Es kann nur einen geben? PROC REPORT von SAS vs. FLEXTABLE von R – ein Vergleich

14. und 15. September 2022 KSFE 2022

Autorin: Fenja Seither, fenja.seither@mainanalytics.de

Gliederung

Einleitung

Technische Voraussetzungen

Gegenüberstellung beider Ansätze

Vorteile SAS PROC REPORT

Vorteile R FLEXTABLE

Das Beste aus beiden Ansätzen

Einleitung

SAS

- Jahrzehntelang bevorzugte Statistiksoftware
- Effektive Funktionen wie PROC REPORT/TABULATE

R

- Immer größere Verbreitung der Open-Source-Software
- Flexibles Paket FLEXTABLE

Gliederung

Technische Voraussetzungen

- R / R Studio (Pakete: flextable, officer, magrittr)
- SAS 9.4 (Base)

Gliederung

Einleitung

Technische Voraussetzungen

Gegenüberstellung beider Ansätze

Vorteile SAS PROC REPORT

Vorteile R FLEXTABLE

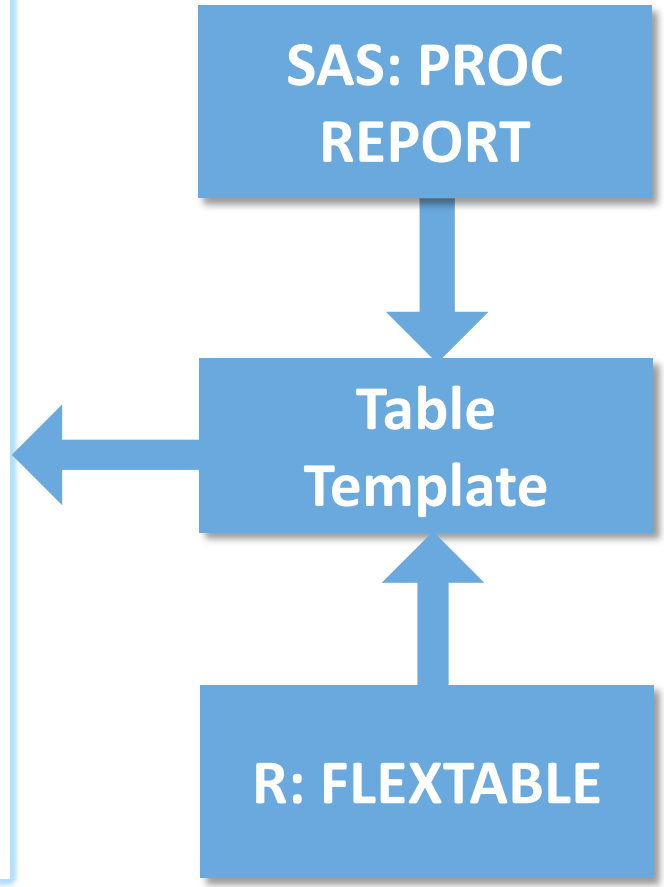
Das Beste aus beiden Ansätzen

Gegenüberstellung beider Ansätze

x.1: Time to first event: Cox Regression of outcome <event> by groups

Group	*Placebo*	*Treat1*	*<Treat k>*
<Group1>			
Number of subjects (N, %)	xxx (100 %)	xxx (100 %)	xxx (100 %)
Subjects with events (N, %)	xxx (xx.x %)	xxx (xx.x %)	xxx (xx.x %)
Subjects with censors (N, %)	xxx (xx.x %)	xxx (xx.x %)	xxx (xx.x %)
Comparison vs Placebo			
Hazard Ratio (adjusted#)		xx.xx	xx.xx
95% Confidence interval		(xx.xx,xx.xx)	(xx.xx,xx.xx)
P-value		x.xxxx	x.xxxx
< Group2……..>			
Number of subjects (N, %)	xxx (100 %)	xxx (100 %)	xxx (100 %)
Subjects with events (N, %)	xxx (xx.x %)	xxx (xx.x %)	xxx (xx.x %)
Subjects with censors (N, %)	xxx (xx.x %)	xxx (xx.x %)	xxx (xx.x %)
Comparison vs Placebo			
Hazard Ratio (adjusted#)		xx.xx	xx.xx
95% Confidence interval		(xx.xx,xx.xx)	(xx.xx,xx.xx)
P-value		x.xxxx	x.xxxx

#Adjustment is for sex, age and treatment.



