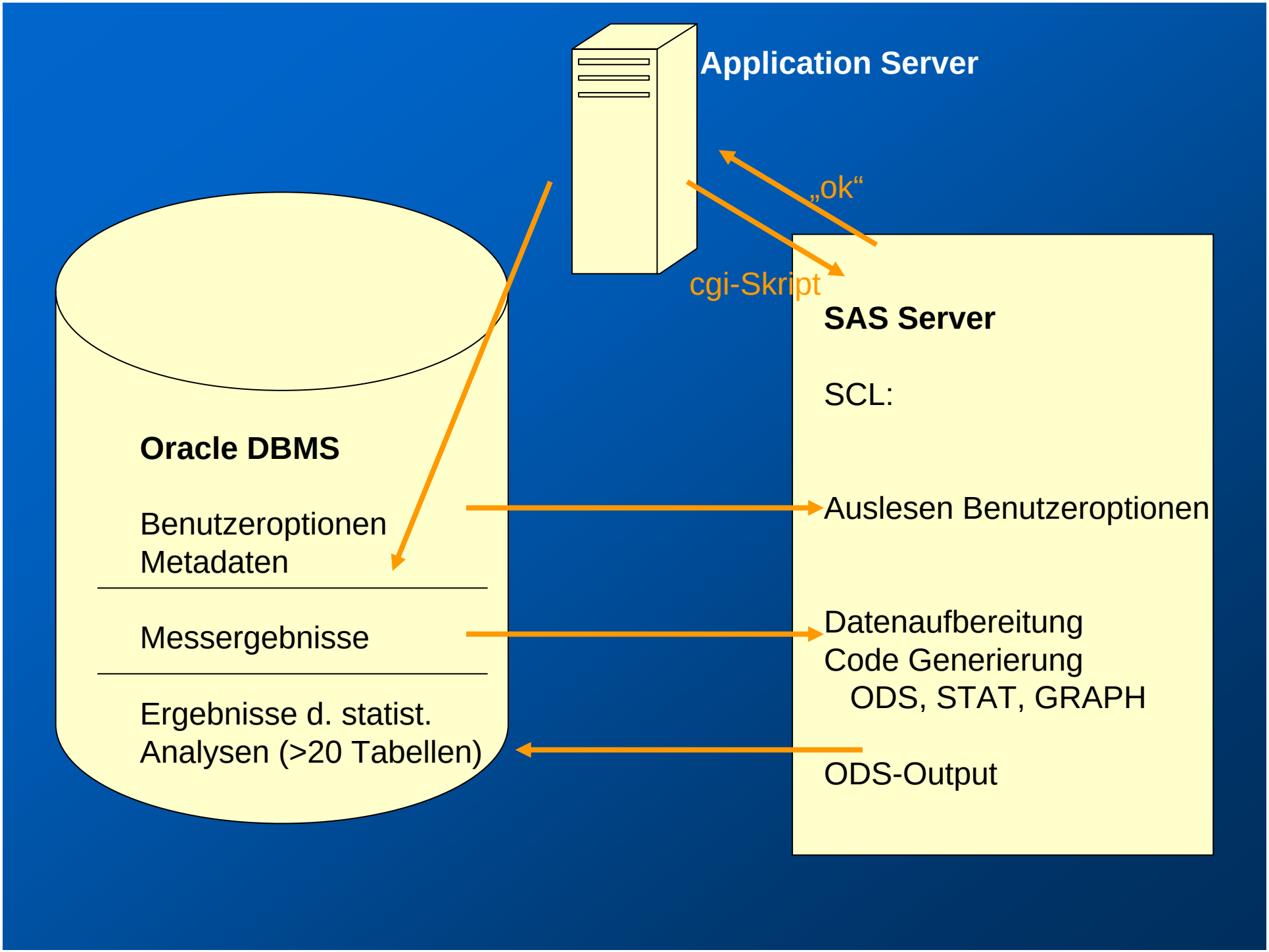
Ein System zur Erfassung, Speicherung und Auswertung landwirtschaftlicher Versuche – Einsatz von SAS in Datenbankzugriff und statistischer Analyse

Teil II: die SAS-Komponente

01.03.2002 Dr. Catherine Janson Info Ware GmbH


Application Server
„ok“
cgi-Skript
SAS Server
SCL:
Oracle DBMS
Benutzeroptionen
Metadaten
Auslesen Benutzeroptionen
Messergebnisse
Datenaufbereitung
Code Generierung
ODS, STAT, GRAPH
Ergebnisse d. statist.
Analysen (>20 Tabellen)
ODS-Output

Funktionalität

GLM / MIXED
> 28 varianzanalytische Modelle:
1-, 2-,3-faktoriell,
vollständig randomisierte Anlage, randomisierte Blockanlage, Spaltanlage(n), Streifenanlage(n)
LSMEANS-Test
Multiple Mittelwertvergleiche: Bonferroni, Scheffe, Tukey, Dunnett, Student-Newman-Keuls

CORR – Korrelation
REGR – Regression
NPAR1WAY – Wilcoxon / Kruskal-Wallis Vergleich d. Treatments
FREQ – Häufigkeitsanalyse

GPLOT – Box-Wkisker-Plots (CGM, RTF)

Permutationstests (SAS/IML Makros Prof. E. Schumacher, Univ. Hohenheim)

