

Inhalt

Vorwort	V
1 Übersicht zu Band 1 und 2	1
2 Fokus: Missings	19
2.1 Von Anfang an: Definition und Definieren von Missings.....	22
2.2 Abfragen für Missings.....	26
2.3 Missings in Aggregierungsfunktionen	30
2.4 Mögliche unerwünschte Ergebnisse bei (u.a.) WHERE, GROUP und ORDER.....	35
2.5 Mögliche unerwünschte Ergebnisse bei Joins	41
2.5.1 Beispiele I: Self-Joins bei einer Tabelle.....	41
2.5.2 Beispiele II: Missings in mehreren Tabellen.....	44
2.6 Suchen und Ersetzen von Missings	49
2.6.1 Suchen von Missings (Screening).....	49
2.6.2 Suchen und Ersetzen von Missings (Konvertieren).....	51
2.7 Prädikate (ALL/ANY/SOME, EXISTS, IN, IS, LIKE usw.).....	62
3 Fokus: Datenqualität mit PROC SQL	67
3.1 Integrity Constraints (Prüfregeln) und Audit Trails	68
3.1.1 Integrity Constraints (Prüfregeln).....	69
3.1.2 Audit Trails (Prüfprotokolle).....	81
3.2 Identifizieren und Filtern von mehrfachen Werten (Doppelte)	91
3.3 Identifizieren und Filtern von Ausreißern	101
3.4 Identifizieren, Filtern oder Ersetzen un(erwünschter) Zeichen(ketten).....	108
4 Fokus: Makroprogrammierung mit SQL	117
4.1 Makrovariablen	120
4.1.1 Automatische SAS Makrovariablen.....	122
4.1.2 Automatische SQL Makrovariablen.....	126
4.1.3 Anwenderdefinierte SAS Makrovariablen (INTO)	131

4.1.4	Makrovariablen, INTO und möglicher Präzisionsverlust.....	134
4.1.5	Die vielen Wege zu einer (Makro)Variablen (Bezüge).....	136
4.2	Makroprogramme mit PROC SQL	137
4.2.1	Was sind Makroprogramme?	138
4.2.2	Let's do it: Erste SAS Makros mit dem %LET-Statement.....	143
4.2.3	Listenweises Ausführen von Befehlen	146
4.2.4	Bedingungsgeleitetes Ausführen von Befehlen	154
4.2.5	Tipps für den Einsatz von Makros	159
4.3	Elemente der SAS Makrosprache	163
4.3.1	SAS Makrofunktionen	163
4.3.2	SAS Makro Statements (z.B. %DO..., %IF...)	182
4.3.3	Schnittstellen I: Von der SAS Macro Facility zu PROC SQL bzw. DATA Step...	190
4.3.4	Schnittstellen II: Von PROC SQL in die SAS Macro Facility (INTO:)	198
4.4	Anwendung 1: Zeilenweises Aktualisieren von Daten mit Sicherheitsprüfung.....	204
4.5	Anwendung 2: Arbeiten mit mehreren Dateien (Aufteilen).....	207
4.5.1	Aufteilen eines Datensatzes in einheitlich gefilterte Subdatensätze (Split-Variable ist vom Typ „String“)	207
4.5.2	Aufteilen eines Datensatzes in einheitlich gefilterte Subdatensätze (Split-Variable ist vom Typ „numerisch“)	210
4.6	Anwendung 3: „Drehen“ einer SAS Tabelle (stack/unstack).....	213
4.6.1	Umwandeln von stack nach unstack („aus 1 mach 3“)	215
4.6.2	Umwandeln von unstack nach stack („aus 3 mach 1“)	222
4.7	Anwendung 4: Makros für den Abruf von Systeminformationen.....	229
4.8	Anwendung 5: Anlegen von Verzeichnissen für das Ablegen von Daten.....	235
4.9	Anwendung 6: Fortlaufende „exotische“ Spaltennamen („2010“, „2011“, ...)	237
4.10	Anwendung 7: Konvertieren ganzer Variablenlisten von String nach numerisch..	240
5	Fokus: SQL für Geodaten	245
5.1	Geodaten und Distanzen	245
5.1.1	Distanzen im zweidimensionalen Raum (Basis: Metrische Koordinaten)	247
5.1.2	Distanzen im sphärischen Raum (Basis: Längen- und Breitengrade)	251
5.1.3	Drei Beispiele für SQL Abfragen für Koordinatendaten.....	256
5.2	SQL und Karten	259
6	Fokus: Hashing als Alternative zu SQL	267
6.1	Was ist Hash Programmieren?	268
6.2	Arbeiten innerhalb einer Tabelle.....	270
6.2.1	Aggregieren mit SQL, SUMMARY und Hash	270
6.2.2	Sortieren mit SORT, SQL und Hash.....	276

6.2.3	Subsetting: Filtern und zufälliges Ziehen	277
6.2.4	Eliminieren von doppelten Keys	283
6.2.5	Abfragen von Werten (Retrieval).....	285
6.3	Arbeiten mit mehreren Tabellen.....	291
6.3.1	Joins (Inner, Outer) mit zwei Tabellen.....	292
6.3.2	Fuzzy-Join mit zwei Tabellen	297
6.3.3	Splitten einer (nicht) sortierten SAS Tabelle.....	301
6.4	Übersicht: Elemente des Hash Programming	304
7	Fokus: Performanz und Effizienz	307
7.1	Eine Strategie als SAS Programm.....	310
7.2	Weniger ist mehr: Eingrenzen großer Tabellen auf das Wesentliche.....	313
7.3	Noch mehr Luft rauslassen: Verkürzen und Komprimieren.....	319
7.4	Sortieren? Je weniger, desto besser	323
7.5	Beschleunigen: Besondere Tricks für besondere Anlässe (SQL und mehr).....	329
7.5.1	Grundlegende Techniken	329
7.5.2	SQL-spezifische Techniken.....	332
7.5.3	Weitere Techniken	336
7.6	Datenverarbeitung in SAS oder im DBMS: Abstimmen von SQL auf DBMS	339
7.7	Schritt für Schritt zu mehr Performanz: Performanz als Strategie	344
8	Hilfen, Tipps und Tricks	349
8.1	Laufzeit als Schlüssel zur Performanz.....	349
8.2	Bescheid wissen: SAS Dictionaries.....	359
8.3	Datenhandling und Datenstrukturierung	366
8.4	Aktualisieren von Tabellen (SQL vs. DATA Step).....	376
9	SAS Syntax – PROC SQL, SAS Funktionen und SAS Routinen	387
9.1	PROC SQL Syntax (Übersicht).....	388
9.2	SAS Funktionen und Funktionsaufrufe (Übersicht).....	393
9.3	Pass-through Facility (Syntax, Besonderheiten).....	430
10	Verwendete SAS Datasets	443
11	Literatur	445

12	Ihre Meinung zu diesem Buch	449
13	Autor	451
	Syntaxverzeichnis	453
	Sachverzeichnis	469