Gegenüberstellung beider Ansätze

```
proc phreg data=coxdata alpha=0.05;
    by BYVAR;
    class TREAT (ref="Placebo") SEX RACE / param=glm order=internal;
    model AVAL*CNSR(0) = TREAT SEX RACE / risklimits ;
    lsmeans TREAT/ pdiff=all cl alpha = 0.05 exp;
run;
```

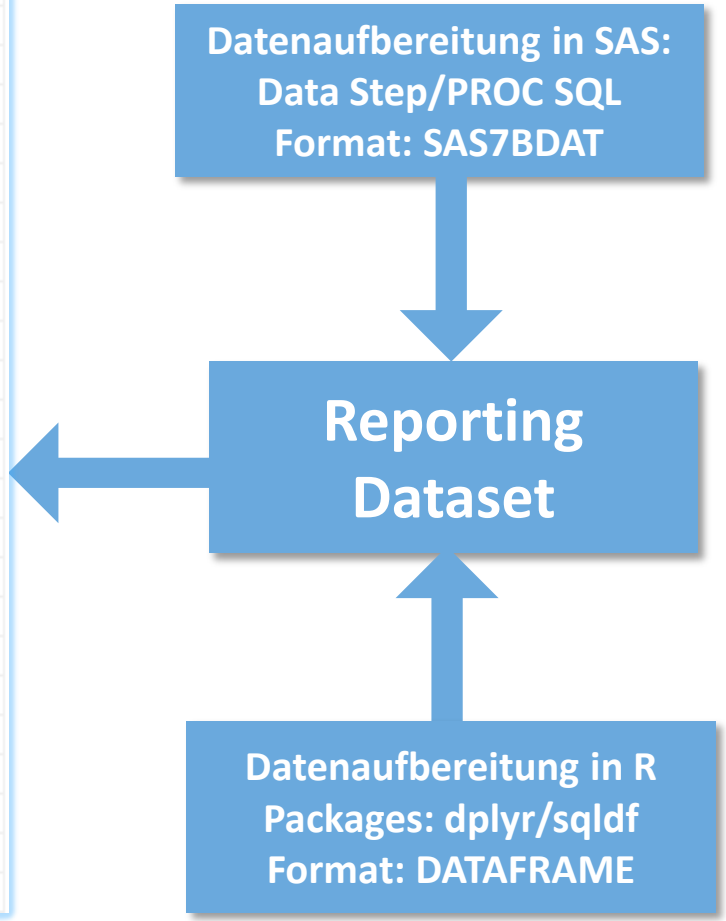
Verwendeter SAS Code

```
install.packages(c("survival", "survminer"))
library(survival)
library(survminer)
coxresult<-  lapply(split(coxdata, coxdata$BYVAR), function(x)
             coxph(Surv(AVAL, CNSR==0) ~ (TREAT + SEX + RACE),data = x))
```

