Ihre Datenqualität ist nicht so gut, wie sie dachten.

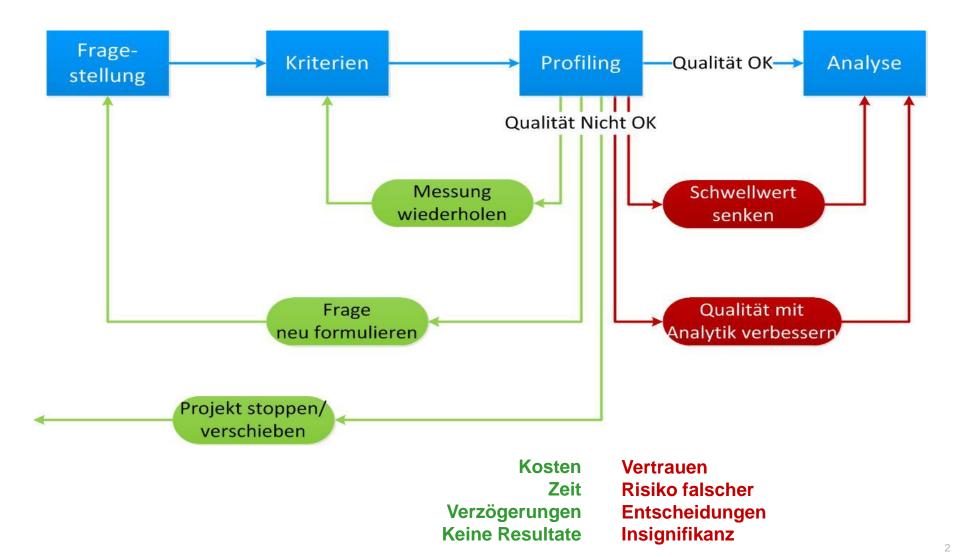
Was sind die Konsequenzen für analytische Auswertungen?

Dr. Gerhard Svolba





#### Konsequenzen schlechter Datenqualität



## Die Konsequenzen der folgenden Effekte wurden untersucht:

- Wie beeinflussen fehlende Werte die Vorhersagekraft? (SIM\_1)
- Variation der Anzahl der verfügbaren Beobachtungen und Ereignisse (SIM\_2)
- Schrittweise Erhöhung der verfügbaren Länge der Datenhistorie (SIM\_3)
- Andere Fragen / Simulationen
  - Entfernen der wichtigsten Variablen aus dem Input-Set
  - Einführung zufälliger und systematischer Verzerrungen in der Input- und Zielgröße im Predictive Modeling
  - Auswirkung von zufälligen und systematischen fehlenden Werte und Fehlern in der Zeitreihen Prognose

## Echtdaten wurden für die Simulationsstudien verwendet

4 Echt-Datensätze aus unterschiedlichen Branchen mit einem binären

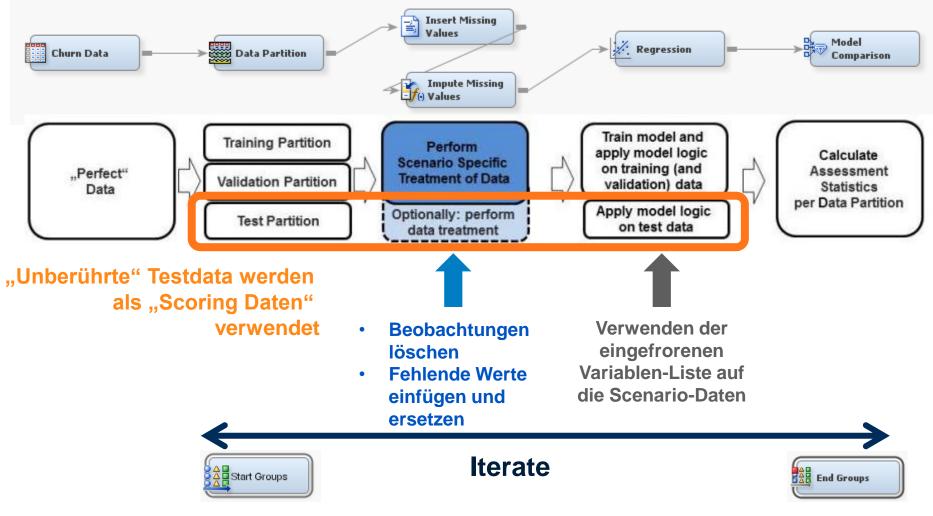
Target wurden verwendet

 Variablen mit > 5 % fehlenden Werten wurden entfernt

- Bei Variablen mit ≤ 5 % fehlenden Werten wurden Beobachtungen mit Missings gefilert
- Durchlauf multipler Modell-Zyklen um ein stabiles Modell mit guter Vorhersagekraft zu bekommen

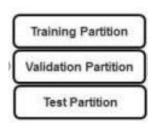
							RETAIN! NPOL		_				
	29Jul 1998		Commercia	24836		29Jul 1998	1	1 Sedan	yes	43	0	11	\$91.44
	06May1994		Commercia	93553		#######		1 SUV	no	44	2	12	\$43.48
	18May1996		Commercia	145326		#######	7	1 Van	no	50	0	- 7	\$106.95
	18Sep1995		Commercia	157851		02Oct1988	7	Van	yes	55	0	11	\$59.16
	30Jan1995	7	Commercia	328102		30Jan1995		1 Panel Tru	icł no	46	0	14	\$59.16
	14Aug1997		Commercia	427058		#######	4	2	no	36	2	12	\$130.5
26934155	19Jan1997	12	Commercia	479082	\$25.810	21Jan1991	6	1 Panel Tru	icł no	41	0	- 7	\$92.8
3969722	12Apr1997	38	Commercia	492179	\$16.640	12May1993		3 Van	yes	39	0	10	\$35.0
7373785	05Sep1995	38	Commercia	536131	\$29.450	27Sep1989	6	2 Van	no	43	2	17	\$145.3
70821537	02Jan1994	48	Commercia	568498	\$9.280	02Jan1994	1	SUV	no	52	0	0	
25895102	24Nov1998	22	Commercia	574750	\$6.450	24Nov1998	1	1 Pickup	no	35	0	- 7	\$14.5
25561360	22Nov1997	5	Commercia	589112	\$29.270	23Nov1993		1 Panel Tru	cl yes	40	0	11	\$57.4
75303717	08Nov1995	62	Commercia	595977	\$4.600	09Nov1989		1 Sedan	yes	31	1	12	\$26.5
3837480	20May1998	35	Commercia	609510	\$19.450	21May1992	6	1 SUV	no	47	0	8	\$18.4
21305754	22Sep1994	24	Commercia	680397	\$23,450	22Sep1994	1	1 Sedan	no	42	0	8	\$52.9
10120026	07Apr1999	14	Commercia	746346	\$28,560	12Mau 1986	13	2 Panel Tru	cl ves	48	0	8	\$52.9
350798	01Jul 1994	31	Commercia	809783	\$33,710	02Jul1991	3	Panel Tru		49	0	9	\$52.9
2990245			Commercia	825520		22Feb1990	4	Panel Tru		46	0	14	\$52.9
	15Mau1996		Commercia	838306		#######	13	2 Pickup	ves	45	0	11	\$61.9
366048			Commercia	891607		06Jun1992	6	Panel Tru		48	0	9	\$61.5
3033252			Commercia	1026106		06Jul 1985	10	1 Sports Ca		44	2	15	\$139.3
	12Jun 1997		Commercia	1049545		15Jun 1987	10	1 Van	no	38	ñ	0	ψ100.0
	30Aug1995		Commercia	1103124		#######	10	1	ves	55	0	0	\$84.9
935828			Commercia	1253235		07Jul 1980		1 Panel Tru		42	ō	9	\$76.2
	28May1997		Commercia	1331376		05Jun1997	1	2 Sedan	no	52	0	9	\$68.9
12131636			Commercia	1344988		#######	16	Panel Tru		43	0	11	\$132.5
3707484			Commercia	1350845	\$7,500		9	1 Pickup	ves	60	0	9	\$125.8
	210ct1995		Commercia	1440036		21Oct1995	3	1 Van	no	37	0	12	\$123.5
	17Mav1997		Commercia	1441076		02Jun1991		3 Panel Tru		47	0	13	\$45.2
209593	11Dec1998		Commercia	1488986		12Dec1993	5			40	3	9	\$75.5
							4	1 Sports Ca				12	
	28Dec1994		Commercia	1597160		16Dec1990		2 Panel Tru		47 33	0		\$104.2
	30Apr1998		Commercia	1668247		#######	1	3 Van	no		1	12	\$111.4
307371	17/Dct1998		Commercia	1699144		19Sep1994		2 Pickup	no	44			\$44.7
	20Apr1998		Commercia	1787809		#######		2 Van	yes	50	0	14	\$44.7
	19Mar1995		Commercia	1816497		30Mar1995	1	2 Pickup	no	51	0	9	\$50.1
	06Dec1994		Commercia	1875494		09Dec1982	12	1 Sports Ca		43	3	14	\$48.
16157391			Commercia	2030290		21Jun1983	11	Pickup	no	49	0	8	\$22.0
252155			Commercia	2035565		#######	7	Pickup	yes	34	1	10	\$23.5
833784			Commercia	2070960		05Nov1989	4	2 Panel Tru		55	0	14	\$55.4
	02Sep1997		Commercia	2137830		05Sep1987	10	1 Pickup	yes	48	0	10	\$39.6
	06Jun1995		Commercia	2219845		07Jun1988	7	1 Van	no	45	3	15	\$23.7
	01Jun1996		Commercia	2235985		25Apr1996	1	2 Panel Tru		41	0	8	\$55.3
	29Aug1993		Commercia	2321817		01Sep1983	10	1 Panel Tru	icł no	55	0	16	\$163.1
429873			Commercia	2364349		21Jan1982	14	1 Sedan	yes	52	0	12	\$24.5
	03Apr1999	45	Commercia	2365874		04Apr1992	7	1 Pickup	no	54	0	15	\$107.8
39176001	20Dec1994	42	Commercia	2392645	\$7.470	22Dec1987	7	1 Pickup	no	39	2	12	\$59.6
84194086	25Jun1995	51	Commercia	2465546	\$6.720	26Jun1991	4	1 Pickup	no	44	0	4	\$146.2
60189132	20Apr1996	10	Commercia	2546478	\$6.120	23Apr1986	10	1 SUV	no	45	3	4	\$1.1
	07(Dct1997		Commercia	2577125		23Oct1984	13	2 Sedan	no	43	0	11	\$1.1