Drei besondere Aspekte der SAS-Programmierung

Einsatz von Steuertabellen

ODS Output

Performance

Einsatz von Steuertabellen

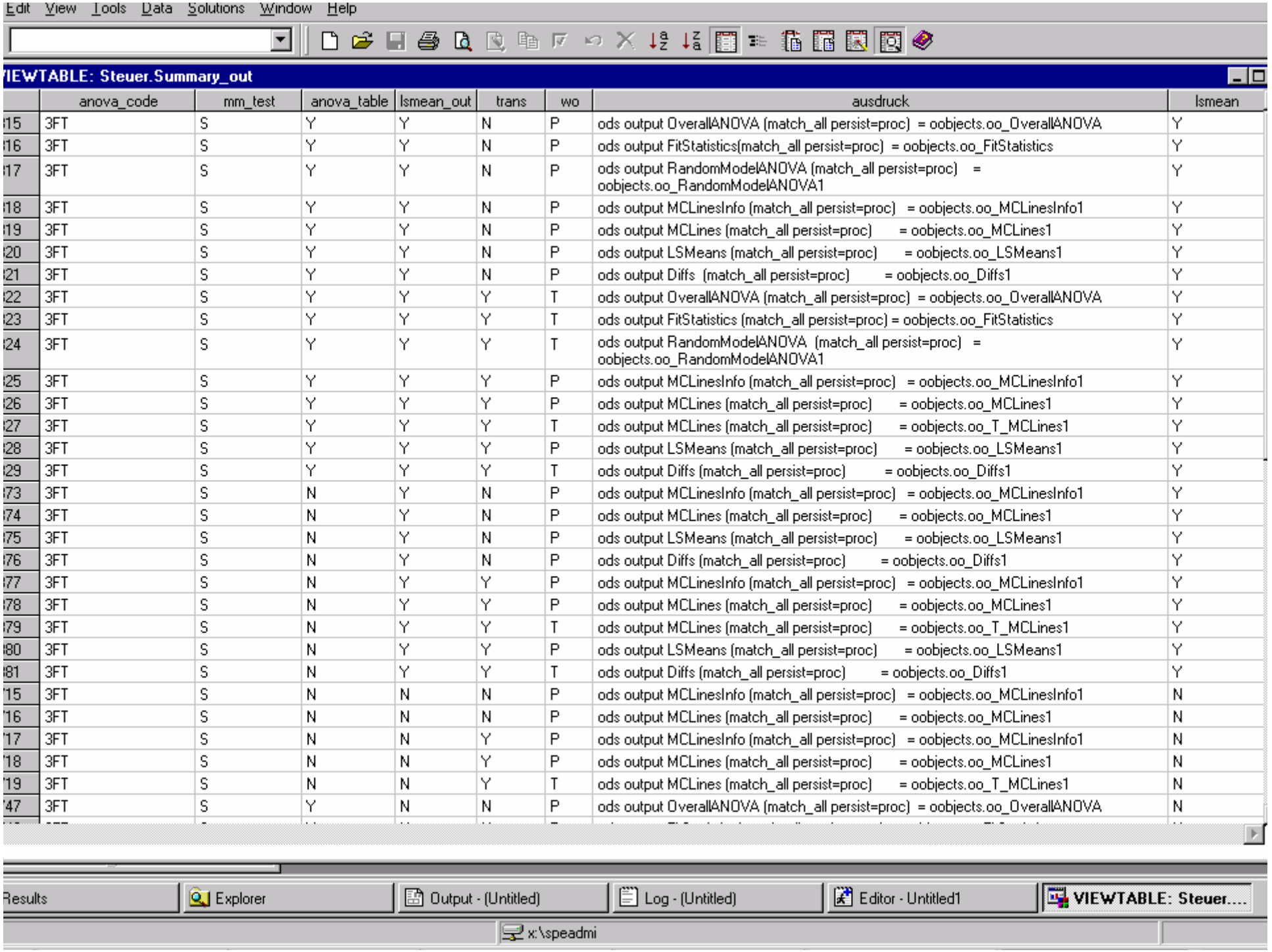
bei extrem komplexer Fallunterscheidung

Programmcode bleibt übersichtlich

