Joe Celko’s
SQL for Smarties
Fourth Edition
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joe Celko's Data, Measurements and Standards in SQL</td>
<td>Joe Celko</td>
</tr>
<tr>
<td>Information Modeling and Relational Databases, 2nd Edition</td>
<td>Terry Halpin, Tony Morgan</td>
</tr>
<tr>
<td>Joe Celko's Thinking in Sets</td>
<td>Joe Celko</td>
</tr>
<tr>
<td>Business Metadata</td>
<td>Bill Inmon, Bonnie O'Neil, Lowell Fryman</td>
</tr>
<tr>
<td>Unleashing Web 2.0</td>
<td>Gottfried Vossen, Stephan Hagemann</td>
</tr>
<tr>
<td>Enterprise Knowledge Management</td>
<td>David Loshin</td>
</tr>
<tr>
<td>IT Manager's Handbook, 2nd Edition</td>
<td>Bill Holtsnider &amp; Brian Jaffe</td>
</tr>
<tr>
<td>Joe Celko's Puzzles and Answers, 2nd Edition</td>
<td>Joe Celko</td>
</tr>
<tr>
<td>Making Shoes for the Cobbler's Children</td>
<td>Charles Betz</td>
</tr>
<tr>
<td>Joe Celko's Analytics and OLAP in SQL</td>
<td>Joe Celko</td>
</tr>
<tr>
<td>Data Preparation for Data Mining Using SAS</td>
<td>Mamdouh Refaat</td>
</tr>
<tr>
<td>Querying XML: XQuery, XPath, and SQL/ XML in Context</td>
<td>Jim Melton and Stephen Buxton</td>
</tr>
<tr>
<td>Data Mining: Concepts and Techniques, 2nd Edition</td>
<td>Jiawei Han and Michelene Kamber</td>
</tr>
<tr>
<td>Foundations of Multidimensional and Metric Data Structures</td>
<td>Hanan Samet</td>
</tr>
<tr>
<td>Joe Celko's SQL, for Smarties: Advanced SQL Programming, 4th Edition</td>
<td>Joe Celko</td>
</tr>
<tr>
<td>Moving Objects Databases</td>
<td>Ralf Hartmut Güting and Markus Schneider</td>
</tr>
<tr>
<td>Joe Celko's SQL Programming Style</td>
<td>Joe Celko</td>
</tr>
<tr>
<td>Data Mining, Second Edition: Concepts and Techniques</td>
<td>Jiawei Han, Michelene Kamber, Jian pei</td>
</tr>
<tr>
<td>Fuzzy Modeling and Genetic Algorithms for Data Mining and Exploration</td>
<td>Earl Cox</td>
</tr>
<tr>
<td>Data Modeling Essentials, 3rd Edition</td>
<td>Graeme C. Simson and Graham C. Witt</td>
</tr>
<tr>
<td>Location-Based Services</td>
<td>Jochen Schiller and Agnès Voisard</td>
</tr>
<tr>
<td>Managing Time in Relational Databases: How to Design, Update and Query</td>
<td>Tom Johnston and Randall Weis</td>
</tr>
<tr>
<td>Temporal Data</td>
<td>Database Modeling with Microsoft Visio for Enterprise Architects</td>
</tr>
<tr>
<td>Terry Halpin, Ken Evans, Patrick Hallock, Bill Maclean</td>
<td>Designing Data-Intensive Web Applications</td>
</tr>
<tr>
<td>Stephano Ceri, Piero Fraternali, Aldo Bongio, Marco Brambilla, Sara Comai, Maristella Materia</td>
<td>Mining the Web: Discovering Knowledge from Hypertext Data</td>
</tr>
<tr>
<td>Soumen Chakrabarti</td>
<td>Advanced SQL: 1999—Understanding Object-Relational and Other Advanced Features Jim Melton</td>
</tr>
<tr>
<td>Database Tuning: Principles, Experiments, and Troubleshooting Techniques</td>
<td>Dennis Shasha, Philippe Bonnet</td>
</tr>
<tr>
<td>SQL:1999—Understanding Relational Language Components</td>
<td>Jim Melton, Alan R. Simon</td>
</tr>
<tr>
<td>Information Visualization in Data Mining and Knowledge Discovery</td>
<td>Edited by Usama Fayyad, Georges G. Grinstein, Andreas Wierse</td>
</tr>
<tr>
<td>Transactional Information Systems</td>