## Simulationsstudien helfen die Konsequenzen schlechter Datenqualität zu quantifizieren



# Prozess für die Szenarios mit fehlenden Werten (SIM\_1)

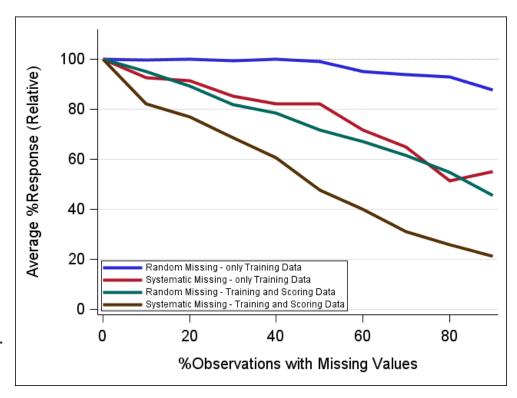
- Für einen bestimmten Anteil von Beobachtungen (10%, ...)
  - Setzen von Intervall-skalierten und nominalen Input-Variablen auf fehlend
    - » Zufällige Auswahl
    - » Systematische Auswahl auf Basis von Segmenten
  - Ersetzen der fehlende Werte durch den Impute Node im SAS Enterprise Miner
- Trainieren des Modells mit dem eingefrorenen Set an Variablen
- Optional: Anwenden der "Behandlung" auch für Scoring-Daten
- Bewerten der Modellqualität