Verwendeter R Code

Gegenüberstellung beider Ansätze

Age Group (N)	Group	Placebo	Xanomeline Dose	Page order
1	Age Group: <65			1
1	Number of subjects	14 (100.0)	19 (100.0)	1
1	Subjects with events	5 (35.7)	17 (89.5)	1
1	Subjects with censors	9 (64.3)	2 (10.5)	1
1	Comparison vs Placebo			1
1	Hazard Ratio (adjusted#)		3.78	1
1	95% Confidence interval		(0.64,22.24)	1
1	P-value		0.1418	1
2	Age Group: 65-80			1
2	Number of subjects	42 (100.0)	102 (100.0)	1
2	Subjects with events	14 (33.3)	79 (77.5)	1
2	Subjects with censors	28 (66.7)	23 (22.5)	1
2	Comparison vs Placebo			1
2	Hazard Ratio (adjusted#)		1.79	1
2	95% Confidence interval		(1.13,2.81)	1
2	P-value		0.0124	1
3	Age Group: >80			2
3	Number of subjects	30 (100.0)	47 (100.0)	2
3	Subjects with events	10 (33.3)	27 (57.4)	2
3	Subjects with censors	20 (66.7)	20 (42.6)	2
3	Comparison vs Placebo			2
3	Hazard Ratio (adjusted#)		2.27	2
3	95% Confidence interval		(1.37,3.77)	2
3	P-value		0.0016	2



Gegenüberstellung beider Ansätze

SAS Methode

```
title "x.1: Time to first event: Cox Regression of outcome event by age groups";
proc report data=final nowd missing headskip headline spacing=2 split='~';
    column byord catc col_0 col_1;

    define byord     /  order order=internal noprint;
    define catc      /  display width=40 left flow spacing=3;
    define col_0     /  order width=22 left flow spacing=1;
    define col_1     /  order width=22 left flow spacing=1;

    break after byord/skip;

    compute before _page_;
          line @1 120*'ƒ';
     endcomp;

     compute after _page_;
          line @1 120*'ƒ';
          line @1 "#Adjustment is for sex, race, and treatment.";
     endcomp;
run;
```

R Methode

```
report <- flextable(final) %>%
  # Width of columns
  width(j = 2:3, width = 2) %>% width(j = 1, width = 2.5) %>%
  # Alignment of columns
  align(j = 2:3, align = "center", part = "all") %>%
  align(j = 1, align = "left", part = "all") %>%
  # Add titles
  add_header_lines(values =
    paste("x.1: Time to first event: Cox Regression of outcome event
          by age groups", sep = "")) %>%
  # Add footnotes
  add_footer_lines(values =
    paste("#Adjustment is for sex, race, and treatment.", sep = "")) %>%
  # Line at bottom of the table
  hline_bottom(border = fp_border(color = "black"), part = "body") %>%
  # Line below of column names
  hline_bottom(border = fp_border(color = "black"), part = "header") %>%
  # Line above column names
  border_inner_h(border = fp_border("black"), part = "header")
```

Gegenüberstellung beider Ansätze

SAS Methode

x.1: Time to first event: Cox Regression of outcome event by age groups

Group	Placebo	Xanomeline Dose
Age Group: <65		
Number of subjects	14 (100.0)	19 (100.0)
Subjects with events	5 (35.7)	17 (89.5)
Subjects with censors	9 (64.3)	2 (10.5)
Comparison vs Placebo		
Hazard Ratio (adjusted#)		3.78
95% Confidence interval		(0.64,22.24)
P-value		0.1418
Age Group: 65-80		
Number of subjects	42 (100.0)	102 (100.0)
Subjects with events	14 (33.3)	79 (77.5)
Subjects with censors	28 (66.7)	23 (22.5)
Comparison vs Placebo		
Hazard Ratio (adjusted#)		1.79
95% Confidence interval		(1.13,2.81)
P-value		0.0124
Age Group: >80		
Number of subjects	30 (100.0)	47 (100.0)
Subjects with events	10 (33.3)	27 (57.4)
Subjects with censors	20 (66.7)	20 (42.6)
Comparison vs Placebo		
Hazard Ratio (adjusted#)		2.27
95% Confidence interval		(1.37,3.77)
P-value		0.0016

#Adjustment is for sex, race, and treatment.