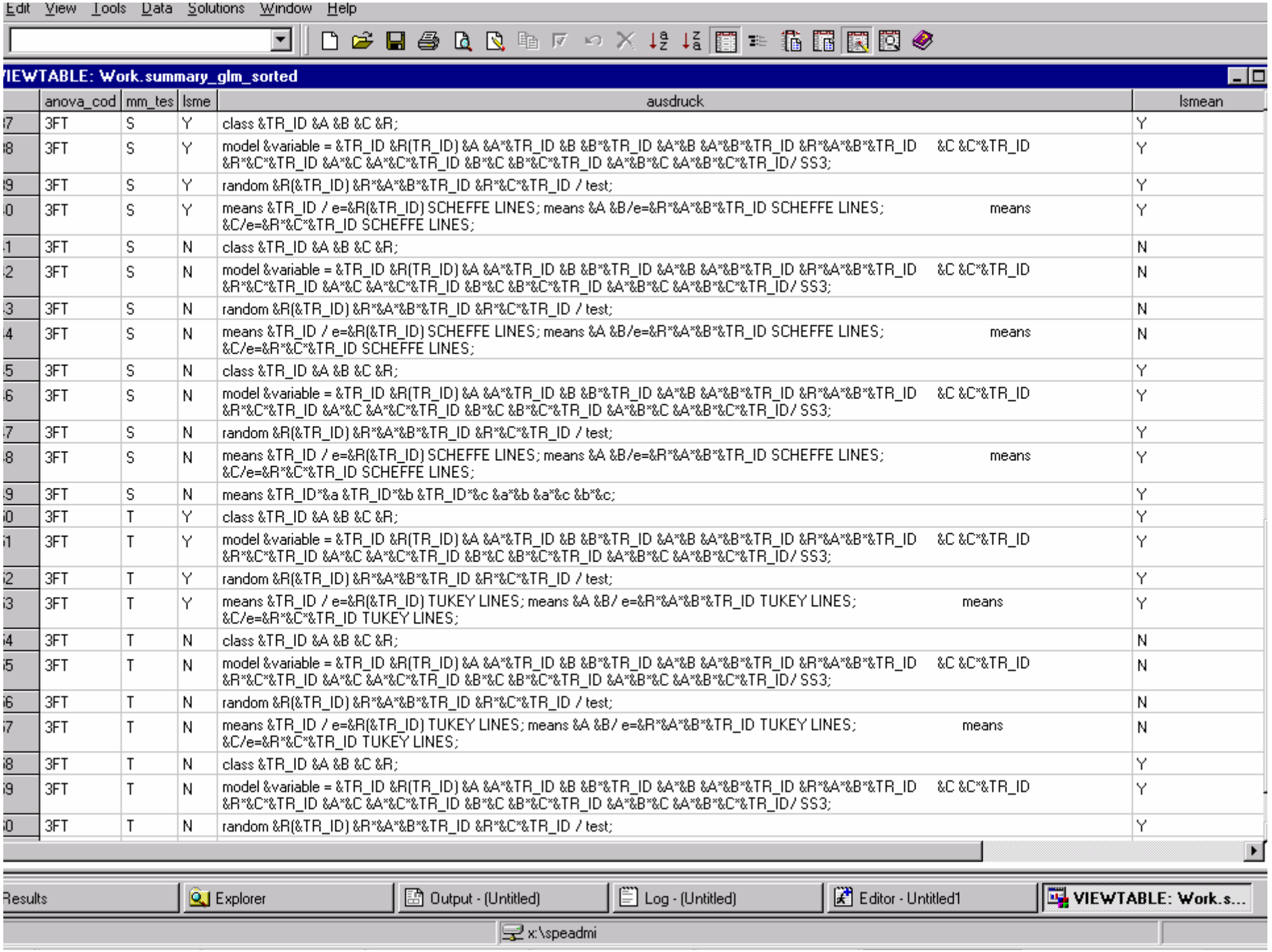
Flexibilität (z.B. Hinzunahme eines neuen Modells)

Steuerung von:

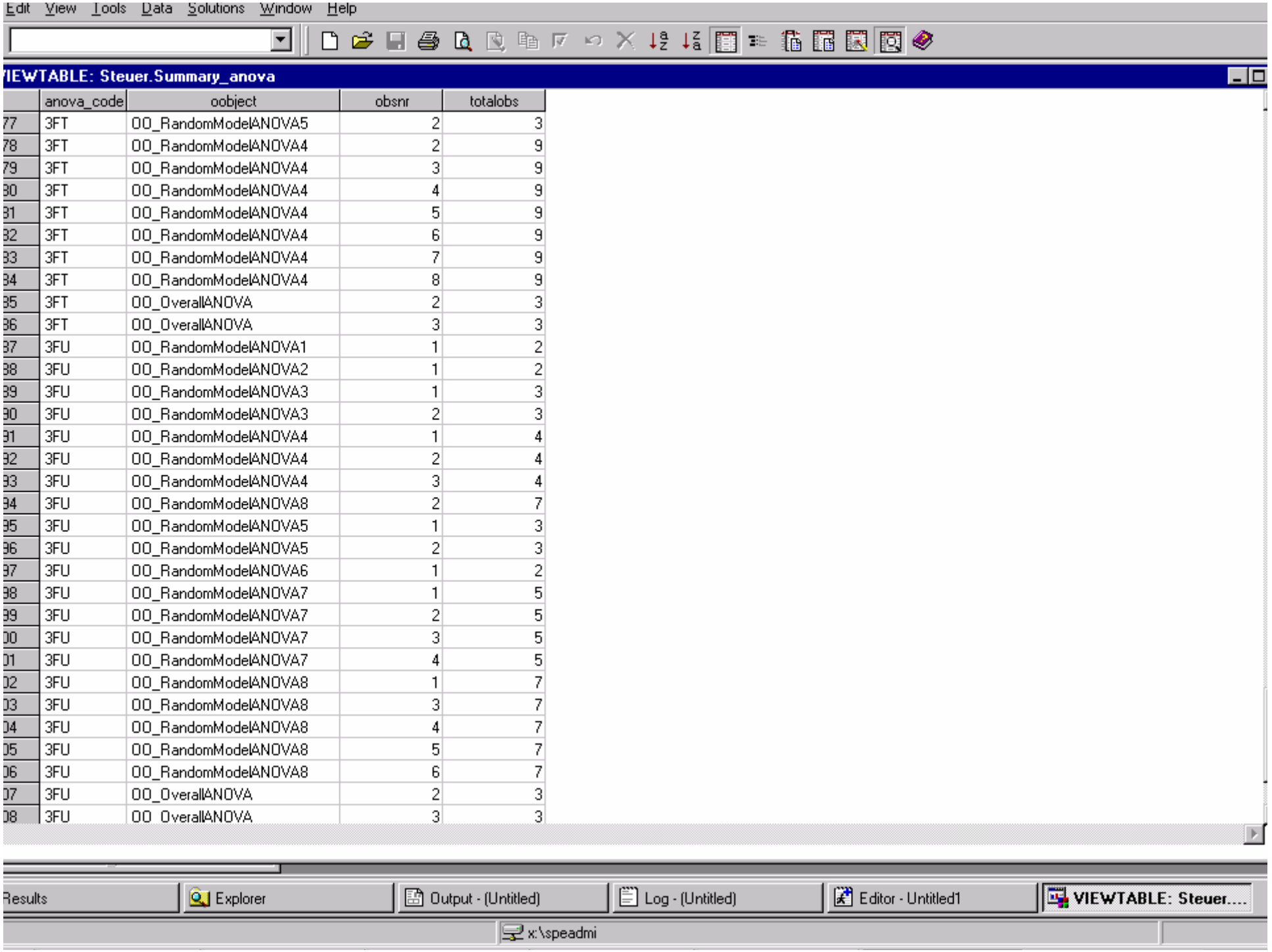
- **ODS-Statements**
- **Code-Generierung SAS/STAT**
- **Zeilenreihenfolge der ANOVA-Ausgabe**