85688650         48         \$18.510         ######## no         40         \$106.952           3764824:         25         \$17.600         02Oct1988         yes         55         \$59.162           89165555:         7         \$17.230         30Jan1995         no         46         \$59.162           3969722:         38         \$16.640         12May1993         yes         39         \$35.081           3969722:         38         \$16.640         12May1993         yes         39         \$35.081           3969722:         38         \$16.640         12May1993         yes         39         \$35.081           3969723:         38         \$29.450         27Sep1994         no         52         \$0           25695100:         22         \$6.450         24Nov1998         no         35         \$14.508           25561360:         23         \$29.270         23Nov1993         yes         40         \$57.474           7530371:         62         \$4.600         99Nov1989         yes         31         \$26.520           21305754:         24         \$23.450         22Sep1994         no         42         \$52.988           39302541:         30	ID T	RAVTIM E	BLUEBOO	INITDATE	RED_C/	AGE	INCOME
\$3764824   25	13276104		\$14.940	29Jul1998	yes		\$91,449
6916555!         7         \$17,230         30Jan1995 no         46         \$59,162           2693415!         12         \$25,810         21Jan1991 no         40         \$59,162           3969722!         38         \$16,640         12May1993 yes         39         \$35,081           7082153.         48         \$9,280         02Jan1994 no         52         \$0           2589510.         22         \$6,450         24Nov1998 no         35         \$14,508           25561360         23         \$29,270         23Nov1993 yes         40         \$57,474           7530371.         62         \$4,600         09Nov1989 yes         31         \$26,520           4012002(         14         \$28,560         12May1986 yes         48         \$52,988           79350798.         31         \$33,710         02Jul1991 no         49         \$52,988           79392445         30         \$10,910         ###################### yes         45         \$61,931           79392541         30         \$10,910         ####################################	8568865	48	\$18.510	#######	no	40	\$106,952
2693415!         12         \$25.810         21Jan1991         no         40         \$59.162           3969722!         38         \$16.640         12May1993         yes         39         \$35.081           1737378!         38         \$29.450         27Sep1989         no         43         \$145.353           70821537         48         \$3.280         02Jan1994         no         52         \$0           25895102         22         \$6.450         24Nov1998         no         35         \$14.508           25561361         23         \$29.270         23Nov1993         yes         40         \$57.474           75303717         62         \$4.600         09Nov1998         yes         31         \$25.598           40120026         14         \$28.560         12May1986         yes         48         \$25.988           9350798         31         \$33.710         02Jul1991         no         49         \$52.988           29902451         23         \$33.320         22Feb1990         yes         40         \$52.988           29902451         30         \$10.910         ####################################	3764824	25				55	\$59,162
3969722   38	69165555		\$17.230	30Jan1995	no	46	\$59,162
1737378  38	26934155	12	\$25.810	21Jan1991	no	40	\$59,162
70821537	3969722		\$16.640	12May1993	yes	39	\$35.081
25895102         22         \$6.450         24Nov1998         no         35         \$14.508           2556136(         23         \$29.270         23Nov1993         yes         40         \$57.474           7530371;         62         \$4.600         09Nov1989         yes         31         \$26.520           21305754         24         \$23.450         22Sep1994         no         42         \$52.988           4012002(         14         \$28.560         12May1986         yes         48         \$52.988           2990245         23         \$33.370         02Jul1991         no         49         \$52.988           2990245         23         \$33.320         22Feb1990         yes         40         \$52.988           2990245         30         \$10.910         ######## yes         45         \$61.931           39392541         30         \$10.910         ######## yes         45         \$61.931           5366048.         40         \$23.230         00Jun1992         yes         42         \$76.226           7048324         36         \$17.230         05Jun1997         no         52         \$59.162           42131636         \$30.150         #####							\$145,353
2556136(         23         \$29,270         23Nov1993         yes         40         \$57,474           75303717         62         \$4,600         09Nov1989         yes         31         \$26,520           21305754         24         \$23,450         22Sep1994         no         42         \$52,988           4012002(         14         \$28,560         12May1986         yes         48         \$52,988           3930798         31         \$33,710         02Jul1991         no         49         \$52,988           2990245         23         \$33,320         22Eb1990         yes         40         \$52,988           3939254         30         \$10,910         ######### yes         45         \$61,931           5366048         40         \$23,230         06Jun1992         yes         48         \$61,509           80332521         35         \$22,050         no         44         \$133,330         44         \$133,330           4660671         41         \$17,470         15Jun1987         no         38         \$0           5935828         18         \$23,390         07Ju1980         yes         40         \$125,893           7048322 <th< th=""><th>70821537</th><th></th><th></th><th></th><th></th><th></th><th>\$0</th></th<>	70821537						\$0
75303717 62 \$4,600 09Nov1989 yes 31 \$26,520 21305754 24 \$23,450 22Sep1994 no 42 \$52,988 4012002( 14 \$28,560 12May1986 yes 48 \$52,988 3930798 31 \$33,710 02Jul1991 no 49 \$52,988 2990245 23 \$33,320 22Feb1990 yes 40 \$52,988 2990245 30 \$10,910 ####### yes 45 \$61,931 5366048. 40 \$23,230 06Jun1992 yes 48 \$61,509 8033252 35 \$22,050 no 44 \$139,330 4660671; 41 \$17,470 15Jun1987 no 38 \$0 5935828 18 \$23,390 07Jul1980 yes 42 \$76,226 1704832 36 \$17,230 05Jun1997 no 52 \$59,162 42131636 8 \$30,150 ####### no 43 \$132,561 3707484 21 \$7,500 29Jul1985 yes 40 \$125,893 7182942( 14 \$17,550 21Dct1995 no 40 \$123,520 5388757; 33 \$29,210 02Jun1991 yes 47 \$45,257 5203593 53 \$13,050 12Dec1993 no 40 \$75,516 5684737, 40 \$36,120 16Dec1990 yes 47 \$104,271 45380631 11 \$17,300 ####### no 33 \$111,427 5804259 23 \$11,620 30Mar1995 no 50 \$111,427 5804259 23 \$11,620 30Mar1995 no 40 \$22,059 93412915 50 \$14,530 09Dec1982 no 43 \$48,184 46157391 24 \$21,990 21Jun1983 no 40 \$22,059 93412915 50 \$14,530 09Dec1982 no 43 \$48,184 46157391 24 \$21,990 21Jun1983 no 40 \$22,059 93412915 50 \$14,530 09Dec1982 no 43 \$48,184 46157391 24 \$21,990 21Jun1983 no 40 \$22,059 93412915 50 \$14,530 09Dec1982 no 55 \$55,409 93412915 50 \$14,530 09Dec1982 no 55 \$55,409 93412915 50 \$14,530 09Dec1982 no 55 \$55,409 5039064 48 \$8,460 05Sep1987 yes 48 \$39,613 19707619 9 \$17,230 07Jun1988 no 45 \$23,773 3457713( 17 \$31,390 25Apr1996 yes 40 \$23,571 5830856: 44 \$2,700 05Nov1989 no 55 \$55,409 5039064: 48 \$8,460 05Sep1987 yes 48 \$39,613 19707619 9 \$17,230 07Jun1988 no 45 \$23,773 3457713( 17 \$31,390 25Apr1996 yes 52 \$59,162 50189132 10 \$6,120 23Apr1986 no 45 \$107,808 39176001 42 \$7,470 22Dec1987 no 39 \$59,685 6429873: 23 \$10,350 21Jan1990 no 55 \$163,158 6429873: 23 \$10,350 21Jan1991 no 44 \$146,267 60189132 10 \$6,120 23Apr1986 no 45 \$1,158 7921960! 32 \$13,160 23Oct1984 no 43 \$1,158							\$14.508
\$21305754   24		23					
\$40120026							
\$350798							
2990245         23         \$33,320         22Feb1990         yes         40         \$52,988           3939254!         30         \$10,910         ####### yes         45         \$61,931           5366048.         40         \$23,230         06Jun1992         yes         48         \$61,503           8033252!         35         \$22,050         no         44         \$139,330           466067!         41         \$17,470         15Jun1987 no         38         \$0           5935828         18         \$23,390         07Jul1980 yes         42         \$76,226           1704832*         36         \$17,230         05Jun1997 no         52         \$59,162           4213166         8         \$30,150         ######## no         43         \$132,561           3707484         21         \$7,500         29Jul1985 yes         40         \$125,833           71829426         14         \$17,550         210c1995 no         40         \$123,520           5388757;         33         \$29,210         02Jun1991 yes         47         \$45,257           5209593         53         \$13,050         12Dec1993 no         40         \$75,516           584737.         4							
3939254  30							
5366048         40         \$23,230         06Jun1992         yes         48         \$61,509           80332521         35         \$22,050         no         44         \$139,330           46606715         41         \$17,470         15Jun1987         no         38         \$0           5935828         18         \$23,390         07Jul1980         yes         42         \$76,226           17048324         36         \$17,230         05Jun1997         no         52         \$59,162           42131636         8         \$30,50         ######## ino         43         \$132,561           3707484         21         \$7,500         29Jul1985         yes         40         \$125,893           71829426         14         \$17,550         21De1995         no         40         \$123,520           5388757:         33         \$29,210         02Jun1991         yes         47         \$45,257           5209593         53         \$13,050         12Dec1993         no         40         \$75,516           5684737.         40         \$36,120         16Dec1990         yes         47         \$104,271           58208610         11         \$17,300         <			•		/		
8033252!         35         \$22.050         no         44         \$13,330           466067!!         41         \$17,470         \$15Jun1987         no         38         \$0           5935828*         18         \$23,330         07Jul1980         yes         42         \$76,226           1704832*         36         \$17,230         05Jun1997         no         52         \$59,162           42131636*         8         \$30,150         ######## ino         43         \$132,561           3707484*         21         \$7,500         29Jul1985         yes         40         \$125,833           71829426*         14         \$17,550         21Det1995         no         40         \$123,520           5388757:         33         \$29,210         02Jun1991         yes         47         \$45,257           5209593         53         \$13,050         12Dec1993         no         40         \$75,516           5684737:         40         \$36,120         16Dec1990         yes         47         \$10,427           5820861         11         \$17,300         ######## no         33         \$111,427           58208661         11         \$17,300         ######## n					/		
466067I:         41         \$17.470         15Jun1987 no         38         \$0           5935828**         18         \$23.390         07Jul1980 yes         42         \$76.226           1704832**         36         \$17.230         05Jun1997 no         52         \$59.162           42131636**         \$30.150         ####### no         43         \$132.561           3707484**         21         \$7.500         29Jul1985 yes         40         \$125.893           71829426**         14         \$17.550         21Oct1995 no         40         \$123.520           5388757**         33         \$29.210         02Jun1991 yes         47         \$45.257           5209593**         53         \$13.050         12Dec1993 no         40         \$75.516           5684737**         40         \$36.120         16Dec1990 yes         47         \$104.271           4538673**         35         \$28.180         ######## no         33         \$111.427           5804610**         11         \$17.300         ######## yes         50         \$111.427           5804259         23         \$11.620         30Mar1995 no         51         \$50.166           93412915         50         \$14.5				06Jun1992			
5935828         18         \$23,390         O7Jul 1980         yes         42         \$76,226           17048324         36         \$17,230         05Jun 1997         no         52         \$59,162           42131636         8         \$30,150         ####### Ino         43         \$132,561           3707484         21         \$7,500         29Jul 1985         yes         40         \$125,893           71829426         14         \$17,550         21Dct1995         no         40         \$123,520           53887571         33         \$29,210         02Jun 1991         yes         47         \$45,257           5209593         53         \$13,050         12Dec1993         no         40         \$75,516           56847377         40         \$36,120         16Dec1990         yes         47         \$104,275           5820861         11         \$17,300         ####### no         33         \$111,427           5804259         23         \$11,620         30Mar 1995         no         51         \$50,166           93412915         50         \$14,530         09Dec1982         no         43         \$48,184           46157391         24         \$21,9							*
1704832  36							
42131636         8         \$30,150         ######## no         43         \$132,551           3707484         21         \$7,500         29Jul1985 yes         40         \$125,893           71829426         14         \$17,550         21Dct1995 no         40         \$123,520           5388757:         33         \$29,210         02Jun1991 yes         47         \$45,257           5209593         53         \$13,050         12Dec1993 no         40         \$75,516           5684737.         40         \$36,120         16Dec1990 yes         47         \$104,271           4538673.         35         \$28,180         ########         no         33         \$111,427           58208610.         11         \$17,300         ########         no         33         \$111,427           58208610.         11         \$17,300         ######## po         33         \$111,427           58208611.         11         \$17,300         ######## po         50         \$111,427           58208612.         50         \$14,530         90Dec1982 no         51         \$50,166           80415915.         50         \$14,530         90Dec1982 no         43         \$48,184           46							
3707484         21         \$7.500         29Jul1985 yes         40         \$125.893           71829426         14         \$17.550         21Dct1995 no         40         \$123.520           5388757:         33         \$29.210         02Jun1991 yes         47         \$45.257           5209593         53         \$13.050         12Dec1993 no         40         \$75.516           5684737.         40         \$36.120         16Dec1990 yes         47         \$104.271           4538673         35         \$28.180         ######## no         33         \$111.427           58208610         11         \$17.300         ######## yes         50         \$111.427           58208610         11         \$17.300         ######## yes         50         \$111.427           58208611         11         \$17.300         ######## yes         50         \$111.427           58208611         11         \$16.20         30Mar1995 no         51         \$50.166           804259         23         \$11.620         30Mar1995 no         51         \$50.166           33412915         50         \$14.530         09Dec1982 no         43         \$48.184           46157391         24							
7182942£ 14 \$17.550 21\(\text{Dct1995}\) no 40 \$123.520  5388757; 33 \$29.210 02\(\text{Jun1991}\) yes 47 \$45.257  5209593 53 \$13.050 12\(\text{Dct1993}\) no 40 \$75.516  5684737; 40 \$36.120 16\(\text{Dct1990}\) yes 47 \$104.271  4536673; 35 \$28.180 ####### no 33 \$111.427  5820861( 11 \$17.300 ######## yes 50 \$111.427  5820861( 11 \$17.300 ######## yes 50 \$111.427  6804259; 23 \$11.620 30\(\text{Mar1995}\) no 51 \$50.166  6804259; 23 \$11.620 30\(\text{Mar1995}\) no 51 \$50.166  46157391 24 \$21.990 21\(\text{Jun1983}\) no 40 \$22.059  22521555; 35 \$12.180 ######## yes 40 \$23.571  6833784		-					