<td>Gerhard Weikum and Gottfried Vossen</td>
</tr>
<tr>
<td>Spatial Databases</td>
<td>Philippe Rigaux, Michel Scholl, and Agnes Voisard</td>
</tr>
<tr>
<td>Managing Reference Data in Enterprise Database</td>
<td>Malcolm Chisholm</td>
</tr>
<tr>
<td>Understanding SQL and Java Together</td>
<td>Jim Melton and Andrew Eisenberg</td>
</tr>
<tr>
<td>Performance, 2nd Edition</td>
<td>Patrick and Elizabeth O'Neil</td>
</tr>
<tr>
<td>The Object Data Standar</td>
<td>Edited by R. G. G. Cattell, Douglas Barry</td>
</tr>
<tr>
<td>Data on the Web: From Relations to Semistructured Data and XML</td>
<td>Serge Abiteboul, Peter Buneman, Dan Suciu</td>
</tr>
<tr>
<td>Data Mining, Third Edition Practical Machine Learning Tools and Techniques with Java Implementations</td>
<td>Ian Witten, Eibe Frank</td>
</tr>
<tr>
<td>Joe Celko's Data and Databases: Concepts in Practice</td>
<td>Joe Celko</td>
</tr>
<tr>
<td>Developing Time-Oriented Database Applications in SQL</td>
<td>Richard T. Snodgrass</td>
</tr>
<tr>
<td>Web Farming for the Data Warehouse</td>
<td>Richard D. Hackathorn</td>
</tr>
<tr>
<td>Management of Heterogeneous and Autonomous Database Systems</td>
<td>Edited by Ahmed Elmagarmid, Marek Rusinkiewicz, Amit Sheth</td>
</tr>
<tr>
<td>Object-Relational DBMSs: 2nd Edition</td>
<td>Michael Stonebraker and Paul Brown, with Dorothy Moore</td>
</tr>
<tr>
<td>Universal Database Management: A Guide to Object/Relational Technology</td>
<td>Cynthia Maro Saracco</td>
</tr>
<tr>
<td>Readings in Database Systems, 3rd Edition</td>
<td>Edited by Michael Stonebraker, Joseph M. Hellerstein</td>
</tr>
<tr>
<td>Principles of Multimedia Database Systems</td>
<td>V. S. Subrahmanian</td>
</tr>
<tr>
<td>Principles of Database Query Processing for Advanced Applications</td>
<td>Clement T. Yu, Weiyi Meng</td>
</tr>
<tr>
<td>Advanced Database Systems</td>
<td>Carlo Zaniolo, Stefano Ceri, Christos Faloutsos, Richard T. Snodgrass, V. S. Subrahmanian, Roberto Zicari</td>
</tr>
<tr>
<td>Transaction Processing, 2nd Edition</td>
<td>Philip A. Bernstein, Eric Newcomer</td>
</tr>
<tr>
<td>Using the New DB2: IBM's Object-Relational Database System</td>
<td>Don Chamberlin</td>
</tr>
<tr>
<td>Distributed Algorithms</td>
<td>Nancy A. Lynch</td>
</tr>
<tr>
<td>Active Database Systems: Triggers and Rules For Advanced Database Processing</td>
<td>Edited by Jennifer Widom, Stefano Ceri</td>
</tr>
<tr>
<td>Migrating Legacy Systems: Gateways, Interfaces, &amp; the Incremental Approach</td>
<td>Michael L. Brodie, Michael Stonebraker</td>
</tr>
<tr>
<td>Atomic Transactions</td>
<td>Nancy Lynch, Michael Merritt, William Weih, Alan Fekete</td>
</tr>
<tr>
<td>Query Processing for Advanced Database Systems</td>
<td>Edited by Johann Christoph Freytag, David Maier, Gottfried Vossen</td>
</tr>
<tr>
<td>Transaction Processing</td>
<td>Jim Gray, Andreas Reuter</td>
</tr>
<tr>
<td>Database Transaction Models for Advanced Applications</td>
<td>Edited by Ahmed K. Elmagarmid</td>
</tr>
<tr>
<td>A Guide to Developing Client/Server SQL Applications</td>
<td>Setrag Khoshafian, Arvola Chan, Anna Wong, Harry K. T. Wong</td>
</tr>
</tbody>
</table>
To Ann and Jackers
# CONTENTS

