Osborne McGraw-Hill

Using SQL

Includes Coverage of Oracle®, DB2™, SQL Server, dBASE IV®, SQL/DS, Ingres™, OS/2® Extended Edition, Informix®, Sybase®, SQLBase®, VAX® SQL™

James R. Groff &
Paul N. Weinberg

Using SQL

James R. Groff and Paul N. Weinberg

Osborne **McGraw-Hill**

Berkeley New York St. Louis San Francisco
Auckland Bogotá Hamburg London Madrid
Mexico City Milan Montreal New Delhi Panama City
Paris São Paulo Singapore Sydney
Tokyo Toronto

Osborne **McGraw-Hill**
2600 Tenth Street
Berkeley, California 94710
U.S.A.

For information on translations and book distributors outside of the U.S.A., please write to Osborne **McGraw-Hill** at the above address.

A complete list of trademarks appears on page 667.

Using SQL

1234567890 DOC 99876543210

ISBN 0-07-881524-X

Acknowledgments

This book would not have happened without the dedication and hard work of many people at Osborne/McGraw-Hill. In particular, we would like to thank Cindy Hudson for her patience, Liz Fisher for keeping the faith, Judy Wohlfrom for her tolerance, Ilene Shapera for keeping the process moving, and Judith Brown, Nancy Beckus, and Ann Spivack for their meticulous attention to detail. Also, special thanks to Joe Brilando at Ashton-Tate, for asking the question that inspired this book.

Preface

Using SQL provides a comprehensive, in-depth treatment of the SQL language for both technical and non-technical users, programmers, data processing professionals and managers who want to understand the impact of SQL in the computer market. This book offers a conceptual framework for understanding and using SQL, describes the history of SQL and SQL standards, and explains the role of SQL in the computer industry today. It will show you, step-by-step, how to use SQL features, with many illustrations and realistic examples to clarify SQL concepts. The book also compares SQL products from leading DBMS vendors, describing their advantages, benefits, and trade-offs, to help you select the right product for your application.

In some of the chapters in this book, the subject matter is explored at two different levels—a fundamental description of the topic, and an advanced discussion intended for computer professionals who need to understand some of the "internals" behind SQL. The more advanced information is covered in sections marked with an asterisk (*). You do not need to read these sections to obtain an understanding of what SQL is and what it does.

How this Book Is Organized

The book is divided into six parts that cover various aspects of the SQL language:

- Part One, "An Overview of SQL," provides an introduction to SQL and a market perspective of its role as a database language. Its four chapters describe the history of SQL, the evolution of SQL standards, and how SQL relates to the relational data model and to earlier database technologies. Part One also contains a quick tour of SQL that briefly illustrates its most important features and provides you with an overview of the entire language early in the book.
- Part Two, "Retrieving Data," describes the features of SQL that allow you to perform database queries. The first chapter in this part describes the basic structure of the SQL language. The next four chapters start with the simplest SQL queries, and progressively build to more complex queries, including multi-table queries, summary queries, and queries that use subqueries.
- Part Three, "Updating Data," shows how you can use SQL to add new data to a database, delete data from a database, and modify existing database data. It also describes the database integrity issues that arise when data is updated, and how SQL addresses these issues. The last of the three chapters in this part discusses the SQL transaction concept and SQL support for multi-user transaction processing.
- Part Four, "Database Structure," deals with creating and administering a SQL-based database. Its four chapters tell you how to create the tables, views, and indexes that form the structure of a relational database. It also describes the SQL security scheme that prevents unauthorized access to data, and the SQL system catalog that describes the structure of a database. This part also discusses the significant differences between the database structures supported by various SQL-based DBMS products.
- Part Five, "Programming with SQL," describes how application programs use SQL for database access. It discusses the embedded SQL specified by the ANSI standard and used by IBM, Oracle, Ingres, Informix, and most other SQL-based DBMS products. It also describes the dynamic SQL interface that is used to build general-purpose database tools, such as report writers and database browsing programs. Finally, this part describes the SQL APIs provided by SQL Server, Oracle, and SQLBase, and contrasts them with the embedded IBM and ANSI interfaces.

- Part Six, "Future Directions," examines the state of SQL-based DBMS products today, the directions that SQL will take over the next decade, and the likely impact of SQL in various segments of the computer market. It describes the intense current activity in distributed databases, the continuing evolution of SQL standards, and the role of SQL-based databases in online transaction processing applications. This part also discusses the battle between IBM, Oracle, Microsoft, and Ashton-Tate in the OS/2 database server market, and the impact that object-oriented databases may have on the evolution of SQL in the 1990s.