5388757:         33         \$29.210         02Jun1991         yes         47         \$45.257           5209593:         53         \$13.050         12Dec1993         no         40         \$75.516           5684737:         40         \$36.120         16Dec1990         yes         47         \$104.271           4538673:         35         \$28.180         ########         no         33         \$111.427           5820861(         11         \$17.300         ########         yes         50         \$111.427           5804259:         23         \$11.620         30Mar1995         no         51         \$50.166           93412915         50         \$14.530         09Dec1982         no         43         \$48.184           46157391         24         \$21.990         21Jun1983         no         40         \$22.059           22521555:         35         \$12.180         ############## yes         40         \$22.059           22521555:         35         \$12.180         ############ yes         40         \$23.571           5039064:         48         \$8.460         05Sep1987         yes         48         \$39.613           79707619         9							
5209593         53         \$13.050         12Dec1993 no         40         \$75.516           5684737.         40         \$36.120         16Dec1990 yes         47         \$104.271           4538673.         35         \$28.180         ####### no         33         \$111.427           58208610.         11         \$17.300         ####### yes         50         \$111.427           58208610.         11         \$17.300         ######## yes         50         \$111.427           58208610.         11         \$17.300         ######## yes         50         \$111.427           58208610.         11         \$17.300         30Mar1995 no         51         \$50.166           93412915.         50         \$14.530         09Dec1982 no         43         \$48.184           46157391.         24         \$21.990         2Juni983 no         40         \$22.059           22521555.         35         \$12.180         ######### yes         40         \$23.571           5039064.         48         \$8.460         05Sep1987 yes         48         \$39.613           19707619.         9         \$17.230         07Juni988 no         45         \$23.773           3457730.         17							
5684737.         40         \$36.120         16Dec1990 yes         47         \$104.271           4538673.         35         \$28.180         ####### no         33         \$111.427           5820861(         11         \$17.300         ####### yes         50         \$111.427           5804259.         23         \$11.620         30Mar1995 no         51         \$50.166           93412915.         50         \$14.530         09Dec1982 no         43         \$48.184           46157391.         24         \$21.990         21Jun1983 no         40         \$22.059           22521555.         35         \$12.180         ######## yes         40         \$23.571           58337844.         45         \$27.890         05Nov1989 no         55         \$55.409           5039064.         48         \$8.460         05Sep1987 yes         48         \$39.613           19707619.         9         \$17.230         07Jun1988 no         45         \$23.773           3457713(							
4538673         35         \$28.180         ####### no         33         \$111.427           58208610         11         \$17.300         ######## yes         50         \$111.427           6804259         23         \$11.620         30Mar1995         no         51         \$50.166           93412915         50         \$14.530         09Dec1982         no         43         \$48.184           46157391         24         \$21.990         21Jun1983         no         40         \$22.059           22521555         35         \$12.180         ######## yes         40         \$23.571           68337841         45         \$27.890         05Nov1989         no         55         \$55.409           5039064.         48         \$8.460         05Sep1987         yes         48         \$39.613           19707619         9         \$17.230         07Jun1988         no         45         \$23.733           34577130         17         \$31.390         25Apr1996         yes         40         \$55.364           8308556:         44         \$23.320         23Jan1990         no         55         \$163.158           54298731         23         \$10.350         23							
5820861(         11         \$17.300         ####### yes         50         \$111.427           6804259:         23         \$11.620         30Mar1995         no         51         \$50.166           93412915         50         \$14.530         09Dec1982         no         43         \$48.184           46157391         24         \$21.990         21Jun1983         no         40         \$22.059           2252155!         35         \$12.180         ######## yes         40         \$23.571           5833784!         45         \$27.890         05Nov1989         no         55         \$55.409           5039064.         48         \$8.460         05Sep1987 yes         48         \$39.613           19707619         9         \$17.230         07Jun1988 no         45         \$23.773           3457713(         17         \$31.390         25Apr1996 yes         40         \$55.364           8308556:         44         \$23.320         23Jan1900 no         55         \$163.158           6429873:         23         \$10.300         21Jan1982 yes         52         \$59.162           9309292:         45         \$27.200         04Apr1992 no         54         \$107.808							
6804259.         23         \$11.620         30Mar1995         no         51         \$50.166           93412915         50         \$14.530         09Dec1982         no         43         \$48.184           46157391         24         \$21.990         2JJun1983         no         40         \$22.059           22521555         35         \$12.180         ######## yes         40         \$23.571           88337841         45         \$27.890         05Nov1983         no         55         \$55.409           5039064         48         \$3.460         05Sep1987         yes         48         \$39.613           19707619         9         \$17.230         07Jun1988         no         45         \$23.773           34577130         17         \$31.390         25Apr1996         yes         40         \$55.364           8308556:         44         \$23.320         2JJan1900         no         55         \$163.158           45         \$27.202         1Jan1992         yes         52         \$59.162           9309292         45         \$27.020         1Jan1992         no         54         \$107.808           39176001         42         \$7.470			¥				
93412915         50         \$14,530         09Dec1982 no         43         \$48,184           46157391         24         \$21,990         21Jun1983 no         40         \$22,059           22521555         35         \$12,180         ######## yes         40         \$23,571           68337841         45         \$27,890         05Nov1989 no         55         \$55,409           5039064         48         \$8,460         055ep1987 yes         48         \$38,613           79707619         9         \$17,230         07Jun1988 no         45         \$23,773           34577130         17         \$31,390         25Apr1996 yes         40         \$55,364           8308556:         44         \$23,320         2Jun1990 no         55         \$163,158           6429873:         23         \$10,350         21Jun1982 yes         52         \$59,655           39176001         42         \$7,470         22Dec1987 no         39         \$59,685           84194086         51         \$6,720         26Jun1991 no         44         \$146,267           60189132         10         \$6,120         23Apr1986 no         45         \$1,158           79219605         32					/		
46157391         24         \$21,990         21Jun1983         no         40         \$22,059           22521555         35         \$12,180         ######## yes         40         \$23,571           68337841         45         \$27,890         05Nov1989         no         55         \$55,409           5039064         48         \$8,460         05Sep1987         yes         48         \$39,613           19707619         9         \$17,230         07Jun1988         no         45         \$23,773           34577130         17         \$31,390         25Apr1996         yes         40         \$55,364           8308556:         44         \$23,320         23Jan1900         no         55         \$163,158           6429873:         23         \$10,350         21Jan1982         yes         52         \$59,162           9309292:         45         \$27,020         04Apr1992         no         54         \$107,808           39176001         42         \$7,470         22Dec1987         no         39         \$59,685           84194086         51         \$6,720         26Jun1991         no         44         \$146,257           50189132         10							
2252155!         35         \$12.180         ######## yes         40         \$23.571           6833784:         45         \$27.890         05Nov1989 no         55         \$55.409           5039064:         48         \$8.460         05Sep1987 yes         48         \$39.613           19707619:         9         \$17.230         07Jun1988 no         45         \$23.773           3457713(         17         \$31.390         25Apr1936 yes         40         \$55.364           8308556:         44         \$23.320         23Jan1900 no         55         \$163.158           6429873:         23         \$10.350         21Jan1982 yes         52         \$59.162           9309292:         45         \$27.020         04Apr1992 no         54         \$107.808           39176001         42         \$7.470         22Dec1987 no         39         \$59.685           84194086         51         \$6.720         25Jun1991 no         44         \$146.267           50189132         10         \$6.120         23Apr1986 no         45         \$1.158           79219605         32         \$13.160         23Oct1984 no         43         \$1.158							
6833784:         45         \$27.890         05Nov1989 no         55         \$55.409           5039064:         48         \$8.460         05Sep1987 yes         48         \$39.613           19707619:         9         \$17.230         07Jun1988 no         45         \$23.773           34577130:         17         \$31.390         25Apr1996 yes         40         \$55.364           8308556:         44         \$23.320         23Jan1900 no         55         \$163.158           6429873:         23         \$10.350         21Jan1982 yes         52         \$59.162           93092921:         45         \$27.020         04Apr1992 no         54         \$107.808           793176001:         42         \$7.470         22Dec1987 no         39         \$55.685           84194086:         51         \$6.720         25Jun1991 no         44         \$146.267           80189132:         10         \$6.120         23Apr1986 no         45         \$1.158           79219605:         32         \$13.160         23Oct1984 no         43         \$1.158							
5039064:         48         \$8,460         05Sep1987 yes         48         \$39,613           79707619         9         \$17,230         07Jun1988 no         45         \$23,773           3457713(         17         \$31,390         25Apr1996 yes         40         \$55,364           8308556:         44         \$23,320         23Jan1900 no         55         \$163,158           6429873!         23         \$10,350         21Jan1982 yes         52         \$59,162           9309292!         45         \$27,020         04Apr1992 no         54         \$107,808           39176001         42         \$7,470         22Dec1987 no         39         \$55,685           84194086         51         \$6,720         25Jun1991 no         44         \$146,267           80189132         10         \$6,120         23Apr1986 no         45         \$1,58           79219605         32         \$13,160         23Oct1984 no         43         \$1,158					/		
19707619         9         \$17.230         07Jun1988 no         45         \$23.773           3457713(         17         \$31.390         25Apr1996 yes         40         \$55.364           8308556:         44         \$23.320         23Jan1900 no         55         \$163.158           6429873!         23         \$10.350         21Jan1982 yes         52         \$59.162           9309292!         45         \$27.020         04Apr1992 no         54         \$107.808           39176001         42         \$7.470         22Dec1987 no         39         \$59.685           84194086         51         \$6.720         26Jun1991 no         44         \$146.267           60189132         10         \$6.120         23Apr1986 no         45         \$1.158           79219605         32         \$13.160         23Oct1984 no         43         \$1.158							
3457713(         17         \$31.390         25Apr1996 yes         40         \$55.364           8308556:         44         \$23.320         23Jan1900 no         55         \$163.158           6429873!         23         \$10.350         21Jan1982 yes         52         \$59.162           9309292!         45         \$27.020         04Apr1992 no         54         \$107.808           39176001         42         \$7.470         22Dec1987 no         39         \$59.685           84194086         51         \$6.720         26Jun1991 no         44         \$145.267           60189132         10         \$6.120         23Apr1986 no         45         \$1.158           79219605         32         \$13.160         23Oct1984 no         43         \$1.158	_						
8308556:         44         \$23.320         23Jan1900 no         55         \$163.158           6429873!         23         \$10.350         21Jan1982 yes         52         \$59.162           9309292!         45         \$27.020         04Apr1992 no         54         \$107.808           39176001         42         \$7.470         22Dec1987 no         39         \$59.685           84194086         51         \$6.720         26Jun1991 no         44         \$146.267           60189132         10         \$6.120         23Apr1986 no         45         \$1.158           79219605         32         \$13.160         23Oct1984 no         43         \$1.158		_					
6429873!         23         \$10.350         21Jan1982 yes         52         \$59.162           9309292!         45         \$27.020         04Apr1992 no         54         \$107.808           39176001         42         \$7.470         22Dec1987 no         39         \$59.685           84194086         51         \$6.720         26Jun1991 no         44         \$146.267           60189132         10         \$6.120         23Apr1986 no         45         \$1.158           79219605         32         \$13.160         23Oct1984 no         43         \$1.158							
9309292I         45         \$27,020         04Apr1992 no         54         \$107,808           39176001         42         \$7,470         22Dec1987 no         39         \$59,685           84194086         51         \$6,720         26Jun1991 no         44         \$146,267           60189132         10         \$6,120         23Apr1986 no         45         \$1,158           79219605         32         \$13,160         23Oct1984 no         43         \$1,158							
39176001         42         \$7.470         22Dec1987 no         39         \$59.685           84194086         51         \$6.720         26Jun1991 no         44         \$146.267           60189132         10         \$6.120         23Apr1986 no         45         \$1.158           79219605         32         \$13.160         23Oct1984 no         43         \$1.158							
84194086         51         \$6.720         26Jun1991 no         44         \$146.267           60189132         10         \$6.120         23Apr1986 no         45         \$1.158           79219605         32         \$13.160         23Oct1984 no         43         \$1.158							\$59,685
60189132         10         \$6.120         23Apr1986 no         45         \$1.158           79219605         32         \$13.160         23Oct1984 no         43         \$1.158							
<b>7921960</b> 5 32 \$13.160 2300ct1984 no 43 \$1.158							\$1.158
							\$1,158
			* /==				