About the Author ................................................................. xix
Introduction to the Fourth Edition ................................. xxi

## Chapter 1  Databases versus File Systems ......................... 1
1.1 Tables as Entities ......................................................... 4
1.2 Tables as Relationships .............................................. 4
1.3 Rows versus Records .................................................. 5
1.4 Columns versus Fields ............................................... 6
1.5 Schema Objects ......................................................... 7
1.6 `CREATE SCHEMA` Statement ...................................... 7

## Chapter 2  Transactions and Concurrency Control ............. 11
2.1 Sessions ................................................................. 11
2.2 Transactions and ACID ............................................... 12
2.3 Concurrency Control .................................................. 14
2.4 Pessimistic Concurrency Control ................................. 18
2.5 `SNAPSHOT Isolation and Optimistic Concurrency` .......... 19
2.6 Logical Concurrency Control ....................................... 21
2.7 Deadlock and Livelocks .............................................. 21

## Chapter 3  Schema Level Objects .................................. 23
3.1 `CREATE SCHEMA` Statement .................................... 23
3.2 `CREATE PROCEDURE`, `CREATE FUNCTION`, and
    `CREATE TRIGGER` .................................................. 24
3.3 `CREATE DOMAIN` Statement ...................................... 24
3.4 `CREATE SEQUENCE` .................................................. 25
3.5 `CREATE ASSERTION` ................................................ 26
3.6 Character Set Related Constructs ................................. 31

## Chapter 4  Locating Data and Special Numbers ............... 35
4.1 Exposed Physical Locators ......................................... 35
4.2 Generated Identifiers ............................................... 40
4.3 Sequence Generator Functions ................................... 42
4.4 Preallocated Values .................................................. 43
4.5 Special Series ......................................................... 44
# Chapter 5  Base Tables and Related Elements

5.1 CREATE TABLE Statement ........................................... 53
5.2 Nested UNIQUE Constraints ....................................... 62
5.3 CREATE ASSERTION Constraints ................................. 76
5.4 TEMPORARY Tables .................................................. 77
5.5 Manipulating Tables ............................................... 78
5.6 Avoiding Attribute Splitting ...................................... 81
5.7 Modeling Class Hierarchies in DDL .............................. 83
5.8 Exposed Physical Locators ....................................... 85
5.9 Auto-Incrementing Columns ...................................... 85
5.10 Generated Identifiers ............................................. 90
5.11 A Remark on Duplicate Rows ................................. 94
5.12 Other Schema Objects ........................................... 96
5.13 Temporary Tables ................................................ 97
5.14 CREATE DOMAIN Statement .................................... 98
5.15 CREATE TRIGGER Statement ............................... 99
5.16 CREATE PROCEDURE Statement ............................. 99
5.17 DECLARE CURSOR Statement .............................. 100