Conventions Used in this Book

Using SQL describes the SQL features and functions that are available in the most popular SQL-based DBMS products and those that are described in the ANSI/ISO SQL standard. Whenever possible, the SQL statement syntax described in the book and used in the examples applies to all dialects of SQL. When the dialects differ, the differences are pointed out in the text, and the examples follow the most common practice. In these cases, you may have to modify the SQL statements in the examples slightly to suit your particular brand of DBMS.

Throughout the book, technical terms appear in italics the first time that they are used and defined. SQL language elements, including SQL keywords, table and column names, and sample SQL statements appear in an uppercase monospace font. SQL API function names appear in a lowercase monospace font. Program listings also appear in monospace font, and use the normal case conventions for the particular programming language (uppercase for COBOL and FORTRAN, lowercase for C). Note that these conventions are used solely to improve readability; most SQL implementations will accept either uppercase or lowercase statements. Many of the SQL examples include query results, which appear immediately following the SQL statement as they would in an interactive SQL session. In some cases, long query results are truncated after a few rows; this is indicated by a vertical ellipsis (. . .) following the last row of query results.

About the Authors

James R. Groff

Jim Groff is the cofounder and President of Network Innovations Corporation, a developer of SQL-based networking software that links personal computers to corporate databases. Founded in 1984, Network Innovations was acquired by Apple Computer in 1988. Prior to founding Network Innovations, Groff held marketing positions at Hewlett-Packard and Plexus Computers, a manufacturer of UNIX-based microcomputer systems. He is a frequent speaker and author of technical articles on SQL and UNIX topics. Groff holds a B.S. degree in mathematics from the Massachussetts Institute of Technology and an M.B.A. degree from Harvard University.

Paul N. Weinberg

Paul Weinberg is the cofounder and Vice President of Network Innovations Corporation. Prior to founding Network Innovations, he held software development and marketing positions at Bell Laboratories, Hewlett-Packard, and Plexus Computers. Weinberg is the coauthor, with Jim Groff, of *Understanding UNIX: A Conceptual Guide*, a best-selling overview of the UNIX operating system for managers and data processing professionals. He holds a B.S. degree from the University of Michigan and an M.S. degree from Stanford University, both in computer science. While at Stanford, Weinberg collaborated on *The Simple Solution to Rubik's Cube*, the best-selling book of 1981, with more than seven million copies in print.

Contents at a Glance

Part One An Overview of SQL *3*

1 Introduction *5*

2 A Quick Tour of SQL *15*

3 SQL in Perspective *27*

4 Relational Databases *49*

Part Two Retrieving Data *71*

5 SQL Basics *73*

6 Simple Queries *95*

7 Multi-Table Queries (Joins) *141*

8 Summary Queries *181*

9 Subqueries *213*

Part Three Updating Data *245*

10 Database Updates *247*

11 Data Integrity *269*

12 Transaction Processing *297*

Part Four Database Structure *329*

13 Creating a Database *331*

14 Views *367*

15 SQL Security *389*

16 The System Catalog *415*

Part Five Programming with SQL *435*

17 Embedded SQL *437*

18 Dynamic SQL *493*

19 SQL APIs *539*

Part Six Future Directions *583*

20 Distributed Database Management *585*

21 The Future of SQL *611*

A The Sample Database *631*

B DBMS Vendor Profiles *639*

C Company and Product List *653*

D ANSI/ISO SQL Syntax *661*

Index *671*

Contents

Preface ..... ***xxiii***

Why this Book Is for You ***1***

Part One An Overview of SQL ***3***

1 Introduction ***5***
The SQL Language 6
The Role of SQL 8
SQL Features and Benefits 10
Vendor Independence 11
Portability Across Computer Systems 11
SQL Standards 11
IBM Endorsement 12
Relational Foundation 12
High-Level, English-Like Structure 12
Interactive, Ad Hoc Queries 12
Programmatic Database Access 13
Multiple Views of Data 13
Complete Database Language 13
Dynamic Data Definition 14
Client/Server Architecture 14

2 A Quick Tour of SQL ***15***
A Simple Database 15
Retrieving Data 17
Summarizing Data 19
Adding Data to the Database 21
Deleting Data 21
Updating the Database 22

Protecting Data 22
Creating a Database 23
Summary 24

3 *SQL in Perspective* 27
SQL and Database Management 27
A Brief History of SQL 28
The Early Years 28
Early Relational Products 30
IBM Products 31
Commercial Acceptance 31
SQL Standards 32
The ANSI/ISO Standard 33
Other SQL Standards 34
The Portability Myth 35
SQL and Networking 37
Centralized Architecture 37
File Server Architecture 38
Client/Server Architecture 39
The Impact of SQL 41
SQL and IBM's SAA 41
SQL on Minicomputers 42
SQL on UNIX Systems 43
SQL and Transaction Processing 43
SQL on Personal Computers 44
SQL and OS/2 46
Summary 47