	anova_code	mm_test	anova_table	lsmean_out	trans	wo	ausdruck	lsmean
15	3FT	S	Y	Y	N	P	ods output OverallANOVA (match_all persist=proc) = oobjects.oo_OverallANOVA	Y
16	3FT	S	Y	Y	N	P	ods output FitStatistics(match_all persist=proc) = oobjects.oo_FitStatistics	Y
17	3FT	S	Y	Y	N	P	ods output RandomModelANOVA (match_all persist=proc) = oobjects.oo_RandomModelANOVA1	Y
18	3FT	S	Y	Y	N	P	ods output MCLinesInfo (match_all persist=proc) = oobjects.oo_MCLinesInfo1	Y
19	3FT	S	Y	Y	N	P	ods output MCLines (match_all persist=proc) = oobjects.oo_MCLines1	Y
20	3FT	S	Y	Y	N	P	ods output LSMeans (match_all persist=proc) = oobjects.oo_LSMeans1	Y
21	3FT	S	Y	Y	N	P	ods output Diffs (match_all persist=proc) = oobjects.oo_Diffs1	Y
22	3FT	S	Y	Y	Y	T	ods output OverallANOVA (match_all persist=proc) = oobjects.oo_OverallANOVA	Y
23	3FT	S	Y	Y	Y	T	ods output FitStatistics (match_all persist=proc) = oobjects.oo_FitStatistics	Y
24	3FT	S	Y	Y	Y	T	ods output RandomModelANOVA (match_all persist=proc) = oobjects.oo_RandomModelANOVA1	Y
25	3FT	S	Y	Y	Y	P	ods output MCLinesInfo (match_all persist=proc) = oobjects.oo_MCLinesInfo1	Y
26	3FT	S	Y	Y	Y	P	ods output MCLines (match_all persist=proc) = oobjects.oo_MCLines1	Y
27	3FT	S	Y	Y	Y	T	ods output MCLines (match_all persist=proc) = oobjects.oo_T_MCLines1	Y
28	3FT	S	Y	Y	Y	P	ods output LSMeans (match_all persist=proc) = oobjects.oo_LSMeans1	Y
29	3FT	S	Y	Y	Y	T	ods output Diffs (match_all persist=proc) = oobjects.oo_Diffs1	Y
73	3FT	S	N	Y	N	P	ods output MCLinesInfo (match_all persist=proc) = oobjects.oo_MCLinesInfo1	Y
74	3FT	S	N	Y	N	P	ods output MCLines (match_all persist=proc) = oobjects.oo_MCLines1	Y
75	3FT	S	N	Y	N	P	ods output LSMeans (match_all persist=proc) = oobjects.oo_LSMeans1	Y
76	3FT	S	N	Y	N	P	ods output Diffs (match_all persist=proc) = oobjects.oo_Diffs1	Y
77	3FT	S	N	Y	Y	P	ods output MCLinesInfo (match_all persist=proc) = oobjects.oo_MCLinesInfo1	Y
78	3FT	S	N	Y	Y	P	ods output MCLines (match_all persist=proc) = oobjects.oo_MCLines1	Y
79	3FT	S	N	Y	Y	T	ods output MCLines (match_all persist=proc) = oobjects.oo_T_MCLines1	Y
80	3FT	S	N	Y	Y	P	ods output LSMeans (match_all persist=proc) = oobjects.oo_LSMeans1	Y
81	3FT	S	N	Y	Y	T	ods output Diffs (match_all persist=proc) = oobjects.oo_Diffs1	Y
15	3FT	S	N	N	N	P	ods output MCLinesInfo (match_all persist=proc) = oobjects.oo_MCLinesInfo1	N
16	3FT	S	N	N	N	P	ods output MCLines (match_all persist=proc) = oobjects.oo_MCLines1	N
17	3FT	S	N	N	Y	P	ods output MCLinesInfo (match_all persist=proc) = oobjects.oo_MCLinesInfo1	N
18	3FT	S	N	N	Y	P	ods output MCLines (match_all persist=proc) = oobjects.oo_MCLines1	N
19	3FT	S	N	N	Y	T	ods output MCLines (match_all persist=proc) = oobjects.oo_T_MCLines1	N
47	3FT	S	Y	N	N	P	ods output OverallANOVA (match_all persist=proc) = oobjects.oo_OverallANOVA	N



