

THIRD EDITION

Learning SQL

Generate, Manipulate, and Retrieve Data

Alan Beaulieu

Beijing • Boston • Farnham • Sebastopol • Tokyo

O'REILLY®

Table of Contents

Preface.....	xi
1. A Little Background.....	1
Introduction to Databases	1
Nonrelational Database Systems	2
The Relational Model	5
Some Terminology	7
What Is SQL?	8
SQL Statement Classes	9
SQL: A Nonprocedural Language	10
SQL Examples	11
What Is MySQL?	13
SQL Unplugged	14
What's in Store	15
2. Creating and Populating a Database.....	17
Creating a MySQL Database	17
Using the mysql Command-Line Tool	18
MySQL Data Types	20
Character Data	20
Numeric Data	23
Temporal Data	25
Table Creation	27
Step 1: Design	27
Step 2: Refinement	28
Step 3: Building SQL Schema Statements	30
Populating and Modifying Tables	33
Inserting Data	33

Updating Data	38
Deleting Data	39
When Good Statements Go Bad	39
Nonunique Primary Key	39
Nonexistent Foreign Key	40
Column Value Violations	40
Invalid Date Conversions	40
The Sakila Database	41
3. Query Primer.....	45
Query Mechanics	45
Query Clauses	47
The select Clause	48
Column Aliases	50
Removing Duplicates	51
The from Clause	53
Tables	53
Table Links	56
Defining Table Aliases	57
The where Clause	58
The group by and having Clauses	60
The order by Clause	61
Ascending Versus Descending Sort Order	63
Sorting via Numeric Placeholders	64
Test Your Knowledge	65
Exercise 3-1	65
Exercise 3-2	65
Exercise 3-3	65
Exercise 3-4	65
4. Filtering.....	67
Condition Evaluation	67
Using Parentheses	68
Using the not Operator	69
Building a Condition	70
Condition Types	71
Equality Conditions	71
Range Conditions	73
Membership Conditions	77
Matching Conditions	79
Null: That Four-Letter Word	82
Test Your Knowledge	85

Exercise 4-1	86
Exercise 4-2	86
Exercise 4-3	86
Exercise 4-4	86
5. Querying Multiple Tables.....	87
What Is a Join?	87
Cartesian Product	88
Inner Joins	89
The ANSI Join Syntax	91
Joining Three or More Tables	93
Using Subqueries as Tables	95
Using the Same Table Twice	96
Self-Joins	98
Test Your Knowledge	99
Exercise 5-1	99
Exercise 5-2	99
Exercise 5-3	100
6. Working with Sets.....	101
Set Theory Primer	101
Set Theory in Practice	104
Set Operators	105
The union Operator	106
The intersect Operator	108
The except Operator	109
Set Operation Rules	111
Sorting Compound Query Results	111
Set Operation Precedence	112
Test Your Knowledge	114
Exercise 6-1	114
Exercise 6-2	114
Exercise 6-3	114
7. Data Generation, Manipulation, and Conversion.....	115
Working with String Data	115
String Generation	116
String Manipulation	121
Working with Numeric Data	129
Performing Arithmetic Functions	129
Controlling Number Precision	131
Handling Signed Data	133

Working with Temporal Data	134
Dealing with Time Zones	134
Generating Temporal Data	136
Manipulating Temporal Data	140
Conversion Functions	144
Test Your Knowledge	145
Exercise 7-1	145
Exercise 7-2	145
Exercise 7-3	145
8. Grouping and Aggregates.....	147
Grouping Concepts	147
Aggregate Functions	150
Implicit Versus Explicit Groups	151
Counting Distinct Values	152
Using Expressions	153
How Nulls Are Handled	153
Generating Groups	155
Single-Column Grouping	155
Multicolumn Grouping	156
Grouping via Expressions	157
Generating Rollups	157
Group Filter Conditions	159
Test Your Knowledge	160
Exercise 8-1	160
Exercise 8-2	160
Exercise 8-3	160
9. Subqueries.....	161
What Is a Subquery?	161
Subquery Types	163
Noncorrelated Subqueries	163
Multiple-Row, Single-Column Subqueries	164
Multicolumn Subqueries	169
Correlated Subqueries	171
The exists Operator	173
Data Manipulation Using Correlated Subqueries	174
When to Use Subqueries	175
Subqueries as Data Sources	176
Subqueries as Expression Generators	182
Subquery Wrap-Up	184
Test Your Knowledge	185

Exercise 9-1	185
Exercise 9-2	185
Exercise 9-3	185
10. Joins Revisited.....	187
Outer Joins	187
Left Versus Right Outer Joins	190
Three-Way Outer Joins	191
Cross Joins	192
Natural Joins	198
Test Your Knowledge	199
Exercise 10-1	200
Exercise 10-2	200
Exercise 10-3 (Extra Credit)	200
11. Conditional Logic.....	201
What Is Conditional Logic?	201
The case Expression	202
Searched case Expressions	202
Simple case Expressions	204
Examples of case Expressions	205
Result Set Transformations	205
Checking for Existence	206
Division-by-Zero Errors	208
Conditional Updates	209
Handling Null Values	210
Test Your Knowledge	211
Exercise 11-1	211
Exercise 11-2	211
12. Transactions.....	213
Multiuser Databases	213
Locking	214
Lock Granularities	214
What Is a Transaction?	215
Starting a Transaction	217
Ending a Transaction	218
Transaction Savepoints	219
Test Your Knowledge	222
Exercise 12-1	222

13. Indexes and Constraints.....	223
Indexes	223
Index Creation	224
Types of Indexes	229
How Indexes Are Used	231
The Downside of Indexes	232
Constraints	233
Constraint Creation	234
Test Your Knowledge	237
Exercise 13-1	237
Exercise 13-2	237
14. Views.....	239
What Are Views?	239
Why Use Views?	242
Data Security	242
Data Aggregation	243
Hiding Complexity	244
Joining Partitioned Data	244
Updatable Views	245
Updating Simple Views	246
Updating Complex Views	247
Test Your Knowledge	249
Exercise 14-1	249
Exercise 14-2	250
15. Metadata.....	251
Data About Data	251
information_schema	252
Working with Metadata	257
Schema Generation Scripts	257
Deployment Verification	260
Dynamic SQL Generation	261
Test Your Knowledge	265
Exercise 15-1	265
Exercise 15-2	265
16. Analytic Functions.....	267
Analytic Function Concepts	267
Data Windows	268
Localized Sorting	269
Ranking	270

Ranking Functions	271
Generating Multiple Rankings	274
Reporting Functions	277
Window Frames	279
Lag and Lead	281
Column Value Concatenation	283
Test Your Knowledge	284
Exercise 16-1	284
Exercise 16-2	285
Exercise 16-3	285
17. Working with Large Databases.....	287
Partitioning	287
Partitioning Concepts	288
Table Partitioning	288
Index Partitioning	289
Partitioning Methods	289
Partitioning Benefits	297
Clustering	297
Sharding	298
Big Data	299
Hadoop	299
NoSQL and Document Databases	300
Cloud Computing	300
Conclusion	301
18. SQL and Big Data.....	303
Introduction to Apache Drill	303
Querying Files Using Drill	304
Querying MySQL Using Drill	306
Querying MongoDB Using Drill	309
Drill with Multiple Data Sources	315
Future of SQL	317
A. ER Diagram for Example Database.....	319
B. Solutions to Exercises.....	321
Index.....	349