R Methode

x.1: Time to first event: Cox Regression of outcome event by age groups

Group	Placebo	Xanomeline Dose
Age Group: <65		
Number of subjects	14 (100.0)	19 (100.0)
Subjects with events	5 (35.7)	17 (89.5)
Subjects with censors	9 (64.3)	2 (10.5)
Comparison vs Placebo		
Hazard Ratio (adjusted#)		3.78
95% Confidence interval		(0.64,22.24)
P-value		0.1418
Age Group: 65-80		
Number of subjects	42 (100.0)	102 (100.0)
Subjects with events	14 (33.3)	79 (77.5)
Subjects with censors	28 (66.7)	23 (22.5)
Comparison vs Placebo		
Hazard Ratio (adjusted#)		1.79
95% Confidence interval		(1.13,2.81)
P-value		0.0124
Age Group: >80		
Number of subjects	30 (100.0)	47 (100.0)
Subjects with events	10 (33.3)	27 (57.4)
Subjects with censors	20 (66.7)	20 (42.6)
Comparison vs Placebo		
Hazard Ratio (adjusted#)		2.27
95% Confidence interval		(1.37,3.77)
P-value		0.0016

#Adjustment is for sex, race, and treatment.

Gegenüberstellung beider Ansätze

Scope	PROC REPORT option/statement	FLEXTABLE function
Aligning columns	center/left/right options in the define statement	align function
Width of columns	width option in the define statement	width function
Line above column names	'___' in the column statement (see proc report on page 2)	border_inner_h function
Line below column names	headline option in proc report statement	
Line at the bottom of the table	compute after _page_/<page- variable> block	hline_bottom function
Add footnotes/titles	footnoten/titlen statement /line statement in a compute after block (footnotes only)	add_header_lines (titles) add_footer_lines (footnotes)
Multicolumn header	Embed columns in brackets: ("header" col1 col2)	add_header_row using the colwidths option
Skip lines	break after <var> / skip	none
Page break	break after <var> / page	body_add_break (from the officer package)

Vergleichbare Optionen gibt es in FLEXTABLE nicht

Gliederung

Einleitung

Technische Voraussetzungen

Gegenüberstellung beider Ansätze

Vorteile SAS PROC REPORT

Vorteile R FLEXTABLE

Das Beste aus beiden Ansätzen

Vorteile von SAS PROC REPORT

PROC REPORT

```
break after pagen/PAGE;
```

Weitere Funktion nötig: FLEXTABLE mit OFFICER Package

```
page.break <- function(data, pagevar = "page"){
  #Loop for number of pages
  for (i in 1:n) {

    #Filter object dataset for per page
    subdata<- subset(data, page==i)
    report <- flextable(subdata)

    # Write first page
    if (i == 1) {
      doc <- read_docx() %>% body_add_flextable(value = report)
      print(doc, target = "Table_cox.docx")}
    #Add pages with previous pages
    else {
      doc <- read_docx("Table_cox.docx") %>%
      body_add_break() %>% body_add_flextable(value = report)
      print(doc, target = "Table_cox.docx")} }}
```

Vorteile von SAS PROC REPORT

PROC REPORT

```
proc report data=disbp nowd missing headskip headline;
    column avisitn avisit grouplabel trtp
           disbp,("--" N MIN MEAN STD MEDIAN MAX);

    define avisitn     /  group noprint;
    define avisit      /  group "Week" width=12;
    define grouplabel  /  group "Types of value" width=20;
    define trtp        /  group width=20;
    define disbp       /  analysis N width=11 format=8.1;

    break after grouplabel/skip;
    break after avisitn/skip;
run;
```

FLEXTABLE

- Berechnungen innerhalb der Funktion nicht möglich
- Alle Analysen müssen vor der Erstellung des Reports durchgeführt werden