#### Ergebnisse aus den Missing Value Szenarios

- Zufällig fehlende Werte nur in den Trainingsdaten haben nur beschränkte Auswirkungen.
- Fehlende Werte in den Scoringdaten beeinflussen das Ergebnis viel stärker.
- Systematische fehlende Werte haben einen viel stärkeren Effekt
- Punkte die wichtig sind:
  - Nicht nur der Anteil fehlender Werte, sondern insbesondere der Тур
  - Fehlende Werte in den Scoring Daten



## Quantifizieren der Ergebnisse der Missing Value Szenarios

Durchführen eines allgemeinen linearen Modells:

Response = f(%missing, Systematic\_YN, ScoringData\_YN)

Parameter	Value	Interpretation
Intercept	19.29	Response ohne fehlende Werte
%missing	- 0.1	10 % fehlend ~ 1% weniger Response
Systematic_YN	- 3.6	Systematische Fehler resultieren in 3.6 % weniger Response
Scoring_YN	- 4.23	Fehlende Werte in den Scoring Daten resultieren in 4.23 % weniger Response

Analyse des Effekts der Datenquantität in der Ereignisvorhersage (SIM\_2) 13276104 122 1514.940 233415 188851 1888 18551 1888855 1888 18551 1888855 1888 18551 1888855 1888 18551 1888855 1888 18551 1888855 1888 18551 1888855 1888 18551 1888855 1888 18551 1888855 1888 18551 1888855 1888 18551 1888855 1888 18551 1888855 1888 18551 1888855 1888 18551 1888855 1888 18551 1888855 1888 18551 1888855 1888 18551 1888855 1888 18551 1888855 1888 18551 1888855 1888 18551 18551

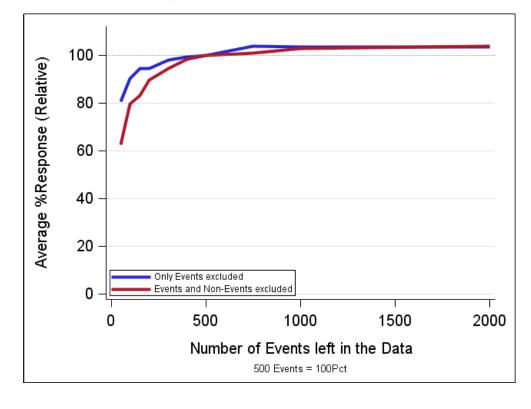
- Zufällig ausgewählte Beobachtungen wurden aus den Daten entfernt
- Zusätzliche Beobachtungen und Ereignisse resultieren in einer höheren %Correct Response Rate
- Aber:
  - Linearer oder nicht-linearer Effekt
  - Wie kann dieser Effekt quantifiziert werden?
  - Tragen auch Non-Events zu einer Steigerung bei?
  - Zahlt es sich aus, auf weitere Events zu warten?

