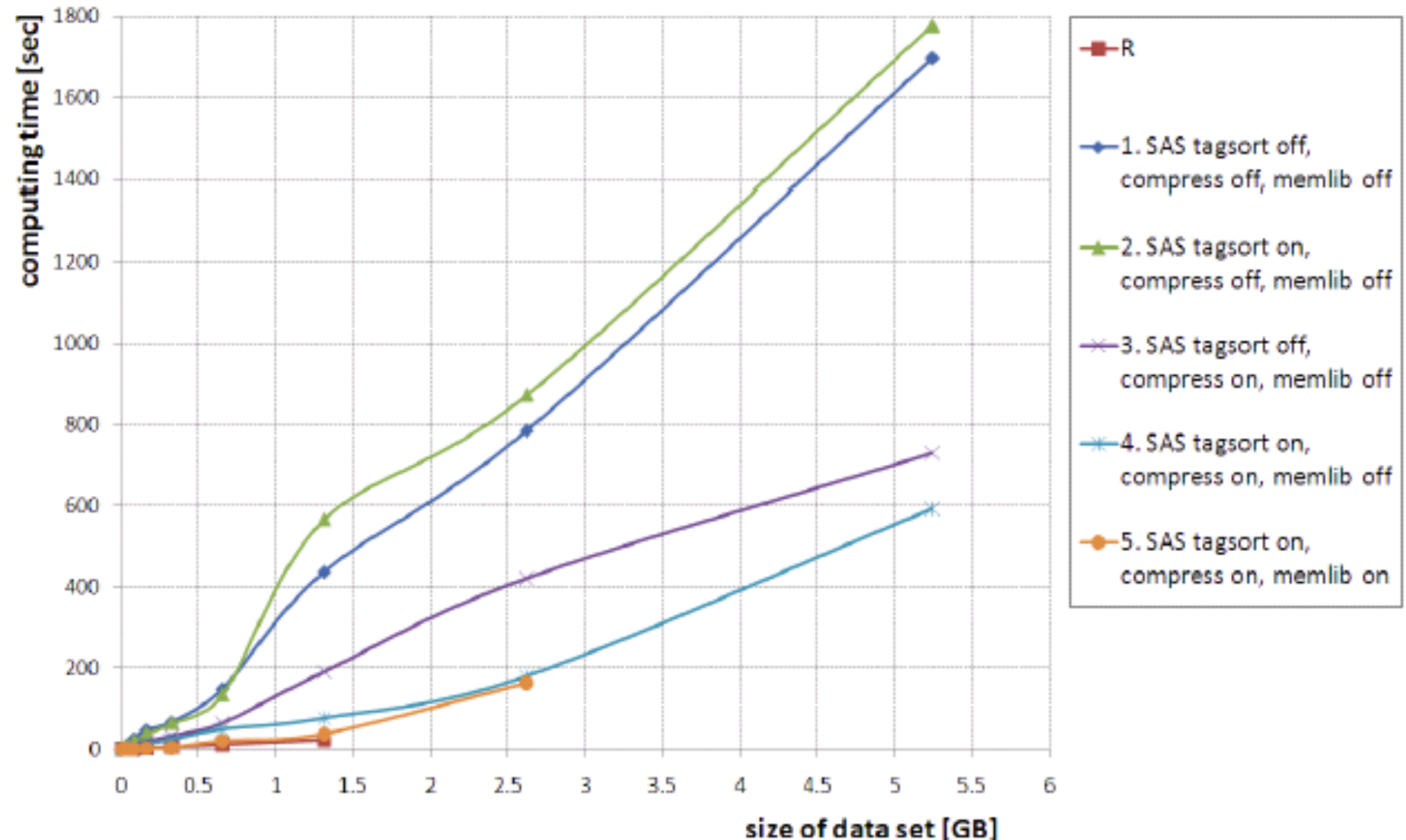
Error-Message from R: non-numeric argument to binary operator

Limitations

- No option to tell SAS which R to use.
- Large data sets are limited by memory size
 - In R the data are load and processed in the main memory. The same is true for SAS IML.
- R comes along with absolutely no warranty
 - take this especially into account for the large amount of add-on-packages.
- Not all data types from R can be converted
 - every data type has to be converted either to a *data.frame* or a *matrix* before transferring data to SAS

Questions during PhUSE...

- What is faster SAS or R?