	anova_cod	mm_tes	lsme	ausdruck	lsmean
7	3FT	S	Y	class &TR_ID &A &B &C &R;	Y
8	3FT	S	Y	model &variable = &TR_ID &R(TR_ID) &A &A*&TR_ID &B &B*&TR_ID &A*&B &A*&B*&TR_ID &R*&A*&B*&TR_ID &C &C*&TR_ID &R*&C*&TR_ID &A*&C &A*&C*&TR_ID &B*&C &B*&C*&TR_ID &A*&B*&C &A*&B*&C*&TR_ID/ SS3;	Y
9	3FT	S	Y	random &R(&TR_ID) &R*&A*&B*&TR_ID &R*&C*&TR_ID / test;	Y
0	3FT	S	Y	means &TR_ID / e=&R(&TR_ID) SCHEFFE LINES; means &A &B/e=&R*&A*&B*&TR_ID SCHEFFE LINES; means &C/e=&R*&C*&TR_ID SCHEFFE LINES;	Y
1	3FT	S	N	class &TR_ID &A &B &C &R;	N
2	3FT	S	N	model &variable = &TR_ID &R(TR_ID) &A &A*&TR_ID &B &B*&TR_ID &A*&B &A*&B*&TR_ID &R*&A*&B*&TR_ID &C &C*&TR_ID &R*&C*&TR_ID &A*&C &A*&C*&TR_ID &B*&C &B*&C*&TR_ID &A*&B*&C &A*&B*&C*&TR_ID/ SS3;	N
3	3FT	S	N	random &R(&TR_ID) &R*&A*&B*&TR_ID &R*&C*&TR_ID / test;	N
4	3FT	S	N	means &TR_ID / e=&R(&TR_ID) SCHEFFE LINES; means &A &B/e=&R*&A*&B*&TR_ID SCHEFFE LINES; means &C/e=&R*&C*&TR_ID SCHEFFE LINES;	N
5	3FT	S	N	class &TR_ID &A &B &C &R;	Y
6	3FT	S	N	model &variable = &TR_ID &R(TR_ID) &A &A*&TR_ID &B &B*&TR_ID &A*&B &A*&B*&TR_ID &R*&A*&B*&TR_ID &C &C*&TR_ID &R*&C*&TR_ID &A*&C &A*&C*&TR_ID &B*&C &B*&C*&TR_ID &A*&B*&C &A*&B*&C*&TR_ID/ SS3;	Y
7	3FT	S	N	random &R(&TR_ID) &R*&A*&B*&TR_ID &R*&C*&TR_ID / test;	Y
8	3FT	S	N	means &TR_ID / e=&R(&TR_ID) SCHEFFE LINES; means &A &B/e=&R*&A*&B*&TR_ID SCHEFFE LINES; means &C/e=&R*&C*&TR_ID SCHEFFE LINES;	Y
9	3FT	S	N	means &TR_ID*&a &TR_ID*&b &TR_ID*&c &a*&b &a*&c &b*&c;	Y
0	3FT	T	Y	class &TR_ID &A &B &C &R;	Y
1	3FT	T	Y	model &variable = &TR_ID &R(TR_ID) &A &A*&TR_ID &B &B*&TR_ID &A*&B &A*&B*&TR_ID &R*&A*&B*&TR_ID &C &C*&TR_ID &R*&C*&TR_ID &A*&C &A*&C*&TR_ID &B*&C &B*&C*&TR_ID &A*&B*&C &A*&B*&C*&TR_ID/ SS3;	Y
2	3FT	T	Y	random &R(&TR_ID) &R*&A*&B*&TR_ID &R*&C*&TR_ID / test;	Y
3	3FT	T	Y	means &TR_ID / e=&R(&TR_ID) TUKEY LINES; means &A &B/ e=&R*&A*&B*&TR_ID TUKEY LINES; means &C/e=&R*&C*&TR_ID TUKEY LINES;	Y
4	3FT	T	N	class &TR_ID &A &B &C &R;	N
5	3FT	T	N	model &variable = &TR_ID &R(TR_ID) &A &A*&TR_ID &B &B*&TR_ID &A*&B &A*&B*&TR_ID &R*&A*&B*&TR_ID &C &C*&TR_ID &R*&C*&TR_ID &A*&C &A*&C*&TR_ID &B*&C &B*&C*&TR_ID &A*&B*&C &A*&B*&C*&TR_ID/ SS3;	N
6	3FT	T	N	random &R(&TR_ID) &R*&A*&B*&TR_ID &R*&C*&TR_ID / test;	N
7	3FT	T	N	means &TR_ID / e=&R(&TR_ID) TUKEY LINES; means &A &B/ e=&R*&A*&B*&TR_ID TUKEY LINES; means &C/e=&R*&C*&TR_ID TUKEY LINES;	N
8	3FT	T	N	class &TR_ID &A &B &C &R;	Y
9	3FT	T	N	model &variable = &TR_ID &R(TR_ID) &A &A*&TR_ID &B &B*&TR_ID &A*&B &A*&B*&TR_ID &R*&A*&B*&TR_ID &C &C*&TR_ID &R*&C*&TR_ID &A*&C &A*&C*&TR_ID &B*&C &B*&C*&TR_ID &A*&B*&C &A*&B*&C*&TR_ID/ SS3;	Y
0	3FT	T	N	random &R(&TR_ID) &R*&A*&B*&TR_ID &R*&C*&TR_ID / test;	Y