# Chapter 6  Procedural, Semiprocedural, and Declarative Programming

6.1 Basics of Software Engineering ................................ 105
6.2 Cohesion ............................................................ 105
6.3 Coupling ............................................................. 106
6.4 The Big Leap ......................................................... 107
6.5 Rewriting Tricks .................................................... 114
6.6 Functions for Predicates ......................................... 118
6.7 Procedural versus Logical Decomposition .................... 119

# Chapter 7  Procedural Constructs

7.1 CREATE PROCEDURE ............................................... 123
7.2 CREATE TRIGGER .................................................. 124
7.3 CURSORS ............................................................ 127
7.4 SEQUENCEs ........................................................ 141
7.5 Generated Columns ............................................... 142
7.6 Table Functions .................................................... 143
Chapter 8  Auxiliary Tables. ................................................................. 145
8.1 The Series Table ................................................................. 145
8.2 Lookup Auxiliary Tables ..................................................... 151
8.3 Auxiliary Function Tables ..................................................... 159
8.4 Global Constants Tables ....................................................... 169
8.5 A Note on Converting Procedural Code to Tables ...................... 175

Chapter 9  Normalization ................................................................. 181
9.1 Functional and Multivalued Dependencies ................................. 183
9.2 First Normal Form (1NF) ......................................................... 184
9.3 Second Normal Form (2NF) ..................................................... 188
9.4 Third Normal Form (3NF) ....................................................... 189
9.5 Elementary Key Normal Form (EKNF) ....................................... 191
9.6 Boyce-Codd Normal Form (BCNF) .......................................... 192
9.7 Fourth Normal Form (4NF) ..................................................... 194
9.8 Fifth Normal Form (5NF) ....................................................... 194
9.9 Domain-Key Normal Form (DKNF) ............................................. 196
9.10 Practical Hints for Normalization ............................................ 204
9.11 Key Types ........................................................................ 205
9.12 Practical Hints for Denormalization ......................................... 208

Chapter 10 Numeric Data Types ...................................................... 215
10.1 Numeric Types ................................................................. 215
10.2 Numeric Type Conversion .................................................... 220
10.3 Four Function Arithmetic .................................................... 222
10.4 Arithmetic and NULLs ......................................................... 224
10.5 Converting Values to and from NULL ..................................... 225
10.6 Mathematical Functions ...................................................... 228
10.7 Unique Value Generators .................................................... 232
10.8 IP Addresses .................................................................. 235