13276104         22         \$14.940         29.Jul 1998 yes         43         \$91.44           85688651         48         \$18.510         ####### no         50         \$106.95           3764824.         25         \$17.600         02Oct1988 yes         55         \$59.16           26934155         12         \$25.810         21Jan1991 no         41         \$92.84           3963722.         38         \$16.640         12May1993 yes         39         \$35.06           7082153.         48         \$9.280         02Jan1994 no         52         \$           7082153.         48         \$9.280         02Jan1994 no         52         \$           2556136.         5         \$29.270         23Nov1993 yes         31         \$26.52           2556136.         5         \$29.270         23Nov1998 yes         31         \$26.52           2556136.         5         \$29.270         23Nov1998 yes         31         \$26.52           2556756.         24         \$23.450         22Sep1994 no         42         \$52.98           4012002.         14         \$28.560         12May1986 yes         48         \$52.98           3930798.         31         \$33.710	ID 🐧	TRAVTIM	BLUEBOO	INITDATI	RED C	AGE	INCOM <u>e</u>
8568865          48         \$18.510         ####### no         50         \$106.95           3764824:         25         \$17.600         02Dct1988 yes         55         \$59.16           2693415:         12         \$25.810         21Jan1991 no         41         \$92.84           3969722:         38         \$16.640         12May1993 yes         39         \$35.06           7082153:         48         \$9.20         02Jan1994 no         52         \$           7550377:         62         \$4.600         09Nov1989 yes         40         \$57.47           7530377:         62         \$4.600         09Nov1989 yes         41         \$52.98           40120026:         14         \$28.560         12May1986 yes         48         \$52.98           7930377:         62         \$33.320         22Feb1990 yes         46         \$52.98           7930459:         31         \$33.710         02Jul1991 no         49         \$52.98           7930459:         30         \$10.910         ####################################							\$91,449
2693415!         12         \$25,810         21Jan1991 no         41         \$92,84           3969722:         38         \$16,640         12May1993 yes         39         \$35,08           7737378!         38         \$29,450         27Sep1989 no         43         \$145,35           7082153;         48         \$9,280         02Jan1994 no         52         \$           2556136[         5         \$29,270         23Nov1993 yes         40         \$57,47           7530371;         62         \$4,600         09Nov1989 yes         31         \$26,52           21305754         24         \$23,450         22Sep1994 no         42         \$52,98           4012002[         14         \$28,860         12May1986 yes         48         \$52,98           3530798         31         \$33,710         02Jul1991 no         49         \$52,98           3939254!         30         \$10,910         ######## yes         45         \$61,93           3939254!         30         \$10,910         ########## yes         45         \$61,93           401366         8         \$30,150         ############# no         43         \$132,56           3707484         21         \$7,500 <th>8568865</th> <th>48</th> <th>\$18.510</th> <th>#######</th> <th>no</th> <th>50</th> <th>\$106,952</th>	8568865	48	\$18.510	#######	no	50	\$106,952
3969722   38	3764824	25	\$17.600	02Oct1988	yes	55	\$59,162
1737378  38	26934155	12	\$25.810	21Jan1991	no	41	\$92.842
7082153.         48         \$9.280         02Jan1994 no         52         \$           2556136(         5         \$29.270         23Nov1993 yes         40         \$57.47           7530371.         62         \$4.600         09Nov1989 yes         31         \$26.52           21305754         24         \$23.450         22Sep1994 no         42         \$52.98           4012002(         14         \$28.560         12May1986 yes         48         \$52.98           4012002(         14         \$28.560         12May1986 yes         48         \$52.98           3930254(         30         \$10.910         ######### yes         45         \$61.93           3939254(         30         \$10.910         ######## yes         45         \$61.93           7074832(         36         \$15.120         05Jun1997 no         52         \$68.93           42131636         8         \$30.150         ######### no         43         \$132.56           3707484         21         \$7.500         29Jul1985 yes         60         \$125.99           7182942(         14         \$17.550         21Oct1995 no         37         \$123.52           5398757(         33         \$29.210 <th>3969722</th> <th>38</th> <th>\$16.640</th> <th>12May1993</th> <th>yes</th> <th>39</th> <th>\$35.08</th>	3969722	38	\$16.640	12May1993	yes	39	\$35.08
2556136(         5         \$29.270         23Nov1993 yes         40         \$57.47           7530371;         62         \$4.600         09Nov1989 yes         31         \$26.52           21305754         24         \$23.450         22Sep1994 no         42         \$52.98           4012002(         14         \$28.560         12May1986 yes         48         \$52.98           3930798         31         \$33.710         02Ju11991 no         49         \$52.98           29902451         22         \$33.320         22Feb1990 yes         46         \$52.98           39392541         30         \$10.910         ######## yes         45         \$61.93           5935828         18         \$23.390         07Ju1980 yes         42         \$76.22           77048324         36         \$15.120         05Jun1997 no         52         \$68.93           42131636         8         \$30.150         ######## no         43         \$132.55           71829426         14         \$17.550         21Clet1995 no         37         \$123.52           53887571         33         \$29.210         02Jun1991 yes         47         \$45.25           5209593         53         \$13.050<	17373785	38	\$29,450	27Sep1989	no	43	\$145,353
7530371; 62 \$4.600 09Nov1989 yes 31 \$26.52 21305754 24 \$23.450 22Sep1994 no 42 \$52.98 4012002( 14 \$28.560 12May1986 yes 48 \$52.98 9350798. 31 \$33.710 02Ju11991 no 49 \$52.98 2990245) 22 \$33.320 22Feb1990 yes 46 \$52.98 3939254( 30 \$10.910 ######## yes 45 \$61.93 5935828 18 \$23.390 07Ju11980 yes 42 \$76.22 42131636 8 \$30.150 ####### no 43 \$132.56 3707484 21 \$7.500 29Ju11987 po 43 \$132.56 3707484 21 \$7.500 29Ju11985 yes 60 \$125.99 7182942( 14 \$17.550 21Cuc11995 no 37 \$123.52 5388757( 33 \$29.210 02Jun1991 yes 47 \$45.25 5209593 53 \$13.050 12Dec1993 no 40 \$75.51 5684737; 40 \$36.120 16Dec1990 yes 47 \$104.27 5820861( 11 \$17.300 ######## yes 50 \$111.42 6804259, 32 \$11.620 30Mar1995 no 51 \$50.16 93412915 50 \$14.530 09Lec1982 no 43 \$48.18 46157391 24 \$1.780 21Jun1983 no 49 \$22.05 6833784( 45 \$27.890 05Nov1989 no 55 \$55.40 5039064( 48 \$8.460 05Sep1987 yes 48 \$39.61 9707619 9 \$34.510 07Jun1988 no 45 \$23.77 8308556: 44 \$23.320 01Sep1983 no 55 \$163.15 6429873; 59 \$10.350 21Jan1982 yes 52 \$24.59 9309292( 45 \$27.020 04Apr1992 no 54 \$107.80 39176001 42 \$7.470 22Dec1987 no 39 \$59.68 60189132 10 \$6.120 23Apr1986 no 45 \$1.15 64194086 51 \$6.720 20Apr1991 no 54 \$107.80 39176001 42 \$7.470 22Dec1987 no 39 \$59.68 60189132 10 \$6.120 23Apr1986 no 55 \$163.15 6429873; 59 \$10.350 21Jan1982 yes 52 \$24.59 93092921 45 \$27.020 04Apr1992 no 54 \$107.80 3457713( 17 \$31.390 25Apr1996 yes 41 \$55.36 6429873; 59 \$10.350 21Jan1982 yes 52 \$24.59 93092921 45 \$27.020 04Apr1992 no 54 \$107.80 3457713( 17 \$31.390 25Apr1996 yes 41 \$55.36 6429873; 59 \$10.350 21Jan1982 yes 52 \$24.59 93092921 45 \$27.020 04Apr1992 no 54 \$107.80 39176001 42 \$7.470 22Dec1987 no 39 \$59.68	70821537	48	\$9.280	02Jan1994	no	52	\$0
21305754         24         \$23.450         22Sep1994 no         42         \$52.98           40120026         14         \$28.560         12May1986 yes         48         \$52.98           9350798         31         \$33.710         02Jul1991 no         49         \$52.98           2990245         22         \$33.320         22Feb1990 yes         46         \$52.98           3933254         30         \$10.910         ######## yes         45         \$61.93           5935828         18         \$23.390         07Jul1980 yes         42         \$76.22           17048324         36         \$15.120         05Jun1997 no         52         \$68.93           42131636         8         \$30.150         ######## no         43         \$132.56           42131636         8         \$30.150         ####### no         43         \$132.56           71829426         14         \$17.550         21Cct1995 no         37         \$123.52           53887571         33         \$29.210         02Jun1991 yes         47         \$145.25           5209593         53         \$13.050         \$12Dec1993 no         40         \$75.51           58208616         11         \$17.300 <th>25561360</th> <th>5</th> <th>\$29.270</th> <th>23Nov1993</th> <th>yes</th> <th>40</th> <th>\$57.474</th>	25561360	5	\$29.270	23Nov1993	yes	40	\$57.474
4012002£         14         \$28,560         12May 1986         yes         48         \$52,98           9350798         31         \$33,710         02Jul 1991         no         49         \$52,98           29902451         22         \$33,320         22Feb1990         yes         46         \$52,98           39332541         30         \$10,910         ####### yes         45         \$61,93           5935828         18         \$23,390         07Jul 1980         yes         42         \$76,22           17048324         36         \$15,120         05Jun 1997         no         52         \$68,93           42131636         8         \$30,150         ######## no         43         \$132,56           42131636         8         \$30,150         ####### no         43         \$132,52           71829426         14         \$17,550         21Clc1995         no         37         \$123,52           53887571         33         \$29,210         02Jun 1991         yes         47         \$45,25           5209593         53         \$13,050         12Dec1993         no         40         \$75,51           58208611         11         \$17,300         #########	75303717	62	\$4.600	09Nov1989	yes	31	\$26.520
\$330798   31   \$33.710   \$02.0u11991   no   49   \$52.98	21305754	24				42	\$52.988
2990245i         22         \$33,320         22Feb1990 yes         46         \$52,98           3939254!         30         \$10,910         ######## yes         45         \$61,93           5935828         18         \$23,390         07Jul1980 yes         42         \$76,22           17048324         36         \$15,120         05Jun1997 no         52         \$68,93           42131636         8         \$30,150         ######## no         43         \$132,56           3707484         21         \$7,500         29Jul1985 yes         60         \$125,89           71829426         14         \$17,550         210ct1995 no         37         \$123,52           5209593         53         \$29,210         02Jun1991 yes         47         \$45,25           5209593         53         \$13,050         \$12Dec1993 no         40         \$75,51           5684737;         40         \$36,120         \$16Dec1990 yes         47         \$104,27           5820861(         11         \$17,300         ############# yes         50         \$111,42           6804259;         32         \$11,620         30Mar1995 no         51         \$50,16           93412915         50 <th< th=""><th>40120026</th><th></th><th>\$28.560</th><th>12May1986</th><th>yes</th><th>48</th><th>\$52.988</th></th<>	40120026		\$28.560	12May1986	yes	48	\$52.988
3939254!         30         \$10.910         ########   yes         45         \$61.93           5935828         18         \$23.390         O7Jul1980         yes         42         \$76.22           17048324         36         \$15.120         05Jun1997         no         52         \$68.93           42131636         8         \$30.150         ########         no         43         \$132.56           3707484         21         \$7.500         29Jul1985         yes         60         \$125.89           71829426         14         \$17.550         210ct1995         no         37         \$123.55           53887571         33         \$29.210         02Jun1991         yes         47         \$45.25           5209593         53         \$13.050         12Dec1993         no         40         \$75.51           5684737.         40         \$36.120         16Dec1990         yes         47         \$104.27           58208610         11         \$17.300         ####################################							\$52.988
5935828         18         \$23,390         O7,Jul 1980         yes         42         \$76,22           17048324         36         \$15,120         05,Jun 1997         no         52         \$68,89           42131636         8         \$30,150         #######         no         43         \$132,56           3707484         21         \$7,500         29,Jul 1985         yes         60         \$125,89           71829426         14         \$17,550         210ct1995         no         37         \$123,52           53887571         33         \$29,210         02,Jun 1991         yes         47         \$45,25           5209593         53         \$13,050         12Dec1993         no         40         \$75,51           56847371         40         \$36,120         16Dec1990         yes         47         \$104,27           58208611         11         \$17,300         ####################################	2990245						\$52.988
1704832  36							\$61.931
42131636         8         \$30.150         ######## no         43         \$132.56           3707484         21         \$7.500         29Jul 1985         yes         60         \$125.89           7182942(         14         \$17.550         21Oct 1995         no         37         \$123.52           5388757(         33         \$29.210         02Jun 1991         yes         47         \$45.25           5209593         53         \$13.050         12Dec1993         no         40         \$75.51           5684737(         40         \$36.120         16Dec1990         yes         47         \$104.27           5820861(         11         \$17.300         #########         yes         50         \$111.42           6804259(         32         \$11.620         30Mar1995         no         51         \$50.16           93412915         50         \$14.530         09Dec1982         no         43         \$42.05           46157391         24         \$21.990         21Jun1983         no         49         \$22.05           6833784(         45         \$27.890         05Nov1989         no         55         \$55.40           5039064(         48							\$76.226
3707484         21         \$7.500         29Jul 1985 yes         60         \$125.89           71829426         14         \$17.550         21Oct 1995 no         37         \$123.52           53887571         33         \$29.210         02Jun 1991 yes         47         \$45.25           5209593         53         \$13.050         12Dec1993 no         40         \$75.51           5684737.         40         \$36.120         16Dec1990 yes         47         \$104.27           58208610         11         \$17.300         ######## yes         50         \$111.42           6804259.         32         \$11.620         30Mar 1995 no         51         \$50.16           93412915         50         \$14.530         09Dec1982 no         43         \$48.18           46157391         24         \$21.990         21Jun 1983 no         49         \$22.05           525555         35         \$12.180         ####################################							\$68.992
71829426 14 \$17.550 21Oct1995 no 37 \$123.52 53887571 33 \$29.210 02Jun1991 yes 47 \$45.25 5209593 53 \$13.050 12Dec1993 no 40 \$75.51 56847371 40 \$36.120 16Dec1990 yes 47 \$104.27 5820861( 11 \$17.300 ######## yes 50 \$111.42 6804259; 32 \$11.620 30Mar1995 no 51 \$50.16 93412915 50 \$14.530 09Dec1982 no 43 \$48.18 46157391 24 \$21.990 21Jun1983 no 49 \$22.05 22521551 35 \$12.180 ####### yes 34 \$23.55 6833784 45 \$27.890 05Nov1989 no 55 \$55.40 5039064 48 \$8.60 05Sep1987 yes 48 \$39.61 9707619 9 \$34.510 07Jun1988 no 45 \$23.77 8308556: 44 \$23.320 01Sep1983 no 55 \$163.15 6429873: 59 \$10.350 21Jan1982 yes 52 \$24.59 9309292 45 \$27.020 04Apr1992 no 54 \$107.80 39176001 42 \$7.470 22Dec1987 no 39 \$59.68 60189132 10 \$6.120 23Apr1986 no 45 \$1.15 79219601: 32 \$13.160 23Oct1984 no 43 \$1.15 3457713( 17 \$31.390 25Apr1996 yes 41 \$55.36 6429873: 59 \$10.350 21Jan1982 yes 52 \$24.59 93092921 45 \$27.020 04Apr1992 no 54 \$107.80 53176001 42 \$7.470 22Dec1987 no 39 \$59.68 6429873: 59 \$10.350 21Jan1982 yes 52 \$24.59 93092921 45 \$27.020 04Apr1992 no 54 \$107.80 53457713( 17 \$31.390 25Apr1996 yes 41 \$55.36 5429873: 59 \$10.350 21Jan1982 yes 52 \$24.59 93092921 45 \$27.020 04Apr1992 no 54 \$107.80 3457713( 17 \$31.390 25Apr1996 yes 41 \$55.36 5429873: 59 \$10.350 21Jan1982 yes 52 \$24.59 93092921 45 \$27.020 04Apr1992 no 54 \$107.80 3457713( 17 \$31.90 25Apr1996 no 55 \$163.15 5429873: 59 \$10.350 21Jan1982 yes 52 \$24.59 93092921 45 \$27.020 04Apr1992 no 54 \$107.80 39176001 42 \$7.470 22Dec1987 no 39 \$59.68		_					\$132.561
5388757!         33         \$29.210         02Jun1991 yes         47         \$45.25           5209593         53         \$13.050         12Dec1993 no         40         \$75.51           5684737!         40         \$36.120         16Dec1990 yes         47         \$104.27           5820861[         11         \$17.300         ######## yes         50         \$111.42           6804259!         32         \$11.620         30Mar1995 no         51         \$50.16           93412915         50         \$14.530         09Dec1982 no         43         \$48.18           46157391         24         \$21.990         21Jun1983 no         49         \$22.05           22521555:         35         \$12.180         ######## yes         34         \$23.57           5039064:         48         \$8.460         05Sep1987 yes         48         \$39.61           49707619         9         \$34.510         07Jun1988 no         45         \$23.77           8308556:         44         \$23.320         01Sep1983 no         55         \$163.15           6429873!         59         \$10.350         21Jan1982 yes         52         \$24.59           9309292!         45         \$27.							
5209593         53         \$13.050         12Dec1993 no         40         \$75.51           5684737;         40         \$36.120         16Dec1990 yes         47         \$104.27           5820861(         11         \$17.300         ####### yes         50         \$111.42           6804259;         32         \$11.620         30Mar1995 no         51         \$50.16           93412915         50         \$14.530         09Dec1982 no         43         \$48.18           46157391         24         \$21.990         21Jun1983 no         49         \$22.05           2252155;         35         \$12.180         ########## yes         34         \$23.57           5033064;         48         \$27.890         05Nov1989 no         55         \$55.40           5039064;         48         \$8.460         05Sep1987 yes         48         \$39.61           79707619         9         \$34.510         07Jun1988 no         45         \$23.77           8308556;         44         \$23.320         01Sep1983 no         55         \$163.15           6429873;         59         \$10.350         21Jan1982 yes         52         \$24.59           9309292;         45         \$27.0							