	anova_code	oobject	obsnr	totalobs
77	3FT	00_RandomModelANOVA5	2	3
78	3FT	00_RandomModelANOVA4	2	9
79	3FT	00_RandomModelANOVA4	3	9
80	3FT	00_RandomModelANOVA4	4	9
81	3FT	00_RandomModelANOVA4	5	9
82	3FT	00_RandomModelANOVA4	6	9
83	3FT	00_RandomModelANOVA4	7	9
84	3FT	00_RandomModelANOVA4	8	9
85	3FT	00_OverallANOVA	2	3
86	3FT	00_OverallANOVA	3	3
87	3FU	00_RandomModelANOVA1	1	2
88	3FU	00_RandomModelANOVA2	1	2
89	3FU	00_RandomModelANOVA3	1	3
90	3FU	00_RandomModelANOVA3	2	3
91	3FU	00_RandomModelANOVA4	1	4
92	3FU	00_RandomModelANOVA4	2	4
93	3FU	00_RandomModelANOVA4	3	4
94	3FU	00_RandomModelANOVA8	2	7
95	3FU	00_RandomModelANOVA5	1	3
96	3FU	00_RandomModelANOVA5	2	3
97	3FU	00_RandomModelANOVA6	1	2
98	3FU	00_RandomModelANOVA7	1	5
99	3FU	00_RandomModelANOVA7	2	5
00	3FU	00_RandomModelANOVA7	3	5
01	3FU	00_RandomModelANOVA7	4	5
02	3FU	00_RandomModelANOVA8	1	7
03	3FU	00_RandomModelANOVA8	3	7
04	3FU	00_RandomModelANOVA8	4	7
05	3FU	00_RandomModelANOVA8	5	7
06	3FU	00_RandomModelANOVA8	6	7
07	3FU	00_OverallANOVA	2	3
08	3FU	00_OverallANOVA	3	3

Einsatz von Steuertabellen - Bsp

```
d_glm = open("steuer."||steuerglm||"(where=(ANOVA_CODE='"||upcase(ANOVA_CODE)||"'
              and "LSMEAN='"||upcase(LSMEAN)||"' and . . . . . ;

n_ausdr=varnum(d_glm,'AUSDRUCK');
obs_exist=attrn(d_glm,'ANY');
dcl char (250) ausdruck;

if obs_exist then do;
  submit;
    PROC GLM data=&datei;
  endsubmit;
  link subcode;
end;
rc = close_(d_glm);

SUBCODE:
do while (not fetch(d_glm));
  ausdruck=getvarc(d_glm,n_ausdr);

  submit;
    &ausdruck
  endsubmit;

end;

submit continue;
  run;
endsubmit;
return;
```

ODS Output

Name der Output-Objekte im SAS-LOG :

```
ODS trace on;
```

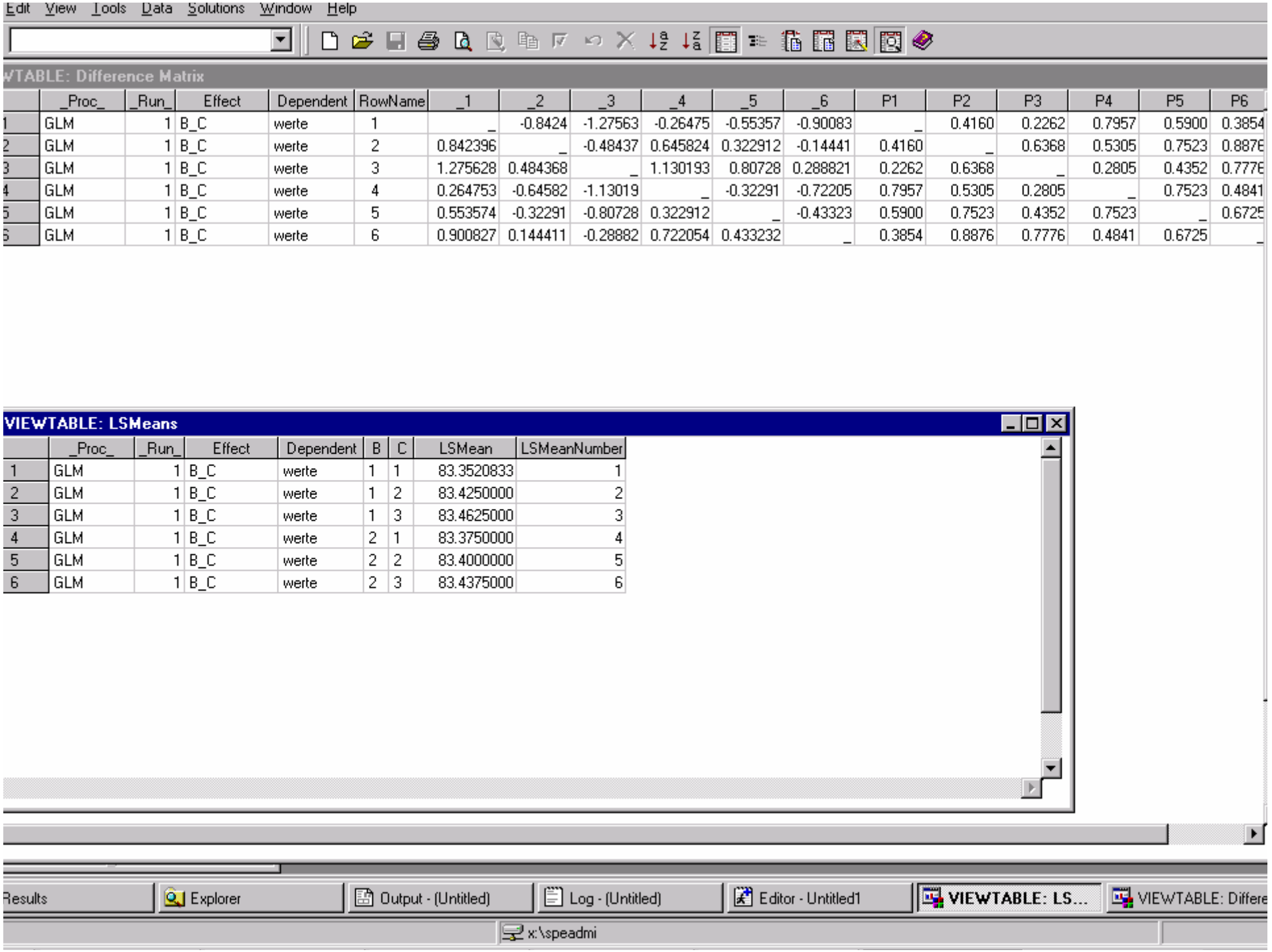
Beispiel:

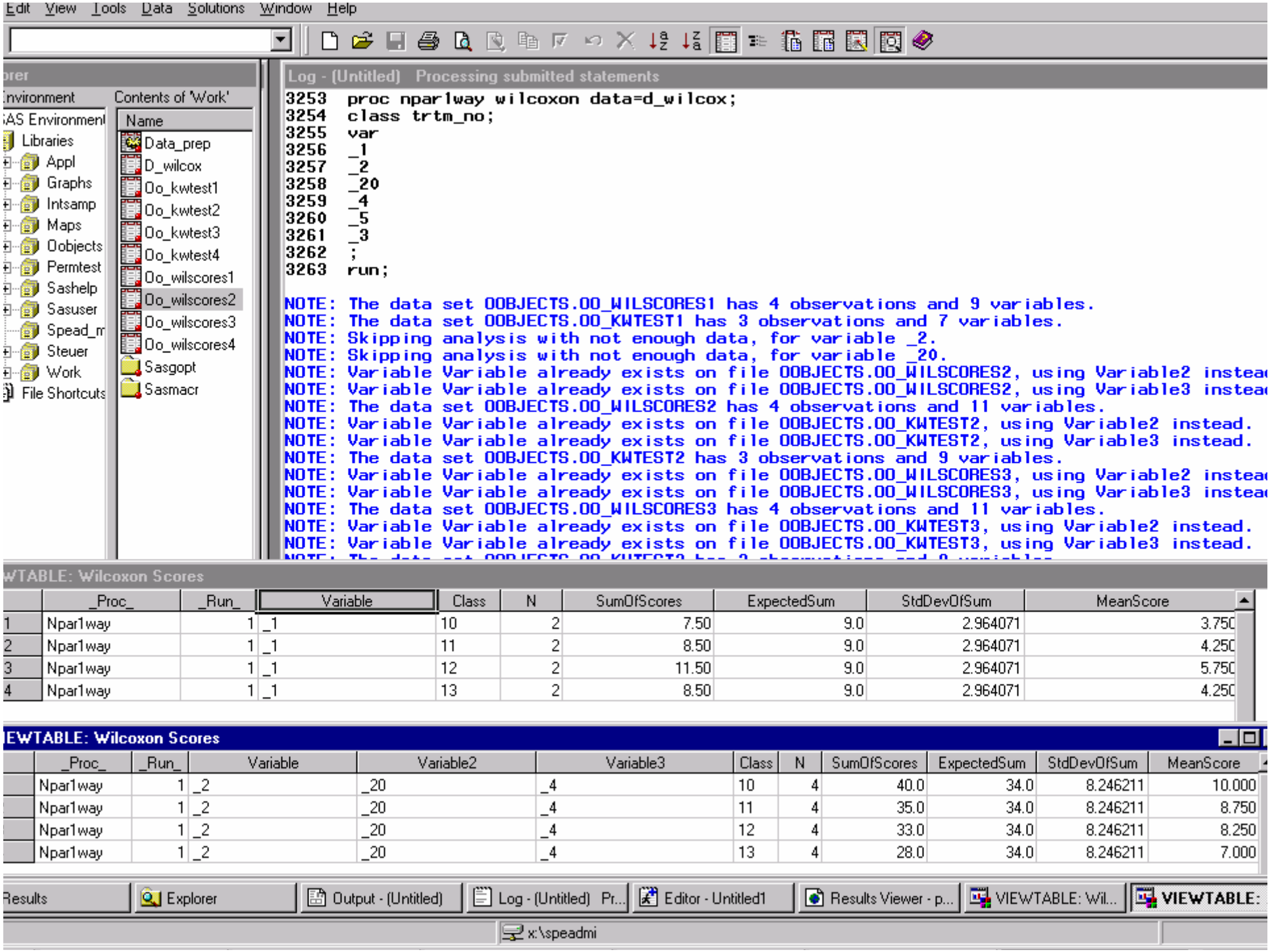
```
ODS output clear;

ODS output OverallANOVA (match_all persist=proc)  =
   oobjects.oo_OverallANOVA;
ODS output MCLinesInfo (match_all persist=proc)   =
   oobjects.oo_MCLinesInfo1;
```