Chapter 11 Temporal Data Types .................................................. 237
11.1 Notes on Calendar Standards ............................................... 237
11.2 SQL Temporal Data Types .................................................. 240
11.3 INTERVAL Data Types ......................................................... 246
<table>
<thead>
<tr>
<th>Chapter 12</th>
<th>Character Data Types</th>
<th>255</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.1</td>
<td>Problems with SQL Strings</td>
<td>255</td>
</tr>
<tr>
<td>12.2</td>
<td>Standard String Functions</td>
<td>258</td>
</tr>
<tr>
<td>12.3</td>
<td>Common Vendor Extensions</td>
<td>259</td>
</tr>
<tr>
<td>12.4</td>
<td>Cutter Tables</td>
<td>268</td>
</tr>
<tr>
<td>12.5</td>
<td>Nested Replacement</td>
<td>269</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 13</th>
<th>NULLs: Missing Data in SQL</th>
<th>271</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.1</td>
<td>Empty and Missing Tables</td>
<td>272</td>
</tr>
<tr>
<td>13.2</td>
<td>Missing Values in Columns</td>
<td>273</td>
</tr>
<tr>
<td>13.3</td>
<td>Context and Missing Values</td>
<td>275</td>
</tr>
<tr>
<td>13.4</td>
<td>Comparing NULLs</td>
<td>276</td>
</tr>
<tr>
<td>13.5</td>
<td>NULLs and Logic</td>
<td>277</td>
</tr>
<tr>
<td>13.6</td>
<td>Math and NULLs</td>
<td>281</td>
</tr>
<tr>
<td>13.7</td>
<td>Functions and NULLs</td>
<td>281</td>
</tr>
<tr>
<td>13.8</td>
<td>NULLs and Host Languages</td>
<td>281</td>
</tr>
<tr>
<td>13.9</td>
<td>Design Advice for NULLs</td>
<td>282</td>
</tr>
<tr>
<td>13.10</td>
<td>A Note on Multiple NULL Values</td>
<td>285</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 14</th>
<th>Multiple Column Data Elements</th>
<th>289</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.1</td>
<td>Distance Functions</td>
<td>289</td>
</tr>
<tr>
<td>14.2</td>
<td>Storing an IPv4 Address in SQL</td>
<td>291</td>
</tr>
<tr>
<td>14.3</td>
<td>Storing an IPv6 Address in SQL</td>
<td>293</td>
</tr>
<tr>
<td>14.4</td>
<td>Currency and Other Unit Conversions</td>
<td>294</td>
</tr>
<tr>
<td>14.5</td>
<td>Social Security Numbers</td>
<td>295</td>
</tr>
<tr>
<td>14.6</td>
<td>Rational Numbers</td>
<td>298</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 15</th>
<th>Table Operations</th>
<th>299</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.1</td>
<td>DELETE FROM Statement</td>
<td>299</td>
</tr>
<tr>
<td>15.2</td>
<td>INSERT INTO Statement</td>
<td>307</td>
</tr>
<tr>
<td>15.3</td>
<td>The UPDATE Statement</td>
<td>310</td>
</tr>
<tr>
<td>15.4</td>
<td>A Note on Flaws in a Common Vendor Extension</td>
<td>317</td>
</tr>
<tr>
<td>15.5</td>
<td>MERGE Statement</td>
<td>319</td>
</tr>
</tbody>
</table>
Chapter 16  Comparison or Theta Operators .............................................. 323
  16.1 Converting Data Types ................................................................. 323
  16.2 Row Comparisons in SQL .............................................................. 326
  16.3 IS [NOT] DISTINCT FROM Operator ............................................... 328

Chapter 17  Valued Predicates ................................................................. 329
  17.1 IS NULL ......................................................................................... 329
  17.2 IS [NOT]{TRUE | FALSE | UNKNOWN} Predicate ............................... 330
  17.3 IS [NOT] NORMALIZED Predicate .................................................. 332

Chapter 18  CASE Expressions ................................................................. 333
  18.1 The CASE Expression ................................................................. 333
  18.2 Subquery Expressions and Constants ............................................ 341
  18.3 Rozenshtein Characteristic Functions .......................................... 342

Chapter 19  LIKE and SIMILAR TO Predicates ........................................ 345
  19.1 Tricks with Patterns ........................................................................ 346
  19.2 Results with NULL Values and Empty Strings ............................... 347
  19.3 LIKE Is Not Equality ...................................................................... 348
  19.4 Avoiding the LIKE Predicate with a Join ..................................... 348
  19.5 CASE Expressions and LIKE Search Conditions ............................ 349
  19.6 SIMILAR TO Predicates ................................................................. 350
  19.7 Tricks with Strings .......................................................................... 352

Chapter 20  BETWEEN and OVERLAPS Predicates ..................................... 355
  20.1 The BETWEEN Predicate ............................................................. 355
  20.2 OVERLAPS Predicate ................................................................. 358

Chapter 21  The [NOT] IN() Predicate ....................................................... 369
  21.1 Optimizing the IN() Predicate ....................................................... 370
  21.2 Replacing ORs with the IN() Predicate .......................................... 373
  21.3 NULLs and the IN() Predicate ......................................................... 374
  21.4 IN() Predicate and Referential Constraints .................................... 376
  21.5 IN() Predicate and Scalar Queries ............................................... 377
**Chapter 22**  EXISTS() Predicate .................................................. 381
  22.1  EXISTS and NULLs ......................................................... 382
  22.2  EXISTS and INNER JOINs ................................................ 384
  22.3  NOT EXISTS and OUTER JOINs .......................................... 385
  22.4  EXISTS() and Quantifiers ............................................... 385
  22.5  EXISTS() and Referential Constraints ............................... 386
  22.6  EXISTS and Three-Valued Logic ..................................... 387