Gliederung

Einleitung

Technische Voraussetzungen

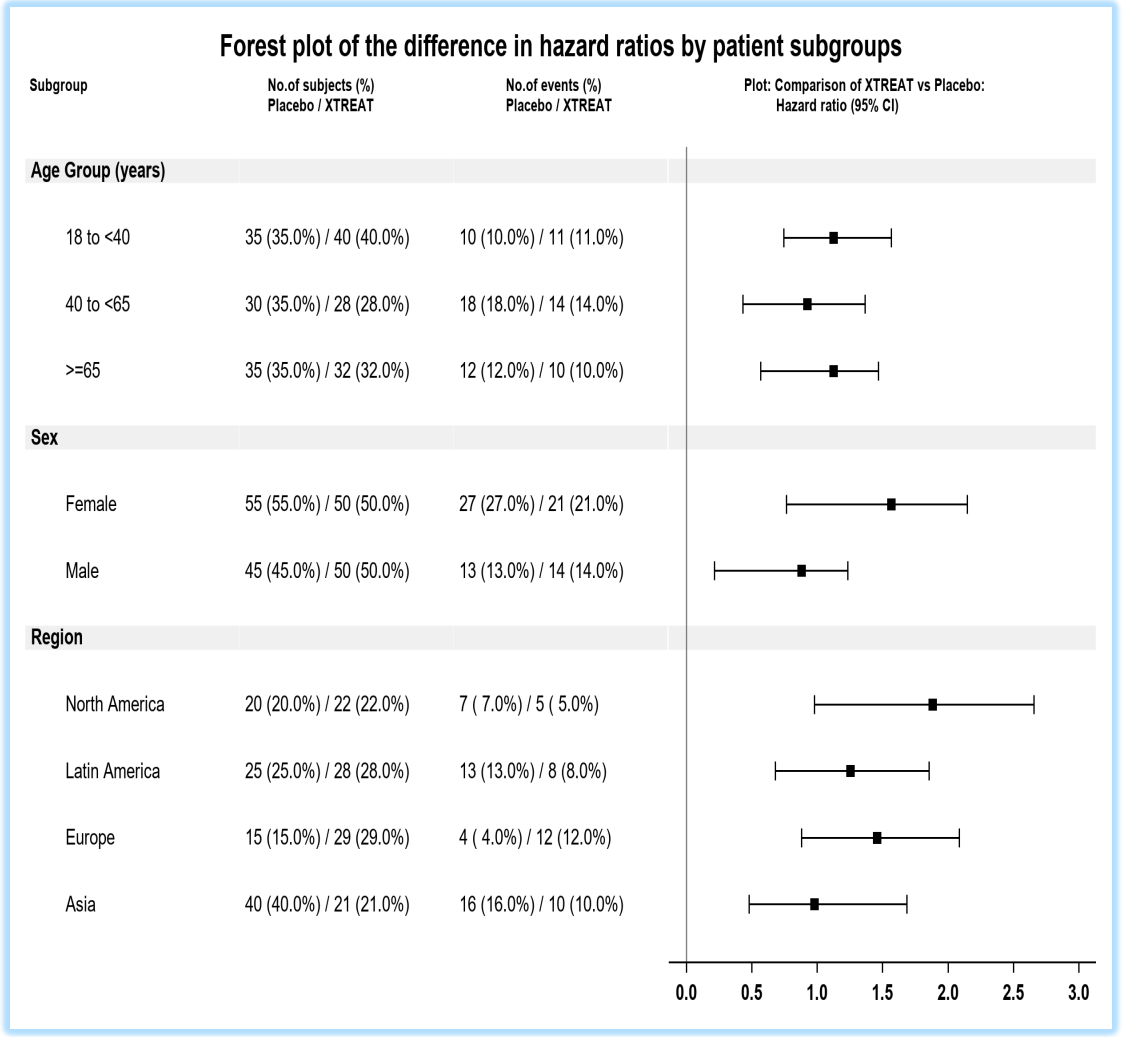
Gegenüberstellung beider Ansätze

Vorteile SAS PROC REPORT

Vorteile R FLEXTABLE

Das Beste aus beiden Ansätzen

Vorteile von FLEXTABLE in R

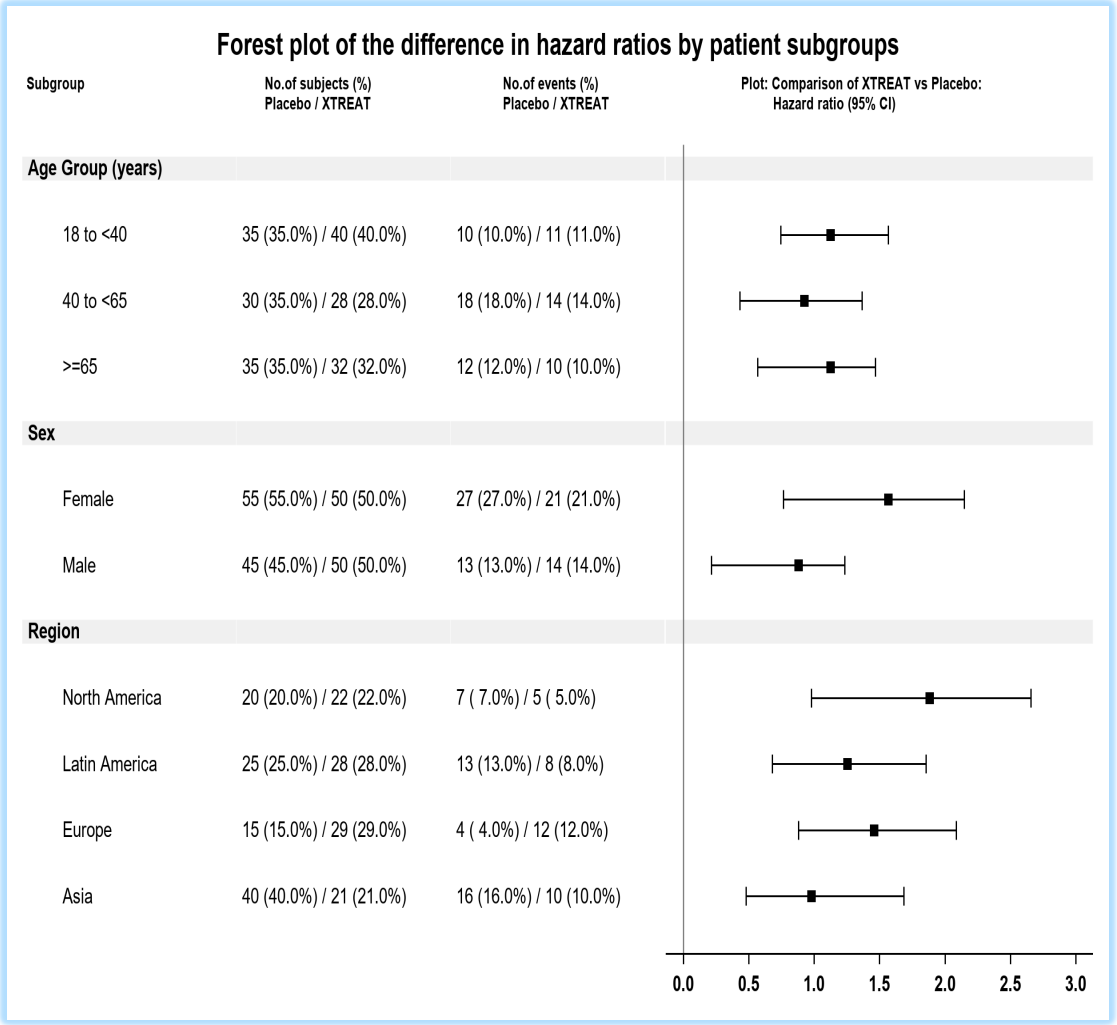


FLEXTABLE with ggplot2

```
dat <- forestdt
dat <- dat %>%
  tidyr::nest(data = -c(group, n, nevents )) %>%
  dplyr::mutate(
    #Create forest plots with ggplots
    forest = lapply(data, function(x)
             gg(x,
                cols = c("estimate", "lower", "upper", "groupn")))) %>%
    dplyr::select(-data)

ft <- flextable(dat[, c("group", "n", "nevents", "forest")])
ft <- compose(ft,
              #Render Forest Plots
              j = c("forest"),
              value = as_paragraph
                      (gg_chunk(value = ., height =1, width = 3)),
              use_dot = TRUE)
```

Vorteile von FLEXTABLE in R



Forest plot of the difference in hazard ratios by patient subgroups

Subgroup	No.of subjects (%) Placebo / XTREAT	No.of events (%) Placebo / XTREAT