erzeugt SAS-Datasets aus den Output-Objekten

Struktur des ODS-Datasets: undokumentiert


Edit View Tools Data Solutions Window Help
WTABLE: Difference Matrix
Proc _Run_ Effect Dependent RowName _1 _2 _3 _4 _5 _6 P1 P2 P3 P4 P5 P6
1 GLM 1 B_C werte 1 _ -0.8424 -1.27563 -0.26475 -0.55357 -0.90083 _ 0.4160 0.2262 0.7957 0.5900 0.3854
2 GLM 1 B_C werte 2 0.842396 _ -0.48437 0.645824 0.322912 -0.14441 0.4160 _ 0.6368 0.5305 0.7523 0.8876
3 GLM 1 B_C werte 3 1.275628 0.484368 _ 1.130193 0.80728 0.288821 0.2262 0.6368 _ 0.2805 0.4352 0.7776
4 GLM 1 B_C werte 4 0.264753 -0.64582 -1.13019 _ -0.32291 -0.72205 0.7957 0.5305 0.2805 _ 0.7523 0.4841
5 GLM 1 B_C werte 5 0.553574 -0.32291 -0.80728 0.322912 _ -0.43323 0.5900 0.7523 0.4352 0.7523 _ 0.6725
6 GLM 1 B_C werte 6 0.900827 0.144411 -0.28882 0.722054 0.433232 _ 0.3854 0.8876 0.7776 0.4841 0.6725 _
VIEWTABLE: LSMeans
Proc _Run_ Effect Dependent B C LSMean LSMeanNumber
1 GLM 1 B_C werte 1 1 83.3520833 1
2 GLM 1 B_C werte 1 2 83.4250000 2
3 GLM 1 B_C werte 1 3 83.4625000 3
4 GLM 1 B_C werte 2 1 83.3750000 4
5 GLM 1 B_C werte 2 2 83.4000000 5
6 GLM 1 B_C werte 2 3 83.4375000 6
Results
Explorer
Output - (Untitled)
Log - (Untitled)
Editor - Untitled1
VIEWTABLE: LS...
VIEWTABLE: Differe
x:\speadmi


Edit View Tools Data Solutions Window Help
Contents of 'Work'
Name
Data_prep
D_wilcox
Oo_kwtest1
Oo_kwtest2
Oo_kwtest3
Oo_kwtest4
Oo_wilscores1
Oo_wilscores2
Oo_wilscores3
Oo_wilscores4
Sasgopt
Sasmacr
Libraries
Appl
Graphs
Intsamp
Maps
Oobjects
Permtest
Sashelp
Sasuser
Steuer
Work
File Shortcuts
Log - (Untitled) Processing submitted statements
3253 proc nparlway wilcoxon data=d_wilcox;
3254 class trtm_no;
3255 var
3256 _1
3257 _2
3258 _20
3259 _4
3260 _5
3261 _3
3262 ;
3263 run;
NOTE: The data set OOBJECTS.OO_WILSCORES1 has 4 observations and 9 variables.
NOTE: The data set OOBJECTS.OO_KWTEST1 has 3 observations and 7 variables.
NOTE: Skipping analysis with not enough data, for variable _2.
NOTE: Skipping analysis with not enough data, for variable _20.
NOTE: Variable Variable already exists on file OOBJECTS.OO_WILSCORES2, using Variable2 instead
NOTE: Variable Variable already exists on file OOBJECTS.OO_WILSCORES2, using Variable3 instead
NOTE: The data set OOBJECTS.OO_WILSCORES2 has 4 observations and 11 variables.
NOTE: Variable Variable already exists on file OOBJECTS.OO_KWTEST2, using Variable2 instead.
NOTE: Variable Variable already exists on file OOBJECTS.OO_KWTEST2, using Variable3 instead.
NOTE: The data set OOBJECTS.OO_KWTEST2 has 3 observations and 9 variables.
NOTE: Variable Variable already exists on file OOBJECTS.OO_WILSCORES3, using Variable2 instead
NOTE: Variable Variable already exists on file OOBJECTS.OO_WILSCORES3, using Variable3 instead
NOTE: The data set OOBJECTS.OO_WILSCORES3 has 4 observations and 11 variables.
NOTE: Variable Variable already exists on file OOBJECTS.OO_KWTEST3, using Variable2 instead.
NOTE: Variable Variable already exists on file OOBJECTS.OO_KWTEST3, using Variable3 instead.
WTABLE: Wilcoxon Scores
Proc _Run_ Variable Class N SumOfScores ExpectedSum StdDevOfSum MeanScore
1 Npar1way 1 _1 10 2 7.50 9.0 2.964071 3.750
2 Npar1way 1 _1 11 2 8.50 9.0 2.964071 4.250
3 Npar1way 1 _1 12 2 11.50 9.0 2.964071 5.750
4 Npar1way 1 _1 13 2 8.50 9.0 2.964071 4.250
EWTABLE: Wilcoxon Scores
Proc _Run_ Variable Variable2 Variable3 Class N SumOfScores ExpectedSum StdDevOfSum MeanScore
Npar1way 1 _2 _20 _4 10 4 40.0 34.0 8.246211 10.000
Npar1way 1 _2 _20 _4 11 4 35.0 34.0 8.246211 8.750
Npar1way 1 _2 _20 _4 12 4 33.0 34.0 8.246211 8.250
Npar1way 1 _2 _20 _4 13 4 28.0 34.0 8.246211 7.000
Results
Explorer
Output - (Untitled)
Log - (Untitled) Pr...
Editor - Untitled1
Results Viewer - p...
VIEWTABLE: Wil...
VIEWTABLE:
x:\speadmi

Performance

Untersuchungwerkzeuge:

System-Optionen

STIMER FULLSTIMER (OS/390: MEMRPT, FULLSTATS)
– keine Anzeige für SCL

Interaktiv: Performance Analyzer (AF C=… SCLPROF=TIMER)

Timestamps in Programmcode, anschl. Auswertung des LOG

(V8: Application Response Measurement – ARM)

Performance

ORACLE-Zugriff

Tests/Vergleiche :
1. SQL-passthrough
2. Libname engine, proc sql … insert
3. Libname engine, PROC Append
4. Libname engine, scl-Funktion append()

PROC MIXED

- siehe Hinweise aus Online Dokumentation
- Einsatz nur wenn unbedingt erforderlich