**Chapter 23**  Quantified Subquery Predicates ................................. 389
  23.1  Scalar Subquery Comparisons ......................................... 389
  23.2  Quantifiers and Missing Data ......................................... 391
  23.3  The ALL Predicate and Extrema Functions .......................... 393
  23.4  The UNIQUE Predicate .................................................. 394
  23.5  The DISTINCT Predicate ................................................. 395

**Chapter 24**  The Simple SELECT Statement .................................. 397
  24.1  SELECT Statement Execution Order .................................. 397
  24.2  One-Level SELECT Statement ......................................... 397

**Chapter 25**  Advanced SELECT Statements .................................. 407
  25.1  Correlated Subqueries .................................................. 407
  25.2  Infixed INNER JOINs ..................................................... 411
  25.3  OUTER JOINs .............................................................. 413
  25.4  UNION JOIN Operators ................................................... 425
  25.5  Scalar SELECT Expressions ............................................. 426
  25.6  Old versus New JOIN Syntax .......................................... 427
  25.7  Constrained JOINs ....................................................... 428
  25.8  Dr. Codd’s T-Join ......................................................... 437

**Chapter 26**  Virtual Tables: VIEWs, Derived Tables, CTEs,  
and MQTs ........................................................................... 445
  26.1  VIEWs in Queries ........................................................... 445
  26.2  Updatable and Read-Only VIEWs ....................................... 446
  26.3  Types of VIEWs ............................................................... 448
  26.4  How VIEWs Are Handled in the Database Engine .................. 453
  26.5  WITH CHECK OPTION Clause .......................................... 457
30.5 Ordinal Functions .................................................. 547
30.6 Vendor Extensions ............................................... 550
30.7 A Bit of History .................................................. 553

Chapter 31 Descriptive Statistics in SQL ......................... 555
31.1 The Mode .......................................................... 555
31.2 The $\text{AVG}(\cdot)$ Function .................................. 556
31.3 The Median ....................................................... 557
31.4 Variance and Standard Deviation .............................. 572
31.5 Average Deviation .............................................. 573
31.6 Cumulative Statistics ......................................... 573
31.7 Cross Tabulations .............................................. 582
31.8 Harmonic Mean and Geometric Mean ....................... 588
31.9 Multivariable Descriptive Statistics in SQL ............... 589
31.10 Statistical Functions in SQL:2006 ......................... 591

Chapter 32 Subsequences, Regions, Runs, Gaps, and Islands .... 595
32.1 Finding Subregions of Size (n) ............................... 596
32.2 Numbering Regions ........................................... 597
32.3 Finding Regions of Maximum Size ........................... 598
32.4 Bound Queries .................................................. 602
32.5 Run and Sequence Queries .................................. 603
32.6 Summation of a Series ....................................... 607
32.7 Swapping and Sliding Values in a List ...................... 610
32.8 Condensing a List of Numbers ............................... 612
32.9 Folding a List of Numbers ................................... 612
32.10 Coverings ....................................................... 613

Chapter 33 Matrices in SQL ........................................ 617
33.1 Arrays via Named Columns ................................... 617
33.2 Arrays via Subscript Columns ................................. 621
33.3 Matrix Operations in SQL .................................... 622
33.4 Flattening a Table into an Array ............................... 627
33.5 Comparing Arrays in Table Format ......................... 628
38.8 Calendar Auxiliary Table ........................................ 729
38.9 Problems with the Year 2000. ...................................... 731