4 *Relational Databases* 49
Early Data Models 49
File Management Systems 50
Hierarchical Databases 50
Network Databases 53
The Relational Data Model 56
The Sample Database 57
Tables 59
Primary Keys 61
Relationships 62

Foreign Keys 64
Codd's Twelve Rules* 66
Summary 69

Part Two Retrieving Data ***71***

5 SQL Basics ***73***
Statements 73
Names 77
Table Names 77
Column Names 78
Data Types 79
Extended Data Types 80
Data Type Differences 81
Constants 85
Numeric Constants 85
String Constants 86
Date and Time Constants 87
Symbolic Constants 88
Expressions 90
Built-In Functions 91
Missing Data (NULL Values) 92
Summary 93

6 Simple Queries ***95***
The SELECT Statement 95
The SELECT Clause 98
The FROM Clause 99
Query Results 99
Simple Queries 102
Calculated Columns 102
Selecting All Columns (SELECT *) 106
Duplicate Rows (DISTINCT) 107
Row Selection (WHERE Clause) 109
Search Conditions 111
Comparison Test (=, <>, <, <=, >, >=) 112
Range Test (BETWEEN) 115

Set Membership Test (IN) 118
Pattern Matching Test (LIKE) 120
Null Value Test (IS NULL) 123
Compound Search Conditions (AND, OR, and NOT) . 124
Sorting Query Results (ORDER BY Clause) 128
Rules for Single-Table Query Processing 130
Combining Query Results (UNION) * 130
Unions and Duplicate Rows * 134
Unions and Sorting * 135
Multiple Unions * 136
Summary 138

7 *Multi-Table Queries (Joins)* 141
A Two-Table Query Example 142
Simple Joins (Equi-Joins) 143
Parent/Child Queries 146
Joins with Row Selection Criteria 149
Multiple Matching Columns 149
Queries With Three or More Tables 150
Other Equi-Joins 153
Non-Equi-Joins 156
SQL Considerations for Multi-Table Queries 157
Qualified Column Names 158
All-Column Selections 159
Self-Joins 160
Table Aliases 164
Multi-Table Query Performance 165
The Structure of a Join 166
Table Multiplication 167
Rules for Multi-Table Query Processing 168
Outer Joins * 170
Left and Right Outer Joins * 175
Outer Join Notation * 176
Summary 178

8 *Summary Queries* 181
Column Functions 182
Computing a Column Total (SUM) 184

Computing a Column Average (AVG) 185
Finding Extreme Values (MIN and MAX) 186
Counting Data Values (COUNT) 187
Column Functions in the Select List 189
NULL Values and Column Functions 191
Duplicate Row Elimination (DISTINCT) 194
Grouped Queries (GROUP BY Clause) 195
Multiple Grouping Columns 198
Restrictions on Grouped Queries 202
NULL Values in Grouping Columns 204
Group Search Conditions (HAVING Clause) 206
Restrictions on Group Search Conditions 210
NULL Values and Group Search Conditions 211
HAVING Without GROUP BY 211
Summary 212

9 *Subqueries* 213
Using Subqueries 214
What Is a Subquery? 215
Subqueries in the WHERE Clause 215
Outer References 217
Subquery Search Conditions 218
Subquery Comparison Test (=, <>, <, <=, >, >=) 219
Set Membership Test (IN) 221
Existence Test (EXISTS) 223
Quantified Tests (ANY and ALL) * 226
Subqueries and Joins 232
Nested Subqueries 234
Correlated Subqueries * 235
Subqueries in the HAVING Clause * 238
Summary 241
SQL Queries—A Final Summary 241

***Part Three Updating Data* 245**

10 *Database Updates* 247
Adding Data to the Database 248

The Single-Row INSERT Statement 248
The Multi-Row INSERT Statement. 253
Bulk Load Utilities 256
Deleting Data from the Database 257
The DELETE Statement 257
Deleting All Rows 259
DELETE with Subquery * 260
Modifying Data in the Database 262
The UPDATE Statement 263
Updating All Rows 266
UPDATE with Subquery * 266
Summary 268

11 Data Integrity 269
What Is Data Integrity? 270
Required Data 271
Validity Checking 272
Entity Integrity 274
Other Uniqueness Constraints 274
Uniqueness and NULL Values 275
Referential Integrity 276
Referential Integrity Problems 277
Delete Rules 279
Cascaded Deletes * 282
Referential Cycles * 283
Foreign Keys and NULL Values * 288
Business Rules 290
What Is a Trigger? 291
Triggers and Referential Integrity 292
The Future of Triggers 293
Summary 295

12 Transaction Processing 297
What Is a Transaction? 298
COMMIT and ROLLBACK 299
The ANSI/ISO Transaction Model 302
Other Transaction Models 304
Transactions: Behind the Scenes * 307