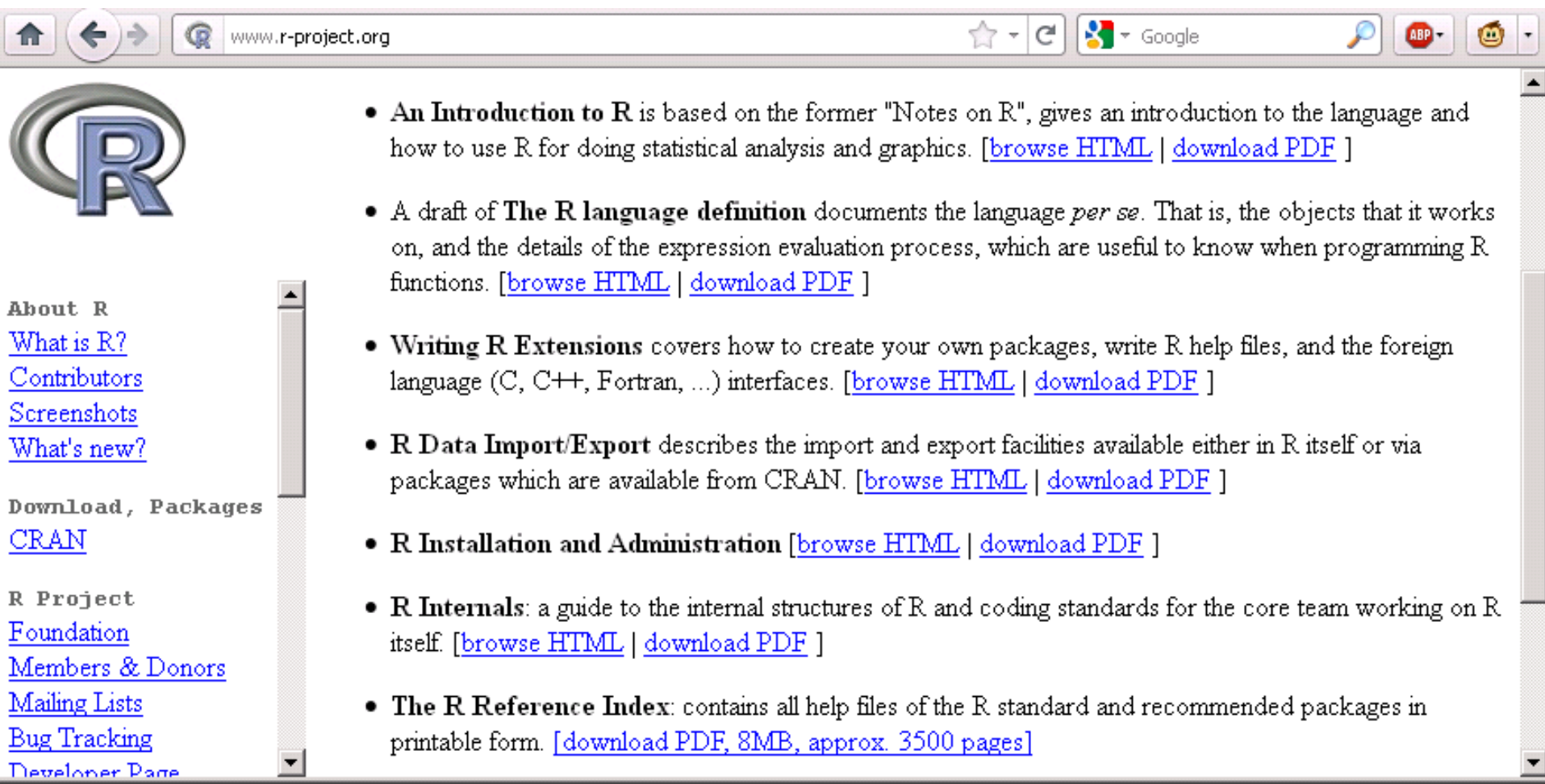
Questions during PhUSE...

- Is PROC IML the only way to run R programs?

```
OPTIONS XWAIT XSYNC;  
X "r.exe --no-save --quiet  
  <""c:\temp\program.r""  
  > ""c:\temp\program.log""";
```

Questions during PhUSE...

- Is R easy to learn?