Chapter 39 Optimizing SQL ............................................. 737
39.1 Access Methods .................................................. 738
39.2 How to Index ....................................................... 740
39.3 Give Extra Information ............................................ 744
39.4 Index Multiple Columns Carefully ............................... 745
39.5 Watch the IN Predicate ........................................... 746
39.6 Avoid UNIONS ..................................................... 748
39.7 Prefer Joins over Nested Queries ................................. 748
39.8 Use Fewer Statements ............................................. 749
39.9 Avoid Sorting ....................................................... 750
39.10 Avoid CROSS JOINS ............................................ 754
39.11 Know Your Optimizer ............................................. 754
39.12 Recompile Static SQL after Schema Changes ............... 756
39.13 Temporary Tables Are Sometimes Handy ..................... 757
39.14 Update Statistics ................................................ 759
39.15 Do Not Trust Newer Features ................................... 760

References ............................................................... 763
General References ...................................................... 763
Logic ................................................................. 763
Mathematical Techniques .............................................. 763
Random Numbers ....................................................... 763
Scales and Measurements ............................................. 764
Missing Values ........................................................ 765
Regular Expressions ..................................................... 765
Graph Theory .......................................................... 766
Introductory SQL Books ............................................... 766
Optimizing Queries ..................................................... 767
Temporal Data and the Year 2000 Problem ......................... 767
SQL Programming Techniques ....................................... 768
Classics ............................................................... 768
Updatable Views ......................................................... 769
Theory, Normalization, and Advanced Database Topics ....... 769
Books on SQL-92 and SQL-99 .................................................. 770
Standards and Related Groups ............................................. 770
Web Sites Related to SQL ...................................................... 771
Statistics ............................................................................ 771
Temporal Databases ............................................................ 771
Miscellaneous Citations ....................................................... 772
Index .................................................................................. 775
About the Author

Joe Celko served 10 years on ANSI/ISO SQL Standards Committee and contributed to the SQL-89 and SQL-92 Standards.

He has written over 900 columns in the computer trade and academic press, mostly dealing with data and databases, and has authored seven other books on SQL for Morgan Kaufmann:

- *Data and Databases* (1999)
- *SQL Programming Style* (2005)
- *Analytics and OLAP in SQL* (2005)

Mr. Celko’s past columns include:
- Columns for Simple Talk (Redgate Software)
- “CELKO,” *Intelligent Enterprise* magazine (CMP)
- BMC’s DBAazine.com e-magazine (BMC Software)
- “SQL Explorer,” *DBMS* (Miller Freeman)
- “Celko on SQL,” *Database Programming and Design* (Miller Freeman)
- “WATCOM SQL Corner,” *Powerbuilder Developers’ Journal* (SysCon)
- “SQL Puzzle,” *Boxes and Arrows* (Frank Sweet Publishing)
- “Data Points,” *PC Techniques* (Coriolis Group)
- “Celko on Software,” *Computing* (VNC Publications, UK)
- “SELECT * FROM Austin” (Array Publications, The Netherlands)

In addition, Mr. Celko was editor for the “Puzzles & Problems” section of ABACUS (Springer Verlag) and he ran the CASEFORUM section 18, “Celko on SQL,” on CompuServe.
INTRODUCTION TO THE FOURTH EDITION

This book, like the first, second, and third editions before it, is for the working SQL programmer who wants to pick up some advanced programming tips and techniques. It assumes that the reader is an SQL programmer with a year or more of actual experience. This is not an introductory book, so let's not have any gripes in the amazon.com reviews about that like we did with the prior editions.

The first edition was published 10 years ago, and became a minor classic among working SQL programmers. I have seen copies of this book on the desks of real programmers in real programming shops almost everywhere I have been. The true compliment are the Post-it® notes sticking out of the top. People really use it often enough to put stickies in it! Wow!