5684737:         40         \$36.120         16Dec1990 yes         47         \$104.27           5820861(         11         \$17.300         ######## yes         50         \$111.42           6804259:         32         \$11.620         30Mar1995 no         51         \$50.16           93412915         50         \$14.530         09Dec1982 no         43         \$48.18           46157391         24         \$21.990         21Jun1983 no         49         \$22.05           22521555:         35         \$12.180         ######## yes         34         \$23.57           68337841         45         \$27.890         05Nov1989 no         55         \$55.40           5039064.         48         \$8.460         05Sep1987 yes         48         \$39.61           79707619         9         \$34.510         07Jun1988 no         45         \$23.77           8308556:         44         \$23.320         01Sep1983 no         55         \$163.15           6429873:         59         \$10.350         21Jan1982 yes         52         \$24.59           9309292:         45         \$27.020         04Apr1992 no         54         \$107.80           39176001         42         \$7.							
5820861(         11         \$17.300         ######## yes         50         \$11.42           6804259;         32         \$11.620         30Mar1995 no         51         \$50.16           93412915         50         \$14.530         09Dec1982 no         43         \$48.18           46157391         24         \$21.990         21Jun1983 no         49         \$22.05           2252155;         35         \$12.180         ######## yes         34         \$23.57           6833784i         45         \$27.890         05Nov1989 no         55         \$55.40           5039064;         48         \$8.460         05Sep1987 yes         48         \$39.61           79707619         9         \$34.510         07Jun1988 no         45         \$23.77           8308556;         44         \$23.320         01Sep1983 no         55         \$163.15           6429873!         59         \$10.350         21Jan1982 yes         52         \$24.59           9309292!         45         \$27.020         04Apr1992 no         54         \$107.80           39176001         42         \$7.470         22Dec1987 no         39         \$59.68           60189132         10         \$6.120 </th <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>							
6804259;         32         \$11,620         30Mar1995 no         51         \$50,16           93412915         50         \$14,530         09Dec1982 no         43         \$48,18           46157391         24         \$21,990         21Jun1983 no         49         \$22,05           22521555;         35         \$12,180         ######## yes         34         \$23,57           6833784!         45         \$27,890         05Nov1989 no         55         \$55,40           5039064;         48         \$8,460         05Sep1987 yes         48         \$39,61           19707619         9         \$34,510         07Jun1988 no         45         \$23,77           8308556;         44         \$23,320         01Sep1983 no         55         \$163,15           6429873!         59         \$10,350         21Jan1982 yes         52         \$24,59           39176001         42         \$7,470         22Dec1987 no         54         \$107,80           39176001         42         \$7,470         22Dec1987 no         45         \$1,15           3457713         17         \$31,390         25Apr1996 yes         41         \$55,36           30919605         32         \$13,550 </th <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>							
93412915         50         \$14,530         09Dec1982 no         43         \$48,18           46157391         24         \$21,990         21Jun1983 no         49         \$22,05           22521555         35         \$12,180         ######## yes         34         \$23,57           68337841         45         \$27,890         05Nov1989 no         55         \$55,40           5039064         48         \$8,460         05Sep1987 yes         48         \$39,61           79707619         9         \$34,510         07Jun1988 no         45         \$23,77           8308556         44         \$23,320         01Sep1983 no         55         \$163,15           6429873!         59         \$10,350         21Jan1982 yes         52         \$24,59           39176001         42         \$7,470         22Dec1987 no         39         \$59,68           60189132         10         \$6,120         23Apr1986 no         45         \$1,15           79219605         32         \$13,160         23Cu1984 no         43         \$1,15           3457713         17         \$31,390         25Apr1996 yes         41         \$55,36           542973!         59         \$10,350							
46157391         24         \$21,990         21Jun1983 no         49         \$22,05           22521555         35         \$12,180         ######## yes         34         \$23,57           68337841         45         \$27,890         05Nov1989 no         55         \$55,40           5039064.         48         \$8,460         05Sep1987 yes         48         \$39,61           79707619         9         \$34,510         07Jun1988 no         45         \$23,77           8308556.         44         \$23,320         01Sep1983 no         55         \$163,15           6429873!         59         \$10,350         21Jan1982 yes         52         \$24,59           39309292!         45         \$27,020         04Apr1992 no         54         \$107,80           39176001         42         \$7,470         22Dec1987 no         39         \$59,68           60189132         10         \$6,120         23Apr1986 no         45         \$1,15           79219605         32         \$13,160         23Cet1984 no         43         \$1,15           79219605         32         \$13,160         23Cet1984 no         45         \$15,36           6429873!         59         \$10,350 <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>							
2252155!         35         \$12.180         ######## yes         34         \$23.57           6833784!         45         \$27.890         05Nov1989 no         55         \$55.40           5039064!         48         \$8.460         05Sep1987 yes         48         \$39.61           19707619         9         \$34.510         07Jun1988 no         45         \$23.77           8308556:         44         \$23.320         01Sep1983 no         55         \$163.15           6429873!         59         \$10.350         21Jan1982 yes         52         \$24.59           9309292!         45         \$27.020         04Apr1992 no         54         \$107.80           39176001         42         \$7.470         22Dec1987 no         39         \$59.68           60189132         10         \$6.120         23Apr1986 no         45         \$1.15           7921960!         32         \$13.160         23Cet1984 no         43         \$1.15           3457713          17         \$31.390         25Apr1996 yes         41         \$55.36           5429873          59         \$10.350         21Jan1982 yes         52         \$24.59           9309292          45         \$27.020 </th <th></th> <th></th> <th>*</th> <th></th> <th></th> <th></th> <th></th>			*				
6833784!         45         \$27.890         05Nov1989 no         55         \$55.40           5039064.         48         \$8.460         05Sep1987 yes         48         \$39.61           19707619         9         \$34.510         07Jun1988 no         45         \$23.77           8308556.         44         \$23.320         01Sep1983 no         55         \$163.15           6429873!         59         \$10.350         21Jan1982 yes         52         \$24.59           9309292!         45         \$27.020         04Apr1992 no         54         \$107.80           39176001         42         \$7.470         22Dec1987 no         39         \$59.68           60189132         10         \$6.120         23Apr1986 no         45         \$1.15           79219605         32         \$13.160         23Cet1984 no         43         \$1.15           34577131         17         \$31.390         25Apr1996 yes         41         \$55.36           8429873!         59         \$10.350         21Jan1982 yes         52         \$24.59           9303292!         45         \$27.020         04Apr1992 no         54         \$107.80           9303292!         45         \$27.020<							
5039064:         48         \$8.460         05Sep1987 yes         48         \$39.61           19707619         9         \$34.510         07Jun1988 no         45         \$23.77           8308556:         44         \$23.320         01Sep1983 no         55         \$163.15           6429873:         59         \$10.350         21Jan1982 yes         52         \$24.59           9309292:         45         \$27.020         04Apr1992 no         54         \$107.80           39176001         42         \$7.470         22Dec1987 no         39         \$59.68           60189132         10         \$6.120         23Apr1986 no         45         \$1.15           7921960!         32         \$13.160         23Cct1984 no         43         \$1.15           3457713(         17         \$31.390         25Apr1996 yes         41         \$55.36           8308556:         44         \$23.320         01Sep1983 no         55         \$163.15           6429873:         59         \$10.350         21Jan1982 yes         52         \$24.59           9309292:         45         \$27.020         04Apr1992 no         54         \$107.80           39176001         42         \$7.470<					/		
19707619         9         \$34.510         07Jun1988 no         45         \$23.77           8308556:         44         \$23.320         01Sep1983 no         55         \$163.15           6429873!         59         \$10.350         21Jan1982 yes         52         \$24.59           9309292!         45         \$27.020         04Apr1992 no         54         \$107.80           39176001         42         \$7.470         22Dec1987 no         39         \$59.68           60189132         10         \$6.120         23Apr1986 no         45         \$1.15           7921960!         32         \$13.160         23Cet1984 no         43         \$1.15           3457713(         17         \$31.390         25Apr1996 yes         41         \$55.36           8429873!         59         \$10.350         21Jan1982 yes         52         \$24.59           9309292          45         \$27.020         04Apr1992 no         54         \$107.80           39176001         42         \$7.470         22Dec1987 no         39         \$59.68           84194086         51         \$6.720         26Jun1991 no         44         \$146.26           60189132         10         \$6.120 <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>							
8308556:         44         \$23.320         01Sep1983 no         55         \$163.15           6429873!         59         \$10.350         21Jan1982 yes         52         \$24.59           9309292!         45         \$27.020         04Apr1992 no         54         \$107.80           39176001         42         \$7.470         22Dec1987 no         39         \$59.68           60189132         10         \$6.120         23Apr1986 no         45         \$1.15           7921960!         32         \$13.160         23Cet1984 no         43         \$1.15           3457713(         17         \$31.390         25Apr1996 yes         41         \$55.36           8308556:         44         \$23.320         01Sep1983 no         55         \$163.15           6429873!         59         \$10.350         21Jan1982 yes         52         \$24.59           3309292!         45         \$27.020         04Apr1992 no         54         \$107.80           39176001         42         \$7.470         22Dec1987 no         39         \$59.68           84194086         51         \$6.720         26Jun1991 no         44         \$146.26           60189132         10         \$6.120<							
6429873!         59         \$10.350         21Jan1982 yes         52         \$24.59           9309292!         45         \$27.020         04Apr1992 no         54         \$107.80           39176001         42         \$7.470         22Dec1987 no         39         \$59.68           60189132         10         \$6.120         23Apr1986 no         45         \$1.15           7921960!         32         \$13.160         23Cet1984 no         43         \$1.15           3457713(         17         \$31.390         25Apr1996 yes         41         \$55.36           8308556:         44         \$23.320         01Sep1983 no         55         \$163.15           6429873!         59         \$10.350         21Jan1982 yes         52         \$24.59           39309292!         45         \$27.020         04Apr1992 no         54         \$107.80           39176001         42         \$7.470         22Dec1987 no         33         \$59.68           84194086         51         \$6.720         26Jun1991 no         44         \$146.26           60189132         10         \$6.120         23Apr1986 no         45         \$1.15		_					
9309292         45         \$27,020         04Apr1992 no         54         \$107,80           39176001         42         \$7,470         22Dec1987 no         39         \$59,68           60189132         10         \$6,120         23Apr1986 no         45         \$1,15           7921960!         32         \$13,160         23Oct1984 no         43         \$1,15           3457713(         17         \$31,390         25Apr1996 yes         41         \$55,36           8308556:         44         \$23,320         01Sep1983 no         55         \$163,15           6429873:         59         \$10,350         21Jan1982 yes         52         \$24,59           9309292:         45         \$27,020         04Apr1992 no         54         \$107,80           39176001         42         \$7,470         22Dec1987 no         39         \$59,68           8419408(         51         \$6,720         26Jun1991 no         44         \$146,26           60189132         10         \$6,120         23Apr1986 no         45         \$1,15							
39176001         42         \$7.470         22Dec1987 no         39         \$59.68           60189132         10         \$6.120         23Apr1986 no         45         \$1.15           79219605         32         \$13.160         23Oct1984 no         43         \$1.15           34577130         17         \$31.390         25Apr1996 yes         41         \$55.36           8308556:         44         \$23.320         01Sep1983 no         55         \$163.15           6429873:         59         \$10.350         21Jan1982 yes         52         \$24.59           9309292i         45         \$27.020         04Apr1992 no         54         \$107.80           39176001         42         \$7.470         22Dec1987 no         39         \$59.68           84194086         51         \$6.720         26Jun1991 no         44         \$146.26           60189132         10         \$6.120         23Apr1986 no         45         \$1.15							
60189132         10         \$6.120         23Apr1986 no         45         \$1.15           7921960!         32         \$13.160         23Clct1984 no         43         \$1.15           3457713(         17         \$31.390         25Apr1996 yes         41         \$55.36           8308556:         44         \$23.320         01Sep1983 no         55         \$163.15           6429873!         59         \$10.350         21Jan1982 yes         52         \$24.59           9309292!         45         \$27.020         04Apr1992 no         54         \$107.80           39176001         42         \$7.470         22Dec1987 no         39         \$59.68           84194086         51         \$6.720         26Jun1991 no         44         \$146.26           60189132         10         \$6.120         23Apr1986 no         45         \$1.15							
7921960!         32         \$13.160         23Oct1984 no         43         \$1.15           3457713(         17         \$31.390         25Apri996 yes         41         \$55.36           8308556:         44         \$23.320         01Sep1983 no         55         \$163.15           6429873:         59         \$10.350         21Jan1982 yes         52         \$24.59           9309292:         45         \$27.020         04Apri992 no         54         \$107.80           39176001         42         \$7.470         22Dec1987 no         39         \$59.68           8419408(         51         \$6.720         26Jun1991 no         44         \$146.26           60189132         10         \$6.120         23Apri986 no         45         \$1.15							
34577131         17         \$31.390         25Apr1996 yes         41         \$55.36           8308556:         44         \$23.320         01Sep1983 no         55         \$163.15           6429873!         59         \$10.350         21Jan1982 yes         52         \$24.59           9309292!         45         \$27.020         04Apr1992 no         54         \$107.80           39176001         42         \$7.470         22Dec1987 no         39         \$59.68           84194086         51         \$6.720         26Jun1991 no         44         \$146.26           60189132         10         \$6.120         23Apr1986 no         45         \$1.15							
8308556:         44         \$23.320         01Sep1983 no         55         \$163.15           6429873!         59         \$10.350         21Jan1982 yes         52         \$24.59           9309292!         45         \$27.020         04Apr1992 no         54         \$107.80           39176001         42         \$7.470         22Dec1987 no         39         \$59.68           84194086         51         \$6.720         26Jun1991 no         44         \$146.26           60189132         10         \$6.120         23Apr1986 no         45         \$1.15							
6429873!         59         \$10.350         21Jan1982 yes         52         \$24.59           9309292!         45         \$27.020         04Apr1992 no         54         \$107.80           39176001         42         \$7.470         22Dec1987 no         39         \$59.68           84194086         51         \$6.720         26Jun1991 no         44         \$146.26           60189132         10         \$6.120         23Apr1986 no         45         \$1.15							
9309292I         45         \$27,020         04Apr 1992 no         54         \$107,80           39176001         42         \$7,470         22Dec1987 no         39         \$59,68           84194086         51         \$6,720         26Jun1991 no         44         \$146,26           60189132         10         \$6,120         23Apr 1986 no         45         \$1,15							
39176001         42         \$7.470         22Dec1987 no         39         \$59.68           84194086         51         \$6.720         26Jun1991 no         44         \$146.26           760189132         10         \$6.120         23Apr1986 no         45         \$1.15							
84194086         51         \$6.720         26Jun1991 no         44         \$146.26           760189132         10         \$6.120         23Apri 1986 no         45         \$1.15							•
<b>60189132</b> 10 \$6.120 23Apr1986 no 45 \$1.15							
	r 32 (300)	32	φ13, 100	200001004	I IO	43	φ1. IJ0

