

Contents

About This Book	xv
About The Author	xix
Acknowledgments	xxi
Chapter 1: Designing Database Tables	1
Introduction	2
Database Design	2
Conceptual View	2
Table Definitions	3
Redundant Information	3
Normalization	4
Normalization Strategies	5
Column Names and Reserved Words	7
ANSI SQL Reserved Words.....	8
SQL Code.....	8
Data Integrity	8
Referential Integrity	9
Database Tables Used in This Book	9
CUSTOMERS Table	9
INVENTORY Table	10
INVOICE Table.....	10
MANUFACTURERS Table	10
PRODUCTS Table	11
PURCHASES Table	11
Table Contents	12
The Database Structure	14
Sample Database Tables	14
Summary	21

Chapter 2: Working with Data in PROC SQL	23
Introduction	24
Overview of Data Types.....	24
Numeric Data	24
Date and Time Column Definitions	27
Character Data.....	28
Missing Values and NULL	28
Arithmetic and Missing Data	29
SQL Keywords	32
SQL Operators and Functions	35
Comparison Operators.....	35
Logical Operators	36
Arithmetic Operators.....	38
Character String Operators and Functions.....	40
Summarizing Data	58
Predicates.....	62
Dictionary Tables	72
Dictionary Tables and Metadata	72
Displaying Dictionary Table Definitions	74
Dictionary Table Column Names	75
Accessing a Dictionary Table's Contents	78
Summary	89
Chapter 3: Formatting Output.....	91
Introduction	92
Formatting Output.....	92
Writing a Blank Line between Each Row	92
Displaying Row Numbers.....	93
Using the FORMAT= Column Modifier to Format Output	96
Concatenating Character Strings	97
Inserting Text and Constants between Columns	99
Using Scalar Expressions with Selected Columns	101
Ordering Output by Columns	104
Grouping Data with Summary Functions	107
Grouping Data and Sorting.....	109

Subsetting Groups with the HAVING Clause	110
Formatting Output with the Output Delivery System	112
ODS and Output Formats.....	113
Sending Output to a SAS Data Set	114
Converting Output to Rich Text Format	115
Exporting Data and Output to Excel	116
Delivering Results to the Web	118
Summary	119
Chapter 4: Coding PROC SQL Logic	121
Introduction	122
Conditional Logic	122
SQL Code.....	122
SQL Code.....	122
SQL Code.....	123
SQL Code.....	123
CASE Expressions.....	123
Simple Case Expression	124
Searched CASE Expression.....	137
Case Logic versus COALESCE Expression	142
Assigning Labels and Grouping Data	143
Logic and Nulls	146
Interfacing PROC SQL with the Macro Language	148
Exploring Macro Variables and Values.....	149
Creating Multiple Macro Variables	153
Using Automatic Macro Variables to Control Processing.....	156
Building Macro Tools and Applications	158
Creating Simple Macro Tools.....	158
Cross-Referencing Columns	158
Determining the Number of Rows in a Table.....	159
Identifying Duplicate Rows in a Table	160
Summary	161
Chapter 5: Creating, Populating, and Deleting Tables	163
Introduction	164
Creating Tables	164

Creating a Table Using Column-Definition Lists	165
Creating a Table Using the LIKE Clause	169
Deriving a Table and Data from an Existing Table	170
Populating Tables.....	172
Adding Data to a Table with a SET Clause	173
Adding Data to All of the Columns in a Row.....	176
Adding Data to Some of the Columns in a Row	181
Adding Data with a SELECT Query	185
Bulk Loading Data from Microsoft Excel	186
Integrity Constraints	192
Defining Integrity Constraints	192
Types of Integrity Constraints	192
Preventing Null Values with a NOT NULL Constraint	192
Enforcing Unique Values with a UNIQUE Constraint.....	195
Validating Column Values with a CHECK Constraint.....	196
Referential Integrity Constraints	197
Establishing a Primary Key.....	198
Establishing a Foreign Key	199
Displaying Integrity Constraints.....	202
Deleting Rows in a Table	203
Deleting a Single Row in a Table	203
Deleting More Than One Row in a Table.....	204
Deleting All Rows in a Table	204
Deleting Tables.....	205
Deleting a Single Table	205
Deleting Multiple Tables	206
Deleting Tables That Contain Integrity Constraints.....	206
Summary	208
Chapter 6: Modifying and Updating Tables and Indexes	209
Introduction	210
Modifying Tables	210
Adding New Columns.....	210
Controlling the Position of Columns in a Table.....	212
Changing a Column's Length	214
Changing a Column's Format.....	218