What Changed in Ten Years

Hierarchical and network databases still run vital legacy systems in major corporations. SQL people do not like to admit that IMS and traditional files are still out there in the Fortune 500. But SQL people can be proud of the gains SQL-based systems have made over the decades. We have all the new applications and all the important smaller databases.

OO programming is firmly in place, but may give ground to functional programming in the next decade. Object and object-relational databases found niche markets, but never caught on with the mainstream.

XML is no longer a fad in 2010. Technically, it is syntax for describing and moving data from one platform to another, but its support tools allow searching and reformatting. There is an SQL/XML subcommittee in INCITS H2 (the current name of the original ANSI X3H2 Database Standards Committee) making sure they can work together.

Data warehousing is no longer an exotic luxury only for major corporations. Thanks to the declining prices of hardware and software, medium-sized companies now use the technology. Writing OLAP queries is different from OLTP queries and probably needs its own “Smarties” book now.
Open Source databases are doing quite well and are gaining more and more Standards conformance. The LAMP platform (Linux, Apache, MySQL, and Python/PHP) has most of the web sites. Ingres, Postgres, Firebird, and other products have the ANSI SQL-92 features, most of the SQL-99, and some of the SQL:2003 features.

Columnar databases, parallelism, and Optimistic Concurrency are all showing up in commercial product instead of the laboratory. The SQL Standards have changed over time, but not always for the better. Parts of it have become more relational and set-oriented while other parts put in things that clearly are procedural, deal with nonrelational data, and are based on file system models. To quote David McGoveran, “A committee never met a feature it did not like.” And he seems to be quite right.

But with all the turmoil the ANSI/ISO Standard SQL-92 was the common subset that will port across SQL products to do useful work. In fact, years ago, the US government described the SQL-99 standard as “a standard in progress” and required SQL-92 conformance for federal contracts.

We had the FIPS-127 conformance test suite in place during the development of SQL-92, so all the vendors could move in the same direction. Unfortunately, the Clinton administration canceled the program and conformance began to drift. Michael M. Gorman, President of Whitemarsh Information Systems Corporation and secretary of INCITS H2 for over 20 years, has a great essay on this and other political aspects of SQL’s history at Wiscorp.com that is worth reading.

Today, the SQL-99 standard is the one to use for portable code on the greatest number of platforms. But vendors are adding SQL:2003 features so rapidly, I do not feel that I have to stick to a minimal standard.

**New in This Edition**


I have moved and greatly expanded techniques for trees and hierarchies into their own book (*Trees and Hierarchies in SQL*, ISBN 13:978-1558609204) because there was enough material to justify it. There is a short mention of some techniques here, but not to the detailed level in the other book.

is an advanced programmer's book and I assume that the reader is now writing real SQL, not some dialect or his or her native programming language in a thin disguise. I also assume that the reader can translate Standard SQL into his or her local dialect without much effort.

I have tried to provide comments with the solutions, to explain why they work. I hope this will help the reader see underlying principles that can be used in other situations.

A lot of people have contributed material, either directly or via Newsgroups and I cannot thank all of them. But I made a real effort to put names in the text next to the code. In case I missed anyone, I got material or ideas from Aaron Bertrand, Alejandro Mesa, Anith Sen, Craig Mullins (who has done the tech reads on several editions), Daniel A. Morgan, David Portas, David Cressey, Dawn M. Wolthuis, Don Burleson, Erland Sommarskog, Itzak Ben-Gan, John Gilson, Knut Stolze, Ken Henderson, Louis Davidson, Dan Guzman, Hugo Kornelis, Richard Romley, Serge Rielau, Steve Kass, Tom Moreau, Troels Arvin, Vadim Tropashko, Plamen Ratchev, Gert-Jan Strik, and probably a dozen others I am forgetting.

**Corrections and Additions**

Please send any corrections, additions, suggestions, improvements, or alternative solutions to me or to the publisher. Especially if you have a better way of doing something.

www.mkp.com