Age Group (years)		
18 to <40	35 (35.0%) / 40 (40.0%)	10 (10.0%) / 11 (11.0%)
40 to <65	30 (35.0%) / 28 (28.0%)	18 (18.0%) / 14 (14.0%)
>=65	35 (35.0%) / 32 (32.0%)	12 (12.0%) / 10 (10.0%)
Sex		
Female	55 (55.0%) / 50 (50.0%)	27 (27.0%) / 21 (21.0%)
Male	45 (45.0%) / 50 (50.0%)	13 (13.0%) / 14 (14.0%)
Region		
North America	20 (20.0%) / 22 (22.0%)	7 (7.0%) / 5 (5.0%)
Latin America	25 (25.0%) / 28 (28.0%)	13 (13.0%) / 8 (8.0%)
Europe	15 (15.0%) / 29 (29.0%)	4 (4.0%) / 12 (12.0%)
Asia	40 (40.0%) / 21 (21.0%)	16 (16.0%) / 10 (10.0%)

Plot: Comparison of XTREAT vs Placebo: Hazard ratio (95% CI)

0.0 0.5 1.0 1.5 2.0 2.5 3.0

PROC REPORT

- Erstellen von Grafiken nicht möglich
- Einsatz der SAS Graph Template Language (GTL) nötig

Gliederung

Einleitung

Technische Voraussetzungen

Gegenüberstellung beider Ansätze

Vorteile SAS PROC REPORT

Vorteile R FLEXTABLE

Das Beste aus beiden Ansätzen

Das Beste aus beiden Ansätzen

Entscheidung zwischen PROC REPORT und FLEXTABLE abhängig von