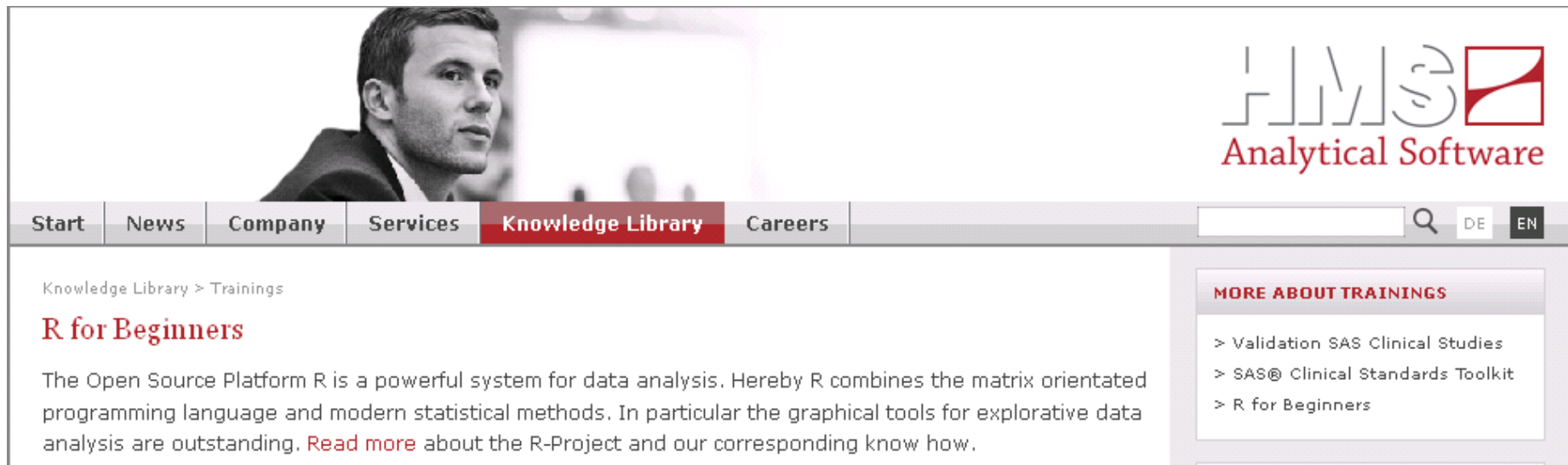
The screenshot shows the R Project website in a web browser. The browser's address bar displays 'www.r-project.org'. The website features a large R logo on the left. A sidebar on the left contains links for 'About R', 'What is R?', 'Contributors', 'Screenshots', 'What's new?', 'Download, Packages', 'CRAN', 'R Project', 'Foundation', 'Members & Donors', 'Mailing Lists', 'Bug Tracking', and 'Developer Page'. The main content area lists several resources with links to browse HTML or download PDF files.

- **An Introduction to R** is based on the former "Notes on R", gives an introduction to the language and how to use R for doing statistical analysis and graphics. [[browse HTML](#) | [download PDF](#)]
- A draft of **The R language definition** documents the language *per se*. That is, the objects that it works on, and the details of the expression evaluation process, which are useful to know when programming R functions. [[browse HTML](#) | [download PDF](#)]
- **Writing R Extensions** covers how to create your own packages, write R help files, and the foreign language (C, C++, Fortran, ...) interfaces. [[browse HTML](#) | [download PDF](#)]
- **R Data Import/Export** describes the import and export facilities available either in R itself or via packages which are available from CRAN. [[browse HTML](#) | [download PDF](#)]
- **R Installation and Administration** [[browse HTML](#) | [download PDF](#)]
- **R Internals**: a guide to the internal structures of R and coding standards for the core team working on R itself. [[browse HTML](#) | [download PDF](#)]
- **The R Reference Index**: contains all help files of the R standard and recommended packages in printable form. [[download PDF, 8MB, approx. 3500 pages](#)]

Questions during PhUSE...

- Is R easy to learn?

2-day-training course, HMS



The screenshot shows the HMS Analytical Software website. The header features a navigation bar with links: Start, News, Company, Services, Knowledge Library (highlighted), and Careers. A search bar and language selectors (DE, EN) are also present. The main content area is titled 'Knowledge Library > Trainings' and features a section for 'R for Beginners'. The text describes R as a powerful system for data analysis, combining matrix-oriented programming with modern statistical methods. A 'Read more' link is provided. A sidebar on the right titled 'MORE ABOUT TRAININGS' lists three courses: 'Validation SAS Clinical Studies', 'SAS® Clinical Standards Toolkit', and 'R for Beginners'.

Questions during PhUSE...

Is are a validated environment?

- no, it is NOT: R comes along with absolutely no warranty



Conclusion

- The best of two worlds.
- SAS implements a powerful interface to R within PROC IML.
- SAS data sets can be send to R and back.
- Some SAS formats are converted to R counterpart.
- R graphics can be integrated into SAS PDF-reports.
- Different ways to handle errors from R in SAS.

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