## Ergebnisse aus den Datenquantitäts Scenarios

- Grenznutzen wird im Bereich von 500 bis 1000 Ereignissen flach
- Auch Nicht-Ereignisse bieten zusätzliche Informationen, insbesondere im Bereich von bis zu 500 Ereignissen

Varying the number of events and non-events



### Kontinuierliches Erweitern der verfügbaren Länge der Historie im Zeitreihen-Forecasting

#### Fachliche Fragen

- Ist es möglich, Zeitreihen-Analyse zu starten, wenn nur 18 Monate Historie verfügbar sind?
- Soll auf weitere Historien-Monate gewartet werden?
- Was ist der Vorteil zusätzlichen Daten-Management-Aufwands?
- Was ist die beste Länge der Datenhistorie für Zeitreihen Prognose?

#### Methoden

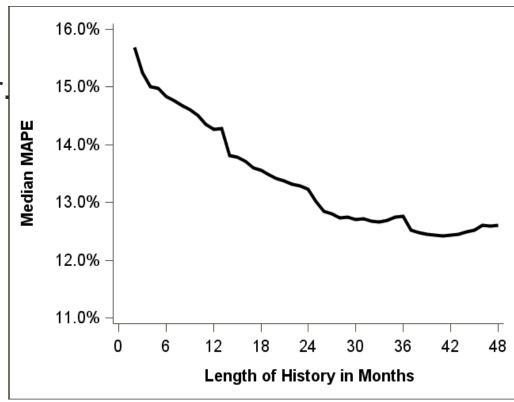
- Simulationsumgebung mit SAS High Performance Forecasting
- Auf Basis von 788 Zeitreihen auf monatlichen Daten aus verschiedenen Branchen
- Minimum Historie für jede Zeitreihe: 48 Monate
- Eingeschränkt auf Prognoseverfahren "exponentielle Glättung
- Validierung auf MAPE auf 12 Zukunftsmonate
- Iteration durch Verschieben der "Zero-Time"





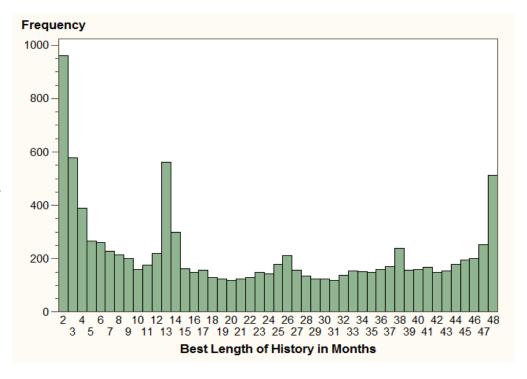
#### Wie weit sollten wir uns "zurückerinnern"?

- Die erwartete Abnahme des MAPE mit zunehmender Historienlänge ist erkennbar.
- Es gibt eine exponentielle Abnahme der zusätzlichen Werts weiterer Monate.
- Stärkere Sprünge nach 12, 24 und 36 Monaten sind sichtbar.



# Was ist die beste Länge der Datenhistorie für Zeitreihen-Forecasting?

- Methode: für jede Zeitreihe wurde ermittelt, welche Anzahl von Historien-Monaten den kleinsten Fehler für die künftigen 12 Monate ergibt.
- Ergebnis: nicht in allen Fällen ist es vorteilhaft die lange Zeithistorie zu verwenden.



#### **Abschließende Kommentare**

- Datenqualität für Analytik ist mehr
  - Mehr Anforderungen
  - Mehr Möglichkeiten
- Gehen Sie ins Detail!
  - Zufällige oder systematische Fehler?
  - Permanenten oder historische/temporäre Qualitätsprobleme?
- Datenmenge macht einen Unterschied!
  - Aber balancieren Sie Aufwand und Nutzen!
- SAS unterstützt beim
  - Profiling, Verbessern, Bewerten, Simulieren der Datenqualität.