- **Gewünschter Darstellung**
- **Programmieraufwand**

Das Beste aus beiden Ansätzen

SAS

- Vorteile von PROC IML nutzbar

```
PROC IML;
    SUBMIT / R;
        a <- 2
        paste("a:",a)
        b <- 3
        paste("b:",b)
        c <- a * b
        paste("c:",c);
    ENDSUBMIT;
QUIT;
```

R

- Konsolenbefehle nutzen

```
callSAS <- function (f, path=getwd(), workDir=sprintf("E:", Sys.info()["user"])) {
  # Usage:
  # > callSAS("getClaimsMed.sas")
  if (grepl(".sas$", f)) {
    exeFile <- "E:\\Apps\\SASHome\\SASFoundation\\9.4\\sas.exe"
    sasFile <- sprintf("%s\\%s", path, f)
    logFile <- sprintf("%s\\%s", path, gsub(".sas$", ".log", f))
    lstFile <- sprintf("%s\\%s", path, gsub(".sas$", ".lst", f))
    cmd <- sprintf("\"%s\" -sysin \"%s\" -log \"%s\" -print \"%s\"",
                   exeFile,
                   sasFile,
                   logFile,
                   lstFile
                   )
    system(cmd, invisible=FALSE)
    cat(readLines(sasFile), sep="\n")
  }
}
```

Danke für Ihre Aufmerksamkeit