Changing a Column's Label.....	219
Renaming a Column	219
Renaming a Table	221
Indexes	222
Designing Indexes	224
Cardinality.....	225
Index Selectivity.....	225
Defining Indexes	227
Creating a Simple Index	228
Creating a Composite Index.....	228
Preventing Duplicate Values in an Index.....	229
Modifying Columns Containing Indexes	229
Deleting (Dropping) Indexes	229
Updating Data in a Table	230
Summary	232
Chapter 7: Coding Complex Queries	233
Introduction	234
Introducing Complex Queries	234
Joins	235
Why Joins Are Important.....	235
Information Retrieval Based on Relationships	235
DATA Step Merges versus PROC SQL Joins.....	236
Types of Complex Queries.....	236
Demystifying Join Algorithms	239
Influencing Joins with a Little Magic	239
Cartesian Product Joins	242
Inner Joins	242
Equijoins	243
Non-Equijoins.....	245
Reflexive or Self Joins	247
Using Table Aliases in Joins.....	249
Performing Computations in Joins	250
Joins with Three Tables	251
Joins with More Than Three Tables	253

Outer Joins.....	255
Left Outer Joins	255
Right Outer Joins	258
Full Outer Joins	259
Subqueries	261
Alternate Approaches to Subqueries	261
Passing a Single Value with a Subquery	262
Passing More Than One Row with a Subquery	266
Comparing a Set of Values	267
Correlated Subqueries	269
Set Operations.....	271
Rules for Set Operators	271
Set Operators and Precedence.....	272
Accessing Rows from the Intersection of Two Queries	272
Accessing Rows from the Combination of Two Queries	274
Concatenating Rows from Two Queries	276
Comparing Rows from Two Queries	278
Complex Query Applications	280
One-to-One, One-to-Many, Many-to-One, and Many-to-Many Relationships.....	280
Processing First, Last, and Between Rows for BY-and Groups.....	285
Determining the Number of Rows in an Input Table.....	290
Identifying Tables with the Most Indexes	291
Summary.....	293
Chapter 8: Working with Views	295
Introduction	296
Views—Windows to Your Data	296
What Views Aren't.....	297
Types of Views	297
Creating Views	299
Displaying a View's Contents	300
Describing View Definitions.....	301
Creating and Using Views in SAS	302
Views and SAS Procedures	303
Views and DATA Steps.....	305
Eliminating Redundancy.....	307

Restricting Data Access— Security	307
Hiding Logic Complexities	308
Nesting Views	310
Updatable Views.....	312
Inserting New Rows of Data.....	313
Updating Existing Rows of Data	317
Deleting Rows of Data	320
Deleting Views	321
Summary	322
Chapter 9: Troubleshooting and Debugging	323
Introduction	323
The World of Bugs.....	324
The Debugging Process	324
Types of Problems	326
Troubleshooting and Debugging Techniques	327
Validating Queries with the VALIDATE Statement	327
Documented PROC SQL Options and Statement	328
Undocumented PROC SQL Options.....	342
Macro Variables	343
Troubleshooting and Debugging Examples.....	345
Summary	350
Chapter 10: Tuning for Performance and Efficiency.....	351
Introduction	351
Understanding Performance Tuning	352
Sorting and Performance	352
User-Specified Sorting (SORTPGM= System Options)	353
Automatic Sorting.....	353
Grouping and Performance.....	354
Splitting Tables.....	354
Indexes and Performance	354
Reviewing CONTENTS Output and System Messages	355
Optimizing WHERE Clause Processing with Indexes	358
Constructing Efficient Logic Conditions	359
Avoiding UNIONS	361

Summary	365
Index	367